

In contrast to the nematode and fruit fly all 9 intron positions of the sea anemone lamin gene are conserved in human lamin genes

Alexander Zimek*, Klaus Weber

Max Planck Institute for Biophysical Chemistry, Am Fassberg, D-37077 Goettingen, Germany

Received 25 November 2007; received in revised form 11 January 2008; accepted 16 January 2008

Abstract

We identified the single gene for nuclear lamin in the genome draft of the sea anemone *Nematostella vectensis*, a member of the cnidaria, a very old metazoan phylum. The gene consists of 10 exons and 9 introns. Strikingly all 9 intron positions are conserved in the human lamin B genes, which have only 1 (lamin B1) or 2 (lamin B2) additional introns. Using the information on neighboring genes we propose that the human lamin B1 gene on chromosome 5 is the true homolog of the *Nematostella* lamin gene, while the lamin B2 gene on chromosome 19 arose during vertebrate evolution. In marked contrast to this conservation of gene structure are the results in the rapidly evolving genomes of *Drosophila* and *Caenorhabditis elegans*. Here the lamin genes have much fewer introns and these occur often at novel positions. In the single nematode lamin gene and the *Drosophila* lamin Dmo gene no intron position coincides with an intron in the sea anemone lamin gene. © 2008 Elsevier GmbH. All rights reserved.

Keywords: Bilateria; Cnidaria; Cytoplasmic intermediate filaments; Gene organization; Metazoan phylogeny; *Nematostella vectensis*; Nuclear lamin

Introduction

Eumetazoa cover the numerous phyla of bilateral animals and the radiata, which consist of the cnidaria and the ctenophora. Recent evidence suggests that the genomes of the nematode and the fruit fly are not typical for invertebrates because they arose under strong evolutionary drift, resulting in drastic changes in gene organization. Using genomic data and expressed sequence tags from the marine annelid *Platynereis* (Raible et al., 2005) showed that about two-thirds of the human introns predate the bilateral radiation, but were to a large extent lost from insect and nematode genomes. A comparison of coding exon sequences confirmed the

ancestral nature of the annelid and human genes. A drawback of this work was that it covered only a very small fraction of the genome. Therefore the results of the complete genome draft of the sea anemone *Nematostella* as a representative of the oldest eumetazoan phylum, the cnidaria, are even more striking. The sea anemone genome is complex, with a gene repertoire, exon-intron structures and large-scale gene linkages much more similar to vertebrates than to that of flies and nematodes (Putnam et al., 2007).

The lamin provides stability to the nuclear envelope in *Drosophila* (Lenz-Bohme et al., 1997), and the single lamin of the nematode *C. elegans* is essential for embryogenesis (Liu et al., 2000). Lamin mutants are responsible for various human diseases often called laminopathies (reviewed in Schirmer and Foissner, 2007; Worman and Bonne, 2007). Nuclear lamins are not restricted to bilateral animals, but also occur in cnidarians. A prominent band in

*Corresponding author. Tel.: +49 551 201 1157; fax: +49 551 201 1175.

E-mail address: azimek@gwdg.de (A. Zimek).

electrophoretically separated polypeptides of cytoskeletal preparations of *Hydra* and *Tealia* was identified as lamin. Tryptic peptides from this band were instrumental in obtaining complete cDNA clones of cnidarian lamins (Erber et al., 1999).

Because of our interest in metazoan lamin evolution (Erber et al., 1999; Riemer et al., 1993, 1995, 2000) we identified the lamin gene of the sea anemone genome and show that it consists of 10 exons and 9 introns. Strikingly all 9 intron positions are conserved in the human B-type lamin genes which have either 1 (lamin B1) or 2 (lamin B2) additional introns. Using the information on neighboring genes in the sea anemone and human genome, we propose that the human lamin B1 gene is the direct homolog of the cnidarian lamin gene while the B2 lamin gene arose later during vertebrate evolution.

Materials and methods

We analyzed the genome assembly of the sea anemone (*Nematostella vectensis*). The *Nematostella* genome assembly 1.0 was produced by the Joint Genome Institute (JGI). It has a 7.8-fold coverage in 10,804 scaffolds with a total sequence length of 356 Mbp of sequence. Roughly half of the genome is contained in 181 scaffolds all at least 473 kb in length. The lamin sequence was compared with genomic sequences from the urochordate *Ciona intestinalis* (March 2005 (ci2) assembly), the sea urchin *Strongylocentrotus purpuratus* (September 2006 (strPur2) assembly) and human (March 2006 human reference sequence; NCBI Build 36.1). Those are available on the website of the University of California, Santa Cruz (<http://genome.ucsc.edu/>). Extracted sequences were analyzed with the GCG sequence analysis software (GCG Version 11.1, Accelrys Inc., San Diego, CA). Graphics were created with the open source DTP software Scribus 1.3.3.9 (<http://www.scribus.net>). The *Nematostella* lamin was compared with other metazoan lamin genes using the gap program of the GCG package. The lamin sequences used were: *Caenorhabditis elegans* (X74027), *D. melanogaster* Dmo (X07278), human Lamin AC (NM_170707), human lamin B1 (NM_005573), human lamin B2 (NM_032737), *Hydra attenuata* (AJ005934) and *Strongylocentrotus purpuratus* (NM_214500).

Results and discussion

Conservation of lamin gene structure in *Nematostella* and man

The sea anemone has a single lamin gene, which is situated on scaffold 145 (nt 322747–325106). It consists

of 10 exons and 9 introns (Fig. 1). Strikingly all 9 intron positions are conserved in the human lamin B genes, which have either 1 (lamin B1) or 2 (lamin B2) additional introns (Fig. 2). This situation is drastically different in the nematode *C. elegans* and in *Drosophila* (Fig. 2). The single lamin gene of *C. elegans* displays none of the 9 intron positions conserved between *Nematostella* and the human lamin genes. Similarly the generally expressed *Drosophila* lamin gene Dmo retains none of the 9 intron positions common to the sea anemone and man. The second *Drosophila* lamin gene, lamin C, which is expressed in differentiated cells (Riemer et al., 1995) keeps only 2 of the 9 intron positions conserved between *Nematostella* and human lamins.

As shown in Fig. 1 the introns of the sea anemone are strikingly shorter than their human counterparts. Thus the first intron is with 1.2 kb the largest intron of the sea anemone gene while its human counterpart in lamin B1 measures 26.9 kb. The other introns of the sea anemone gene are only 0.2–0.8 kb in length.

The neighboring genes identify human lamin B1 as the homolog of the *Nematostella* lamin

One of the striking features of the sea anemone genome was the discovery of blocks of DNA that have essentially the same complement of genes as in the human genome. Individual genes may have swapped places but often they remained linked (Putnam et al., 2007). We have explored this feature of the sea anemone genome for the lamin gene. Directly upstream of the lamin gene on scaffold 145 lies the homolog for the membrane-associated RING finger protein 3 (nt 322747–325106). This also describes a single gene in the human genome. Strikingly it lies on chromosome 5 directly upstream of the gene for human lamin B1 (nt 126140732–126200608). Thus, it seems that the human lamin B1 gene is the true homolog of the sea anemone lamin while the B2 gene on chromosome 19 arose later with the vertebrate lineage.

Other metazoan genomes

Using a cDNA earlier reported (Holy et al., 1995) we explored the recently established genome draft of the sea urchin *Strongylocentrotus purpuratus* and located the lamin gene on scaffold 96 (nt 12338–34282). As shown in Fig. 2 the sea urchin lamin gene keeps 8 of the 9 intron positions conserved between the sea anemone and man. It has acquired one additional intron, which is slightly moved downstream from the intron common to all human lamin genes. Our results support the notion that the sea urchin genome did not evolve under strong drift. We also looked at the genome draft of the tunicate



Fig. 1. Comparison of the genes for the sea anemone lamin (lower lines) and the human lamin B1 (upper lines). Intron sequences are indicated by lower case letters and the size of the introns is given in numbers of nucleotides (kb). Although the introns of the sea anemone are shorter, the two genes have 9 identically placed introns. One additional intron is formed in the human gene. Identical nucleotides are marked by vertical lines.

Ciona intestinalis and located the lamin gene on chromosome scaffold 12p (nt 906430–916232). As already known from cDNA cloning and the ultrastructure of the recombinant protein the gene is unique in that it lacks the Ig-like subdomain in the middle of the tail region (Karabinos et al., 2003, 2004; Riemer et al., 2000). In spite of this deletion that removes two usually conserved intron positions the gene still shares 7 intron positions with the sea anemone and human lamin genes and has acquired three additional intron positions (Fig. 2). Currently we do not know whether the unique deletion in the tunicate gene can be taken as an indication of pronounced evolutionary drift of the genome. A broader analysis of many tunicate

genes may settle this problem (see also Raible et al., 2005).

We also determined the sequence identities of the various lamin proteins versus the sea anemone lamin. The human lamins B1 and B2 show 40% identity and a similar value of 41% holds for the sea urchin protein. In contrast the *C. elegans* lamin shows only 30% identity. While these results parallel the different gene structures (see above and Fig. 2) the *Drosophila* lamin protein Dmo shows with 37% a value between those of *C. elegans* and the human lamins. Surprisingly the two cnidarian lamins i.e. the lamins of *Hydra* and the sea anemone share 43% identity indicating a substantial drift between hydrozoa and anthozoa. This observation

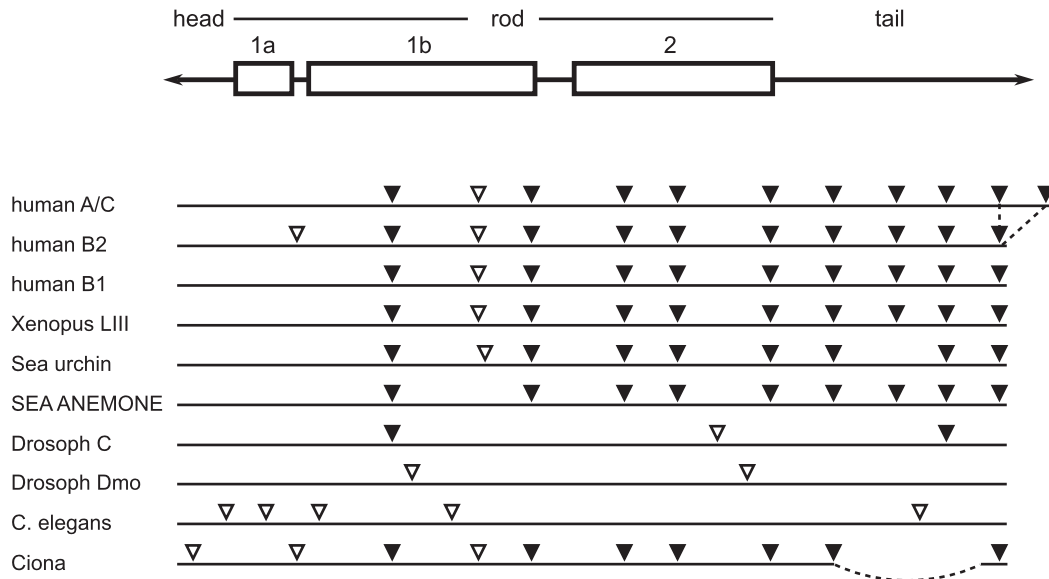


Fig. 2. Summary of intron positions in metazoan lamin genes. The tripartite domain organization of the nuclear lamins is shown at the top. Filled and open triangles mark common and unique intron positions, respectively. The gene of the sea anemone is derived from Fig. 1. Other lamin genes shown are the three human genes B1, B2 and A/C. The additional exon and intron of lamin A (Lin and Worman, 1993) are inserted into the penultimate exon of the lamin B1 gene. Note that lamin A/C is only expressed in differentiated cells (Rober et al., 1989). Also shown is the oocyte-specific *Xenopus* lamin LIII (Döring and Stick, 1990) which is lacking in mammalian genomes. Note the high degree of conservation of intron positions in the sea anemone and vertebrate genes. Thus all human lamin genes conserve the intron positions of the sea anemone gene. This is strikingly lost in the genes from *Drosophila* (Dmo and C) and *C. elegans*. The ubiquitously expressed Dmo and the single lamin gene of the nematode show no common intron position. Dmo and C are from (Osman et al., 1990) and (Riemer et al., 1995), respectively. For the *C. elegans* gene see (Riemer et al., 1993). For the lamin genes of the sea urchin *Strongylocentrotus purpuratus* and the tunicate *Ciona intestinalis* (see text). Note that the sea urchin lamin gene lacks one of the 9 intron positions conserved in the human and sea anemone genes. Its additional intron is slightly moved from the position of a common human intron. The wavy dotted line in the *Ciona* genome marks a unique deletion, which removes the Ig-like subdomain in the middle of the tail region (see Karabinos et al., 2004).

is interesting in view of our previous finding that the *Hydra* lamin gene has only 3 introns (Erber et al., 1999). Their positions correspond to the last three introns of the sea anemone gene. While we cannot exclude a mistake in our previous PCR mapping of the *Hydra* gene (Erber et al., 1999) a further large-scale gene comparison of *Hydra* and *Nematostella* may clarify the relation.

Search for genes encoding cytoplasmic intermediate filament (IF) proteins

Nuclear lamins and cytoplasmic IF proteins are related in sequence, domain organization and assembly features. This relation is particularly striking for the currently known cytoplasmic IF proteins from invertebrates (Dodemont et al., 1990; Döring and Stick, 1990). For a recent review of invertebrate cytoplasmic IF proteins see Zimek and Weber (2002). We therefore screened the genome of the sea anemone with probes corresponding to a keratin II (mouse K8) and a keratin I (human K18). We also used the full cDNA sequence of the *Saccoglossus* IF protein as well as the sea anemone

lamin and the IF consensus sequence around coil 1a. Additional probes were a type III (mouse vimentin) and a type IV (mouse NF-M) cDNA. None of these probes led to the identification of an IF gene other than that for the nuclear lamin. Thus it seems that the sea anemone lacks genes encoding a cytoplasmic IF protein. A somewhat similar situation was earlier recognized in *Drosophila* and other arthropods. Here bundles of microtubules occupy cellular positions, which in other phyla display bundles of cytoplasmic IF proteins. For discussion see (Karabinos et al., 2001).

The sea anemone genome is complex, with a gene repertoire, exon–intron structures, and large-scale gene linkage more similar to vertebrates than to that of *Drosophila* and *C. elegans* (Putnam et al., 2007). This point is well illustrated by our results on the genes for nuclear lamins. Thus there is a true need for a typical genome from a protostomic invertebrate, i.e. a lophotrochozoan protostomic animal (Aguinaldo et al., 1997). This could well be the complete genome of the marine annelid *Platynereis dumerilii* for which promising, but very preliminary data were collected (Raible et al., 2005). We expect from such a genome that the conservation of gene organization documented for the

sea anemone and human lamins extend to the annelids. Currently it is unclear whether the rapidly evolving genomes of *Drosophila* and *C. elegans* are a feature of all ecdysozoa. Although the features are shared by the insects *Drosophila*, *Anopheles* and *Apis* (Raible et al., 2005), it is not known whether they apply to the entire large phylum of the arthropods. Our results argue against a rapidly evolving genome of the sea urchin but tentatively raise this possibility for the tunicates. Nuclear lamins are identified in various bilateral animals and the cnidaria. Formal proof for their existence in the ctenophoras is still lacking. It will probably need a whole genome draft of a sponge to decide whether lamin genes already exist in the most primitive animals, the porifera.

References

- Aguinaldo, A.M., Turbeville, J.M., Linford, L.S., Rivera, M.C., Garey, J.R., Raff, R.A., Lake, J.A., 1997. Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* 387, 489–493.
- Dodemont, H., Riemer, D., Weber, K., 1990. Structure of an invertebrate gene encoding cytoplasmic intermediate filament (IF) proteins: implications for the origin and the diversification of IF proteins. *EMBO J.* 9, 4083–4094.
- Döring, V., Stick, R., 1990. Gene structure of nuclear lamin LIII of *Xenopus laevis*; a model for the evolution of IF proteins from a lamin-like ancestor. *EMBO J.* 9, 4073–4081.
- Erber, A., Riemer, D., Hofemeister, H., Bovenschulte, M., Stick, R., Panopoulou, G., Lehrach, H., Weber, K., 1999. Characterization of the *Hydra* lamin and its gene: a molecular phylogeny of metazoan lamins. *J. Mol. Evol.* 49, 260–271.
- Holy, J., Wessel, G., Berg, L., Gregg, R.G., Schatten, G., 1995. Molecular characterization and expression patterns of a B-type nuclear lamin during sea urchin embryogenesis. *Dev. Biol.* 168, 464–478.
- Karabinos, A., Schmidt, H., Harborth, J., Schnabel, R., Weber, K., 2001. Essential roles for four cytoplasmic intermediate filament proteins in *Caenorhabditis elegans* development. *Proc. Natl. Acad. Sci. USA* 98, 7863–7868.
- Karabinos, A., Schünemann, J., Meyer, M., Aebi, U., Weber, K., 2003. The single nuclear lamin of *Caenorhabditis elegans* forms in vitro stable intermediate filaments and paracrystals with a reduced axial periodicity. *J. Mol. Biol.* 325, 241–247.
- Karabinos, A., Zimek, A., Weber, K., 2004. The genome of the early chordate *Ciona intestinalis* encodes only five cytoplasmic intermediate filament proteins including a single type I and type II keratin and a unique IF-annexin fusion protein. *Gene* 326, 123–129.
- Lenz-Bohme, B., Wismar, J., Fuchs, S., Reifegerste, R., Buchner, E., Betz, H., Schmitt, B., 1997. Insertional mutation of the *Drosophila* nuclear lamin Dmo gene results in defective nuclear envelopes, clustering of nuclear pore complexes, and accumulation of annulate lamellae. *J. Cell Biol.* 137, 1001–1016.
- Lin, F., Worman, H.J., 1993. Structural organization of the human gene encoding nuclear lamin A and nuclear lamin C. *J. Biol. Chem.* 268, 16321–16326.
- Liu, J., Rolef Ben-Shahar, T., Riemer, D., Treinin, M., Spann, P., Weber, K., Fire, A., Gruenbaum, Y., 2000. Essential roles for *Caenorhabditis elegans* lamin gene in nuclear organization, cell cycle progression, and spatial organization of nuclear pore complexes. *Mol. Biol. Cell* 11, 3937–3947.
- Osman, M., Paz, M., Landesman, Y., Fainsod, A., Gruenbaum, Y., 1990. Molecular analysis of the *Drosophila* nuclear lamin gene. *Genomics* 8, 217–224.
- Putnam, N.H., Srivastava, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., Terry, A., Shapiro, H., Lindquist, E., Kapitonov, V.V., Jurka, J., Genikhovich, G., Grigoriev, I.V., Lucas, S.M., Steele, R.E., Finnerty, J.R., Technau, U., Martindale, M.Q., Rokhsar, D.S., 2007. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* 317, 86–94.
- Raible, F., Tessmar-Raible, K., Osoegawa, K., Wincker, P., Jubin, C., Balavoine, G., Ferrier, D., Benes, V., de Jong, P., Weissenbach, J., Bork, P., Arendt, D., 2005. Vertebrate-type intron-rich genes in the marine annelid *Platynereis dumerilii*. *Science* 310, 1325–1326.
- Riemer, D., Dodemont, H., Weber, K., 1993. A nuclear lamin of the nematode *Caenorhabditis elegans* with unusual structural features; cDNA cloning and gene organization. *Eur. J. Cell Biol.* 62, 214–223.
- Riemer, D., Stuurman, N., Berrios, M., Hunter, C., Fisher, P.A., Weber, K., 1995. Expression of *Drosophila* lamin C is developmentally regulated: analogies with vertebrate A-type lamins. *J. Cell Sci.* 108, 3189–3198.
- Riemer, D., Wang, J., Zimek, A., Swalla, B.J., Weber, K., 2000. Tunicates have unusual nuclear lamins with a large deletion in the carboxyterminal tail domain. *Gene* 255, 317–325.
- Rober, R.A., Weber, K., Osborn, M., 1989. Differential timing of nuclear lamin A/C expression in the various organs of the mouse embryo and the young animal: a developmental study. *Development* 105, 365–378.
- Schirmer, E.C., Foisner, R., 2007. Proteins that associate with lamins: many faces, many functions. *Exp. Cell Res.* 313, 2167–2179.
- Worman, H.J., Bonne, G., 2007. “Laminopathies”: a wide spectrum of human diseases. *Exp. Cell Res.* 313, 2121–2133.
- Zimek, A., Weber, K., 2002. The gene for a cytoplasmic intermediate filament (IF) protein of the hemichordate *Saccoglossus kowalevskii*; definition of the unique features of chordate IF proteins. *Gene* 288, 187–193.