Evolutionary changes in lamin expression in the vertebrate lineage

Reimer Stick and Annette Peter
FB2 Biology/Chemistry, University of Bremen, Bremen, Germany

ABSTRACT
The nuclear lamina is involved in fundamental nuclear functions and provides mechanical stability to the nucleus. Lamin filaments form a meshwork closely apposed to the inner nuclear membrane and a small fraction of lamins exist in the nuclear interior. Mutations in lamin genes cause severe hereditary diseases, the laminopathies. During vertebrate evolution the lamin protein family has expanded. While most vertebrate genomes contain 4 lamin genes, encoding the lamins A, B1, B2, and LIII, the majority of non-vertebrate genomes harbor only a single lamin gene. We have collected lamin gene and cDNA sequence information for representatives of the major vertebrate lineages. With the help of RNA-seq data we have determined relative lamin expression levels for representative tissues for species of 9 different gnathostome lineages. Here we report that the level of lamin A expression is low in cartilaginous fishes and ancient fishes and increases toward the mammals. Lamin B1 expression shows an inverse tendency to that of lamin A. Possible implications for the change in the lamin A to B ratio is discussed in the light of its role in nuclear mechanics.

KEYWORDS
B-type lamins; hagfish lamins; lamin expression; lamina; lamin A; vertebrate evolution

Introduction

Nuclear lamin proteins are main constituents of the nuclear skeleton, the lamina. Lamins together with their associated proteins are involved in fundamental functions in nuclear mechanics and genome regulation. They provide anchoring sites for chromatin, aiding in spatial organization of interphase chromosomes and are involved in DNA replication, transcription, and nuclear positioning. Mutations in lamin genes and lamin-associated genes cause severe hereditary diseases, collectively called laminopathies or envelopopathies. For the human LMNA gene a particularly large number of mutations is known, which give rise to a broad spectrum of clinical manifestations.

Lamins are intermediate filament (IF) proteins. All IF proteins share a tripartite domain structure in which a short head and a longer tail domain flank a coiled-coil-forming domain of conserved sub-structure. Lamin proteins, in addition, harbor in their tail domain a nuclear localization signal (NLS), an immunoglobulin-like domain (Ig-domain) and a C-terminal CaaX-motif, required for membrane targeting. Lamins are the founding members of the IF family. Initially they were thought to be restricted to metazoans but a lamin-like protein has recently been characterized in the amoeba Dictyostelium. Recently an extensive sequence database search identified bona fide lamin sequences not only in further clades of the unikont branch but also in oomycetes, which belong to the bikont branch of the eukaryotes. Depending on the rooting of the eukaryotic tree this would place the appearance of the lamins at or very near the root of eukaryotic evolution. It remains to be clarified why lamins are absent in many branches of the eukaryotic tree and why bona fide lamin homologs have been identified so far in just a single sub-branch of the bikonts.

Most non-vertebrate genomes encode a single lamin gene. However, in some phylogenetically distant groups 2 lamin genes are present which most probably originated by independent group-specific gene duplications. The situation in vertebrates is more complex. Vertebrate genomes have undergone 2 rounds of whole genome duplication (2R WGD), which gave rise to
families of related proteins providing the basis for increasingly complex developmental programs.\textsuperscript{13-15} The major vertebrate clades diverged in the latest Precambrian and in the Paleozoic era, approximately 595–350 million years ago.\textsuperscript{16} Phylogenetic classification divides vertebrates into 2 monophyletic groups, the jawless (Agnatha) and the jawed vertebrates (Gnathostomata). The former house the lampreys and hagfish the latter all other vertebrate species. Gnathostomes are divided into the cartilaginous fish (Chondrichthyes: chimeras, sharks, skates and rays) and the bony fish (Osteichthyes). The latter are further grouped into the ray-finned fish (Actinopterygii) and the lobe-finned fish (Sarcopterygii) (Fig. 1A). The phylogenetic relations within the ray-finned fish, which house the bichirs, sturgeons,gars, bowfins, and teleost fish and the lobe-finned fish, housing the coelacanths, lungfish and tetrapods are shown in Fig. 1A.\textsuperscript{17} Tetrapods are further subdivided into amphibians, sauropsids (birds, crocodiles, turtles, lizards and snakes), and mammals which contain the monotremes, marsupials and placental mammals (Fig. 1B).\textsuperscript{16}

In jawed vertebrates 4 lamin sub-types are found, namely B1, B2, LIII, and A. They show a dichotomous relationship, in which sub-types B1 and B2 are more closely related to each other than to the sub-types LIII and A, which is consistent with their emergence in the course of 2 consecutive WGDs (Fig. 2).\textsuperscript{6,12} Beside sequence similarity, sub-type classification is based on features that are confined to individual lamin sub-types.\textsuperscript{6} A-lamins contain an extra domain within their tail domain. This A-specific domain harbors a proteolytic cleavage site necessary for processing of pre-lamin A to mature lamin A. Lamin LIII genes encode 2 alternatively spliced CaaX-encoding exons involved in membrane targeting.\textsuperscript{18,19} Our knowledge of vertebrate lamins is mainly based on studies in humans and model organisms like mouse, rat, chicken, Xenopus, and zebrafish. Sequence information from an increasing number of sequencing projects of non-model vertebrates enables us to assess the extent of evolutionary divergence within the vertebrate lamin family.

We collected lamin sequence data for representatives of all vertebrate lineages from publicly available databases. With the help of RNA-seq data we determined relative lamin expression levels for representative tissues in the different gnathostome lineages. We find that expression levels of lamin A have changed significantly in the course of evolution. Expression of lamin A is low in cartilaginous fish and increases toward the mammals. Lamin B1 expression shows an inverse tendency to that of lamin A. The observed changes in the lamin A to B-type ratio is discussed in the light of possible alterations of nuclear mechanics in the course of gnathostome evolution.

![Figure 1. Cladogram of extant vertebrate lineages and their respective lamin sub-types. (A) The cladogram shows the sister group relationship between the extant vertebrate lineages. (B) Cladogram of major tetrapod subgroups. Presence of sequence information for the respective lamins is indicated by the names of the corresponding lamin proteins. Lamin C is an alternative splice product of the LMNA gene in mammals. The tetrapod male germline-specific lamins are alternative splice products of their respective genes (see main text). LIII* in Fig. 1B indicates that LIII of platypus is a truncated, atypical lamin. Arrows mark the first appearance of the respective lamin subtypes in the course of evolution.](image-url)
Results and discussion

Vertebrate lamin proteins

Jawed vertebrates (Gnathostomata) represent the major group of living vertebrates. Their emergence from jawless vertebrates (Agnatha) was accompanied by major innovations, including jaws, paired appendages, and an adaptive immune system. Previous analyses of the lamprey (Petromyzon marinus), a member of the Agnatha, revealed that the lamprey genome encodes 3 functional lamin genes. One of these genes exhibits the features typical for a teleost/tetrapod LIII gene while the other 2 lack features that would allow assignment of orthology to teleost/tetrapod lamin genes. RNA-seq data of the hagfish, Eptatretus cirrhatus, a member of the sister group of the lampreys allowed us to assemble 3 complete lamin transcripts (Supplementary sequence information). The 3 hagfish lamins, LmnI, LmnII, and LIII resemble those of the lamprey. In a phylogenetic tree of vertebrate lamins the hagfish lamins group together with the respective sequences of the lamprey (Fig. 2). These data support our previous findings that lamin LIII is present in both, agnathans and gnathostomes, while orthology of the other 2 agnathan lamins to one of the gnathostome lamin sub-types cannot be assigned.

Lamin LmnI of both the lamprey and the hagfish lacks a CaaX-motif. Moreover, also in the hagfish there is no evidence for the presence of an A-type lamin. This strengthens the view that the lamna gene is a synapomorphic character state of the gnathostome lineage and that the evolution of the lamin family went separate paths in the 2 vertebrate lineages.

Lamin sequence information for several major gnathostome clades, e.g. bichirs, sturgeons, bowfins, and lungfish, was either incomplete or entirely missing. Moreover, automatic lamin gene annotations present in public databases often need manual re-evaluation. To achieve a more comprehensive picture of the gnathostome lamin family we searched publicly available databases for lamin sequences using lamin sequences from various species as queries. Additionally, we performed de novo assembly of lamin transcripts from RNA-seq reads available in the NCBI sequence read.
Cartilaginous fish (Chondrichthyes) are the oldest group of extant jawed vertebrates. They diverged about 457 (443–464) Ma.22 Chondrichthyes are divided into 2 subgroups, the Holocephali (chimeras) and the Elasmobranchia (sharks, skates, and rays). For the chimera Callorhinchus milii and for the small-spotted catshark Scyliorhinus canicula we obtained complete sequence information for lamins B1, B2, LIII, and A (Supplementary Sequence Information). Orthology to the respective teleost/tetrapod lamin sub-types B1, B2, LIII, and A can be inferred based on sequence similarity. Sequence identity values of individual sub-types between species are generally higher than identity values between sub-types within a particular species and in phylogenetic tree reconstructions the chondrichthyan lamins group within their respective sub-groups, in contrast to the 3 agnathan lamins, which all group outside the 4 subtypes (Fig. 2).21 The presence of an additional exon with conserved lamin A-specific sequence features in the A-lamins of both chondrichthyan species further indicates that the lamin A gene had already evolved with the emergence of the Chondrichthyes. The exon-intron pattern of vertebrate lamin genes is highly conserved.6 This holds also for the lamin genes of Callorhinchus and Scyliorhinus, with 2 notable exceptions. (i) The Lmnb2 genes of both species lack the lamin B2-specific intron typical for all teleost/tetrapod species. This intron interrupts the region encoding linker 1 of the rod domain.23 It is already present in the Lmnb2 gene of the gar, Lepisosteus oculatus (see below). (ii) The lamprey LIII gene as well as the LIII genes of the teleosts/tetrapods encode 2 alternatively spliced CaaX-encoding exons.18,19,21 The relative order of these 2 exons is highly conserved, i.e. the LIIIb-CaaX exon, which encodes additional membrane targeting motifs, is always located upstream of the LIIIa-CaaX exon. In contrast, Callorhinchus expresses only a single lamin LIII variant; a LIIIb-specific exon is missing in the respective genomic region. Whether loss of this exon is limited to the holocephalian branch of the Chondrichthyes has to await genomic sequence information for elasmobranch species. So far we have not found a LIIIb-specific transcript for an elasmobranch species.

The second group of the jawed vertebrates are the bony fish (Osteichthyes). These comprise the ray-finned fish (Actinopterygii) and the lobe-finned fish (Sarcopterygii). Actinopterygii are further divided into 5 clades: bichirs (Cladistia), sturgeons (Chondrostei), gars (Lepisoteiformes), bowfins (Amiiformes), and teleost fish (Teleostei). These 5 clades are each monophyletic, however the relationship among the clades has not been fully solved.17,24 Teleosts, which include several model organisms for genomic and developmental biology form the by far most species-rich subgroup of ray-finned fish, while the remainder 4 groups together comprise only approximately 50 species.24,25 However, from an evolutionary perspective these ‘ancient fish’ are of special interest since branching of the 4 clades took place early in evolution.24,25 We used RNA-seq data from one representative each of the 4 clades of ancient fish for de novo assembly of lamin transcripts. The selected species were Polypterus senegalus, Acipenser baerii, Lepisosteus oculatus, and Amia calva. We were able to assemble cDNAs for all 4 lamin sub-types of the 4 species. Additional search of EST and RNA-seq databases provided evidence that both splice variants of lamin LIII are expressed in all 4 ancient fish. For the spotted gar (Lepisosteus oculatus) sufficient genomic sequence information is available, that allowed us to establish the conserved genomic organization of 3 of its lamin genes, namely Lmnb1, Lmnb2, and LIII. It could also be proven that the relative order of the 2 alternatively spliced exons in the LIII gene is conserved in this species at the base of the osteichthyan lineage.

Genomes of quite several teleost fish have been sequenced and annotated.26 Sequences retrieved from Ensembl were inspected and manually edited where necessary.21 The common ancestor of all teleost fish has experienced a WGD. It is termed the teleost-specific WGD.27 Interestingly, the lamin genes experienced different fates after the teleost-specific WGD. Only one copy of the lamin B1, B2, and LIII genes has been retained after the teleost-specific WGD. In contrast, many teleost genomes encode 2 lamin A genes. The lamin A-subtypes, termed lamin A1 and A2, differ in the length of their A-specific domain; they show overall sequence identity values between 75 and 80%. Of note, expression of A1 markedly exceeds that of A2 (results not shown), pointing to sub-functionalization of the 2 teleost A lamins. In some teleost branches additional WGDs have occurred. The genome of the
Atlantic salmon (Salmo salar) encodes 3 lamin A genes, 2 of which, lamin A1–1 and A1–2, are 96% identical and probably result from the Salmonid-specific WGD. The third A-lamin, A2 shares only 73% sequence identity with the 2 A1 lamins. Remarkably, the 2 nearly identical lamins A1–1 and A1–2 are both highly expressed, while expression of A2 is significantly lower.

Coelacanths (Actinistia), lungfish (Dipnoi), and tetrapods (Tetrapoda) form the group of lobe-finned fish (Sarcopterygii). Although coelacanths and lungfish were abundant in the Devonian, only 2 species of coelacanths and 6 species of lungfish are extant. Due to morphological features, which they share with their fossil ancestors, they are commonly referred to as ‘living fossils’. For the coelacanth Latimeria chalumnae a whole genome assembly and transcriptome data are available. The complete set of the 4 lamin genes is present in L. chalumnae.

From an evolutionary perspective the presence of lungs and other morphological and physiologic characters make lungfish particularly interesting. Several of these features are considered as prerequisites for the transition from aquatic to terrestrial life. Moreover, lungfish are the closest extant relatives to the tetrapods. However, their genome sizes are by far the largest among vertebrates, as their c-values are up to hundred times larger than c-values of other vertebrates, which hampers whole genome assembly. Transcriptome assembly however allowed us to ensemble complete cDNA sequences for lamin B1, B2, and A of the lungfish Protopterus annectens. Within more then 800 million paired-end reads we did not detect any sequences related to lamin LIII. This is particularly notable since the pool of reads originated from various tissues, including brain and male and female gonads in which expression of LIII is particularly high in other species. It must await additional sequence information to decide whether this indicates that lungfish have lost the lamin LIII gene or whether LIII transcripts have so far gone undetected.

Lamin sequences and their respective genes have been described previously for all major groups of tetrapods. Three aspects of tetrapod lamin evolution are particularly noteworthy. Marsupials and placental mammals express an additional lamin: lamin C. Lamin C is an alternative splice product of the Lmna gene. Secondly, marsupial and placental mammals have lost the lamin LIII gene. In contrast, the platypus, a member of the monotreme mammals, has retained the LIII gene, albeit in a truncated version, giving rise to an atypical lamin. Its expression is confined to the testis. Third, atypical male germ line-specific lamins have also been found in other tetrapods. They assemble into short rather than long filaments and are thought to alter the properties of the nuclear lamina of spermatogenic cells. All result from alternative splicing, each from a different gene e.g., lamin LIV is a variant of the LIII gene in Xenopus, lamin C2 is a variant of the Lmna gene in all mammals, and lamin B3 is a variant of the Lmnb2 gene in placental mammals.

Sequence similarities between vertebrate lamin proteins

Previous sequence alignments and phylogenetic tree building of vertebrate lamins revealed a dichotomous relationship of the 4 lamin sub-types where sub-types B1 and B2 are more closely related to each other than to the group made of sub-types LIII and A. To estimate the degree of sequence variation within each of the 4 gnathostome lamin sub-types multiple sequence alignments were done for each sub-type separately. Lamin proteins of 14 species, each representing a different vertebrate lineage, were included in these alignments. From the resulting 4 multiple alignments amino acid identity and similarity values for each pair of lamin proteins were calculated (Supplementary Tables 1–4). Identity values were grouped in 10% intervals and the number of the respective identity values within each interval was plotted (Fig. 3). The average sequence identity values of the different lamin sub-types are 66% for sub-type A and nearly 70% for the sub-types B1 and B2. Variations within these 3 sub-types are relatively small (Fig. 3). In contrast, lamin LIII sequences diverge to a much larger extent; the average sequence identity of LIII proteins is from 17% to nearly 20% lower compared with that of the 3 other sub-types. This might indicate a relaxation from functional constraints for this lamin sub-type.

Expression analysis of lamins in representative gnathostome species

Lamin expression analyses have been most extensively performed for a large number of human tissues using immunohistochemistry combined with immunoblotting, and immunoprecipitation. Similar but
less extensive studies were done for several model organisms like mouse, rat, chicken, zebrafish, and Xenopus. \textsuperscript{18,40-42} From these data it can be deduced that lamin A is absent in early development but is predominant in most adult mammalian tissues. However, low levels of lamin A/C protein are present in murine embryonic stem cells and in the inner cell mass of blastocysts. \textsuperscript{43} For non-model vertebrates information about lamin protein expression is almost completely missing. This is mainly due to lack of suitable anti-lamin antibodies for these species. RNA-seq sequencing projects have provided a wealth of sequence information for representatives of the major groups of extant vertebrates. With adequate calculation methods these RNA-seq data can be used to calculate the relative abundance of lamin transcripts. Despite the fact that transcript levels do not necessarily reflect protein levels in all cases, they may indicate possible trends in tissue-specific expression of lamins in the course of vertebrate evolution. To this end we have selected species from different gnathostome groups for a comparative study. Criterion for the selection of the respective species was the existence of a complete set of lamin sequences and the availability of RNA-seq data for at least 6 different tissues, namely brain, heart, kidney, liver, testis, and ovary. For 9 species, i.e., Callorhinichus, Lepisosteus, Amia, Danio, Xenopus, Gallus, Ornithorhynchus, Monodelphis, and Homo RNA-seq reads were mapped onto the respective lamin open reading frames. Mapped reads were extracted for the individual tissues and the relative abundance of the different lamin transcripts was calculated. We used the transcript per million (TPM) calculation method introduced by Wagner et al. instead of the reads per kilobase per million reads (RPKM) method. \textsuperscript{44,45} In RPKM calculations inconsistencies across samples may arise due to differences in sequencing depth since in the RPKM method normalization is done with the total number of reads of a particular RNA-seq run. \textsuperscript{44} The TPM calculation eliminates this inconsistency. It provides relative abundances and allows to compare results obtained from different sequencing projects. \textsuperscript{45} This is a prerequisite for the present study, for which data from many different sequencing projects had to be recruited. Reads were mapped against the lamin open reading frames excluding the 5'- and 3'-UTRs, to avoid a potential source of variability that could be caused by the UTR.

**Figure 3.** Distribution of amino acid identity values of the 4 gnathostome lamin sub-types. Lamin protein sequences of 14 gnathostome species, each representing a different gnathostome lineage, i.e., Callorhinichus milii (Holocephali), Scyliorhinus canicula (Elsambranchii), Polypterus senegalus (Cladistia), Acipenser baeri (Chondrostei), Lepisosteus oculatus (Lepisosteiformes), Amia calva (Amiiformes), Danio rerio (Teleosteii), Latimeria chalumnae (Actinistia), Protoperus annectens (Dipnoi), and 5 tetrapods: Xenopus tropicalis (Amphibia), Gallus gallus (Sauropsida), Ornithorhynchus anatinus (Prototheria), Monodelphis domestica (Metatheria), and Homo sapiens (Eutheria) were aligned. Multiple alignments were done for each sub-type separately using MultAlin software. Amino acid identity values were calculated using the Ident/Sim function of the sequence manipulation suite (Supplementary Tables 1–4). Identity values were binned in 10% intervals. The number of identity values within each interval is plotted against the respective interval. Note that the total number of identity values for sub-type LIII is lower than the numbers for B1, B2, and A respectively (55 versus 91 values), due to the lack of sequence information for LII of Protoperus and the absence of LII genes in Meta- and Eutheria (Monodelphis and Homo).
The respective TPM values can be converted directly into percent values for each lamin sub-type. This allowed us to estimate the relative expression levels of the individual lamin sub-types and to compare these values between tissues of the same species as well as between evolutionarily distantly related species. Where possible, data obtained in this way were compared with previously published RNA-seq data and were found to closely match. Validation of the RNA-seq data by independent methods as for example by RT-qPCR was not possible in the frame of this analysis since suitable RNA samples for the respective species were not available to us. However, RNA-seq quantifications are widely used and have been validated by e.g., RT-qPCR in several reports. To more clearly contrast the expression of lamin A with that of the 3 B-type lamins B1, B2, and LIII the values of the B-type lamins were cumulated and presented as stacked diagrams (Fig. 4).

The most obvious finding is that lamin A is expressed at very low levels in the chimera Callorhinichus and in the 2 ancient fishes Lepisosteus and Amia, respectively. In striking contrast to the low level of lamin A expression in cartilaginous fish and the group of the ancient fish, in most mammalian tissues lamin A transcripts are by far the predominant lamin transcripts. This holds for all placental mammals that were analyzed, i.e., man (Fig. 4) and another 5 primate species and the mouse (results not shown), and similarly for the opossum, Monodelphis domestica (Fig. 4), and, with the exception of the testis, also for platypus, Ornithorhynchus anatinus. Platypus is the only mammal retaining an LIII gene. It encodes an atypical lamin protein. Its expression is restricted to the testis where its transcripts make up more than 75% of all lamin transcripts (Fig. 4). In Callorhinchus only very few LIII transcripts have been detected, these are found in the ovary (Fig. 4). Expression of the LIII gene in the other non-mammalian vertebrates is found in nearly all tissues, with high expression levels in testes and ovaries and, in some species, also in the brain. Based on earlier studies, which mainly focused on early amphibian development, the expression of lamin LIII was thought to be mainly restricted to oocytes, early embryos, and a few somatic cell types. With the help of RT-PCR LIII transcripts were detected in oocytes, testis, and, to a lesser extent, in the heart but not in brain, liver, and kidney of Xenopus. Differences between these reports and the RNA-seq data presented here, might be due to a lower sensitivity obtained in the previous RT-PCR

Figure 4. Lamin transcript levels in tissues of representatives of 9 gnathostome groups. RNA-seq reads were mapped to the respective lamin open reading frames, mapped reads were extracted, and TPM values were calculated for each tissue. TPM values were converted into percent values for the individual lamins. The percent TPM values for B-type lamins B1 (dark green), B2 (light green), and LIII (blue) are shown as stacked columns each right to the lamin A columns (red). Corresponding tissues are vertically aligned. The tissue type is indicated at the bottom. Species names and the respective systematic groups are given above each panel. For phylogenetic relationships see Fig. 1.
experiments that were done primarily to detect the LIV splice variant of the LIII gene. The RNA-seq data point to a more widespread expression of LIII in many species.

Expression ratios of B1 relative to B2 transcripts do vary to some extent between tissues and species. However, minor differences should be interpreted with some caution, as the data were collected from different studies. Despite these possible limitations, the data presented here make 2 trends particularly clear: there is a marked increase in lamin A expression along the course of gnathostome evolution and this increase has taken place mainly in expense of lamin B1 expression (Fig. 4). This trend becomes even more obvious when the TPM values for the 6 different tissues are pooled and compared with each other (Fig. 5). Plotting the pooled TPM of individual lamin sub-types against the relative phylogenetic position of the respective species and statistical analysis reveals that these correlations are significant with Pearson’s coefficient of $r = 0.92842$ for the increase in lamin A expression and $r = -0.86893$ for the decrease of lamin B1 expression (Fig. 5B). In contrast, changes in the expression levels of lamin B2 and LIII do not correlate with phylogenetic position. For lamin B2 a Pearson’s coefficient of $r = -0.49312$ is obtained and the corresponding value for LIII is 0.2688 (Supplementary Fig. 1A and 1B). Limited by the available sequence information this study includes only a limited number of species. While each of these species represents a separate systematic group, not all of these groups are of the same phylogenetic rank. In addition, the abscissa does not exactly correspond to an evolutionary time scale. Insertion of data from additional representatives of vertebrate groups might modify this graph in the future and will have some influence on correlation coefficients, but this will not change the principal result. With the data at hand, we cannot prove yet whether the

![Figure 5](image-url)

**Figure 5.** Changes of lamin expression in the course of gnathostome evolution. (A) Data presented in Fig. 4 are presented in a different way: Percent lamin TPM values of the 6 tissues (brain, heart, kidney, liver, testis, and ovary) were combined and depicted as percent lamin TPM for each species. Values for B-type lamins are shown as stacked columns. Color code is as in Figure 4. Species abbreviations are: Cm: Callorhinchus milii, Lo: Lepisosteus oculatus, Ac: Amia calva, Dr: Danio rerio, Xt: Xenopus tropicalis, Gg: Gallus gallus, Oa: Ornithorhynchus anatinus, Md: Monodelphis domestica, Hs: Homo sapiens. (B, C) Line plots showing percent TPM values of lamin A and lamin B1, respectively, plotted against the order in which the respective species occurs in vertebrate evolution (compare Fig. 1). Note that the distances on the abscissa are arbitrary, they do not reflect the evolutionary time scale.
increase in lamin A expression is a gradual one or whether changes coincide with the appearance of certain systematic groups.

The strength of this finding critically depends on the number of species analyzed for each systematic group. While a sufficiently large number of RNA-seq data are available for mammalian species, which are all in good agreement with each other, comparatively few data are available for more basal vertebrates. Only few additional RNA-seq data are currently available for other cartilaginous fish, e.g., the small-spotted catshark. These data closely resemble those of the chimera. Nonetheless, we must await further RNA-sequencing results for more species of cartilaginous fish to underpin our findings. The stickleback Gasterosteus aculeatus and the African cichlids Oreochromis niloticus and Astatotilapia burtoni, all members of the teleost fish, show lamin A expression values in accordance with their respective phylogenetic position (results not shown).

**Functional aspects**

In light of the high sequence conservation between lamin A proteins of the diverse gnathostome species it is unlikely that the biological function of lamin A has been subject to extensive changes during gnathostome evolution. In general, sequence differences reflect the phylogenetic distances rather than going parallel to the observed differences in the expression level.

Tissue mechanics are influenced by the intracellular skeleton as well as by extracellular matrix components. The nuclear lamina provides mechanical stability to the nucleus and lamin filaments are linked via integral membrane proteins of the inner and outer nuclear membranes to the cytoskeleton. Lamin A is of particular importance for nuclear mechanics. There is growing evidence that the viscoelastic properties of the nucleus mainly depend on lamin A. Lamin A filaments are thicker than filaments made from B-type lamins and can form multilayered filament bundles superimposed on the network of B-lamins. The majority of the mutations of the lamin A gene affect nuclear rigidity and deformability. Lamin A expression in mammalian tissues scales with stiffness of the respective tissue. Higher A to B-type lamin ratios result in stiffer tissues.

**Materials and methods**

**Sequence retrieval and sequence assembly**

Lamin sequences were either retrieved from NCBI (http://www.ncbi.nlm.nih.gov), Ensembl (http://www.ensembl.org), or the Elephant Shark Genome Project (http://esharkgenome.imcb.a-star.edu.sg). Sequences not present in these databases were de novo assembled from RNA-seq reads present in the NCBI sequence read archive (SRA) (https://www.ncbi.nlm.nih.gov/sra/). SRA sequences were retrieved via the European nucleotide archive (ENA) (http://www.ebi.ac.uk/ena) and assembled using CLC Genomic Workbench 7.5 software (CLC Bio, Aarhus, Denmark) as described. Accession numbers of the proteins, genes, cDNA, and SRA sequences, respectively, are listed in Supplementary Table 5 together with information, how individual lamin sequences were assembled.

**Protein alignment, protein analysis, and phylogenetic reconstruction**

Protein sequences were aligned using MultAlin software (http://multalin.toulouse.inra.fr/multalin/) with the Blosum62–12 amino acid substitution matrix. Sequence identity and similarity values were calculated using the Ident and Sim function of the sequence manipulation suite (http://www.bioinformatics.org/SMS/index.html) with the following groups of similar amino acids for the similarity calculations: ILV, FWY, KRH, DE, GAS, TNQM.

Evolutionary analysis was conducted in MEGA7. Details are given in the legend to Fig. 2.

**Expression analysis**

Expression analysis was done by mapping RNA-seq reads (Supplementary Table 5, column SRR for TPM) to the respective open reading frames using the tool
‘map reads to contigs’ of the CLC Genomic Workbench with the parameters: No masking, mismatch cost = 2, insertion cost = 3, deletion cost = 3, length fraction = 0.95–1.0, similarity fraction = 0.99–1.0. Mapped reads were extracted and the TPM values were calculated for individual tissues from the respective RNA-seq runs. The TPM values were directly converted into percent values for the individual lamins. Calculations were done with Excel Mac2011 (Microsoft Corporation, Redmond, WA, USA), graphics were done with Apple Numbers, version 2.3 (Apple Inc. Cupertino, CA, USA) and Adobe Illustrator CS5 (Adobe Systems Software Ireland Ltd., Dublin, Ireland). Statistical analysis was done with Past 3.1.1.

Abbreviations

- EST: expressed sequence tag
- IF: intermediate filament
- Ma: million years ago
- RNA-seq: RNA sequencing
- RT-qPCR: quantitative reverse transcription PCR
- SRA: sequence read archive
- TPM: transcripts per million
- UTR: untranslated region
- WGD: whole genome duplication

Acknowledgments

We would like to thank Uwe Nehls for generously providing us access to computer facilities and Thomas Hurek for support in the use of the bioinformatics programs and for critical discussion of the results. We thank Wendy Gerber and Irm Huttenlauch for critically reading the manuscript.

References

[1] Gruenbaum Y, Foisner R. Lamins: Nuclear Intermediate Filament Proteins with Fundamental Functions in Nuclear Mechanics and Genome Regulation. Annu Rev Biochem 2015; 84:131–64
[2] Worman HJ. Nuclear lamins and laminopathies. J Pathol 2012; 226:316–25; PMID:21953297; https://doi.org/10.1002/path.2999
[3] Worman HJ, Bonne G. “Laminopathies: a wide spectrum of human diseases. Exp Cell Res 2007; 313:2121–33; PMID:17467691; https://doi.org/10.1016/j.yexcr.2007.03.028
[4] Zaremba-Czogala M, Dubinska-Magiera M, Rzepceki R. Laminopathies: the molecular background of the disease and the prospects for its treatment. Cell Mol Biol Lett 2011; 16:114–48; https://doi.org/10.2478/s11658-010-0038-9
[5] Herrmann H, Aebi U. Intermediate filaments: molecular structure, assembly mechanism, and integration into functionally distinct intracellular Scaffolds. Annu Rev Biochem 2004; 73:749–89; PMID:15189158; https://doi.org/10.1146/annurev.biochem.73.011303.073823
[6] Peter A, Stick R. Evolution of the lamin protein family: What introns can tell. Nucleus 2012; 3:34–59; PMID:22156746; https://doi.org/10.4161/ncl.18927
[7] Weber K, Plessmann U, Ulrich W. Cytoplasmic intermediate filament proteins of invertebrates are closer to nuclear lamins than are vertebrate intermediate filament proteins; sequence characterization of two muscle proteins of a nematode. EMBO J 1989; 8:3221–7; PMID:2583097
[8] Dodemont H, Riemer D, Weber K. Structure of an invertebrate gene encoding cytoplasmic intermediate filament (IF) proteins: implications for the origin and the diversification of IF proteins. EMBO J 1990; 9:4083–94; PMID:2249666
[9] Döring V, Stick R. Gene structure of nuclear lamin LII of Xenopus laevis; a model for the evolution of IF proteins from a lamin-like ancestor. EMBO J 1990; 9:4073–81; PMID:2249665
[10] Krüger A, Batsios P, Baumann O, Luckert E, Schwarz H, Stick R, Meyer I, Gräf R. Characterization of NE81, the first lamin-like nucleoskeleton protein in a unicellular organism. Mol Biol Cell 2012; 23:360–70; PMID:22090348; https://doi.org/10.1091/mbc.E11-07-0595
[11] Kollmar M. Polyphyly of nuclear lamin genes indicates an early eukaryotic origin of the metazoan-type intermediate filament proteins. Sci Rep 2015; 5:10652; PMID:26024016; https://doi.org/10.1038/srep10652
[12] Peter A, Stick R. Evolutionary aspects in intermediate filament proteins. Curr Opin Cell Biol 2015; 32C:48–55; https://doi.org/10.1016/j.cceb.2014.12.009
[13] Ohno S. Gene duplication and the uniqueness of vertebrate genomes circa 1970–1999. Semin Cell Dev Biol 1999; 10:517–22; https://doi.org/10.1006/scdb.1999.0332
[14] Holland PW, Garcia-Fernandez J, Williams NA, Sidor A. Gene duplications and the origins of vertebrate development. Development 1994;125–33.
[15] Kasahara M. The 2R hypothesis: an update. Current opinion in immunology 2007; 19:547–52; PMID:17707623; https://doi.org/10.1016/j.coi.2007.07.009
[16] Hedges SB, Marin J, Suleski M, Paymer M, Kumar S. Tree of life reveals clock-like speciation and diversification. Molecular biology and evolution 2015; 32:835–45; PMID:25739733; https://doi.org/10.1093/molbev/msv037
[17] Betancur-R R, Broughton RE, Wiley EO, Carpenter K, Lopez JA, Li C, Holcroft NI, Arcila D, Sanciangco M, Cureton II JC, et al. The tree of life and a new classification of bony fishes. PLoS currents 2013; 18:5.
[18] Hofemeister H, Kuhn C, Franke WW, Weber K, Stick R. Conservation of the gene structure and membrane-targeting signals of germ cell-specific lamin LIII in...
amphibians and fish. European journal of cell biology 2002; 81:51-60; PMID:11893082; https://doi.org/10.1078/0171-9335-00229

[19] Hofmeister H, Weber K, Stick R. Association of prenylated proteins with the plasma membrane and the inner nuclear membrane is mediated by the same membrane-targeting motifs. Molecular biology of the cell 2000; 11:3233-46; PMID:10982413; https://doi.org/10.1091/mbc.11.9.3233

[20] Venkatesh B, Lee AP, Ravi V, Maurya AK, Lian MM, Swann JB, Ohta Y, Flajnik MF, Sutoh Y, Kasahara M, et al. Elephant shark genome provides unique insights into gnathostome evolution. Nature 2014; 505:174-9; PMID:24402279; https://doi.org/10.1038/nature12826

[21] Schilf P, Peter A, Hurek T, Stick R. Lamins of the sea lamprey (Petromyzon marinus) and the evolution of the vertebrate lamin protein family. Eur J Cell Biol 2014; 93 (7):308-21; PMID:25059907

[22] Inoue JG, Miya M, Lam K, Tay BH, Danks JA, Bell J, Walker TJ, Venkatesh B. Evolutionary origin and phylogeny of the modern holocephalans (Chondrichthyes: Chimaeriformes): a mitogenomic perspective. Mol Biol Evol 2010; 27:2576-86; PMID:20551041; https://doi.org/10.1093/molbev/msq147

[23] Zewe M, Hoger TH, Fink T, Lichter P, Krohge G, Franke WW. Gene structure and chromosomal localization of the murine lamin B2 gene. Eur J Cell Biol 1991; 56:342-50; PMID:1802718

[24] Near TJ, Miya M. Ray-finned fishes (Actinopterygii). In: Hedges SB, Kumar S, eds. The tree of life. Oxford: Oxford University Press, 2009.

[25] Near TJ, Eytan RI, Dornburg A, Kuhn KL, Moore JA, Davis MP, Wainwright PC, Friedman M, Smith WL. Resolution of ray-finned fish phylogeny and timing of diversification. Proceedings of the National Academy of Sciences of the United States of America 2012; 109:13698-703; PMID:22869754; https://doi.org/10.1073/pnas.1206625109

[26] Yates A, Akanni W, Amode MR, Barrell D, Billis K, Carvalho-Silva D, Cummings C, Clapham P, Fitzgerald S, Gil L, et al. Ensembl 2016. Nucleic Acids Research 2016; 44: D710-6; PMID:26687719; https://doi.org/10.1093/nar/gkv1157

[27] Glasauer SM, Neuhaus SC. Whole-genome duplication in teleost fishes and its evolutionary consequences. Molecular genetics and genomics : MGG 2014; 289:1045-60; PMID:25092473; https://doi.org/10.1007/s00438-014-0889-2

[28] Amemiya CT, Alfoldi J, Lee AP, Fan S, Philippe H, MacCallum I, Braasch I, Manousaki T, Schneider I, Rohner N, et al. The African coelacanth genome provides insights into tetrapod evolution. Nature 2013; 496:311-6; PMID:23598338; https://doi.org/10.1038/nature12027

[29] Gregory TR. Animal genome size database. http://www.genomesize.com, 2016.

[30] Biscotti MA, Gerdol M, Canapa A, Forconi M, Olmo E, Pallavicini A, Barucca M, Schartl M. The Lungfish Transcriptome: A Glimpse into Molecular Evolution Events at the Transition from Water to Land. Scientific Rep 2016; 6:21571; PMID:26908371; https://doi.org/10.1038/srep21571

[31] Peter A, Khandekar S, Deakin JE, Stick R. A peculiar lamin in a peculiar mammal: Expression of lamin LIII in platypus (Ornithorhynchus anatinus). Eur J Cell Biol 2015; 94:522-30; PMID:26213206; https://doi.org/10.1016/j.ejcb.2015.07.001

[32] Hesse M, Zimek A, Weber K, Magin TM. Comprehensive analysis of keratin gene clusters in humans and rodents. Eur J Cell Biol 2004; 83:19-26; PMID:15085952; https://doi.org/10.1074/ehu.071-9335-00354

[33] von Moeller F, Barendziak T, Apte K, Goldberg MW, Stick R. Molecular characterization of Xenopus lamin LIV reveals differences in the lamin composition of sperms in amphibians and mammals. Nuclear 2010; 1:85-95; PMID:21327107; https://doi.org/10.4161/ncli.1.1.10517

[34] Alsheimer M, Benavente R. Change of karyoskeleton in teleosts and its evolutionary consequences. J Mol Evol 2013; 76(6):311-6; PMID:2347:1260419; PMID:25613900; https://doi.org/10.1126/science.1260419

[35] Alsheimer M, von Glashenapp E, Hock R, Benavente R. Architecture of the nuclear periphery of rat pachyten spermatoocytes: distribution of nuclear envelope proteins in relation to synaptonemal complex attachment sites. Mol Biol Cell 1999; 10:1235-45; PMID:10198069; https://doi.org/10.1091/mbc.10.4.1235

[36] Furukawa K, Hotta Y. cDNA cloning of a germ cell specific lamin B3 from mouse spermatocytes and analysis of its function by ectopic expression in somatic cells. EMBO J 1993; 12:97-106; PMID:8094052.

[37] Furukawa K, Inagaki H, Hotta Y. Identification and cloning of an mRNA coding for a germ cell-specific A-type lamin in mice. Exp Cell Res 1994; 212:426-30; PMID:8187835; https://doi.org/10.1006/excr.1994.1164

[38] Broers JL, Machiels BM, Kuijpers HJ, Smedts F, van den Kieboom R, Raymond Y, Ramaekers FC. A- and B-type lamins in the ejaculated sperm. Exp Cell Res 1996; 228:181-8; PMID:8912709; https://doi.org/10.1006/excr.1996.0315

[39] Lukanidi P, Gareiß M, Eberhardt K, Kruger E, Kandert S, Bohm C, Zentgraf H, Müller CR, dabauvalle MC. Emerin expression in early development of Xenopus laevis. Eur J Cell
[43] Eckersley-Maslin MA, Bergmann JH, Lazar Z, Spector DL. Lamin A/C is expressed in pluripotent mouse embryonic stem cells. Nucleus 2013; 4:53-60; PMID:18516045; https://doi.org/10.4161/nucl.23384

[44] Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 2008; 5:621-8; PMID:18516045; https://doi.org/10.1038/nmeth.1226

[45] Wagner GP, Kin K, Lynch VJ. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theor Biosci 2012; 131:281-5; https://doi.org/10.1007/s12064-012-0162-3

[46] Fang Z, Cui X. Design and validation issues in RNA-seq experiments. Brief Bioinfor 2011; 12:280-7; PMID:21498551; https://doi.org/10.1093/bib/bbr004

[47] Brawand D, Soumillon M, Necsulea A, Julien P, Csardi G, Harrigan P, Weier M, Liechti A, Aximu-Petri A, Kircher M, et al. The evolution of gene expression levels in mammalian organs. Nature 2011; 478:343-8; PMID:22012392; https://doi.org/10.1038/nature10532

[48] Necsulea A, Soumillon M, Warnefors M, Liechti A, Daish T, Zeller U, Baker JC, Grützner F, Kaessmann H. The evolution of IncRNA repertoires and expression patterns in tetrapods. Nature 2014; 505:635-40; PMID:24463194; https://doi.org/10.1038/nature12943

[49] Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M. The transcriptional landscape of the yeast genome defined by RNA sequencing. Science 2008; 320:1344-9; PMID:18451266; https://doi.org/10.1126/science.1158441

[50] Camarena L, Bruno V, Euskirchen G, Poggio S, Snyder M. Molecular mechanisms of ethanol-induced pathogenesis revealed by RNA-sequencing. PLoS Pathog 2010; 6: e1000834; PMID:20368969; https://doi.org/10.1371/journal.ppat.1000834

[51] Stick R, Hauser P. Changes in the nuclear lamina composition during early development of Xenopus laevis. Cell 1985; 41:191-200; PMID:3995581; https://doi.org/10.1016/0028-6874(85)90073-X

[52] Benavente R, Krohne G. Change of karyoskeleton during spermatogenesis of Xenopus: expression of lamin LIV, a nuclear lamina protein specific for the male germ line. Proc Natl Acad Sci U S A 1985; 82:6176-80; PMID:3862126; https://doi.org/10.1073/pnas.82.18.6176

[53] Dahl KN, Kahn SM, Wilson KL, Discher DE. The nuclear envelope lamina network has elasticity and a compressibility limit suggestive of a molecular shock absorber. J Cell Sci 2004; 117:4779-86; PMID:15331638; https://doi.org/10.1242/jcs.01357

[54] Swift J, Discher DE. The nuclear lamina is mechanoresponsive to ECM elasticity in mature tissue. J Cell Sci 2014; 127:3005-15; PMID:24963133; https://doi.org/10.1242/jcs.149203

[55] Zwerger M, Ho CY, Lammerding J. Nuclear mechanics in disease. Annu Rev Biomed Eng 2011; 13:397-428; PMID:21756143; https://doi.org/10.1146/annurev-bioeng-071910-124736

[56] Banerjee A, Rathee V, Krishnaswamy R, Bhattacharjee P, Ray P, Sood AK, Sengupta K. Viscoelastic behavior of human lamin A proteins in the context of dilated cardiomyopathy. PloS one 2013; 8:e83410; PMID:24386194; https://doi.org/10.1371/journal.pone.0083410

[57] Dutta S, Bhattacharyya M, Sengupta K. Implications and Assessment of the Elastic Behavior of Lamins in Lamino-pathies. Cells 2016; 5:pii:E37.

[58] Goldberg MW, Huttenlauch I, Hutchison CJ, Stick R. Filaments made from A- and B-type lamins differ in structure and organization. J Cell Sci 2008; 121:215-25; PMID:18187453; https://doi.org/10.1242/jcs.022020

[59] Vandebergh W, Bossuyt F. Radiation and functional diversification of alpha keratins during early vertebrate evolution. Mol Biol Evol 2012; 29:995-1004; PMID:22046002; https://doi.org/10.1093/molbev/msr269

[60] Corpet F. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res 1988; 16:10881-90; PMID:2849754; https://doi.org/10.1093/nar/16.24.10881

[61] Stothard P. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. BioTechniques 2000; 28:1102.

[62] Kumar S, Stecher G, Tamura K. MEGAS: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 2016; 33:1870–4; PMID:27004904; https://doi.org/10.1093/molbev/msw054

[63] Hammer Ø, Harper DAT, Ryan PD. Paleontological statistics software package for education and data analysis. Palaeontologia Electronica 2001; 4:9.

[64] Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. Computer applications in the biosciences : CABIOS 1992; 8:275-82; PMID:1633570