Evolution of the Globin Gene Family in Deuterostomes: Lineage-Specific Patterns of Diversification and Attrition

Federico G. Hoffmann,*[1,2] Juan C. Opazo,[3] David Hoogewijs,[4] Thomas Hankeln,[5] Bettina Ebner,[5] Serge N. Vinogradov,[6] Xavier Bailly,[7] and Jay F. Storz[8]

[1]Department of Biochemistry and Molecular Biology, Mississippi State University

[2]Institute for Genomics, Biocomputing and Biotechnology, Mississippi State University

[3]Instituto de Ecología y Evolución, Facultad de Ciencias, Universidad Austral de Chile, Valdivia, Chile

[4]Institute of Physiology and Zürich Center for Integrative Human Physiology, University of Zürich, Zürich, Switzerland

[5]Institute of Molecular Genetics, Johannes Gutenberg-University, Mainz, Germany

[6]Department of Biochemistry and Molecular Biology, Wayne State University School of Medicine

[7]Station Biologique de Roscoff, Roscoff, France

[8]School of Biological Sciences, University of Nebraska

***Corresponding author:** E-mail: federico.g.hoffmann@gmail.com.

Associate editor: Yoko Satta

Research article

Abstract

In the Metazoa, globin proteins display an underlying unity in tertiary structure that belies an extraordinary diversity in primary structures, biochemical properties, and physiological functions. Phylogenetic reconstructions can reveal which of these functions represent novel, lineage-specific innovations, and which represent ancestral functions that are shared with homologous globin proteins in other eukaryotes and even prokaryotes. To date, our understanding of globin diversity in deuterostomes has been hindered by a dearth of genomic sequence data from the Ambulacraria (echinoderms + hemichordates), the sister group of chordates, and the phylum Xenacoelomorpha, which includes xenoturbellids, acoelomorphs, and nemertodermatids. Here, we report the results of a phylogenetic and comparative genomic analysis of the globin gene repertoire of deuterostomes. We first characterized the globin genes of the acorn worm, *Saccoglossus kowalevskii*, a representative of the phylum Hemichordata. We then integrated genomic sequence data from the acorn worm into a comprehensive analysis of conserved synteny and phylogenetic relationships among globin genes from representatives of the eight lineages that comprise the superphylum Deuterostomia. The primary aims were 1) to unravel the evolutionary history of the globin gene superfamily in deuterostomes and 2) to use the estimated phylogeny to gain insights into the functional evolution of deuterostome globins. Results of our analyses indicate that the deuterostome common ancestor possessed a repertoire of at least four distinct globin paralogs and that different subsets of these ancestral genes have been retained in each of the descendant organismal lineages. In each major deuterostome group, a different subset of ancestral precursor genes underwent lineage-specific expansions of functional diversity through repeated rounds of gene duplication and divergence. By integrating results of the phylogenetic analysis with available functional data, we discovered that circulating oxygen-transport hemoglobins evolved independently in several deuterostome lineages and that intracellular nerve globins evolved independently in chordates and acoelomorph worms.

Key words: acorn worm, Ambulacraria, chordates, gene family evolution, hemoglobin, neuroglobin.

Introduction

The structural unity of animal globins is reflected in the highly conserved "globin fold," a characteristic three-dimensional folding of six to eight α-helices that enclose a noncovalently bound heme group in a hydrophobic pocket. This diagnostic structural feature is highly conserved among homologous globins from all kingdoms of life in spite of extensive divergence in amino acid sequence (Lesk and Chothia 1980; Lecomte et al. 2005). In animals, globins are expressed in a diverse range of tissues and cell types including muscle and nerve cells as well as circulating red blood cells, and in many invertebrates, extracellular globins are dissolved in vascular, coelomic, and/or perienteric fluids (Weber and Vinogradov 2001). This diversity in anatomical and cellular sites of expression is associated with a correspondingly wide variety of quaternary structures. Animal globins function as single-domain monomers, single-domain multisubunit proteins, or multi-domain multisubunit proteins with 2–18 covalently linked globin domains per chain (Vinogradov 1985; Riggs 1991; Vinogradov et al. 1993; Terwilliger 1998; Weber and Vinogradov 2001; Projecto-Garcia et al. 2010). Some of the most remarkable elaborations of quaternary structure are found in the extracellular globins of invertebrates. The monomeric myoglobin (Mb; ~17 kDa) and the tetrameric hemoglobin (Hb; ~64 kDa) of gnathostome vertebrates are dwarfed by the multisubunit "hexagonal bilayer" globins of annelid and vestimentiferan worms (~3,600 kDa), which

© The Author 2012. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

are composed of 144 globin chains and 36 linker peptides (Lamy et al. 1996), and the multidomain multisubunit globins of bivalve molluscs (~800–12,000 kDa) (Terwilliger and Terwilliger 1978).

The diversity of primary and quaternary structures displayed by animal globins underlies an equally rich diversity of function, ranging from reversible oxygen binding that is central to the familiar oxygen-transport and storage functions of vertebrate Hb and Mb, as well as functions related to facilitated oxygen diffusion, oxygen sensing, the scavenging of reactive oxygen and nitrogen species, redox signaling, and enzymatic functions involving NO oxygenase and reductase activities (Weber and Vinogradov 2001; Pesce et al. 2002; Fago et al. 2004; Hankeln et al. 2005; Burmester and Hankeln 2009). Phylogenetic analysis is required to determine which of these functions represent novel lineage-specific innovations, and which represent ancestral functions that are shared with homologous globin proteins in other eukaryotes and even prokaryotes (Hardison 1996, 1998; Vinogradov et al. 2005, 2006, 2007; Vinogradov and Moens 2008; Kakar et al. 2010). The primary functions of many animal globins have yet to be elucidated, and insights into their evolutionary origins and phylogenetic affinities may suggest hypotheses about their physiological roles. Although the majority of vertebrate globins are products of gene duplication or whole-genome duplication events that occurred in the stem lineage of vertebrates (Hoffmann et al. 2011, 2012; Storz et al. 2011), two monomeric globins, neuroglobin (*Ngb*) and globin X (*GbX*), have a more ancient history, as they derive from duplications that predate the origin of deuterostomes (Roesner et al. 2005; Ebner et al. 2010). The great antiquity and broad phyletic distribution of these two globins suggest that they may be performing physiological functions related to very fundamental aspects of cellular metabolism. The case of GbX is particularly interesting because it has been recently reported that GbX is a membrane-associated protein that may have an antioxidant function (Blank et al. 2011). Thus, identification of *Ngb* and *GbX* orthologs in nonchordate deuterostome taxa may provide additional insights into the physiological functions of these proteins.

Despite the significant advances in our understanding of globin gene family evolution in chordates (Ebner et al. 2003, 2010; Hoffmann, Opazo, et al. 2010, 2011, 2012; Hoffmann, Storz, et al. 2010; Storz et al. 2011), our understanding of globin diversity in deuterostomes as a whole has been hindered by a dearth of genomic sequence data for the Ambulacraria (echinoderms + hemichordates), the sister group of chordates, and the phylum Xenacoelomorpha, which includes xenoturbellids, acoelomorphs, and nemertodermatids. Globins were originally thought to have a very limited phyletic distribution among nonchordate deuterostomes (Vinogradov et al. 2006; Vinogradov and Moens 2008). A number of oxygen-transport globins have been functionally characterized in echinoderms, including sea cucumbers (class Holothuroidea) and brittle stars (class Ophiuroidea; Terwilliger and Read 1972; Bonaventura et al. 1976; Suzuki 1989; Mauri et al. 1991, McDonald et al. 1992; Baker and Terwilliger 1993; Kitto et al. 1998, Christensen et al. 2003), and bioinformatic searches in the genome of the sea urchin (class Echinoidea) revealed the presence of a 34-exon globin gene that encodes a single polypeptide with 16 covalently linked globin domains (Bailly and Vinogradov 2008). Here, we report the results of a phylogenetic and comparative genomic analysis of the globin gene set of deuterostomes. As a first step, we characterized the globin gene repertoire of the acorn worm, *Saccoglossus kowalevskii*, a representative of the phylum Hemichordata. We then integrated genomic sequence data from the acorn worm into a comprehensive analysis of conserved synteny and phylogenetic relationships among globin genes from representatives of the eight main lineages that comprise the superphylum Deuterostomia. The primary aims were 1) to unravel the evolutionary history of the globin gene superfamily in deuterostomes and 2) to use the estimated phylogeny to gain insights into the functional evolution of deuterostome globins. Results of our analyses indicate that the deuterostome common ancestor possessed a repertoire of at least four distinct globin paralogs and that different subsets of these ancestral genes have been retained in each of the descendant organismal lineages that survive to the present day. The phyletic distribution of globin genes among deuterostome taxa is clearly attributable to a winnowing of ancestral diversity such that different sets of structurally distinct paralogs have been retained in different lineages, and each of these paralogs seeded different lineage-specific expansions of functional diversity through repeated rounds of gene duplication and divergence.

Materials and Methods

Bioinformatic Analyses

We used bioinformatic tools to assemble a data set combining the complete globin gene repertoires of 18 deuterostome taxa, including 16 chordates (13 vertebrates, 2 urochordates, and 1 cephalochordate) plus 1 hemichordate and 1 echinoderm as representatives of Ambulacraria. The vertebrate species included five teleost fish (fugu, *Takifugu rubripes*; medaka, *Oryzias latipes*; pufferfish, *Tetraodon nigroviridis*; three-spined stickleback, *Gasterosteus aculeatus*; and zebrafish, *Danio rerio*), one amphibian (western clawed frog, *Xenopus tropicalis*), one squamate reptile (green anole lizard, *Anolis carolinensis*), three birds (chicken, *Gallus gallus*; turkey, *Meleagris gallopavo*; and zebra finch, *Taeniopygia guttata*), and three mammals (human, *Homo sapiens*; gray short-tailed opossum, *Monodelphis domestica*; and platypus, *Ornithorhynchus anatinus*). The nonvertebrate species included two sea squirts (*Ciona intestinalis* and *C. savignyi*, Urochordata), amphioxus (*Branchiostoma floridae*, Cephalochordata), the purple sea urchin (*Strongylocentrotus purpuratus*, Echinodermata), and the acorn worm (*S. kowalevskii*, Hemichordata). The acorn worm data were obtained from the Acorn Worm Genome Project

hosted by the Human Genome Sequence Center from Baylor College of Medicine. Because the sea urchin possesses a globin gene that encodes 16 putative globin domains (Bailly and Vinogradov 2008), we considered each of the domains as separately alignable globin sequences in all analyses.

In addition to retrieving the complete globin gene repertoires of the species listed above, we also included individual globin sequences from several additional deuterostome taxa. For echinoderms, we included sequences from two subunits of the coelomic red blood cell Hbs of the sea cucumber, *Caudina arenicola*. For cartilaginous fish, we obtained α- and β-*Hb* sequences from the red stingray (*Dasyatis akajei*) and gummy houndshark (*Mustelus antarcticus*), as well as *Mb* sequences from the gummy houndshark and the Port Jackson shark (*Heterodontus portusjacksoni*). For birds, we included the *globin E* sequence from the preliminary release of the mallard duck (*Anas platyrhynchos*) genome, available from the Ensembl-Pre database. Most of the gnathostome vertebrates included in this study possess multiple paralogous copies of α- and β-like globin genes. Since the monophyly of the α- and β-globin gene families has been well established (Goodman et al. 1975, 1987), we only included a representative subset of the α- and β-globins from each species in our analyses. In the genomic analyses of the various chordate taxa and the sea urchin, our bioinformatic searches identified the same complement of globin genes that had been annotated previously. Because no information was available about the globin repertoire of the acorn worm genome, we used a representative set of the previously characterized deuterostome globins to seed bioinformatic searches for homologous sequences in the genome of this organism using the BLASTP, PSI-BLAST, and TBLASTN programs (Altschul et al. 1990, 1997).

Current estimates of deuterostome systematic relationships place xenoturbellids, acoelomorph worms, and nemertodermatids in the phylum Xenacoelomorpha, which is regarded as sister to Ambulacraria (Bourlat et al. 2006; Philippe et al. 2011). Thus, in order to include representatives from each of the main deuterostome lineages, we obtained two globin sequences from the acoel *Symsagittifera roscoffensis*, three globin sequences from the nemertodermatid *Nemertoderma westbladi*, plus 1 sequence from the xenoturbellid *Xenoturbella bocki*. The acoel globins derive from the acoel genome project, whereas the nemertodermatid and xenorurbellid globins were identified bioinformatically in the Trace Archive from the National Center for Biotechnology Information. The full list of sequences and the corresponding accession numbers are presented in supplementary table S1 (Supplementary Material online). Finally, based on the results from Hoogewijs et al. (2012), we added nine globin sequences from plants, which were included as outgroup sequences in phylogenetic analyses. The list of plant sequences and the corresponding accession numbers are presented in supplementary table S2 (Supplementary Material online).

Phylogenetic Analyses

We conducted our phylogenetic analyses in two stages. In the first stage, we aligned all sequences using MAFFT with automatic strategy selection (Katoh and Toh 2008), and we used Neighbor-Joining phylogenetic reconstructions as implemented in MEGA version 5 (Tamura et al. 2011) to identify and remove redundant sequences. In the second stage, we used the pruned data set to build multiple alignments using a variety of approaches: Kalign2 (Lassmann et al. 2009), the E-INS-i, G-INS-i, and L-INS-i strategies from Mafft v6.83 (Katoh and Toh 2008; Katoh et al. 2009), Muscle (Edgar 2004), PROMALS3d (Pei et al. 2008), and T-coffee (Notredame et al. 2000). In addition, we also tested a profile alignment approach in which the globin sequences from acorn worm, sea urchin, sea cucumber, and *C. savignyi*, plus those of the plant globins were added to a previously assembled alignment of globin sequences from vertebrates, amphioxus, and *C. intestinalis* (Hoffmann et al. 2011). We then employed MUMSA (Lassmann and Sonnhammer 2005, 2006) to rank the multiple alignments and then used the suiteMSA package ver 1.2.06 (Anderson et al. 2011) to visually compare between the two alignments with the highest MUMSA scores and selected the alignment from the profile alignment approach for structural reasons. Subsequently, we used the selected alignment to estimate phylogenetic relationships using Bayesian methods as implemented in MrBayes version 3.1.2 (Ronquist and Huelsenbeck 2003) under a mixed model of amino acid substitution. We set two independent runs of six simultaneous chains for 10,000,000 generations, sampling every 2,500 generations, and using default priors. Once convergence was verified, support for the nodes and parameter estimates were derived from a majority rule consensus of the last 2,500 trees. All trees were rooted with plant *Hbs*.

Results and Discussion

The Globin Gene Repertoire of Hemichordates

Our bioinformatic surveys of genomic and transcriptomic databases for the acorn worm revealed the presence of 17 putative globin genes that ranged in size from 151 to 243 codons (table 1). Sixteen of these sequences were derived from the acorn worm draft genome and one was derived from an mRNA record. Because the latter sequence (ACY92586) was identical to one of the genome-derived records (NP_001161601), we excluded the mRNA-derived sequence from all further analyses. The 16 acorn worm globin sequences were found on eight different genomic scaffolds, and bioinformatic predictions indicated that 11 of them have the 3-exon–2-intron structure typical of vertebrate-specific globins, 1 has a 4-exon–3-intron structure, and the remaining 4 are intron-less genes.

The 16 acorn worm globins identified appear to be bona fide globins based on several criteria. The coding sequence lengths were within the range expected for functional globin proteins and were similar in size to the globins of sea squirts (Ebner et al. 2003) and amphioxus (Ebner et al. 2010). Moreover, the amino acid sequence alignment

Table 1. Globin Genes Identified in the Acorn Worm Genome Assembly.

Label	Protein Record	Gene ID	Number of Codons	Exons	Genomic Location
Acorn worm *Gb1*	NP_001161601	LOC100313670	155	3	scaffold_29426
Acorn worm *Gb2*	XP_002732485	LOC100378901	183	3	scaffold_9907
Acorn worm *Gb3*	XP_002732488	LOC100366547	151	3	scaffold_9907
Acorn worm *Gb4*	XP_002732489	LOC100366699	151	3	scaffold_9907
Acorn worm *Gb5*	XP_002732490	LOC100366851	224	4	scaffold_9907
Acorn worm *Gb6*	XP_002733371	LOC100374916	183	3	scaffold_14411
Acorn worm *Gb7*	XP_002739222	LOC100375093	204	3	scaffold_38908
Acorn worm *Gb8*	XP_002739227	LOC100375836	243	3	scaffold_38908
Acorn worm *Gb9*	XP_002739228	LOC100375983	180	3	scaffold_38908
Acorn worm *Gb10*	XP_002739229	LOC100376130	163	3	scaffold_38908
Acorn worm *Gb11*	N/A	N/A	193	1	scaffold_25907
Acorn worm *Gb12*	N/A	N/A	189	1	scaffold_38407
Acorn worm *Gb13*	N/A	N/A	191	1	scaffold_44414
Acorn worm *Gb14*	N/A	N/A	200	1	scaffold_3524
Acorn worm *Gb15*	N/A	N/A	242	3	scaffold_38908
Acorn worm *Gb16*	N/A	N/A	174	3	scaffold_38908

NOTE.—Acorn worm *Gbs* 11–16 were manually annotated and do not have an associated protein record or gene ID number. N/A, not applicable.

revealed that functionally important residues are conserved in all acorn worm globin sequences (fig. 1). The conserved sites include the CD1 phenylalanine, and the proximal (F8) and distal (E7) histidines, which are involved in heme-group coordination and/or Fe^{2+}-ligand stabilization, respectively.

A Phylogenetic Tree of Deuterostome Globins

Current estimates of the deuterostome phylogeny (see the inset tree in fig. 2) place the group Ambulacraria, which includes echinoderms and hemichordates, as sister to the phylum Xenacoelomorpha, which includes acoels, nemertodermatids, and xenoturbellids (Bourlat et al. 2006; Philippe et al. 2011). To reconstruct the evolutionary history of the deuterostome globin gene superfamily, we combined the complete set of globins from the acorn worm with a diverse array of globins from echinoderms in the subphylum Echinozoa (the purple sea urchin [class Echinoidea] and sea cucumber [class Holothuroidea]) plus representatives of each of the major chordate and xenacoelomorph lineages.

The initial data set of 180 sequences included an overrepresentation of vertebrate globins relative to those from nonvertebrate deuterostomes, as well as redundant sequences (the full list of sequences and the corresponding accession numbers are presented in supplementary tables S1 and S2, Supplementary Material online). In order to achieve a more balanced taxonomic representation without sacrificing phylogenetic coverage of globin diversity, we first generated a multiple sequence alignment using all sequences and we then built a neighbor joining tree (supplementary fig. S1, Supplementary Material online). The resultant tree allowed us to identify redundant and highly similar sequences, which we then removed to prune our initial data set to 110 sequences. We then aligned these sequences with eight alternative strategies (see Material and Methods), ranked the resulting alignments base on their MUMSA score, and compared the two alignments with the highest MUMSA scores (supplementary table S3, Supplementary Material online) using the suiteMSA package ver 1.2.06 (Anderson et al. 2011). For structural reasons, we selected the one from the profile alignment approach for all subsequent phylogenetic analyses (supplementary data file S1, Supplementary Material online).

Bayesian phylogenies arranged all deuterostome globins other than amphioxus *Gb8* and the acorn worm intron-less globins (acorn worm *Gb11*, *Gb12*, *Gb13*, and *Gb14*) into four well-supported clades (fig. 2), each of them defined by the presence of both chordate and ambulacrarian globins. The first of these groups, clade 1, contains all vertebrate-specific globins (i. e., *Cygb* [including cyclostome *Hbs*], *Mb*, *α-* and *β-Hb*, *GbE*, and *GbY*), plus two separate groups of amphioxus globins, all urochordate globins, a clade of acorn worm globins, which is sister to a clade that includes nemertodermatid and xenoturbellid globins, and the two Hb subunit isoforms of the sea cucumber, which are sister to acoel *Gb1*. The second group, clade 2, contains acorn worm *Gbs 7–10*, *16*, acoel *Gb2*, and vertebrate *GbX* and is sister to the third group, clade 3, which contains acorn worm *Gb6* and *Gb15*, and amphioxus *Gbs 3*, *6*, *12–14*. Finally, the fourth group, clade 4, includes the sea urchin subdomains, plus vertebrate *Ngb* and amphioxus *Gb4*. Relationships among these four clades are well resolved, with the Ngb-containing clade 4 sister to the remaining three clades.

These phylogenies placed the acorn worm globins in four separate clades and placed the six xenacoelomorph globins in three separate clades (fig. 2). Acorn worm *Gb6* and *Gb15* were placed in globin clade 3, together with amphioxus *Gbs 3*, *6*, and *12–14* with strong support, and the clade containing acorn worms *Gbs 7–10* and *Gb16* was placed in globin clade 2, together with acoel *Gb2* and vertebrate *GbX* sequences also with strong support. Acorn worm *Gbs 1–5* were placed in globin clade 1, in a clade sister to the globins from *Nemertoderma* and *Xenoturbella*, whereas the intron-less acorn worm globins (acorn worm *Gbs11–14*) were placed in a deeply diverged clade with

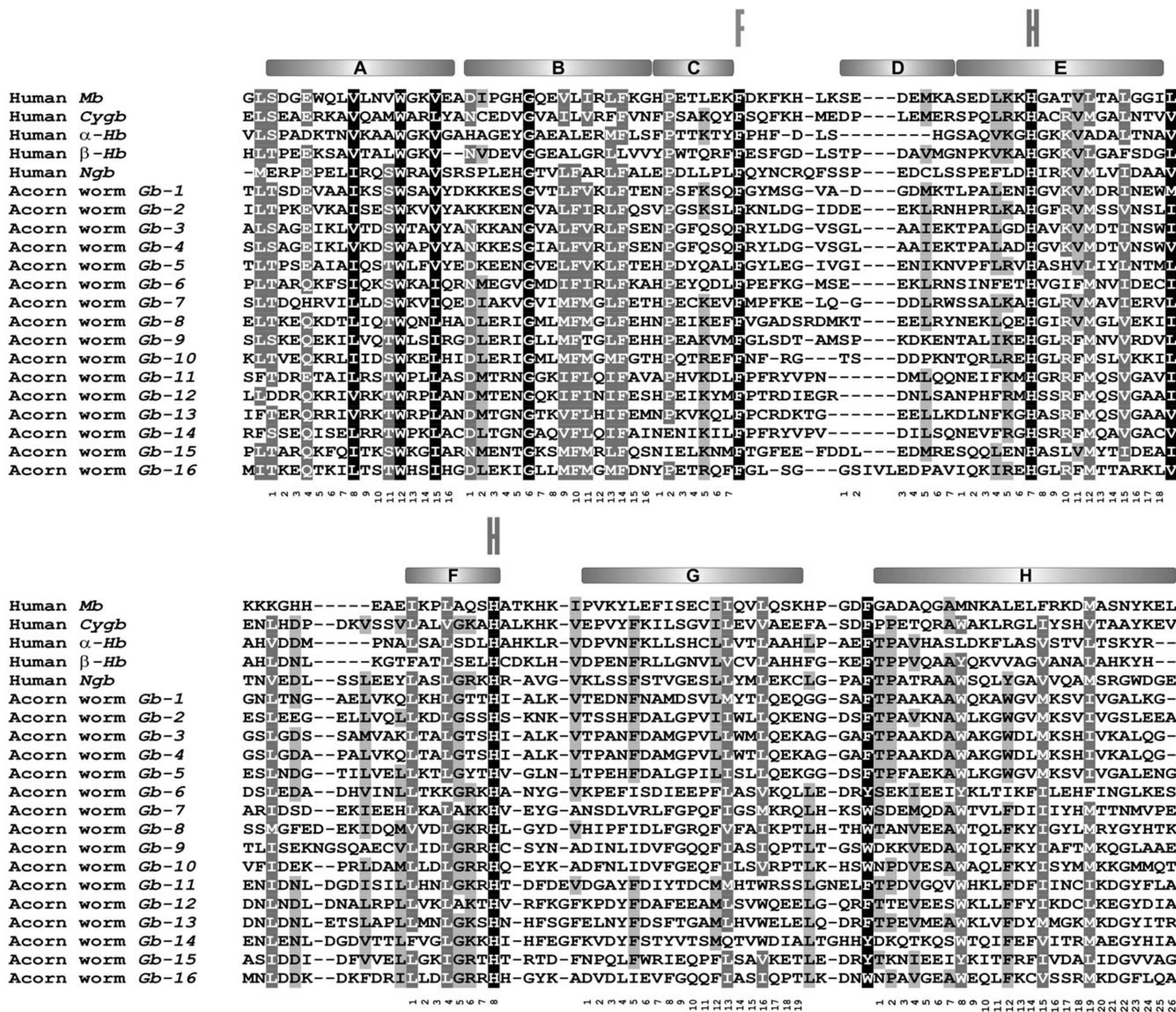


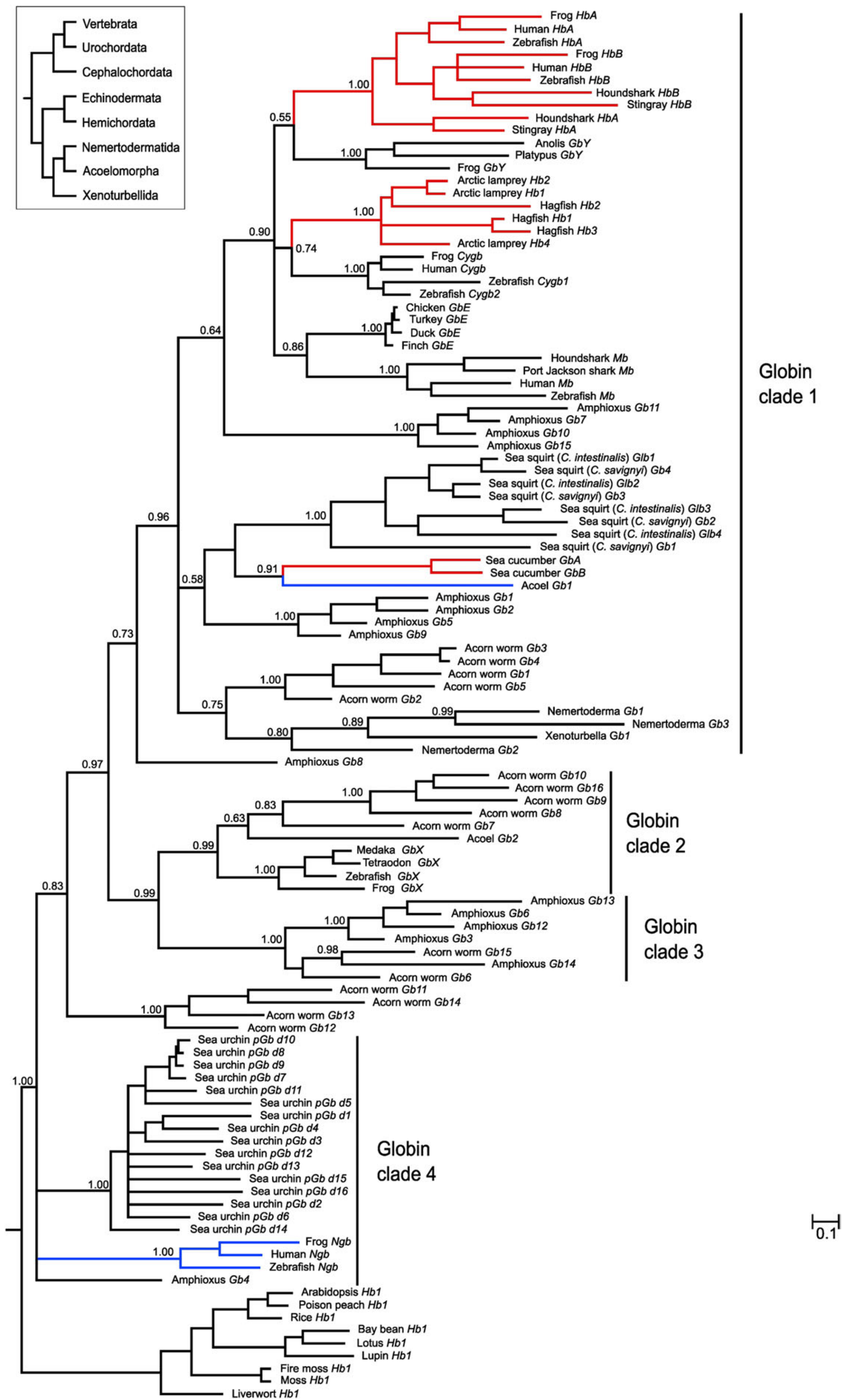
Fig. 1. Structural alignment of acorn worm globins to Human Mb, Cygb, α-Hb, β-Hb, and Ngb sequences. The α-helical structure is shown on top of the alignments. The functionally important phenylalanine in position CD1 and the distal and proximal histidines in positions E7 and F8 are indicated. Numbers correspond to the Human Mb sequence, which was used as a reference. N- and C-terminal extensions of the alignment are not shown.

uncertain phylogenetic affinities. Finally, the remaining acoel globin, acoel *Gb1*, which is functionally similar to vertebrate Ngb (Bailly X, unpublished data) is also placed in globin clade 1, but sister to the two Hb subunit isoforms of the sea cucumber. In all cases, xenacoelomorph globins are placed sister to ambulacrarian globins, as expected under the systematic arrangement shown in the inset of figure 2.

The phylogenetic arrangement in figure 2 is consistent with previous studies in arranging vertebrate globins into those that derive from vertebrate-specific duplications (*Cygb*, *GbE*, *GbY*, *Mb*, and the *Hbs* from jawed and jawless vertebrates) and those that derive from duplications that predate the divergence between deuterostomes and protostomes (*GbX* and *Ngb*; Roesner et al. 2005; Hoffmann, Opazo, et al. 2010; Dröge and Makałowski 2011; Hoffmann et al. 2011, 2012; Hoogewijs et al. 2012; Storz et al. 2011). Furthermore, as in Hoogewijs et al. (2012), our results also identify the clade containing vertebrate *Ngb* as the deepest split among deuterostome globins.

Patterns of Conserved Synteny Corroborate the Phylogenetic Reconstructions

Because globins are short proteins, they provide a limited number of informative sites for resolving phylogenetic relationships. We therefore used comparisons of conserved synteny to provide a second line of evidence concerning the duplicative origins of deuterostome globins. This analysis of conserved synteny revealed that the genomic locations of the acorn worm globins are concordant with the inferred phylogenies. The five acorn worm globins that formed a clade sister to vertebrate *GbX* (*Gbs 7–10, 16*) are arranged in tandem along scaffold_38908, which also includes acorn worm *Gb15*. Four of the five acorn worm genes in globin clade 1 (*Gbs 2–5*) are also arranged in


Vertebrata
Urochordata
Cephalochordata
Echinodermata
Hemichordata
Nemertodermatida
Acoelomorpha
Xenoturbellida
Frog HbA
Human HbA
Zebrafish HbA
Frog HbB
Human HbB
Zebrafish HbB
Houndshark HbB
Stingray HbB
Houndshark HbA
Stingray HbA
Anolis GbY
Platypus GbY
Frog GbY
Arctic lamprey Hb2
Arctic lamprey Hb1
Hagfish Hb2
Hagfish Hb1
Hagfish Hb3
Arctic lamprey Hb4
Frog Cygb
Human Cygb
Zebrafish Cygb1
Zebrafish Cygb2
Chicken GbE
Turkey GbE
Duck GbE
Finch GbE
Houndshark Mb
Port Jackson shark Mb
Human Mb
Zebrafish Mb
Amphioxus Gb11
Amphioxus Gb7
Amphioxus Gb10
Amphioxus Gb15
Sea squirt (C. intestinalis) Glb1
Sea squirt (C. savignyi) Gb4
Sea squirt (C. intestinalis) Glb2
Sea squirt (C. savignyi) Gb3
Sea squirt (C. intestinalis) Glb3
Sea squirt (C. savignyi) Gb2
Sea squirt (C. intestinalis) Glb4
Sea squirt (C. savignyi) Gb1
Sea cucumber GbA
Sea cucumber GbB
Acoel Gb1
Amphioxus Gb1
Amphioxus Gb2
Amphioxus Gb5
Amphioxus Gb9
Acorn worm Gb3
Acorn worm Gb4
Acorn worm Gb1
Acorn worm Gb5
Acorn worm Gb2
Nemertoderma Gb1
Nemertoderma Gb3
Xenoturbella Gb1
Nemertoderma Gb2
Amphioxus Gb8
Globin clade 1
Acorn worm Gb10
Acorn worm Gb16
Acorn worm Gb9
Acorn worm Gb8
Acorn worm Gb7
Acoel Gb2
Medaka GbX
Tetraodon GbX
Zebrafish GbX
Frog GbX
Globin clade 2
Amphioxus Gb13
Amphioxus Gb6
Amphioxus Gb12
Amphioxus Gb3
Acorn worm Gb15
Amphioxus Gb14
Acorn worm Gb6
Globin clade 3
Acorn worm Gb11
Acorn worm Gb14
Acorn worm Gb13
Acorn worm Gb12
Sea urchin pGb d10
Sea urchin pGb d8
Sea urchin pGb d9
Sea urchin pGb d7
Sea urchin pGb d11
Sea urchin pGb d5
Sea urchin pGb d1
Sea urchin pGb d4
Sea urchin pGb d3
Sea urchin pGb d12
Sea urchin pGb d13
Sea urchin pGb d15
Sea urchin pGb d16
Sea urchin pGb d2
Sea urchin pGb d6
Sea urchin pGb d14
Globin clade 4
Frog Ngb
Human Ngb
Zebrafish Ngb
Amphioxus Gb4
Arabidopsis Hb1
Poison peach Hb1
Rice Hb1
Bay bean Hb1
Lotus Hb1
Lupin Hb1
Fire moss Hb1
Moss Hb1
Liverwort Hb1
0.1
1.00
0.55
1.00
1.00
0.90
0.74
1.00
0.64
1.00
0.86
1.00
1.00
0.96
1.00
0.91
0.58
1.00
0.73
1.00
0.75
0.99
0.89
0.80
0.97
1.00
0.83
0.63
0.99
1.00
0.83
0.99
1.00
1.00
0.98
1.00
1.00
1.00
1.00

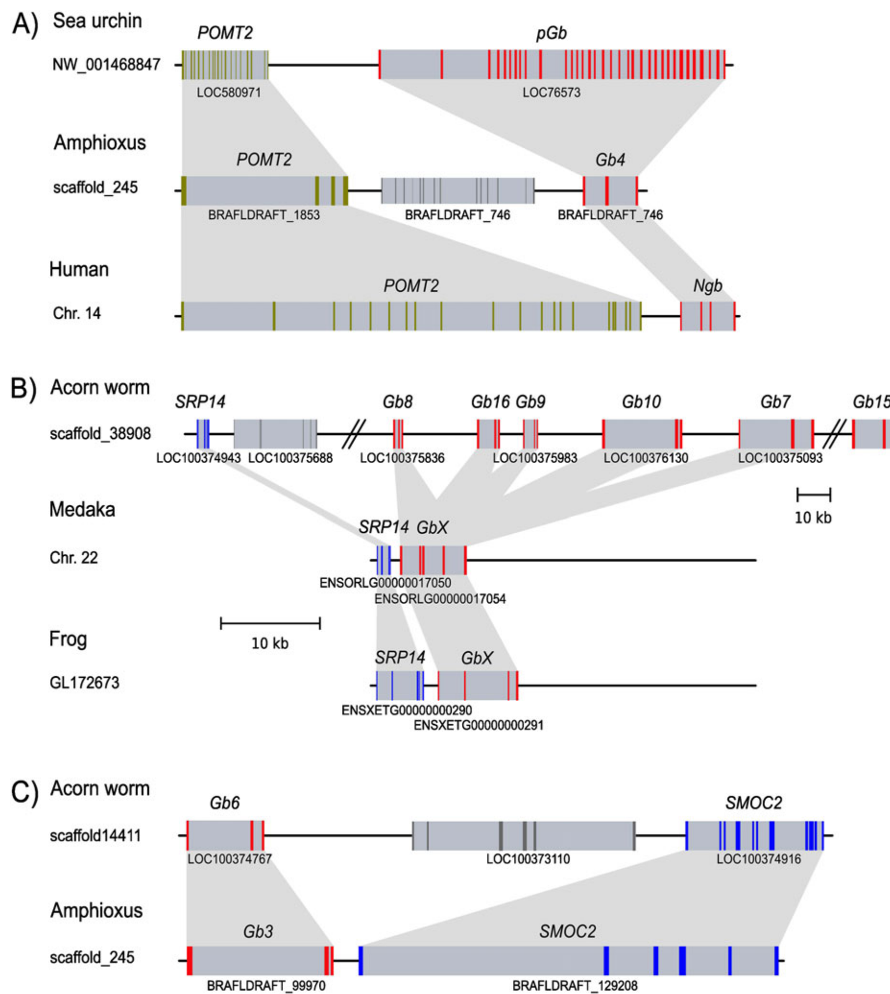


FIG. 3. Patterns of conserved synteny in the genomic regions that harbor paralogous genes from deuterostome globin clades 2–4. (*A*) Synteny comparisons of the *Ngb* region in acorn worm, amphioxus, and vertebrates (as represented by human). (*B*) Synteny comparisons of the *GbX* region in acorn worm, medaka, and frog. (*C*) Synteny comparisons between homologous chromosomal regions that contain the acorn worm *Gb6* and amphioxus *Gb3* genes. Nonglobin genes have been labeled according to their human ortholog. Intervening genes that do not contribute to conserved synteny are shown in gray. Horizontal lines denote orthologous relationships. The small scale applies to the acorn worm fragment that harbors four globin genes that are coorthologous to vertebrate *GbX*.

tandem on scaffold_9907, and the remaining acorn worm globins are located in separate scaffolds (table 1).

Genomic comparisons of conserved synteny provide additional support for globin clades 2–4 in the phylogeny of deuterostome globins (fig. 3). For example, in the case of globin clade 4, vertebrate *Ngb*, amphioxus *Gb4*, and the multidomain globin gene of the purple sea urchin (*pGb*) are each found adjacent to an ortholog of the human gene *POMT2* (fig. 3A), which is consistent with results reported by Ebner et al. (2010). In the case of globin clade 2, the five acorn worm globins that are coorthologs of vertebrate *GbX* are arranged in a cluster that is flanked by orthologs of the human *SRP14* gene, which is also located in the flanking region of *GbX* in frog, zebrafish, stickleback, medaka, and fugu (fig. 3B). Synteny comparisons for globin clade 3 are complicated by the fact that the four amphioxus globin genes in this clade are found on separate scaffolds (Ebner et al. 2010), and the same applies to the two acorn worm globins in this clade. Interestingly, both acorn worm *Gb6* and amphioxus *Gb3* are found next to an ortholog of the human *SMOC2* gene (fig. 3C). Previous synteny comparisons had identified similarities in genomic context between vertebrate *GbX* and amphioxus *Gb3*, suggesting they might be orthologs (Ebner et al. 2010).

Posttranslational Modifications in Globin Clades 2 and 3

Several acorn worm globins contain N- and C-terminal extensions as observed for nematode, fruitfly, and amphioxus globins as well as vertebrate Cygb and GbX (Burmester and

⟵
FIG. 2. Bayesian phylogram describing relationships among globin genes from representative deuterostome taxa, rooted with plant *Hbs*. Numbers next to the nodes correspond to Bayesian posterior probabilities. Circulating oxygen-transport globins are shown in red, whereas noncirculating intracellular nerve globins are shown in blue. The inset on top shows the organismal phylogeny.

Hankeln 1999; Burmester et al. 2002, 2006; Roesner et al. 2005; Hoogewijs et al. 2008; Ebner et al. 2010). Interestingly, recent evidence suggests that vertebrate GbX is a membrane-bound protein, and that the myristoylation and palmitoylation sites in the N-terminal extension of this protein are critical for its correct subcellular localization (Blank et al. 2011). Both N-myristoylation and S-palmitoylation are posttranslational modifications characteristic of membrane-associated proteins (Resh 1999; Farazi et al. 2001; Linder and Deschenes 2007), and thus, the presence of potential N-myristoylation and S-palmitoylation sites might provide clues to the subcellular localization and functional role of the analyzed proteins. Using computational prediction tools (Myristoylator [Bologna et al. 2004], MYR Predictor [Maurer-Stroh et al. 2002], and CSS-Palm [Ren et al. 2008]), we identified a putative N-terminal myristoylation site (at Gly2) in ten of the acorn worm globins and a palmitoylation site (at Cys3) in nine of them (supplementary table S4, Supplementary Material online). Moreover, all the vertebrate GbX sequences were predicted to have both myristoylation and palmitoylation sites, as was the case for the majority of the globins included in clades 2 and 3. These results suggest that these genes might encode membrane-associated globins with physiological roles similar to that of vertebrate GbX.

The Globin Repertoire of the Deuterostome Common Ancestor

For the purpose of reconstructing the complete complement of globin genes in the deuterostome common ancestor, we reasoned that monophyletic groups containing sequences from representatives of both Ambulacraria (echinoderms and hemichordates) and Chordata (vertebrates, cephalochordates, and urochordates) must trace their ancestry back to the stem lineage of deuterostomes. According to this criterion, we inferred that the last common ancestor of extant deuterostomes possessed a repertoire of at least four paralogous globin genes. These four globin genes may well be older than the divergence of Protostomia and Deuterostomia, and some of them may have originated prior to the split between animals and plants. It should be noted that this inferred number of ancestral globin genes represents a minimal estimate, as the addition of complete genome sequence data from additional deuterostomes might result in the discovery of additional globin lineages that could be traced to the deuterostome common ancestor. In fact, the phylogenetic relationships within globin clades 1 and 3 could reflect the presence of additional globin genes in the common ancestor of deuterostomes, and a similar case can be made for the divergent amphioxus *Gb8* sequence and for the intron-less acorn worm globins.

Gains and Losses of Deuterostome Globin Genes

The phyletic distribution of the four defined globin lineages and the phylogenetic relationships within each of the corresponding clades reveal a complex history of lineage-specific gene duplications and deletions. None of the five deuterostome phyla for which we have whole-genome sequence data (Cephalochordata, Echinodermata, Hemichordata, Urochordata, and Vertebrata) appear to have retained representatives of all four ancestral deuterostome globins. The species included in this study exhibit tremendous variability in the size and membership composition of their globin gene repertoires. For example, the globin repertoires of the sea urchin and sea squirts descend from two distinct paralogs that were present in the deuterostome common ancestor, and each paralogous gene lineage diversified independently in the two separate organismal lineages via repeated rounds of duplication and divergence. The limited data available suggest that the globins from the sea cucumber also derive from several rounds of lineage-specific duplication and divergence. However, there are several Hb subunits from this species for which we have no sequence data (Kitto et al. 1998). Thus, additional genomic sequence data will be required to obtain more refined insights into the origins of the sea cucumber globins.

In contrast to the patterns observed in sea urchins and sea squirts, representatives of hemichordates, cephalochordates, and vertebrates have retained orthologous copies of at least three distinct genes that were present in the deuterostome common ancestor. These ancestral gene sets diversified independently in each of the different organismal lineages. Perhaps the most striking pattern is the one observed in globin clade 4, which contains vertebrate *Ngb*s. In this case, orthologs of *Ngb* have been secondarily lost in urochordates and hemichordates and have been retained as single copy genes in cephalochordates and in vertebrates. Remarkably, the *Ngb* ortholog in sea urchin has undergone multiple rounds of internal domain duplication so that it encodes a polypeptide with 16 covalently linked globin domains (Bailly and Vinogradov 2008).

Convergent Evolution of Deuterostome Globins

Phylogenetic analyses of vertebrate globins have revealed the origins of novel functional properties and have also revealed cases in which evolution has fashioned structurally distinct design solutions to similar physiological problems (Berenbrink et al. 2005; Berenbrink 2007; Opazo et al. 2008; Hoffmann, Storz, et al. 2010; Hoffmann, Opazo, et al. 2010; Storz et al. 2011). Our results demonstrate that within deuterostomes, circulating intracellular globins with oxygen-transport functions evolved at three times independently, in echinoderms, in cyclostome vertebrates, and in gnathostome vertebrates. Similarly, the functional Ngb of acoel, acoel *Gb1*, is clearly not orthologous to chordate *Ngbs*, indicating that these nerve globins evolved independently. In fact, acoel *Gb1* is placed sister to the oxygen-carrying sea cucumber Hbs, in globin clade 1. A similar case of recruitment of an Hb-type globin as a nerve globin has been reported for the mollusc *Spisula solidissima* (Dewilde et al. 2006). In the case of circulating oxygen-transport globins, the underlying structural mechanisms of cooperative oxygen binding also evolved independently. Noncirculating globins, such as the gnathostome Cygb,

GbE, GbX, and Ngb proteins are present in clades 1, 2, and 4, and circulating intracellular Hbs with cooperative oxygen binding are found on three different lineages within clade 1 (in gnathostomes [jawed vertebrates], cyclostomes [lampreys and hagfish], and sea cucumbers). In most gnathostomes, the cooperativity of tetrameric $\alpha_2\beta_2$ Hb stems from an oxygenation-linked change in the Fe^{2+}-F8(His) bond that is propagated to the $\alpha_1\beta_2$ contacts. The resultant transition in quaternary structure between the oxy and the deoxy conformations involves a reordering of hydrogen bonds at the intersubunit contacts formed by the C, G, and H helices (Perutz et al. 1987). By contrast, in the oxygen-transport Hbs of cyclostomes, cooperativity stems from an oxygenation-linked dissociation of homo- and/or heterodimers into oxy-state monomers (e.g., Brittain and Wells 1986; Brittain et al. 1989; Fago and Weber 1995; Qiu et al. 2000; Fago et al. 2001). This ligand-dependent dissociation is mediated by intersubunit contacts involving the E and F helices such that the heme groups are in almost direct contact in the deoxy state. Similar modes of ligand-dependent dissociation/association dynamics are observed in the intracellular and extracellular Hbs of invertebrates (Kitto et al. 1998; Riggs 1998; Weber and Vinogradov 2001). For example, in the intracellular Hbs of the sea cucumber, *Caudina arenicola*, cooperativity stems from ligand-linked association into deoxygenated tetramers and higher level polymers (Bonaventura and Kitto 1973; Mitchell et al. 1995; Kitto et al. 1998). Similar to the Hbs of cyclostomes and those in many invertebrate taxa, the intersubunit contacts in the coelomic red blood cell Hbs of *Caudina* are formed by the E and F helices. However, crystal structures of *Caudina* Hbs indicate that the intersubunit contacts involve a unique set of residues (Mitchell et al. 1995; Kitto et al. 1998).

Results of our phylogenetic survey also shed light on the mechanism involved in the evolution of multidomain globins. Whereas Ngb is a single-domain monomer in all chordates examined to date, the Ngb ortholog in the purple sea urchin has been elaborated into a chimeric 16-domain protein through a combination of gene fusion and internal domain duplication (Bailly and Vinogradov 2008). As genome sequencing projects reveal globin gene structures for an increasing diversity of deuterostome and protostome taxa, it should be possible to evaluate the relative roles of internal domain duplication, domain shuffling, and gene fusion in the evolution of complex multidomain globins.

Supplementary Material

Supplementary tables S1–S4, figure S1, and data file S1 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We thank the Acorn Worm Genome Sequencing Consortium for making their data publicly available and the Human Genome Sequence Center from Baylor College of Medicine for providing access to the Acorn Worm Genome Project. We also thank Y. Satta and two anonymous reviewers for helpful comments and suggestions. F.G.H. acknowledges grant support from the National Science Foundation (EPS-0903787), J.F.S. acknowledges grant support from the National Institutes of Health/NHLBI (R01 HL087216 and HL087216-S1) and the National Science Foundation (IOS-0949931), J.C.O. acknowledges grant support from the Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT 11080181), D.H. is supported by a Marie Curie IEF Fellowship from the European Commission, and T.H. acknowledges support by a Johannes Gutenberg University internal funding grant (SRFN).

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol.* 215:403–410.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped Blast and PSI-Blast: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389–3402.

Anderson CL, Strope CL, Moriyama EN. 2011. SuiteMSA: visual tools for multiple sequence alignment comparison and molecular sequence simulation. *BMC Bioinformatics.* 12:184.

Bailly X, Vinogradov S. 2008. The bilatarian sea urchin and the radial starlet sea anemone globins share strong homologies with vertebrate neuroglobins. In: Bolognesi M, di Prisco G, Verde C, editors. Dioxygen binding and sensing proteins. New York: Springer Verlag. p. 191–201.

Baker S, Terwilliger N. 1993. Hemoglobin structure and function in the rat tailed sea cucumber, *Paracaudina chilensis*. *Biol Bull.* 185:115–122.

Berenbrink M. 2007. Historical reconstructions of evolving physiological complexity: O_2 secretion in the eye and swimbladder of fishes. *J Exp Biol.* 210:1641–1652.

Berenbrink M, Koldkjaer P, Kepp O, Cossins AR. 2005. Evolution of oxygen secretion in fishes and the emergence of a complex physiological system. *Science* 307:1752–1757.

Blank M, Wollberg J, Gerlach F, Reimann K, Roesner A, Hankeln T, Fago A, Weber RE, Burmester T. 2011. A membrane-bound vertebrate globin. *PLoS One* 6:e25292.

Bologna G, Yvon C, Duvaud S, Veuthey AL. 2004. N-Terminal myristoylation predictions by ensembles of neural networks. *Proteomics* 4:1626–1632.

Bonaventura C, Bonaventura J, Kitto B, Brunori M, Antonini E. 1976. Functional consequences of ligand-linked dissociation in hemoglobin from the sea cucumber *Molpadia arenicola*. *Biochim Biophys Acta.* 428:779–786.

Bonaventura J, Kitto GB. 1973. Ligand-linked dissociation of some invertebrate hemoglobins. In: Bolis L, Schmidt-Nielson K, Maddress SHP, editors. Comparative physiology. Amsterdam: North-Holland Publishing Company. p. 493–506.

Bourlat SJ, Juliusdottir T, Lowe CJ, et al. (14 co-authors). 2006. Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida. *Nature* 444:85–88.

Brittain T, O'Brien AJ, Wells RMG, Baldwin J. 1989. A study of the role of subunit aggregation in the expression of cooperative ligand binding in the haemoglobin of the lamprey *Mordacia mordax*. *Comp Biochem Physiol.* 93B:549–554.

Brittain T, Wells RM. 1986. Characterization of the changes in the state of aggregation induced by ligand binding in the hemoglobin system of a primitive vertebrate, the hagfish *Eptatretus cirrhatus*. *Comp Biochem Physiol A Comp Physiol.* 85:785–790.

Burmester T, Ebner B, Weich B, Hankeln T. 2002. Cytoglobin: a novel globin type ubiquitously expressed in vertebrate tissues. *Mol Biol Evol.* 19:416–421.

Burmester T, Hankeln T. 1999. A globin gene of *Drosophila melanogaster*. *Mol Biol Evol.* 16:1809–1811.

Burmester T, Hankeln T. 2009. What is the function of neuroglobin? *J Exp Biol.* 212:1423–1428.

Burmester T, Storf J, Hasenjäger A, Klawitter S, Hankeln T. 2006. The hemoglobin genes of *Drosophila*. *FEBS J.* 273:468–480.

Christensen AB, Colacino JM, Bonaventura C. 2003. Functional and biochemical properties of the hemoglobins of the burrowing brittle star *Hemipholis elongata* Say (Echinodermata, Ophiuroidea). *Biol Bull.* 205:54–65.

Dewilde S, Ebner B, Vinck E, et al. (15 co-authors). 2006. The nerve hemoglobin of the bivalve mollusc Spisula solidissima: molecular cloning, ligand binding studies, and phylogenetic analysis. *J Biol Chem.* 281:5364–5372.

Dröge J, Makałowski W. 2011. Phylogenetic analysis reveals wide distribution of globin X. *Biol Direct.* 6:54.

Ebner B, Burmester T, Hankeln T. 2003. Globin genes are present in *Ciona intestinalis*. *Mol Biol Evol.* 20:1521–1525.

Ebner B, Panopoulou G, Vinogradov SN, Kiger L, Marden MC, Burmester T, Hankeln T. 2010. The globin gene family of the cephalochordate amphioxus: implications for chordate globin evolution. *BMC Evol Biol.* 10:370.

Edgar RC. 2004. Muscle: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792–1797.

Fago A, Giangiacomo L, D'Avino R, Carratore V, Romano M, Boffi A, Chiancone E. 2001. Hagfish hemoglobins: structure, function, and oxygen-linked association. *J Biol Chem.* 276:27415–27423.

Fago A, Hundahl C, Malte H, Weber RE. 2004. Functional properties of neuroglobin and cytoglobin. Insights into the ancestral physiological roles of globins. *IUBMB Life* 56:689–696.

Fago A, Weber RE. 1995. The hemoglobin system of the hagfish *Myxine glutinosa*: aggregation state and functional properties. *Biochim Biophys Acta.* 1249:109–115.

Farazi TA, Waksman G, Gordon JI. 2001. The biology and enzymology of protein N-myristoylation. *J Biol Chem.* 276:39501–39504.

Goodman M, Czelusniak J, Koop BF, Tagle DA, Slightom JL. 1987. Globins: a case study in molecular phylogeny. *Cold Spring Harbor Symp Quant Biol.* 52:875–890.

Goodman M, Moore GW, Matsuda G. 1975. Darwinian evolution in the genealogy of haemoglobin. *Nature* 253:603–608.

Hankeln T, Ebner B, Fuchs C, et al. (23 co-authors). 2005. Neuroglobin and cytoglobin in search of their role in the vertebrate globin family. *J Inorg Biochem.* 99:110–119.

Hardison RC. 1996. A brief history of hemoglobins: plant, animal, protist, and bacteria. *Proc Natl Acad Sci U S A.* 93:5675–5679.

Hardison RC. 1998. Hemoglobins from bacteria to man: evolution of different patterns of gene expression. *J Exp Biol.* 201:1099–1117.

Hoffmann FG, Opazo JC, Storz JF. 2010. Gene cooption and convergent evolution of oxygen transport hemoglobins in jawed and jawless vertebrates. *Proc Natl Acad Sci U S A.* 107: 14274–14279.

Hoffmann FG, Opazo JC, Storz JF. 2011. Differential loss and retention of cytoglobin, myoglobin, and globin-E during the radiation of vertebrates. *Genome Biol Evol.* 3:588–600.

Hoffmann FG, Opazo JC, Storz JF. 2012. Whole-genome duplications spurred the functional diversification of the globin gene superfamily in vertebrates. *Mol Biol Evol.* 29:303–312.

Hoffmann FG, Storz JF, Gorr TA, Opazo JC. 2010. Lineage-specific patterns of functional diversification in the α- and β-globin gene families of tetrapod vertebrates. *Mol Biol Evol.* 27:1126–1138.

Hoogewijs D, De Henau S, Dewilde S, Moens L, Couvreur M, Borgonie G, Vinogradov SN, Roy SW, Vanfleteren JR. 2008. The *Caenorhabditis* globin gene family reveals extensive nematode-specific radiation and diversification. *BMC Evol Biol.* 8:279.

Hoogewijs D, Ebner B, Germani F, et al. (11 co-authors). 2012. Androglobin: a chimeric globin in metazoans that is preferentially expressed in mammalian testes. *Mol Biol Evol.* 29(4):1105–1114.

Kakar S, Hoffman FG, Storz JF, Fabian M, Hargrove MS. 2010. Structure and reactivity of hexacoordinate hemoglobins. *Biophys Chem.* 152:1–14.

Katoh K, Asimenos G, Toh H. 2009. Multiple alignment of DNA sequences with MAFFT. *Methods Mol Biol.* 537:39–64.

Katoh K, Toh H. 2008. Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform.* 9:286–298.

Kitto GB, Thomas PW, Hackert ML. 1998. Evolution of cooperativity in hemoglobins: what can invertebrate hemoglobins tell us? *J Exp Zool.* 282:120–126.

Lamy JN, Green BN, Toulmond A, Wall JS, Weber RE, Vinogradov SN. 1996. The giant hexagonal bilayer hemoglobins. *Chem Rev.* 96:3113–3124.

Lassmann T, Frings O, Sonnhammer ELL. 2009. Kalign2: high-performance multiple alignment of protein and nucleotide sequences allowing external features. *Nucleic Acids Res.* 37:858–865.

Lassmann T, Sonnhammer ELL. 2005. Automatic assessment of alignment quality. *Nucleic Acids Res.* 33:7120–7128.

Lassmann T, Sonnhammer ELL. 2006. Kalign, kalignvu and mumsa: web servers for multiple sequence alignment. *Nucleic Acids Res.* 34:W596–W599.

Lecomte JT, Vuletich DA, Lesk AM. 2005. Structural divergence and distant relationships in proteins: evolution of the globins. *Curr Opin Struct Biol.* 15:290–301.

Lesk AM, 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.

Linder ME, Deschenes RJ. 2007. Palmitoylation: policing protein stability and traffic. *Nat Rev Mol Cell Biol.* 8:74–84.

Maurer-Stroh S, Eisenhaber B, Eisenhaber F. 2002. N-terminal N-myristoylation of proteins: prediction of substrate proteins from amino acid sequence. *J Mol Biol.* 317:541–557.

Mauri F, Omnaas J, Davidson L, Whitfill C, Kitto GB. 1991. Amino acid sequence of a globin from the sea cucumber *Caudina (Molpadia) arenicola*. *Biochim Biophys Acta.* 1078:63–67.

McDonald GD, Davidson L, Kitto GB. 1992. Amino acid sequence of the coelomic C globin from the sea cucumber *Caudina (Molpadia) arenicola*. *J Protein Chem.* 11:29–37.

Mitchell DT, Ernst SR, Wu WX, Hackert ML. 1995. Three-dimensional structure of a hemichrome hemoglobin from *Caudina arenicola*. *Acta Crystallogr D Biol Crystallogr.* 51:647–653.

Notredame C, Higgins DG, Heringa J. 2000. T-coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol.* 302:205–217.

Opazo JC, Hoffmann FG, Storz JF. 2008. Genomic evidence for independent origins of β-like globin genes in monotremes and therian mammals. *Proc Natl Acad Sci U S A.* 105:1590–1595.

Pei J, Tang M, Grishin NV. 2008. Promals3d web server for accurate multiple protein sequence and structure alignments. *Nucleic Acids Res.* 36:W30–W44.

Perutz MF, Fermi G, Luisi B, Shaanan B, Liddington RC. 1987. Stereochemistry of cooperative mechanisms in hemoglobin. *Cold Spring Harb Symp Quant Biol.* 52:555–565.

Pesce A, Bolognesi M, Bocedi A, Ascenzi P, Dewilde S, Moens L, Hankeln T, Burmester T. 2002. Neuroglobin and cytoglobin. Fresh blood for the vertebrate globin family. *EMBO Rep.* 3:1146–1151.

Philippe H, Brinkmann H, Copley RR, Moroz LL, Nakano H, Poustka AJ, Wallberg A, Peterson KJ, Telford MJ. 2011.

Acoelomorph flatworms are deuterostomes related to *Xenoturbella*. *Nature* 470:255–258.

Projecto-Garcia J, Zorn N, Jollivet D, Schaeffer SW, Lallier FH, Hourdez S. 2010. Origin and evolution of the unique tetra-domain hemoglobin from the hydrothermal vent scale worm *Branchipolynoe*. *Mol Biol Evol.* 27:143–152.

Qiu Y, Maillett DH, Knapp J, Olson JS, Riggs AF. 2000. Lamprey hemoglobin. Structural basis of the Bohr effect. *J Biol Chem.* 275:13517–13528.

Ren J, Wen L, Gao X, Jin C, Xue Y, Yao X. 2008. CSS-Palm 2.0: an updated software for palmitoylation sites prediction. *Protein Eng Des Sel.* 21:639–644.

Resh MD. 1999. Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. *Biochim Biophys Acta.* 1451:1–16.

Riggs AF. 1991. Aspects of the origin and evolution of non-vertebrate hemoglobins. *Am Zool.* 31:535–545.

Riggs AF. 1998. Self-association, cooperativity and supercooperativity of oxygen binding by hemoglobins. *J Exp Biol.* 201:1073–1084.

Roesner A, Fuchs C, Hankeln T, Burmester T. 2005. A globin gene of ancient evolutionary origin in lower vertebrates: evidence for two distinct globin families in animals. *Mol Biol Evol.* 22:12–20.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.

Storz JF, Opazo JC, Hoffmann FG. 2011. Phylogenetic diversification of the globin gene superfamily in chordates. *IUBMB Life* 63:313–322.

Suzuki T. 1989. Amino acid sequence of a major globin from the sea cucumber *Paracaudina chilensis*. *Biochim Biophys Acta.* 998:292–296.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol.* 28:2731–2739.

Terwilliger NB. 1998. Functional adaptations of oxygen-transport proteins. *J Exp Biol.* 201:1085–1098.

Terwilliger NB, Terwilliger RC. 1978. Oxygen binding domains of a clam (*Cardita borealis*) extracellular hemoglobin. *Biochim Biophys Acta.* 537:77–85.

Terwilliger RC, Read KR. 1972. The hemoglobin of the holothurian echinoderm, *Molpadia oölitica* Pourtales. *Comp Biochem Physiol B.* 42:65–72.

Vinogradov SN. 1985. The structure of invertebrate extracellular hemoglobins (erythrocruorins and chlorocruorins). *Comp Biochem Physiol B.* 82:1–15.

Vinogradov SN, Hoogewijs D, Bailly X, Arredondo-Peter R, Gough J, Dewilde S, Moens L, Vanfleteren JR. 2006. A phylogenomic profile of globins. *BMC Evol Biol.* 6:31.

Vinogradov SN, Hoogewijs D, Bailly X, Arredondo-Peter R, Guertin M, Gough J, Dewilde S, Moens L, Vanfleteren JR. 2005. Three globin lineages belonging to two structural classes in genomes from the three kingdoms of life. *Proc Natl Acad Sci U S A.* 102:11385–11389.

Vinogradov SN, Hoogewijs D, Bailly X, Mizuguchi K, Dewilde S, Moens L, Vanfleteren JR. 2007. A model of globin evolution. *Gene* 398:132–142.

Vinogradov SN, Moens L. 2008. Diversity of globin function: enzymatic, transport, storage, and sensing. *J Biol Chem.* 283: 8773–8777.

Vinogradov SN, Walz DA, Pohajdak B, Moens L, Kapp OH, Suzuki T, Trotman CN. 1993. Adventitious variability? The amino acid sequences of nonvertebrate globins. *Comp Biochem Physiol B.* 106:1–26.

Weber RE, Vinogradov SN. 2001. Nonvertebrate hemoglobins: functions and molecular adaptations. *Physiol Rev.* 81:569–628.