[Cell Cycle 4:2, 297-299; February 2005]; ©2005 Landes Bioscience

Views and Commentaries

Glimpsing over the Event Horizon

Evolution of Nuclear Pores and Envelope

Gáspár Jékely

Correspondence to: Gáspár Jékely; Meyerhofstrasse 1; Heidelberg 69117 Germany; Tel.: +49.6221.387623; Fax: +49.6221.387166; Email: jekely@embl.de

Received 12/06/04; Accepted 12/06/04

Previously published online as a *Cell Cycle* E-publication: http://www.landesbioscience.com/journals/cc/abstract.php?id=1426

KEY WORDS

comparative genomics, nuclear envelope, nuclear pore complex, eukaryotic evolution

ABBREVIATIONS

NPC	nuclear pore complex
NE	nuclear envelope
ER	endoplasmic reticulum
LECA	last eukaryotic common ancestor
LTD	lamin-tail domain

ACKNOWLEDGEMENTS

The author is supported by a FEBS Long Term Fellowship.

ABSTRACT

The origin of eukaryotes from prokaryotic ancestors is one of the major evolutionary transitions in the history of life. The nucleus, a membrane bound compartment for confining the genome, is a central feature of eukaryotic cells and its origin also has to be a central feature of any workable theory that ventures to explain eukaryotic origins. Recent bioinformatic analyses of components of the nuclear pore complex (NPC), the nuclear envelope (NE), and the nuclear transport systems revealed exciting evolutionary connections (e.g., between NPC and coated vesicles) and provided a useful record of the phyletic distribution and history of NPC and NE components. These analyses allow us to refine theories on the origin and evolution of the nucleus, and consequently, of the eukaryotic cell.

A central feature that distinguishes eukaryotes from prokaryotes is the presence in eukaryotes of a membrane-bound compartment to confine the genome, transcription, RNA maturation, and ribosome assembly. The eukaryotic nucleus, together with other intracellular membrane compartments, originated during one of life's major evolutionary transition[1] and is unparalleled in the prokaryotic world. This remains true even if some bacteria have internal membranes. Members of a distinct group of bacteria, the Planctomycetes, for example store their chromosomes within an internal membranous structure, resembling a true nucleus.[2] These nuclear bodies even have pore-like openings reminiscent of eukaryotic NPCs. However, the genome sequence of *Pirellula sp* revealed no orthologs of eukaryotic NE or NPC components.[3] The endomembrane structures of Planctomycetes then, however fascinating they are, probably evolved independently, and are not informative when we try to understand the origin of the eukaryotic nucleus.

If prokaryotic endomembranes do not represent primitive stages in the development of the eukaryotic nucleus we can only learn about the provenance of this compartment by looking at its present composition. In a recent paper Mans et al.[3] set out to do this in a systematic manner by undertaking an exhaustive bioinformatic analysis of the known components of the nuclear pore complex (NPC), the nuclear envelope (NE), and the nuclear transport systems.[3] Since the origin of a functioning nucleus cannot be envisaged without parallel evolution of these subsystems (e.g., no closed NE can exist without functioning NPCs) such an analysis is indispensable to formulate a detailed evolutionary scenario. Mans et al. base their analysis on the 28 available eukaryotic genome sequences and the catalogs of NE and NPC components obtained by genetic and proteomic analyses.[4-7] The insights they gain into the origin of some of the NPC and NE proteins provide us with invaluable information for finding out how the nucleus evolved.

As a first step Mans et al. look at the phyletic distribution of orthologous genes coding for components of NPCs, the NE, and the nuclear lamina. Based on this comparative-genomic analysis they then try to reconstruct the protein complement of the Last Eukaryotic Common Ancestor (LECA) and to build parsimonious scenarios for the gain and loss of NE and NPC components during the evolution of different eukaryotic lineages. The reconstruction of the ancestral gene set allows one to have inferences about the structure and biology of the ancestral NPC and NE (assuming that no major functional changes occurred and thus functional data obtained mostly for yeast and animal cells apply throughout eukaryotes). The gene gain-loss scenarios on the other hand reveal lineage specific changes, such as evolutionary innovations at the origin of metazoa.

The main difficulty with the reconstructions is that the results largely depend on where the root of the eukaryotic tree is positioned. Since no rooting is generally accepted Mans et al. reconstruct parsimonious scenarios of the gain and loss of NPC and NE components using two alternative topologies.[8,9] The pictures obtained with the two rootings are markedly different. Assuming that the root lies between plants and animals, the LECA had

a very complex NPC and transport apparatus, equipped with most of its present components (59 proteins), many of which have been lost in Diplomonads and Kinetoplastids (*Giardia* and *Leishmania* have NPCs with only 20 and 21 proteins, respectively). If, on the other hand, Diplomonads, Kinetoplastids, and Alveolates are considered as basal groups, splitting off before the plant-animal divergence, one can see a progressive increase in complexity with the LECA having only 19 NPC components, the common ancestor of Alveolates and animals/plants 36, and the animal/plant common ancestor, again, 59.

Mans et al. adopt the latter topology and discuss most of their findings accordingly. However, the reader should keep in mind that if the alternative rooting turns out to be the correct one, as suggested by cell morphology,[10] a derived gene fusion,[8] concatenated mitochondrial protein trees,[11] and Hsp90 trees[12] (see also ref. 13), all proteins referred to as specific for the 'crown group' had also been present in the LECA (i.e., the 59 NPC proteins, etc.). Irrespective of the rooting, however, it is clear that the NPC and the nuclear transport apparatus already formed a functional, multicomponent molecular machine in the LECA. The NPC had all three structural units (central pore, cytoplasmic filaments, nuclear basket), and the transport machinery was already operating with the core set of karyopherins, the Ran cycle, and the Rix system for ribosomal subunit export. An ancestral Src1p/Man1 protein was also present and presumably tethered the NE to chromatin.[3]

Since the basic structural units of the NPC and of the nuclear transport system were present in the LECA their origin is beyond the reach of comparative genomics. The important evolutionary steps occurred in the eukaryotic stem lineage (i.e., the lineage leading to the LECA and leaving descendants only through LECA) and no intermediary forms survived. Nevertheless, there are other ways to gain insights into the origin of those genes that largely lie, as Mans et al. put it, "beyond our event horizon". To achieve this, one has to look for prokaryotic homologs or distant eukaryotic relatives of the relevant molecules. Mans et al. do this for the entire dataset and identify surprising connections that give us hints about the origin of some of the NPC and NE components and suggest cell biological scenarios about the origin and evolution of the nucleus.

Members of the karyopherin family play a central role in nucleocytoplasmic transport. Different members can have different directionalities and substrate specificities. Several karyopherins can be traced back to the LECA and their evolution might give insights into the diversification of karyopherin cargo specificities. The karyopherin snurportin1, for example, functions as a spliceosomal U snRNP-specific nuclear import receptor. It has an N-terminal IBB domain (or K-N module), similar to that of importin-α but differs from all other karyopherins in having a C-terminal m3G-cap-binding region.[14] Mans et al. found that this m3G-cap-binding domain is homologous to the guanylyl transferase domain of the mRNA capping enzyme. The substrate binding residues characteristic of the family are conserved in snurportin1, however, it lacks the catalytic residues, suggesting that it is inactive. This similarity reveals how the specific function of snurportin1 in the import of U snRNPs evolved by the fusion of an IBB and a guanylyl transferase domain.

A combination of comparative genomics and similarity searches revealed interesting aspects about the origin of the nuclear lamina as well. Lamins, the primary components of the nuclear lamina are specific for metazoa. These intermediate filament proteins are composed of a long coiled-coil domain followed by a C-terminal lamin-tail domain (LTD). When looking for homologs of the LTD, Mans et al. found several uncharacterized proteins in diverse prokaryotes but not in eukaryotes outside metazoa. This suggests that the precursor of the animal LTD had been acquired from bacteria via horizontal gene transfer. Mans et al. argue that the evolution of the nuclear lamina (following the fusion of a coiled-coiled domain from an intermediate filament protein with the LTD) might have been an adaptation to the greater contractile forces experienced by animal cells.

Another example with interesting prokaryotic connections is the Ran import factor, NTF2, and some other related nuclear proteins having an NTF2 domain (Mtr2, Tap/Mex67, Bre5p). NTF2 is a member of a structural superfamily the members of which are widespread in eubacteria and include proteins with diverse enzymatic and small-molecule-binding properties (e.g., scytalone dehydratase, 3-oxo-Delta(5)-steroid isomerase, the β-subunit of naphthalene dioxygenase).[15] Structural comparisons suggest that the closest relatives of the eukaryotic NTF2 family are the eubacterial steroid isomerases. It seems that the versatility of this family in binding small molecules was exploited for peptide recognition during the evolution of eukaryotic NTF2.[3]

Steroid isomerases are prevalent in α-proteobacteria, and NTF2 may have originated from the proto-mitochondrial endosymbiont.[3] The uptake of an α-proteobacterium, the ancestor of mitochondria, by the proto-eukaryotic (prekaryotic) host provided ample opportunities for the transfer of α-proteobacterial genes to the host genome. Though it is not possible to carry out phylogenetic analysis in most cases, Mans et al. suggest that many of the NPC components are of α-proteobacterial origin. This could mean that the endosymbiotic event predated the formation of the nucleus. This may also be true for the perfection of mitosis since eukaryotic separases, proteases required for sister chromatid segregation, also have close relatives in α-proteobacteria.[16]

The uptake of an α-proteobacterium may have predated, or even triggered, the segregation of the nuclear compartment. The nuclear envelope then evolved from primitive secretory endomembranes already present, together with the ability to phagocytose other cells, in the prekaryotic cell. This scenario is now further substantiated by the identification of evolutionary links between components of the NPC and secretory endomembranes. Mans et al. found that the nuclear pore protein Nup107 (Nup84 in yeast) contains a conserved domain that is present in other Nups and the Sec31p (Web1p) family of proteins constituting subunits of the COPII coat, a complex that envelops transport vesicles at the endoplasmic reticulum (ER). The presence of a shared domain in some NPC proteins and the vesicular coat proteins of the ER suggests a common origin for the structural components of the NPC and ER-derived vesicular coats. An even broader evolutionary relatedness of NPC proteins and components of coated vesicles has been recognized by others in a recent paper presenting a structural analysis of seven proteins of the yNup84/vNup107-160 subcomplex, a core building block of the NPC.[17] Computational and biochemical studies revealed close similarities between yNup84/vNup107-160 proteins and components of all of the major types of vesicle coating complexes, including clathrin/adaptin, COPI, and COPII coats. These findings again pointing at a common origin of coated vesicles and the NPC. The primordial NPC may thus be envisaged as a defective vesicle budding complex enveloping and curving membranes linked to the chromatin. The original function of such a complex—being unable to form carrier vesicles—might have been to prevent complete fusion of membranes around chromatin.[10] These exciting novel findings leave

little doubt that the NE and its NPCs evolved autogenously by the modification of secretory endomembranes and vesicle coat complexes, and present a real challenge for symbiogenetic theories on the origin of the nuclear compartment.[17]

The autogenous origin of the nucleus has to be interpreted within the general framework of the evolution of eukaryotic endomembranes.[10] Some of the steps of endomembrane evolution, including the origin of the nucleus, may be traced by inspecting the history of Ras-family small GTPases.[18] As a member of the family, Ran is essential for the definition of the nucleus. It regulates the assembly of the NE and NPCs after mitosis, nuclear trafficking in interphase cells, and the formation of a mitotic spindle.[19-22] Proteins of the Ran cycle comprise the most highly conserved unit among all NPC components. Ran, RanGAP, NTF2, RanBP1 and RanGEF are nearly ubiquitous in eukaryotes.[3] Given its central importance in nuclear function it is likely that the origin of the nucleus and mitosis were associated with the advent and evolution of the Ran-cycle machinery. As Mans et al. emphasize Ran differs from all Ras-like GTPases in having a conserved proline followed by a C-terminal acidic tail, instead of having conserved cysteins which undergo fatty-acid modification in most of the other Ras-family GTPases. The loss of this membrane targeting signal and the recruitment of Ran or RanGEF to the chromatin established the Ran system as a positional marker for the chromosomes and was probably an early and major step in the evolution of the nucleus.[18]

The data presented by Mans et al. clearly represent an indispensable resource for cell biologists and besides prompting them to view beyond their selected model organism will help them to formulate experimentally testable predictions. The data will also be useful for evolutionary biologists who are interested in reconstructing the ancestral endomembrane architecture of prekaryotic cells and to formulate cell biological scenarios about how the nuclear compartment (and the first eukaryotic cell) evolved. Mans et al., though very rigorous in their analyses, are extremely cautious with speculations about things that happened beyond the event horizon (i.e., before the LECA). It is evident, however, that these early steps can only be tackled by scenario building, by evolutionary constrained speculations. Such attempts will profit tremendously from the novel constraints introduced in the form of unexpected histories and homologies by this careful and extensive bioinformatic study.

The work by Mans et al. is another example to demonstrate that an evolutionary perspective is indispensable if we want to understand the make-up of highly complex cellular machines. As Eugene Koonin wrote: "Since the organisms and their genomes have been shaped by the haphazard, and yet, highly efficient process of evolution, any understanding of these systems attained without formulating the principles and reconstructing the specifics of evolution will be shallow at best, and more likely, misleading."[23]

References

1. Maynard Smith J, Szathmáry E. The Major transitions in evolution. New York: Oxford University Press, 1997.
2. Lindsay MR, Webb RI, Strous M, Jetten MS, Butler MK, Forde RJ, Fuerst JA. Cell compartmentalisation in planctomycetes: Novel types of structural organisation for the bacterial cell. Arch Microbiol 2001; 175:413-429.
3. Mans B, Anantharaman V, Aravind L, Koonin EV. Comparative genomics, evolution, and origins of the nuclear envelope and nuclear pore complex. Cell Cycle 2004; 3:1612-37.
4. Cronshaw JM, Krutchinsky AN, Zhang W, Chait BT, Matunis MJ. Proteomic analysis of the mammalian nuclear pore complex. J Cell Biol 2002; 158:915-27.
5. Dreger M, Bengtsson L, Schoneberg T, Otto H, Hucho F. Nuclear envelope proteomics: Novel integral membrane proteins of the inner nuclear membrane. Proc Natl Acad Sci USA 2001; 98:11943-8.
6. Mounkes L, Kozlov S, Burke B, Stewart CL. The laminopathies: Nuclear structure meets disease. Curr Opin Genet Dev 2003; 13:223-30.
7. Starr DA, Han M. Role of ANC-1 in tethering nuclei to the actin cytoskeleton. Science 2002; 298:406-9.
8. Stechmann A, Cavalier-Smith T. Rooting the eukaryote tree by using a derived gene fusion. Science 2002; 297:89-91.
9. Templeton TJ, Iyer LM, Anantharaman V, Enomoto S, Abrahante JE, Subramanian GM, Hoffman SL, Abrahamsen MS, Aravind L. Comparative analysis of apicomplexa and genomic diversity in eukaryotes. Genome Res 2004; 14:1686-95.
10. Cavalier-Smith T. The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. Int J Syst Evol Microbiol 2002; 52:297-354.
11. Lang BF, O'Kelly C, Nerad T, Gray MW, Burger G. The closest unicellular relatives of animals. Curr Biol 2002; 12:1773-8.
12. Stechmann A, Cavalier-Smith T. Phylogenetic analysis of eukaryotes using heat-shock protein Hsp90. J Mol Evol 2003; 57:408-19.
13. Arisue N, Hasegawa M, Hashimoto T. Root of the eukaryota tree as inferred from combined maximum likelihood analyses of multiple molecular sequence data. Mol Biol Evol 2004; Epub ahead of print.
14. Huber J, Cronshagen U, Kadokura M, Marshallsay C, Wada T, Sekine M, Luhrmann R. Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure. EMBO J 1998; 17:4114-26.
15. Nagata Y, Mori K, Takagi M, Murzin AG, Damborsky J. Identification of protein fold and catalytic residues of gamma-hexachlorocyclohexane dehydrochlorinase LinA. Proteins 2001; 45:471-7.
16. Aravind L, Koonin EV. Classification of the caspase-hemoglobinase fold: Detection of new families and implications for the origin of the eukaryotic separins. Proteins 2002; 46:355-67.
17. Devos D, Dokudovskaya S, Alber F, Williams R, Chait BT, Sali A, Rout MP. Components of coated vesicles and nuclear pore complexes share a common molecular architecture. PLoS Biol 2004; 2:e380.
18. Jékely G. Small GTPases and the evolution of the eukaryotic cell. Bioessays 2003; 25:1129-1138.
19. Carazo-Salas RE, Guarguaglini G, Gruss OJ, Segref A, Karsenti E, Mattaj IW. Generation of GTP-bound Ran by RCC1 is required for chromatin-induced mitotic spindle formation. Nature 1999; 400:178-81.
20. Hetzer M, Bilbao-Cortes D, Walther TC, Gruss OJ, Mattaj IW. GTP hydrolysis by Ran is required for nuclear envelope assembly. Mol Cell 2000; 5:1013-24.
21. Kuersten S, Ohno M, Mattaj IW. Nucleocytoplasmic transport: Ran, beta and beyond. Trends Cell Biol 2001; 11:497-503.
22. Walther TC, Askjaer P, Gentzel M, Habermann A, Griffiths G, Wilm M, Mattaj IW, Hetzer M. RanGTP mediates nuclear pore complex assembly. Nature 2003; 424:689-94.
23. Koonin E. Interview with Eugene Koonin. Public Awareness Program of the FEBS-IUBMB Conference 2005; In press.