Evolution and diversification of the nuclear pore complex

Alexandr A. Makarov1,*, Norma E. Padilla-Mejia1,* and Mark C. Field1,2

1School of Life Sciences, University of Dundee, Dundee DD1 5EH, U.K.; 2Institute of Parasitology, Biology Centre, Czech Academy of Sciences, 37005 České Budějovice, Czech Republic

Correspondence: Mark C. Field (mfield@mac.com)

The nuclear pore complex (NPC) is responsible for transport between the cytoplasm and nucleoplasm and one of the more intricate structures of eukaryotic cells. Typically composed of over 300 polypeptides, the NPC shares evolutionary origins with endo-membrane and intraflagellar transport system complexes. The modern NPC was fully established by the time of the last eukaryotic common ancestor and, hence, prior to eukaryote diversification. Despite the complexity, the NPC structure is surprisingly flexible with considerable variation between lineages. Here, we review diversification of the NPC in major taxa in view of recent advances in genomic and structural characterisation of plant, protist and nucleomorph NPCs and discuss the implications for NPC evolution. Furthermore, we highlight these changes in the context of mRNA export and consider how this process may have influenced NPC diversity. We reveal the NPC as a platform for continual evolution and adaptation.

Introduction

The nuclear pore complex (NPC), responsible for bidirectional transport of proteins and RNA between the cytosol and nucleus, is an octagonally symmetric structure consisting of multiple co-axial rings, each built of eight spokes (Figure 1). The inner (IR) and outer cytoplasmic and nuclear rings (CR and NR) form the core scaffold, anchored to membrane ring (MR) trans-membrane domain proteins and housing the bulk of FG-nucleoporins (Nups) that maintain the NPC permeability barrier. The nuclear basket (NB) and cytoplasmic filaments (CFs) are attached to the core scaffold and contribute towards both protein and mRNA transport [1].

The protocoatomer hypothesis [2] was proposed in recognition of common architectures between core proteins of multiple complexes within eukaryotic cells, including the NPC. Protocoatomers are membrane-deforming proteins consisting of β-propellers and α-solenoids, and while the primary structure between family members is frequently poorly retained, the secondary structure is considerably better conserved. Inferred from this is that an archetypal membrane coating complex evolved in the earliest eukaryotes (Figure 1A), supported by the presence of β-propeller and α-solenoid-encoding genes within the closest known prokaryotic ancestors of eukaryotes, the archaea. The model proposes that through paralogous duplication, extensive type I and II coat families arose: Type I coats feature an N-terminal β-propeller followed by a continuous α-solenoid, while type II coats bear a perversion within the α-solenoid [3,4]. NPCs contain proteins of both type I and II architectures, suggesting evolution followed the establishment of the major coat types. In Saccharomyces cerevisiae two clearly paralogous columns parallel to the NPC axis form each IR spoke and each spoke is, in turn, duplicated vertically. Similarly, the spokes in the CR and NR are built of two columns each. The column building block is an amalgam of type I and II coat proteins, and this subcomplex, via paralogous expansions, possibly populated the NPC architecture in the LECA (Figure 1B).
Figure 1. Evolutionary origins of the NPC.

(A) Protocoatomer hypothesis states that a single protocoatomer (left, dark purple line) originated sometime early in eukaryotic evolution and gave rise to the two major coat protein families — type I (red lines) and type II (blue lines). Type I and type II coat proteins jointly formed the Proto-NPC — progenitor of the LECA NPC (right) — the concentric assembly of octagonally symmetric inner (IR, purple and dark blue), outer (OR, orange) and membrane (MR, beige) rings anchored in the NE pore that houses the nuclear basket (light blue), the cytoplasmic export platform (pink and burgundy) and FG-repeat nucleoporins (green). (B) An assembly of type I and II coat proteins formed an NPC protomer that populated the double inner ring via oligomerisation and stoichiometric duplication, and further — via paralogous duplication — the outer rings of LECA NPC. (C) Further diversification post-LECA gave rise to many NPC architectures in major taxa of eukaryotic tree (top) that principally differ in the stoichiometry and symmetry of outer cytoplasmic and nucleoplasmic rings (CR and NR), the manner of outer ring attachment to the inner rings and presence and symmetry of specific FG-nup and MR elements.
Molecular data are available for NPCs from many lineages, including metazoa (*Homo sapiens*) [5–8], fungi (*S. cerevisiae, Schizosaccharomyces pombe, Chaetomium thermophila*) [9–16], algae (*Chlamydomonas reinhardtii*) [17], higher plants (*Arabidopsis thaliana, Malus domestica*) [18,19], alveolates (*Tetrahymena thermophila*) [20] and excavates (*Trypanosoma brucei*) [21]. While not particularly deep in taxon coverage, these data span a considerable proportion of eukaryotic diversity. Here we consider the structural variations that are evident between NPCs, consider export mechanisms for transiting the pore and how these systems co-evolved.

**Stoichiometry in outer rings**

The outer rings anchor the cytoplasmic export platform and nuclear basket [6,10,11,22–24]. The Nup85 (yeast nomenclature) or Y-complex is the outer ring building block and contains up to nine components. All are β-propeller, α-solenoid or β-propeller/α-solenoid proteins, archetypal for membrane coating complexes and thus likely to share a common evolutionary origin [1,25–27]. At least six Y-complex components are broadly conserved and five are scaffold nucleoporins, orthologs of HsNup75, HsNup96, HsNup107, HsNup133 and HsNup160 [28]. Notable exceptions are *T. brucei* and *T. thermophila*; each possess novel or highly divergent β/α-proteins TbNup109 [21] and TtNup185 (albeit with some similarity to HsNup133, 155 and 160), respectively [20]. The remaining components are β-propeller nucleoporins, the widely conserved Sec13, Seh1 (absent from excavates), HsNup43 (absent from fungi and TSAR) and HsNup37 (restricted to animals and some fungi); all of which suggests considerable evolutionary flexibility (Figure 2).

Notably, a pronounced difference in NPC organisation across taxa lies in the number of these rings and stoichiometry of Y-complex components. While *S. cerevisiae* has a single CR and NR [10], *H. sapiens* enjoys two copies of each [6] and curiously *C. reinhardtii* has two NR but only one CR [17] (Figures 1C and 2). Additional components are present for interconnection of these duplicate rings, including the vertebrate-specific CF component RanBP2/Nup358 [7,29] and inner ring components — Nup155, 188 and 205 [6–8,17,29]. Deviating further are *S. pombe* and *T. thermophila* (Figures 1C and 2). *S. pombe* rings are ’split’: SpNup107 and SpNup132 localise exclusively to the NR, while six remaining nucleoporins, including SpNup131 (diverse paralog of SpNup132), are CR-exclusive. Deepening this uneven distribution is the overall comparatively low abundance of SpNup107 and SpNup131 [12,13]. *T. thermophila* NPCs, by contrast, demonstrate an uneven distribution but with comparatively higher abundance of 2.5- to 4-fold for Y-complex subunits in the micronucleus (MIC) [20]. Whether this is related to the transcriptional activity differences is yet to be explored. *M. domestica* (apple) differs from *Arabidopsis* as it possesses multiple Y-complex subunit paralogs (albeit likely results of genome duplication) [18,19] (Figure 2), while *T. brucei* acquired an additional complex of three FG-Nups, symmetrically positioned on both sides of the NE [21]; the exact symmetry and stoichiometry of the outer ring in plants and excavates remains to be established.

ELYS, a β/α-nucleoporin with a nucleosome-binding domain, originally identified as a transcription factor [30], interacts with both the Y-complex and pore membrane proteins (POMs). ELYS mediates post-mitotic NPC assembly/anchoring, chromatin compaction and NPC/lamina associations, but most of these functions are metazoan and/or open mitosis-specific [31,32]. Paradoxically ELYS is present in most taxa, except excavates or TSAR, suggesting a probable ancient origin [12,17–21]. As neither plants nor fungi share a lamina with metazoa, ELYS functions are unclear [33,34]. Even more unusual is that the *S. pombe* ELYS ortholog is located on the cytoplasmic side of the NPC, potentially precluding obvious nucleoplasmic roles [13].

**Expansion, loss and diversification in the inner core**

Similar to ORs, main IR components are large α-solenoid (Nup93/Nic96, Nup188, Nup205 orthologs) or β-propeller/α-solenoid proteins (Nup155 orthologs) and again with clear origins in membrane coating complexes [1,25–27]. The IR is well conserved but sports surprising functional diversity (Figure 2). IR in some taxa rely on nucleoporins containing a membrane-binding domain (MMB) for anchorage to the pore membrane. Of these ScNup53 and ScNup59 in *S. cerevisiae* and MdNup35a and MdNup35b in *M. domestica* exemplify a paralog duplication absent elsewhere and presumably occurred independently. However, *T. brucei* sports TbNup65, a single ortholog to ScNup53/59, possessing a conventional *trans*-membrane domain instead of an amphipathic lipid-packing sensor (ALPS) MMB present in ScNup53/59 [21]. *T. thermophila*, by contrast, has no identifiable orthologs to ScNup53/59 and possibly relies on TtNup155, an ortholog to ScNup157/170, and additional POMs for membrane anchoring [20].

Connections between the IR, OR, NB and CFs are also variant. In *S. cerevisiae* connector FG-nups ScNup145N/ScNup116/ScNup100 asymmetrically connect IR with outer rings, with ScNup145N extending...
Figure 2. Comparisons of nucleoporins across species in selected taxa.

Data collated for H. Sapiens [8], S. cerevisiae [10], S. pombe [12], C. thermophilum [14], C. reinhardtii [17], A. thaliana [18], M. domestica [19], T. thermophila [20], T. brucei [21] species and Chlorarachniophytes and Cryptophyceae species clades [125]. Nucleoporins are listed according to their complex and ring disposition in NPC compartments: the inner ring (IR), outer cytoplasmic and nucleoplasmic rings (CR and NR), membrane ring (Mem. Ring, MR), cytoplasmic filaments (CF) and nuclear basket (Nuc. Basket, NB). Nucleoporins are additionally coloured by type: FG-nups and linker FG-nups in IR, CR and NR, CF and NB in swamp green; Nup93/Nic96 complex core in IR in blue; core IR scaffold Nups in purple (also present in ORs as connector Nups); OR Nups in orange; pore membrane proteins (POMs) in MR in beige; Cytoplasmic Export platform.
scaffold in CF in pink; mRNP Remodelling in CF in burgundy, and basket scaffold Nups in NB in light blue. Each row represents an orthologous nucleoporin group. Nucleoporins absent in an organism are indicated by a dash (-). In cases of paralogous duplication within an organism — multiple nucleoporins are listed in square brackets, i.e. [ScNup167 ScNup170]. Alternative Y-complex component distribution in S. pombe is additionally indicated by colour: red for CR-specific and blue for NR-specific. Macronuclei/micronuclei- (MAC-/MIC-) specificity for nucleoporins in Tetrahymena is denoted by Mac- or Mic-prefixes in gene names or * and ** for POMs. Trypanosoma-specific Multi-Complex FG-Nup and Nup76 complexes are shown by additional brackets. In ‘architecture’ column (left) given are the schematic protein fold architectures common for the orthologous Nup groups. Additional unique architectures are indicated for TbNup65 that sports a trans-membrane helix uniquely in its orthologous group (demonstrated on the right), and for lineage-specific Nups TtPOM82, TtNup185, TbNup64, TbNup75, TbNup98, TbNup140 and TbNup149.

from the central NPC towards NR and NB, and ScNup116 and ScNup100 binding IR to CR and the export complex [10] (Figure 2). Furthermore, ScNup157 and 170 stabilise the spokes within the inner ring, while ScNup188 and ScNup192 act as buttresses within the spokes and neither interact with the OR. However, to facilitate connections with duplicate ORs in H. sapiens and C. reinhardtii eight additional copies of HsNup155/CrNup155 and HsNup188/CrNup188 (respective orthologs of ScNup157/170 and ScNup188) form pillars, one per spoke, on each side of the NE when there are duplicate ORs [6–8,17]. Interestingly, in the Xenopus laevis NPC XNup205 (ortholog of ScNup192) replaces Nup188 [29], which suggests considerable flexibility, even within vertebrates.

Unsurprisingly, excavate and TSAR NPCs are organised distinctly to metazoan and fungal complexes (Figure 2). T. brucei has a paralog pair, TbNup144 (orthologous to ScNup157/170) and TbNup119 (similar to ScNup170) [21]. Notably, pullouts revealed TbNup144 to have weak interactions with TbNup89, the ortholog of HsNup75/ScNup85 of the Y-complex. In contrast TbNup119 pulled down the entire T. brucei OR as well as the IR core α-solenoid TbNup225, the ortholog of HsNup205/ScNup192. Notably, neither TbNup225, nor other components of the T. brucei IR interact with the OR, suggesting that TbNup119 bridges IR and OR and is more akin to ScNup145N/ScNup116/ScNup100. Simultaneously no data exists on whether TbNup158, an ortholog to HsNup98/(ScNup145N/ScNup116/ScNup100), is present in the IR or contributes to IR-OR connection. Instead, TbNup158 appears as a constituent component of the Y-complex in T. brucei anchoring the TbNup76 complex and multi-complex FG-Nups in the ORs [21]. Finally, T. thermophila sports single orthologs to each HsNup155/(ScNup157/ScNup170)/(TbNup144/TbNup199) and HsNup205/ScNup192/TbNup225, but up to four orthologs to HsNup98/(ScNup145N/ScNup116/ScNup100)/TbNup158. Two localise exclusively to the macronuclei and carry GLFG repeats, while two are micronuclei-specific and carry NIFN repeats, likely functioning to differentially regulate MAC/MIC-specific transport [20].

**Pore membrane proteins**

The MR is an integral structure of the NE [35], composed of POMs [9] and is possibly the least conserved NPC subcomplex (Figure 2). Only three POMs are candidates for a LECA origin: Ndc1 and Gp210, present broadly but partially lost from TSAR and excavates, with Gp210 also lost from algae and fungi [36]; and the more widely found POM121 [20,36]. All other POMs appear narrowly conserved, leading to structural deviation: Metazoan and fungal NPC ultrastructures are very different within the NE lumen, despite being comprised of structurally homologous domains of Gp210 or POM152, respectively [10,37–39], and no such similarities can be expected in algae or excavates lacking Gp210 [17,21]. Similarly, a particular function is difficult to assign to the core structural POMs. POM152, and the closed-mitosis fungal-specific POM34 [36], form a complex interaction network with NPC components but are non-essential in yeast [40]. However, Gp210 is not expressed in several tissues of mice and human primary embryos [41,42] and reports are conflicting on the phenotype of Gp210 deletions in vertebrates and nematodes [43,44]. POM121 is dispensable in somatic human cell lines but critical in Xenopus embryos [44,45]. In Tetrahymena, POM121 is MAC-specific and distributes towards the NR side, but it is unknown if POM121 acts with MIC- and/or cytoplasmic side-specific TtPOM82 to alter OR stoichiometry [20]. Ndc1 depletion causes severe NPC assembly defects in metazoan and yeast, where it additionally functions in embedding spindle pole bodies into the NE [46–49]. Thus, the MR is a highly divergent structure and may lack a single defined architecture or function.
The nuclear basket and mRNA export

The NB in Metazoa consists of three highly conserved nucleoporins (Figures 2 and 3): Tpr, Nup50 and Nup153 [50], with yeast Mlp1/2, Nup2 and Nup1/Nup60 as respective homologues [51–53]. Tpr/Mlp proteins are involved in RNA biogenesis, spindle pole assembly, regulation of transcription, chromatin remodelling and mRNA export [54]. Although highly conserved within Metazoa, NB proteins suffered a significant diversification in other taxa, suggesting adaptations to organismal-specific roles. Tpr/Mlp homologues are more conserved than Nup153 and Nup50 and several organisms bear lineage-specific NB proteins [55–57].

Tpr, a coiled-coil homodimer, is the NB scaffold [58]. Nup153, a protein with RNA-binding properties [59], anchors to Tpr and the FG-repeats of Nup153 can reach into the NPC core [56,60]. It is unclear if Nup153 is essential for Tpr attachment to the NPC as data are conflicting [56,60–62]. It is likely that Nup153 is essential for Tpr recruitment during post-mitotic NPC assembly but not for anchoring Tpr already NPC localised, suggesting that an additional Tpr-binding site is present [63]. A. thaliana and M. domestica present divergent Tpr homologues (Tpr/NUA) [64–66] and also possess AtNup1/136, homologue of yeast Nup1/human Nup153 [18,67]. AtNup1 and AtTpr localise in the nuclear periphery during interphase with AtTpr localised in the vicinity of the spindle in pro-metaphase [66]. Significantly, C. thermophilum possesses two novel NB proteins, CtNup152 and CtNup56, which bear partial similarity to metazoan Nup50 and yeast Nup2 [14]. In trypanosomes, and conserved across kinetoplastids, TbNup92 and TbNup110 (coiled-coil proteins) constitute the NB and likely evolved independently from Mlp/Tpr [21].

The NB is a platform for the initial stages of RNA export, docking mRNPs and facilitating transport [10,68]. Tpr is crucial for recruiting TREX-2 complex components and hence efficient export of mRNA [61,63]. In yeast, Mlp1/Mlp2 provide docking sites for mRNPs and nuclear export adaptors, while Nup60 bears quality control capacity as Nup60 deletion causes Mlp1 mislocalisation and defects to pre-mRNAs retention [69]. In Arabidopsis, AtTpr and AtNup1 participate in polyA transcript export [18,64–67,70]. In T. brucei, NB proteins impact transcription of some RNA-binding proteins but precise roles remain undetermined [71].

The Mex67–Mtr2 family

mRNA export pathways employ primarily the Nxf1–Nxt1 complex in mammals and their orthologs Mex67–Mtr2 in yeast. Sequence conservation between NFX1–NXT1 and Mex67–Mtr2 is low, albeit retaining functional complementarity [72,73]. Mex67–Mtr2 is the main mature mRNA carrier, an interaction essential for export [74,75]. Translocation of mRNPs across the NPC is mediated by interactions between Mex67–Mtr2 and FG-Nups [76–78]. Localisation of Mex67 at the NB is highly dependent on FG-repeats and deletions cause severe loss of Mex67 from the nuclear periphery [74]. Moreover, in humans, NXF1 presents additional tissue-specific isofoms [79] and ubiquitination regulates recruitment of Mex67 to nascent transcripts [80], adding further complexity to the regulation of mRNA export.

Mex67 has an RNA-recognition motif (RRM), a leucine-rich repeat (LRR), a nuclear transport factor 2-like (NTF2L) and ubiquitin-associated (UBA) domains [74,75,81]. The RRM binds interactors such as the TREX complex [82] and splicing factors [83,84]. Mtr2 also possesses an NTF2L domain, which contacts the Mex67 NTF2L region, forming a heterodimer. Although Mtr2 colocalises with the NPC, it does not directly contact the FG-repeats [76,77,85].

Orthologs of Mex67 and Mtr2 are both present in trypanosomes and essential for mRNA export [21,86–90]. Trypanosome Mex67 retains the LRR and the NTF2L domains, but has a divergent N-terminus with a CCCH-type Zn²⁺ finger replacing the typical RRM domain [21,91], while TbMtr2 also retains an NTF2L domain [21,91]. Additional TbMex67 paralogs are also present but their functions are unknown [90]. Importantly, TbMex67 interacts directly with Ran, suggesting that mRNA export is Ran-dependant and hence mechanistically distinct from animals and fungi [90,91]. In T. gondii, TgZnf2, a nuclear protein containing a C²H₂ zinc finger, functions in mRNA export and cell cycle regulation and is highly conserved across Apicomplexa. TgZnf2 interacts with TgNtx1, a probable ortholog of Mtr2 [92]. In plants, Mex67 has not been found.

Transcription-export complex 2

The TREX-2 complex has a crucial role in genome stability and transcription, providing an essential platform to recruit the mRNA processing machinery [93]; subunit deletions lead to nuclear mRNA accumulation [94]. TREX-2 is loaded onto mRNPs to aid translocation to the NPC and facilitates export by increasing the entry
Figure 3. mRNA export machinery evolution.

(A) Schematic summarising interactions and participants in the global RNA-independent mRNA export machinery. Nascent mRNA complexes with THO/TREX and after displacement of UAP56, export factor Nxf1 is recruited. mRNP is delivered to TREX-2 which facilitates the entry to the NPC. Nxf1 interacts with FG-repeats along the NPC and components at the cytoplasmic side.
Figure 3. mRNA export machinery evolution. Part 2 of 2

cytoplasmic face catalase the release of mRNA. (B) TREX and TREX-2 complexes, export factors, nuclear basket and cytoplasmic NPC components across taxa are summarised. Compared components are shown on the left, depicted with small coloured circles. All the experimental characterised components are written. With a dash, components are certainly absent, dotted circles, components identified in silico, but not experimentally characterised yet. Components marked with a double black star have been identified, but participation in mRNA export has not been proven. n.i. component not identified yet.

efficiency of mRNPs into the NPC transport channel [93,95]. In metazoa, TREX-2 interacts with the NPC in a highly stable manner [61].

In S. cerevisiae TREX-2 consists of six subunits (Sac3, Thp1, Sus1, Cdc31 and Sem1) with Sac3 acting as the scaffold [93]. In animals, the complex consists of a GANP scaffold, PCID2, two copies of ENY2, Centrin2 and DSS1 [96] as respective homologues. TREX-2, through FG-repeats in Sac3, interacts with the mRNP and Mex67/Mtr2, and through Sac3–Sus1–Cdc31 with Nup1 [96–99]. Homologues of TREX-2 components are present in A. thaliana, including Thp1, two Sac3 paralogs and orthologs of Cdc31 and Sem1 with physical interactions between them [67]. However, the Sus1 ortholog, AtENY2, is not a TREX-2 constituent [99,100] and rather is a component of the SAGA transcription complex [101]. As in other systems, depletion of Sac3 and Thp1 in A. thaliana causes mRNA [67] and miRNA accumulation [102]. To our knowledge, no components of TREX-2 have been functionally characterised in other taxa, with the possibility that they are present but too divergent to be identified by sequence comparison. In Apicomplexa, a potential Cdc31 has only been identified in silico [103].

TREX

The TREX complex is co-transcriptionally recruited to nascent mRNAs and regulates splicing and export [104]. TREX consists of the multi-subunit THO complex, a conserved DEAD-box RNA helicase Sub2 and an export adaptor Yra1 (Figure 3), to which other components assemble. Sub2 has conserved functions promoting splicing, mRNA export and recruitment of Yra1. In the assembled TREX-mRNP complex, Sub2 together with Yra1 may load Mex67–Mtr2. Remarkably, although THO complex composition is clearly divergent between animals and fungi (Figure 3), recent data suggests that the overall tertiary structure, multimerization and flexibility of TREX are strikingly conserved [105]. Yra1 is essential in S. cerevisiae [106], but not metazoa [107,108] and overexpression in A. thaliana lacks an obvious phenotype [109]. This highlights that, although overall structure and mechanisms seem conserved, adaptations can contribute towards evolutionary context-dependent essentiality.

THO is conserved in plants but A. thaliana possesses additional THO and Yra1 paralogs (Figure 3), indicating complex diversity [110,111]. Trypanosomes have a Sub2 ortholog, an essential protein associated with mRNA transcription/processing sites and export [112], while in T. gondii a highly divergent ortholog has been characterised [113]. The remainder of the TREX complex in these organisms awaits discovery [114].

Cytoplasmic filaments and mRNA export

At the cytoplasmic face of the NPC in yeast lies the Nup82 complex (Nup82, Nup159, Nsp1), a part of the export platform and attached to the cytoplasmic OR facing the central channel. Nup82 also helps Gle1, Dbp5 and Nup42 in organising the last stages of export. Dbp5 is an RNA-binding DEAD-box helices involved in transcription, mRNA export and termination of transcription [115,116]. During mRNA export, Dbp5 triggers remodelling of mRNAs emerging into the cytoplasm in the final export steps. Yeast Dbp5 is regulated by Gle1, an interaction stabilised by inositol-hexakisphosphate (IP₆), which catalyses the release of RNA-binding proteins to ensure directional transport from the nucleus [117,118]. However, in humans IP₆ binding may be dispensable [119], suggesting diverse mechanisms. Adding additional levels of regulation of mRNA export pathways, multiple Gle1 isoforms have been found [120].

In Arabidopsis, Gle1 is highly conserved [121], essential [122] and stimulates the Dbp5 homologue AtLOS4 [18,122]. Interestingly, in Lotus japonicus, Gle1 has evolved to promote a symbiotic relationship with mycorrhiza Mesorhizobium loti for symbiotic nitrogen fixation [123]. Trypanosomes lack orthologs of Gle1 and Dbp5 indicating a distinct mechanism [21] and consistent with simplified trans-splicing.
Gle2 (mammal/plant Rae1) is involved in mRNA export as inactivation/mislocalisation leads to nuclear accumulation of mRNA [124]. Interestingly, Gle2 is retained in trypanosomes and nucleomorph nuclei [21,125] (Figures 2 and 3), suggesting an ancient origin in the eukaryotic lineage and possibly a central role.

**Functional and genomic constraints moulding NPC evolution**

The NPC is a remarkable example of co-evolution as mutations could result in complex effects produced by impacting many interacting partners [126,127]. Interestingly, Nups show different evolutionary rates which may reflect distinct evolutionary pressures [128]. Distinct functional constraints (Table 1) may influence this process and include protein–protein and protein–nucleic acid interactions [127,129]. Moreover, ubiquitously expressed proteins tend to evolve slower than tissue-specific proteins [129,130] and HsNup50, Tpr and Gp210 show differential mRNA and protein expression levels in different tissues, suggesting altered NPC composition between cell types [5,131].

An interesting further example is the autocatalytic domain of HsNup98/Sc145C. HsNup98 and HsNup96 are expressed as a single fusion protein which undergoes autoproteolytic processing [132], an event essential for localisation [133]. Interestingly, this mechanism is conserved in S. cerevisiae Nup145 [134], but absent from other organisms such as trypanosomes [21].

**Dynamic properties of the NPC**

The inner and outer rings are the most conserved NPC subcomplexes and the α-solenoids, β-propellers and coiled-coil bundles within the ring nucleoporins account for the majority of mass density in cryo-EM structures. Conserved architectures trace NPC core evolution back to type I and type II coats (Figure 4A), but NPC modularity permits a variety of architectures (Figure 4B). It is important to understand that the NPC core is dynamic and there is a noticeable variation of the central channel diameter between 60 and 40 nm, respectively, depending on species and imaging method [7,10,17,135,136], representing the capacity of the NPC for constriction/dilation in response to environment [137], presence of transport factors/cargo [138–140], cell cycle stage

<table>
<thead>
<tr>
<th>Constraint class</th>
<th>Constraint</th>
<th>Possible impact</th>
<th>Examples within the NPC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Functional</strong></td>
<td>Structural environment of catalytic amino acids</td>
<td>Residues within the catalytic core are under greater pressure of being conserved. Residues are substituted in ways that the overall stabilities of structure are maintained [127,129]</td>
<td>Nups with enzymatic activity: Gle1, helicases Dpb5/DHX9. Autoproteolytic domain found in human Nup98/Nup96 (yeast 145N/145C) [135]</td>
</tr>
<tr>
<td></td>
<td>Protein–nucleic acid interactions</td>
<td>Protein–nucleic acid recognition/ interfaces tend to be highly conserved [127]</td>
<td>Rae1, Elys, Nup153 and ScNup157 capacity to bind nucleic acid [32,59,137]</td>
</tr>
<tr>
<td><strong>Genomic</strong></td>
<td>Gene expression: differential expression in tissues</td>
<td>Ubiquitously expressed proteins tend to evolve slower than tissue-specific proteins [129,130]</td>
<td>HsNup50, Tpr, Nup214, Aladin, Gp210, Pom121 and Nup37 levels (transcript and protein) are different in different tissues, suggesting rearrangements of NPC stoichiometry across cell types [5,131]</td>
</tr>
<tr>
<td></td>
<td>Epigenetics: chromatin remodelling</td>
<td>Chromatin remodelling and epigenetic marks impact gene expression and therefore, protein evolution [126,130]</td>
<td>Nuclear basket roles in chromatin remodelling and epigenetic regulation (reviewed in [139])</td>
</tr>
</tbody>
</table>
Figure 4. Outstanding questions in NPC evolution and known NPC diversifications. Part 1 of 2

(A) While the structural core of LECA NPC — its inner and outer rings — can be convincingly traced to type I and type II coat proteins (left), it is yet unclear how the NPC acquired its other components — responsible for mRNA processing and general permselective function (right). As individual FG-nups/transport factor combinations were shown sufficient to execute partial per-selective function, and POMs — to form pores in lipid bilayers,
Figure 4. Outstanding questions in NPC evolution and known NPC diversifications. Part 2 of 2

the possibility arises that these components formed a separate simple pore before incorporation into NPC. (B) Summary of principal variations and evolutionary events in NPC architectures post-LECA. Principal architectural changes are named for each compartment. Principal evolutionary events are named and indicated. Paralogous duplication events indicated by a — a,b pairs with slight colour change. Gene loss — by dashed lines and letters. Gene gain — by a — c pairs and colour change. Changes in stoichiometry — shown schematically. Principal complex gains are named.

[141] and mechanical force [142]. Such flexibility is thought to depend on several structural characteristics of the NPC. Firstly, within the central channel allosteric coupling exists between structured and intrinsically disordered FG-Nup domains in interactions with transport factors [139]. Furthermore, the lateral spoke interconnections within the inner ring are small [10] and contributing to the stability of the inner ring are ‘flexible connectors’, intrinsically disordered sequences interconnect NPC subcomplexes. For example, in S. cerevisiae these connectors are present in Nic96 that bridges Nup192 and the FG-Nups of the Nsp1 complex. Furthermore, Nup145N/Nup100/Nup116 interconnects inner and outer rings [16]. Notably, these interactions can be allosterically coupled by transport factors: Nup53 interaction with Kap121 destabilises Nic96 and Nup157 binding [140], potentially contributing to loosening of the inner ring and pore dilation to allow transport of cargo otherwise excluded from the NPC. Conservation of these interaction sites and allosteric interactions between animals and fungi [16,140] suggest such mechanisms are intrinsic to NPC function.

The NPC during mitosis

Eukaryotes perform cell division by closed, open or semi-open mitosis [143]; in all cases, nucleoporins affect mitotic progression. Metazoans and higher plants (here represented by H. Sapiens, A. thaliana and M. domestica) undergo open mitosis that includes full breakdown of NE (NEBD) and NPC disassembly [143,144]. The remaining organisms, i.e. C. reinhardtii, T. thermophila and T. brucei perform close mitosis [145–147] which leaves the NE intact during cell division.

Although these models of nuclear division seem radically different, the mechanisms for disassembly of the NPC are strikingly similar, occurring in a highly synchronised manner. In animal cells, NPC disassembly occurs during NEBD and after phosphorylation of Nup98, while in S. pombe NPCs are gradually lost from the anaphase bridge connecting daughter nuclei. In open and closed division peripheral Nups are disassembled first, followed by the central scaffold and finally POMs [148–151]. If differences in NPC architecture influence the mitotic mode is unknown.

Moonlighting NPC components

Moonlighting, or multi-functionality, is common amongst NPC and NPC-associated components and inevitably impacts selection constraints. NPC components are involved in a variety of pathways, located in additional compartments as individual nucleoporins or entire NPC subcomplexes. Sec13 (Y-complex) is a COPII component [152] and, together with Seh1 (also Y-complex) [153], are part of the SEA vacuolar complex. Significantly Nup62 and Nup188 are at mammalian centrosomes [154,155], and the entire mammalian Y-complex at kinetochores and, together with Seh1, recruit the chromosomal passenger complex [156–160]. Multiple nucleoporins promote chromatin decondensation, transcriptional activation [161–166] or epigenetic silencing [163] via direct localisation in the nucleoplasm or tethering chromosomal regions to assembled NPCs [167–173]. Very little is known in terms of conservation of moonlighting functions, although it is clear that trypanosomes do have similar processes, albeit with likely distinct evolutionary origins, suggesting possible convergence.

One emerging example of NPC component moonlighting is the ‘ciliary pore complex’ [174]. NPC components can localise at the base of the cilia and in human cells, these are the cytoplasmic filament Nup214, inner ring Nup35, Nup62 and Nup93, the outer ring Nup37 [174] and Nup75/Nup85 [175]. Additionally, Nup93 and multi-ring Nup188 [176] and, potentially, Nup205 [177] are thought to localise to the cilia base in Xenopus. These structures appear to support embryonic cilia formation and intrflagellar transport (IFT). Additional similarities between NPC and IFT, such as the ciliary localisation signal (CLS) likeness to NLS [178] and the requirement of CLS-recognising nuclear transport factor importin-β2 and a RanGTP/GDP gradient for the ciliary transport of several proteins [179,180], have added to the concept of a ciliary pore complex [174]. However, structural details and ubiquity remain to emerge and imaging of Nup188 in Xenopus shows structures incompatible with typical NPC organisation [174]. Among FG-nups, only Nup98 was found at the cilia [175] and ciliary...
transport appears insensitive to classic inhibitors of the NPC barrier [181]. The similarity in form and function between the NPC and ciliary pore complex is, therefore, unclear; however, the constraints applied by potential co-evolution between the nuclear and ciliary transport are intriguing. Interestingly, there is further mutual moonlighting between the NPC and cilia as centrin2, critical for centriole/centrosome, and thus cilia organisation [182], is also present in animal and fungal NPCs, contributing to mRNA and protein export [183].

**NPC evolution: a minimal pore?**

The predictive and experimental structural analysis finds that the ~30 proteins of the NPC core are composed predominantly of just eight structural fold types [25,26]. Of these, the three most frequent, α-solenoid, FG-repeat and β-propeller, account for >80% of residues. Notably, while sequence similarity is generally low, comparative genomics has identified over 20 nucleoporins from all three groups as conserved across all eukaryotic taxa [36] and thus likely represented in the LECA. Assuming that the NPC evolved incrementally with increased complexity, what forms did the NPC take during the transition from the FECA to the LECA (Figure 4A)? Phylogenetic and genetic analysis may be currently insufficient to resolve this question, but insights from other studies are valuable here.

Firstly, nucleomorphs, drastically reduced nuclei from red and green algae endosymbiotically absorbed into chlorarachniophytes and cryptophyceans [125], reveal a near-complete secondary reduction in algae NPC genes. The only NPC proteins that could participate in pore formation [184] are orthologs of Nup98 (in chlorarachniophytes) and Rae1 (in chlorarachniophytes and cryptophyceans) (Figure 2). However, there is no ultrastructural evidence for a pore-associated structure at the nucleomorph membrane and the retention of these subunits may indicate recruitment to other processes. Secondly, attempts to create an artificial pore, using truncated versions of yeast Nsp1 and Nup100 [185] or a designer ‘NupX’ [186], produced pores with selective permeability, suggesting minimal requirements to achieve gating. A self-assembling pore of just two POMs, Ndc1 and an FG-containing Pom121 in a lipid bilayer has also been achieved [187]. While these artificial or derived states almost definitely do not recapitulate the NPC during the FECA to LECA transition, they do indicate considerable potential for flexibility and great simplicity in mediating selective nucleocytoplasmic transport.

**Perspectives**

- There are examples of clear divergence in NPC structure and likely functions across eukaryotes; more examples are needed to understand the basis for these changes.

- Mapping structure to function remains a major goal for NPC research, but understanding diversity and how these connect to biological aspects are also critical.

- Gaining deeper insight in the organisation of nuclear pores in nucleomorphs and into pore formation by POMs may aid reconstructions of the origins of permeability-selective NPCs.

- There is an urgent need to characterise mRNA export platforms (TREX, Mex67) and the manner in which these interface with the NPC, especially in non-canonical organisms.

**Competing Interests**

The authors declare that there are no competing interests associated with the manuscript.

**Author Contributions**

A.A.M. and N.E.P.-M.; investigation and preparation of manuscript and figures. A.A.M., N.E.P.-M. and M.C.F.; writing, review and editing; M.C.F.; conceptualisation and supervision.
Acknowledgements

Work in our laboratory is supported by the Wellcome Trust (204697/Z/16/Z to M.C.F.).

Abbreviations

CF, cytoplasmic filament; IFT, intraflagellar transport; LRR, leucine-rich repeat; MIC, micronucleus; MMB, membrane-binding domain; MR, membrane ring; NB, nuclear basket; NPC, nuclear pore complex; NTF2L, nuclear transport factor 2-like; POMs, pore membrane proteins; RRM, RNA-recognition motif.

References


175 Endicott, S.J. and Brueckner, M. (2018) NUP98 sets the size-exclusion diffusion limit through the ciliary base.


