Phylogenetic analysis reveals an apparent duplication of the non-symbiotic hemoglobin 1 gene early in the evolution of monocotyledonous plants

Gustavo Rodríguez-Alonso, Raúl Arredondo-Peter*

Laboratorio de Biofísica y Biología Molecular, Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias, Universidad Autónoma del Estado de Morelos, Ave. Universidad 1001, Col. Chamilpa, 62210 Cuernavaca, Morelos, México

*Author for correspondence: Raúl Arredondo-Peter, email: ra@uaem.mx
Received 25 Jun 2012; Accepted 28 Jul 2012; Available Online 28 Jul 2012

Abstract

Two types of non-symbiotic hemoglobins (nsHbs) have been detected in angiosperms: nsHb-1 and nsHb-2. The origin of nsHb-1 and nsHb-2 prior to the monocot-dicot divergence was a major event in the evolution of land plant nsHbs. Because nsHb-2 has not been detected in monocots, apparently only nsHb-1 exists in monocots. Here, we report the phylogenetic analysis and in silico characterization of novel monocot (barley, Brachypodium, foxtail millet, maize, rice, sorghum, switchgrass, and wheat) nsHb sequences deposited in databases. Results suggest that only nsHb-1 evolved within monocots; that nsHb-1 duplicated early in the evolution of monocots, originating clade I and clade II nsHbs (nsHb-1 and nsHb-2, respectively); that nsHbs-I correspond to dicot nsHbs-1; and that nsHbs-II diversified into regular nsHbs-II, post-helix H-containing nsHbs-II, and 11 amino acids deletion-containing nsHbs-II. Molecular modeling showed that monocot nsHbs-II may fold into the myoglobin-fold and that Fe-heme is hexacoordinate in the predicted barley, Brachypodium, maize, and sorghum nsHb-II and pentacoordinate in the predicted foxtail millet, rice, and switchgrass nsHb-II.

Keywords: Evolution; Gene duplication; Hemoglobin; Monocots; Non-symbiotic; Plants

Abbreviations: Hb, Hemoglobin; Mb, myoglobin; nsHb, Non-symbiotic hemoglobin; sHb, symbiotic hemoglobin; tHb, truncated-like hemoglobin

1. Introduction

Hemoglobins (Hbs) are O₂-binding/transporting proteins that have been detected in a wide variety of organisms, including prokaryotes and eukaryotes [1, 2]. In land plants, apparently two Hb lineages exist: non-symbiotic (nsHb)/symbiotic (sHb) and truncated-like (tHb) Hb lineages [3-5]. Based on both O₂-affinity and sequence similarity nsHbs are classified into type 1 and type 2 (nsHbs-1 and nsHbs-2, respectively) [6, 7]. Symbiotic Hbs (or leghemoglobins when isolated from legume species) are localized and function only within nodules of N₂-fixing plants, including legumes and a variety of non-legume species [8-12]. Apparently, sHbs facilitate O₂-diffusion to symbiotic bacteria for aerobic respiration [9, 13]. In contrast, nsHbs and tHbs are widely distributed in land plants, from primitive bryophytes to evolved monocots and dicots, and are localized in symbiotic and non-symbiotic plant organs [3, 14-18]. It has been postulated that nsHbs-1 modulate levels of NO and redox potentials [19-22] and that nsHbs-2 transport O₂ [7, 23]. The function of land plant tHbs remains unknown, however it has been proposed that these proteins transport O₂ [18] and are involved in aspects of NO metabolism [3].

In the last few years, the outline of plant Hb evolution subsequent to land colonization has been clarified [4, 5, 24-27]. A major event in the evolution of land plant nsHbs was the duplication of an ancestral nsHb into nsHb-1 and nsHb-2 prior to the monocot-dicot divergence at ca. 140 million years ago (mya) [27, 28]. Analyses of sequences deposited in databases revealed that both nsHb-1 and nsHb-2 exist in dicots, but apparently only nsHb-1 exists in monocots [3, 4, 29, 30]. Previously, Garrocho-Villegas et al. [31] reported the existence of a nsHb (nsHb5) divergent from rice (nsHb1 through 4) nsHbs-1 and suggested that nsHbs divergent from nsHbs-1 evolved within monocots. The accumulation of sequences generated by gene cloning and genome sequencing programs in the last few years permitted the identification of novel land plant Hbs [3-5, 27], including novel monocot Hbs. Here, we report the phylogenetic analysis and in silico characterization of monocot nsHbs detected in databases. Results show that nsHb-2 was apparently lost in monocots shortly after duplication of the ancestral nsHb and that nsHb-1 duplicated early in the evolution of monocots, originating clade I and clade II nsHbs.

2. Experimental Details

2.1. Database search

Rice nsHb1 through 5 protein sequences (GenBank accession numbers AAK72229.1, AAK72228.1, AAK72230.1, AAK72231.1 and ABN45744.1, respectively) were used as probes to search for nsHb sequences in monocot genomes in the Phytozone (http://www.phytozone.org/, last accessed on December 2011) and GenBank (http://www.ncbi.nlm.nih.gov, last accessed on December 2011) databases using the Blast tool [32]. The putative nsHb sequences were subjected to a FUGE analysis (http://www-cryst.bioc.cam.ac.uk, last accessed on December 2011) to determine the most similar globin structure and the presence of proximal His at the myoglobin (Mb)-fold position F8. The putative nsHbs had to

Simplex Academic Publishers
www.simplex.academic-publishers.com
satisfy the following criteria to be considered authentic globins: length higher than 100 amino acids, a FUGE Z score higher than 6 (which corresponds to 99% specificity [33]) with globin structures, and the presence of proximal His at position F8.

2.2. Sequence alignment and phenetic analysis

Multiple sequence alignments of nsHb polypeptides were performed using the ClustalX program [34]. Sequence alignments were manually verified using the procedure described by Kapp et al. [35] based on the Mb-fold [36]. Phenogram was obtained from aligned sequences using the MEGA program (version 5.05) [37] and the Neighbor-Joining method [38] with 10000 bootstrap replicates.

2.3. Molecular modeling and analysis of predicted tertiary structures

The tertiary structure of detected monocot nsHbs-II was modeled using the automated mode of the I-TASSER server (http://zhanglab.ccmb.med.umich.edu/I-TASSER/last accessed on March 2012) [39-41] and the crystal structure of rice HbI [42] (PDB ID 1D8U) as a template. Model reliability was evaluated using the Verify3D program (http://nihserver.mbi.ucla.edu/Verify_3D/, last accessed on March 2012) [43]. The model was edited using the VMD program [44] and Adobe Photoshop® software. The distances of amino acid residues at the heme prosthetic group were calculated using the distance tool of the SwissPDBViewer program as described by Gopalasubramaniam et al. [45].

3. Results and Discussion

3.1. Detection and phylogenetic analysis of monocot nsHbs

Our search in databases using rice nsHb1 through 5 protein sequences as probes showed that a number of nsHb-like sequences exist in barley, Brachypodium, foxtail millet, maize, rice, sorghum, switchgrass, and wheat genomes. After sequence evaluation, only the nsHb sequences shown in Table 1 satisfied the criteria to be considered authentic globins (see subsection 2.1). The highest number of nsHbs (5) was detected in Oryza sativa var. indica and O. sativa var. japonica, whereas 1 to 3 nsHbs were detected in barley, Brachypodium, foxtail millet, maize, O. glaberrima, O. rufipogon, sorghum, switchgrass, and wheat. Comparisons among aligned sequences (not shown) revealed that monocot nsHbs are highly conserved: identity and similarity values were 45 to 100% and 57 to 100%, respectively.

A phenogram was constructed from the above sequence alignment (Figure 1). Phenogram topology showed that monocot nsHbs cluster into two clades: clade I and clade II (nsHb-I and nsHb-II, respectively). Clade I contains the hexacoordinate and high O2-affinity barley [46], rice [45, 47], and maize [48-50] nsHbs-1, as well as Brachypodium, foxtail millet, sorghum, switchgrass, and wheat nsHbs-I. Clade II contains rice nsHb5 (see section 1) [31] and nsHb5-like barley, Brachypodium, foxtail millet, maize, sorghum, and switchgrass nsHbs. With the exception of maize, O. glaberrima, and O. rufipogon (whose all hb copies remain unidentified because their genome sequencing are still in progress) all monocots analyzed in this work contain nsHb-I and nsHb-II (Table 1).

### Table 1. Binomial and database accession number for monocot (clade I and II) nsHbs and outgroup moss (Physcomitrella patens) nsHb1.

<table>
<thead>
<tr>
<th>Species/common name</th>
<th>Protein</th>
<th>Clade</th>
<th>Accession no.</th>
<th>Database</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oryza sativa var. indica (rice)</td>
<td>nsHb1</td>
<td>I</td>
<td>AAC49882.1</td>
<td>GenBank</td>
</tr>
<tr>
<td></td>
<td>nsHb2</td>
<td>I</td>
<td>AAC49881.1</td>
<td>GenBank</td>
</tr>
<tr>
<td></td>
<td>nsHb3</td>
<td>I</td>
<td>EAY89159.1</td>
<td>GenBank</td>
</tr>
<tr>
<td></td>
<td>nsHb4</td>
<td>I</td>
<td>EAY89160.1</td>
<td>GenBank</td>
</tr>
<tr>
<td></td>
<td>nsHb5</td>
<td>II</td>
<td>ABN45744.1</td>
<td>GenBank</td>
</tr>
<tr>
<td>Oryza sativa var. japonica (rice)</td>
<td>nsHb1</td>
<td>I</td>
<td>AAK7229.1</td>
<td>GenBank</td>
</tr>
<tr>
<td></td>
<td>nsHb2</td>
<td>I</td>
<td>AAK7228.1</td>
<td>GenBank</td>
</tr>
<tr>
<td></td>
<td>nsHb3</td>
<td>I</td>
<td>AAK7230.1</td>
<td>GenBank</td>
</tr>
<tr>
<td></td>
<td>nsHb4</td>
<td>I</td>
<td>AAK7231.1</td>
<td>GenBank</td>
</tr>
<tr>
<td></td>
<td>nsHb5</td>
<td>II</td>
<td>AAT01321.1</td>
<td>GenBank</td>
</tr>
<tr>
<td>Oryza glaberrima (rice)</td>
<td>nsHb</td>
<td>I</td>
<td>AC210470.1</td>
<td>GenBank</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(position19510-20342)</td>
<td>GenBank</td>
</tr>
<tr>
<td>Oryza rufipogon (rice)</td>
<td>nsHb</td>
<td>I</td>
<td>CU046294.1</td>
<td>GenBank</td>
</tr>
<tr>
<td>Brachypodium distachyon</td>
<td>nsHb</td>
<td>I</td>
<td>Bradi1g69320.1</td>
<td>Phytozone</td>
</tr>
<tr>
<td></td>
<td>nsHb</td>
<td>II</td>
<td>Bradi2g19690.1</td>
<td>Phytozone</td>
</tr>
<tr>
<td>Hordeum vulgare (barley)</td>
<td>nsHb</td>
<td>I</td>
<td>AAB70097.1</td>
<td>GenBank</td>
</tr>
<tr>
<td></td>
<td>nsHb</td>
<td>II</td>
<td>NIASHV31021D</td>
<td>barley full-length cDNA Database</td>
</tr>
<tr>
<td>Panicum virgatum (switchgrass)</td>
<td>nsHb1</td>
<td>I</td>
<td>Pavirv0001133m</td>
<td>Phytozone</td>
</tr>
<tr>
<td></td>
<td>nsHb2</td>
<td>I</td>
<td>Pavirv0004386m</td>
<td>Phytozone</td>
</tr>
<tr>
<td></td>
<td>nsHb</td>
<td>II</td>
<td>sg0.contigit01192.2007...21247</td>
<td>Phytozone</td>
</tr>
<tr>
<td>Sorghum bicolor (sorghum)</td>
<td>nsHb</td>
<td>I</td>
<td>Sb01g42260.1</td>
<td>Phytozone</td>
</tr>
<tr>
<td></td>
<td>nsHb</td>
<td>II</td>
<td>Sb09g025730.2</td>
<td>Phytozone</td>
</tr>
<tr>
<td>Zea mays ssp. mays (maize)</td>
<td>nsHb</td>
<td>I</td>
<td>GRMZM2G067402_T02</td>
<td>Phytozone</td>
</tr>
<tr>
<td></td>
<td>nsHb</td>
<td>II</td>
<td>GRMZM2G168898_T01</td>
<td>Phytozone</td>
</tr>
<tr>
<td>Setaria italica (foxtail millet)</td>
<td>nsHb</td>
<td>I</td>
<td>Si03749676(49-210)</td>
<td>Phytozone</td>
</tr>
<tr>
<td></td>
<td>nsHb</td>
<td>II</td>
<td>Si023403m</td>
<td>Phytozone</td>
</tr>
<tr>
<td>Triticum aestivum (wheat)</td>
<td>nsHb</td>
<td>I</td>
<td>AAN85432.1</td>
<td>Genbank</td>
</tr>
<tr>
<td>Physcomitrella patens</td>
<td>nsHb1</td>
<td>Moss (outgroup)</td>
<td>XP_001764902.1</td>
<td>Genbank</td>
</tr>
</tbody>
</table>
The existence of clade I and clade II nsHbs in monocots suggests that nsHbs-I and nsHbs-II evolved from nsHbs-1 and nsHbs-2, respectively. We evaluated this possibility by quantitating and comparing the similarities and identities of monocot nsHbs-I and nsHbs-II to nsHbs-1 and nsHbs-2 using the Blast tool and sequences deposited in the GenBank database. Monocot nsHbs-I exhibit higher similarity and identity to nsHbs-1 (~82% and ~60%, respectively) than to nsHbs-2 (~61% and ~46%, respectively). Thus, apparently monocot nsHbs-I correspond to nsHbs-1. Representatives of monocot nsHbs-I and dicot nsHbs-I have been characterized with detail. For example, the crystal structures of rice Hb1 and barley Hb have been elucidated [29, 42], and the O₂-binding kinetics were reported for rice Hb1 [47], barley Hb [46], Arabidopsis AHB1 [7], and tomato Hb [51]. However, nsHbs-2 have not been identified in monocots [30]. Blast analysis of clade II O. sativa var. indica nsHb5 and B. distachyon nsHb-II (representatives of 11 amino acids deletion-containing nsHbs-II and regular nsHbs-II, respectively (see subsection 3.2)) with sequences deposited in the GenBank database show that monocot nsHbs-II are more similar and identical to nsHbs-1 (76 to 89% and 62 to 85%, respectively) than to nsHbs-2 (67 to 70% and 50 to 59% respectively) (Table 2). This observation indicates that monocot nsHbs-II are closer to nsHbs-1 than to nsHbs-2, which suggests that the direct ancestor of monocot nsHbs-II was a nsHb-1. This finding implies that only nsHb-1 evolved within monocots and that an ancestral nsHb-1 duplicated early in the evolution of monocots (i.e. before radiation of the monocot species analyzed in this work) originating nsHbs-II.

3.2. Sequence alignment and analysis of monocot nsHbs-II
Comparisons among aligned sequences (Figure 2) showed that monocot nsHbs-II are highly conserved: similarity and identity values were 69 to 100% and 64 to 97%, respectively. Monocot nsHbs-II contain the highly conserved distal (H60/66) and proximal (H101) His as well as Phe (F30) B10 and Phe (F44) CD1. However, in contrast to monocot nsHbs-I [42, 49, 50, 52], nsHbs-II lack the pre-helix A (Figures 2 and 3). Thus, monocot nsHbs-II are shorter than nsHbs-I at the N-terminal region. A characteristic of rice var. indica, rice var. japonica, foxtail millet, and switchgrass nsHb-II is the existence of an 11 amino acids deletion at helix E. Also, sorghum and maize nsHb-II are longer than known plant nsHbs because of the existence of additional amino acids at the
Table 2. Sequence similarity and identity among selected monocot (*O. sativa* var. indica and *B. distachyon*) nsHbs-II and land plant nsHbs-1 and nsHbs-2. Identity and similarity values were obtained from Blastp analysis of query sequences with sequences deposited in the GenBank database. Only (subject) nsHb-1 and nsHb-2 sequences with highest scores are shown. Similarity values indicate amino acid position with identical polarity (negative, positive, or non-polar) in aligned sequences. Identity values indicate identical amino acids in aligned sequences.

<table>
<thead>
<tr>
<th>Query</th>
<th>Subject</th>
<th>GenBank accession no.</th>
<th>nsHb type</th>
<th>Identity (%)</th>
<th>Similarity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>O. sativa</em> var. indica nsHb5</td>
<td><em>B. distachyon</em> nsHb2</td>
<td>XP_003568090.1</td>
<td>1</td>
<td>80</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td><em>Zea mays</em> nsHb2</td>
<td>NP_001105819.1</td>
<td>1</td>
<td>79</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td><em>Hordeum vulgare</em> nsHb1</td>
<td>Q42831.1</td>
<td>1</td>
<td>64</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td><em>Citrus unshii</em></td>
<td>AAK07675.1</td>
<td>1</td>
<td>62</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td><em>Gossypium hirsutum</em></td>
<td>Q93Y92.1</td>
<td>2</td>
<td>53</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td><em>Solanum lycopersicum</em></td>
<td>NP_001234111.1</td>
<td>2</td>
<td>53</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td><em>Arabidopsis thaliana</em></td>
<td>NP_187663.1</td>
<td>2</td>
<td>51</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td><em>Brassica napus</em></td>
<td>Q941Q2.1</td>
<td>2</td>
<td>50</td>
<td>67</td>
</tr>
<tr>
<td><em>B. distachyon</em> nsHb-II</td>
<td><em>Zea mays</em> nsHb2</td>
<td>NP_001105819.1</td>
<td>1</td>
<td>85</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td><em>Oryza sativa</em> var. indica nsHb5</td>
<td>EAY98709.1</td>
<td>1</td>
<td>82</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td><em>Rheum australe</em></td>
<td>ACH63214.1</td>
<td>1</td>
<td>62</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td><em>Citrus unshii</em></td>
<td>AAK07675.1</td>
<td>1</td>
<td>66</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td><em>Gossypium hirsutum</em></td>
<td>Q93Y92.1</td>
<td>2</td>
<td>59</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td><em>Solanum lycopersicum</em></td>
<td>NP_001234111.1</td>
<td>2</td>
<td>55</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td><em>Arabidopsis thaliana</em></td>
<td>NP_187663.1</td>
<td>2</td>
<td>53</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td><em>Brassica napus</em></td>
<td>Q941Q2.1</td>
<td>2</td>
<td>53</td>
<td>68</td>
</tr>
</tbody>
</table>

Figure 2. Sequence alignment of known monocot nsHbs-II. †, distal (H60/66) and proximal (H101) His; *, Phe B10 and CD1; gray background, highly conserved amino acids. Helix position corresponds to helices from the rice Hb1 structure [42].
Figure 3. Overlay of predicted known monocot nsHbs-II (blue) and (clade I) native rice Hb1 (PDB ID 1D8U) (yellow) tertiary structure. (A), rice var. indica nsHb5; (B) foxtail millet nsHb-II; (C) switchgrass nsHb-II; (D) Brachypodium nsHb-II; (E) barley nsHb-II; (F) sorghum nsHb-II; (G) maize nsHb-II. Helices are indicated with letters A through H. Proximal and distal His are shown with green and orange color, respectively. The predicted structure of barley, Brachypodium, maize, foxtail millet, rice, sorghum, and switchgrass nsHb-II is deposited in the Protein Model Database (http://mi.caspur.it/PMDB/, last accessed on May 2012) under ID numbers PM0078169, PM0078168, PM0078174, PM0078172, PM0078170, PM0078173, and PM0078171, respectively.
C-terminal region. We named this region “post-helix H” (Figure 2). Blast analysis revealed no similarities between post-helix H and protein sequences deposited in the GenBank database. Thus, it was not possible to propose a role for post-helix H within the sorghum and maize nsHb-II structure. These observations show that (at least) the following nsHbs-II exist in monocots: nsHbs-II containing an 11 amino acids deletion at helix E (represented by rice var. indica, rice var. japonica, foxtail millet, and switchgrass nsHb-II) and regular nsHbs-II (represented by barley, Brachypodium, maize, and sorghum nsHb-II), which include the post-helix H-containing (maize and sorghum) nsHbs-II.

3.3. Analysis of predicted monocot nsHbs-II tertiary structure

Elucidating the tertiary structure of nsHbs is of interest in order to understand the evolution of these proteins in land plants. We modeled the tertiary structure of predicted monocot nsHbs-II using the automated mode of the I-TASSER server and the crystal structure of (clade I) rice Hb1 [42] (PDB ID 1D8U) as a template. Evaluation using the Verify3D program indicated that the model for monocot nsHbs-II was reliable, as values along the sequences were higher than 0. Figure 3 shows that the predicted structures of monocot nsHbs-II and the native structure of rice Hb1 are similar, thus suggesting that monocot nsHbs-II may fold into the Mb-fold. Major differences between monocot nsHbs-II and rice Hb1 were primarily detected at the CD-loop and the N- and C-terminal regions. With the exception of rice nsHb5, the size of the CD-loop is reduced in monocot nsHbs-II compared to that in rice Hb1 (Figure 3B to G). This observation suggests that the mobility of helix E (where distal His is located) is restricted in monocot nsHbs-II. The major difference between rice nsHb5 and Hb1 is that the length of the CD-loop is unusually long in predicted nsHb5 (Figure 3A) [31]. Thus, the mobility of helix E in rice nsHb5 may be greater than in other nsHbs. Sorghum and maize nsHb-II are longer at the C-terminal region than known plant nsHbs because of the existence of post-helix H (Figure 2). The predicted structure of maize and sorghum nsHb-II post-helix H corresponds to an unfolded region (Figure 3F and G). However, as indicated above (see subsection 3.2), it is still not possible to propose a role for post-helix H within the sorghum and maize nsHb-II tertiary structure.

An examination of amino acids essential to binding ligands to the Fe-heme showed that the distance between proximal His and Fe-heme is similar when comparing predicted monocot nsHbs-II to native rice Hb1 (Table 3), whereas the distances between Phe B10 and Fe-heme and between Phe CD1 and Fe-heme vary (Table 3). For example, the overall distances between Phe B10 and Fe-heme and between Phe CD1 and Fe-heme are quite similar when comparing switchgrass and maize nsHb-II to rice Hb1, but are ~1.5 to ~2.5 times longer in barley, Brachypodium, foxtail millet, and rice nsHb-II than in rice Hb1. This observation suggests that access of ligands to the heme pocket is easier in barley, Brachypodium, foxtail millet, and rice nsHb-II than in rice Hb1.

Table 3 also shows that the distance between distal His and Fe-heme is similar when comparing native rice Hb1 to the nsHb-II of predicted barley, Brachypodium, maize, and sorghum nsHb-II. In contrast, distal His is farther away from Fe-heme in predicted foxtail millet, rice, and switchgrass nsHb-II. These observations suggest that Fe-heme is hexacoordinate in barley, Brachypodium, maize, and sorghum nsHb-II and pentacoordinate in foxtail millet, rice, and switchgrass nsHb-II. Based on these observations, we predict that the O₂-binding properties of barley, Brachypodium, maize, and sorghum nsHb-II are similar to those of rice Hb1, i.e. they exhibit a very high O₂-affinity because of an extremely low O₂-dissociation rate constant [47]. In addition, we predict that foxtail millet, rice, and switchgrass nsHb-II exhibit low to moderate O₂-affinity because of high O₂-association and –dissociation rate constants.

### 4. Conclusions

The major results from this work show that (i) monocot nsHbs-I correspond to nsHbs-1, (ii) nsHbs-2 did not evolve within monocots, and (iii) only nsHbs-I were conserved in monocots, originating nsHbs-I and nsHbs-II after gene duplication. Figure 4 shows the events that we postulate occurred during the evolution of known monocot and dicot nsHbs. These events include: the duplication of an ancestral (pre-angiosperm) nsHb that originated nsHb-1 and nsHb-2 at ca.
140 mya; the evolution of dicot nshb-1 and nshbs-2 into hexacoordinate/high \(O_2\)-affinity nshb-I and pentacoordinate/moderate \(O_2\)-affinity nshb-II; the duplication of a nshb-1 into clade I and clade II nshbs early in the evolution of monocots; and the diversification of clade II nshbs into (at least) regular nshb-II, post-helix H-containing nshbs-II and 11 amino acids deletion-containing nshbs-II. An end result of these events could be the origin of clade I and clade II nshbs that code for primarily hexacoordinate and high \(O_2\)-affinity nshb-I and hexa- and pentacoordinate and high to moderate \(O_2\)-affinity nshb-II, respectively.

Acknowledgements

Authors wish to express their gratitude to Miss Gillian Klucas for English corrections. Work in the authors’ laboratory was funded by SEP-PROMEP (grant number UAMOR-PTC-01-01/PTC23) and Consejo Nacional de Ciencia y Tecnología (CoNaCyT grants numbers 25229N and 42873Q). México. G. R.-A. is a graduate student financed by CoNaCyT.

References


Additional information:
This is an Open Access article published under a Creative Commons Attribution 2.0 Generic License.
To view a copy of this license, visit http://creativecommons.org/licenses/by/2.0/

Cite this article as: