Uniquely designed nuclear structures of lower eukaryotes
Masaaki Iwamoto¹, Yasushi Hiraoka¹,²,³ and Tokuko Haraguchi¹,²,³

The nuclear structures of lower eukaryotes, specifically protists, often vary from those of yeasts and metazoans. Several studies have demonstrated the unique and fascinating features of these nuclear structures, such as a histone-independent condensed chromatin in dinoflagellates and two structurally distinct nuclear pore complexes in ciliates. Despite their unique molecular/structural features, functions required for formation of their cognate molecules/structures are highly conserved. This provides important information about the structure–function relationship of the nuclear structures. In this review, we highlight characteristic nuclear structures found in lower eukaryotes, and discuss their attractiveness as potential biological systems for studying nuclear structures.

Addresses
¹ Advanced ICT Research Institute, National Institute of Information and Communications Technology (NICT), Kobe, Japan
² Graduate School of Frontier Biosciences, Osaka University, Suita, Japan
³ Graduate School of Science, Osaka University, Toyonaka, Japan

Corresponding author: Haraguchi, Tokuko (tokuko@nict.go.jp)

Current Opinion in Cell Biology 2016, 40:66–73
This review comes from a themed issue on Cell nucleus
Edited by Ulrike Kutay and Orna Cohen-Fix
For a complete overview see the Issue and the Editorial
Available online 7th March 2016
http://dx.doi.org/10.1016/jceb.2016.02.019
0955-0674/© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Introduction
Most of our knowledge about the ‘typical’ nuclear structures of eukaryotes, such as chromatin, the nuclear envelope (NE), and the nuclear pore complex (NPC), has been acquired from studies in classical model organisms such as yeasts and humans, which belong to the group Opisthokonta (Figure 1). However, recent technical advances in genomic DNA sequencing and protein analysis have facilitated the study of organisms other than Opisthokonta, unveiling astonishingly unusual nuclear structures that had not previously been observed. One example is the ‘dinokaryon’ chromatin structure of dinoflagellates (Alveolata, SAR) that lacks bulk histones. Other examples are the NEs in the Filastera Capaspora owczarzaki (Opisthokonta) and the Dictyostelium discoideum (Amoebozoa) that share common NE proteins with the NEs of multicellular organisms [1]. Yet another example is the structurally distinct NPCs formed in the ciliated protozoa Tetrahymena (Alveolata, SAR), which possesses two structurally and functionally distinct nuclei within a single cell. These unicellular eukaryotes are generally called protists across all phylogenetic supergroups. In this review, we highlight the unique nuclear structures of protists and discuss their attractiveness as potential biological systems for studying nuclear structures.

Virus-derived non-histone proteins packed into highly condensed chromatin in dinoflagellates
Dinoflagellates, which belong to the Alveolata superphylum (Figure 1) [2], contain a strikingly different chromatin structure from most eukaryotes. The chromatin in dinoflagellates remains largely condensed, even in interphase [3]. In most eukaryotes, chromatin is mainly composed of DNA and histone proteins within a nucleosome, which efficiently packs the DNA into a condensed form. In contrast, dinoflagellates have an unusual chromatin structure that does not contain bulk histones [4]. As such, they were once considered an evolutionary intermediate species between histone-less prokaryotes and eukaryotes [4,5]. However, recent genomic and transcriptomic analyses of the dinoflagellates Lingulodinium demonstrated that they have a set of genes encoding all core histones [6], and the genes are expressed as mRNAs, although proteins produced from the genes are below detectable levels [7]. Interestingly, biochemical and transcriptomic analyses of the toxigenic marine dinoflagellate Hematodinium sp. identified no histones, but revealed a novel non-histone basic protein named dinoflagellate/viral nucleoprotein (DVNP) that functions as a bulk protein for packing the genomic DNA (Figure 2) [8]. DVNPs originated from the algal virus protein ExV-152 (Figure 2) [8]. This discovery strongly suggests that dinoflagellates stopped using histones to package the DNA because they gained new proteins with superseding functions, not because they lost histone genes. The genome size of dinoflagellates that evolved DVNPs is much larger than that of related histone-bearing unicellular eukaryotes, such as Perkinsus marinus, a parasite of marine mollusks (4800 ± 500 Mbp in Hematodinium sp. vs 58 ± 9 Mbp in Perkinsus marinus, about an 80-fold increase) [3,9,10]. This coincident increase in genome size with the acquisition of DVNPs leads to the supposition...
Eukaryotic evolutionary tree according to Adl et al. [63]. The five supergroups are represented by different line colors. Groups containing the species described in the text are underlined. Dotted lines indicate uncategorized groups. LECA means the last eukaryotic common ancestor.

Molecular features of the dinoflagellate *Hematodinium* DVNP.5 (GenBank accession number JX839700), phycodnavirus EsV-152 (NP_077537), and human core histones. Orange and green boxes represent α-helices and β-strands, respectively, predicted by psipred (http://bioinf.cs.ucl.ac.uk/psipred/). Vertical lines indicate the positions of basic residues; black and blue lines represent lysine and arginine, respectively. Red diamonds indicate serine/threonine phosphorylation sites. The phosphorylation sites of DVNP.5 and EsV-152 were predicted by Motif Scan (http://myhits.isb-sib.ch/cgi-bin/motif_scan). pI values were calculated at http://web.expasy.org/compute_pi/. E-values were obtained by PSI-BLAST applied against DVNP.5 with the default settings of algorithm parameters. Although DVNP.5 and EsV-152 have a helix-turn-helix motif similar to that of the histone fold, they are evolutionarily distinct from histones.
that condensation with DVNPs results in more compact packaging of their large genome. It remains unknown how DNA is packed with DVNPs. Understanding the structure of DNA condensed with DVNPs requires future investigation. It is also worth investigating how acquisition of DVNPs and the subsequent condensed chromatin structure is advantageous for maintaining the remarkably large sized genome in the extremely diversified dinoflagellate species. To address this last question, finding new species that exhibit intermediate-usage between histones and DVNPs is required. Further studies on dinoflagellates will provide additional information about the molecular basis of genome organization.

Unexpectedly common NE proteins in unicellular eukaryotes

The discovery of 67 novel kinds of nuclear membrane proteins in rodent cells via subtractive proteomics analysis [11] paved the way for the subsequent detection of several hundred NE transmembrane proteins (NETs). These NETs are differentially expressed in different tissues and at different times in the same cells [12,13]. Compared to the successful identification of NETs in higher eukaryotic cells, knowledge of NETs in lower eukaryotes has been limited. Several NE proteins of the yeast *Schizosaccharomyces pombe* have been shown to be conserved in metazoans: the LEM domain proteins Lem2 and Man1, the SUN domain protein Sad1, and the KASH domain protein Kms1, among others [14–17]. In addition, two SUN domain proteins have been identified in *Dictostelium discoideum* (*Dictyostelia, Amoebozoaa*) [18–20]; however, no other NE proteins have been experimentally identified in unicellular eukaryotes other than Opisthokonta, although *in silico* analysis of genomic sequencing predicts the presence of conserved NE proteins in unicellular organisms [21].

Because many of the NE proteins are associated with tissue-specific diseases [13,22,23] — for example, mutations in the A-type lamin gene and the lamin-interacting genes emerin and barrier-to-autointegration factor cause various tissue-specific diseases in humans — how and when these NE proteins appeared during evolution from unicellular to multicellular organisms is an important question to address. Recently, the unicellular Opisthokonta *Capsaspora owczarzaki* (Filastera), was proposed to be a relative of a direct unicellular ancestor of metazoans (multicellular animals) [24] because its genome sequence [25*] contains the gene predicted to encode the lamin protein (XP_004365259) (Figure 3). Lamins (classified

**Figure 3**

Molecular features of lamins and lamin-like proteins. Structural features of human lamin isoforms of lamin A/C (GenBank accession number NP_733821), lamin B1 (AAC37575), and lamin B2 (NP_116126), *Capsaspora* lamin (XP_004365259), and *Dictostelium* lamin-like protein NE81 (Q54HI5) are shown. The positions of secondary structures are predicted by psipred as described in the legend of Figure 2. Orange and green boxes represent α-helices and β-strands, respectively. Conserved domains are represented by gray boxes. LTD is an abbreviation for lamin tail domain. Each human isoform was compared to *Capsaspora* lamin by PSI-BLAST with the default settings of algorithm parameters. The results are indicated beneath the respective protein names.
as type V intermediate filament proteins [26,27]) are major components of the nuclear lamina. In addition, C. owczarzaki contains genes predicted to be associated with the lamina-interacting NE proteins emerin (KJE94080) and barrier-to-autointegration factor (RAF, XP_004343428). These proteins, as well as lamins, were believed to be metazoan-specific proteins. However, recent reports suggest that these proteins were acquired in unicellular eukaryotes [24,25**], at least in ancestral Opisthokonts, after the diversification of the fungi groups. Therefore, C. owczarzaki and its related species could be potential models for studying the structure, function, and evolution of NEs as well as their role in acquiring open mitosis from lower species performing closed mitosis.

It has recently been reported that D. discoideum (Dictyostelium, Amoebozoa), Phytophthora ramorum (Stramenopiles, SAR), and Corallomyxa tenera (Rhizaria, SAR), which belong to groups other than Opisthokonta and thus were believed to lack lamins, possess lamina-like intermediate filaments [1*,28**]. Dictyostelium NE81 protein is characterized as a lamin-like protein from its structural and functional similarity to lamins [28**,29] (see Figure 3). The intermittent appearance of these lamin-like proteins in the various branches of the phylogenetic tree (see Figure 1 and also see Figure 3 of reference [1*]) suggests that these species may have acquired these genes through horizontal gene transfer.

Functionally analogous but structurally distinct proteins from lamins have been reported in Trypanosoma (Discoba, Excavata), the coiled-coil protein NUP-1 [30,31,32**], and in plants, the nuclear matrix constituent proteins (NMCPs) [33]. The evolutionary relationship between these proteins and Opisthokonta lamins is uncertain.

NPCs separate distinctive nuclei in ciliates

The NPC is an NE-embedded structure that is indispensable for all eukaryotes since it functions as a gateway for molecular transport between the cytoplasm and the nucleus. Comprehensive studies have revealed that approximately 30 different kinds of protein components, known as nucleoporins (Nups), constitute the NPC [34–37]. The NPC exhibits rotational symmetry with 8-fold compositional units. The main scaffold body is built up with three ring structures: cytoplasmic, inner, and nuclear rings [38*]. Symmetrically located cytoplasmic and nuclear rings are composed of sixteen Nup107-160 subcomplexes (Y-complexes), and the inner ring structure is composed of Nup93 subcomplexes [39*,40,41]. Another group of Nups that possess disordered regions studded with phenylalanine-glycine (FG) repeats is located on the peripheral region of the NPC. The FG repeats interact with nuclear transport receptors, and thus play a critical role in nuclear-cytoplasmic transport [42,43]. These structures are anchored to the NE by transmembrane Nups. Recent studies demonstrate that the composition of the NPC varies between cell lines and depending on physiological conditions [44–48,49*,50]. However, it is still largely unknown whether these slight differences in NPC composition are functionally related to their nuclear and cellular functions. Notably, ciliates demonstrate that compositionally different NPCs are able to generate separate nuclear functions within a single cell.

Ciliates, characterized by numerous motile cilia on their cell surface, are a taxonomic group belonging to Alveolata (Figure 1). Another biological characteristic of ciliates is the presence of dimorphic nuclei in the same cytoplasm: a macronucleus (MAC) with somatic functions and a micronucleus (MIC) with germline functions [51,52]. The MAC is transcriptionally active and expresses all the proteins in proliferating cells. Its genomic DNA is fragmented into a large number of small-sized chromosomes that lack centromeres. Therefore, the chromosomes in the MAC are divided into daughter cells by amitotic division without the use of kinetochores. In contrast, the MIC is transcriptionally inert. The chromosomes in the MIC divide mitotically using centromeres. The progression of the cell cycle in the MAC and MIC are surprisingly independent of each other (Figure 4a). This dimorphic nuclear system of ciliates is unique and clearly different from other protists exhibiting dimorphic nuclei, such as Giardia intestinalis (Metamonada, Excavata), in which functionally and morphologically equivalent nuclei co-exist [53].

In order to separately regulate such functionally distinctive nuclei, ciliates differentiate the NPC composition of the MAC and MIC. In Tetrahymena thermophila, four paralogs of Nup98, an FG-repeat-bearing Nup, exist exclusively in either the MAC or the MIC, characterizing these two nuclei [54,55]. In contrast, the other Nups are located in both the MAC and MIC NPCs [Figure 4b] [55]. The major signature of the FG-repeats in Nup98 of metazoans and yeasts is Gly-Leu-Phe-Gly (GLFG). Similarly, the two MAC-specific Nup98 paralogs (MacN-up98s) possess the typical GLFG repeat; however, the two MIC-specific Nup98 paralogs (MicN-up98s) possess Asn-Ile-Phe-Asn (NIFN) repeats as an eccentric derivative of the FG-repeat [55,56]. In vivo swapping of these distinctive repeat regions between MacN-up98 and MicN-up98 reduced correct nuclear transport of MIC-specific and MAC-specific linker histones, suggesting that nucleus-specific Nup98 paralogs with distinct repeat signatures are involved in the nucleus-selective transport of nuclear proteins [55]. Tetrahymena clearly demonstrates that the repeat region of Nup98 acts as a selective barrier to nuclear transport. Two distinct nuclei with distinct nucleoporins provide an opportunity to compare their transport activity in the same cytoplasmic space and time, suggesting that ciliates can serve as an excellent experimental system for understanding the function of each nucleoporin.
It has been known that the zygotic nucleus derived from the MIC differentiates into a MAC and MIC after post-zygotic division in ciliates. Thus, MIC-specific Nup98s must be replaced with MAC-specific Nup98s during nuclear differentiation into the MAC. Our recent microscopic study showed that two MAC-specific Nup98s appear only in the presumptive new MAC prior to or at the very early stages of nuclear differentiation, immediately after the last nuclear division of the zygotic nucleus [57**]. Nuclear import of the argonaute-family protein Twi1p, which is required for large-scale chromatin organization for MAC differentiation [58], occurs 20–30 min after the appearance of MacNup98s in presumptive new MACs [57**]. Thus, the assembly of MAC-type NPCs appears to be a primary determinant leading to MAC differentiation involving drastic genome rearrangement [59]. Differences in NPC composition may explain the function of nucleoporins in cell differentiation in multicellular organisms such as humans, as demonstrated in the nuclear differentiation of ciliates.

**Unique NPC structures in other protists**

From the perspective view that NPC remodeling with different Nup98 paralogs acts as a master switch for nuclear differentiation, *Naegleria gruberi* (Discoba, Excavata) is very interesting. This organism performs morphological transformation between amoeba and...
flagellate forms depending on the nutrient condition: the amoeba form in nutrient-rich conditions and the flagellate form in starved conditions [60,61]. Like *Tetrahymena*, *N. gruberi* has two Nup98 paralog genes. Although *N. gruberi* has only a single nucleus, one form has an FG-rich repeat (XP_002679847) and the other has an FN-rich repeat (XP_002677377). This implies that *N. gruberi* switches between these two Nup98 paralogs depending on the physiological conditions. Therefore, these Nups may act as a master switch for morphological transformation by changing the quality of nuclear transport through the NPC and/or by modifying gene expression.

So far, only a few protist species have been used for the comprehensive study of NPC components. In order to understand the evolutionary commonality of NPC architecture and function, various protists extending throughout every supergroup should be analyzed. Common traits of the NPC demonstrated by all supergroups indicate the true origins of the NPC at the last eukaryotic common ancestor (LECA). However, protists could possibly demonstrate unexpected individuality acquired through the various life styles of unique species. For example, although the fundamental NPC scaffold architectures constituted with structural Nups are conserved in the pathogenic protist *Trypanosoma brucei* (Discoba, Exacvata), which causes African sleeping sickness, the primary structures of FG-repeat-containing Nups largely diverge from those of other species [62]. It is possible that parasites or infectious species have a different evolutionary rate that is affected by different selection pressures received from their respective environments. Therefore, *T. brucei* is probably demonstrating both evolutionary commonality and taxa-specific characteristics of the NPC architecture.

**Concluding remarks**

Protists are primarily single-celled eukaryotic organisms that are present in all phylogenetic supergroups (Figure 1). Because each of them evolved over the same period of time to arrive at their present morphologies and biological systems, they cannot simply be considered ‘lower’ eukaryotes, but should be considered ‘higher’ eukaryotes that have evolved in different directions from humans. From this point of view, protists may provide information for unique, yet possibly more efficient, nuclear structures that are different from that of humans. Thus, studies of the nuclear structures of protists are important for understanding the biological significance and functions of the nuclear structure of all eukaryotes, including humans.

**Acknowledgements**

We also thank Dr. D.B. Alexander for critical reading of this paper. We are grateful for funding from JST to T.H. and JSPS Kakenhi Grant Numbers, 24570227, 15K07066 to M.I., 26116511, 26251037 to Y.H. and 25116006, 26291007 to T.H.

**References and recommended reading**

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest


This paper demonstrates, by *in silico* analysis, that genes encoding lamin and lamin-like intermediate filament proteins exist in unicellular eukaryotes. The origins of lamin and the nuclear envelope in eukaryotic evolution are proposed.


This study reports the discovery of a novel non-histone DNA packaging protein. This protein has sequence homology to the algal viral nucleo-protein; the proteins are named as dinoflagellate/viral nucleoprotein (DVNPs). The dinoflagellates switched their DNA packaging proteins from histones to DVNPs at an early stage of their evolution.


Cell nucleus


This paper shows the complete genome sequence data of the filasterean Capsaspora owczarzaki. This unicellular opisthokont is one of metazoan’s closest relatives. Their genome includes many genes encoding metazoan-biased protein domains including nuclear structure-related proteins.


This is the first report of a lamin-like protein in a protist. An NE81 protein discovered as one of the proteins allocated to centrosome-associated structures in Dictyostelium is identified as a lamin-like protein by its structural similarities to lamins and its localization in the nuclear envelope and its functions in nuclear integrity and chromatin organization.


This study characterizes Trypanosoma brucei NUP-1 as a lamin analog protein by demonstrating its functions in nuclear morphology, chromatin organization, and gene silencing at the nuclear periphery. Although NUP-1 is a coiled-coil protein and constitutes the nuclear lamina of Trypanosoma, it is not sequentially related to lamin proteins.


This paper resolves the structure of the human nuclear pore complex at a resolution of 3.2 nm using cryo-electron microscopy.


Using a combination of a cryo-electron tomography with mass spectrometry and other technologies, this paper reveals the interactions of the Y-complex with neighboring Y-complexes and with other scaffold components. The data explain how two kinds of inner and outer Y-complex rings with different diameters can be formed with the same structural blocks.


This paper shows that the composition of the nuclear pore complex varies in different human tissues and in disease conditions. There are significant differences in the expression levels of several NPC components among different cell lines, allowing the variable stoichiometry of the protein complex to functionally adapt to different cellular conditions.


This paper shows that type switching of the nuclear pore complex leads to nuclear differentiation in the ciliate *Tetrahymena*. Live cell imaging demonstrates that MAC-type NPCs are assembled only to the presumptive new MAC prior to large-scale MAC differentiation.


This paper describes the most recent revision of the eukaryote phylogenetic tree.