原発性の骨格構成要素についての研究
The current understanding of the origin and evolution of the genetic cassette for the vertebrate skeletal system is reviewed. Molecular phylogenetic analyses of fibrillar collagen genes, which encode the main component of both cartilage and mineralized bone, suggest that genome duplications in vertebrate ancestors were essential for producing distinct collagen fibers for cartilage and mineralized bone. Several data indicate co-expression of the ancestral copy of fibrillar collagen with the SoxE and Runx transcription factors. Therefore, the genetic cassette may have already existed in protochordate ancestors, and may operate in the development of the pharyngeal gill skeleton. Accompanied by genome duplications in vertebrate ancestors, this genetic cassette may have also been duplicated and co-opted for cartilage and bone. Subsequently, the genetic cassette for cartilage recruited novel genetic material via domain shuffling. Aggrecan, acquired by means of domain shuffling, performs an essential role in cartilage as a shock absorber. In contrast, the cassette for bone recruited new genetic material produced by tandem duplication of the SPARC/osteonectin genes. Some of the duplicated copies of SPARC/osteonectin became secretory Ca-binding phosphoproteins (SCPPs) performing a central role in mineralization by regulating the calcium phosphate concentration. Comparative genome analysis revealed similar molecular evolutionary histories for the genetic cassettes for cartilage and bone, namely duplication of the ancestral genetic cassette and recruitment of novel genetic material.

Key words: ascidian, amphioxus, collagen, Sox, Runx, SCPP, aggrecan, domain shuffling, cartilage, bone

INTRODUCTION

The skeleton is one of the most important vertebrate innovations. Vertebrates have two distinct types of skeleton: cartilage and mineralized bone. These skeletal tissues are derived from distinct precursor cells, chondrocytes and osteocytes, respectively. Differentiation mechanisms in these cell types have been extensively examined, thus facilitating an understanding of the molecular evolutionary history of these mechanisms.

Differentiation of vertebrate chondrocytes is primarily controlled by Sox transcription factors, such as Sox9, Sox5, and Sox6, reviewed by Crombrugghe et al. (2000). These transcription factors directly regulate the genetic expression of cartilage matrix components, including col2a1, aggrecan, and probably col11a2 (Bell et al., 1997; Bridgewater et al., 1998; Sekiya et al., 2000). Similarly, the differentiation of osteocytes is controlled by transcription factors such as Runx2/CBFA1 (Ducy et al., 1997; Komori, 2002; Komori et al., 1997; Otto et al., 1997) and osterix/SP7 (Nakashima et al., 2002), and these transcription factors directly regulate the genetic expression of bone matrix, including col1a1, col1a2, and some other matrix genes (Kern et al., 2001; Komori, 2002).

Vertebrates produce these two types of skeletal tissue from several distinct cell lineages. For example, most of the cranial skeleton is derived from neural crest cells, while the vertebral and limb skeletons are derived from paraxial mesoderm and lateral plate mesoderm, respectively. Thus, after the genetic regulatory cassettes were established for cartilage and mineralized bone, these cassettes were transferred to some other cell lineages.

In this article, I present some new data on ascidian gene expression patterns and review our current understanding of how the genetic cassette for vertebrate skeletal tissues evolved at the molecular level.

GENETIC DUPLICATION OF FIBRILLAR COLLAGEN

As fibrillar collagen is the main matrix component of both cartilage and bone, its origin was the first target in examining the origin of the skeleton. Vertebrates have seven types of fibrillar collagens, encoded by 11 genes. Types I, V, and XXIV are components of mineralized bone, whereas types II, XI, and XXVII are components of cartilage. Molecular phylogenetic analyses of deuterostome fibrillar collagen genes, including one amphioxus gene and four ascidian genes, detected three clades (Fig. 1) (Wada et al., 2006), and thus the common ancestor of deuterostomes possessed three fibrillar collagen genes. Moreover, all of these clades include both one or more genes utilized for cartilage and one or more genes utilized for bone, indicating that genetic duplications in vertebrate ancestors produced subsets of genes for both cartilage and bone (Fig. 1).
Because the clade-A fibrillar collagen genes are linked with Hox clusters (Bailey et al., 1997), the fibrillar collagen gene duplications almost certainly accompanied two rounds of whole-genome duplications.

These molecular phylogenetic analyses imply that from the aspect of fibrillar collagen, the production of cartilage and bone as distinct skeletal elements did not occur before the genome duplications in vertebrates. Therefore, ascidians and amphioxus should possess the ancestral skeletal structure, such as existed before the differentiation of cartilage and bone. The notochord expresses all of the fibrillar collagen genes analyzed so far (Wada et al., 2006; Meulemans and Bronner-Fraser, 2007) and is a candidate for the origin of the vertebrate skeleton. Apart from the notochord, amphioxus and acorn worms also possess skeletal rods in the pharynx. Fibrillar collagen genes are expressed in cells surrounding the pharyngeal skeleton, a structure also suggested to have been the origin of the vertebrate skeleton (Rychel and Swalla, 2007; Hecht et al., 2008). The notochord and pharyngeal skeleton as possible origins of the vertebrate skeleton are discussed in more detail below.

**DOMAIN SHUFFLING AND THE EVOLUTION OF AGGREGCAN**

Aggrecan is the core protein of proteoglycan and is the most abundant non-collagenous cartilage protein. Aggrecan has an N-terminal V-set domain followed by two pairs of Xlink domains. Aggrecan binds, via Xlink domains, to hyaluronic acid (HA), a glycosaminoglycan (GAG) that is also abundant in cartilage (Doege et al., 1987). The middle of the aggrecan molecule contains regions that are modified with keratan sulfate and chondroitin sulfate, while the C-terminus contains the EGF, C-type lectin, and sushi domains (Doege et al., 1987; Fig. 2). The interaction between aggrecan and HA is supported by the link protein/HAPLN1, which also has Xlink domains but lacks the C-terminal domains and sequences modified with keratan sulfate and chondroitin sulfate (Doege et al., 1986). The complex consisting of HA, link protein/HAPLN1, and aggrecan provides the tensile strength for cartilage to absorb shocks and to resist compression in joints. Mutations in aggrecan or the link protein/HAPLN1 cause severe defects in cartilage and result in dwarfism (Watanabe et al., 1994; Watanabe and Yamada, 1999).

No genes similar in structure to aggrecan or link protein/HAPLN1 were found in either the amphioxus or the ascidian genomes. Instead, several genes were found that possess the Xlink domains (Kawashima et al., 2009). These observations suggest that aggrecan is a novel gene acquired by means of domain shuffling. Indeed, aggrecan was recovered in a comprehensive analysis of genes built up by means of domain shuffling in vertebrate lineages (Kawashima et al., 2009). Evolution of the aggrecan gene by domain shuffling was essential for vertebrate cartilage to possess characters distinct from bone and to function as a shock absorber.

One of the ascidian genes (Ci-Link1) encodes a single Xlink domain and an F5/8-type C domain, and is expressed in some blood cells in juveniles (Kawashima et al., 2009). In vertebrates, the Xlink domain is found not only in aggrecan 2009). These observations suggest that aggrecan is a novel gene acquired by means of domain shuffling. Indeed, aggrecan was recovered in a comprehensive analysis of genes built up by means of domain shuffling in vertebrate lineages (Kawashima et al., 2009). Evolution of the aggrecan gene by domain shuffling was essential for vertebrate cartilage to possess characters distinct from bone and to function as a shock absorber.

One of the ascidian genes (Ci-Link1) encodes a single Xlink domain and an F5/8-type C domain, and is expressed in some blood cells in juveniles (Kawashima et al., 2009). In vertebrates, the Xlink domain is found not only in aggrecan
and its paralogs, but also in CD44, TNFAIP6 (also known as TSG6), LYVE1, and stablin 1. These proteins contain single Xlink domains (Fig. 2B) and are involved in lymphocyte migration (Kzhyshkowska et al., 2006; Ponta et al., 2003). In molecular structure and function, Ci-Link1 more closely resembles these single-link proteins than it does aggrecan. The Xlink domain was probably a component of surface molecules on blood cells in protochordate ancestors and combined with other domains, such as the V-set domain, in a vertebrate ancestor to give a protein that was a component of cartilage — a unique vertebrate structure.

Cartilage acidic protein 1 (CRTAC1/CEP-68) was also created by combining domains, including the ASPIC-and-UnbV and calcium-binding EGF domains (Kawashima et al., 2009). Although cartilage acidic protein is specifically expressed in chondrocytes and the brain (Steck et al., 2007), its function has not been well elucidated. Because this gene is conserved in structure from teleosts to mammals, it likely performs an important role in cartilage formation.

**TANDEM DUPLICATIONS OF OSTEONECTIN/SPARC AND THE EVOLUTION OF BONE**

In addition to fibrillar collagen, mineralized vertebrate bone also possesses characteristic protein components. Some secretory Ca-binding phosphoproteins (SCPPs) are involved in the mineralization of vertebrate bone. Kawasaki et al. (2004) and Kawasaki and Weiss (2006) extensively analyzed the molecular evolution of SCPPs and presented data showing that they were derived from SPARC/osteonectin genes by means of tandem duplications. Ascidians possess a single copy of the SPARC/osteonectin genes by means of tandem duplications. Ascidians possess a D/E-rich sequence, which is utilized for regulating calcium phosphate concentrations in vertebrates. This suggests that the Ciona homolog is also involved in calcium phosphate regulation, which might have been co-opted for bone mineralization in vertebrates. A genomic survey of amphioxus genomic sequences revealed that amphioxus possesses two copies of SPARC, whose expression has not yet been analyzed.

**INSIGHTS FROM EXPRESSION OF THE TRANSCRIPTION FACTORS SOX AND RUNX2/CBFA1**

Transcriptional regulation has been extensively analyzed in both chondrocytes and osteocytes. In chondrocytes, genes of the SoxE subfamily (including Sox9 and Sox10) and SoxD subfamily (including Sox5 and Sox6) perform central roles in differentiation (Crombrugghe et al., 2000). These genes are expressed in chondrocytes derived from neural crest cells for pharyngeal chondrocytes, sclerome cells for axial chondrocytes, and lateral plate mesoderm cells for limb chondrocytes. The Sox genes directly regulate the transcription of several cartilage effector genes such as col2a1, col11a2, and aggrecan (Bell et al., 1997; Bridgewater et al., 1998; Sekiya et al., 2000). Both ascidians and amphioxus possess single copies for each of the SoxE and SoxD subfamilies. Neither SoxD nor SoxE is expressed in the ascidian notochord. Rather, the ascidian SoxD is expressed in some mesenchymal cells underlying the adhesive palp (Fig. 3B). Ascidian SoxE shows transient expression in pyrolic gland cells during metamorphosis (Fig. 3C, D). In amphioxus, SoxD is expressed in the notochord, while SoxE is expressed in the neural tube and dorsal wall of the anterior gut (Meulemans and Bronner-Fraser, 2007). Overall, Sox gene expression is quite divergent in protochordates, and overlapping expression with collagen genes is seen only in the amphioxus SoxD gene. It is noteworthy that SoxD is expressed in the lamprey notochord (Ohtani et al., 2008). Thus, SoxD may regulate the notochord expression of collagen genes in protochordates.

In contrast, overlapping expression of SoxE and fibrilar collagen genes has been reported in the pharyngeal skeletal rods of amphioxus and acorn worms (Hecht et al., 2008; Rychel and Swalla, 2007). Based on this observation, these studies imply that a genetic cassette, comparable to that for vertebrate cartilage, exists for the differentiation of the gill skeleton in amphioxus and acorn worms (Hecht et al., 2008; Rychel and Swalla, 2007).

**CBFA1/Runx2 and osterix/SP7** perform central functions in osteoblast differentiation (Komori, 2002). In addition, Runx2 is also involved in the early differentiation of chondrocytes (Stricker et al., 2002; Takeda et al., 2001). Ascidians and amphioxus possess single copies of cognates for CBFA1/Runx2, whose ancestral copy was duplicated into three vertebrate Runx genes (Runx1–3). Although embryonic expression has not been reported in either ascidians or amphioxus, amphioxus Runx was reported to be expressed in the adult pharyngeal skeleton (Hecht et al., 2008).

The molecular phylogenetic history of osterix/SP7 is not well elucidated. In the zinc finger gene

---

**Fig. 3.** Expression of SPARC and Sox genes in ascidians. (A) Expression of Ciona SPARC in the notochord of a *C. intestinalis* tailbud embryo. (B) Expression of Ciona SoxD in mesenchymal cells underlying the anterior adhesive palp in a *C. intestinalis* larva (arrow). (C) Expression of SoxE during metamorphosis in *Ciona intestinalis*, in a cell mass comprising presumptive pyloric gland cells (arrow) when the tail is being absorbed (arrowhead). (D) SoxE expression from the stomach to the pyloric gland in a *C. intestinalis* juvenile.
The notochord and cartilage have similar histological structures. Cells are surrounded by an extracellular matrix (ECM) rich in fibrillar collagen, although the ECM is much thicker in cartilage. Indeed, in chordates, the same fibrillar collagen is utilized in both the notochord and cartilage. In addition, SoxE and SoxD transcription factors are expressed in the vertebrate notochord (Barriónuevo et al., 2006; Smits and Lefebvre, 2003). Thus, one may reasonably conclude that the transcriptional regulatory system for the notochord was co-opted for cartilage. However, as noted above, except for amphioxus SoxD, no other Sox genes (neither ascidian SoxD and SoxE nor amphioxus SoxE) are expressed in the notochord. Therefore, although the same fibrillar collagen is utilized as a matrix component, gene expression might be regulated by distinct mechanisms. Alternatively, because SoxD and the fibrillar collagen genes are co-expressed in the lamprey notochord (Ohtani et al., 2008), genetic regulation of the fibrillar collagen gene by SoxD may have already existed in the protochordate notochord. Analysis of the genetic mechanisms regulating the notochord expression of collagen genes in species of amphioxus or lamprey might provide further information.

PHARYNGEAL SKELETON OF AMPHIHOXUS AND ACORN WORMS

In contrast to the notochord, some evidence supports that the genetic regulatory cassette for collagen gene expression operates in the differentiation of the pharyngeal cartilage in amphioxus and acorn worms (Rychel and Swalla, 2007; Hecht et al., 2008). In addition to Sox, Runx is also expressed in the pharyngeal endoderm (Rychel and Swalla, 2007; Hecht et al., 2008). However, some issues must be considered. First, the pharyngeal skeletons are acellular in both amphioxus and acorn worms. This is in contrast to the structure of vertebrate cartilage in which cells are embedded in the collagen-rich ECM. A second, and more serious, consideration is that we do not know which cells are responsible for production of the acellular skeleton. Although Rychel and Swalla (2007) suggested that endodermal cells secrete the ECM, we still do not have any clear evidence data. Note also that gene expression was examined in fully developed adults, in which we do not know whether production of the skeletal structure is ongoing. Therefore, we need to examine with more care the genetic mechanisms regulating the development of pharyngeal cartilage, and it is highly desirable to examine gene expression in specimens in which the skeleton is definitely being produced.

CONCLUDING REMARKS

In this article, I sought to elaborate on the origin of vertebrate cartilage and bone from the aspect of genetic evolution. In acorn worms and amphioxus, SoxE is expressed in the pharyngeal endoderm, which might be responsible for the pharyngeal skeleton (Rychel and Swalla, 2007; Hecht et al., 2008). Runx is expressed in the amphioxus pharyngeal endoderm (Hecht et al., 2008). These observations, together with collagen expression in those tissues, suggest that the genetic cascade of Sox and Runx regulation in the expression of collagen might already have existed in the protochordate pharyngeal skeleton, although some issues should be addressed as mentioned above. The function of SoxD should be further examined to assess the contribution of a putative SoxD-collagen cassette to the evolution of the vertebrate skeleton. Noting that the fibrillar collagen genes of protochordates represent the pre-duplicated copy from which respective copies arose for cartilage and bone (Wada et al., 2006), one might imagine that the genetic network was duplicated into two distinct cassettes after genome duplication. SoxD and SoxE were utilized mainly for chondrocyte differentiation, and Runx kept its function in the differentiation of both types of the vertebrate skeleton. This hypothesis can be tested by examining the transcriptional regulation of collagen in the pharyngeal skeleton of amphioxus or acorn worms.

Wada et al. (2006) argued that fibrillar collagen genes were co-opted for cartilage and bone independently in three distinct lineages. We reported that cis elements may have been able to evolve rapidly enough for three distinct genes to be regulated by chondrocyte transcription factors (Wada et al., 2006). Moreover, nearly the same process must have occurred with respect to the copies utilized for bone. However, if the expression of three ancestral copies of collagen was already regulated by Sox and/or Runx in the pharyngeal skeleton, we do not have to assume such a rapid evolution of cis elements. Rather, genetic regulation might have been subfunctionalized into chondrocytes and osteocytes after genome duplication.

Duplication of the genetic cassette was not sufficient for the vertebrate skeleton to have distinct characters. After the genetic cassette was duplicated, both cassettes recruited new genetic material as effectors. Aggrecan, which was built up by means of domain shuffling (aggrecan for cartilage) and tandem gene duplication (SCPPs for bone).
emerged from tandem duplications of SPARC/osteonectin, were recruited as new matrix components of bone, and this recruitment of SCPPs was essential for mineralization (Kawasaki et al., 2004; Kawasaki and Weiss, 2006). Genomic analysis of the origin of skeletal genes has revealed that our body evolved via a dramatic molecular evolutionary history involving duplications of the genetic cassette, together with the recruitment of novel genes that evolved via domain shuffling and gene duplications.

ACKNOWLEDGMENTS

I thank N. Satoh, Y. Satou, and K. Hirayama for providing Ciona intestinalis specimens through the National Bio-Resource Project of the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. This work was supported by a KAKENHI (Grant-in-Aid for Scientific Research) in the Priority Area “Comparative Genomics” from MEXT.

REFERENCES


Smits P, Lefebvre E (2003) Sox5 and Sox6 are required for notochord extracellular matrix sheath formation, notochord cell survival and development of the nucleus pulposus of intervertebral discs. Development 130: 1135–1148


(Received September 30, 2009 / Accepted October 7, 2009)