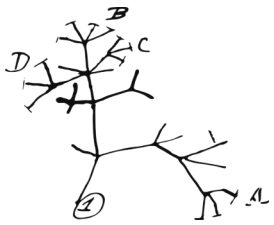
EVOLUTION OF 'CILIA' PROTEINS' GENE REGULATORY MECHANISMS

ALEXANDER EWERLING



Institute of Molecular Physiology

Faculty of Biology

Johannes Gutenberg University

Mainz

Mainz, August 2023

EVOLUTION OF 'CILIA' PROTEINS' GENE REGULATORY MECHANISMS

Institute of Molecular Physiology

Faculty of Biology

Johannes Gutenberg University

Mainz

Thesis submitted by

ALEXANDER EWERLING

born 01.09.1993 in Lebach, Germany

for the degree of

DOCTOR RERUM NATURALIUM (DR. RER. NAT.)

Mainz, August 2023

DEAN: Prof. Dr. Eckhard Thines
Institute of Molecular Physiology
Johannes Gutenberg University, Mainz, Germany

1ST EXAMINER: Prof. Dr. Helen Louise May-Simera
Institute of Molecular Physiology
Johannes Gutenberg University, Mainz, Germany

2ND EXAMINER: Prof. Dr. Martin Kaltenpoth
Department of Insect Symbiosis
Max Planck Institute for Chemical Ecology, Jena, Germany

Date of oral examination: 24/11/2023

ABSTRACT – ENGLISH

Eukaryotic cilia are evolutionarily conserved organelles that enable cellular motility and sensory functions and were present in the last eukaryotic common ancestor (LECA). Cilia are cell appendages composed of microtubules, emerging from a basal body that anchors the cilium within the cell, facilitating both movement and perception of external signals. While single-celled eukaryotes possess motile cilia that can also exhibit sensory functions, multicellular organisms exhibit two main types: motile cilia, aiding fluid transport, and immotile primary cilia, which serve as sensory signalling hubs for tissue development and homeostasis. Mutations in ciliary proteins lead to distinct clinical pathogenic phenotypes, stemming from impaired signalling in primary cilia or reduced motility of motile cilia. These classes of syndromes are termed ciliopathies, of which the Bardet-Biedl Syndrome (BBS) is considered the archetypical ciliopathy. The BBSome is a highly conserved octameric complex (BBS1, 2, 7, 9, 4, 8, 5, and 18) facilitating ciliary cargo transport by linking cargo proteins with transport complexes. Research in cilia biology has traditionally focussed on the roles of BBSome and chaperonin-like BBS proteins (BBS6, 10, 12 – required in the assembly of the BBsome) in cilia assembly. Recent findings hint at novel functions for ciliary proteins, including vesicular transport, mitotic regulation, and cytokinesis, and nuclear-associated tasks like chromatin modulation via histones, gene regulation via transcription factor transport, and epigenetic modifications via DNA methylation. In this thesis we comprehensively analysed human BBS protein nuclear localisation and their conservation across eukaryotes through phylogenetic reconstructions. We found that BBS proteins are highly conserved from LECA, and that some are even retained despite absence of other BBS proteins. By signal sequence prediction and permutation analysis, the propensity of BBS proteins to enter eukaryotic nuclei was assessed both across different taxa as well as different BBS proteins. We found that the ability to enter the nucleus is unaffected by mitotic nuclear envelope breakdown. Exploring potential gene regulatory roles, non-model organisms, in this case insects, were considered given their limited cilia expression as well as possible phenotypes linked to the regulatory roles of BBS proteins. Honeybees provide promising BBS protein candidates for influencing gene expression. Furthermore, this thesis examines the co-evolution of organelles during early eukaryogenesis by integrating insights into ciliary proteins, nuclear pore complexes, and karyopherins. This sheds light into potential processes that could have lead to the acquisition of nuclear functions during the transition to LECA, and provides a basis to identify possible conserved functions of BBS proteins in otherwise disparate eukaryotic organisms. With this work, we gained new insights into the evolution of nuclear localisation of BBS proteins across the entire eukaryotic Tree of Life. The results obtained have implications for evolutionary and developmental biology alike, as they highlight the versatility of ‘ciliary’ proteins to participate in diverse biological processes apart from ciliary transport. These new functions might help explain ciliopathy phenotypes and elucidate disease mechanisms, and aid in understanding early eukaryogenesis.

ABSTRACT – DEUTSCH

Eukaryotische Zilien sind konservierte Organellen, die Motilität und sensorische Funktionen ermöglichen und beim letzten gemeinsamen Vorfahren der Eukaryoten (LECA) vorhanden waren. Zilien sind Zellfortsätze, die aus Mikrotubuli bestehen und aus einem Basalkörper hervorgehen, der das Zilium in der Zelle verankert und sowohl die Bewegung als auch die Wahrnehmung externer Signale vermittelt. Während einzellige Eukaryoten bewegliche Zilien besitzen, die auch sensorische Funktionen haben können, gibt es bei mehrzelligen Organismen zwei Haupttypen: bewegliche Zilien zum Flüssigkeitstransport, und unbewegliche Primärzilien als sensorische Signalzentren für die Gewebeentwicklung und Homöostase. Mutationen in Zilienproteinen führen zu unterschiedlichen pathologischen Phänotypen, die auf eine gestörte Signalübertragung oder eine verminderte Beweglichkeit der Zilien zurückzuführen sind. Diese werden als Ziliopathien bezeichnet, von denen das Bardet-Biedl-Syndrom (BBS) als Archetyp gilt. Das BBSome ist ein hochkonservierter Komplex (BBS1, 2, 7, 9, 4, 8, 5 und 18), der den ziliären Ladungstransport durch die Verknüpfung von Frachtproteinen mit Transportkomplexen erleichtert. Die Forschung im Bereich der Zilienbiologie hat sich traditionell auf die Rolle des BBSome und der Chaperonin-ähnlichen BBS-Proteine (BBS6, 10, 12 - die für den Aufbau des BBsome erforderlich sind) konzentriert. Jüngste Erkenntnisse deuten auf neue Funktionen für ziliäre Proteine hin, darunter Vesikeltransport, mitotische Regulation und Zytokinese sowie kernassoziierte Aufgaben wie Chromatinmodulation über Histone, Genregulation über Transkriptionsfaktortransport und epigenetische Modifikationen über DNA-Methylierung. In dieser Arbeit wurde die Kernlokalisierung menschlicher BBS-Proteine und ihre Erhaltung bei Eukaryoten durch phylogenetische Rekonstruktionen umfassend analysiert. BBS-Proteine sind seit LECA konserviert, und in manchen Fällen trotz des Fehlens anderer BBS-Proteine erhalten. Durch Signalsequenzvorhersage wurde die Neigung von BBS-Proteinen, in eukaryotischen Zellkernen aufzutreten, untersucht. Die Fähigkeit, in den Zellkern einzudringen, ist nicht durch den Abbau der mitotischen Kernhülle beeinflusst. Bei der Untersuchung potenzieller Kernfunktionen wurden Nicht-Modellorganismen, in diesem Fall Insekten, in Betracht gezogen, da sie nur begrenzte Zilienexpression aufweisen, aber manche Phänotypen durch genregulatorische Effekte von BBS-Protein beeinflusst werden könnten. Honigbienen bieten vielversprechende BBS-Protein-Kandidaten für die Beeinflussung der Genexpression. Zudem wird die Koevolution von Organellen während der frühen Eukaryogenese untersucht, indem Erkenntnisse über ziliäre Proteine, Kernporenkomplexe und Karyopherine vereint werden. Dies beleuchtet mögliche Prozesse, die zum Erwerb von Kernfunktionen während des Übergangs zu LECA geführt haben könnten, und bietet eine Grundlage für die Identifizierung möglicher konservierter eukaryotischer Kernfunktionen von BBS-Proteinen. Dies ist für Evolutions- und Entwicklungsbiologie von Bedeutung, da die Vielseitigkeit der "ziliären" Proteine aufgezeigt wird, die neben dem ziliären Transport an verschiedenen biologischen Prozessen beteiligt sind. Diese neuen Funktionen könnten dazu beitragen, die Phänotypen der Ziliopathie zu erklären, Krankheitsmechanismen aufzuklären und die frühe Eukaryogenese besser zu verstehen.

CONTENTS

ABBREVIATIONS

ADP/ATP – adenosine di-/triphosphate

AP – adaptor protein, adaptin

Arf – ADP rybosylation factor

Arl – Arf-like

BB – basal body

BBS – Bardet-Biedl Syndrome

BLAST – Basic Local Alignment Search Tool

BLASTp – protein BLAST

BS – bootstrap

CCT – chaperonin-containing TCP

COP – coatomer protein

CORVET – class C core vacuole-endosome tethering

CPC – ciliary pore complex

DAPI – 4′,6-diamidino-2-phenylindole

DMEM – Dulbecco's modified Eagle's medium

DNA – deoxyribonucleic acid

DNMT – DNA (cytosine-5)-methyltransferase

E-value – expectation value

ER – endoplasmic reticulum

ESCRT – endosomal sorting complex required for transport

(e)ToL – (eukaryotic) Tree of Life

FBS – fetal bovine serum

FECA – first eukaryotic common ancestor

FIJI – Fiji is just ImageJ

g – gravitational acceleration [9.81 m × s^{-2}]

g – gramme(s)

GA – Golgi apparatus

GAE – ɣ-adaptin ear

GAPDH – glyceraldehyde 3-phosphate dehydrogenase

GDP/GTP – guanosine di-/triphosphate

GPCR – G-protein coupled receptor

HEAT – Huntingtin, elongation factor 3, protein phosphatase 2A, target of rapamycin 1

HEPES – 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HEK293T – human embryonic kidney cells expressing SV40 large T antigen

HMM – hidden Markov model

HOPS – homotypic fusion and vacuole protein sorting

IFT – intraflagellar transport

IFT-A/B – IFT complex A/B

kDa – kilo-Dalton

l – litre(s)

LECA – last eukaryotic common ancestor

LG – Le-Gascuel

LUCA – last universal common ancestor

M – molar [mol × l^{-1}]

MAFFT – Multiple Alignment using Fast Fourier Transform

MKKS – McKusick-Kaufman Syndrome

MKS – Meckel-Gruber Syndrome

ML – maximum likelihood

mRNA – messenger RNA

mRNP – messenger ribonucleoprotein

MT – microtubule

NE – nuclear envelope

NES – nuclear exit/export signal

NLS – nuclear localisation signal

NPC – nuclear pore complex

NPHP – Nephronophthisis

Nup – nucleoporin

OMV – outer membrane vesicle

p-value – probability value

PBS – phosphate-buffered saline

pH – *potentia Hydrogenii*

PH – pleckstrin homology

PRC – polycomb repressive complex

PVDF – polyvinylidene difluoride

PY-NLS – proline-tyrosine NLS

RNA – ribonucleic acid

rRNA – ribosomal RNA

SEA – Seh1-associated

SH – Shimodeira-Hasegawa

SV40 – simian virus 40

TBS – Tris-buffered saline

TCP – T-complex protein

TPR – TTC repeat

TREX2 – three prime repair exonuclease 2

Tris – tris(hydroxymethyl)aminomethane

TTC – tetratricopeptide

WAG – Whelan and Goldman

WB – Western blot

FIGURES

TABLES

1. INTRODUCTION

1.1. THE DOMAIN *EUKARYA*

The story of eukaryotic life on earth is one of the most successful stories of life and considered one of the major transitions of cellular life forms (Szathmáry and Smith, 1995). Eukaryotes comprise some of the most complex organisms on earth, with a highly specialised compartmentalisation of the cytosol that allows many different milieus for biochemical reactions to occur in a small space. They are also the only taxon to contain multicellular life, at least from what is known so far. The unifying theme behind eukaryotes is their defining feature, the nucleus. The nucleus (from Latin *nucleus*, 'kernel, seed') contains the genetic information in eukaryotes (from Greek εὖ, 'proper, true' and κάρυον, 'kernel, seed') and is separated from the remaining cytosol, in contrast to prokaryotes where the genetic material is freely accessible in the cytosol. The separation of genetic material from the rest of the eukaryotic cell came with the advantage of sophisticated, spatiotemporal regulation of gene expression and therefore improved tractability of life cycle stages. Through evolution, eukaryotes have diversified into many different groups inhabiting most ecological niches (even with extreme environmental conditions) that are summarised in the eukaryotic Tree of Life.

The phylogenetic basis for the Tree of Life is the assumption that all eukaryotic lifeforms evolved from a single common ancestor, LECA. With the radiation of eukaryotic lifeforms, new species began occupying an astonishing spectrum of ecological niches that lead to the current plethora of eukaryotes. The eukaryotic Tree of Life (eToL) is the basis for all modern phylogenetic reconstructions of relationships between species of eukaryotes. Its topology is dominated by supergroups, organisms that cluster together based on similar proteomic and genomic analyses. Initially eukaryotes were grouped into plants, animals, fungi, and protists (which in principle encompassed every organism that did not fit the previous groups) (Whittaker, 1969), but this classification did not properly resolve the relationships within the so-called protists. With the advance of genomics, proteomics and molecular phylogenetics, incremental improvements have led to a better resolved tree that is currently composed of eight supergroups: TSAR (telonemids, stramenopiles, alveolates, and Rhizaria), Haptista, Cryptista, Archaeplastida, Amorphea, CruMs (collodictyonids, Rigifilida, *Mantamonas*), Discoba, Metamonada, Hemimastigophora, and orphan clades with uncertain placement (Fig. 1) (Adl et al., 2019, 2005; Baldauf, 2003; Burki et al., 2020; Keeling et al., 2005; Roger and Simpson, 2009; Simpson and Roger, 2004, 2002).

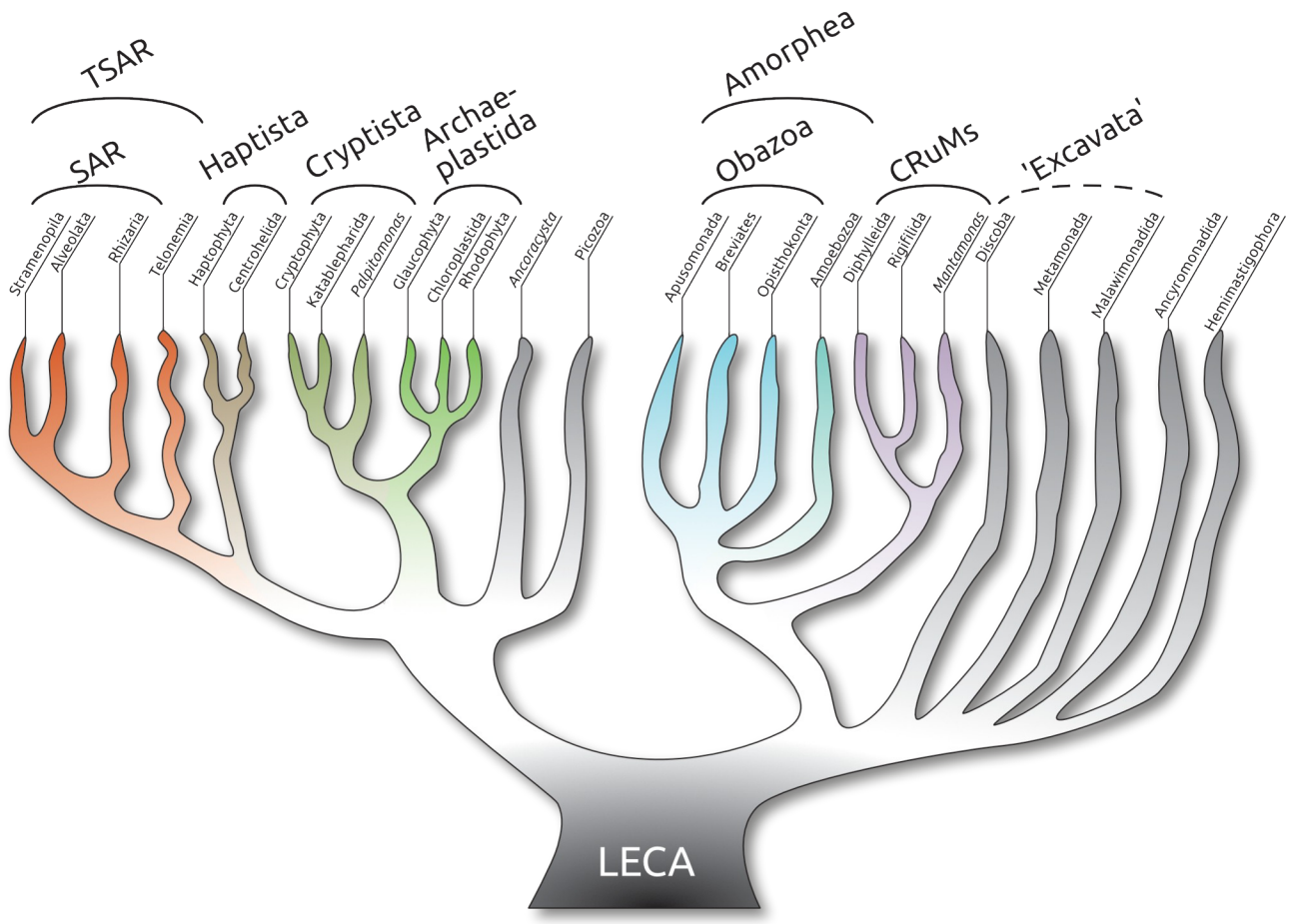


Figure 1: The eukaryotic Tree of Life. Branches represent current crown taxa and are grouped into 'supergroups' with correspondingly coloured branches. Grey branches are 'orphan' clades with uncertain placement within the tree, mostly due to lacking data. Adapted from Burki et al., 2020.

Eukaryotes are a diverse group of organisms ranging from single-celled to multicellular, but have a core set of defining features that is shared between them. Through endosymbiosis, eukaryotes gained the ability to generate energy within their cytoplasm facilitated by mitochondria. Mitochondria themselves can be found in various forms, e.g. as aerobic or anaerobic mitochondria, hydrogenosomes or mitosomes (Müller et al., 2012), depending on their host lifestyle, and are also regulators of various other functions such as apoptosis (Roger et al., 2017). Eukaryotes maintain endomembrane organelles which are needed to spatially separate biochemical reactions and cellular processes (Fig. 2). Two of them, the endoplasmic reticulum (ER) and Golgi apparatus (GA), provide an interface for sequential protein modification, quality control and trafficking. Eukaryotes separate transcription and translation and confined them in two respective organelles, the nucleus and the ER. This allowed for better spatiotemporal management of gene regulation (Martin and Koonin, 2006) and likely contributed to evolvability of signal transduction cascades (Poole et al.,

2003). The nucleus itself serves as a major regulator through selective permeability for proteins and molecules: through nuclear pore complexes (NPCs) and cargo adaptors, the cell can regulate which proteins enter the nucleus at specific times, resulting in an intricate balance of gene activation or repression through various steps in signal transduction cascades. Further, eukaryotes developed a novel structure termed the cilium that is used in locomotion and signal recognition throughout the eToL (Carvalho-Santos et al., 2011; Hodges et al., 2010). The cilium is a microtubule-based structure that extends from the cell body into the extracellular space and moves the cell by beating, or, in the case of specialised primary cilia in animals, remains stationary. In all forms, it serves as a signalling hub for the cell's surroundings, emitting and receiving signals in various forms (Bloodgood, 2010; Mitchell, 2007). Cilia show high conservation not only in function, but also proteomic makeup. This is best exemplified when looking at the movement of proteins within the cilium. In cilia, this is called intraflagellar trafficking (IFT) and mediated through motor proteins and cargo adaptor complexes, the IFT-A/B complexes and BBSome.

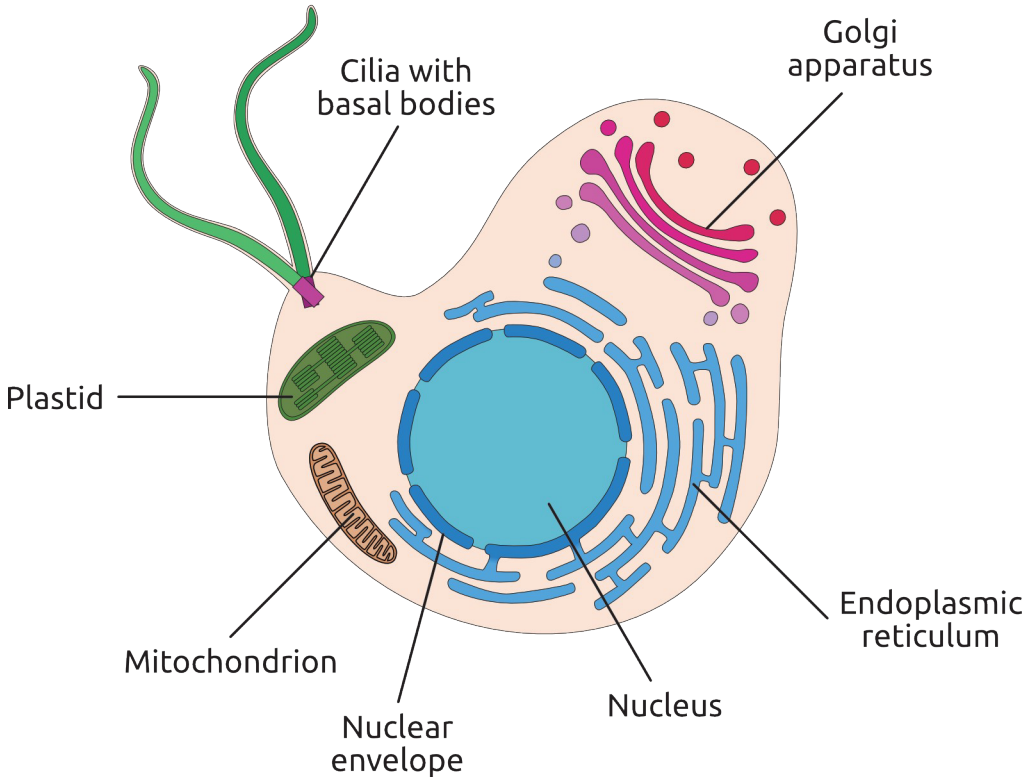


Figure 2: Overview of an eukaryotic cell with organelles. Note that not all organelles are present in all eukaryotes. Adapted from Ewerling et al., 2023b.

All these organelles most likely arose sequentially, although the exact order of emergence is currently unknown. This means however that the proteomic environment in the evolving proto-eukaryote was in constant flux, and without the boundaries of organelles had ample opportunity to interact. These interactions lie at the heart of co-evolution, and are best exemplified by two organelles: the nucleus and the cilium. Cilia influence gene expression after signal perception from receptors on the ciliary surface, and a several ciliary proteins could be found in the nucleus (Ewerling et al., 2023b; Gascue et al., 2012; Horwitz and Birk, 2021; Scott et al., 2017; Shi et al., 2018). On the other hand, cilia utilise mechanisms for nucleocytoplasmic translocation for ciliary import and export (Dishinger et al., 2010; Huang et al., 2021), and *vice versa* nuclear proteins have been found in cilia proteomes (Kee et al., 2012; Marquez et al., 2021). Why do these two organelles use a shared set of proteins? Could they have co-evolved during eukaryogenesis? To answer these questions, we need to examine the evolutionary history of both organelles, and analyse the steps that were necessary to arrive at the *status quo* of cilia proteins localising to the nucleus.

1.2. THE NUCLEUS

Despite being ubiquitous in eukaryotes, the nucleus and its NPCs display significant variability (Bapteste et al., 2005). Across the eToL, both have undergone lineage-specific adaptations (Fernandez-Martinez and Rout, 2021), reflecting their adaptability to various biological niches. The NPC architecture remains consistent among eukaryotes (Fig. 3): nucleoporins (Nups) constitute a core scaffold with linked outer and inner rings on the cytosolic and nucleosolic sides of the pore. The nucleoplasmic face is integrated into the nuclear lamina. Extending in both directions, cytoplasmic filaments form on the cytosolic face, while other proteins create the nuclear basket on the nuclear face. A central channel is decorated with phenylalanine-glycine (FG) repeat Nups, forming a brush-like structure that impedes free diffusion of molecules. While this structure is foundational across eukaryotes, Nups themselves exhibit substantial structural, but not sequence conservation (Alber et al., 2007; Asakawa et al., 2014; Cronshaw et al., 2002; Eibauer et al., 2015; Fischer et al., 2015; Iwamoto et al., 2010; Kim et al., 2018; Kosinski et al., 2016; Mosalaganti et al., 2018; Obado et al., 2016; Ori et al., 2013; Rout et al., 2000; Tamura et al., 2010). They share the same α-solenoid/β-propeller structure seen in descendants of the hypothetical protocoatomer. Nups likely existed in LECA (Neumann et al., 2010) and expanded through paralogous duplication before eukaryotic lineages diverged (Alber et al., 2007; Beck et al., 2018; Devos et al., 2004; Kim et al., 2018; Kosinski et al., 2016). Given their limited sequence conservation, the focus was on their capacity to form the same complex in diverse ways rather than exact amino acid composition. Nups

share relations with COP I and COP II family membrane coating proteins (Kim et al., 2018), along with adaptin-like proteins (Alber et al., 2007; Hoelz et al., 2016; Knockenhauer and Schwartz, 2016; Kosinski et al., 2016), and show modest sequence conservation across lineages (Bapteste et al., 2005; DeGrasse et al., 2009; Mans et al., 2004). This counters a "nucleus-first" evolutionary scenario and suggests potential co-evolution of the nucleus with the ER and GA, as both COP I and COP II coats are vital for transport between these organelles and the plasma membrane. Given their presence in the nucleus, primitive forms of these organelles could have existed during NPC evolution. The evolutionary roots of Nups (and other protocoatomer-derived proteins) likely trace back to prokaryotic and archaeal α-solenoid and β-propeller domain proteins (Field and Rout, 2019). The expansion of this protein family gave rise to the diverse membrane coats shaping organelle identity and contributing to intricate multiprotein complexes in eukaryotic cells. Similarly, early Nup precursor proteins' association with membranes may have led to primitive repetitive structures resembling proto-NPCs within the (potentially incomplete) nuclear envelope (NE). Through duplications, the NPC diversified in composition while retaining interaction sites with other NPC subunits.

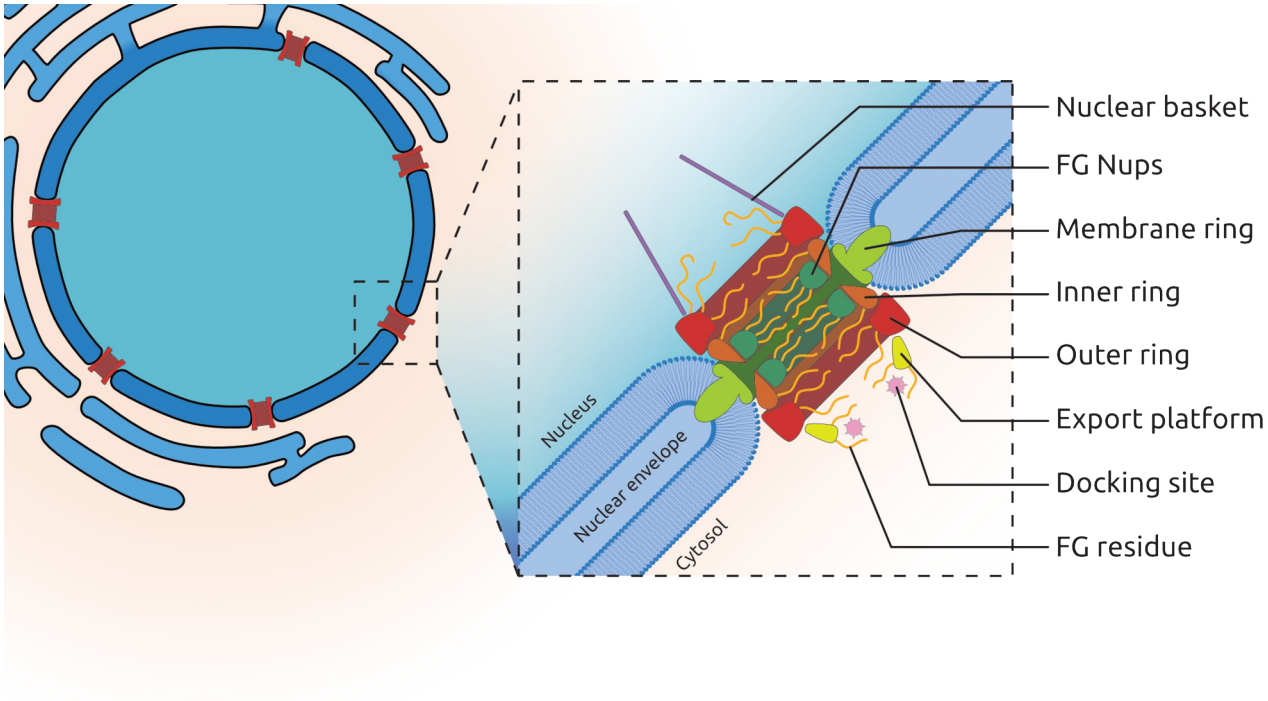


Figure 3: Detailed structure of a nuclear pore complex (NPC). NPCs have eightfold radial symmetry. The NPC is anchored to the NE membrane by transmembrane nucleoporins (Nups) (membrane ring). Membrane, inner, and outer rings associate with each other to form a barrel-like structure with a central pore. Cargoes enter the nuclear pore at the docking site and are transported by electro- and hydrostatic interactions with phenylalanine-glycine (FG) residues. FG residues convey selective transport through the nuclear pore. Cargo complexes leave the pore through the nuclear basket. Transport into the cytoplasm is analogous, with cargoes exiting from the export platforms. Adapted from Field and Rout, 2019.

NPCs are the gatekeeper to the nucleoplasm. They safeguard against undesired molecule entry with phenylalanine-glycine (FG) moieties stemming from FG-Nups on both faces of the NPC (Alber et al., 2007; Rout et al., 2000). FG Nups are intrinsically disordered proteins with substantial hydrodynamic radii that effectively further pore size (Denning et al., 2003). Inter- and intramolecular electrostatic interactions among these FG-Nups (Gehrke et al., 1997; Kawai et al., 2003) prevent random passage through nuclear pore, maintaining high entropy within the NPC. FG-Nups play a role in regulated entry and exit of proteins and molecules across the NPC by interaction with transporter proteins, the karyopherins, that facilitate nucleocytoplasmic transport of the proteins they associate with. Phylogenetic analysis suggests that LECA likely possessed a diverse range of karyopherins (Kaps) (O'Reilly et al., 2011). These proteins predominantly adopt α-solenoid/β-propeller structures, with HEAT and ARM repeats being the prominent domains (Bayliss et al., 2002, 2000), and share a probable proto-coatomer origin akin to Nups and other membrane coats. Functionally, they are often referred to as importins and exportins, depending on their roles in respective pathways. Both importins and exportins serve as cargo adapters for nuclear proteins. The interaction of Kaps with FG residues within NPCs through hydrophobic interactions enables the Kap-cargo complex to surmount the entropy barrier that typically impedes free cargo diffusion

(Iovine et al., 1995). Despite the possibility of diffusion in the presence of FG repeats, specialised import and export proteins compete for binding sites within NPCs, effectively preventing nonspecific passage through higher-affinity interactions (Jovanovic-Talisman et al., 2009; Zilman, 2009; Zilman et al., 2007). The pathways also rely on the RanGTP/RanGDP gradient across the NE to ensure proper transport directionality. Importins, when bound to cargo, move through the NPC (Izaurralde et al., 1997), releasing cargo within the nucleus before binding nucleoplasmic RanGTP and undergoing export again. In the cytoplasm, hydrolysis of RanGTP to RanGDP frees importin from the export complex for another cycle. For nuclear export, exportin associates with cargo and nuclear RanGTP, forming the export complex, which releases cargo in the cytoplasm through RanGTP hydrolysis. This gradient holds critical importance for nucleocytoplasmic transport (Lui and Huang, 2009). Interestingly, Nups, a RanGTP/RanGDP gradient, and karyopherins are also present in cilia and fulfil in principle the same functions there as in the nucleus (Dishinger et al., 2010; Huang et al., 2021; Marquez et al., 2021). Given their shared hypothetical ancestor, this argues strongly for an evolutionary relationship between cilia and the nucleus, and the proteins needed for their functionality.

1.3. EUKARYOTIC CILIA

While both archaea and bacteria possess superficially similar structures (archella and flagella, respectively), the eukaryotic cilium is a structure whose building blocks do not have homologues in the bacterial or achaeal proteomes from what we know so far. Cilia are an eukaryotic innovation dating back to LECA and represent one of its defining features (Koumandou et al., 2013). While cilia can be found across all major eukaryotic lineages, secondary losses have occurred frequently (Carvalho-Santos et al., 2011; Hodges et al., 2010). The eukaryotic cilium functions as a specialised signalling structure abundant in receptors and contributing to locomotion (Bloodgood, 2010; Mitchell, 2007; Satir et al., 2008). Although often labelled as an "organelle," this term can be misleading since cilia lack a continuous membrane that would define compartmentalisation. Anchored by the basal body, a barrel-like structure composed of nine symmetrically arranged MT triplets, cilia extend into the extracellular space via the ciliary axoneme, which consists of MT doublets sheathed by the plasma membrane, with motile cilia containing a central pair (9 + 2 architecture) that is lacking in immotile cilia (9 + 0) (Fig. 4). The ciliary base is enclosed by the transition zone, a specialised protein compartment regulating ciliary molecule sorting, entry, and exit (Hu et al., 2010). Contemporary eukaryotes showcase diverse cilia forms, ranging from the immotile primary cilia specific to Metazoa to the prevalence of motile cilia in most single-celled

eukaryotes. However, all these variants likely trace back to the LECA cilium, which performed similar functions and employed comparable molecular mechanisms for assembly, disassembly, and maintenance, as seen in present-day eukaryotes. As cilia and basal body structures are conserved, so are the proteomes, with essential components largely consistent in most ciliated species (Hodges et al., 2011; Wickstead and Gull, 2007). The A- and B-tubules of axonemal doublets serve as highways for ciliary trafficking essential for cilium assembly, disassembly, and maintenance. Highly conserved IFT complexes IFT-A, IFT-B, and the BBSome move along these tracks using kinesin-2 (anterograde) or dynein 1b/cytoplasmic dynein 2 (retrograde) motor proteins. These motor proteins link with ciliary cargo proteins, such as transmembrane receptors, via the BBSome as an adaptor complex (Prevo et al., 2017; Taschner and Lorentzen, 2016). This well-conserved machinery likely evolved from protocoatomer diversification during the transition from FECA to LECA, contributing to the shaping of the organellar landscape (van Dam et al., 2013).

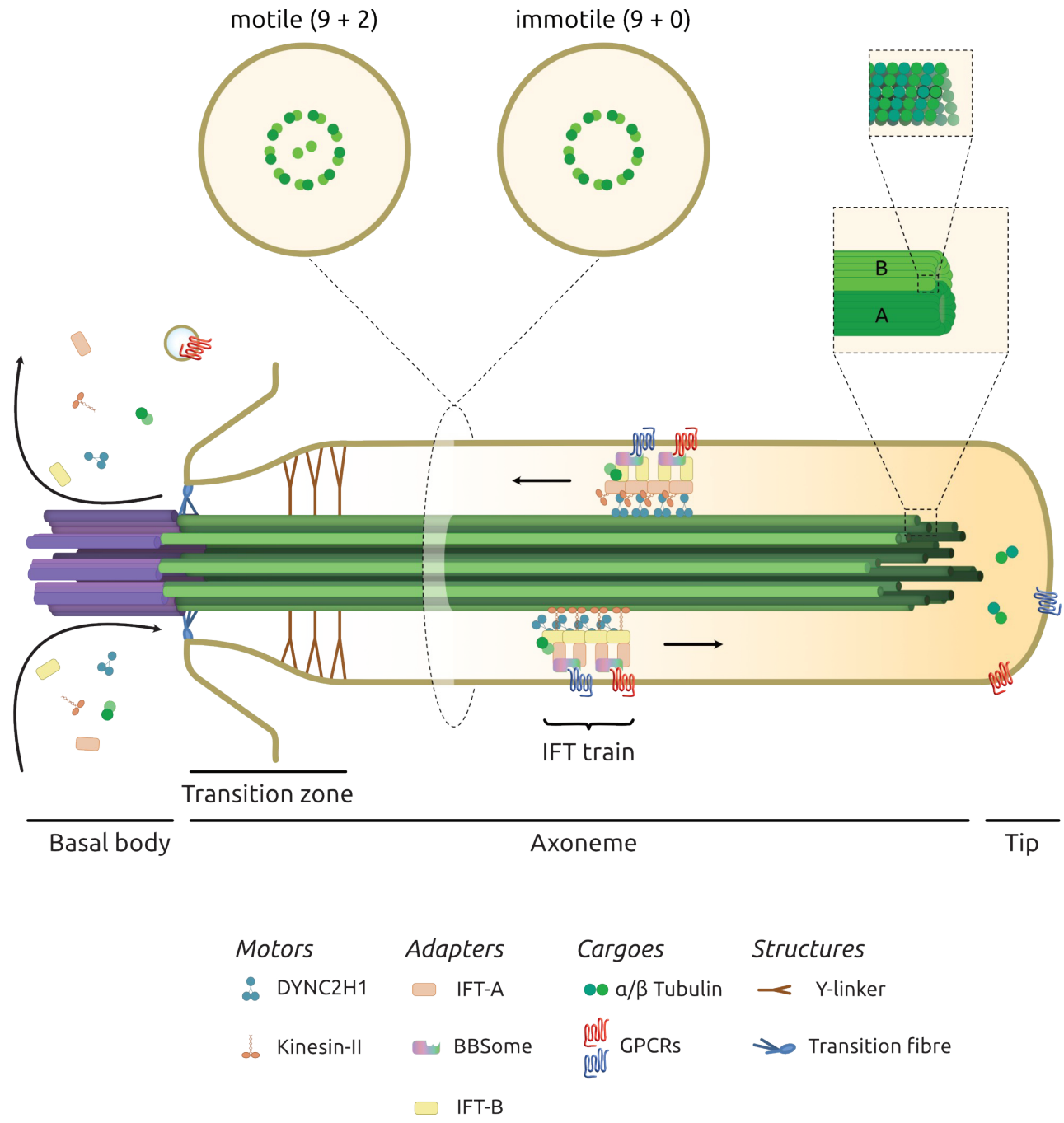


Figure 4: Overview of an eukaryotic cilium. Typical architecture of motile (9 + 2) and immotile (9 + 0) cilia. Cargo proteins enter and exit the cilium through the transition zone at the ciliary base. Selectivity is facilitated by transition fibres, Y-linkers, and protein complexes at the transition zone. The ciliary axoneme consists of microtubular A- and B-tubules, which polymerise from αβ-tubulin dimers.

1.3.1. CILIOPATHIES

Extant cilia serve two major function. On the one hand, cilia either help organisms and cells move through their surroundings (protists, sperm cells) and move fluids across surfaces (ciliated epithelia); on the other hand, they are enriched with receptors and sense environmental cues, from

both biotic and abiotic sources. The immotile primary cilia of animals have specialised in signalling functions and play a crucial part in development and tissue maturation (Anvarian et al., 2019; Blitzer et al., 2011; Luo et al., 2014; Veland et al., 2009). As other eukaryotic cilia, they rely on IFT for assembly, maintenance, and disassembly. Conversely, any disruption in IFT leads to a malfunction in ciliary function. In humans, these malfunctions cause distinct clinical disease phenotypes, fittingly called ciliopathies (Badano et al., 2006). Ciliopathies arise from mutations in cilia-related genes, and proteins in the same functional pathway cause overlapping disease phenotypes (Table 1). Within this ever expanding class of diseases, the predominant phenotypes include retinopathy, obesity, polydactyly, renal failure, *situs inversus*, and cognitive impairment (Braun and Hildebrandt, 2017; Kenny and Beales, 2013). The archetypical ciliopathy is the Bardet-Biedl Syndrome (BBS) (Bardet, 1920; Biedl, 1922) as patients display all hallmark symptoms of other ciliopathies (Forsythe and Beales, 2013). The causative mutations for this syndrome can be found in the BBS proteins. BBS proteins form the BBSome (BBS1, 2, 7, 9, 4, 8, 5, 18) (Nachury et al., 2007), aid in folding the BBSome (BBS6, 10, 12) (Seo et al., 2010), or fulfil other function in the vicinity of the cilium. To date, 24 different BBS proteins are known (Florea et al., 2021), and the number continues to grow. There is limited genotype to phenotype correlational with affected siblings often experiencing differing symptoms, and symptoms can arise at variable stages during life. Clearly there must be underlying mechanisms that cause this variability. Research into the non-ciliary functions of BBS proteins might be a promising avenue to explore this phenomenon. For some cilia proteins, studies have been conducted in mammalian cells and could partially explain certain aspects of the disease phenotype (Gascue et al., 2012; Scott et al., 2017). However, comprehensive approaches are still missing. Exploring this new topic in cilia biology could potentially harbour answers to how ciliopathies can be so phenotypically plastic while the confounding factors are the same.

Table 1: Ciliopathies and associated phenotypes. Adapted from Kenny and Beales, 2013. Abbreviations: BBS, Bardet-Biedl Syndrome; ALMS, Alström Syndrome; JATD, Jeune Syndrome; JBTS, Joubert Syndrome; LCA, Leber's congenital amaurosis; MKKS, McKusick-Kaufman Syndrome; MKS, Meckel-Gruber Syndrome; NPHP, Nephronophthisis; OFD, Oro-facial-digital syndrome Type 1; SLS, Senior-Løken Syndrome.

	BBS	ALMS	JATD	JBTS	LCA	MKKS	MKS	NPHP	OFD1	SLS
Retinopathy	✔	✔	✔	✔	✔		✔	✔		✔
Obesity	✔	✔								
Polydactyly	✔		✔	✔		✔	✔		✔	✔
Renal failure	✔		✔	✔			✔	✔	✔	✔
Situs inversus	✔			✔			✔	✔		
Cognitive impairment	✔	✔		✔	✔		✔	✔	✔	

1.4. Origin of eukaryotic life

Although eukaryotes are an incredibly diverse group, they all likely trace back to a single common ancestral lifeform, LECA. This hypothetical ancestor separated from Archaea potentially ~1.8 billion years ago (Betts et al., 2018). Recent advances in phylogenetic reconstruction of several protein families have started to paint a clearer picture of what this organism may have looked like (Bremer et al., 2023, 2022; Skejo et al., 2021) and what it was capable of. However, it is mandatory to take a closer look at how LECA came to be in the first place, and then trace back its features and evolution.

Together with Eubacteria and Archaebacteria (or simply Archaea), eukaryotes comprise the three extant domains of life. Two models for the ontology of the Tree of Life exist: in its crudest form, the three-domain-model was proposed first by Ernst Haeckel in 1866 (Haeckel, 1866) and re-defined in the late 1970s (Woese and Fox, 1977) based on rRNA sequence comparisons. It postulated that besides the typical bacteria, there is a branch of methanogenic bacteria called archaebacteria and a

branch of “urkaryotes” that would have been the progenitors of modern eukaryotes. Interestingly, Woese and Fox already concluded that what they termed archaebacteria “appear to be no more related to typical bacteria than they are to eukaryotic cytoplasms”, meaning they are most likely sister lineages rather than a subset of eubacteria. They further stated “[o]ne of the three may represent a far earlier bifurcation than the other two, making there in effect only two urkingdoms” and in principle establishing the final three-domain-model. And eukaryotes are indeed more closely related to archaeal lifeforms in terms of transcriptional and translational machineries, suggesting they are more closely related than eubacteria and eukaryotes.

The second model states a two-domain structure after the bifurcation of eubacteria and archaea, and that eukaryotes arose from within archae and are monophyletic with them. This theory was dismissed by some researchers (Forterre, 2013) that argued that the ancestor of eukaryotes arose from an endosymbiotic event between a bacterium from the PVC (Planctomycetes, Verrucomicrobia, Chlamydiae) phylum and an engulfed thaumarchaeal cell. Hallmark features would have been contributed through archaea, bacteria, and viruses. Subsequent studies however deemed this scenario to be convergent evolution rather than homology, refuting this as the origin of endomembrane organelles (McInerney et al., 2011).

Both models existed for a long time without any obvious reasons for either to be more reasonable than the other (Forterre, 2011; Gribaldo et al., 2010). And while the initial three-domain-model acknowledges a close proximity of Archaea and Eukarya, the lack of data and modern phylogenetic methods impeded the development of the theory which might have led to the closely related, current hypothesis of eukaryogenesis. The two-domain model gained new momentum when in 2015, a group of researchers ventured into the deep, more specifically to a deep-sea vent named Loki’s Castle located between Norway and Greenland. They recovered samples of a novel archaeal lineaged they called ‘Lokiarchaeota’ (Spang et al., 2015) that showed astonishing similarities to eukaryotes on a protein level. The novel lineage was found to be part of a superphylum containing Thaumarchaeota, Aigarchaeota, Crenarchaeota, and Korarchaeota (TACK) and, later, to form the group of Asgardarchaea, together with Thorarchaeota (Seitz et al., 2016), Odinarchaeota, and Heimdallarchaeota (Zaremba-Niedzwiedzka et al., 2017). The Asgardarchaea beared signatures of eukaryotic signal proteins (ESPs) across all phyla, namely an extensively expanded family of small GTPases, actin-related proteins, ‘proto’-ESCRT proteins related to eukaryotic ESCRT I, II, and II complexes required for membrane coating, components of the ubiquitylation system, and glycosylation and translocation pathway homologues (Spang et al., 2017; Zaremba-Niedzwiedzka et

al., 2017). Phylogenetic analyses confidently placed the root of eukaryotes within this group, based on ribosomal RNA comparisons (Zaremba-Niedzwiedzka et al., 2017). This made the Asgardarchaea the closest relatives of eukaryotes to this date, and solidifies the two-domain model where eukaryotes arose from within archaea.

But given that eukaryotes arose from within archaea, how did they arrive at their current morphology? How was the archaeal isoprene-based lipid ether membrane(Langworthy et al., 1972) replaced by the phospholipid bilayer seen in extant eukaryotes? How did the numerous membrane-bounded organelles like ER, GA, and endosomes, mitochondria and chloroplasts, and the namesake of the domain, the nucleus, arise? The answer to those questions lies in the endosymbiosis theory, a concept first put forward in the early 20th century (Mereschkowsky, 1905), neglected, rediscovered, and refined over the following decades (De Duve, 1969; John and Whatley, 1975; Margulis, 1970; Martin, 2017, 1999; Raven, 1970; Sagan, 1967; Wallin, 1923).

1.5. EUKARYOGENESIS: EVOLUTION FROM FECA TO LECA

Eukaryotes most likely arose from an endosymbiotic event. The most tenable hypothesis today is that an ancestral prokaryotic cell (likely a relative of Lokiarchaea, based on phylogenetic reconstructions) engulfed an anaerobic α-Proteobacterium that would later become the mitochondrion of all subsequent eukaryotic lineages (Fig. 5). In the case of Archaeplastida (green plants), Rhodophyta (red algae), Excavata (including e.g. Euglenids) and Chromalveolata (including e.g. brown algae), one or multiple secondary acquisitions of photosynthetic cyanobacteria (and even other eukaryotes) have led to a complex manifold of plastids with a diverse range of functions (Keeling, 2010; Strassert et al., 2021). Over the years, many other theories have been put forward to explain the current set of eukaryotic organelles and machineries, all with their own advantages and shortcomings (de Duve, 2007; Devos and Reynaud, 2010; Forterre, 2002; Kurland et al., 2006; Van Valen and Maiorana, 1980). Currently, the most likely evolutionary scenario now is the acquisition of an endosymbiotic bacterium by an archaeon and the establishment of the endomembrane system, resulting in a rather complex LECA.

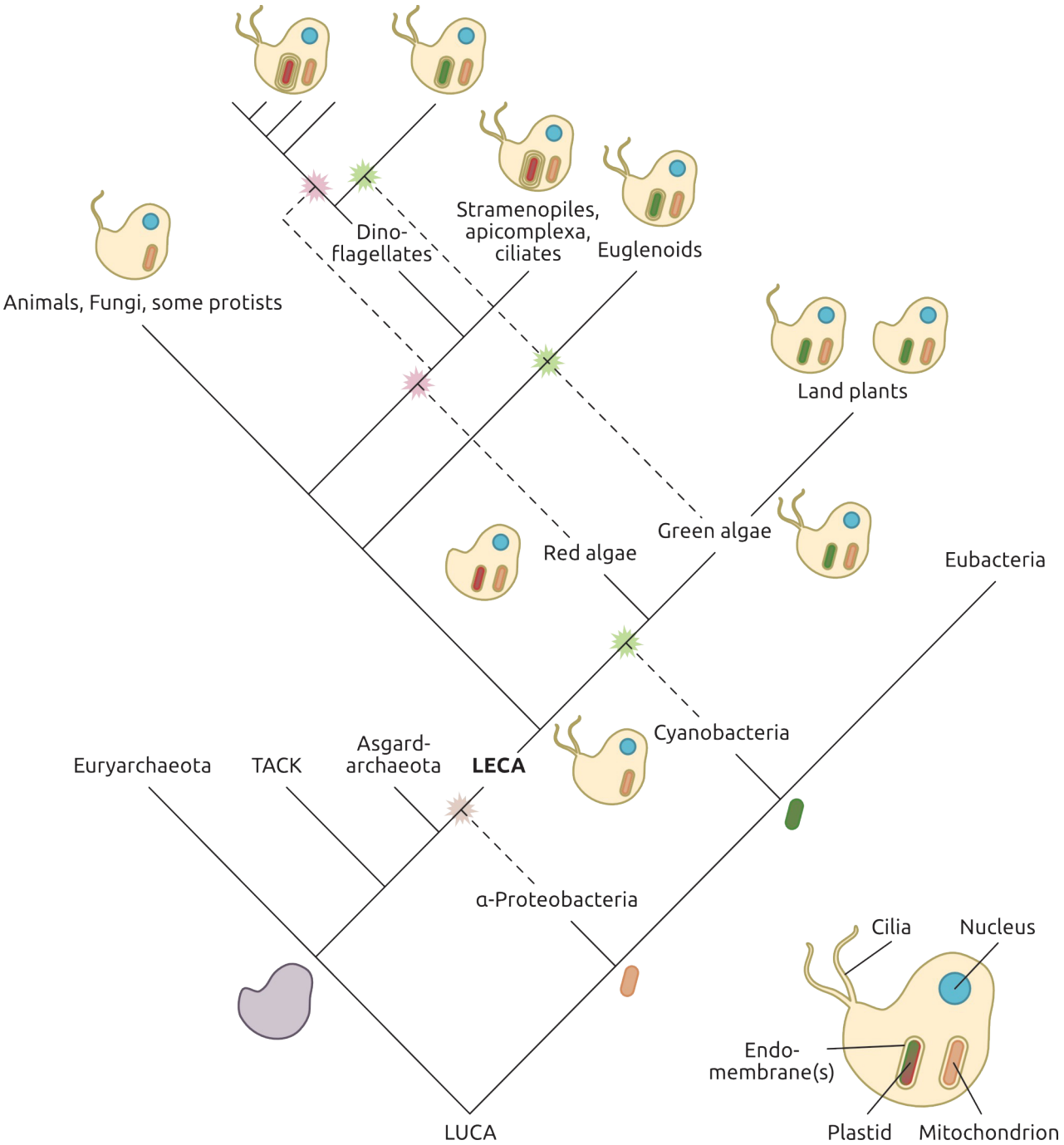


Figure 5: The Tree of Life from an endosymbiotic point of view. The last universal common ancestor (LUCA) split into the eubacterial and archaeal lineage. The dawn of eukaryotes began with the engulfment of an α-Proteobacterium by a relative of Asgardarchaea, resulting in the last eukaryotic common ancestor (LECA). A second endosymbiosis event occurred between a descendent of LECA and a cyanobacterium, resulting in green and red algae. Several secondary and tertiary endosymbioses of red and green algae and previous symbiotic eukaryotes by other eukaryotes resulted in multi-layered plastids seen in several lineages. LECA, last eukaryotic common ancestor; LUCA, last universal common ancestar; TACK, Thaumarchaeota, Aigarchaeota, Crenarchaeota, Korarchaeota. Adapted from Keeling, 2010, McFadden, 2001, and Rout and Field, 2017.

1.5.1. FECA

The first eukaryotic common ancestor (FECA) was most likely related to Lokiarchaea and the vessel from which eukaryotes evolved (Spang et al., 2015; Zaremba-Niedzwiedzka et al., 2017). It contained relatives of crucial information processing tools found in extant eukaryotes, including homologues to the eukaryotic translation initiation factor EIF1, Topoisomerase IB, RNA polymerase G (Spang et al., 2015), DNA polymerase ε-like, RNA polymerase A (Spang et al., 2017), and ancestrally related versions of the proteins that compact and bind DNA in eukaryotes, histones (Sandman and Reeve, 2006, 1998). This ancestral cell additionally contained membrane-deforming proteins that are likely the ancestors of eukaryotic membrane coat proteins (Samson et al., 2008), but will be discussed later. This cell then went on to acquire an endosymbiotic α-Proteobacterium, although it is currently not clear by which mechanism. Many argued for endocytosis as the basis for the engulfment and subsequent incorporation of the bacterium into the archaeal cytosol based on the theory of Margulis, then named Sagan (Cavalier-Smith, 2002; Doolittle, 1998; Poole and Gribaldo, 2014; Roger et al., 2017; Sagan, 1967). The incorporated bacterium would then in principle be an unprocessed food vesicle, “an undigested meal” (Bremer et al., 2022). Recent advances in phylogenetic reconstruction (Bremer et al., 2022) and the consideration of environmental factors (like low oxygenation) during early eukaryotic evolution (Mills et al., 2022) now diverge from the previous theories by tracing back mitochondria, but not phagocytosis (or even phagotrophy), to LECA. Additional biochemical studies show that phagocytosis and membrane-dependent ATP synthesis are mutually exclusive, and that a phagocytotic lifestyle would not have been beneficial to the archaeal host (Martin et al., 2017). Phagocytosis is however not needed to acquire mitochondria in the first place, as recent studies show that membrane fusion events can be seen in extant archaea without homologues to phagocytotic processes (Naor and Gophna, 2013). It is then conceivable to imagine a scenario where free-living relatives of facultatively aerobic α-Proteobacteria were in such close proximity to the future anaerobic archaeal host (Martin and Müller, 1998) that the larger host’s cytoplasm could engulf the bacterium at one point, rendering the bacterium unable to escape, and thus leading to replication of the bacterium within the new host cell (Fig. 6). This theory does not necessitate an explanation of how the bacterium was not digested, and paves the way to a metabolism-based symbiotic origin of mitochondria: under anaerobic conditions within the host, the bacterium could produce hydrogen, delivering a source of food to the host. The relationship between host and symbiont could then be further developed on this basis, resulting in genome reduction of the

endosymbiotic bacterium and lateral gene transfer to the host genome. Additionally, this hypothesis explains how the archaeal isoprene ether membrane could be replaced by a fatty acid ester membrane found in eukaryotes and bacteria (Gould et al., 2016). This early-mitochondrion-hypothesis provides a robust biochemical and genetic base for further innovations in the developing eukaryote on its way to LECA.

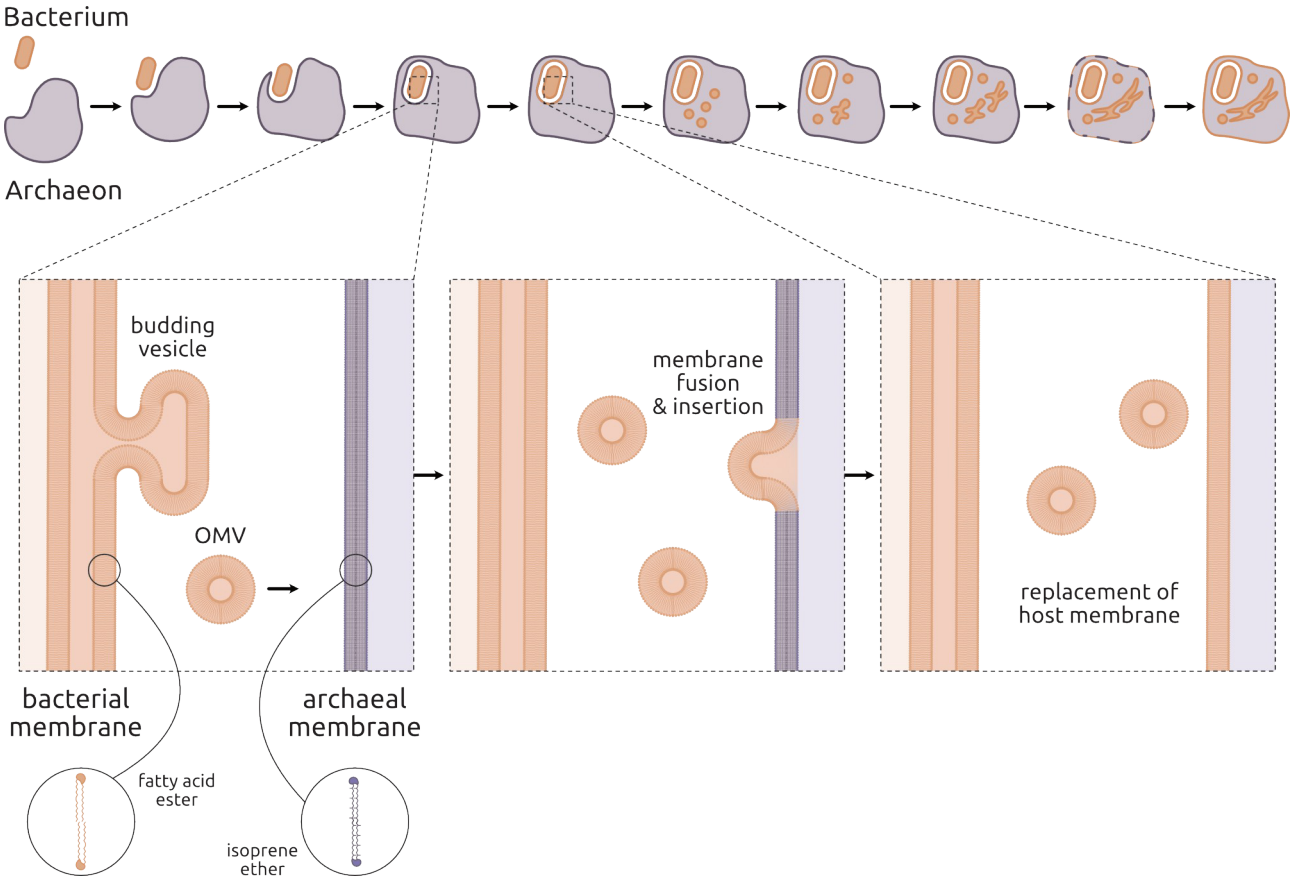


Figure 6: Theory for endosymbiotic replacement of archaeal isoprene ethers with bacterial fatty acid esters. Secreted outer membrane vesicles (OMVs) composed of fatty acid esters originating from the engulfed α-Proteobacterium probably first fused with the host membrane surrounding the bacterium, slowly replacing the isoprene ethers. Further secretion into the host's cytosol led to accumulation of OMVs, fusion thereof, and resulted in a crude endoplasmic reticulum (ER). OMVs then further integrated into the host's plasma membrane, and with them the vesicles from the proto-ER, probably along with proteins anchored in the fatty acid ester layer. Evolutionary advantage then led to a gradual loss of the isoprene ether membrane. Adapted from Gould et al., 2016.

1.5.2. THE TRANSITIONAL PROTO-EUKARYOTE

The endosymbiosis theory of archaeal host and bacterial symbiont results in a quandary regarding cell morphology: While it explains the energy metabolic advantages of such a relationship to early proto-eukaryotes, it also results in a non-eukaryotic cell membrane. The archaeal membrane is comprised of isoprene ethers, while eukaryotic membranes are fatty acid esters. Interestingly, mitochondria, bacteria and archaea all possess the ability to secrete vesicles from their outer

membrane (Benaim et al., 1990; Braschi et al., 2010; Cook et al., 2014; Deatherage and Cookson, 2012; Grimm et al., 1998; Lee et al., 2009; Schwechheimer and Kuehn, 2015; Soubannier et al., 2012). This ability might have shaped the early proto-eukaryote's cytosol and outer membrane: With nowhere else to go, outer membrane vesicles (OMVs) of the bacterial endosymbiont could only accumulate in the host's cytoplasm, eventually fusing and forming a structure akin to a proto-ER, or travelling to the host's membrane and gradually replacing the ethers with esters (Gould et al., 2016). This establishes the basic functionality and topology of the ER in extant eukaryotes, without the need of evolutionary innovation other than the endosymbiotic transition.

With a proto-ER in place and a source for energy at hand, the proteomic makeup of the archaeal host provides a solid ground for evolutionary innovation of existing systems. As mentioned before, archaea possess homologues to eukaryotic information processing. With a source of fatty acid esters that is the bacterial OMV-derived proto-ER, the remaining eukaryotic membrane-bounded organelles could have evolved as a consequence of co-evolution of archaeal and bacterial secretion systems and membrane-deforming proteins. And indeed, many of the proteins associated with vesicular transport and membrane-defining proteins have homologues in both archaea and bacteria (Dacks and Field, 2007, 2007; Field et al., 2011; Field and Dacks, 2009; Leung et al., 2008; Rout and Field, 2017; Schlacht, 2016).

1.5.3. LECA

The intricate interplay between archaeal and bacterial proteins and systems likely catalysed the innovation that enabled eukaryotic cells to thrive in diverse ecological niches. Emerging research proposes that LECA was characterised by distinctive features such as the cilium and efficient recombination of genetic information through sexual reproduction (Bremer et al., 2023; Kontou et al., 2022; Koumandou et al., 2013; O'Malley et al., 2019; Skejo et al., 2021). LECA likely existed as a syncytial cell, affording it the ability to buffer against instabilities and deleterious mutations that could be disadvantageous for single-celled organisms (Skejo et al., 2021). Phylogenetic analysis of LECA reveals a surprisingly complex entity equipped with characteristic organelles seen in contemporary eukaryotes, boasting proteomic diversity that often equals or exceeds that of present-day eukaryotic organisms (Koumandou et al., 2013). A compelling focus emerges on membrane coating complexes, dynamic entities that have profoundly influenced the identity of organelles during eukaryotic evolution and continue to evolve in present-day species. In order to manage the intricacies of multiple organelles, LECA featured a well-defined array of distinct

trafficking pathways to and from organelles and the plasma membrane, ensuring accurate cargo delivery and precise transport directionality.

1.6. MEMBRANE COATING COMPLEXES

On its way to becoming LECA, the proto-eukaryote not only remodelled a source for energy metabolism (via the acquisition of mitochondria), it also acquired a sophisticated endomembrane system, with distinct organelles for specialised functions such as protein synthesis, import and export of vesicles and biomolecules, and digestion, as well as a container for genetic information. The organellar constitution of extant eukaryotes encompasses the ER, GA, endo-/lysosomes, and the nucleus, a set that was likely present in LECA (Dacks and Field, 2007; Elias et al., 2012; Field et al., 2007; Field and Dacks, 2009; Koumandou et al., 2007). All of these compartments require a scaffold that deforms and stabilises membranes in a reversible fashion, i.e. to not only form vesicles or curved membranes, but also to release them again (Fig. 7). Different classes of membrane coats act in a concerted manner to fulfil these tasks and are predominantly found in only one type of membrane, although exceptions to this rule exist. In the vane of this notion, the organellar paralogy model was formulated, stating that by diversification of an ancient, uniform set of coating proteins, the membrane-bounded organelles were shaped and assigned each a subset of proteins to facilitate directionality in trafficking (Dacks and Field, 2007).

1.6.1. Protocoatomer-derived coats

The endosymbiotic relationship allowed the proto-eukaryote to choose from a wide variety of adaptable protein architectures that were once exclusive to either archaea or bacteria. One class of proteins that evolved into membrane coats are the α-solenoid/β-propeller domain proteins. The basic fold motifs of these proteins can be traced back to prokaryotes, although separately (Chaudhuri et al., 2008; Karpenahalli et al., 2007; Kobe and Kajava, 2000; Smith, 2013); the combination of both is only widespread in eukaryotes, not prokaryotes (Devos et al., 2004). These folds are readily obtainable and have low requirements for sequence conservation, making them malleable templates to adapt to different functions and drive evolution (Marcotte et al., 1999). Due to their high sequence variability, definitive ancestors to proteins of these families are hard to trace back before LECA, but it is

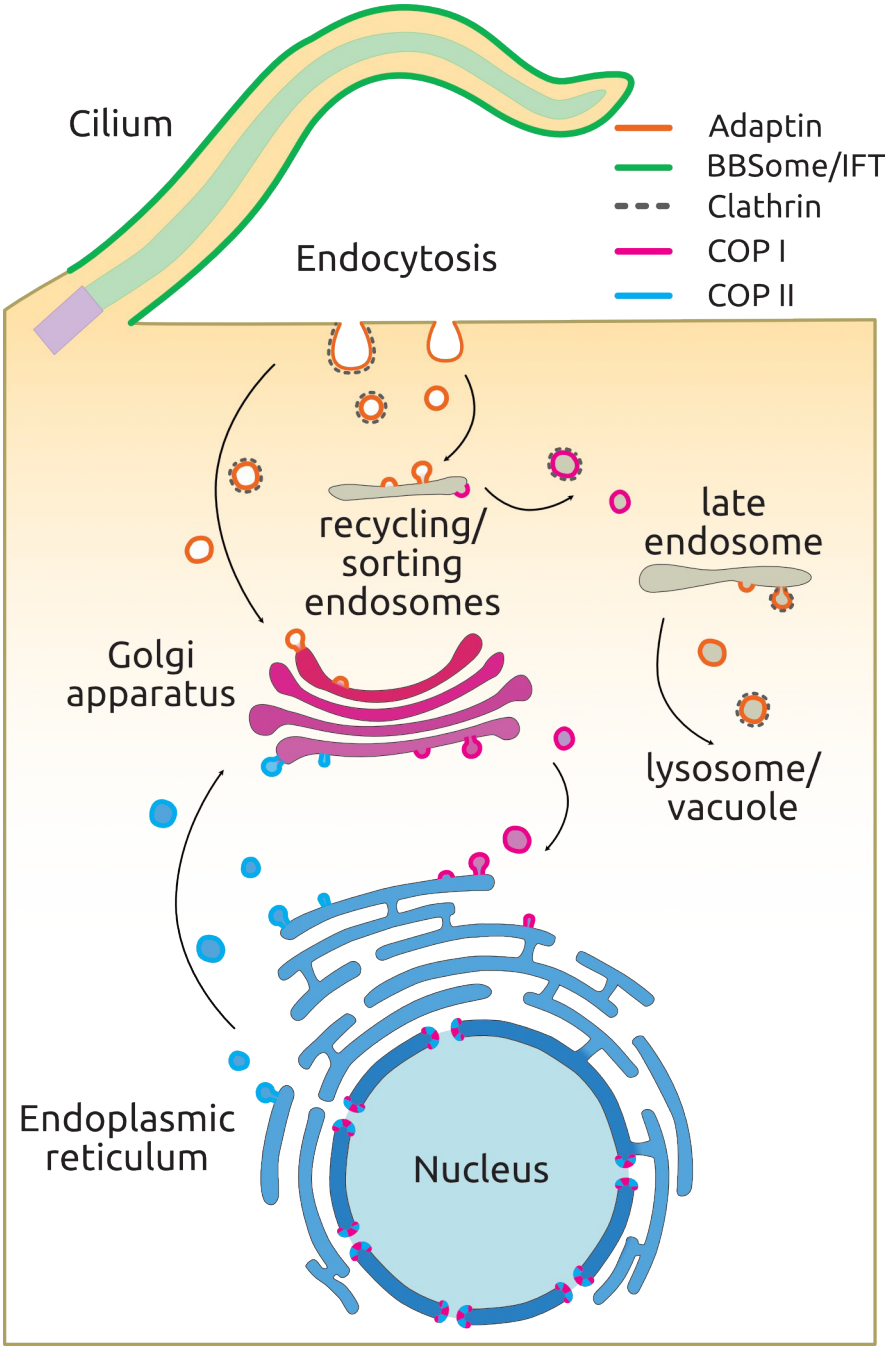


Figure 7: Overview of protocoatomer-derived vesicle coats in eukaryotes. Coats can roughly be grouped into type I coats (adaptin, BBSome/IFT, clathrin, COP I) and type II coats (COP II). While type I coats mostly enable inward trafficking from plasma membrane to endosomes and Golgi apparatus (GA), and GA to endoplasmic reticulum (ER), type II coats facilitate trafficking from ER to GA. Note that IFT/BBSome proteins are technically not vesicle coats, but are related via their protocoatomer ancestor. Adapted from Field et al., 2011.

likely that each of the groups was already present in LECA (Field et al., 2011). In fact, the combination is thought to have emerged rather early in the transition period after FECA, with Armadillo and HEAT repeats being widely present throughout eyukaryotes (Andrade et al., 2001; Cingolani et al., 1999; Smith et al., 1999). The relatively low evolutionary pressure on sequence conservation led to a rapid diversification of these proteins and novel roles in various locations (Andrade et al., 2001; Malik et al., 1997; Smith et al., 1999). Many extant membrane coats have this arrangement of domains, which lead to the formulation of the protocoatomer hypothesis (Devos et al., 2006, 2004; Dokudovskaya et al., 2011; Field and Dacks, 2009; van Dam et al., 2013). The protocoatomer is a hypothetical complex of membrane-deforming protein coats that gave rise to three distinct groups of membrane coats seen in extant eukaryotes: the type I coats, type II coats, and the IFT/BBSome proteins (Beck et al., 2018; Dacks and Robinson, 2017; Jékely and Arendt, 2006; Jin et al., 2010; Satir et al., 2008; van Dam et al., 2013). Other notable members are HOPS/CORVET, SEA and clathrin family proteins (Field et al., 2011), but are less relevant for the topic of this work; we will focus on the closely related type I, type II, and IFT coats and their entangled evolutionary history.

1.6.1.1. TYPE I COATS

The type I family of coats consists of COPI, adaptins, and clathrin. In the secretory pathway from ER to GA and the plasma membrane/endosomes, COPI can be found on vesicles that originate from the *cis*-GA and travel towards the ER (Beck et al., 2009; Duden, 2003; Lee et al., 2004; Spang, 2009). They can also be found on some endosomal cargoes, and are required for lipid homeostasis, mRNA transport, and NE breakdown (Aniento et al., 1996; Beller et al., 2008; Bi et al., 2007; Daro et al., 1997; Gabriely et al., 2007; Gu et al., 1997; Guo et al., 1994; Liu et al., 2003; Prunuske et al., 2006; Razi et al., 2009; Soni et al., 2009; Thiam et al., 2013; Todd et al., 2013; Trautwein et al., 2004; Wilfling et al., 2014; Zabezhinsky et al., 2016). Adaptins bridge the gap of retrograde trafficking of vesicles from the plasma membrane inwards to the *trans*-GA and to early and recycling endosomes. They further sort cargo to the late endosome or lysosome and thus play crucial roles in homeostasis. Adaptins also serve as cargo adaptors of clathrin-coated vesicles and plasma membrane invaginations (Edeling et al., 2006; Paczkowski et al., 2015). Clathrin is recruited to vesicles budding from the plasma membrane, the *trans*-Golgi network, and endosomes, and forms a cage coating the vesicle (Vigers et al., 1986).

1.6.1.2. TYPE II COATS

Type II family coats are COPII and HOPS /CORVET (Koumandou et al., 2007). COPII proteins facilitate the anterograde transport of vesicles from the ER towards Golgi, while the HOPS/CORVET coats are also involved in early to late endosomal sorting, and late endosome to lysosome/vacuole transport.

1.6.1.3. IFT AND BBSOME

Although technically derived from a membrane-deforming ancestor complex (van Dam et al., 2013), the proteins facilitating IFT in eukaryotic cilia are not able to deform membranes by themselves (Jin et al., 2010). However, like NPCs, they are associated with membranes. The ancestral structure of IFT and BBSome proteins BBS4 and BBS8 is related to COPIε (van Dam et al., 2013) and first split into the proto-BBSome and proto-IFT, the latter of which then split into IFT-A and IFT-B (van Dam et al., 2013). IFT complexes are heterooligomeric complexes, with IFT-A comprised of six different proteins (IFT-144, 140, 139, 122, 121, and 43), while IFT-B contains 16 proteins (IFT-172, 88, 81, 80, 74, 70, 57, 56, 54, 52, 46, 38, 27, 25, 22, and 20). The BBSome consists of eight proteins from four subclasses: proteins containing a β-propeller and an α-solenoid followed by a ɣ-adaptin ear (BBS1) and followed by an α/β-platform and α-helix (BBS2, BBS7, and BBS9); TPR proteins (BBS4, 8); PH-domain proteins with C-terminal α-helices (BBS5); and short α-helical linker proteins (BBS18) (Fig. 8). The BBSome itself is assembled in a step-wise process (Klink et al., 2020) aided in animals by the CCT/TRiC-associated chaperonin-like BBS proteins BBS6, 10, and 12 (Seo et al., 2010; Zhang et al., 2012). While parts of IFT and the BBSome share the overall α-solenoid/β-propeller structure of other protocoatomer-derived coats (van Dam et al., 2013), they are also strikingly conserved on a sequence level across eukaryotic species (Ewerling et al., 2023b; Hodges et al., 2011, 2010). This is most likely owed to their highly conserved functions in ciliary trafficking. IFT complexes link ciliary cargo proteins to motors like cytoplasmic dynein 2/dynein 1b and kinesin-2 to facilitate movement to and within the ciliary lumen (Ou et al., 2005; Prevo et al., 2017), where kinesin-2 moves towards the ciliary tip, and dynein 1b towards the ciliary base. The BBSome on the other hand links the IFT particles to ciliary cargo, like GPCRs to traverse the diffusion barrier at the base of cilia. All complexes are needed for classical ciliogenesis and functionality of cilia, albeit to a differing extent. While in mammals, loss of BBS proteins has detrimental effects and results in ciliopathies (Badano et al., 2006; Beales, 2005; Braun and Hildebrandt, 2017), the green alga *Chlamydomonas reinhardtii* has a reduced role for the BBSome, and is viable with mutated BBS proteins (Lechtreck et al., 2009).

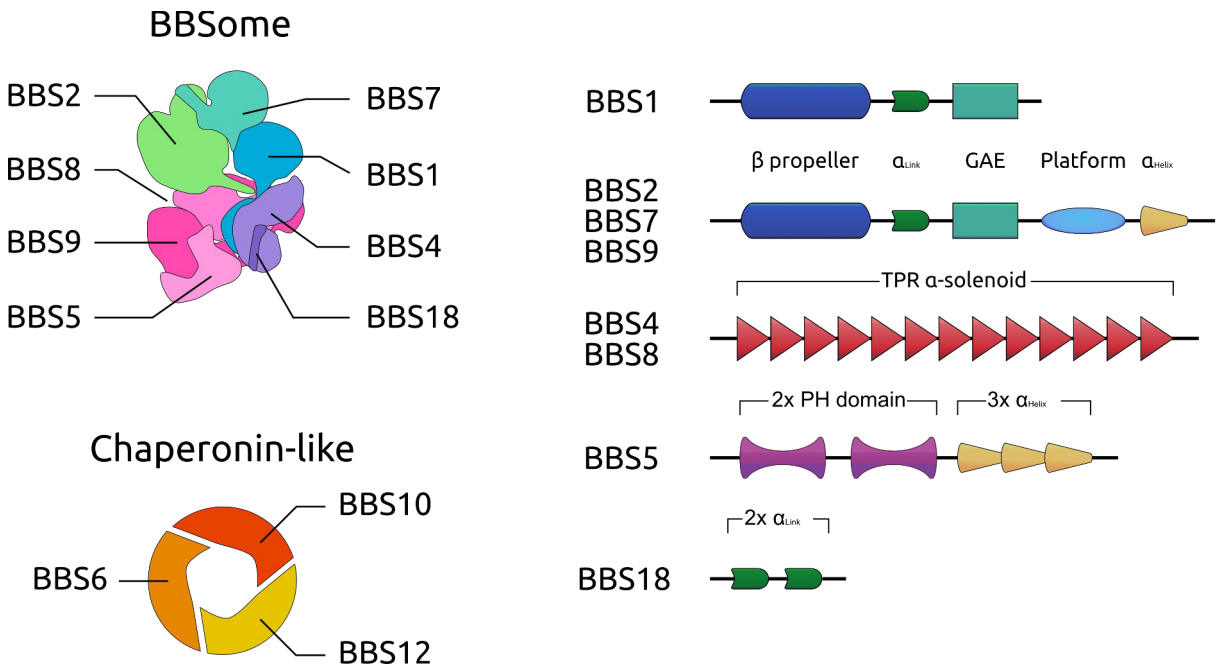


Figure 8: The BBSome, its protein families, and chaperonin-like BBS proteins. The BBSome assembles into a heterooctameric complex, with BBS1, 2, and 7 forming a 'head', and BBS4, 5, 8, 9, and 18 forming the 'body'. Proteins from the BBSome can be grouped according to their protein architecture (right). The chaperonin-like BBS proteins BBS6, 10, and 12 have less strictly defined protein domains, but share a basic palindromic architecture of equatorial domain, intermediate domain, apical domain, intermediate domain, equatorial domain. Adapted from Ewerling et al., 2023b.

1.7. NON-CILIARY FUNCTIONS OF CILIARY PROTEINS

Recent studies could show that parts of the ciliary trafficking complexes also localise to other compartments and structures than the cilium. Interestingly, they do so on their own rather than as a complex (Asiedu et al., 2009; Bachmann-Gagescu et al., 2011; Chaki et al., 2012; Choi et al., 2013; den Dulk et al., 2008; Ewerling et al., 2023b; Gascue et al., 2012; Horwitz and Birk, 2021; Hsiao et al., 2009; Kim et al., 2005; May-Simera et al., 2009; Robert et al., 2007; Scott et al., 2017; Shi et al., 2018; Zhu et al., 2015), indicating that their non-ciliary function depends on the protein itself and is likely independent of their function in their respective complexes. Ciliary proteins are therefore most likely multifunctional proteins. While their function and relevance are most commonly studied in a ciliary context, the non-ciliary functions are however not to be neglected. In mammals, the non-ciliary functions may even be causative of multisystemic diseases (Scott et al., 2017), and may contribute to phenotypic variations in different ciliopathies as suggested above. While the ciliary functions might be well documented, examining the non-ciliary functions may shed light on the evolutionary relationship between different organelles and by extension, early eukaryogenesis. Most of the information about non-ciliary functions currently comes from studies

in animal systems, owed to their relevance to understanding the disease phenotype of various ciliopathies. However, given the conserved character of ciliary proteins, it is highly likely that other members of this group also have additional functions outside of Metazoa, offering alternative study systems to explore these functions..

1.7.1. MICROTUBULAR CARGO TRANSPORT

Ciliary trafficking originated from cytosolic MT trafficking. As such, it is reasonable to find ciliary proteins taking up a role in cytosolic vesicle trafficking. During mammalian ciliogenesis, the BB moves from the vicinity of the nucleus into close proximity of the plasma membrane where preciliary vesicles attach to the distal end of the BB. This process, along with endocytotic activity, is disrupted when Ahi1 is not present in murine cells (Hsiao et al., 2009). Ahi1 is a scaffolding protein required for Rab8 localisation to the BB during ciliogenesis. Mutations in the human AHI1 homologue lead to Joubert syndrome, a severe ciliopathy with predominantly neural and kidney phenotypes. There is obviously a dual role for Ahi1 in ciliary and cytosolic vesicle transport. A similar function can be ascribed to the BB protein cc2d2a, a ciliary protein involved in the correct localisation of transmembrane receptors of photoreceptor outer segments in zebrafish: mutations in the gene underlying the protein lead to Joubert syndrome by a similar mechanism, affecting vesicle trafficking in photoreceptor cells and leading to the mislocalisation of Rab8 (Bachmann-Gagescu et al., 2011). A recent study has implied that translocation of the BB to the immune synapse relies on BBS1 in human T-cells (Cassioli et al., 2021), one of the few cell types to be known that does not assemble cilia.

1.7.2. CELL CYCLE CONTROL AND MITOSIS

Cilia and cell cycle are intrinsically linked. In mammals, only cells in interphase display cilia, and disassemble them prior to cell cycle entry. The reason behind this is that the ciliary BB doubles as the centriole in animals, and can exert its functions as spindle organiser during mitosis only after the cilium is absent. In mammals, BBS6 localises to the material surrounding the centriole and is a crucial factor during cytokinesis, the process of daughter cell separation. Loss of function leads to mislocalised BBS6 during cytokinesis and results in multinucleated and multicentrosomal cells (Kim et al., 2005). The essential IFT component IFT88 normally localises to IFT trains and is involved in protein transport in the cilium; it can be also found at centrosomes in mammals (Robert et al., 2007). Overexpression of IFT88 leads to cell cycle disruption and apoptotic cell death, while silencing increases transition from G_1 into S, G_2 and M phases, clearly indicating a regulatory role.

The MT structure of the cilium often leads to dual functions of ciliary proteins in mitosis and mitotic spindle functionality: another protein mutated in Joubert syndrome, CSPP, is responsible for the correct orientation of the spindle, and attachment to kinetochores (Asiedu et al., 2009; Zhu et al., 2015). PCM1, an important constituent of the pericentriolar matrix and scaffold of several centriole-associated proteins, also aggregates at cytosolic MTs which are assembled and disassembled in a cell-cycle-dependent manner (Balczon et al., 1994; Kubo and Tsukita, 2003). Loss of PCM1 function leads to cell cycle arrest (Balczon et al., 2002).

1.7.3. NUCLEUS-ASSOCIATED FUNCTIONS

The regulatory functions of cilia proteins in the nucleus are the most critically understudied aspect of non-ciliary functions. Anecdotal evidence for nuclear localisation exists, like IFT proteins in the nucleus of Paramecium tetraurelia (Shi et al., 2018), but has only been extended recently through comprehensive studies in mammalian systems (Ewerling et al., 2023b). The nuclear interactome of ciliary proteins is however still overwhelmingly elusive, although some studies elucidated this aspect of cilia biology in recent years. One of the first discovered cases of possible gene regulatory functions is BBS6 which interacts with DNMT1, thereby potentially modulating the epigenetic signature of cells to an unknown extent (May-Simera et al., 2009). Further evidence hints at a role for certain BBS proteins to facilitate transport of transcription factors across the NE, a process that modulates gene expression of downstream targets in response to factors like cytotoxic stress (Horwitz and Birk, 2021; Scott et al., 2017). Defective interaction of BBS6 with SMARCC1 (SWI/SNF related, matrix associated, actin dependent regulator or chromatin subfamily c, member 1) underlie congenital heart disease, as the nuclear import of chromatin remodelling protein SMARCC1 is regulated by BBS6 (Scott et al., 2017). BBS4 on the other hand is required for translocation of transcription factor ATF6α into the nucleus after induction of ER stress (Horwitz and Birk, 2021). Ciliary proteins also appear in DNA damage signalling cascades, an effect that can most often be found in patients suffering from renal ciliopathies (Chaki et al., 2012; Choi et al., 2013; den Dulk et al., 2008). More recent studies implied multiple nuclear interactors for BBS10, although the exact functions remain unknown (Marchese et al., 2022). BBS7, which normally localises to the BB, interacts with RNF2 in the nucleus (Gascue et al., 2012). RNF2 (E3 ubiquitin ligase RING2) is a member of the polycomb group (PcG) and ubiquitylates histone H2A, altering epigenetic signatures and promoting transcriptional repression of target genes (Bentley et al., 2011; Cao et al., 2005; Gao et al., 2014; Luo et al., 2021; McGinty et al., 2014; Taherbhoy et al., 2015; Wang et al., 2004). Clearly, nuclear roles for ciliary proteins are not just a “happy accident”, but

rather an understudied aspect crucial to understanding the evolution of eukaryotic organelles and with implications for human diseases.

1.8. AIM OF THIS WORK

Ciliary proteins are highly conserved throughout the eukaryotic Tree of Life. The process of building and maintaining cilia is fundamentally the same in all eukaryotic lineages and relies on the same core sets of proteins. For these processes, the BBSome is a critical component as it facilitates movement of effector molecules like receptors to and within the cilium. Hence, BBS proteins found in the BBSome are highly conserved as well, and commonly found in all ciliated species. In some species where cilia assembly is BBSome-independent (like *Toxoplasma gondii*) or where cilia have been abandoned completely, some BBS homologues still persist. Is this a commonality across the eukaryotic tree of life? Have ciliary BBS proteins been repurposed to affect other processes? Many BBS proteins show an affinity to the nucleus, and some have been found to take part in gene regulatory processes, like transcription factor shuttling between cytosol and nucleus, or modulating the half-life of histones via interaction with ubiquiting ligases. This raises three pressing questions:

1. Is this nuclear function a confounding part of BBS protein functions?
2. Have the nuclear functions been established in LECA, or have they been acquired independently?
3. Can the high conservation of BBS proteins help understand nuclear functions between eukaryotes, and ultimately, allow predictions for BBS protein functions in other species?

In this work, we aimed to answer the above questions via both bioinformatic and cell biological approaches. First, we consolidate the preservation of BBS proteins across eukaryotes with state-of-the-art bioinformatical modelling and homologue searches. The identified homologues can then be screened for their propensity to enter the nucleus by predicting nuclear signal sequences. To rule out any influence of open or closed mitosis on the incoporation of BBS proteins into the nucleus of dividing cells, we also statistically test for the effect of mitosis on nuclear localisation likelihood. We transfer our predictions to mammalian models to validate the bioinformatic approach, and to gain insights which proteins can enter the nucleus under normal conditions. With the data collected, we can place the findings in an evolutionary context and make predictions about possible nuclear localisations in other eukaryotes, while pinpointing specific proteins and taxa that are more likely to have evolved nuclear functions. Second, we broaden the current knowledge about the conservation

of BBSome proteins in non-model organisms as species beyond the scope of established laboratory models might give a better framework to study non-canonical roles for ciliary proteins. We chose the insects as a representative clade, as insects build specialised cilia in only two tissues, and BBS1 has been implicated in shaping social phenotypes of ants (Wallberg et al., 2016). Insects in general might therefore be a suitable subject to study non-ciliary functions of ciliary proteins, with emphasis on nuclear roles. Third, with the insights gained, we critically review the evolutionary history of endomembrane organelles and membrane coating complexes, including the BBSome, and integrate our findings into the greater evolutionary context of eukaryogenesis. We elucidate possible scenarios in the transition from prokaryote to eukaryote that could have led to the dual functions of ciliary BBS proteins as nuclear effectors, and draw conclusions about the impact of a possible co-evolution of nucleus and cilia on clinical and evolutionary research.

2. NEOFUNCTIONALISATION OF CILIARY BBS PROTEINS TO NUCLEAR ROLES IS LIKELY A FREQUENT INNOVATION ACROSS EUKARYOTES

2.1. PREFACE

Ciliary BBS proteins have been well studied in regard of their ciliary functions in the BBSome (Akella et al., 2020; Eguether et al., 2014; Klink et al., 2020; Lechtreck et al., 2013, 2009; Liew et al., 2014; Nachury et al., 2007; Singh et al., 2020; Tian et al., 2023) or for their functions in assembling a functional BBSome (in the case of the chaperonin-like BBS proteins) (Seo et al., 2010; Zhang et al., 2012). Their non-ciliary functions have however only recently come to the attention of researchers. BBS proteins have been implicated in regulation of cell cycle progression (Kim et al., 2004), T-cell polarity (Cassioli et al., 2021), and metabolism (Marchese et al., 2022), but could also be found to aid translocation of transcription factors across the NE (Horwitz and Birk, 2021; Scott et al., 2017). Interestingly, they interact with nuclear proteins such as ubiquitin ligase RNF2 (Gascue et al., 2012) and DNMT1 (May-Simera et al., 2009), hinting towards a gene regulatory function. In this work we screened all human BBSome and chaperonin-like BBS proteins for their ability to enter the nucleus of non-ciliated human cells by ectopic expression and found that most of the BBS proteins can actually enter the nucleus. The differential localisation patterns in both subcellular fractionation and immunofluorescence analysis suggest that nuclear import and export are differentially regulated for each protein, and that they exert distinct functions in the nucleus. We predicted homologues in species sampled from major eukaryotic lineages, and found that the BBSome is conserved largely as a set, although exemptions to this rule became apparent. We further investigated if any of the BBS proteins or a specific clade if eukaryotes is more likely to contain NLS or NES sequences, and used this as an approximation for nuclear localisation prediction. We found that BBS4 and BBS9 have a high likelihood to contain NLS, while BBS7 has a NES signature. Phylogenetic reconstruction of ancestral sequences showed that NLS in BBS4 and BBS9 evolved multiple times during eukaryotic radiation, but that BBS7 NES is most likely ancestral and traces back to LECA. To rule out an influence of open or closed mitosis on the incorporation of BBS proteins into the nucleus during mitosis, we employed permutation analysis and found that mitosis does not influence the prevalence of NLS and NES in our datasets. We therefore conclude that most BBS proteins can enter the nucleus in humans, and that this nuclear localisation is potentially conserved throughout eukaryotes.

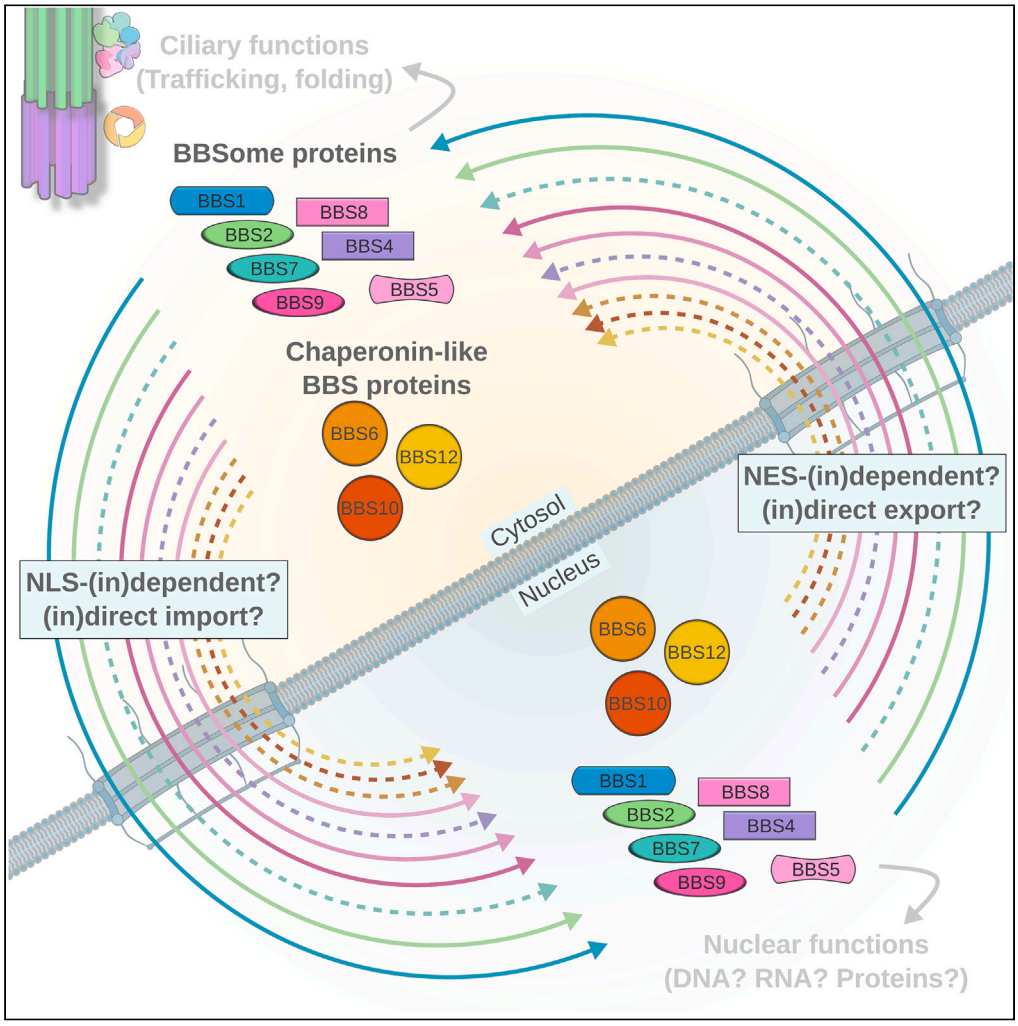
This part of the thesis was published on iScience on March 16, 2023, and appears in Volume 26, Issue 4, 106410, April 21, 2023 (https://doi.org/10.1016/j.isci.2023.106410).

iScience

CellPress
OPEN ACCESS

Article

Neofunctionalization of ciliary BBS proteins to nuclear roles is likely a frequent innovation across eukaryotes



Alexander Ewerling, Vanessa Maissl, Bill Wickstead, Helen Louise May-Simera

may-simera@uni-mainz.de

Highlights

BBSome is conserved as a set from LECA; chaperonin-like BBS arose early in Opimoda

Mode of mitosis does not influence on NLS/NES prevalence in BBS proteins

Nuclear BBS functions potentially arose multiple times independently in eukaryotes

Nuclear localization predictions are partially recapitulated in mammalian cells

Ewerling et al., iScience 26, 106410
April 21, 2023
https://doi.org/10.1016/j.isci.2023.106410

iScience

CellPress
OPEN ACCESS

Article

Neofunctionalization of ciliary BBS proteins to nuclear roles is likely a frequent innovation across eukaryotes

Alexander Ewerling,[1] Vanessa Maissl,[1] Bill Wickstead,[2] and Helen Louise May-Simera[1,3,*]

SUMMARY

The eukaryotic BBSome is a transport complex within cilia and assembled by chaperonin-like BBS proteins. Recent work indicates nuclear functions for BBS proteins in mammals, but it is unclear how common these are in extant proteins or when they evolved. We screened for BBS orthologues across a diverse set of eukaryotes, consolidated nuclear association via signal sequence predictions and permutation analysis, and validated nuclear localization in mammalian cells via fractionation and immunocytochemistry. BBS proteins are—with exceptions—conserved as a set in ciliated species. Predictions highlight five most likely nuclear proteins and suggest that nuclear roles evolved independently of nuclear access during mitosis. Nuclear localization was confirmed in human cells. These findings suggest that nuclear BBS functions are potentially not restricted to mammals, but may be a common frequently co-opted eukaryotic feature. Understanding the functional spectrum of BBS proteins will help elucidating their role in gene regulation, development, and disease.

INTRODUCTION

Eukaryotic cilia and flagella are structurally related organelles of motility and sensation. They were one of the characteristic features of the last eukaryotic common ancestor (LECA) and are conserved in most eukaryotic lineages.[1,2] Cilia/flagella (herein collectively referred to as "cilia") are microtubule-based protrusions that extend from a cytosolic basal body into the extracellular space where they can perform a wide range of functions. Since their emergence in the eukaryotic Tree of Life, cilia have facilitated locomotion[3–5] and reception of extracellular cues even in the earliest single-celled eukaryotes.[6–11] While single-celled eukaryotes show only motile cilia multicellular organisms exhibit two major ciliary subtypes, motile cilia that can move fluids across membrane surfaces,[12–15] and immotile primary cilia that have evolved to serve as sensory signaling organelles crucial for tissue development, maturation, and function.[16–20] Mutations in ciliary proteins result in multiple distinct clinical phenotypes caused either by the impaired signaling or regulatory function of primary cilia, or the diminished motility of motile cilia.[21–23]

The basic construction plan and mechanisms underlying ciliary assembly, disassembly, and maintenance are highly conserved and strikingly similar across otherwise very different eukaryotic taxa.[24–26] Ciliary proteins are found in all major eukaryotic lineages,[8,27–29] and their fundamental functions are homologues in ciliated species.[30] To ensure proper ciliary functions cilia need a highly regulated in- and efflux of transmembrane receptors, soluble effector molecules, and phospholipids. A specialized intraflagellar transport (IFT) comprising three distinct complexes—IFT-A, IFT-B, and the BBSome, a complex of Bardet-Biedl Syndrome (BBS) proteins—mediates much of this movement. The IFT complexes couple with molecular motors of the Kinesin-2 and cytoplasmic dynein 2 (also called dynein 1b) families to mediate anterograde and retrograde transport along the ciliary axoneme, respectively.[31,32] The BBSome serves as an adaptor for ciliary cargo such as transmembrane proteins and links it to the IFT complexes. The human heterooctameric protein complex consists of BBS1, BBS2, and BBS7 ("head"), and BBS4, BBS5, BBS8, BBS9, and BBS18 ("body").[33,34] Its assembly is facilitated by the CCT/TRiC-associated chaperonin-like BBS proteins BBS6/McKusick-Kaufman Syndrome (MKKS), BBS10, and BBS12 before entering the cilium (Figure 1A).

Alongside their highly conserved ciliary roles, many ciliary proteins have been found to localize to other cellular compartments where they function in non-ciliary processes. These include proteins that participate

[1]Institute of Molecular Physiology, Faculty of Biology, Johannes Gutenberg-University Mainz, Mainz, Germany

[2]School of Life Sciences, University of Nottingham, Nottingham, UK

[3]Lead contact

*Correspondence: may-simera@uni-mainz.de

https://doi.org/10.1016/j.isci.2023.106410

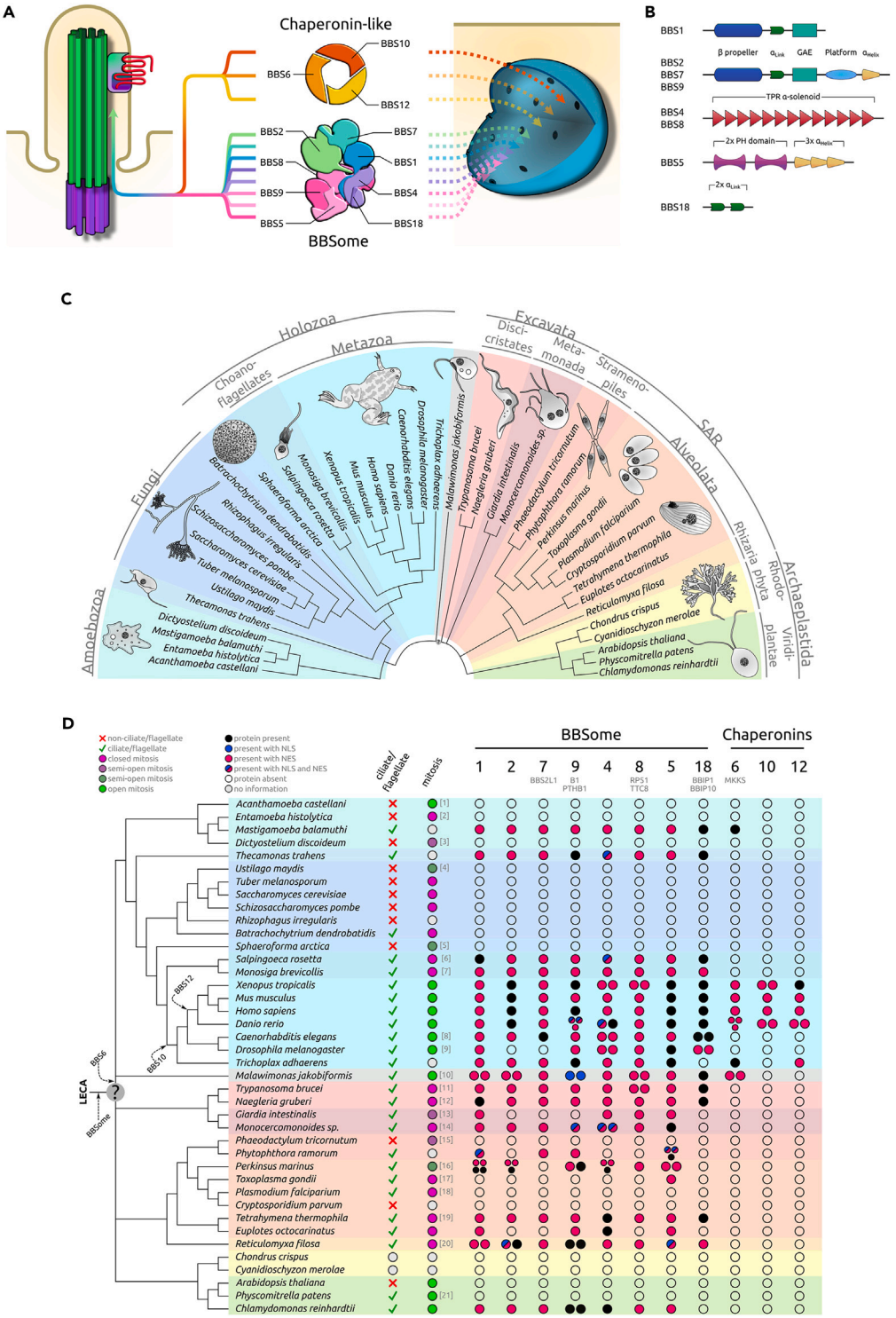


Figure 1. BBS proteins in the eukaryotic Tree of Life

(A) Schematic of BBSome and chaperonin-like BBS complexes (adapted from Singh et al., 2020).

(B) Domain and secondary-structure architecture for BBSome proteins. Domain abbreviations: γ-adaptin ear (GAE), tetratricopeptide repeat (TPR), pleckstrin homology (PH).

(C) Cladogram showing likely phylogenetic relationships between 40 diverse eukaryotes used in this study.

Figure 1. *Continued*
(D) Distribution of BBS proteins across eukaryotes. Presence (filled circle) or absence (open circle) of orthologues of each protein are indicated, along with predictions of NLS and/or NES within each sequence. Multiple circles indicate number of orthologues. Whether the organism is known to build cilia/flagella in some cell types and the mode of mitosis operating (see Table S1) are also indicated.

in mitosis, spindle pole orientation, and cytokinesis[35–38] and other cell-cycle-linked, nuclear processes, such as DNA damage response.[39,40] Ciliary trafficking proteins might also be active in the nucleoplasm; BBS4 has been shown to associate with transcription factors that affect neuronal endoplasmic reticulum stress response,[41] and BBS7 interacts with RNF2, actively regulating the expression of its target genes, likely by modulating the turnover rate of RNF2.[42] Further, the disturbed nucleo-cytoplasmic translocation of BBS6/MKKS in patients has been identified as a contributing factor to congenital heart disease.[43]

To exert a nuclear function, proteins first need to localize to the nucleoplasm. However, the nuclear envelope (NE) forms a highly selective barrier that prevents large molecules from entering the nucleus solely by diffusion.[44] Proteins circumvent this barrier by active transport across the nuclear membrane, allowing precise spatiotemporal management of nuclear interactors. Alternatively, cytoplasmic proteins can potentially access the interior of the nucleus in some lineages as a result of NE breakdown during mitosis. The nucleoplasm and chromatin of vertebrates are highly accessible to cytoplasmic proteins during mitosis from NE breakdown at prometaphase through to telophase. This may allow the BBS proteins to freely diffuse into the former nucleus and interact with their nuclear targets. However, the degree to which the NE becomes permeable differs considerably in different eukaryotic lineages. In animals and land plants mitosis is open with substantial fragmentation of the NE. Hence animal proteins associated with ciliary function (such as BBS4, BBS6, and BBS7) may gain entry to the nucleus during cell division. In contrast, many protists and fungi maintain an apparently intact NE, or show only partial breakdown or fenestration throughout mitosis. Nonetheless, even in organisms with closed mitosis, it appears that ciliary proteins have the potential to gain nuclear functions. For example, in the ciliate *Paramecium tetraurelia* both macro- and micro-nuclei remain intact throughout mitosis, but IFT57 appears to gain access to the macronuclear space.[45]

These findings raise questions as to how and when "ciliary" proteins might have been co-opted into roles in the nucleus. Do these dual functions represent recent neofunctionalization of proteins or are some potentially more ancestral? Is there evidence of additional neofunctionalization in lineages with open mitoses? Alternatively, are particular ciliary proteins more likely to gain nuclear import/export signatures? Can the evolutionary history of ciliary proteins help us predict non-ciliary functions? Here, we look at the conservation of BBS proteins across eukaryotes and search for evidence of potential nuclear localization. We screen a database of forty disparate eukaryotic genomes for BBS protein orthologues by means of phylogenetic inference and correlate them with the organisms' mode of mitosis to pin-point potential co-evolution of nuclear function and NE breakdown. We then predict potential nuclear localization and export signals of those orthologues. Finally, we perform immunocytochemical stainings and fractionations of transfected human embryonic kidney (HEK293T) cells to look for BBS protein entry into the nucleus in mammalian cells.

RESULTS

Distribution of BBS orthologues across the eukaryotic tree of life

To identify BBS orthologues across eukaryotes, we assembled a dataset of 40 predicted proteomes from diverse eukaryotes including organisms that build cilia/flagella during their life cycle and organisms which do not (but might have retained BBS proteins for non-ciliary functions) (Figure 1C). Potential BBS orthologues were identified using BLASTp similarity searches to seed more sensitive hidden markov model-based searches, and the resulting sets checked for potential errors by construction of phylogenetic trees (see STAR Methods). As expected, presence of a complete BBSome strongly correlates with the presence of cilia during an organism's life cycle (Figure 1D). The presence/absence of orthologues also agrees with previous studies that identified BBS proteins in a smaller set of eukaryotes,[26,46,47] demonstrating that our bioinformatic approach is robust against previous analyses (exceptions for single proteins are described in Figure S1).

Our data confirm that BBS proteins normally appear in an "all-or-nothing" fashion, as would be expected for a protein set involved in a common conserved biological process. Notable exceptions here are species that have a reduced set of BBS proteins despite building cilia—including *Drosophila melanogaster*, *Giardia*

intestinalis, *Phytophthora ramorum*, and *Euplotes octocarinatus*—all of which lack detectable copies of at least two core BBSome components. In addition, *Plasmodium falciparum* and the fungus *Batrachochytrium dendrobatidis* appear to lack all BBS-related proteins despite having flagellate stages, while *Toxoplasma gondii* lacks all but BBS5. While *D. melanogaster* is known to maintain neuronal cilia and sperm flagella without a BBSome,[48,49] little is known about composition or presence of a functional BBSome in *Giardia*, *Phytophthora*, and *Euplotes*. Given that there is an already reduced role for cilia in general in *D. melanogaster* and we see a reduction in BBSome components, the gradual loss of these proteins might precede a loss of the BBSome.

One theory for the seemingly "penalty-free" loss of single BBSome proteins might be that some functions of the missing proteins can be substituted by related BBS components. In general, four major groups of BBSome proteins can be distinguished by their tertiary structure (Figure 1B): a) proteins comprised a β-propeller and an α-solenoid followed by a γ-adaptin ear (GAE) domain such as BBS1; b) proteins as in a), but followed by an α/β-platform and α-helix (BBS2, BBS7, and BBS9); c) tetratricopeptide repeat (TPR) proteins (BBS4 and BBS8); and d) proteins with two pleckstrin homology (PH) domains followed by three α-helices (BBS5). To some extent agreeing with this, all BBSome sets identified here contain at least one member of each group present, apart from *Phytophthora* and *Toxoplasma*.

It has been observed that the absence of BBSome proteins often precedes the loss of ciliary structure during eukaryotic evolution.[26] The lack of BBSome components in *Batrachochytrium* and *Plasmodium* are in agreement with this; both *Batrachochytrium* and *Plasmodium* have flagellar assembly and disassembly pathways that happen entirely in the cytoplasm and do not require a functional BBSome.[46,50,51] In contrast to the complete loss of the BBSome components in *Plasmodium*, the closely related *Toxoplasma* species appears to have retained a single orthologue of just one BBSome component, BBS5. The identified protein is annotated as a predicted BBS5 orthologue in the genome database (toxodb.org) and was a strong hit to the BBS5 Hidden Markov model (HMM) here, suggesting this is a true retention and not a mis-identification. Moreover, maximum-likelihood phylogenetic analysis of BBS5 orthologues along with related paralogues shows that the identified *T. gondii* BBS5 (ToxoDB: TGME49_259965) is part of a clade containing known BBS5 orthologues, and the identified BBS5 from *Giardia* (Figures 2A and S2D). This makes a very strong argument for the hit to be a true orthologue of BBS5, despite the absence of other BBSome components.

In contrast to the transport functions of the BBSome, chaperonin-like BBS proteins (BBS6, BBS10, and BBS12) are needed to assemble and fold a functional BBSome. Previous work has shown that these chaperonin-like BBS proteins are related to the chaperonin-containing T-complex protein (TCP) subunit (CCT) proteins, most likely through duplication of a CCT8-like sequence early in eukaryotes.[52,53] These chaperonin-like BBS proteins are mostly conserved in metazoa. However, potential orthologues of BBS6 were also detected in *Mastigamoeba balamuthii* and *Malawimonas jakobiformis*. Both of these appear to be strong BBS6 candidates (Figure S2E), suggesting either an earlier origin for this group of proteins than the base of the metazoa, or gene transfer of metazoan BBS proteins into other lineages.

In addition to likely BBS6 proteins outside of metazoa, a single non-animal BBS12-like sequence was identified by our search method in *Dictyostelium discoideum*. This sequence is not a member of any known CCT family (members of which can be easily identified in *Dictyostelium*; Figure S2Ja) and clusters with good support with *bona fide* BBS12 sequences, albeit at the base of the clade. BLASTp searches using this sequence as a query identify archaeal thermosome subunits as the best matches outside of the clade *Dictyostelia* (data not shown). To test if this might represent a recent gene transfer to *Dictyostelium* rather than a divergent copy of BBS12, we included the three top BLAST hits from the NCBI "nr" database not originating from dictyostelial species into the phylogeny, as well as the hits from searching 48 representative archaeal genomes with the same HMM used for eukaryotic searches. However, no archaeal protein or set of proteins has a strong affinity for this *Dictyostelium* BBS12-like sequence, which is still sister to the animal BBS12 clade (Figure 2B). The phylogeny is thus, consistent with this being a non-animal BBS12 orthologue, and no obvious origin from within either eukaryotic CCT proteins or by lateral gene transfer from prokaryotes can so far be identified. However, given the distribution of BBS proteins and the lack of obvious BBS12 in other organisms that do not build cilia, this should be treated with caution.

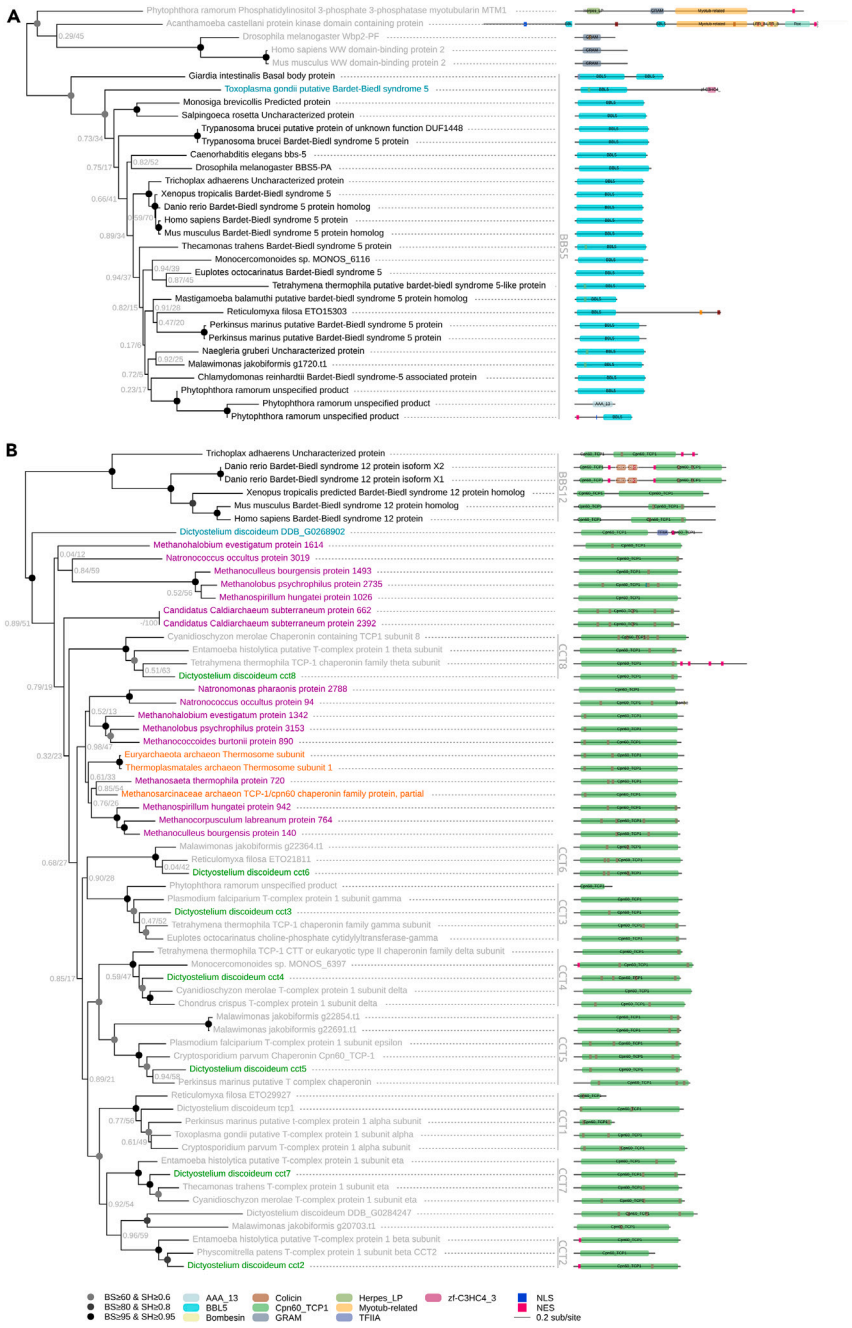


Figure 2. Maximum-likelihood (ML) phylogenetic relationships of BBS5 and BBS12

Bootstrap (BS) and Shimodaira-Hasegawa (SH) branch support values are given at each node (color-coded nodes are darker where support is stronger).

(A) Reconstruction of ML phylogenetic tree for BBS5 orthologues.

(B) Extended phylogeny of BBS12 in the initial database species with chaperonin-containing T-complex protein (TCP) subunit (CCT) orthologues of *Dictyostelium discoideum* (green), top three reciprocal BLASTp hits (orange) for *Dictyostelium discoideum* DDB_G0268902 (cyan; dictybase.org: DDB_G0268902) against NCBI's "nr" database, and orthologous proteins found in an additional HMM search of 48 archaeal proteomes (purple).

BBS orthologues carry nuclear import and export signatures

Recent studies have shown that human BBS proteins BBS6 and BBS7 can localize to the nuclear compartment.[42,43] This raises the question whether the nuclear functions of BBS proteins are ancient or a recently

acquired attribute. We reasoned that if the nuclear role were ancestral to a set of organisms and conserved, there might also be a conserved means of nuclear import/export for that set. To test for potentially conserved nuclear functions, we interrogated the BBS orthologues for predicted nuclear localization signals (NLS) and nuclear exit signals (NES).

Individual BBS orthologues were tested for statistically higher occurrence of predicted NLS or NES scores against sets of equal size made up by sampling from comparison datasets of control proteins. Three comparison datasets were considered: 1) all BBS proteins (to test for individual orthologues being more likely to be nuclear than the BBS set as a whole; 212 sequences in total), 2) all proteins from the proteomes analyzed that are predicted to be cytosolic by WolfPSORT[54] (to test for an individual orthologue being more likely to be nuclear than a "typical" cytoplasmic protein; 91,829 sequences), and 3) a dataset of all proteins (whether nuclear or not) from the predicted proteomes (711,242 sequences; Figure 3A).

In our analysis, no single BBS protein was found to be significantly more likely to have a higher NLS than a random set of the same size of either other BBS proteins, "cytoplasmic" or total proteins from our reference datasets (Figure 3B). The strongest predicted NLS is seen in BBS4 and BBS9 orthologues ($p_{BBS} = 0.104$). In contrast, the predicted NES signal in BBS7 is significantly higher than expected against any of the three control datasets ($p_{BBS} = p_{cyto} = p_{total} = 0.002$; Figure 3C). The BBSome in general is significantly more likely to have a greater predicted NES score than typical proteins predicted to be cytosolic ($p_{cyto} = 0.041$), unlike chaperonin-like BBSs, which are not significantly more likely to have a higher NLS or NES signature than a protein set of the same size sampled from either BBS, cytosolic or whole proteomic datasets (Figure S3). Given that individual BBS proteins, and not all proteins of the complex, have a tendency to have higher nuclear signal sequence scores, these data suggest that the evolution of nuclear functions might involve single BBS proteins and not redistribution of the entire complex.

Presence of nuclear signal sequences is not phylogenetically restricted

To exclude the possibility that certain organisms might have an *a priori* tendency to have stronger predicted nuclear signal sequences—either through some underlying biology or (more likely) reflecting bias in the predictive algorithms—we also tested whether the proteins from certain groups might have stronger signatures than our reference datasets, considering data at different taxonomic depths. While we see no bias when comparing the broadest or finest taxonomic differentiation (Figure 4A/C), BBS orthologues in *Excavata* in general have significantly stronger NES signal than would be expected for typical predicted cytosolic proteins in this group (Figure 4B; $p_{cyto} = 0.042$, permutation test) or for BBS, "cytosolic" or general proteins from across eukaryotes (Figure S4B). This means that BBS orthologues in *Excavata* are not only more likely to bear NES signatures than other excavate proteins, but also compared to all proteins in this analysis. Interestingly, stramenopiles also show a higher-than-average NLS capacity in BBS orthologues compared to other cytosolic proteins (Figure 4C, $p_{cyto} = 0.045$), although it should be noted that prediction of cytosolic localization may not be equally accurate across all eukaryotic lineages. Nonetheless, these data suggest excavates and stramenopiles may be interesting as models for non-ciliary nuclear BBS protein functions.

Nuclear signal sequences in BBSome proteins are ancestrally conserved

By reconstructing the most likely sequence from which extant BBS orthologues diverged, we can infer how well nuclear signal sequences are conserved at hypothetical nodes within the eukaryotic tree. The reconstructed "ancestral" sequences can be used for signal sequence predictions as in previous analyses. Upon ancestral state reconstruction, it appears that most predicted NLS in BBS proteins are protein-specific (Figure 5). As reconstruction only considers shared indels, this method is unable to account for NLS occuring in these regions and therefore misses some predictions in the orthologues sequences themselves (Figure 1C). Interestingly, for BBS4, there is a predicted NLS at the node between choanozoa and metazoa, but the signal in this inferred sequence is not inherited in the majority of the sequences arising from it. These data suggest classical NLS signals arise sporadically across the phylogeny of BBS proteins and are not generally conserved over evolutionary timescales.

In contrast and as previously described, NES are far more widespread and found more readily in the reconstructed sequences. In general, BBSome proteins all have a detectable NES in the ancestral LECA sequence, except for BBS9 and BBS18. Predicted NES in BBS9 are mostly found in extant proteins although weaker NES can be detected at the base of Opimoda and the stramenopiles-alveolates-rhizaria (SAR)

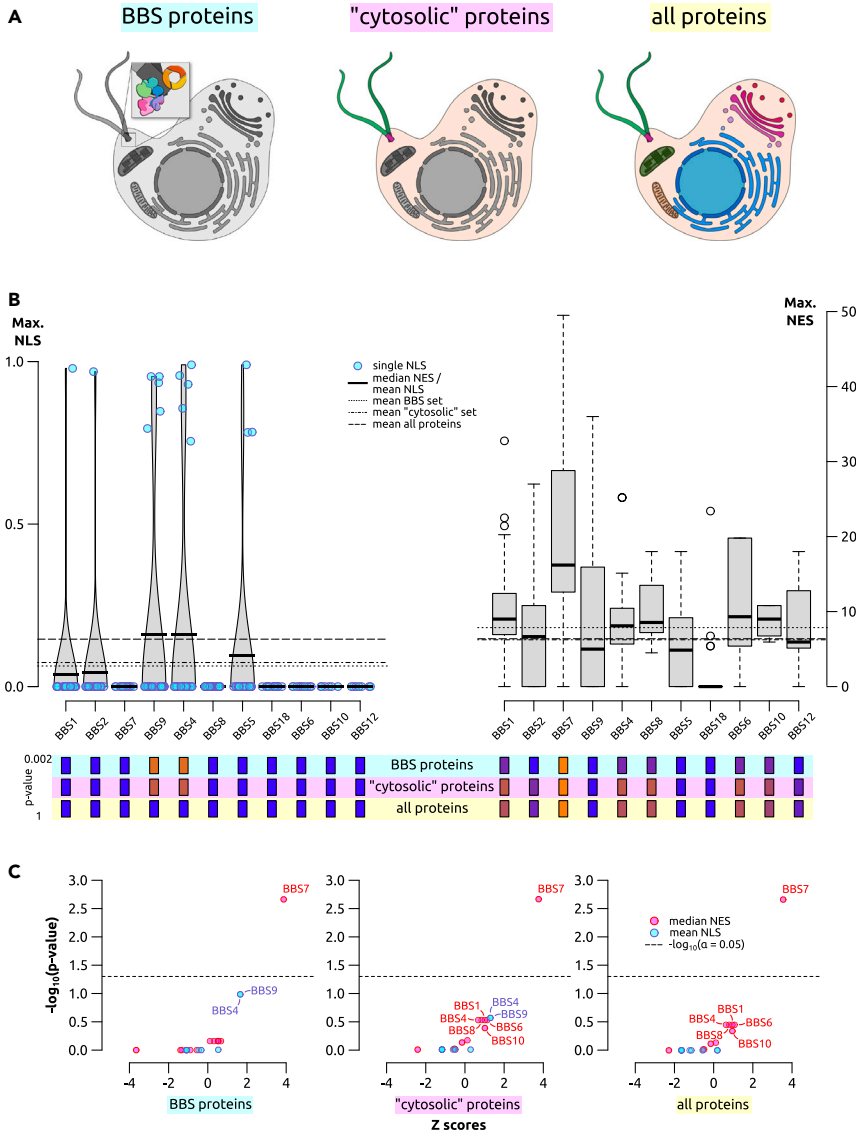


Figure 3. Distribution of predicted nuclear signal sequences in BBS orthologues

(A) Schematic representation of the datasets considered for the prevalence comparison of nuclear signal sequences to those in BBS proteins.

(B) Distribution of NLS (left) and NES (right) in BBS orthologues. *Left*: Violin plots of highest predicted NLS from BBS orthologues. *Right*: Boxplots of highest predicted NES from BBS orthologues of the respective groups indicated. Whiskers and boxes mark quartiles; circles: points outside 1.5 interquartile range.

(C) Volcano plots of Z-scores from BBS protein orthologues against adjusted-$\log_{10}$(p value). Z-scores are calculated as multiples of standard deviations from the reference group mean.

supergroup. Strongest NES at LECA-level was predicted in BBS7 that is in line with previous analyses of extant protein sequences (Figure 3B). Although not as strong, BBS4 and BBS8 show a highly conserved NES signature throughout the reconstructed sequences, followed by BBS1, BBS2, and BBS5. And despite being recent additions, the likely ancestral sequences of chaperonin-like BBS proteins BBS6, BBS10, and BBS12 show predicted NES at all nodes.

Taken together, the ancestral state reconstruction suggests that the nuclear localization of BBS4 is likely a convergent feature of sequences in choanoflagellates and animals. Further it predicts that BBSome proteins most likely have NES since LECA and that they have largely maintained these signal sequences

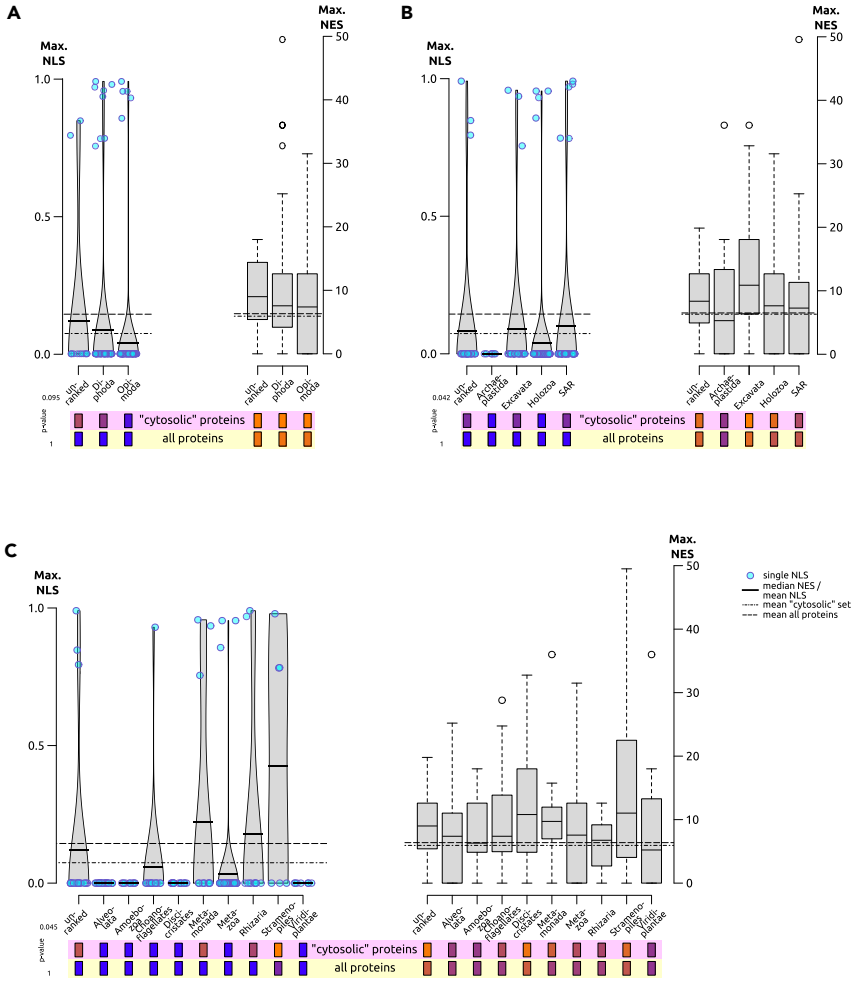


Figure 4. Comparison of BBS protein nuclear signal sequence scores to other proteins of the same taxonomic unit

Tests were carried out for BBS proteins of one higher taxonomic unit vs. "cytosolic" or any proteins of the same taxonomic unit. (A), (B), and (C) show the comparisons at different taxonomic depths.

throughout many divergence events, while chaperonin-like BBS were a later addition but also might have served nuclear functions since their emergence.

Mode of mitosis does not influence nuclear localization signals or nuclear exit signals occurrence

Translocation of a protein from cytoplasm to nucleoplasm can happen either due to active import and export across the NE, or inclusion of proteins in the nascent nucleus during mitosis. The latter would not require NLS, but does require a mitotic breakdown of the NE and may also be associated with NES to promote re-localization of proteins to the cytoplasm. Conversely, we might expect increased NLS signal in species without mitotic NE breakdown if the nuclear functions of BBS are conserved. Lastly, if the mode of mitosis had no influence on the nuclear functions of BBS proteins (i.e., there was no co-evolution of nuclear functions and NE breakdown), we expect no significant difference between signal sequences in species with open vs. closed mitosis.

As the need for import into, or export from the nucleus might be substantially confounded by access granted to the nucleus during open mitosis, we tested whether mode of mitosis influences the distribution of predicted NLS/NES in BBS proteins. As organisms exhibit a range of degrees of NE breakdown during mitosis, organisms were scored based on their position along a closed/open mitosis axis (Figure 6A) to compute a weighted score

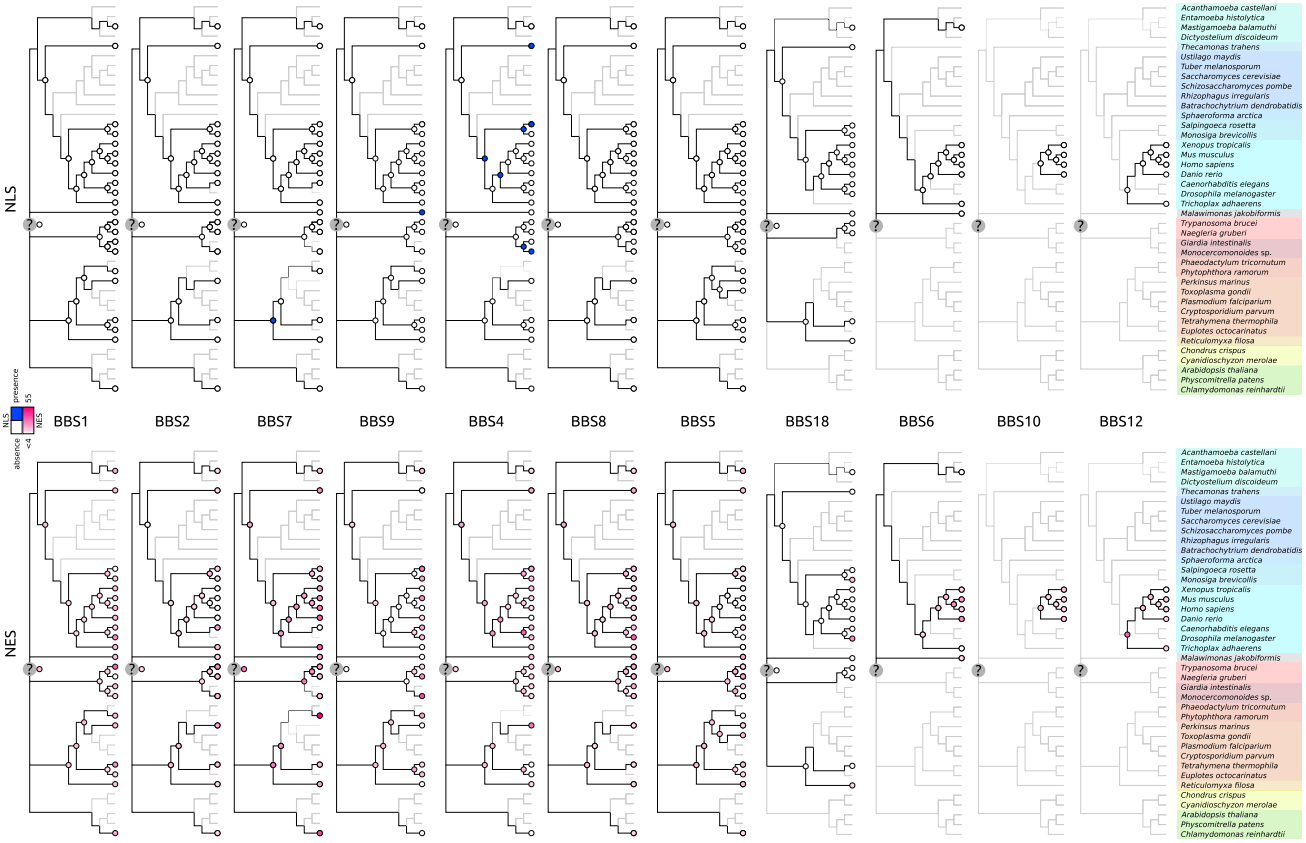


Figure 5. Ancestral sequence reconstruction of BBS proteins

Most likely ancestral sequences were reconstructed to assess their capability to enter the nucleus by prediction of NLS and NES.

sum as a test statistic (see STAR Methods). The statistic was then used in a permutation test for an association of NLS/NES with the mode of mitosis. The proteomic dataset of our previous analyses is well suited for this since it is comprised species representative of the mitotic spectrum (Figure 1D).

Our analysis shows the mean of the weighted NLS sum for all proteins is close to 0, indicating that NLS signals are equally likely in proteins from organisms having either open or closed modes of mitosis. The weighted NLS score for BBS proteins is slightly negative, but not more than expected by chance (p_{NLS} = 0.2108). Interestingly, there is a shift of weighted NES score sums for typical proteins in each proteome toward the more positive end of the open-closed-axis (Figure 6B) showing that higher predicted NES signals are more common for proteins in organisms with more open forms of mitosis, as would be expected—although this is not significantly different from a bias that could be expected by chance. BBS proteins follow this trend no more than an average protein from these organisms (p_{NES} = 0.2626).

Taken together, these results suggest that the presence of NLS and NES signals in BBS proteins are more due to biological adaptations of single eukaryotic lineages other than their mode of mitosis. Statistically, we could show that mitosis alone does not seem to be the evolutionary driver behind the emergence of nuclear signal sequences in eukaryotic BBS proteins. In particular, this means that the higher-than-average NES signature we found in excavates, and the above-average NLS scores in stramenopiles cannot be explained by the organsims' form of mitosis alone. This suggests that the evolution of nuclear functions for BBS proteins is not linked to transitions from closed to open mitosis.

Human BBS proteins differentially localize to the nucleus

Computationally predicted localization signals can aid in understanding the mechanisms of protein-dependent nuclear import and export, but are of course only an approximation for their actual subcellular

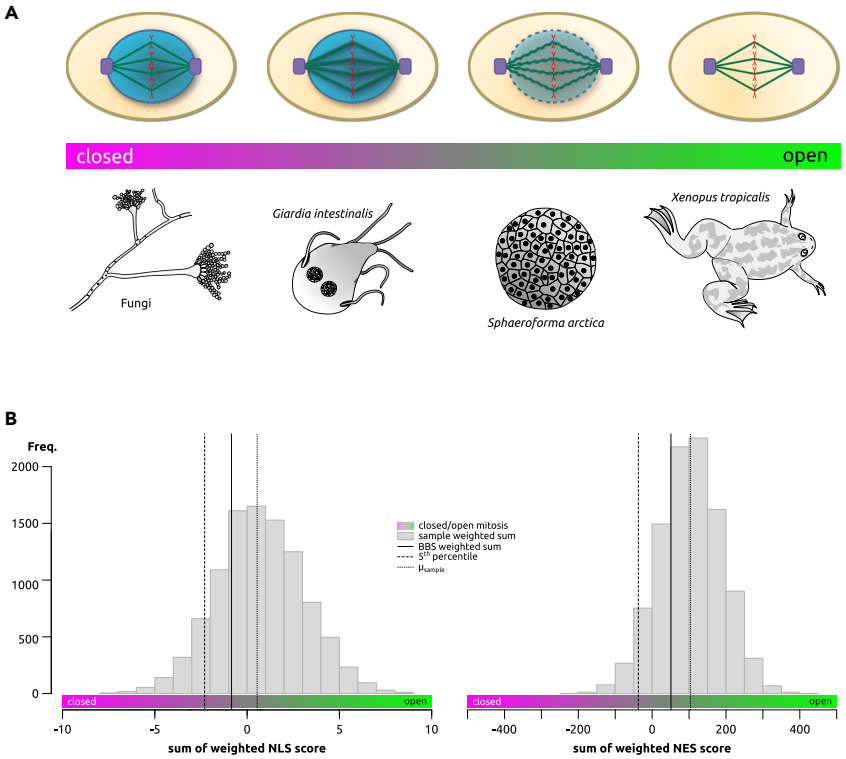


Figure 6. Mode of mitosis and presence of predicted nuclear signal sequences are not correlated

(A) Mitotic spectrum from closed to open mitosis and representative species.

(B) Histogram of samples for permutation testing of nuclear signal sequences with weighted score sums of BBS proteins' predicted NLS (left) and NES (right).

localization. To test predictions for nuclear localization of BBS proteins against actual cell biology, we performed localization studies in transiently transfected HEK293T cells overexpressing FLAG-tagged BBS proteins. We looked at their localization (cytosolic or nucleoplasmic) via subcellular fractionation followed by Western Blotting (WB), and immunocytochemical staining (Figure 7).

All BBS proteins tested were expressed, and can be seen in the cytosolic fraction (Figure 7A). When overexpressed in HEK293T cells, BBS1, BBS2, BBS7 and BBS9, BBS4 and BBS8, and BBS5 also localize to the nucleus. This nuclear localization is most likely the result of active transport between cytosol and nucleus: The size-exclusion barrier that is the nuclear pore does not allow free diffusion of molecules with a mass > 40 kDa[55,56] that all of the nuclear BBS proteins (but BBS5) have. Of the BBS1"family" proteins that localize to the nucleus, BBS1 is most prominently featured in the nuclear fraction, despite lacking a classical NLS. In comparison, BBS2 and BBS9 are found only in trace amounts in the nucleus, while BBS7 has a higher ratio in the nuclear fraction than in the cytosolic fraction. The strong nuclear signal of the WB;however, was only partially reproduced by immunostaining (Figure 7B). Although BBS1 and BBS2 show a weaker nuclear than cytosolic staining, BBS7 is virtually undetectable via immunostaining in the nucleus. BBS9 has the strongest signal, although we detected it in an amount comparable to BBS2 via Western blot. BBS4 and BBS8 on the other hand are about equally present in cytosolic and nuclear fractions. We see the same pattern in immunofluorescence labeling of BBS8-overexpressing cells, but not in their BBS4 counterpart. BBS4 seems to be exclusively localized around the centrosomes, but not anywhere in the nucleoplasm. Despite lacking both classical NES and NLS, we find a signal for BBS5 in the nuclear fraction and in immunocytochemical staining. Theoretically, the two pleckstrin-homology-like domains of BBS5[47] could facilitate an interaction with nuclear phosphoinositides that have been implicated in chromatin remodeling transcriptional regulation.[57]

The only chaperonin-like BBS proteins detectable in a nuclear fraction were BBS10 and, surprisingly, BBS12. While BBS10 has been detected in the nucleus in a recent study with a similar experimental setup,[58]

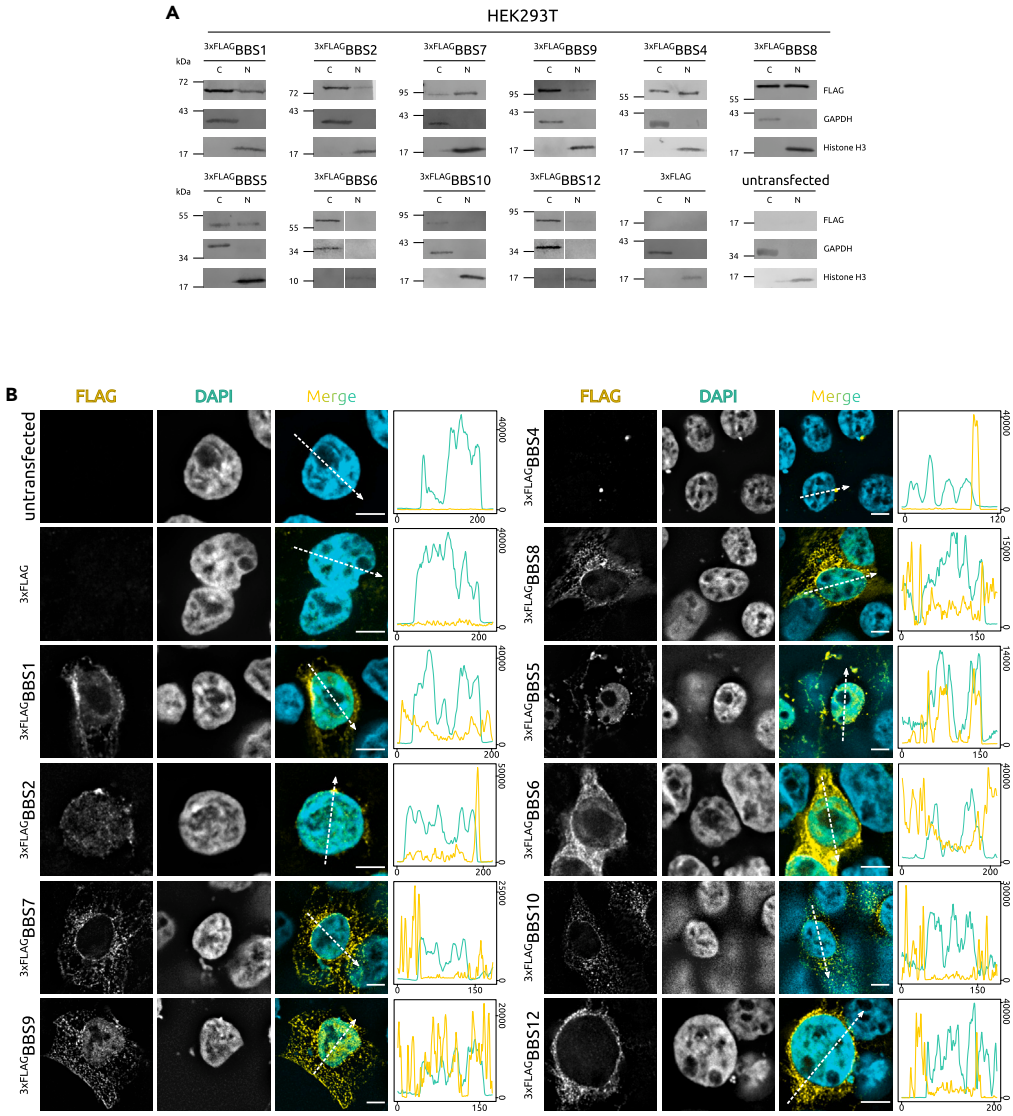


Figure 7. Nuclear localization of different BBS proteins

(A) Western blot after subcellular fractionation of transiently transfected HEK293T cells shows nuclear localization of BBS1, BBS2, BBS7, BBS4, BBS8, and BBS5. All Flag-tagged fusion constructs are expressed and detected at the expected sizes (BBS1: 65 kDa; BBS2: 80 kDa; BBS7: 80 kDa; BBS9: 99 kDa; BBS4 58 kDa; BBS8: 62 kDa; BBS5: 39 kDa; BBS6: 62 kDa; BBS10: 81 kDa; BBS12: 79 kDa). Successful fractionation was determined by presence of marker proteins for the cytosol (GAPDH), and the nucleus (Histone H3). 100 μg protein were loaded per lane except for BBS6 (15 μg) and BBS12 (55 μg).
(B) Immunocytochemical stainings of transiently transfected HEK293T cells expressing FLAG-tagged BBS proteins. Arrows mark the data points used for line plotting in direction of the arrow. Scale bar: 5 μm. Plot X axis: length (pixels); Plot Y axis: pixel intensity (arbitrary units).

BBS6 could not be detected in the nucleus after subcellular fractionation, even though it has been reported to translocate to the nucleus.[43] BBS10 signal is; however, completely absent from the nucleus of immunocytochemically stained HEK293T cells as is BBS12 in contrast to the Western blot. BBS6 can be found in the nucleus in immunostainings, which is congruent with previous studies. The lack of nuclear signal in Western blot analysis might stem from a very transient localization, as it was hypothesized that BBS6 aids in translocation of a SMARCC1-Importin-complex across the nuclear pore.[43] To exclude that the FLAG tag had an effect on the localization, an alternative tag (myc) was used for BBS6 and BBS8. Both proteins showed qualitatively the same pattern in both fractionation and immunostainings as their FLAG-tagged counterparts (Figure S5).

The subnuclear distribution patterns of BBS proteins differ from protein-to-protein. This shows that their import and interaction with other proteins must be differentially regulated and that they may take part in multiple distinct processes. However, we see a discrepancy in the nuclear protein amount when we compare Western blots and immunocytochemical stainings. This emphasizes that by WB, we can detect nuclear signals more sensitively that might go unnoticed with immunocytochemical stainings alone, or that might be ambiguous. On the other hand, immunofluorescence analysis of stained cells reveals a localization of single proteins that is highly specific: BBS7 and BBS8 both localize close to the inner NE (Figure 7B), information which would not be obtainable from Western blots alone. It is also worth noting that in most cases, Western blots and immunostaings agree in their readout (BBS1, BBS2, BBS9, BBS8, and BBS5), although the intensity does not necessarily translate well from one analysis to the other. It is also of note that the ectopic overexpression of BBS proteins may affect the cells differently; we did for example observe that cells overexpressing BBS6 tended to be less viable than others (data not shown). In turn, this means that some of the proteins are harder to detect than others, which is a contributing factor to the different amounts of protein found in the fractions despite being controlled by the same promotor.

DISCUSSION

Cilia and flagella have important roles in many eukaryotic cells. Their proper function is critical for both unicellular life and multicellular organisms, including locomotion, perception, and development. The discovery of ciliary proteins' extraciliary functions in mammalian cells extends the reach of some of their components beyond their ciliary roles. It raises the question whether those functions are conserved in other eukaryotes and therefore ancestral, or if they were co-opted. Our reconstruction of BBSome component distribution across eukaryotes confirms the wide-spread distribution of these proteins and strong phylogenetic association with the presence of cilia. It also suggests that the more phylogenetically restricted chaperonin-like BBS proteins are not exclusive to metazoa, but can also be found in *Malawimonas* and Amebae. We found that BBSome proteins generally appear as a conserved set. This set contains at least one protein from each subset of BBSome proteins, namely the BBS1-like, TPR-domain-containing, and PH-domain-containing protein families (Figure 1B). Exceptions to this rule could be observed in lineages where the most recent species have undergone a gradual process of protein complex reduction and, finally, loss of the cilium. An extreme example is the opisthokont lineage: While the last opisthokont common ancestor was probably still a ciliated single-celled organism with all IFT and BBSome components,[59] the fungal descendants quickly abandoned BBS proteins (as seen in *Batrachochytrium*) and subsequently cilia in general, while the holozoan lineage retained them.

Further sequence analysis of identified orthologues provided insight into the distribution of nuclear localization or export signals (NLS and NES, resp.). After the significant increase of NES in BBS7, BBS1, and BBS8 had the highest probability for high NES scores, while BBS4 and BBS9 had the highest probability for high NLS scores. Preference for stronger nuclear signal sequences in open or closed mitosis was tested revealing no significantly higher likelihood for NLS or NES in BBS proteins in either form of mitosis. This means that nuclear localization of BBS proteins most probably did not co-evolve with the transition to a particular form of mitosis but more likely as a recent adaptation in specific lineages, as appears to be the case for mammalian BBS6.[43]

Through ancestral state reconstructions, we could show that predicted NLS found in BBS proteins are frequent, but most likely lineage-specific. In contrast, predictions of NES—especially for BBS7—can often be found deeper in the evolution of the proteins and tentatively even at the inferred LECA sequence.

Our localization experiments in HEK293T cells revealed that other human BBS proteins than those previously reported[41–43] do indeed also localize to the nucleus. We detect BBS1, BBS2, BBS7, and BBS9; BBS4 and BBS8; BBS5; and BBS10 and BBS12 in the nuclear fraction of transiently transfected and lysed cells, either in soluble form, associated with chromatin or as part of larger complexes. These results could in part be recapitulated in immunocytochemical stainings: we detect nuclear signals for FLAG-tagged BBS1, BBS2 and BBS9; BBS8; BBS5; and BBS6 in our microscopic analysis, but not others. This implicates that the nuclear localization is not simply an artifact produced by overexpression. While mislocalization due to overexpression is a concern, it is also encouraging those previous studies that established lower expression levels in other cell lines do recapitulate a similar subcellular distribution.[35] The distribution of nuclear BBS proteins is not uniform, suggesting that they are associated with different processes in the nucleus. The differences in both abundance (as seen in Western blots), intranuclear localization (as seen in

immunofluorescence stainings), and no preference in either form of mitosis (as seen in permutation analysis) clearly show that BBS proteins' capacity to function in the nucleus differs between specific proteins and is indicative of a specialized role for each BBS protein.

One of the major transitions of life has been the development of a sophisticated, compartmentalized endomembrane system that probably contributed to the success of eukaryotes.[60] While the "original" nuclear membrane was likely freely permeable, the NE has since evolved into a tightly regulated, selective transport barrier.[61–63] During interphase, proteins and nucleic acids must pass the nuclear pore complexes (NPCs) by association with importin- and exportin-class proteins. In some organisms (including green plants and animals), the NE completely breaks down during mitosis, while in others (such as yeasts), the envelope remains intact. The mitotic NE breakdown is however not binary. Many eukaryotes partially break down the NE, allowing microtubules to enter the nuclear space through fenestrated openings, the extent of which is dependent on the species. The breakdown of the NE may serve as an entry point into a compartment that is otherwise inaccessible without an NLS. On the other hand, in an organism with a closed form of mitosis, all nuclear proteins must be actively transported into the nucleus. We reasoned that if nuclear roles for BBS proteins are conserved, there might be higher need for NLS in BBS orthologues in species with a closed form of mitosis and, conversely, a potential increase of NES in BBS orthologues in species with an open mitosis. Our permutation analysis showed no statistical bias toward NLS or NES signals in association with mitotic mode. Hence the appearance of nuclear signal sequences is not coincidental with the evolution of a specific mitotic pattern. However, while we predicted mostly NES signals for human BBS proteins, we still find nuclear localization via both subcellular fractionation and immunofluorescent staining. Our analysis is by necessity based on prediction of classical, basic NLS signals only and non-classical (e.g., PY-NLS) or protein-specific NLSs will not be detected. Moreover, due to the difficulties in prediction even of classical NLSs, computational NLS predictors detect a relatively low percentage of true NLSs with high confidence (37% for NLStradamus against a dataset of non-yeast sequences). There is also the possibility of an "indirect" NLS effect, i.e., association of BBS proteins with interaction partners that are localizing to the nucleus, and BBS proteins hijacking them and being imported into the nucleus passively. Recently it could be shown that CCDC28B, a cilia protein associating with BBS proteins bears a functional NLS in a zebra fish homolog. It could therefore potentially serve as a nucleocytoplasmic shuttle for BBS proteins.[64] Bioinformatic analysis alone is therefore, currently not sufficient to accurately predict nuclear localization for any individual BBS protein—although it does provide a useful tool to assess any underlying trends.

In addition to the nucleus, proteins forming the nuclear pores, nucleoporins, can be found at the base of cilia regulating ciliary import in a similar fashion as they do at nuclear pores.[65] Both localizations share common translocation mechanisms of cargo into the respective compartment, utilizing RanGDP/GTP (Ran-guinosine di-/triphosphate) gradients, importins, and specialized localization sequences. For several ciliary cargoes, including crucial anterograde IFT protein KIF17,[65] the "nuclear" translocation system could be confirmed, but others seem to bypass the need for some parts of the system. Interestingly, primary cilia in particular do not appear to have any nucleoporins at the base.[66] While there are differences in nuclear and ciliary import and export, it is still possible that both systems once shared common ancestral machinery. And indeed, components of both share a potential ancestor that once emerged to facilitate vesicle coats.[26,55–57,63–67] This leaves two possible options for the appearance of ciliary proteins at and in the nucleus: Either I) the cilium developed before the NE, making the ciliary functions ancestral and the nuclear functions co-opted; or II) the NE emerged first, and the close association of the cilium-nucleating centriole led to nuclear proteins "hitch-hiking" to the cilium and adopting ciliary roles. Future analysis comparing the divergence of ciliary vesicle and cargo coat complexes such as the BBSome from nucleoporins, and the hypothetical ancestral protocoatomer needs to be conducted to untangle their shared evolutionary history.

Regulation of differential gene expression during development of tissues is a novel role for BBS proteins that has only been described recently.[42,43,58] But BBS proteins are also involved in short-term responses to more immediate environmental cues. More recent studies have shown that BBS proteins are recruited for translocation of transcription factors after application of cytotoxic stress: BBS4, which is normally localized at centrosomes in cycling cells or the basal body in quiescent cells, translocates from the cytosol to nucleus after induction of endoplasmic reticulum (ER) stress.[41] Here, BBS4 associates with transcription factors that initiate the unfolded protein response (UPR). This is reminiscent of a similar mechanism seen in BBS6 that is disturbed in MKKS, where a mutation disrupts the shuttling of BBS6 in and out of the nucleus, enriching SMARCC1 in the cytoplasm. We found that induction of cyto- and genotoxic stress induces a

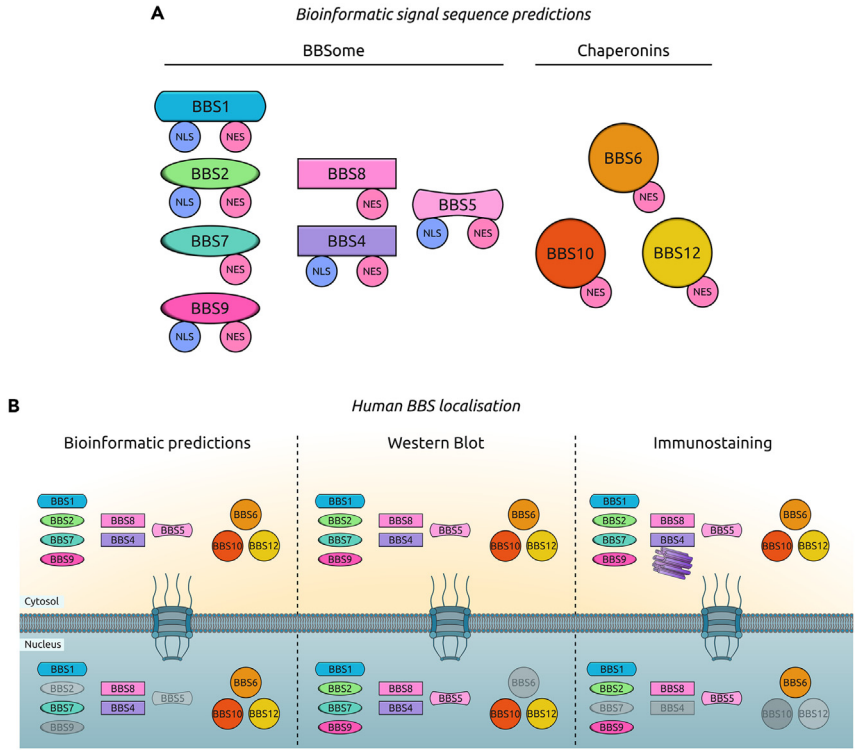


Figure 8. Schematic summary of bioinformatic, molecular and immunological experiments

(A) BBS proteins with predicted NLS and/or NES. A label is added if the respective signal sequence is found in at least one of the predicted eukaryotic orthologues.

(B) Comparison of predicted localization (left), localization after Western blotting (middle), and immunofluorescent staining (right) of FLAG-tagged BBS proteins in *Homo sapiens*.

translocation of BBS4, BBS6, and BBS8 from cytoplasm to nucleus (Patnaik et al., in prep.). BBS proteins therefore, have a pivotal role in both developmental, long-term regulation of genes, and short-term responses to environmental cues by associating with transcription factors and effector proteins and ensuring cytonucleoplasmic translocation.

Our data consolidates that BBS proteins can have roles outside of ciliary trafficking. Our orthologue screen revealed that all BBS proteins in this study have predicted NES signals in at least one species, but classical NLSs are more restricted only being detected in one or more orthologue of BBS1, BBS2, BBS4, and BBS5 in at least one species (Figure 8A). However, as not all BBSome proteins localize to the nucleus (Figure 8B) and since those that do have distinct patterns, it is likely that BBS protein involvement in nuclear processes is independent of their function as a complex. This nuclear localization does not appear to be the result of the BBSome (or individual proteins) being coincidentally incorporated into the nucleus due to evolution toward open mitosis, but is equally likely in all eukaryotic species possessing BBS proteins. This leaves two options for the evolution of BBS proteins' nuclear function: Given that BBS proteins are found across all eukaryotic taxa even on occasion in species without cilia, the BBSome proteins have highly likely already been established in LECA. This might place the root of nuclear functions of BBS proteins potentially quite early during the radiation of eukaryotes. This theory is favored by the widespread NES signature for single BBS proteins we find across the entire eukaryotic tree. Another explanation is an independent convergence scenario where conserved BBS proteins filled an ecological niche for gene regulation on numerous occasions; this would be well explained by the sparse occurences of NLS in single clades of BBS proteins, as well as the absence in ancestral nodes toward the root of the tree. Position-specific tracing of NLS/NES in BBS proteins will clarify whether some signal sequences persisted throughout eukaryotic evolution. Either theory can further be elucidated by investigation of BBS proteins' nuclear interactome and gene regulatory functions in different eukaryotic species. This information is crucial to understand the evolution of the cilium, nucleus, and the origins of regulatory networks in early eukaryotes.

Limitations of the study

Bioinformatical prediction of subcellular localization remains challenging with current predictive algorithms, yet offers a great opportunity to examine protein interactions beyond the scope of their canonical field of action. This is crucial to understand a protein's evolutionary background and—in our case—its role in the etiology of a disease. Although signal sequence predictions by NESmapper and NLStradamus are not to be taken at face value as definitive localization criteria, they allow us to statistically analyze their strength in correlation with other variables, such as clade affiliation or mode of mitosis. We took care not to overestimate the conclusion that can be drawn from our bioinformatical studies and are aware that the predictions do not encompass the full spectrum of possible nuclear import and export mechanisms for proteins. Factors like protein-interaction-mediated (indirect), nucleo-cytoplasmic translocation of proteins, and non-canonical signal sequences (e.g., PY-NLS) are not negligible but virtually impossible to rule out for any comparable approach and need validation beyond bioinformatic detection. We argue that this was not in the scope of this work and that our approach is cautious enough to justify the limitation. To validate our bioinformatic findings, we used a human cell model, HEK293T cells that are explicitly unciliated throughout the course of the experiments. While this reduces the possibility to validate functionality of our tagged proteins through their association with primary cilia, we could ensure that proteins would replicate their non-ciliary localization in unciliated cells as the sequestering effect of a cilium would be ablated. We chose to overexpress tagged BBS proteins in these cells as to enhance detection rather than measuring endogenous levels of protein as reliable antibodies for endogenous BBS proteins are not available. This renders microscopy after indirect immunofluorescent labeling more difficult to interpret as mislocalization after fixation is a known limitation,[68] but the disctint patterns observed after staining differentially transfected cells attest that nuclear localization of BBS proteins is not artifactual. Additional in-depth experiments, including live-cell-imaging with fluorescently labeled proteins and mass spectrometry, would be required to ratify our data further which was beyond the scope of our exploratory study.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Proteomic dataset procurement
  - Homologue prediction/phylogenetic inference
  - Ancestral sequence reconstruction
  - Signal sequence prediction
  - Transfection
  - Subcellular fractionation
  - Polyacrylamide gel electrophoresis
  - Western blot
  - Immunocytochemistry
  - Antibodies
  - Software and packages
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.106410.

ACKNOWLEDGMENTS

We would like to thank Tommy Sroka (Saarland University School of Medicine, Homburg, Germany) for granting us access to an FIJI script and allowing to us to modify it for our specific analysis. We would like to thank Stef Letteboer (Radboud University Medical Center, Nijmegen, The Netherlands), who kindly gifted us the empty vector backbone used in our cloning experiments. We thank Andrew J. Roger

(Dalhousie University, Halifax, Canada) for providing the proteomes of *Malawimonas jakobiformis* and *Monocercomonoides* sp.. We also thank Susanne Gerber (University Medical Center Mainz, Mainz, Germany), and Martin Kaltenpoth (Max Planck Institute for Chemical Ecology, Jena, Germany) for their suggestions and fruitful discussions. This project was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – GRK2526/1 – Projectnr. 407023052.

AUTHOR CONTRIBUTIONS

Conceptualization, B.W. and H.L.M.-S.; Software, A.E. and B.W.; Validation, A.E., V.M., B.W., and H.L.M.-S.; Formal analysis, A.E. and B.W.; Investigation, A.E., V.M., and B.W.; Resources, B.W. and H.L.M.-S.; Data curation, A.E. and B.W.; Writing – Original Draft, A.E.; Writing – Review & Editing, A.E., B.W., and H.L.M.-S.; Visualization, A.E. and B.W.; Supervision, B.W. and H.L.M.-S.; Project administration, H.L.M.-S.; Funding Acquisition, H.L.M.-S.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 12, 2022
Revised: July 20, 2022
Accepted: March 10, 2023
Published: March 16, 2023

REFERENCES

1. Mitchell, D.R. (2007). The evolution of eukaryotic cilia and flagella as motile and sensory organelles. In Eukaryotic Membranes and Cytoskeleton: Origins and Evolution Advances in Experimental Medicine and Biology (Springer), pp. 130–140. https://doi.org/10.1007/978-0-387-74021-8_11.

2. Koumandou, V.L., Wickstead, B., Ginger, M.L., van der Giezen, M., Dacks, J.B., and Field, M.C. (2013). Molecular paleontology and complexity in the last eukaryotic common ancestor. Crit. Rev. Biochem. Mol. Biol. *48*, 373–396. https://doi.org/10.3109/10409238.2013.821444.

3. Bloodgood, R.A. (1981). Flagella-dependent gliding motility inChlamydomonas. Protoplasma *106*, 183–192. https://doi.org/10.1007/BF01275550.

4. Saito, A., Suetomo, Y., Arikawa, M., Omura, G., Khan, M.K., Kakuta, S., Suzaki, E., Kataoka, K., and Suzaki, T. (2003). Gliding movement in Peranema trichophorum is powered by flagellar surface motility. Cell Motil Cytoskeleton *55*, 244–253. https://doi.org/10.1002/cm.10127.

5. Cavalier-Smith, T., Lewis, R., Chao, E.E., Oates, B., and Bass, D. (2009). Helkesimastix marina n. sp. (cercozoa: sainouroidea superfam. N.) a gliding zooflagellate of novel ultrastructure and unusual ciliary behaviour. Protist *160*, 452–479. https://doi.org/10.1016/j.protis.2009.03.003.

6. Yoshimura, K. (1996). A novel type of mechanoreception by the flagella of Chlamydomonas. J. Exp. Biol. *199*, 295–302. https://doi.org/10.1242/jeb.199.2.295.

7. Nauli, S.M., Alenghat, F.J., Luo, Y., Williams, E., Vassilev, P., Li, X., Elia, A.E.H., Lu, W., Brown, E.M., Quinn, S.J., et al. (2003). Polycystins 1 and 2 mediate mechanosensation in the primary cilium of kidney cells. Nat. Genet. *33*, 129–137. https://doi.org/10.1038/ng1076.

8. Pazour, G.J., Agrin, N., Leszyk, J., and Witman, G.B. (2005). Proteomic analysis of a eukaryotic cilium. J. Cell Biol. *170*, 103–113. https://doi.org/10.1083/jcb.200504008.

9. Wakabayashi, K.I., Ide, T., and Kamiya, R. (2009). Calcium-dependent flagellar motility activation in Chlamydomonas reinhardtii in response to mechanical agitation. Cell Motil Cytoskeleton *66*, 736–742. https://doi.org/10.1002/cm.20402.

10. Snell, W.J., and Goodenough, U.W. (2009). Chapter 12 - flagellar adhesion, flagellar-generated signaling, and gamete fusion during mating. In The Chlamydomonas Sourcebook, Second Edition, E.H. Harris, D.B. Stern, and G.B. Witman, eds. (Academic Press), pp. 369–394. https://doi.org/10.1016/B978-0-12-370873-1.00049-6.

11. Bloodgood, R.A. (2010). Sensory reception is an attribute of both primary cilia and motile cilia. J. Cell Sci. *123*, 505–509. https://doi.org/10.1242/jcs.066308.

12. Jeffery, P.K., and Reid, L. (1975). New observations of rat airway epithelium: a quantitative and electron microscopic study. J. Anat. *120*, 295–320.

13. Sanderson, M.J., and Sleigh, M.A. (1981). Ciliary activity of cultured rabbit tracheal epithelium: beat pattern and metachrony. J. Cell Sci. *47*, 331–347. https://doi.org/10.1242/jcs.47.1.331.

14. Nonaka, S., Tanaka, Y., Okada, Y., Takeda, S., Harada, A., Kanai, Y., Kido, M., and Hirokawa, N. (1998). Randomization of left–right asymmetry due to loss of nodal cilia generating leftward flow of extraembryonic fluid in mice lacking KIF3B motor protein. Cell *95*, 829–837. https://doi.org/10.1016/S0092-8674(00)81705-5.

15. Raidt, J., Werner, C., Menchen, T., Dougherty, G.W., Olbrich, H., Loges, N.T., Schmitz, R., Pennekamp, P., and Omran, H. (2015). Ciliary function and motor protein composition of human fallopian tubes. Hum. Reprod. *30*, 2871–2880. https://doi.org/10.1093/humrep/dev227.

16. Essner, J.J., Vogan, K.J., Wagner, M.K., Tabin, C.J., Yost, H.J., and Brueckner, M. (2002). Conserved function for embryonic nodal cilia. Nature *418*, 37–38. https://doi.org/10.1038/418037a.

17. Smith, D.J., Smith, A.A., and Blake, J.R. (2011). Mathematical embryology: the fluid mechanics of nodal cilia. J. Eng. Math. *70*, 255–279. https://doi.org/10.1007/s10665-010-9383-y.

18. May-Simera, H.L., Wan, Q., Jha, B.S., Hartford, J., Khristov, V., Dejene, R., Chang, J., Patnaik, S., Lu, Q., Banerjee, P., et al. (2018). Primary cilium-mediated retinal pigment epithelium maturation is disrupted in ciliopathy patient cells. Cell Rep. *22*, 189–205. https://doi.org/10.1016/j.celrep.2017.12.038.

19. Patnaik, S.R., Kretschmer, V., Brücker, L., Schneider, S., Volz, A.-K., Oancea-Castillo, L.D.R., and May-Simera, H.L. (2019). Bardet-Biedl Syndrome proteins regulate cilia disassembly during tissue maturation. Cell. Mol. Life Sci. *76*, 757–775. https://doi.org/10.1007/s00018-018-2966-x.

20. Anvarian, Z., Mykytyn, K., Mukhopadhyay, S., Pedersen, L.B., and Christensen, S.T. (2019). Cellular signalling by primary cilia in development, organ function and disease.

Nat. Rev. Nephrol. *15*, 199–219. https://doi.org/10.1038/s41581-019-0116-9.

21. Afzelius, B.A. (2004). Cilia-related diseases. J. Pathol. *204*, 470–477. https://doi.org/10.1002/path.1652.

22. Badano, J.L., Mitsuma, N., Beales, P.L., and Katsanis, N. (2006). The ciliopathies: an emerging class of human genetic Disorders. Annu. Rev. Genomics Hum. Genet. *7*, 125–148. https://doi.org/10.1146/annurev.genom.7.080505.115610.

23. Fliegauf, M., Benzing, T., and Omran, H. (2007). When cilia go bad: cilia defects and ciliopathies. Nat. Rev. Mol. Cell Biol. *8*, 880–893. https://doi.org/10.1038/nrm2278.

24. Briggs, L.J., Davidge, J.A., Wickstead, B., Ginger, M.L., and Gull, K. (2004). More than one way to build a flagellum: comparative genomics of parasitic protozoa. Curr. Biol. *14*, R611–R612. https://doi.org/10.1016/j.cub.2004.07.041.

25. Jékely, G., and Arendt, D. (2006). Evolution of intraflagellar transport from coated vesicles and autogenous origin of the eukaryotic cilium. Bioessays *28*, 191–198. https://doi.org/10.1002/bies.20369.

26. van Dam, T.J.P., Townsend, M.J., Turk, M., Schlessinger, A., Sali, A., Field, M.C., and Huynen, M.A. (2013). Evolution of modular intraflagellar transport from a coatomer-like progenitor. Proc. Natl. Acad. Sci. USA *110*, 6943–6948. https://doi.org/10.1073/pnas.1221011110.

27. Li, J.B., Gerdes, J.M., Haycraft, C.J., Fan, Y., Teslovich, T.M., May-Simera, H., Li, H., Blacque, O.E., Li, L., Leitch, C.C., et al. (2004). Comparative genomics identifies a flagellar and basal body proteome that includes the BBS5 human disease gene. Cell *117*, 541–552. https://doi.org/10.1016/S0092-8674(04)00450-7.

28. Wickstead, B., and Gull, K. (2007). Dyneins across eukaryotes: a comparative genomic analysis. Traffic *8*, 1708–1721. https://doi.org/10.1111/j.1600-0854.2007.00646.x.

29. Carvalho-Santos, Z., Machado, P., Branco, P., Tavares-Cadete, F., Rodrigues-Martins, A., Pereira-Leal, J.B., and Bettencourt-Dias, M. (2010). Stepwise evolution of the centriole-assembly pathway. J. Cell Sci. *123*, 1414–1426. https://doi.org/10.1242/jcs.064931.

30. Avidor-Reiss, T., Maer, A.M., Koundakjian, E., Polyanovsky, A., Keil, T., Subramaniam, S., and Zuker, C.S. (2004). Decoding cilia function: Defining specialized genes required for compartmentalized cilia biogenesis. Cell *117*, 527–539. https://doi.org/10.1016/S0092-8674(04)00412-X.

31. Taschner, M., and Lorentzen, E. (2016). The intraflagellar transport machinery. Cold Spring Harb. Perspect. Biol. *8*, a028092. https://doi.org/10.1101/cshperspect.a028092.

32. Prevo, B., Scholey, J.M., and Peterman, E.J.G. (2017). Intraflagellar transport: mechanisms of motor action, cooperation, and cargo delivery. FEBS J. *284*, 2905–2931. https://doi.org/10.1111/febs.14068.

33. Singh, S.K., Gui, M., Koh, F., Yip, M.C., and Brown, A. (2020). Structure and activation mechanism of the BBSome membrane protein trafficking complex. Elife *9*, e53322. https://doi.org/10.7554/eLife.53322.

34. Klink, B.U., Gatsogiannis, C., Hofnagel, O., Wittinghofer, A., and Raunser, S. (2020). Structure of the human BBSome core complex. Elife *9*, e53910. https://doi.org/10.7554/eLife.53910.

35. Kim, J.C., Badano, J.L., Sibold, S., Esmail, M.A., Hill, J., Hoskins, B.E., Leitch, C.C., Venner, K., Ansley, S.J., Ross, A.J., et al. (2004). The Bardet-Biedl protein BBS4 targets cargo to the pericentriolar region and is required for microtubule anchoring and cell cycle progression. Nat. Genet. *36*, 462–470. https://doi.org/10.1038/ng1352.

36. Kim, J.C., Ou, Y.Y., Badano, J.L., Esmail, M.A., Leitch, C.C., Fiedrich, E., Beales, P.L., Archibald, J.M., Katsanis, N., Rattner, J.B., and Leroux, M.R. (2005). MKKS/BBS6, a divergent chaperonin-like protein linked to the obesity disorder Bardet-Biedl syndrome, is a novel centrosomal component required for cytokinesis. J. Cell Sci. *118*, 1007–1020. https://doi.org/10.1242/jcs.01676.

37. Robert, A., Margall-Ducos, G., Guidotti, J.-E., Brégerie, O., Celati, C., Bréchot, C., and Desdouets, C. (2007). The intraflagellar transport component IFT88/polaris is a centrosomal protein regulating G1-S transition in non-ciliated cells. J. Cell Sci. *120*, 628–637. https://doi.org/10.1242/jcs.03366.

38. Delaval, B., Bright, A., Lawson, N.D., and Doxsey, S. (2011). The cilia protein IFT88 is required for spindle orientation in mitosis. Nat. Cell Biol. *13*, 461–468. https://doi.org/10.1038/ncb2202.

39. den Dulk, B., van Eijk, P., de Ruijter, M., Brandsma, J.A., and Brouwer, J. (2008). The NER protein Rad33 shows functional homology to human Centrin2 and is involved in modification of Rad4. DNA Repair *7*, 858–868. https://doi.org/10.1016/j.dnarep.2008.02.004.

40. McClure-Begley, T.D., and Klymkowsky, M.W. (2017). Nuclear roles for cilia-associated proteins. Cilia *6*, 8. https://doi.org/10.1186/s13630-017-0052-x.

41. Horwitz, A., and Birk, R. (2021). BBS4 is essential for nuclear transport of transcription factors mediating neuronal ER stress response. Mol. Neurobiol. *58*, 78–91. https://doi.org/10.1007/s12035-020-02104-z.

42. Gascue, C., Tan, P.L., Cardenas-Rodriguez, M., Libisch, G., Fernandez-Calero, T., Liu, Y.P., Astrada, S., Robello, C., Naya, H., Katsanis, N., and Badano, J.L. (2012). Direct role of Bardet–Biedl syndrome proteins in transcriptional regulation. J. Cell Sci. *125*, 362–375. https://doi.org/10.1242/jcs.089375.

43. Scott, C.A., Marsden, A.N., Rebagliati, M.R., Zhang, Q., Chamling, X., Searby, C.C., Baye, L.M., Sheffield, V.C., and Slusarski, D.C. (2017). Nuclear/cytoplasmic transport defects in BBS6 underlie congenital heart disease through perturbation of a chromatin remodeling protein. PLoS Genet. *13*, e1006936. https://doi.org/10.1371/journal.pgen.1006936.

44. De Magistris, P., and Antonin, W. (2018). The dynamic nature of the nuclear envelope. Curr. Biol. *28*, R487–R497. https://doi.org/10.1016/j.cub.2018.01.073.

45. Shi, B., Xue, M., Wang, Y., Wang, Y., Li, D., Zhao, X., and Li, X. (2018). An improved method for increasing the efficiency of gene transfection and transduction. Int. J. Physiol. Pathophysiol. Pharmacol. *10*, 95–104.

46. Hodges, M.E., Scheumann, N., Wickstead, B., Langdale, J.A., and Gull, K. (2010). Reconstructing the evolutionary history of the centriole from protein components. J. Cell Sci. *123*, 1407–1413. https://doi.org/10.1242/jcs.064873.

47. Nachury, M.V., Loktev, A.V., Zhang, Q., Westlake, C.J., Peränen, J., Merdes, A., Slusarski, D.C., Scheller, R.H., Bazan, J.F., Sheffield, V.C., and Jackson, P.K. (2007). A core complex of BBS proteins cooperates with the GTPase Rab8 to promote ciliary membrane biogenesis. Cell *129*, 1201–1213. https://doi.org/10.1016/j.cell.2007.03.053.

48. Lattao, R., Kovács, L., and Glover, D.M. (2017). The Centrioles,Centrosomes, basal bodies, and cilia of Drosophila melanogaster. Genetics *206*, 33–53. https://doi.org/10.1534/genetics.116.198168.

49. Jafari, S., Henriksson, J., Yan, H., and Alenius, M. (2021). Stress and odorant receptor feedback during a critical period after hatching regulates olfactory sensory neuron differentiation in Drosophila. PLoS Biol. *19*, e3001101. https://doi.org/10.1371/journal.pbio.3001101.

50. Killick-Kendrick, R., and Peters, W. (1978). Rodent Malaria (Academic Press).

51. Venard, C.M., Vasudevan, K.K., and Stearns, T. (2020). Cilium axoneme internalization and degradation in chytrid fungi. Cytoskeleton *77*, 365–378. https://doi.org/10.1002/cm.21637.

52. Mukherjee, K., Conway de Macario, E., Macario, A.J.L., and Brocchieri, L. (2010). Chaperonin genes on the rise: new divergent classes and intense duplication in human and other vertebrate genomes. BMC Evol. Biol. *10*, 64. https://doi.org/10.1186/1471-2148-10-64.

53. Mukherjee, K., and Brocchieri, L. (2013). Ancient origin of chaperonin gene paralogs involved in ciliopathies. J. Phylogenetics Evol. Biol. *1*, 107.

54. Horton, P., Park, K.-J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., and Nakai, K. (2007). WoLF PSORT: protein localization predictor. Nucleic Acids Res. *35*, W585–W587. https://doi.org/10.1093/nar/gkm259.

55. Feldherr, C.M., and Akin, D. (1997). The location of the transport gate in the nuclear pore complex. J. Cell Sci. *110*, 3065–3070. https://doi.org/10.1242/jcs.110.24.3065.

56. Keminer, O., and Peters, R. (1999). Permeability of single nuclear pores. Biophys. J. *77*, 217–228.

57. Keune, W.j., Bultsma, Y., Sommer, L., Jones, D., and Divecha, N. (2011). Phosphoinositide signalling in the nucleus. Adv. Enzyme Regul. *51*, 91–99. https://doi.org/10.1016/j.advenzreg.2010.09.009.

58. Marchese, E., Caterino, M., Viggiano, D., Cevenini, A., Tolone, S., Docimo, L., Di Iorio, V., Del Vecchio Blanco, F., Fedele, R., Simonelli, F., et al. (2022). Metabolomic fingerprinting of renal disease progression in Bardet-Biedl syndrome reveals mitochondrial dysfunction in kidney tubular cells. iScience *25*, 105230. https://doi.org/10.1016/j.isci.2022.105230.

59. Torruella, G., de Mendoza, A., Grau-Bové, X., Antó, M., Chaplin, M.A., del Campo, J., Eme, L., Pérez-Cordón, G., Whipps, C.M., Nichols, K.M., et al. (2015). Phylogenomics reveals convergent evolution of lifestyles in close relatives of animals and fungi. Curr. Biol. *25*, 2404–2410. https://doi.org/10.1016/j.cub.2015.07.053.

60. Rout, M.P., and Field, M.C. (2017). The evolution of organellar coat complexes and organization of the eukaryotic cell. Annu. Rev. Biochem. *86*, 637–657. https://doi.org/10.1146/annurev-biochem-061516-044643.

61. Hoelz, A., Debler, E.W., and Blobel, G. (2011). The structure of the nuclear pore complex. Annu. Rev. Biochem. *80*, 613–643. https://doi.org/10.1146/annurev-biochem-060109-151030.

62. Wilson, K.L., and Dawson, S.C. (2011). Functional evolution of nuclear structure. J. Cell Biol. *195*, 171–181. https://doi.org/10.1083/jcb.201103171.

63. Field, M.C., Koreny, L., and Rout, M.P. (2014). Enriching the pore: splendid complexity from humble origins. Traffic *15*, 141–156. https://doi.org/10.1111/tra.12141.

64. Novas, R., Cardenas-Rodriguez, M., Lepanto, P., Fabregat, M., Rodao, M., Fariello, M.I., Ramos, M., Davison, C., Casanova, G., Alfaya, L., et al. (2018). Kinesin 1 regulates cilia length through an interaction with the Bardet-Biedl syndrome related protein CCDC28B. Sci. Rep. *8*, 3019. https://doi.org/10.1038/s41598-018-21329-6.

65. Dishinger, J.F., Kee, H.L., Jenkins, P.M., Fan, S., Hurd, T.W., Hammond, J.W., Truong, Y.N.-T., Margolis, B., Martens, J.R., and Verhey, K.J. (2010). Ciliary entry of the kinesin-2 motor KIF17 is regulated by importin-β2 and RanGTP. Nat. Cell Biol. *12*, 703–710. https://doi.org/10.1038/ncb2073.

66. Breslow, D.K., Koslover, E.F., Seydel, F., Spakowitz, A.J., and Nachury, M.V. (2013). An in vitro assay for entry into cilia reveals unique properties of the soluble diffusion barrier. J. Cell Biol. *203*, 129–147. https://doi.org/10.1083/jcb.201212024.

67. Obado, S.O., Brillantes, M., Uryu, K., Zhang, W., Ketaren, N.E., Chait, B.T., Field, M.C., and Rout, M.P. (2016). Interactome mapping reveals the evolutionary history of the nuclear pore complex. PLoS Biol. *14*, e1002365. https://doi.org/10.1371/journal.pbio.1002365.

68. Melan, M.A., and Sluder, G. (1992). Redistribution and differential extraction of soluble proteins in permeabilized cultured cells. Implications for immunofluorescence microscopy. J. Cell Sci. *101*, 731–743.

69. Baghirova, S., Hughes, B.G., Hendzel, M.J., and Schulz, R. (2015). Sequential fractionation and isolation of subcellular proteins from tissue or cultured cells. MethodsX *2*, 440–445. https://doi.org/10.1016/j.mex.2015.11.001.

70. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402. https://doi.org/10.1093/nar/25.17.3389.

71. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780. https://doi.org/10.1093/molbev/mst010.

72. Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics *25*, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348.

73. Eddy, S.R. (2011). Accelerated profile HMM searches. PLoS Comput. Biol. *7*, e1002195. https://doi.org/10.1371/journal.pcbi.1002195.

74. Price, M.N., Dehal, P.S., and Arkin, A.P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol. Biol. Evol. *26*, 1641–1650. https://doi.org/10.1093/molbev/msp077.

75. Le, S.Q., and Gascuel, O. (2008). An improved general amino acid replacement matrix. Mol. Biol. Evol. *25*, 1307–1320. https://doi.org/10.1093/molbev/msn067.

76. Pupko, T., Pe'er, I., Shamir, R., and Graur, D. (2000). A fast algorithm for joint reconstruction of ancestral amino acid sequences. Mol. Biol. Evol. *17*, 890–896. https://doi.org/10.1093/oxfordjournals.molbev.a026369.

77. Moshe, A., and Pupko, T. (2019). Ancestral sequence reconstruction: accounting for structural information by averaging over replacement matrices. Bioinformatics *35*, 2562–2568. https://doi.org/10.1093/bioinformatics/bty1031.

78. Nguyen Ba, A.N., Pogoutse, A., Provart, N., and Moses, A.M. (2009). NLStradamus: a simple Hidden Markov Model for nuclear localization signal prediction. BMC Bioinf. *10*, 202. https://doi.org/10.1186/1471-2105-10-202.

79. Kosugi, S., Yanagawa, H., Terauchi, R., and Tabata, S. (2014). NESmapper: accurate prediction of leucine-rich nuclear export signals using activity-based profiles. PLoS Comput. Biol. *10*, e1003841. https://doi.org/10.1371/journal.pcbi.1003841.

80. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682. https://doi.org/10.1038/nmeth.2019.

81. Volker, B. Lif2Tif Macro (Montpellier Ressources Imagerie)

82. R Core Team R: A Language and Environment for Statistical Computing. (R Foundation for Statistical Computing)

83. Adler, D., and Kelly, S.T. Vioplot: Violin Plot

84. Wickham, H., François, R., Henry, L., and Müller, K. (2021). Dplyr: A Grammar of Data Manipulation.

85. Inkscape Project (2020 (Inkscape).

86. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. B *57*, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Monoclonal (clone M2) mouse anti-FLAG	Sigma-Aldrich	Cat# F1804; RRID: AB_262044
Monoclonal (clone D16H11) rabbit anti-GAPDH	Cell Signaling Technology	Cat# 5174; RRID: AB_10622025
Polyclonal rabbit anti-Histone H3	Proteintech	Cat# 17168-1-AP; RRID: AB_2716755
Monoclonal (clone 9B11) mouse anti-myc	Cell Signaling Technology	Cat# 2276; RRID: AB_331783
Polyclonal donkey Alexa Fluor™ Plus 488 anti-mouse	Invitrogen	Cat# A-21202; RRID: AB_141607
Polyclonal donkey IRDye® 800 CW anti-Mouse	LI-COR Biosciences	Cat# 926-32212; RRID: AB_621847
Polyclonal donkey IRDye® 680 RD anti-Mouse	LI-COR Biosciences	Cat# 926-68072; RRID: AB_10953628
Polyclonal donkey IRDye® 800 CW anti-Rabbit	LI-COR Biosciences	Cat# 926-32213; RRID: AB_621848
Polyclonal donkey IRDye® 680 CW anti-Rabbit	LI-COR Biosciences	Cat# 926-68073; RRID: AB_10954442
Chemicals, peptides, and recombinant proteins		
Digitonin	Sigma-Aldrich	Cat# D141; CAS: 11024-24-1; EC: 234-255-6
Hexylene glycol	Sigma-Aldrich	Cat# 112100; CAS: 107-41-5; EC: 203-489-0
Protease and Phosphatase Inhibitor Cocktail (100X)	Thermo Fisher Scientific	Cat# 78440
Benzonase	Merck	Cat# E1014; CAS: 9025-65-4
Deposited data		
Plasmids coding for FLAG-tagged BBS proteins	This paper	figshare.com:
Signal sequence prediction files (Data S2)	This paper	figshare.com: https://doi.org/10.6084/m9.figshare.22140347.v1
Script for statistical analysis (Data S2)	This paper	figshare.com: https://doi.org/10.6084/m9.figshare.22140347.v1
Experimental models: Cell lines		
HEK293T	N/A	RRID: CVCL_0063
Recombinant DNA		
Plasmids coding for FLAG-tagged BBS proteins	This paper	figshare.com: https://doi.org/10.6084/m9.figshare.22140458.v1
Software and algorithms		
R (version 3.6.3)	https://www.R-project.org/	RRID: SCR_001905
FIJI	Schindelin et al.[69]	https://doi.org/10.1038/nmeth.2019; RRID: SCR_002285
FIJI macro for image preparation and micrograph analysis (Data S1)	This paper	figshare.com: https://doi.org/10.6084/m9.figshare.22140278.v1

RESOURCE AVAILABILITY

Lead contact

Requests for further information or regarding resources, reagents, protocols, and data should be directed to and will be handled by the lead contact, Helen Louise May-Simera (may-simera@uni-mainz.de).

Materials availability

Maps and sequence files for plasmids generated in this study are available under the DOI listed in the key resources table. To access plasmids, please refer to the lead contact for further information.

Data and code availability

- Signal sequence prediction data (.nes and.nls) have been deposited at Figshare and are publicly available as of the date of publication. DOIs are listed in the key resources table. Original Western Blot and microscopy files reported in this paper will be shared by the lead contact upon request.

- All original code has been deposited at Figshare and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human embryonic kidney (HEK293T) cells derived from a female were cultivated in Dulbecco's Modified Eagle's Medium (DMEM; Thermo Fisher Scientific GmbH, Dreieich, Germany) supplemented with 10% Fetal Bovine Serum (FBS; Thermo Fisher Scientific GmbH, Dreieich, Germany) at 37°C and 5% CO_2 in high humidity.

METHOD DETAILS

Proteomic dataset procurement

Proteomic datasets were compiled from predicted proteomes of 40 extant eukaryotic species selected to give good distribution across eukaryotes for which data are available. Organisms were selected to give coverage of organisms: a) from six major eukaryotic clades (Amoebozoa, Fungi, Holozoa, Excavata, SAR and Archaeplastida); b) with or without cilia; and c) that are known to undergo open, closed, semi-open, or semi-closed mitosis. Redundancy in predicted proteomes was reduced by removal of sequences with >95% identity for a given species. This reduces skewing during Hidden Markov model (HMM) building toward models reflecting minor differences in organisms with many predicted variants. The list of proteome sources and versions can be found in Table S2.

Homologue prediction/phylogenetic inference

Similarity searches for BBSome proteins BBS1, BBS2, BBS7, BBS9; BBS4, BBS8; BBS5; and BBS9 and for chaperonin-like BBS proteins BBS6, BBS10 and BBS12 were conducted using a combined BLAST and HMM approach. Initially, a BLASTp[70] search (version 2.2.26) of the predicted proteomes was conducted using human BBS proteins as queries. E-value thresholds for each protein were chosen conservatively by manual inspection of hits. Retrieved sequences were then aligned using MAFFT (Multiple Alignment using Fast Fourier Tranform) v7.271 L-INS-i strategy[71] (–maxiterate 1000 –localpair). The resulting alignment was trimmed using TrimAl v1.2[72] with heuristic selection of the automatic method (-automated1). This alignment was then used to build a Hidden Markov Model for the BBS protein and this was used to search the datasets again using HMMER3.[73] Any additional hits from HMM searches were incorporated into a new model, and this process was iterated until no further hits above threshold were achieved.

To check for orthology within hits from HMM searches, hits were aligned against outgroups of likely paralogues from the next best scoring HMM hits below the collection threshold (Figure S2). Alignments and trimming was performed as above. Phylogenetic trees were inferred by the approximately maximum-likelihood method implemented by FastTree 2.1[74] using the Le-Gascuel (LG) amino acid substitution matrix[75] and a discrete gamma model for heterogeneous site evolution with 12 rate categories. Sets of proteins that formed both a consistent set above threshold in iterative HMM searches and also formed a monophyletic group in phylogenies were considered orthologues. The final list of all predicted BBS protein orthologues of human BBS proteins is given in Table S3, along with the sequences and positions of NLS/NES.

Ancestral sequence reconstruction

Most probable ancestral sequences according to the marginal reconstruction were inferred using fastML v3.11[76,77] using the Wheelan and Goldman (WAG) substitution matrix. Alignments for sequence reconstructions were generated from unaligned sets of BBS orthologues (defined as described above). Regions of sequence occurring only in one organism were removed, but other indels were reconstructed using the likelihood-based mixture model implemented by fastML (–indelReconstruction ML). The branching order for sequence evolution was constrained to be the most likely topology of the tree of eukaryotes (Figure 1) with taxa but with inclusion of only organisms in which an orthologue of the specific BBS protein under consideration was detected. Where an organism contained more than one homolog of a specific BBS protein, only the sequence with the highest similarity to orthologues in other organisms analyzed was included to prevent reconstruction being biased by divergent in-paralogues.

Signal sequence prediction

NLS prediction was by NLStradamus v1.8[78] with a minimum probability threshold of 0.7 (37% sensitivity against a dataset of non-yeast sequences). NES prediction was by NESmapper v.1.0[79] with a minimum score threshold of 4 (73% sensitivity against experimentally validated functional Leptomycin-B-responsive NES from yeast proteins).

Transfection

Transfections were performed with jetPRIME Kit (Polyplus, Illkirch, France) according to protocol: For immunofluorescence analysis, 1.5×10^5 cells were seeded into one well of a 6-well plate and transfected with 2 μg plasmid DNA the following day. For fractionation, 4×10^6 cells were seeded in a 15 cm cell culture dish and transfected with 24 μg plasmid DNA the following day. The p3xFLAG-CMV vector backbone modified with Gateway-compatible *attL* sites (gifted from Stef Letteboer, Radboud University Medical Center, Nijmegen, The Netherlands) was used to produce plasmids with N-terminally 3xFLAG-tagged coding sequences of BBS proteins BBS1, BBS2, BBS4, BBS5, BBS6, BBS7, BBS8, BBS9, BBS10, and BBS12. Integration of coding sequences flanked by *attL* sites into the destination plasmid was facilitated via Gateway-cloning with Gateway LR Clonase II Enzyme Mix (Invitrogen GmbH, Karlsruhe, Germany).

Subcellular fractionation

To assess subcellular localization of FLAG-tagged BBS proteins, transfected cells were lysed and fractionated after Baghirova.[69] In this work, the original protocol was used in the following, modified form: A 15 cm cell culture dish was seeded and transfected as described above. Cells were detached 24 h post transfection with 2 mL TrypLE Express (1X) (Fisher Scientific GmbH, Schwerte, Germany). Then 8 mL DMEM with 10% FBS and 1% Penicillin/Streptomycin were added, and the cell suspension collected. Cells were centrifuged at 500 ×*g* for 10 min at 4°C. All following steps were performed at 4°C in a cold room: After discarding the supernatant, the pellet was resuspended in 500 μL ice-cold PBS (137 mM NaCl; 2.7 mM KCl; 10 mM Na_2HPO_4; 2 mM KH_2PO_4; pH 7.4). The cells were again centrifuged at 500 ×*g* for 10 min at 4°C. The supernatant was discarded and 400 μL Lysis Buffer A (150 mM NaCl; 50 mM HEPES pH 7.4; 25 μg/mL Digitonin; 1 M Hexylene glycol; 1% v/v Protease and Phosphatase Inhibitor Cocktail (Thermo Fisher Scientific GmbH, Dreieich, Germany) freshly before use) was added and the pellet resuspended. After incubating for 10 min on an end-over-end rotator cells were centrifuged at 2000 ×*g* for 10 min at 4°C. The supernatant was collected as cytosolic fraction (C). 400 μL Lysis Buffer B (10 mM NaCl; 50 mM HEPES pH 7.4; 1% v/v NP-40; 1 M Hexylene glycol; 1% v/v Protease and Phosphatase Inhibitor Cocktail freshly before use) was added to the pellet and resuspended by vortexing. After incubating for 30 min on ice the tubes were centrifuged at 7000 ×*g* for 10 min at 4°C. The supernatant was collected and set aside as the membrane-bound organelles fraction (M) to keep the nuclear fraction clean. Then 400 μL Lysis Buffer C supplemented with Urea (150 mM NaCl; 50 mM HEPES pH 7.4; 0.5% w/v Sodium deoxycholate; 0.1% w/v Sodium dodecyl sulfate; 1 M Hexylene glycol; 8 M Urea; 0.875 units/μL Benzonase (Merck KGaA, Darmstadt, Germany) and 1% v/v Protease and Phosphatase Inhibitor Cocktail freshly before use) were added, and the pellet was resuspended by pipetting. The highly viscous liquid was incubated on an end-over-end rotator for 30-45 min, resuspended by pipetting and incubated for another for 30-45 min. This suspension contains the nuclear fraction (N).

Polyacrylamide gel electrophoresis

Subcellular fractions were denatured with Laemmli Buffer (final conc. 2% w/v Sodium dodecyl sulfate; 10% v/v Glycerol; 5% v/v β-Mercaptoethanol; 0.002% w/v Bromophenol blue; 62.5 mM Tris-HCl, pH 8.0) before loading onto polyacrylamide gels. 100 μg of protein was loaded per sample to estimate nuclear protein levels compared to the cytosol. Samples were separated at 13.75 V/cm (V~ const.) for 20 min, then 22.5 V/cm for 35 min in Running Buffer (25 mM Tris; 192 mM Glycine; 1% w/v Sodium dodecyl sulfate).

Western blot

Proteins were transferred from SDS gels to polyvinylidene difluoride (PVDF) membranes (pore size 0.45 μm; Merck KGaA, Darmstadt, Germany) using the Western Blot method. After transfer, the membrane was blocked in Milk Blocking Buffer (5% w/v skim milk powder; 0.1% v/v Tween 20; in TBS [20 mM Tris; 150 mM NaCl; pH 7.6]) for 1 h at room temperature. Primary antibodies were added for 12 h at 4°C at their respective dilutions (see "antibodies"). Primary antibodies were then removed, the membranes washed

thrice in TBS-T (0.1% v/v Tween 20; in TBS) and secondary antibodies were added for 1 h at room temperature at a 1:10000 dilution in Milk Blocking Buffer. Secondary antibodies were discarded and the blot was washed twice in TBS-T and once in TBS. The blots were scanned at the respective excitation wavelengths of the secondary antibody fluorophores in an LI-COR Odyssey Fc Imaging System (LI-COR Biosciences GmbH, Bad Homburg, Germany).

Immunocytochemistry

For immunocytochemical staining, medium was removed from transfected HEK293T cells 24 h post transfection. Cells were washed once with PBS. Cells were fixed by incubation with 4% paraformaldehyde for 10 min at room temperature. Cells were washed three times with 1x PBS. Formaldehyde was quenched by addition of 50 mM NH_4Cl for 10 min at room temperature. Coverslips were incubated for 15 min with PBS-TX (0.3% v/v Triton X-; in PBS) for permeabilisation. For blocking, cells were incubated in Fish Blocking Buffer TX (0.1% w/v Ovalbumin; 0.5% w/v Gelatin from coldwater fish; 0.3% v/v Triton X-; in PBS) for 1 h. Anti-FLAG antibody in Fish Blocking Buffer TX was added to the cells and incubated overnight at 4°C under high humidity. Cells were washed three times with PBS-TX. Alexa Fluor Plus 488 Donkey anti-mouse and 4′,6-Diamidino-2-phenylindol (DAPI, 1:8000; both in Fish Blocking Buffer TX) were added to the cells. After 1h incubation cells were washed twice with PBS-TX and once with PBS. Cells were mounted in Fluoromount-G (SouthernBiotech, Birmingham, AL) on a superfrost slide and curated for 24 h before imaging. Images of immunocytochemically stained cells were acquired on a Leica DM6000B microscope (Leica Microsystems GmbH, Wetzlar, Germany) with a sCMOS Microscope Camera K5 (Leica Microsystems GmbH, Wetzlar, Germany). The software used for image visualization was Leica Application Suite X (LAS X) (Leica Microsystems GmbH, Wetzlar, Germany). Fluorescence images were processed in Fiji[80] with the LIF2TIF image converter.[81] The script to prepare and analyze image micrographs is accessible as a supplementary file.

Antibodies

This study used antibodies for indirect immunofluorescent staining (IF) and protein detection after Western Blotting (WB). The following antibodies were used at the respective dilutions: anti-FLAG (host: mouse; Sigma-Aldrich Chemie GmbH, Taufkirchen, Germany; Cat.no. F1804), IF: 1:200, WB: 1:1000; anti-GAPDH (host: rabbit; Cell Signaling Technology Europe, B.V., Leiden, The Netherlands; Cat.no. 5174), WB: 1:2000; anti-Histone H3 (host: rabbit; Proteintech, Martinsried, Germany; Cat.no. 17168-1-AP), WB: 1:1000; anti-myc (host: mouse; Cell Signaling Technology Europe, B.V., Leiden, The Netherlands; Cat.no. 2276), WB: 1:2000; Alexa Fluor Plus 488 Donkey anti-mouse (host: donkey; Invitrogen GmbH, Karlsruhe, Germany; Cat.no. A-21202), IF: 1:400; IRDye 800 CW anti-Mouse IgG (host: donkey; LI-COR Biosciences GmbH, Bad Homburg, Germany; Cat.no. 926-32212), WB: 1:10000; IRDye 680 RD anti-Mouse IgG (host: donkey; LI-COR Biosciences GmbH, Bad Homburg, Germany; Cat.no. 926-68072), WB: 1:10000; IRDye 800 CW anti-Rabbit IgG (host: donkey; LI-COR Biosciences GmbH, Bad Homburg, Germany; Cat.no. 926-32213), WB: 1:10000; IRDye 680 RD anti-Rabbit IgG (host: donkey; LI-COR Biosciences GmbH, Bad Homburg, Germany; Cat.no. 926-68073), WB: 1:10000.

Software and packages

Other than the software mentioned before, the following programs and software were used: Statistical analyses were conducted in R, version 3.6.3[82]. Violin plots were plotted with the 'vioplot' package for R[83], data frames were manipulated in part with 'dplyr'.[84] Code that was adopted from third parties (e.g. via GitHub or stackoverflow) is acknowledged in the respective places in the supplementary R code. Scripts for parsing files in orthologue searches and modeling were provided by B.W. Figures were made in Inkscape.[85]

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical tests were performed by Monte Carlo sampling. Probability distributions were inferred by creating random samples of equal size to the test set from reference sets conforming to the appropriate characteristics (e.g. clade affiliation, subcellular localization, etc.). 10000 samples were taken in all cases without replacement. Correction for multiple testing was carried out using the Benjamini-Hochberg procedure.[86]

To test for association between NLS/NES and mode of mitosis a similar approach was taken by constructing a metric of NLS/NES scores multiplied by a factor representing the mode of mitosis (where 1 indicates fully open mitosis; 0.5 semi-open; −0.5 semi-closed; and −1 fully closed). Factors for species for which mode of mitosis is not known were set to 0. To test the effect of mitotic mode, datasets of equal size as the test set were randomly generated as above, but with permutation of modes of mitosis across the set. 10000 iterations were used to estimate the probability distribution.

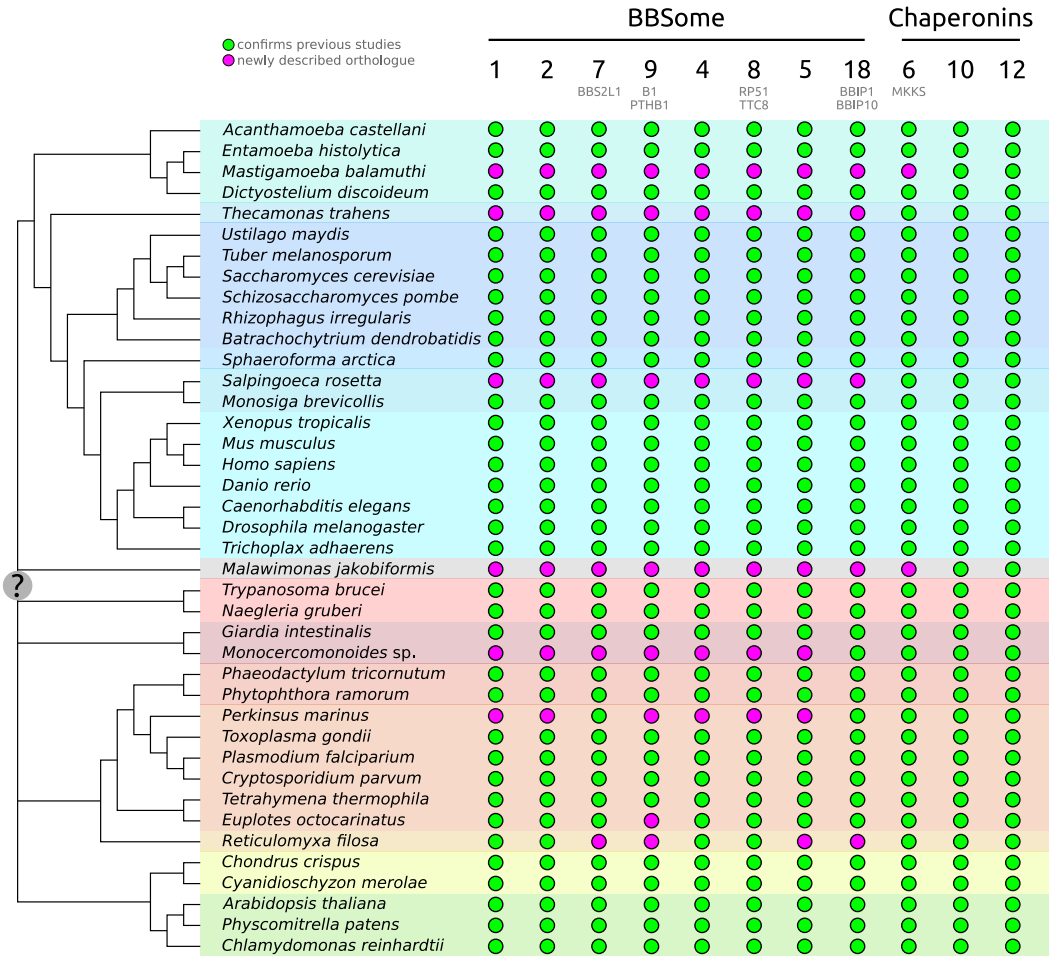
Significance levels were defined as follows: ns: not significant, $0.05 \leq p$; *: $p < 0.05$.

iScience, Volume *26*

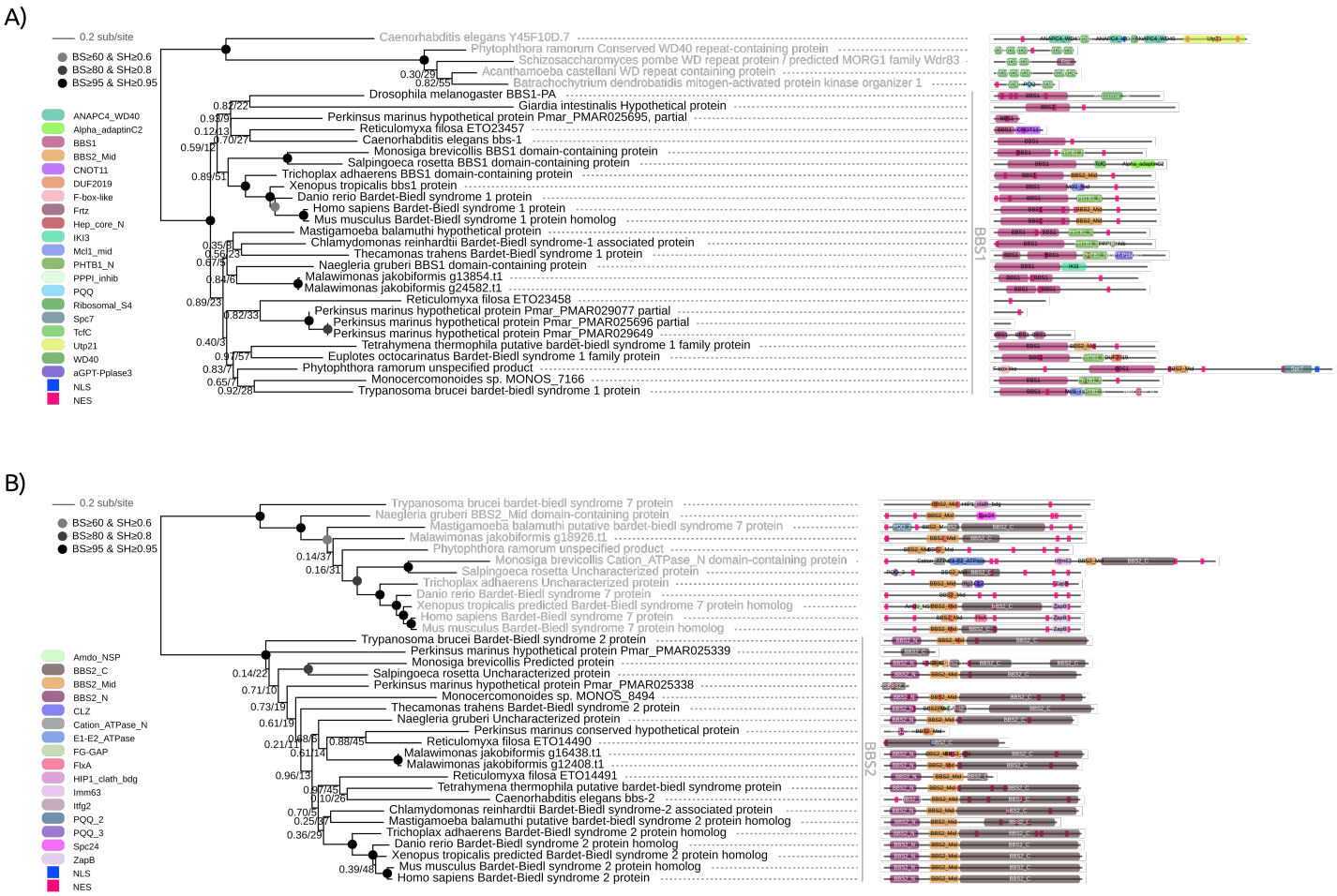
Supplemental information

Neofunctionalization of ciliary BBS proteins to nuclear roles is likely a frequent innovation across eukaryotes

Alexander Ewerling, Vanessa Maissl, Bill Wickstead, and Helen Louise May-Simera

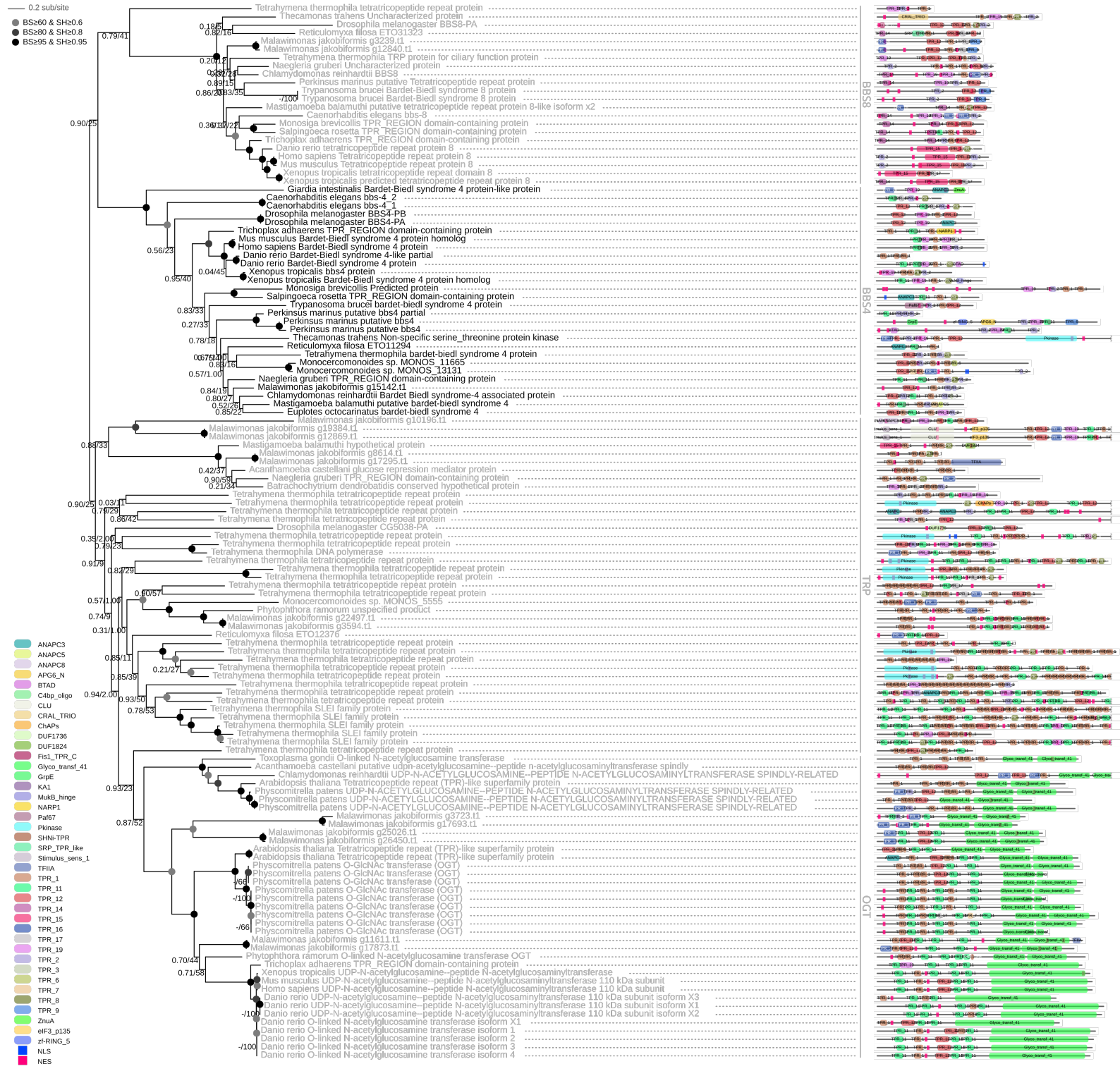

BBSome
Chaperonins
confirms previous studies
newly described orthologue
1
2
7
BBS2L1
9
B1
PTHB1
4
8
RP51
TTC8
5
18
BBIP1
BBIP10
6
MKKS
10
12
Acanthamoeba castellani
Entamoeba histolytica
Mastigamoeba balamuthi
Dictyostelium discoideum
Thecamonas trahens
Ustilago maydis
Tuber melanosporum
Saccharomyces cerevisiae
Schizosaccharomyces pombe
Rhizophagus irregularis
Batrachochytrium dendrobatidis
Sphaeroforma arctica
Salpingoeca rosetta
Monosiga brevicollis
Xenopus tropicalis
Mus musculus
Homo sapiens
Danio rerio
Caenorhabditis elegans
Drosophila melanogaster
Trichoplax adhaerens
Malawimonas jakobiformis
?
Trypanosoma brucei
Naegleria gruberi
Giardia intestinalis
Monocercomonoides sp.
Phaeodactylum tricornutum
Phytophthora ramorum
Perkinsus marinus
Toxoplasma gondii
Plasmodium falciparium
Cryptosporidium parvum
Tetrahymena thermophila
Euplotes octocarinatus
Reticulomyxa filosa
Chondrus crispus
Cyanidioschyzon merolae
Arabidopsis thaliana
Physcomitrella patens
Chlamydomonas reinhardtii

Supp. Fig.1


A)
0.2 sub/site
BS≥60 & SH≥0.6
BS≥80 & SH≥0.8
BS≥95 & SH≥0.95
ANAPC4_WD40
Alpha_adaptinC2
BBS1
BBS2_Mid
CNOT11
DUF2019
F-box-like
Frtz
Hep_core_N
IKI3
Mcl1_mid
PHTB1_N
PPPI_inhib
PQQ
Ribosomal_S4
Spc7
TcfC
Utp21
WD40
aGPT-Pplase3
NLS
NES
Caenorhabditis elegans Y45F10D.7
Phytophthora ramorum Conserved WD40 repeat-containing protein
Schizosaccharomyces pombe WD repeat protein / predicted MORG1 family Wdr83
Acanthamoeba castellani WD repeat containing protein
Batrachochytrium dendrobatidis mitogen-activated protein kinase organizer 1
Drosophila melanogaster BBS1-PA
Giardia intestinalis Hypothetical protein
Perkinsus marinus hypothetical protein Pmar_PMAR025695, partial
Reticulomyxa filosa ETO23457
Caenorhabditis elegans bbs-1
Monosiga brevicollis BBS1 domain-containing protein
Salpingoeca rosetta BBS1 domain-containing protein
Trichoplax adhaerens BBS1 domain-containing protein
Xenopus tropicalis bbs1 protein
Danio rerio Bardet-Biedl syndrome 1 protein
Homo sapiens Bardet-Biedl syndrome 1 protein
Mus musculus Bardet-Biedl syndrome 1 protein homolog
Mastigamoeba balamuthi hypothetical protein
Chlamydomonas reinhardtii Bardet-Biedl syndrome-1 associated protein
Thecamonas trahens Bardet-Biedl syndrome 1 protein
Naegleria gruberi BBS1 domain-containing protein
Malawimonas jakobiformis g13854.t1
Malawimonas jakobiformis g24582.t1
Reticulomyxa filosa ETO23458
Perkinsus marinus hypothetical protein Pmar_PMAR029077 partial
Perkinsus marinus hypothetical protein Pmar_PMAR025696 partial
Perkinsus marinus hypothetical protein Pmar_PMAR029649
Tetrahymena thermophila putative bardet-biedl syndrome 1 family protein
Euplotes octocarinatus Bardet-Biedl syndrome 1 family protein
Phytophthora ramorum unspecified product
Monocercomonoides sp. MONOS_7166
Trypanosoma brucei bardet-biedl syndrome 1 protein
0.30/29
0.82/55
0.82/22
0.95/9
0.12/13
0.70/27
0.59/12
0.89/51
0.35/8
0.56/23
0.67/5
0.84/6
0.89/23
0.82/33
0.40/3
0.97/57
0.83/7
0.65/7
0.92/28
BBS1
B)
0.2 sub/site
BS≥60 & SH≥0.6
BS≥80 & SH≥0.8
BS≥95 & SH≥0.95
Amdo_NSP
BBS2_C
BBS2_Mid
BBS2_N
CLZ
Cation_ATPase_N
E1-E2_ATPase
FG-GAP
FlxA
HIP1_clath_bdg
Imm63
Itfg2
PQQ_2
PQQ_3
Spc24
ZapB
NLS
NES
Trypanosoma brucei bardet-biedl syndrome 7 protein
Naegleria gruberi BBS2_Mid domain-containing protein
Mastigamoeba balamuthi putative bardet-biedl syndrome 7 protein
Malawimonas jakobiformis g18926.t1
Phytophthora ramorum unspecified product
Monosiga brevicollis Cation_ATPase_N domain-containing protein
Salpingoeca rosetta Uncharacterized protein
Trichoplax adhaerens Uncharacterized protein
Danio rerio Bardet-Biedl syndrome 7 protein
Xenopus tropicalis predicted Bardet-Biedl syndrome 7 protein homolog
Homo sapiens Bardet-Biedl syndrome 7 protein
Mus musculus Bardet-Biedl syndrome 7 protein homolog
Trypanosoma brucei Bardet-Biedl syndrome 2 protein
Perkinsus marinus hypothetical protein Pmar_PMAR025339
Monosiga brevicollis Predicted protein
Salpingoeca rosetta Uncharacterized protein
Perkinsus marinus hypothetical protein Pmar_PMAR025338
Monocercomonoides sp. MONOS_8494
Thecamonas trahens Bardet-Biedl syndrome 2 protein
Naegleria gruberi Uncharacterized protein
Perkinsus marinus conserved hypothetical protein
Reticulomyxa filosa ETO14490
Malawimonas jakobiformis g16438.t1
Malawimonas jakobiformis g12408.t1
Reticulomyxa filosa ETO14491
Tetrahymena thermophila putative bardet-biedl syndrome protein
Caenorhabditis elegans bbs-2
Chlamydomonas reinhardtii Bardet-Biedl syndrome-2 associated protein
Mastigamoeba balamuthi putative bardet-biedl syndrome 2 protein homolog
Trichoplax adhaerens Bardet-Biedl syndrome 2 protein homolog
Danio rerio Bardet-Biedl syndrome 2 protein homolog
Xenopus tropicalis predicted Bardet-Biedl syndrome 2 protein homolog
Mus musculus Bardet-Biedl syndrome 2 protein homolog
Homo sapiens Bardet-Biedl syndrome 2 protein
0.14/37
0.16/31
0.14/22
0.71/10
0.73/19
0.61/19
0.88/45
0.61/14
0.96/13
0.97/45
0.10/26
0.70/5
0.25/17
0.36/25
0.39/48
BBS2

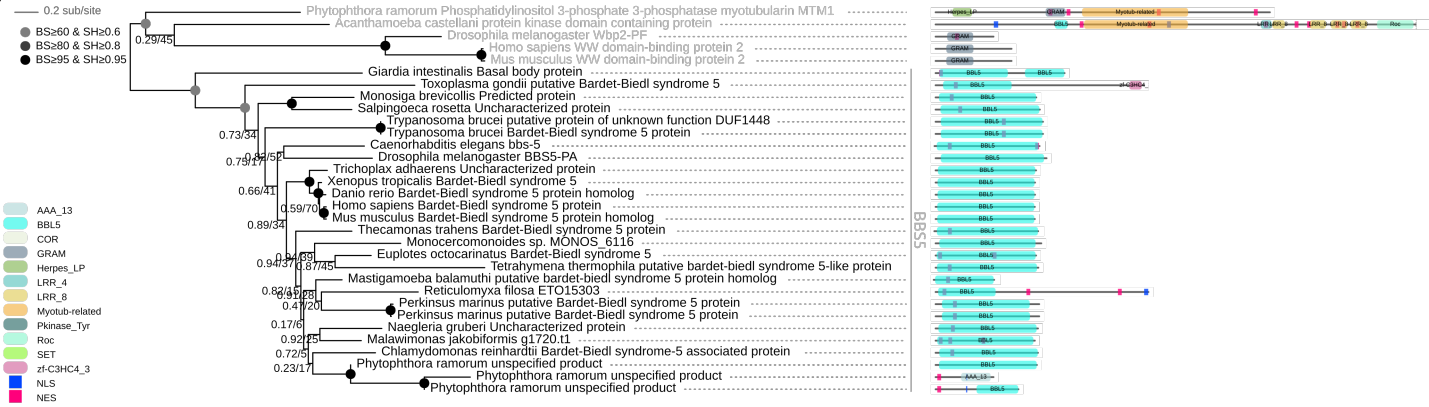
Supp. Fig. 2

C)

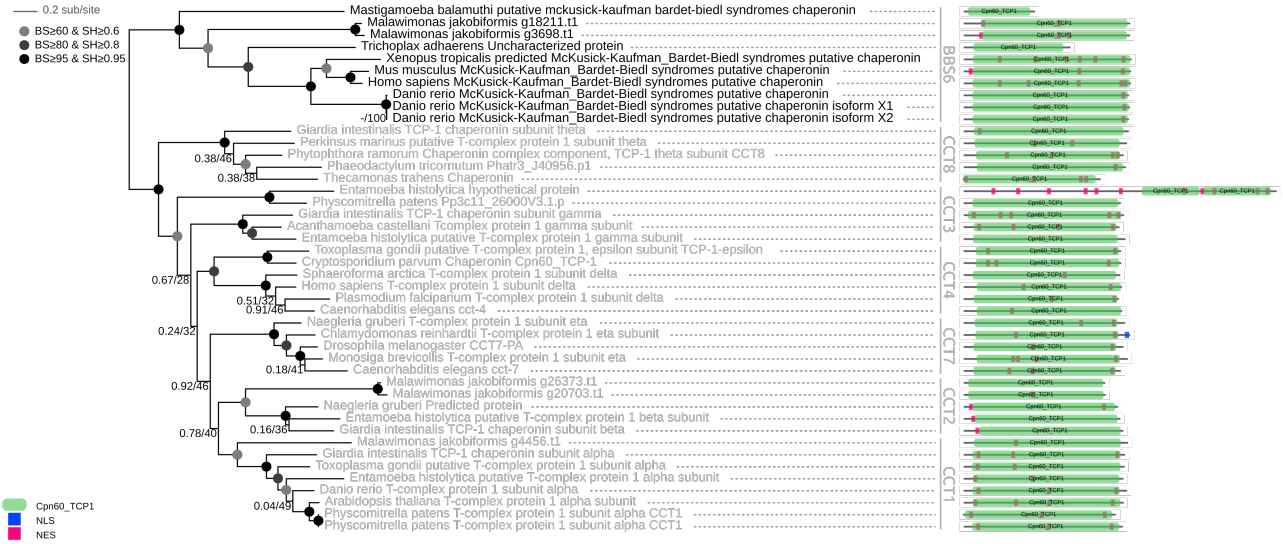

0.2 sub/site
BS≥60 & SH≥0.6
BS≥80 & SH≥0.8
BS≥95 & SH≥0.95
Tetrahymena thermophila tetratricopeptide repeat protein
Thecamonas trahens Uncharacterized protein
Drosophila melanogaster BBS8-PA
Reticulomyxa filosa ETO31323
Malawimonas jakobiformis g3239.t1
Malawimonas jakobiformis g12840.t1
Tetrahymena thermophila TRP protein for ciliary function protein
Naegleria gruberi Uncharacterized protein
Chlamydomonas reinhardtii BBS8
Perkinsus marinus putative Tetratricopeptide repeat protein
Trypanosoma brucei Bardet-Biedl syndrome 8 protein
Trypanosoma brucei Bardet-Biedl syndrome 8 protein
Mastigamoeba balamuthi putative tetratricopeptide repeat protein 8-like isoform x2
Caenorhabditis elegans bbs-8
Monosiga brevicollis TPR_REGION domain-containing protein
Salpingoeca rosetta TPR_REGION domain-containing protein
Trichoplax adhaerens TPR_REGION domain-containing protein
Danio rerio tetratricopeptide repeat protein 8
Homo sapiens Tetratricopeptide repeat protein 8
Mus musculus Tetratricopeptide repeat protein 8
Xenopus tropicalis tetratricopeptide repeat domain 8
Xenopus tropicalis predicted tetratricopeptide repeat protein 8
Giardia intestinalis Bardet-Biedl syndrome 4 protein-like protein
Caenorhabditis elegans bbs-4_2
Caenorhabditis elegans bbs-4_1
Drosophila melanogaster BBS4-PB
Drosophila melanogaster BBS4-PA
Trichoplax adhaerens TPR_REGION domain-containing protein
Mus musculus Bardet-Biedl syndrome 4 protein homolog
Homo sapiens Bardet-Biedl syndrome 4 protein
Danio rerio Bardet-Biedl syndrome 4-like partial
Danio rerio Bardet-Biedl syndrome 4 protein
Xenopus tropicalis bbs4 protein
Xenopus tropicalis Bardet-Biedl syndrome 4 protein homolog
Monosiga brevicollis Predicted protein
Salpingoeca rosetta TPR_REGION domain-containing protein
Trypanosoma brucei bardet-biedl syndrome 4 protein
Perkinsus marinus putative bbs4 partial
Perkinsus marinus putative bbs4
Perkinsus marinus putative bbs4
Thecamonas trahens Non-specific serine_threonine protein kinase
Reticulomyxa filosa ETO11294
Tetrahymena thermophila bardet-biedl syndrome 4 protein
Monocercomonoides sp. MONOS_11665
Monocercomonoides sp. MONOS_13131
Naegleria gruberi TPR_REGION domain-containing protein
Malawimonas jakobiformis g15142.t1
Chlamydomonas reinhardtii Bardet Biedl syndrome-4 associated protein
Mastigamoeba balamuthi putative bardet-biedl syndrome 4
Euplotes octocarinatus bardet-biedl syndrome 4
Malawimonas jakobiformis g10196.t1
Malawimonas jakobiformis g19384.t1
Malawimonas jakobiformis g12869.t1
Mastigamoeba balamuthi hypothetical protein
Malawimonas jakobiformis g8614.t1
Malawimonas jakobiformis g17295.t1
Acanthamoeba castellani glucose repression mediator protein
Naegleria gruberi TPR_REGION domain-containing protein
Batrachochytrium dendrobatidis conserved hypothetical protein
Tetrahymena thermophila tetratricopeptide repeat protein
Tetrahymena thermophila tetratricopeptide repeat protein
Tetrahymena thermophila tetratricopeptide repeat protein
Tetrahymena thermophila tetratricopeptide repeat protein
Drosophila melanogaster CG5038-PA
Tetrahymena thermophila tetratricopeptide repeat protein
Tetrahymena thermophila tetratricopeptide repeat protein
Tetrahymena thermophila DNA polymerase
Tetrahymena thermophila tetratricopeptide repeat protein
Tetrahymena thermophila tetratricopeptide repeat protein
Tetrahymena thermophila tetratricopeptide repeat protein
Tetrahymena thermophila tetratricopeptide repeat protein
Tetrahymena thermophila tetratricopeptide repeat protein
Monocercomonoides sp. MONOS_5555
Phytophthora ramorum unspecified product
Malawimonas jakobiformis g22497.t1
Malawimonas jakobiformis g3594.t1
Reticulomyxa filosa ETO12376
Tetrahymena thermophila tetratricopeptide repeat protein
Tetrahymena thermophila tetratricopeptide repeat protein
Tetrahymena thermophila tetratricopeptide repeat protein
Tetrahymena thermophila tetratricopeptide repeat protein
Tetrahymena thermophila tetratricopeptide repeat protein
Tetrahymena thermophila tetratricopeptide repeat protein
Tetrahymena thermophila tetratricopeptide repeat protein
Tetrahymena thermophila SLEI family protein
Tetrahymena thermophila SLEI family protein
Tetrahymena thermophila SLEI family protein
Tetrahymena thermophila SLEI family protein
Tetrahymena thermophila tetratricopeptide repeat protein
Toxoplasma gondii O-linked N-acetylglucosamine transferase
Acanthamoeba castellani putative udpn-acetylglucosamine--peptide n-acetylglucosaminyltransferase spindly
Chlamydomonas reinhardtii UDP-N-ACETYLGLUCOSAMINE--PEPTIDE N-ACETYLGLUCOSAMINYLTRANSFERASE SPINDLY-RELATED
Arabidopsis thaliana Tetratricopeptide repeat (TPR)-like superfamily protein
Physcomitrella patens UDP-N-ACETYLGLUCOSAMINE--PEPTIDE N-ACETYLGLUCOSAMINYLTRANSFERASE SPINDLY-RELATED
Physcomitrella patens UDP-N-ACETYLGLUCOSAMINE--PEPTIDE N-ACETYLGLUCOSAMINYLTRANSFERASE SPINDLY-RELATED
Physcomitrella patens UDP-N-ACETYLGLUCOSAMINE--PEPTIDE N-ACETYLGLUCOSAMINYLTRANSFERASE SPINDLY-RELATED
Malawimonas jakobiformis g3723.t1
Malawimonas jakobiformis g17693.t1
Malawimonas jakobiformis g25026.t1
Malawimonas jakobiformis g26450.t1
Arabidopsis thaliana Tetratricopeptide repeat (TPR)-like superfamily protein
Arabidopsis thaliana Tetratricopeptide repeat (TPR)-like superfamily protein
Physcomitrella patens O-GlcNAc transferase (OGT)
Physcomitrella patens O-GlcNAc transferase (OGT)
Physcomitrella patens O-GlcNAc transferase (OGT)
Physcomitrella patens O-GlcNAc transferase (OGT)
Physcomitrella patens O-GlcNAc transferase (OGT)
Physcomitrella patens O-GlcNAc transferase (OGT)
Physcomitrella patens O-GlcNAc transferase (OGT)
Malawimonas jakobiformis g11611.t1
Malawimonas jakobiformis g17873.t1
Phytophthora ramorum O-linked N-acetylglucosamine transferase OGT
Trichoplax adhaerens TPR_REGION domain-containing protein
Xenopus tropicalis UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase
Mus musculus UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit
Homo sapiens UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit
Danio rerio UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit isoform X3
Danio rerio UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit isoform X1
Danio rerio UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit isoform X2
Danio rerio O-linked N-acetylglucosamine transferase isoform X1
Danio rerio O-linked N-acetylglucosamine transferase isoform 1
Danio rerio O-linked N-acetylglucosamine transferase isoform 2
Danio rerio O-linked N-acetylglucosamine transferase isoform 3
Danio rerio O-linked N-acetylglucosamine transferase isoform 4
BBS8
BBS4
TRP
OGT
ANAPC3
ANAPC5
ANAPC8
APG6_N
BTAD
C4bp_oligo
CLU
CRAL_TRIO
ChAPs
DUF1736
DUF1824
Fis1_TPR_C
Glyco_transf_41
GrpE
KA1
MukB_hinge
NARP1
Paf67
Pkinase
SHNi-TPR
SRP_TPR_like
Stimulus_sens_1
TFIIA
TPR_1
TPR_11
TPR_12
TPR_14
TPR_15
TPR_16
TPR_17
TPR_19
TPR_2
TPR_3
TPR_6
TPR_7
TPR_8
TPR_9
ZnuA
eIF3_p135
zf-RING_5
NLS
NES

Supp. Fig. 2

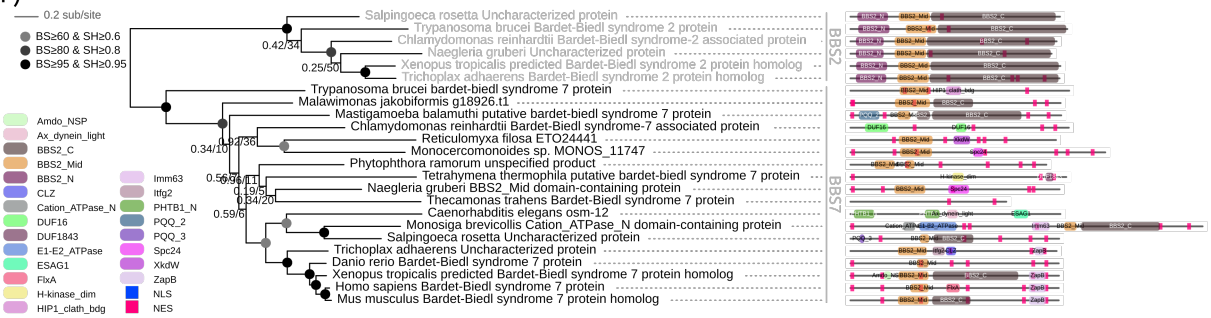
D)



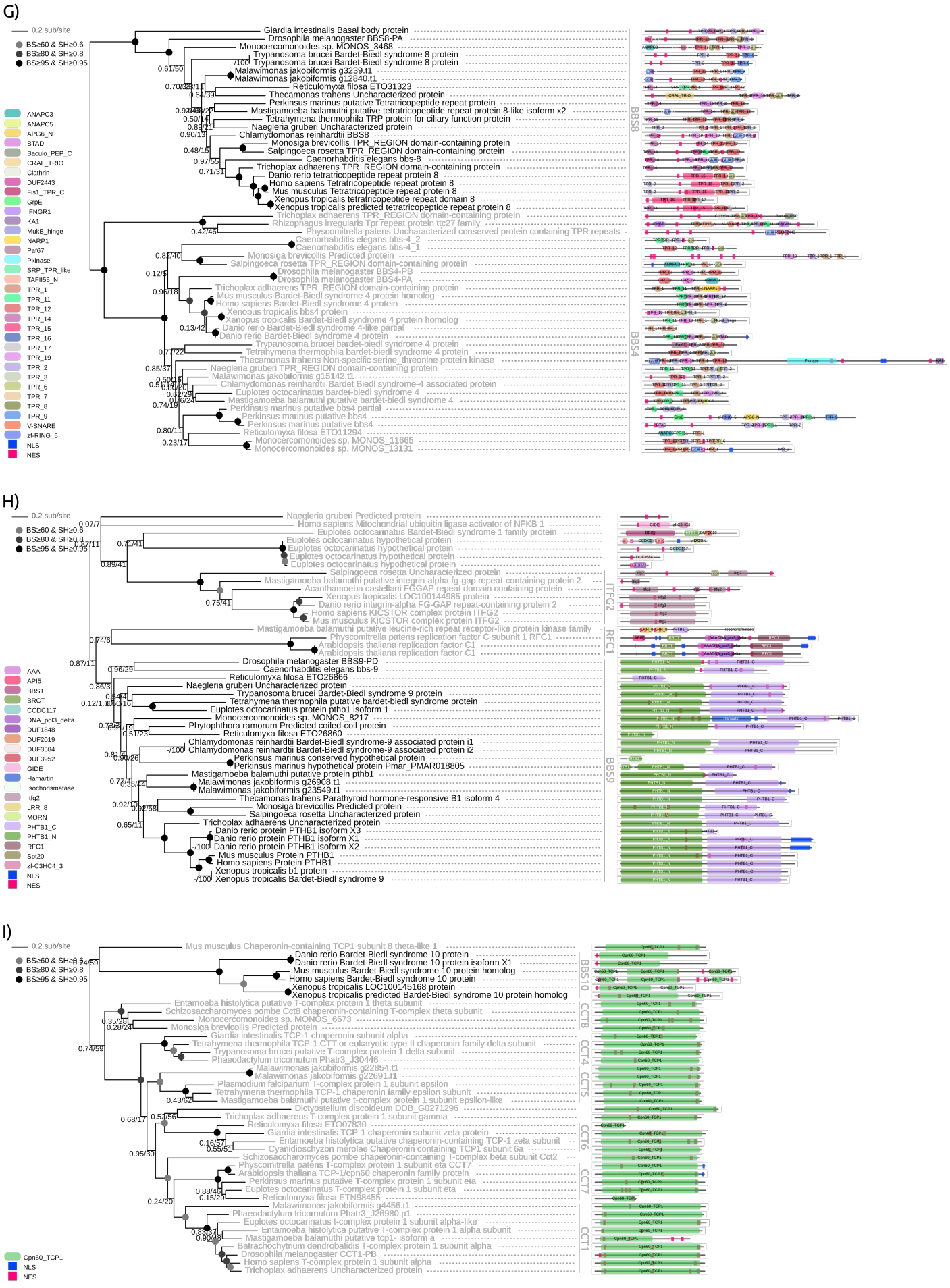
E)



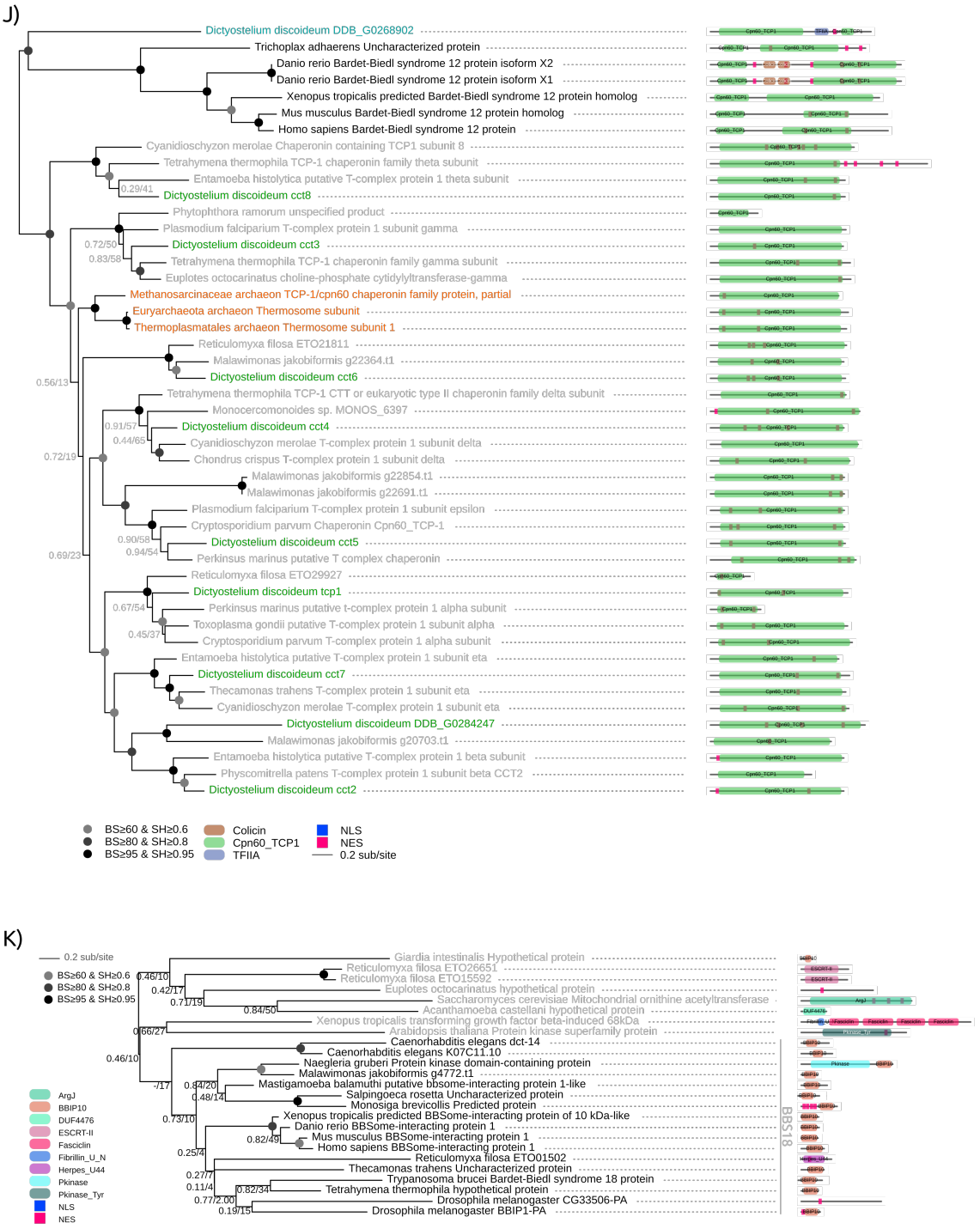
F)



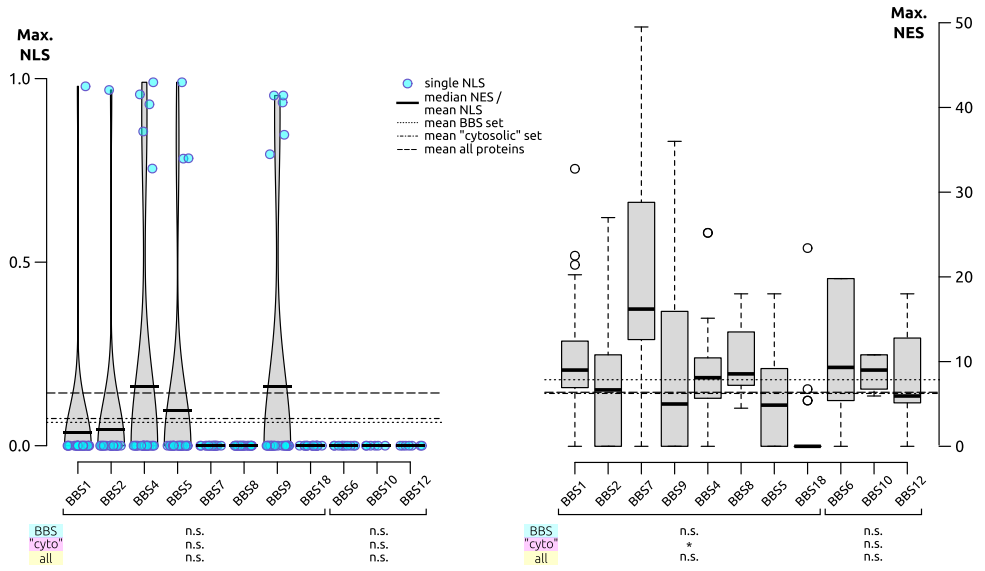
Supp. Fig. 2


G)
0.2 sub/site
BS≥60 & SH≥0.6
BS≥80 & SH≥0.8
BS≥95 & SH≥0.95
ANAPC3
ANAPC5
APG6_N
BTAD
Baculo_PEP_C
CRAL_TRIO
Clathrin
DUF2443
Fis1_TPR_C
GrpE
IFNGR1
KA1
MukB_hinge
NARP1
Paf67
Pkinase
SRP_TPR_like
TAFII55_N
TPR_1
TPR_11
TPR_12
TPR_14
TPR_15
TPR_16
TPR_17
TPR_19
TPR_2
TPR_3
TPR_6
TPR_7
TPR_8
TPR_9
V-SNARE
zf-RING_5
NLS
NES
Giardia intestinalis Basal body protein
Drosophila melanogaster BBS8-PA
Monocercomonoides sp. MONOS_3468
Trypanosoma brucei Bardet-Biedl syndrome 8 protein
Trypanosoma brucei Bardet-Biedl syndrome 8 protein
Malawimonas jakobiformis g3239.t1
Malawimonas jakobiformis g12840.t1
Reticulomyxa filosa ETO31323
Thecamonas trahens Uncharacterized protein
Perkinsus marinus putative Tetratricopeptide repeat protein
Mastigamoeba balamuthi putative tetratricopeptide repeat protein 8-like isoform x2
Tetrahymena thermophila TRP protein for ciliary function protein
Naegleria gruberi Uncharacterized protein
Chlamydomonas reinhardtii BBS8
Monosiga brevicollis TPR_REGION domain-containing protein
Salpingoeca rosetta TPR_REGION domain-containing protein
Caenorhabditis elegans bbs-8
Trichoplax adhaerens TPR_REGION domain-containing protein
Danio rerio tetratricopeptide repeat protein 8
Homo sapiens Tetratricopeptide repeat protein 8
Mus musculus Tetratricopeptide repeat protein 8
Xenopus tropicalis tetratricopeptide repeat domain 8
Xenopus tropicalis predicted tetratricopeptide repeat protein 8
Trichoplax adhaerens TPR_REGION domain-containing protein
Rhizophagus irregularis Tpr repeat protein tic27 family
Physcomitrella patens Uncharacterized conserved protein containing TPR repeats
Caenorhabditis elegans bbs-4_2
Caenorhabditis elegans bbs-4_1
Monosiga brevicollis Predicted protein
Salpingoeca rosetta TPR_REGION domain-containing protein
Drosophila melanogaster BBS4-PB
Drosophila melanogaster BBS4-PA
Trichoplax adhaerens TPR_REGION domain-containing protein
Mus musculus Bardet-Biedl syndrome 4 protein homolog
Homo sapiens Bardet-Biedl syndrome 4 protein
Xenopus tropicalis bbs4 protein
Xenopus tropicalis Bardet-Biedl syndrome 4 protein homolog
Danio rerio Bardet-Biedl syndrome 4-like partial
Danio rerio Bardet-Biedl syndrome 4 protein
Trypanosoma brucei bardet-biedl syndrome 4 protein
Tetrahymena thermophila bardet-biedl syndrome 4 protein
Thecamonas trahens Non-specific serine_threonine protein kinase
Naegleria gruberi TPR_REGION domain-containing protein
Malawimonas jakobiformis g15142.t1
Chlamydomonas reinhardtii Bardet Biedl syndrome-4 associated protein
Euplotes octocarinatus bardet-biedl syndrome 4
Mastigamoeba balamuthi putative bardet-biedl syndrome 4
Perkinsus marinus putative bbs4 partial
Perkinsus marinus putative bbs4
Perkinsus marinus putative bbs4
Reticulomyxa filosa ETO11294
Monocercomonoides sp. MONOS_11665
Monocercomonoides sp. MONOS_13131
BBS8
BBS4
H)
0.2 sub/site
BS≥60 & SH≥0.6
BS≥80 & SH≥0.8
BS≥95 & SH≥0.95
AAA
API5
BBS1
BRCT
CCDC117
DNA_pol3_delta
DUF1848
DUF2019
DUF3584
DUF3952
GIDE
Hamartin
Isochorismatase
Itfg2
LRR_8
MORN
PHTB1_C
PHTB1_N
RFC1
Spt20
zf-C3HC4_3
NLS
NES
Naegleria gruberi Predicted protein
Homo sapiens Mitochondrial ubiquitin ligase activator of NFKB 1
Euplotes octocarinatus Bardet-Biedl syndrome 1 family protein
Euplotes octocarinatus hypothetical protein
Euplotes octocarinatus hypothetical protein
Euplotes octocarinatus hypothetical protein
Euplotes octocarinatus hypothetical protein
Salpingoeca rosetta Uncharacterized protein
Mastigamoeba balamuthi putative integrin-alpha fg-gap repeat-containing protein 2
Acanthamoeba castellani FGGAP repeat domain containing protein
Xenopus tropicalis LOC100144985 protein
Danio rerio integrin-alpha FG-GAP repeat-containing protein 2
Homo sapiens KICSTOR complex protein ITFG2
Mus musculus KICSTOR complex protein ITFG2
Mastigamoeba balamuthi putative leucine-rich repeat receptor-like protein kinase family
Physcomitrella patens replication factor C subunit 1 RFC1
Arabidopsis thaliana replication factor C1
Arabidopsis thaliana replication factor C1
Drosophila melanogaster BBS9-PD
Caenorhabditis elegans bbs-9
Reticulomyxa filosa ETO26866
Naegleria gruberi Uncharacterized protein
Trypanosoma brucei Bardet-Biedl syndrome 9 protein
Tetrahymena thermophila putative bardet-biedl syndrome protein
Euplotes octocarinatus protein pthb1 isoform 1
Monocercomonoides sp. MONOS_8217
Phytophthora ramorum Predicted coiled-coil protein
Reticulomyxa filosa ETO26660
Chlamydomonas reinhardtii Bardet-Biedl syndrome-9 associated protein i1
Chlamydomonas reinhardtii Bardet-Biedl syndrome-9 associated protein i2
Perkinsus marinus conserved hypothetical protein
Perkinsus marinus hypothetical protein Pmar_PMAR018805
Mastigamoeba balamuthi putative protein pthb1
Malawimonas jakobiformis g26908.t1
Malawimonas jakobiformis g23549.t1
Thecamonas trahens Parathyroid hormone-responsive B1 isoform 4
Monosiga brevicollis Predicted protein
Salpingoeca rosetta Uncharacterized protein
Trichoplax adhaerens Uncharacterized protein
Danio rerio protein PTHB1 isoform X3
Danio rerio protein PTHB1 isoform X1
Danio rerio protein PTHB1 isoform X2
Mus musculus Protein PTHB1
Homo sapiens Protein PTHB1
Xenopus tropicalis b1 protein
Xenopus tropicalis Bardet-Biedl syndrome 9
ITFG2
RFC1
BBS9
I)
0.2 sub/site
BS≥60 & SH≥0.6
BS≥80 & SH≥0.8
BS≥95 & SH≥0.95
Cpn60_TCP1
NLS
NES
Mus musculus Chaperonin-containing TCP1 subunit 8 theta-like 1
Danio rerio Bardet-Biedl syndrome 10 protein
Danio rerio Bardet-Biedl syndrome 10 protein isoform X1
Mus musculus Bardet-Biedl syndrome 10 protein homolog
Homo sapiens Bardet-Biedl syndrome 10 protein
Xenopus tropicalis LOC100145168 protein
Xenopus tropicalis predicted Bardet-Biedl syndrome 10 protein homolog
Entamoeba histolytica putative T-complex protein 1 theta subunit
Schizosaccharomyces pombe Cct8 chaperonin-containing T-complex theta subunit
Monocercomonoides sp. MONOS_6673
Monosiga brevicollis Predicted protein
Giardia intestinalis TCP-1 chaperonin subunit alpha
Tetrahymena thermophila TCP-1 CTT or eukaryotic type II chaperonin family delta subunit
Trypanosoma brucei putative T-complex protein 1 delta subunit
Phaeodactylum tricornutum Phatr3_J30446
Malawimonas jakobiformis g22854.t1
Malawimonas jakobiformis g22691.t1
Plasmodium falciparum T-complex protein 1 subunit epsilon
Tetrahymena thermophila TCP-1 chaperonin family epsilon subunit
Mastigamoeba balamuthi putative t-complex protein 1 subunit epsilon-like
Dictyostelium discoideum DDB_G0271296
Trichoplax adhaerens T-complex protein 1 subunit gamma
Reticulomyxa filosa ETO07830
Giardia intestinalis TCP-1 chaperonin subunit zeta protein
Entamoeba histolytica putative chaperonin-containing TCP-1 zeta subunit
Cyanidioschyzon merolae Chaperonin containing TCP1 subunit 6a
Schizosaccharomyces pombe chaperonin-containing T-complex beta subunit Cct2
Physcomitrella patens T-complex protein 1 subunit eta CCT7
Arabidopsis thaliana TCP-1/cpn60 chaperonin family protein
Perkinsus marinus putative T-complex protein 1 subunit eta
Euplotes octocarinatus T-complex protein 1 subunit eta
Reticulomyxa filosa ETN98455
Malawimonas jakobiformis g4456.t1
Phaeodactylum tricornutum Phatr3_J26980.p1
Euplotes octocarinatus t-complex protein 1 subunit alpha-like
Entamoeba histolytica putative T-complex protein 1 alpha subunit
Mastigamoeba balamuthi putative tcp1- isoform a
Batrachochytrium dendrobatidis T-complex protein 1 subunit alpha
Drosophila melanogaster CCT1-PB
Homo sapiens T-complex protein 1 subunit alpha
Trichoplax adhaerens Uncharacterized protein
BBS10
CCT8
CCT4
CCT5
CCT16
CCT7
CCT1

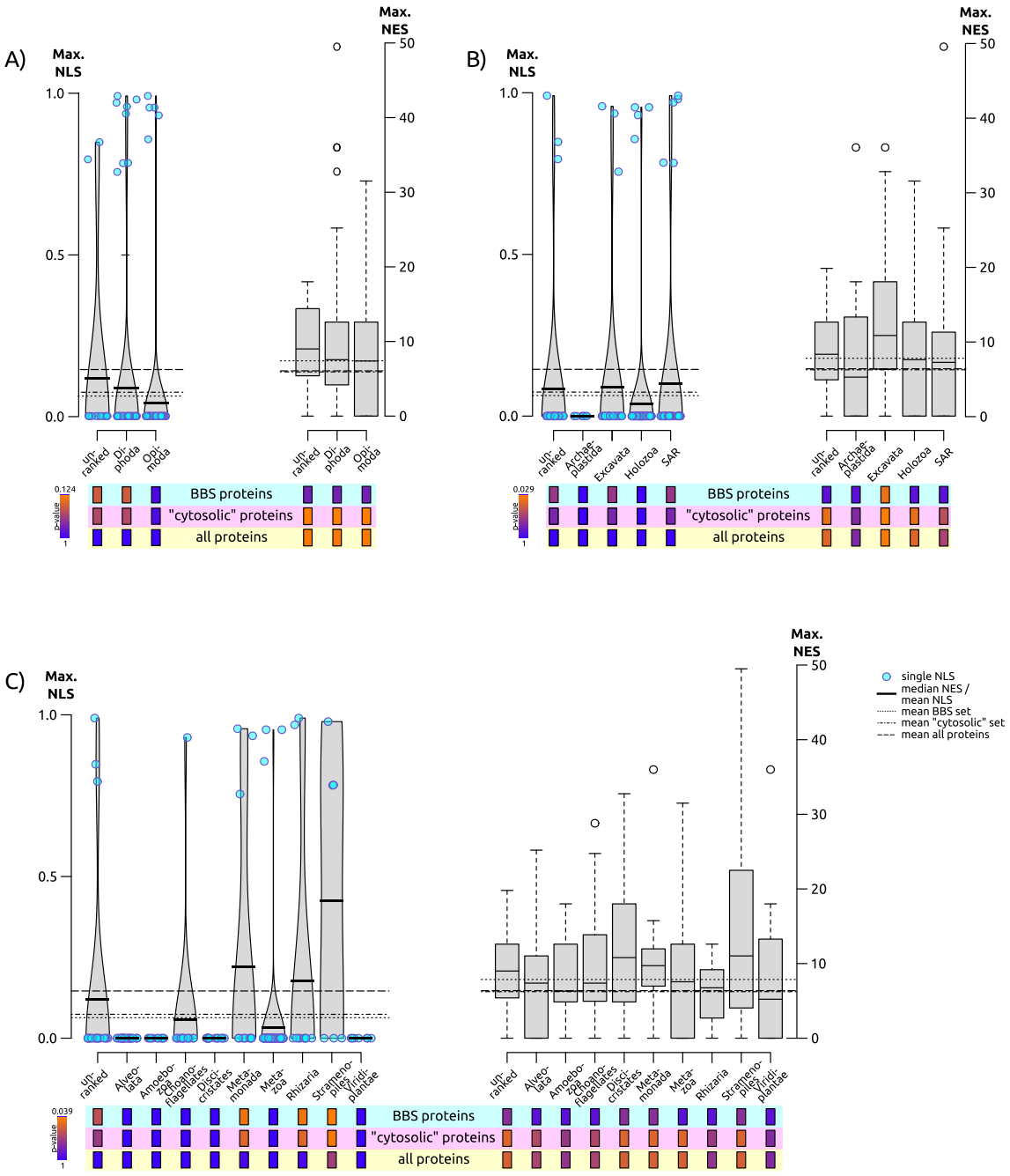
Supp. Fig. 2


J)
Dictyostelium discoideum DDB_G0268902
Trichoplax adhaerens Uncharacterized protein
Danio rerio Bardet-Biedl syndrome 12 protein isoform X2
Danio rerio Bardet-Biedl syndrome 12 protein isoform X1
Xenopus tropicalis predicted Bardet-Biedl syndrome 12 protein homolog
Mus musculus Bardet-Biedl syndrome 12 protein homolog
Homo sapiens Bardet-Biedl syndrome 12 protein
Cyanidioschyzon merolae Chaperonin containing TCP1 subunit 8
Tetrahymena thermophila TCP-1 chaperonin family theta subunit
Entamoeba histolytica putative T-complex protein 1 theta subunit
0.29/41
Dictyostelium discoideum cct8
Phytophthora ramorum unspecified product
Plasmodium falciparum T-complex protein 1 subunit gamma
0.72/50
0.83/58
Dictyostelium discoideum cct3
Tetrahymena thermophila TCP-1 chaperonin family gamma subunit
Euplotes octocarinatus choline-phosphate cytidylyltransferase-gamma
Methanosarcinaceae archaeon TCP-1/cpn60 chaperonin family protein, partial
Euryarchaeota archaeon Thermosome subunit
Thermoplasmatales archaeon Thermosome subunit 1
Reticulomyxa filosa ETO21811
Malawimonas jakobiformis g22364.t1
Dictyostelium discoideum cct6
0.56/13
Tetrahymena thermophila TCP-1 CTT or eukaryotic type II chaperonin family delta subunit
Monocercomonoides sp. MONOS_6397
0.91/57
Dictyostelium discoideum cct4
0.44/65
Cyanidioschyzon merolae T-complex protein 1 subunit delta
0.72/19
Chondrus crispus T-complex protein 1 subunit delta
Malawimonas jakobiformis g22854.t1
Malawimonas jakobiformis g22691.t1
Plasmodium falciparum T-complex protein 1 subunit epsilon
Cryptosporidium parvum Chaperonin Cpn60_TCP-1
0.90/58
Dictyostelium discoideum cct5
0.94/54
0.69/23
Perkinsus marinus putative T complex chaperonin
Reticulomyxa filosa ETO29927
Dictyostelium discoideum tcp1
0.67/54
Perkinsus marinus putative t-complex protein 1 alpha subunit
Toxoplasma gondii putative T-complex protein 1 subunit alpha
0.45/37
Cryptosporidium parvum T-complex protein 1 alpha subunit
Entamoeba histolytica putative T-complex protein 1 subunit eta
Dictyostelium discoideum cct7
Thecamonas trahens T-complex protein 1 subunit eta
Cyanidioschyzon merolae T-complex protein 1 subunit eta
Dictyostelium discoideum DDB_G0284247
Malawimonas jakobiformis g20703.t1
Entamoeba histolytica putative T-complex protein 1 beta subunit
Physcomitrella patens T-complex protein 1 subunit beta CCT2
Dictyostelium discoideum cct2
BS≥60 & SH≥0.6
BS≥80 & SH≥0.8
BS≥95 & SH≥0.95
Colicin
Cpn60_TCP1
TFIIA
NLS
NES
0.2 sub/site
K)
0.2 sub/site
BS≥60 & SH≥0.6
BS≥80 & SH≥0.8
BS≥95 & SH≥0.95
0.46/10
0.42/17
0.71/19
0.84/50
0.66/27
0.46/10
-/17
0.84/20
0.48/14
0.73/10
0.82/45
0.25/4
0.27/7
0.11/4
0.82/34
0.77/2.00
0.19/15
Giardia intestinalis Hypothetical protein
Reticulomyxa filosa ETO26651
Reticulomyxa filosa ETO15592
Euplotes octocarinatus hypothetical protein
Saccharomyces cerevisiae Mitochondrial ornithine acetyltransferase
Acanthamoeba castellani hypothetical protein
Xenopus tropicalis transforming growth factor beta-induced 68kDa
Arabidopsis thaliana Protein kinase superfamily protein
Caenorhabditis elegans dct-14
Caenorhabditis elegans K07C11.10
Naegleria gruberi Protein kinase domain-containing protein
Malawimonas jakobiformis g4772.t1
Mastigamoeba balamuthi putative bbsome-interacting protein 1-like
Salpingoeca rosetta Uncharacterized protein
Monosiga brevicollis Predicted protein
Xenopus tropicalis predicted BBSome-interacting protein of 10 kDa-like
Danio rerio BBSome-interacting protein 1
Mus musculus BBSome-interacting protein 1
Homo sapiens BBSome-interacting protein 1
Reticulomyxa filosa ETO01502
Thecamonas trahens Uncharacterized protein
Trypanosoma brucei Bardet-Biedl syndrome 18 protein
Tetrahymena thermophila hypothetical protein
Drosophila melanogaster CG33506-PA
Drosophila melanogaster BBIP1-PA
ArgJ
BBIP10
DUF4476
ESCRT-II
Fasciclin
Fibrillin_U_N
Herpes_U44
Pkinase
Pkinase_Tyr
NLS
NES
BBS18

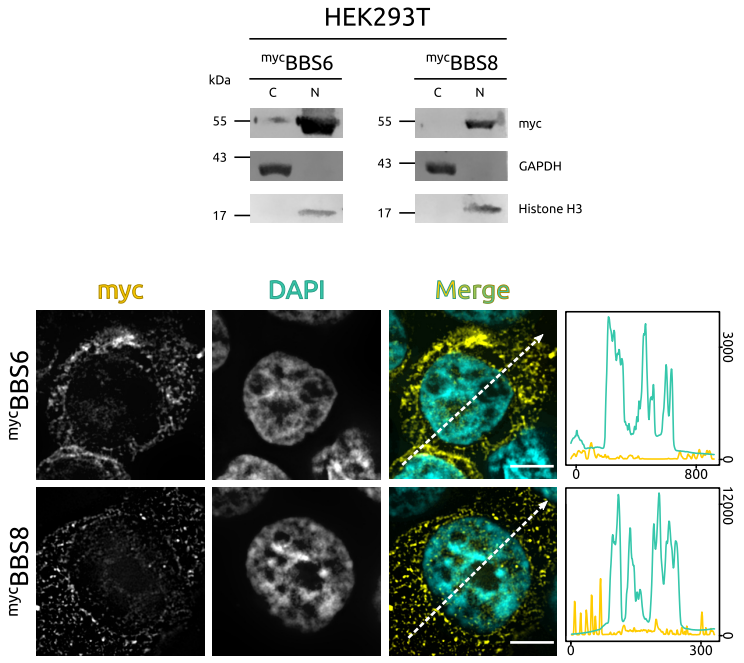
Supp.Fig. 2


Max.
NLS
1.0
0.5
0.0
single NLS
median NES /
mean NLS
mean BBS set
mean "cytosolic" set
mean all proteins
BBS1
BBS2
BBS4
BBS5
BBS7
BBS8
BBS9
BBS18
BBS6
BBS10
BBS12
BBS
"cyto"
all
n.s.
n.s.
n.s.
n.s.
n.s.
n.s.
Max.
NES
50
40
30
20
10
0
BBS1
BBS2
BBS7
BBS9
BBS4
BBS8
BBS5
BBS18
BBS6
BBS10
BBS12
BBS
"cyto"
all
n.s.
*
n.s.
n.s.
n.s.
n.s.

Supp. Fig. 3


A)
Max. NLS
Max. NES
1.0
0.5
0.0
50
40
30
20
10
0
un-ranked
Di-phoda
Opi-moda
0.124
Pvalue
1
BBS proteins
"cytosolic" proteins
all proteins
B)
Max. NLS
Max. NES
un-ranked
Archae-plastida
Excavata
Holozoa
SAR
0.029
Pvalue
1
BBS proteins
"cytosolic" proteins
all proteins
C)
Max. NLS
Max. NES
un-ranked
Alveo-lata
Amoebo-zoa
Crypto-phyta flagellates
Disco-cristates
Meta-monada
Meta-zoa
Rhizaria
Strameno-piles
Viridi-plantae
0.039
Pvalue
1
BBS proteins
"cytosolic" proteins
all proteins
single NLS
median NES / mean NLS
mean BBS set
mean "cytosolic" set
mean all proteins

Supp. Fig. 4


HEK293T
mycBBS6
mycBBS8
kDa
C
N
55
43
17
myc
GAPDH
Histone H3
myc
DAPI
Merge
mycBBS6
mycBBS8
3000
0
800
12000
300

Supp. Fig. 5

Fig. S1 (related to Fig. 1): Distribution of predicted orthologues in comparison to previous work. Presence or absence that was previously known is denoted as green circles, novel hits are denoted by purple circles.

Fig. S2 (related to Fig. 2): Maximum-likelihood phylogenetic trees for each BBS protein used in this work. Sequences included as outgroups are shown in grey. Bootstrap (BS) and Shimodaira-Hasegawa (SH) branch support values are given at each node (colour-coded nodes are darker when support is stronger). Protein domain architectures are based on similarity to models in Pfam-A v32 with E-value ≤ 10^{-3} (84). **A)** BBS1; **B)** BBS2; **C)** BBS4; **D)** BBS5; **E)** BBS6; **F)** BBS7; **G)** BBS8; **H)** BBBS9; **I)** BBS10; **J)** BBS12 tree with top three (orange) hits after reciprocal BLASTp search against the NCBI 'nr' database, and dictyostelial CCT homologues (green); **K)** BBS18.

Fig. S3 (realted to Fig. 3): Distribution of predicted nuclear signals in BBS orthologues and statistical analysis of their prevalence in orthologue groups. Proteins are grouped according to their complex and compared subsequently. Neither BBSome nor chaperonin-like complex have significantly higher NLS (left) or NES (right) signatures than proteins predicted to be in the cytosol, or all predicted proteins encoded in the organisms in the analysis, except for BBSome compared to cytosolic proteins (permutation test; ns: not significant, $0.05 \leq p$; *: $p < 0.05$).

Fig. S4 (related to Fig. 4): Clade-wise comparison of nuclear signal sequence scores for BBS proteins against other protein reference groups. Tests were carried out for BBS proteins of one higher taxonomic unit vs. BBS proteins, "cytosolic" or any proteins of any taxonomic unit. A), B) and C) show the comparisons at different taxonomic depths (permutation test).

Fig. S5 (related to Fig. 7): BBS nuclear localisation is independent of tag. Western blot after subcellular fractionation of transiently transfected HEK293T cells shows nuclear localisation of BBS6 and BBS8. Fusion constructs are expressed and detected at the expected sizes (BBS6 62 kDa; BBS8: 62 kDa). Successful fractionation was determined by presence of marker proteins for the cytosol (GAPDH) and the nucleus (Histone H3). Immunocytochemical stainings of transiently transfected HEK293T cells expressing myc-tagged BBS proteins show a similar nuclear localisation as cells expressing FLAG-tagged BBS6 and BBS8. Arrows mark the data points used for line plotting in direction of the arrow. Scale bar: 5 μm. Plot X-axis: length [pixels]; Plot Y-axis: pixel intensity [arbitrary units].

3. Analysis of human BBS protein homologues in insects support alternative non-ciliary functions

3.1. Preface

Cilia in insects are highly specialised and are only known to appear on two tissue types: neuronal tissue, and the flagellated sperm of males (Keil, 2012; Lattao et al., 2017). However, their assembly and maintenance mechanisms are largely homologous to other ciliated eukaryotes. As fewer cell types exhibit cilia than in other animals, insects might prove a helpful resource to search for non-ciliary functions of BBS proteins. Studies in the fruit fly *Drosophila melanogaster* as a model system have fundamentally shaped our understanding of developmental biology for decades (Jennings, 2011), but the species offers limited opportunity to search for non-ciliary BBS functions as it lacks the genes for two out of eight BBSome proteins (BBS2 and BBS7) (Ewerling et al., 2023b; Hodges et al., 2010). BBS1 however has been implicated in gene regulatory functions in Cape honey bees (Wallberg et al., 2016), so insects might warrant further research into this topic. In the following work, we compiled several insect proteomes from disparate taxa to search for BBS protein homologues to determine if other insects share a reduced set of BBS proteins. We found that *D. melanogaster* is an anomaly in the queried insect proteomes as most species showed clear homologues to all BBSome proteins. Sequence identity across species shows that some BBS proteins are more conserved than others, with BBS5 showing the highest inter-species variance and thus most potential for gain-of-function mutations during evolution. Further analyses of expression profiles of BBS homologues in the honey bee *Apis mellifera* showed that indeed, BBS proteins are most expressed in gonads and neuronal tissue, both of which are known to be ciliated. Most differentially expressed genes are BBS4, 5, 7, and 8, which all have been previously described as nuclear in other species (Ewerling et al., 2023b; Gascue et al., 2012; Horwitz and Birk, 2021). Expression of BBS5 and BBS8 is highest in the brain, and there are significant differences between male and female BBS5 and BBS8 expression on a tissue level. This hints at a role for these proteins beyond their function in the BBSome.and highlights both the importance of broad sampling for phylogenetic homologue searches as well as insects as a valuable resource when studying the interactome of BBS proteins.

The manuscript was written by Alexander Ewerling and Isa Graebling in a co-authorship, with discussion and editing contributions by Susanne Foitzik, Thomas J. Colgan, and Helen May-Simera. This part of the thesis was submitted to Royal Society Open Biology for peer review and is available on a preprint server (bioRxiv, https://doi.org/10.1101/2023.07.28.550953).

Analysis of human BBS protein homologues in insects support alternative non-ciliary functions

Alexander Ewerling[1*], Isa Graebling[2*], Anna Wierczeiko[2], Elisa Kotzurek[3], Susanne Foitzik[3], Susanne Gerber[2], Thomas J. Colgan[3], Helen May-Simera[1]

1 Institute of Molecular Physiology, Johannes Gutenberg University, Mainz, Germany
2 Institute of Human Genetics, University Medical Center, Mainz, Germany
3 Institute of Organismic and Molecular Evolution, Johannes Gutenberg University, Mainz, Germany
* co-authorship

ABSTRACT

Cilia and flagella were one of the characteristic traits of the last eukaryotic common ancestor and as such, are highly conserved among eukaryotes. Their proteomic makeup is consequently remarkably similar throughout all eukaryotic lineages. Recently, one subgroup of ciliary transport proteins in mammalian cells, the Bardet-Biedl Syndrome (BBS) proteins, was shown to have the ability to traverse the nuclear envelope, and to engage in protein-protein-interactions that modulate gene expression, signalling cascades, and cell homeostasis. Insects have been critically understudied in cilia biology because of their highly specialised cilia being localised on only a small subset of cell types. In this study, we present evidence that the BBSome, a hetero-octameric ciliary transport complex of BBS proteins, is largely conserved in multiple insect lineages. Using the honeybee *Apis mellifera* as a study system to explore BBS-associated gene expression, our analyses suggest that not all BBSome-associated genes are expressed equally, indicating possible non-ciliary functions. We also demonstrate that the expression of individual BBS proteins varies significantly between the tissues of queens and males in *A. mellifera*, especially in neuronal tissue. This result raises the question of what role BBS proteins play in these tissues and whether they are involved in gene regulation in insects. The potential gene regulatory function of BBS proteins should be explored in other eukaryotes due to their high degree of conservation.

INTRODUCTION

Cilia are tiny hair-like microtubule-based organelles extending from the surface of most eukaryotic cells. These ancient organelles have a conserved structure, function, and proteome across eukaryotes (1,2). They can be structurally divided into the basal body, which resides in the cytoplasm, the transition zone, the membrane-sheathed axoneme, and the ciliary tip (3) (Fig. 1A). While single-celled eukaryotes generally display motile cilia facilitating locomotion and sensory perception of the environment (2,4–6), immotile (or “primary“) cilia are mostly found as single copies per cell in metazoan cell types (7–9). Acting as a complex signalling centre, primary cilia are essential for several biological processes ranging from chemo- and mechanosensation to transduction of

numerous signalling cascades. Defects in cilia can cause multiple multisystemic diseases, referred to as ciliopathies, which show a wide variety of partly overlapping phenotypes (10–12). One of these diseases, Bardet-Biedl Syndrome (BBS) (13,14), represents a genetically heterogeneous inherited disorder that is considered the archetypical ciliopathy since patients exhibit virtually all symptoms associated with ciliary dysfunction. These include retinal degeneration, kidney disease, obesity, and mental retardation. BBS is caused by pathogenic mutations in genes encoding BBS proteins, which are all involved in facilitating ciliary trafficking which underlies ciliary signalling (15–18) (Fig. 1A).

Three main processes interact to make signalling pathways within cilia possible: trafficking of proteins to cilia from the cytoplasm, selective passage at the base of the cilium (the so-called 'transition zone'), and specific intraflagellar transport (IFT) along the ciliary axoneme (15,19,20). The BBSome, a hetero-octameric complex consisting of the BBS proteins BBS1, BBS2, BBS4, BBS5, BBS7, BBS8, BBS9 and BBS18 (17,18,21), is considered an adaptor that couples ciliary cargo proteins to motor proteins and IFT complexes. In addition to the transport of vesicles to cilia, the BBSome is required to promote the retrieval and export of specific transmembrane proteins from the cilium (16,19,22–24). Therefore, the BBSome plays an important role as a key regulator of cilia composition. A second characteristic group of BBS proteins, so-called chaperonin-like BBS proteins, is comprised of BBS6, BBS10 and BBS12, which show structural homology to the chaperonin-containing t-complex protein 1 (CCT) family of group II chaperonins (25). It was previously demonstrated that these proteins form a hetero-oligomeric complex with other CCT proteins, and that this process plays a key role in the regulation of BBSome assembly (26,27). Cilia structure and function are highly dependent on specific proteins localising to the cilium and given that BBS proteins are required for this, the function of cilia directly depends on their functionality. In addition to the important role BBS and other ciliary proteins play in cilia, they have also been described to function in other cellular processes, such as cell cycle regulation (28) as well as regulation of gene expression by interacting with transcription factors and RNA polymerases (29) (for reviews, see (30–32)). Besides their ciliary functions, the biochemistry and modes of action of BBS proteins outside of cilia are largely unknown.

One approach to further elucidate molecular mechanisms of ciliary proteins in non-ciliary functions is from an evolutionary perspective. Cilia are present in most eukaryotes; however, they were lost independently on multiple occasions in distantly related taxa (33). Despite the absence of a classical cilia, genes coding for putative BBS proteins were shown to still be present in some of the genomes of these organisms. A prime example is the parasite *Toxoplasma gondii* where cilia assembly occurs independently of the BBSome. A single putative BBS5 gene homologue is still actively expressed in a non-flagellate state of the parasite's life cycle (34), suggestive that the BBS-like protein does not perform cilia-associated functions and may have an alternative role. These non-ciliary functions may be crucial for an organism's gene regulation as interactions with transcription factors and RNA

polymerases have been identified (29). Consequently, these proteins might be essential beyond their ancestral ciliary functions. It is therefore of great interest to study the evolutionary patterns of putative BBS homologues and their expression in organisms that lack classical cilia. Indeed, insects possess those cilia in neurons and spermatozoa exclusively (35,36), and thus offer a unique opportunity to identify possible alternative functions of those genes through the determination of their expression outside of cells and tissues associated with the cilia.

As a preliminary step towards identifying possible alternative functions, we first determined whether there are homologues of human BBS proteins in representative members of diverse insect orders (Fig. 1B). More specifically, our study focussed on the BBS proteins that form part of the BBSome, as well as the genes encoding the chaperonin-like BBS proteins. The BBSome is ancestrally conserved in ciliated species and was predicted to be present in the last eukaryotic common ancestor (LECA) while the chaperonin-like BBS proteins are restricted to the unikont lineage (which includes animals and fungi) and the basal eukaryote *Malawimonas jakobiformis* (37), but are mostly found in metazoa. As mere presence or absence of protein homologues is a poor indicator of functionality, we also screened transcriptomic data of the well-studied honeybee *Apis mellifera* for expression of BBS proteins. Honeybees represent one of the most important pollinators, essential for the maintenance of agricultural yields and wildflower diversity (38). The genome was one of the earliest insect genomes to be sequenced (39) and assembled with the species having a wide range of transcriptomic and proteomic resources available (40,41). From an evolutionary perspective, honeybees are also part of the social Hymenoptera whereby caste differentiation has evolved within the female sex resulting in the expression of morphologically, behaviourally, and physiologically distinct phenotypes (42). Using publicly available datasets, we investigated sex- and tissue-specific expression differences to elucidate potential novel roles for BBS proteins in insects.

RESULTS AND DISCUSSION

Homology-based search reveals conservation of BBSome-associated proteins across insects

Our aim in this study was to investigate the occurrence and distribution of BBS proteins across highly divergent insect lineages. We started by studying the transcriptomes and genomes of different taxa and, in a second step, took a closer look at the proteomes of some insect species. To achieve this, we compiled a broad spectrum of predicted insect proteomes (n = 11 species across six orders; Fig. 1B). Divergence at the genome level is often so high that it is difficult to identify homologous genes with a high degree of certainty. However, since the protein sequence is much more conserved, protein-based analyses allow a closer look at the occurrence of homologous BBS proteins even in distant taxa.

In *Drosophila melanogaster*, cilia and flagella are only found on sperm cells and neurons (36). They still possess a functioning, relatively intact BBSome, albeit in a reduced form, with BBS2 and BBS7 missing from the otherwise complete complex found in other metazoa (37,43,44). We therefore investigated if a partial reduction (or even complete loss) is common across other insect taxa, or if the loss of single components identified in *D. melanogaster* is the onset of a taxonomically-restricted gradual loss of cilia. We screened the different insect species' genome and transcriptome assemblies to compare conservation on nucleotide and peptide levels. Generally, over large evolutionary distances, gene-based orthology searches are only a weak indicator of presence or absence of a given gene, as codon bias might heavily skew the analysis, whilst proteins are highly conserved on an amino acid level, but not necessarily on nucleotide level. When BLAST-searching insect genomes with human nucleotide seeds we found a high rate of false-positive, and that only strict E-value cut-offs result in true hits (Fig. 2A). The case is similar when searching for transcribed BBS genes: almost none of the BBS genes appear to be transcribed in queried species, apart from BBS8. In addition to its role in ciliary trafficking, BBS8 has been shown to play a crucial role in centrosome stability and cell division (45–47). This suggests that: 1) the putative non-ciliary functions of BBS8 are independent of tissue, as inferred transcripts cannot be confidently assigned to a specific tissue of a given species; and 2) when searching for non-ciliary functions of BBS proteins, tissue-dependent expression plays a pivotal role in detectability. Considering the first point, it is interesting to find BBS4, BBS5, and BBS7 homologues also being expressed in some insect species. Previous studies showed a high probability for the proteins to be localised to the nucleus at some point, and also showed a nuclear localisation for the human orthologues (29,37). This merits further investigation into where these proteins might be expressed, especially in respect to tissue- and sex-dependent differences. As these gene products are most commonly found in the transcriptomic datasets across different species, the underlying basis for this conserved expression might be a conserved, non-ciliary function.

To compare genome- and transcriptome-based searches with protein homologue searches, we compiled corresponding predicted proteomes from NCBI to conduct an all-versus-all BLAST-based homology search using OrthoFinder (48). We found that homologues of genes encoding BBS proteins that build the BBSome in humans (BBS1, BBS2, BBS7, BBS9, BBS4, BBS8, BBS5, BBS18) are present in almost all insect species, with the exception of BBS18 (Fig. 2B, Supp. Table S1). None of the queried species had homologues for chaperonin-like BBS proteins BBS6, BBS10 or BBS12. This lack of detection may be due to loss of genes in insects, gain of genes in mammals, or divergence of homologues beyond the point of detection. As cilia are built in highly specialised cell types in insects, there might not be a need for specialised CCT-derived BBS proteins in these cells to aid the folding of the BBSome. Overall, the high degree of conservation of the BBSome is reflected in most insect lineages, with *D. melanogaster* being a notable exception. BBS1-like proteins BBS2 and BBS7 were found to be missing in *D. melanogaster*, supporting previous studies that used

different approaches such as reciprocal best BLAST (RBB) and Hidden Markov Model (HMM) searches (37,43). Interestingly, the only BBSome protein missing in *A. mellifera* was BBS1. Given that BBS1 is missing while other BBS1-like proteins are conserved in honeybees, it is likely that the function of BBS1 is now performed to some extent by BBS2, BBS7, or BBS9, which is also hypothesised for the *D. melanogaster* BBSome missing BBS2 and BBS7 (37). BBS1, BBS2, BBS7, and BBS9 all have similar protein architecture (N-terminal β-propeller, followed by an α-helical linker and ɣ-adaptin-ear domain); BBS1 differs in that it is missing the platform- and C-terminal α-helix domains present in BBS2, BBS7 and BBS9. Ultrastructure of the mammalian BBSome was recently established (18), which suggests that BBS1 mediates the interaction between ARL6 and the BBsome needed for membrane attachment. Interaction is facilitated by the ɣ-adaptin-ear domain of BBS1 but that domain is also present in the other 'BBS1-like' proteins, so BBS1 itself may be substitutable in the BBSome of some species. This is supported by the low sequence similarity of honeybee BBS1 to human BBS1 (Fig. 3A), which could result from relaxed purifying selection or positive selection acting in one or both species driving divergence.

Overall, there was significant variation in terms of percentage identity shared between specific human BBS proteins and their insect homologues (min: 35% (BBS9); max: 54.8% (BBS5); Fig. 3A,B), which may reflect differences in terms of evolutionary conservation. While there is no single group of insects standing out in terms of conservation, there is a trend for individual BBS proteins to be generally more conserved than others. For example, the lowest sequence similarity to human BBS proteins is seen in BBS1 and BBS9, and highest in BBS5 (Fig. 3A). BBS8 has the most variation in sequence similarity of all BBSome proteins across insects (Fig. 3B). Low sequence similarities could reflect reduced constraints on functionality derived from specific amino acids, leaving room for evolutionary divergence and acquisition of novel functions. However, non-conserved structures may be amenable to loss, as could be the case for BBS1 in honeybees. The strikingly high sequence conservation of BBS5 on the other hand hints at potentially conserved functions, possibly outside cilia.

BBS protein expression is different in tissues of queens, workers and males

To understand variation in gene expression profiles of BBS insect homologues, we examined differential gene expression of putative BBS homologues across sexes and tissues of *A. mellifera* (Supp. Table S2). Given that honeybee queens and males – referred to as drones – differ extensively in terms of physiology, morphology, and behaviour, yet share the same genome, such observed phenotypic differences are regulated via differential gene expression. Using datasets consisting of somatic and germline tissues from honeybee queens and males (49,50), we identified the expression of BBS4, BBS5, BBS7, and BBS8 homologues in *A. mellifera*, with BBS4 being the most highly expressed overall (Fig. 4A). While in terms of overall expression there were no significant differences between the sexes (Wilcoxon test, $p > 0.05$; Fig. 4B), we found significantly higher expression of

both BBS5 and BBS8 in the brain compared to gonad (Wilcoxon test, $p_{BBS5,\ brain\text{-}gonad}$ = 0.0001, $p_{BBS8,\ brain\text{-}gonad}$ = 0.0451; Fig. 4C). This might reflect the high degree of specialisation insect cilia exhibit: in *D. melanogaster*, cilia are only expressed in (sensory) neurons and sperm cells. The fact that not all BBS genes are expressed higher in these tissues in honeybees is however intriguing and might hint at a more specialised role for these genes and their associated proteins, independent of cilia. We further compared the tissues of queens and drones (both within the same sex and versus the other sex; Fig. 4D) finding that BBS5 and BBS8 are significantly higher expressed in drone brains compared to drone gonads (Wilcoxon test, $p_{BBS5,\ drone,\ brain\text{-}gonad}$ = $p_{BBS8,\ drone,\ brain\text{-}gonad}$ = 0.0317; Benjamini-Hochberg corrected). The case is the same for BBS8 expressed in queen brains versus queen gonads (Wilcoxon test, $p_{BBS8,\ queen,\ brain\text{-}gonad}$ = 0.0357; Benjamini-Hochberg corrected). When comparing expression in the brain, BBS8 is significantly lower expressed in queens compared to drones (Wilcoxon test, $p_{BBS8,\ brain,\ queen\text{-}drone}$ = 0.0357; Benjamini-Hochberg corrected), and BBS5 is significantly higher expressed in queen gonads compared to drone gonads (Wilcoxon test, $p_{BBS8,\ gonad,\ queen\text{-}drone}$ = 0.0357; Benjamini-Hochberg corrected). BBS5 and BBS8 both localise to mammalian nuclei (37) and both proteins have been shown to interact with E3-ubiquitin-protein ligase RING2 (RNF2) (29), a protein of the polycomb group (PcG) repressor complex 1 (PRC1). This complex is mainly responsible for histone 2A (H2A) monoubiquitylation (51,52), leading to repression of proteins crucial during development, such as those encoded by homeobox genes (53,54). Given the high degree of sequence conservation, this could also be the case in honeybees.

An alternative, larger tissue dataset consisting of different tissues from honeybee workers, and here especially from nurses and foragers, shows a similar picture: BBS4 is expressed quite frequently (and relatively strongly) in many tissue types, especially in glandular tissues and antennae (Fig. 4E). The consistently high expression of BBS4 in all honeybee glands studied, including the hypopharyngeal and mandibular glands in the head of the worker bees and the Nasonov and sting glands in their abdomen, suggests a specific function of this gene in glandular physiology, and possibly in chemical communication. Interestingly, the other BBS proteins are also relatively highly expressed in the antennae. Along with the eyes, the antennae are the most important sensory organs of the honeybee and have functions in chemosensory, tactile, temperature and acoustic perception. Type I and II cilia have been described for the Johnston organ of the honeybee antenna pedicel (55), and cilia have also been found in the odour-detecting sensilla placodea at the tip of the antennae (56). Higher expression of BBS proteins in antennae could therefore be due to their cilia function. To fully understand the possibly conserved role of BBS proteins in gene regulation, further studies of the potential target genes are required.

CONCLUSION

The findings of this analysis predominantly validate earlier studies regarding the existence of BBS homologues in insects, such as *A. gambiae, A. mellifera,* and *D. melanogaster (43,44)*. While former

studies focused on the evolution of molecular intraflagellar transport (44) and the evolutionary history of the centriole from protein components (43), our study investigated the evolutionary conservation and expression of BBS protein homologues across insect species. We were able to show that proteins forming the BBSome are potentially more conserved across insects than previously thought. Our study reveals that limiting homologue searches to few model organisms does not necessarily reflect the *status quo* across a larger set of species from the same family and might therefore skew conclusions drawn. In previous studies, homologues of all BBSome genes (with exception of BBS2 and BBS7) were described at the genomic level in *A. mellifera* and *D. melanogaster* (37,43,44). Interestingly, *A. mellifera* lacks a putative BBS1 homologue using our search criteria, despite being found in another study (43). This is seemingly a contradiction to previous work conducted on a Cape honeybee population (57), where the authors concluded LOC102655146, a locus annotated as BBS1-protein like, is under selective pressure for the social parasitism phenotype. Given that the protein architecture of BBS1 and BBS2 (also BBS7 and BBS9) are similar, it could be that Wallberg *et al.* erroneously 'identified' a BBS1-like protein as BBS1. It is still interesting to conduct further analyses towards the identified gene to shed light on possible altered protein interactions upon loss of one essential BBSome constituent.

The comparison between genomic/transcriptomic searches and proteomic approaches clearly shows that with seeds from distantly related species (human vs. insects in this case), orthology searches that focus on conservation of amino acid sequences are clearly superior, both in sensitivity and sensibility. Upon BLAST analysis, sequences of human BBS proteins were found both in the genome and in the transcriptome in each insect. The predicted expression of BBS8 in all but one queried species suggests that many insects transcribe BBS proteins, despite the absence of classical cilia in most tissues. This leads to the question which alternative function this BBS protein might perform in insects.

It is exciting to find a theoretically fully functional BBSome in insects despite their deviation from the 'classical' function of (primary) cilia throughout different tissues seen in other animals. Finding proteins differentially expressed between tissues in honeybees indicates that there are probably cilia-independent functions. The differential expression of BBSome proteins in ciliated tissues, such as the brains of queens and drones, suggests that these proteins may have additional functions beyond their role in ciliary transport.

FIGURE LEGENDS

Fig.1: Ciliary functions are conserved in eukaryotes. A) BBS proteins can be found at the ciliary basal body where the BBSome is assembled with chaperonin-like BBS proteins in concert with CCT/TRiC proteins. Ciliary cargo proteins are then transported along the axoneme with the help of IFT particles and motor proteins (adapted from (18,37,58)). B) Although highly specialised, cilia can be found in insect neurons and sperm flagella. For orthologue search we assembled proteomes from several insect clades, and human and mouse as reference/outgroup.

Fig.2: Evolutionary conservation of core BBSome genes across insects. A) Proteomic search has uncovered orthologues to most human BBSome proteins. BBS18 and chaperonin-like proteins BBS6, BBS10 and BBS12 are however restricted to humans in this analysis. B) While genome searches generally reflect more the proteomic searches, they are easily hitting false-positives (hits for almost all chaperonin-like BBS proteins).

Fig. 3: Variation in sequence similarity across insect BBS proteins. A) Scatterplots for each identified BBS protein across insect genomes displaying percentage identity at the protein level between human BBS proteins (species 1) and putative homologues in insects (n = 11) and mouse (species 2). Each species is represented by an individual colour. B) Boxplot displaying the distribution of percentage identify at the protein level for each BBS protein across all insect species examined (n = 11 species).

Fig.4: Expression differences of expressed honeybee BBS homologues from queens and drones, and across different tissues. A) Variance-stabilising transformed (VST) gene level counts of BBS proteins in somatic and germline tissues of honeybee queens and males. B) Expression profiles of BBS proteins in the brains and gonads of honeybee queens and drones. C) Expression profiles of BBS proteins across different tissues (Wilcoxon test, *: $p <= 0.05$; ***: $p <= 0.001$). D) Expression profiles of BBS proteins across tissues of different sexes (Wilcoxon test, *: $p <= 0.05$). E) VST gene level counts of BBS proteins in somatic tissues of honeybee foragers and nurse. Q: queen, D: drone; B: brain, G: gonad.

MATERIALS AND METHODS

Data procurement

The data used for BLAST analyses of transcriptomes and genomes was obtained from the National Center for Biotechnology Information nucleotide database. For each insect, the representative genome and transcriptome GenBank assemblies were used for this purpose. A list of the GenBank assembly accessions can be found in Supp. Table S3. The human BBS accession IDs used as a seed can be found in Supp. Table S4.

BLAST analysis and homologue search

In terms of BLAST analysis (59) a hit is defined by a detected significant similarity between the sequences compared with respect to the chosen parameter values. Hits can differ in aspects like position, length, e-value and score. In this study the e-value was chosen as representation value to classify the significance and display hits. The BLASTn analyses were carried out via the NCBI BLAST online tool (general parameters: expect threshold = 0.05, word size = 11; scoring parameters: match/mismatch scores= 2,-3, gap costs = existence: 5 extension: 2).

To determine homology of BBS-associated proteins across diverse insect orders, predicted RefSeq proteomes were downloaded from NCBI for the following species: Diptera, *Drosophila melanogaster*, *Anopheles gambiae*; Hymenoptera, *Apis mellifera*, *Bombus terrestris*, *Nasonia vitripennis*, *Solenopsis invicta*, *Ooceraea biroi*; Lepidoptera, *Manduca sexta*; Hemiptera, *Acyrthosiphon pisum*; Phthiraptera, *Pediculus humanus*; Coleoptera, *Tribolium castaneum*; and Mammalia, *Homo sapiens* and *Mus musculus*. From each, the longest predicted peptide was extracted using the OrthoFinder (v.2.5.4) script primary_transcript.py (48). OrthoFinder was subsequently used (default settings) to perform pairwise alignments between each species allowing for the identification and generation of orthogroups, which are groups consisting of potential homologues across species of interest. Such orthogroups were parsed for the determination of insect homologues of described human BBS-associated genes.

Differential gene expression analysis

For the purposes of examining gene expression of putative BBS-associated homologues, we examined expression profiles in somatic and germline tissues of the Western honeybee, *A. mellifera (60)*, BioProject ID: PRJNA243651; (49), BioProject ID: PRJNA386859; (50), BioProject ID: PRJNA689223). We downloaded publicly available transcriptomic datasets from the NCBI Short Read Archive using the sra-toolkit. For each library, we extracted the data in FASTQ format and performed quality assessments using FastQC v.0.11.9 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Data were filtered using fastp v.0.23.2 (61) to remove low quality reads and trim adaptors. We then pseudoaligned each sample against an indexed predicted transcriptome using kallisto v.0.44.0 (62) providing a transcript-level quantification

of expression. Using these estimates, we generated gene-level counts using tximport v1.26.1 (63) and for each dataset, generated variance-stabilised transformed data using DESeq2 v1.38.3 (64).

Statistical testing

Datasets used for statistical testing can be found in Supp. Table S2. Tests were performed in R v.4.0.3 (65), with packages tidyverse (66) (including dplyr (67), ggplot2 (68) and stringr (69)), reshape2 (70), and FSA (71). The accompanying RMarkdown notebook is available figshare (https://doi.org/10.6084/m9.figshare.23780871). Datasets were tested for normality by Shapiro-Wilk testing (where normality was assumed if $p > 0.05$) and depending on outcome, different tests were used to compare subsets. Comparisons were done by Wilcoxon rank-sum test ($p > 0.05$: not significant; $0.05 >= p > 0.01$: *; $0.01 >= p > 0.001$: **; $0.001 >= p > 0.0001$: ***). Multiple testing correction was done using the Benjamini-Hochberg procedure (72).

Additional software

Figures were prepared with ggpubr (73) and Inkscape (74).

ACKNOWLEDGEMENTS

This project was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – GRK2526/1 – Projectnr. 407023052.

REFERENCES

1. Pazour GJ, Agrin N, Leszyk J, Witman GB. Proteomic analysis of a eukaryotic cilium. J Cell Biol. 2005 Jul 4;170(1):103–13.

2. Mitchell DR. The Evolution of Eukaryotic Cilia and Flagella as Motile and Sensory Organelles. In: Eukaryotic Membranes and Cytoskeleton: Origins and Evolution [Internet]. New York, NY: Springer; 2007 [cited 2021 Feb 5]. p. 130–40. (Advances in Experimental Medicine and Biology). Available from: https://doi.org/10.1007/978-0-387-74021-8_11

3. Ounjai P, Kim KD, Liu H, Dong M, Tauscher AN, Witkowska HE, et al. Architectural Insights into a Ciliary Partition. Current Biology. 2013 Feb 18;23(4):339–44.

4. Bloodgood RA. Flagella-dependent gliding motility inChlamydomonas. Protoplasma. 1981 Sep 1;106(3):183–92.

5. Saito A, Suetomo Y, Arikawa M, Omura G, Mostafa Kamal Khan S m., Kakuta S, et al. Gliding movement in Peranema trichophorum is powered by flagellar surface motility. Cell Motility. 2003;55(4):244–53.

6. Cavalier-Smith T, Lewis R, Chao EE, Oates B, Bass D. Helkesimastix marina n. sp. (Cercozoa: Sainouroidea superfam. n.) a Gliding Zooflagellate of Novel Ultrastructure and Unusual Ciliary Behaviour. Protist. 2009 Aug 1;160(3):452–79.

7. Bloodgood RA. Chapter 1 - From Central to Rudimentary to Primary: The History of an Underappreciated Organelle Whose Time Has Come.The Primary Cilium. In: Sloboda RD, editor. Methods in Cell Biology [Internet]. Academic Press; 2009 [cited 2019 Aug 19]. p. 2–52.

(Primary Cilia; vol. 94). Available from: http://www.sciencedirect.com/science/article/pii/S0091679X08940012

8. Bloodgood RA. Sensory reception is an attribute of both primary cilia and motile cilia. Journal of Cell Science. 2010 Feb 15;123(4):505–9.

9. Anvarian Z, Mykytyn K, Mukhopadhyay S, Pedersen LB, Christensen ST. Cellular signalling by primary cilia in development, organ function and disease. Nature Reviews Nephrology. 2019 Apr;15(4):199–219.

10. Afzelius BA. Cilia-related diseases. The Journal of Pathology. 2004;204(4):470–7.

11. Badano JL, Mitsuma N, Beales PL, Katsanis N. The Ciliopathies: An Emerging Class of Human Genetic Disorders. Annual Review of Genomics and Human Genetics. 2006;7(1):125–48.

12. Fliegauf M, Benzing T, Omran H. When cilia go bad: cilia defects and ciliopathies. Nature Reviews Molecular Cell Biology. 2007 Nov;8(11):880–93.

13. Bardet G. Sur un syndrome d'obesite congenitale avec polydactylie et retinite pigmentaire (contribution a l'etude des former cliniques de l'obesite hypophysaire). Amedee Le Grand [Internet]. 1920 [cited 2019 Aug 6];470. Available from: https://ci.nii.ac.jp/naid/20001129230/

14. Biedl A. Ein Geschwister mit adiposogenitaler Dystrophie. Deutsche medicinische Wochenschrift. 1922;48:1630.

15. Nakayama K, Katoh Y. Ciliary protein trafficking mediated by IFT and BBSome complexes with the aid of kinesin-2 and dynein-2 motors. J Biochem. 2018 Mar 1;163(3):155–64.

16. Wingfield JL, Lechtreck KF, Lorentzen E. Trafficking of ciliary membrane proteins by the intraflagellar transport/BBSome machinery. Wakefield JG, Moores CA, editors. Essays in Biochemistry. 2018 Oct 4;62(6):753–63.

17. Singh SK, Gui M, Koh F, Yip MC, Brown A. Structure and activation mechanism of the BBSome membrane protein trafficking complex. Carter AP, Kuriyan J, Lorentzen E, Kikkawa M, editors. eLife. 2020 Jan 15;9:e53322.

18. Klink BU, Gatsogiannis C, Hofnagel O, Wittinghofer A, Raunser S. Structure of the human BBSome core complex. Carter AP, Kuriyan J, Kikkawa M, editors. eLife. 2020 Jan 17;9:e53910.

19. Lechtreck KF, Johnson EC, Sakai T, Cochran D, Ballif BA, Rush J, et al. The Chlamydomonas reinhardtii BBSome is an IFT cargo required for export of specific signaling proteins from flagella. Journal of Cell Biology. 2009 Dec 28;187(7):1117–32.

20. Lechtreck KF. IFT–Cargo Interactions and Protein Transport in Cilia. Trends in Biochemical Sciences. 2015 Dec 1;40(12):765–78.

21. Nachury MV, Loktev AV, Zhang Q, Westlake CJ, Peränen J, Merdes A, et al. A Core Complex of BBS Proteins Cooperates with the GTPase Rab8 to Promote Ciliary Membrane Biogenesis. Cell. 2007 Jun 15;129(6):1201–13.

22. Lechtreck KF, Brown JM, Sampaio JL, Craft JM, Shevchenko A, Evans JE, et al. Cycling of the signaling protein phospholipase D through cilia requires the BBSome only for the export phase. Journal of Cell Biology. 2013 Apr 15;201(2):249–61.

23. Jin H, White SR, Shida T, Schulz S, Aguiar M, Gygi SP, et al. The Conserved Bardet-Biedl Syndrome Proteins Assemble a Coat that Traffics Membrane Proteins to Cilia. Cell. 2010 Jun 25;141(7):1208–19.

24. Nozaki S, Katoh Y, Kobayashi T, Nakayama K. BBS1 is involved in retrograde trafficking of ciliary GPCRs in the context of the BBSome complex. PLOS ONE. 2018 Mar 28;13(3):e0195005.

25. Kim JC, Ou YY, Badano JL, Esmail MA, Leitch CC, Fiedrich E, et al. MKKS/BBS6, a divergent chaperonin-like protein linked to the obesity disorder Bardet-Biedl syndrome, is a novel centrosomal component required for cytokinesis. Journal of Cell Science. 2005 Mar 1;118(5):1007–20.

26. Seo S, Baye LM, Schulz NP, Beck JS, Zhang Q, Slusarski DC, et al. BBS6, BBS10, and BBS12 form a complex with CCT/TRiC family chaperonins and mediate BBSome assembly. PNAS. 2010 Jan 26;107(4):1488–93.

27. Álvarez-Satta M, Castro-Sánchez S, Valverde D. Bardet-Biedl Syndrome as a Chaperonopathy: Dissecting the Major Role of Chaperonin-Like BBS Proteins (BBS6-BBS10-BBS12). Front Mol Biosci [Internet]. 2017 [cited 2019 Aug 27];4. Available from: https://www.frontiersin.org/articles/10.3389/fmolb.2017.00055/full

28. Kim JC, Badano JL, Sibold S, Esmail MA, Hill J, Hoskins BE, et al. The Bardet-Biedl protein BBS4 targets cargo to the pericentriolar region and is required for microtubule anchoring and cell cycle progression. Nature Genetics. 2004 May;36(5):462.

29. Gascue C, Tan PL, Cardenas-Rodriguez M, Libisch G, Fernandez-Calero T, Liu YP, et al. Direct role of Bardet–Biedl syndrome proteins in transcriptional regulation. J Cell Sci. 2012 Jan 15;125(2):362–75.

30. McClure-Begley TD, Klymkowsky MW. Nuclear roles for cilia-associated proteins. Cilia. 2017 May 25;6(1):8.

31. Hua K, Ferland RJ. Primary cilia proteins: ciliary and extraciliary sites and functions. Cell Mol Life Sci. 2018 May;75(9):1521–40.

32. Doornbos C, Roepman R. Moonlighting of mitotic regulators in cilium disassembly. Cell Mol Life Sci [Internet]. 2021 Apr 15 [cited 2021 Jun 10]; Available from: https://doi.org/10.1007/s00018-021-03827-5

33. Hodges ME, Wickstead B, Gull K, Langdale JA. Conservation of ciliary proteins in plants with no cilia. BMC Plant Biology. 2011 Dec 30;11(1):185.

34. Ramakrishnan C, Maier S, Walker RA, Rehrauer H, Joekel DE, Winiger RR, et al. An experimental genetically attenuated live vaccine to prevent transmission of Toxoplasma gondii by cats. Sci Rep. 2019 Feb 6;9(1):1474.

35. Keil TA. Sensory cilia in arthropods. Arthropod Structure & Development. 2012 Nov 1;41(6):515–34.

36. Lattao R, Kovács L, Glover DM. The Centrioles,Centrosomes, Basal Bodies, and Cilia of *Drosophila melanogaster*. Genetics. 2017 May 1;206(1):33–53.

37. Ewerling A, Maissl V, Wickstead B, May-Simera HL. Neofunctionalization of ciliary BBS proteins to nuclear roles is likely a frequent innovation across eukaryotes. iScience. 2023 Apr 21;26(4):106410.

38. Southwick EE, Southwick L Jr. Estimating the Economic Value of Honey Bees (Hymenoptera: Apidae) as Agricultural Pollinators in the United States. Journal of Economic Entomology. 1992 Jun 1;85(3):621–33.

39. Weinstock GM, Robinson GE, Gibbs RA, Weinstock GM, Weinstock GM, Robinson GE, et al. Insights into social insects from the genome of the honeybee *Apis mellifera*. Nature. 2006 Oct;443(7114):931–49.

40. Trapp J, McAfee A, Foster LJ. Genomics, transcriptomics and proteomics: enabling insights into social evolution and disease challenges for managed and wild bees. Mol Ecol. 2017 Feb;26(3):718–39.

41. Grozinger CM, Zayed A. Improving bee health through genomics. Nat Rev Genet. 2020 May;21(5):277–91.

42. Colgan TJ, Carolan JC, Bridgett SJ, Sumner S, Blaxter ML, Brown MJ. Polyphenism in social insects: insights from a transcriptome-wide analysis of gene expression in the life stages of the key pollinator, *Bombus terrestris*. BMC Genomics. 2011 Dec 20;12(1):623.

43. Hodges ME, Scheumann N, Wickstead B, Langdale JA, Gull K. Reconstructing the evolutionary history of the centriole from protein components. J Cell Sci. 2010 May 1;123(9):1407–13.

44. van Dam TJP, Townsend MJ, Turk M, Schlessinger A, Sali A, Field MC, et al. Evolution of modular intraflagellar transport from a coatomer-like progenitor. PNAS. 2013 Apr 23;110(17):6943–8.

45. Ansley SJ, Badano JL, Blacque OE, Hill J, Hoskins BE, Leitch CC, et al. Basal body dysfunction is a likely cause of pleiotropic Bardet–Biedl syndrome. Nature. 2003 Oct;425(6958):628–33.

46. Blacque OE, Leroux MR. Bardet-Biedl syndrome: an emerging pathomechanism of intracellular transport. Cell Mol Life Sci. 2006 Sep 1;63(18):2145–61.

47. May-Simera HL, Kai M, Hernandez V, Osborn DPS, Tada M, Beales PL. Bbs8, together with the planar cell polarity protein Vangl2, is required to establish left–right asymmetry in zebrafish. Developmental Biology. 2010 Sep 15;345(2):215–25.

48. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biology. 2019 Nov 14;20(1):238.

49. Lewis SH, Quarles KA, Yang Y, Tanguy M, Frézal L, Smith SA, et al. Pan-arthropod analysis reveals somatic piRNAs as an ancestral defence against transposable elements. Nat Ecol Evol. 2018 Jan;2(1):174–81.

50. Slater GP, Dapper AL, Harpur BA. Haploid and Sexual Selection Shape the Rate of Evolution of Genes across the Honey Bee (*Apis mellifera L.*) Genome. Genome Biol Evol. 2022 May 31;14(6):evac063.

51. Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, et al. Role of histone H2A ubiquitination in Polycomb silencing. Nature. 2004 Oct 14;431(7010):873–8.

52. Schwartz YB, Pirrotta V. Polycomb silencing mechanisms and the management of genomic programmes. Nat Rev Genet. 2007 Jan;8(1):9–22.

53. Cao R, Tsukada YI, Zhang Y. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Mol Cell. 2005 Dec 22;20(6):845–54.

54. Taherbhoy AM, Huang OW, Cochran AG. BMI1-RING1B is an autoinhibited RING E3 ubiquitin ligase. Nat Commun. 2015 Jul 7;6:7621.

55. Tsujiuchi S, Sivan-Loukianova E, Eberl DF, Kitagawa Y, Kadowaki T. Dynamic Range Compression in the Honey Bee Auditory System toward Waggle Dance Sounds. PLOS ONE. 2007 Feb 21;2(2):e234.

56. Slifer EH, Sekhon SS. The fine structure of the plate organs on the antenna of the honey bee, *Apis mellifera linnaeus.* Experimental Cell Research. 1960 Mar 1;19(2):410–4.

57. Wallberg A, Pirk CW, Allsopp MH, Webster MT. Identification of Multiple Loci Associated with Social Parasitism in Honeybees. PLOS Genetics. 2016 Jun 9;12(6):e1006097.

58. Taschner M, Lorentzen E. The Intraflagellar Transport Machinery. Cold Spring Harb Perspect Biol. 2016 Oct;8(10):a028092.

59. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990 Oct 5;215(3):403–10.

60. Jasper WC, Linksvayer TA, Atallah J, Friedman D, Chiu JC, Johnson BR. Large-Scale Coding Sequence Change Underlies the Evolution of Postdevelopmental Novelty in Honey Bees. Molecular Biology and Evolution. 2015 Feb 1;32(2):334–46.

61. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018 Sep 1;34(17):i884–90.

62. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol. 2016 May;34(5):525–7.

63. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res. 2015;4:1521.

64. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

65. R Core Team. R: A language and environment for statistical computing. [Internet]. Vienna, Austria: R Foundation for Statistical Computing; Available from: https://www.R-project.org/

66. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the Tidyverse. Journal of Open Source Software. 2019 Nov 21;4(43):1686.

67. Wickham H, François R, Henry L, Müller K. dplyr: A Grammar of Data Manipulation [Internet]. 2021. Available from: https://dplyr.tidyverse.org, https://github.com/tidyverse/dplyr

68. Wickham H. ggplot2: Elegant Graphics for Data Analysis [Internet]. Springer-Verlag New York; 2016. Available from: https://ggplot2.tidyverse.org

69. Wickham H. stringr: Simple, Consistent Wrappers for Common String Operations [Internet]. 2022. Available from: https://stringr.tidyverse.org, https://github.com/tidyverse/stringr

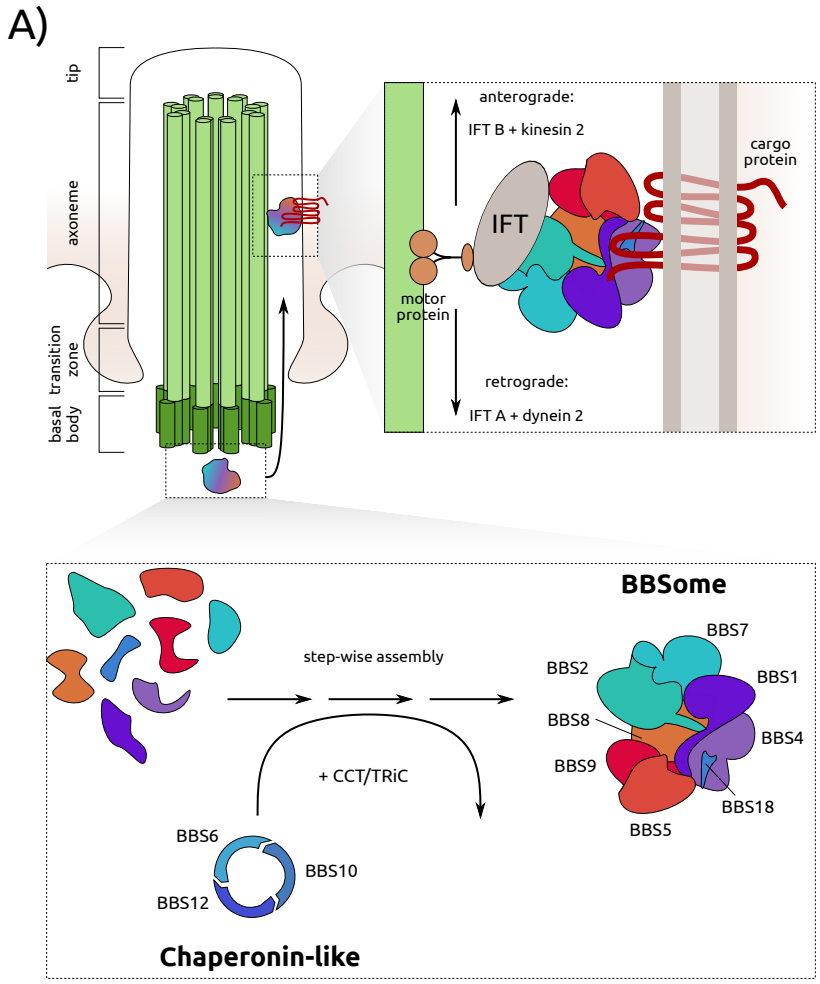
70. Wickham H. Reshaping Data with the reshape Package. Journal of Statistical Software. 2007 Nov 13;21:1–20.

71. Ogle DH, Doll JC, Wheeler AP, Dinno A. FSA: Simple Fisheries Stock Assessment Method [Internet]. 2023. Available from: https://fishr-core-team.github.io/FSA/

72. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological). 1995;57(1):289–300.

73. Kassambara A. ggpubr: 'ggplot2' Based Publication Ready Plots [Internet]. 2023. Available from: https://rpkgs.datanovia.com/ggpubr/

74. Inkscape Project. Inkscape [Internet]. 2020. Available from: https://inkscape.org



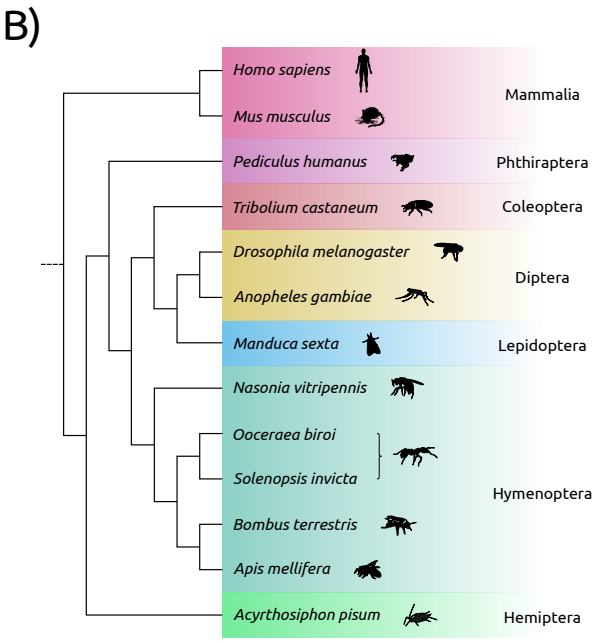
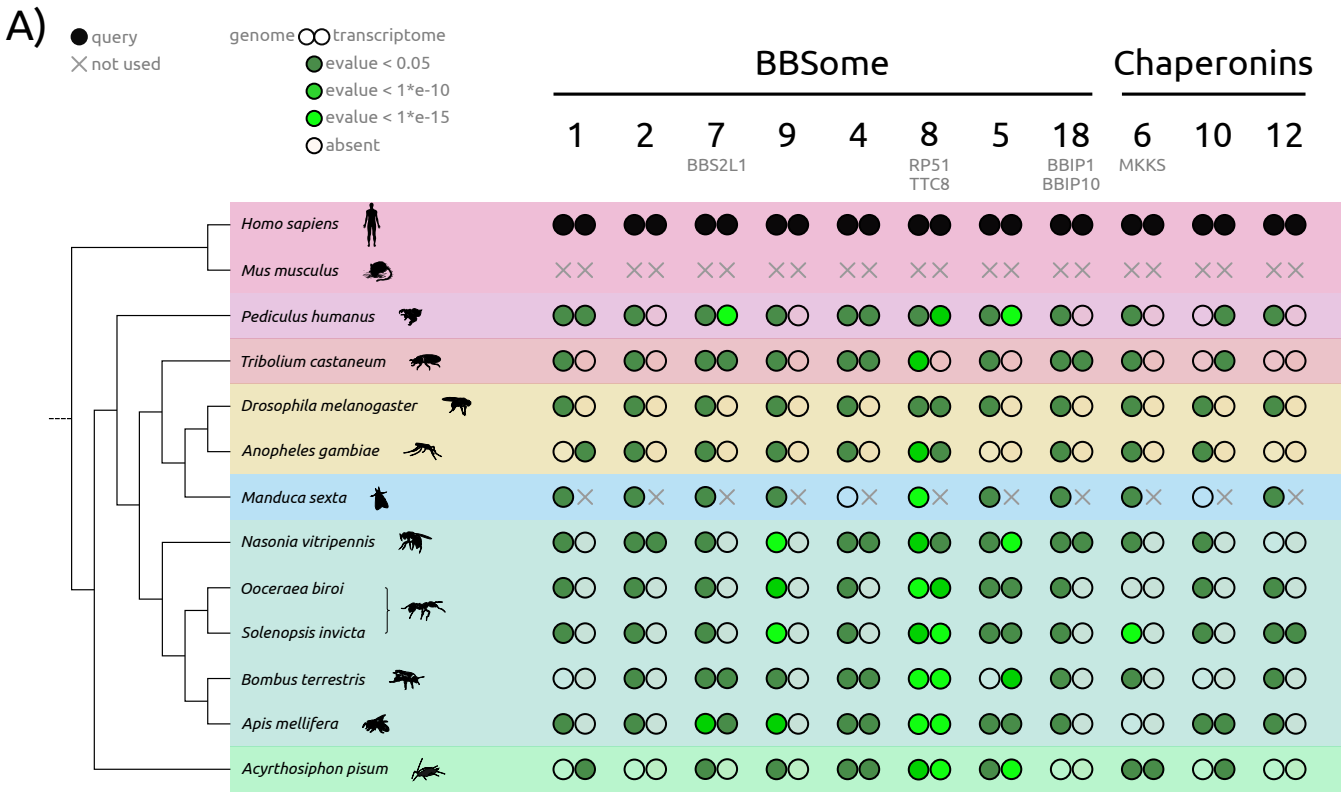


Fig 1



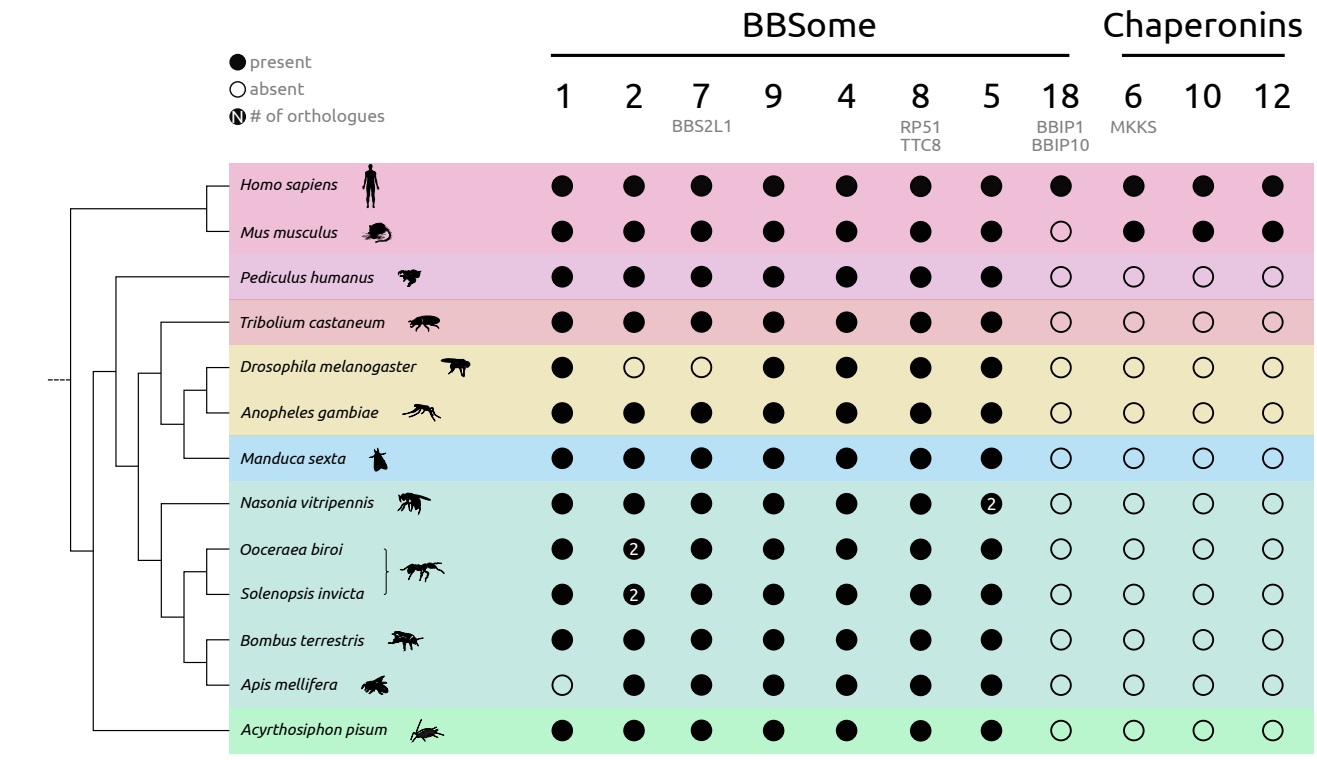
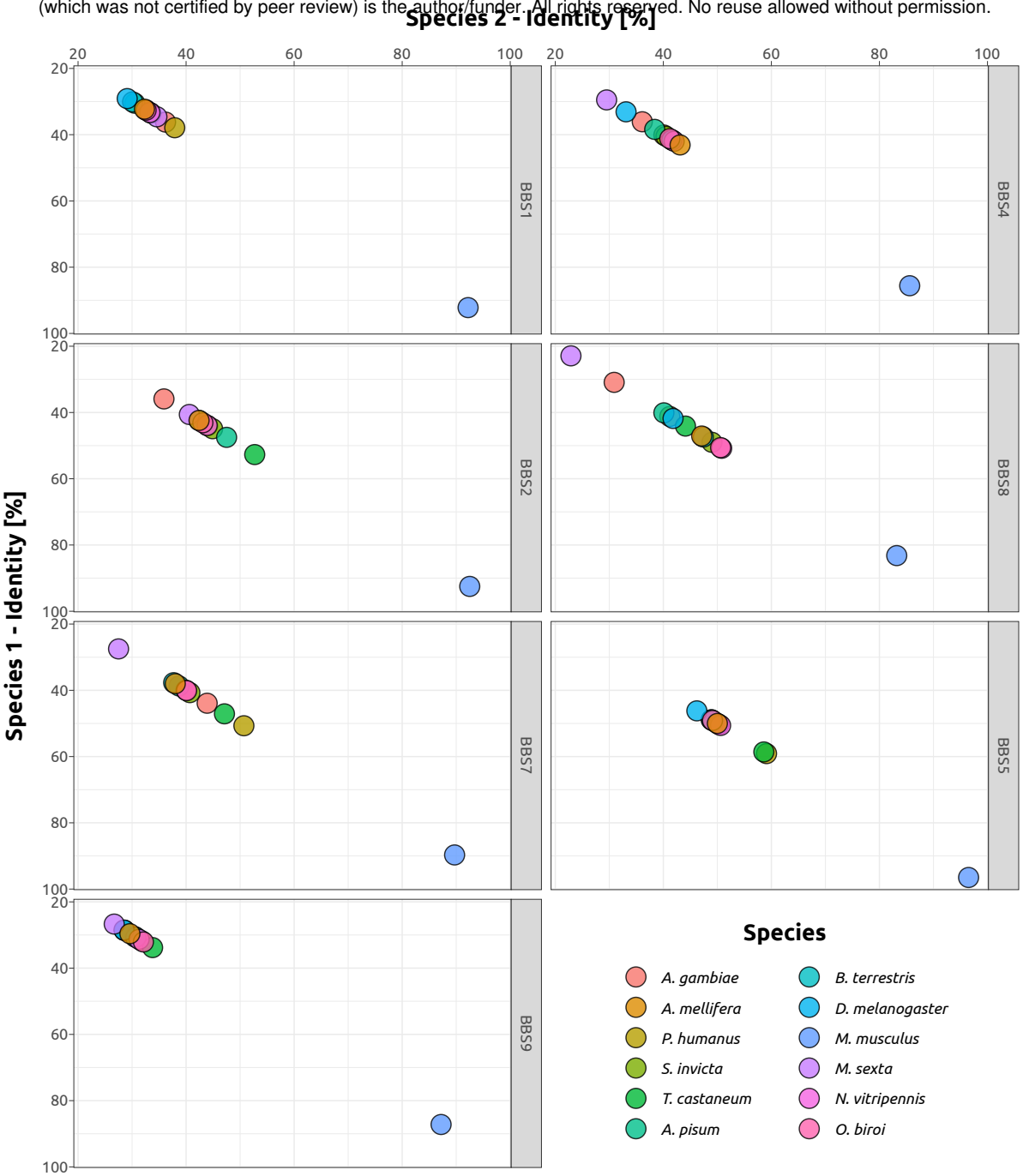


Fig 2

A)



B)

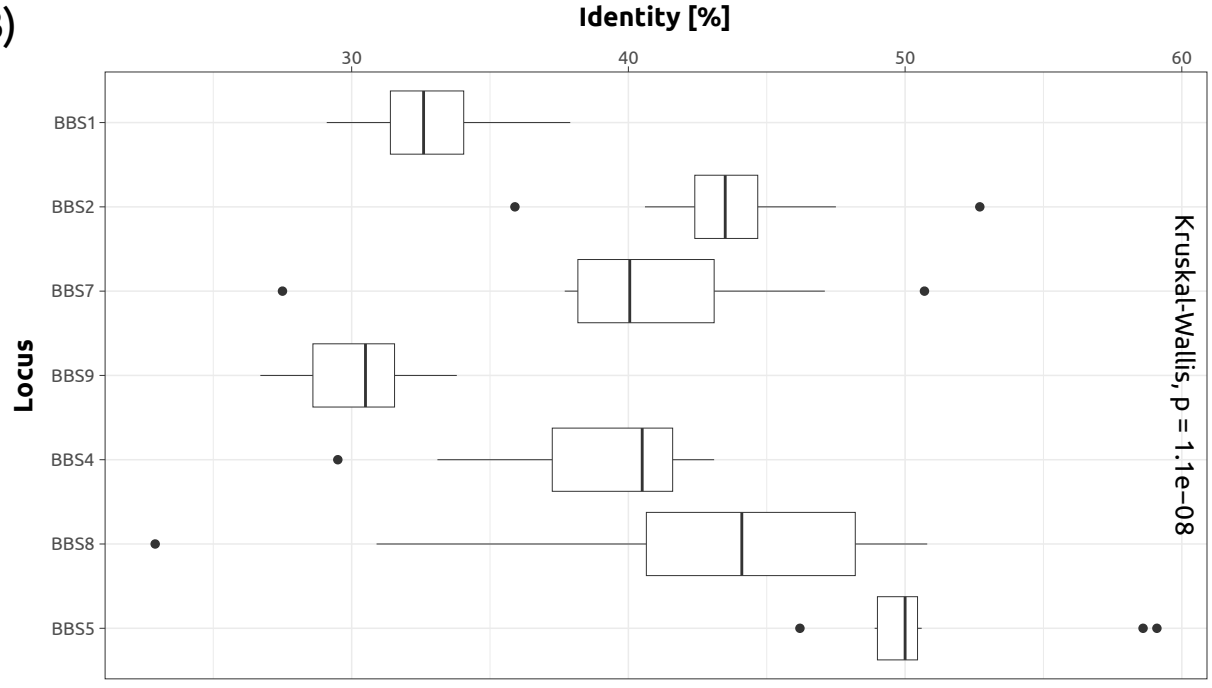


Fig 3

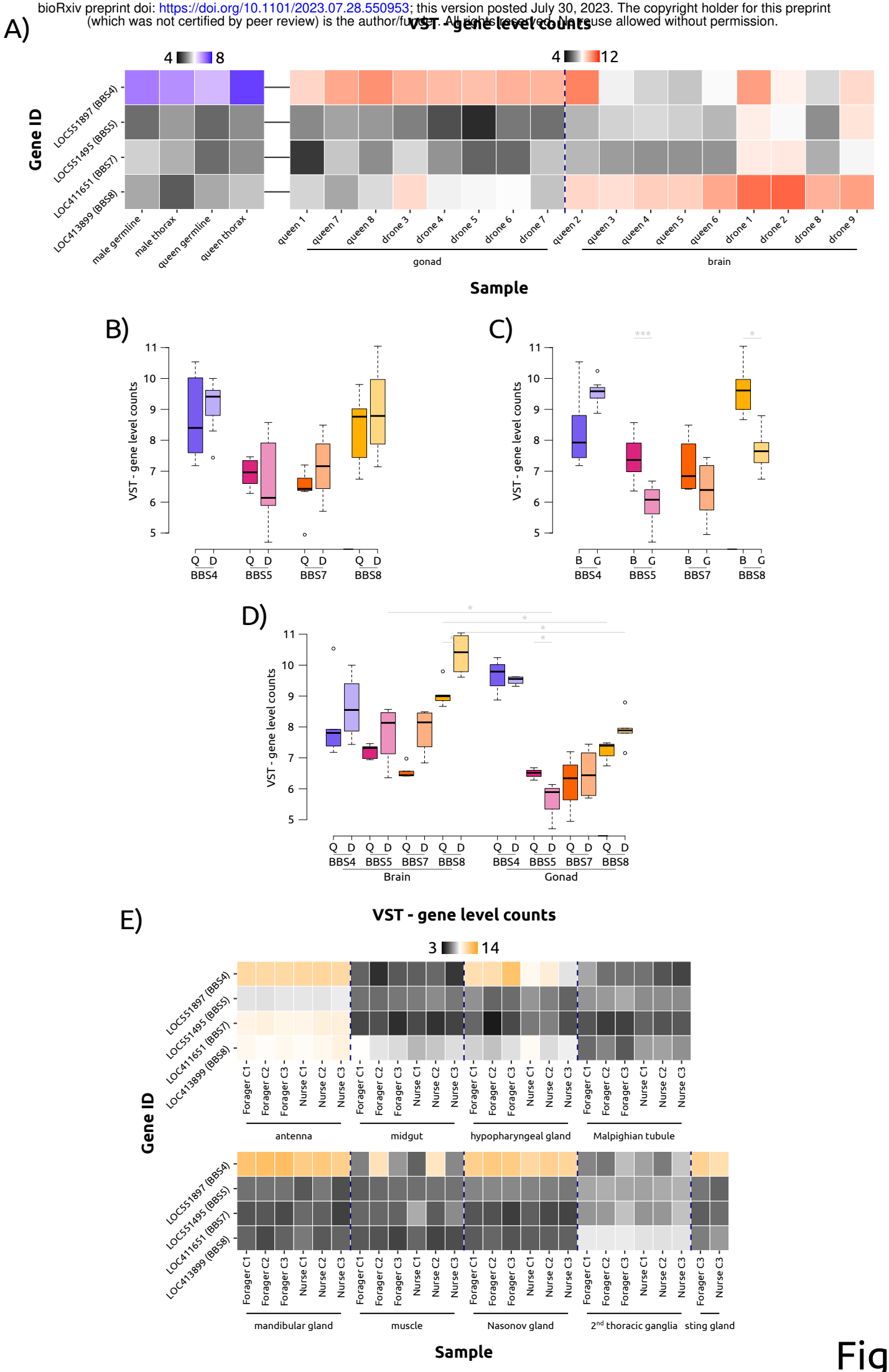


Fig 4

4. EVOLUTIONARY TRAJECTORY FOR NUCLEAR FUNCTIONS OF CILIARY PROTEINS

4.1. PREFACE

This part of the thesis is written in the style of a review with a concluding hypothesis. The topic of the manuscript is the evolution of ciliary trafficking complex proteins in the light of early and pre-eukaryotic evolution. Although it is widely accepted that LECA was a highly complex organism (Koumandou et al., 2013) with extensively expanded families of RabGTPases (Klinger et al., 2016), membrane coats (Dacks and Robinson, 2017), a nucleus with NPCs (Field and Rout, 2019), and at least one cilium (Carvalho-Santos et al., 2011), it is far more contentious in which order these organelles arose after the endosymbiotic event that marked the beginning of eukaryotic evolution. In this work, we review the latest findings regarding a subset of organelles that likely arose through paralogous expansion of an ancient membrane coating complex: the protocoatomer. This complex nucleated many extant membrane coats, namely the type I and type II families, and the IFT/BBSome complexes. The NPC is an amalgamation of different type I and type II coats and therefore shares an evolutionary relationship with IFT/BBSome proteins. Their shared history may have implications for the observed nuclear phenotypes of otherwise ciliary proteins from the IFT and BBSome complexes: In the transition between FECA and LECA where individual organelles were not yet established, there was ample room for co-evolution of primordial NPCs and IFT/BBSome complexes that may have led to the incorporation of ‘ciliary’ proteins into the nucleus. With increasing complexity of regulatory networks, some IFT/BBSome proteins might have lost the ability to enter the nucleus due to the presence of proteins that could transduce signals into the nucleus. Some may have retained their initial nuclear functions and were therefore selected for such purposes. A highly interesting case is the apicomplexan parasite *Toxoplasma gondii* that lost the canonical pathway of cilia assembly. It is missing most of the BBSome but BBS5 as the single remaining BBS protein of that complex (Ewerling et al., 2023b), and RNAseq data confirms expression during a non-ciliated stage during its life cycle (Ramakrishnan et al., 2019). In this review, we give an explanation for how this could be common for other ciliary proteins. We conclude with an exciting perspective where cilia and nucleus co-evolved, and where ciliary proteins had dual functions as nuclear interactors.

The manuscript was written by Alexander Ewerling with discussion and editing contributions by Helen May-Simera. The manuscript has yet to be submitted for peer review.

Evolutionary trajectory for nuclear functions of Bardet-Biedl Syndrome proteins

Alexander Ewerling[1], Helen Louise May-Simera[1]

[1]: Institute of Molecular Physiology, Faculty of Biology, Johannes Gutenberg-University, Mainz, Germany

Abstract

Cilia and the nucleus were two defining features of the last eukaryotic common ancestor. In early eukaryotic evolution, these structures evolved through diversification of a common membrane-coating ancestor, the protocoatomer. While in cilia, the descendants of this protein complex evolved into parts of the intraflagellar transport complexes IFT and BBSome, the nucleus gained its selectivity by recruiting protocoatomer-like proteins to the nuclear envelope to form the selective nuclear pore complexes. Recent studies show a growing number of proteins shared between the proteomes of the respective organelles, and it is currently unknown how ciliary proteins could acquire nuclear functions, and vice versa. Nuclear functions of ciliary proteins are still observable today and remain relevant for the understanding of disease mechanisms behind ciliopathies. In this work, we review the evolutionary history of cilia and nucleus and their respective defining proteins, and integrate current knowledge into theories for early eukaryotic evolution. We postulate a scenario where both compartments co-evolved and that fits current models of eukaryotic evolution, explaining how ciliary proteins and nucleoporins acquired their dual functions.

Introduction

The last eukaryotic common ancestor (LECA) was a remarkably complex organism as inferred from phylogenetic reconstruction. The consensus is that it was a flagellated (Carvalho-Santos *et al.*, 2011; Lindemann, 2022), sexually dividing cell (Villeneuve and Hillers, 2001; Loidl, 2016) unable to phagocytose (Bremer *et al.*, 2022), with mitochondria (Bremer *et al.*, 2022), an endoplasmic reticulum (ER) (Kontou *et al.*, 2022) and the namesake of this kingdom of life, a nucleus (Mans *et al.*, 2004; Bapteste *et al.*, 2005; Neumann *et al.*, 2010). It is still hotly debated in which order the previously mentioned organelles arose, and if they (especially the nucleus) came to be autogenously (Cavalier-Smith, 1987; Jékely, 2003) or via a symbiotic relationship between the Asgardarchaeal eukaryotic ancestor and an engulfed α-proteobacterium (Gould *et al.*, 2016) or a myxobacterium and euryarchaeote (López-García and Moreira, 2006). However LECA has become more and more tangible through better modelling and phylogenetic inference in recent years.

To accommodate this plethora of membrane-bounded organelles and ensure their function, the means to facilitate directional and unambiguous transport must have already been established. The

paralogous expansion of ancestral coat protein families likely shaped the organellar identity (organellar paralogy model (Dacks and Field, 2007)) (Cavalier-Smith, 2002; Dacks and Doolittle, 2002; Jékely, 2003; Araç *et al.*, 2005; Yoshizawa *et al.*, 2006; Sanderfoot, 2007), thereby giving rise to the complexity we observe in many lineages today. This provided a platform for lineage-specific expansion or reduction of the trafficking scaffold (e.g. rhoptry-specific RONs in *Toxoplasma gondii* (Alexander *et al.*, 2005), or caveolin and flotillin in metazoa for clathrin-independent endocytosis (Field *et al.*, 2007)).

One of the defining features of LECA was the cilium (Koumandou *et al.*, 2013). The eukaryotic cilium is a highly specialised signalling 'organelle' rich in receptors and able to move and facilitate locomotion (Mitchell, 2007) (Fig. 1). The term 'organelle' has been used unanimously in the field but is misleading in that cilia are not compartmentalised through a continuos membrane, but rather extend from the plasma membrane into the extracellular space and are shielded by a specialised compartment at the base that allows sorting and regulated entry and exit of ciliary molecules (Hu *et al.*, 2010). While in extant eukaryotes we see many different forms of cilia, be they immotile primary cilia specific to Metaozoa or the abundance of motile cilia (more often called flagella) in almost all eukaryotes, they all likely trace back to the cilium found in LECA. The LECA cilium likely fulfilled the same functions as cilia we see today and utilised the same molecular machinery for assembly, disassembly and maintenance as extant eukaryotes, namely the intraflagellar transport system, short IFT. The IFT machinery consists of three distinct complexes (IFT-A, IFT-B and the BBSome), all of which trace back to a common ancestral complex (van Dam *et al.*, 2013). This 'proto-IFT' is itself derived from a hypothetical protocoatomer complex that gave rise to paralogous membrane coating families that shaped the organelles found in modern eukaryotes. Although lineage-specific adaptations occurred during evolution of the complexes involved in IFT (e.g. cytosolic assembly and disassembly of cilia in *Batrachochytrium dendrobatidis* and *Plasmodium falciparum* (Killick-Kendrick and Peters, 1978; Venard *et al.*, 2020) which lack the BBSome), these complexes are among the most conserved proteins derived from the protocoatomer. While the complexes are mostly known for their involvement in cilia function and dysfunction, studies show that single proteins, but not the entire complexes, play a role in different processes such as cell cycle progression and cytokinesis (Kim *et al.*, 2005; Robert *et al.*, 2007; Delaval *et al.*, 2011; Vitre *et al.*, 2020), establishment of the immune synapse (Finetti *et al.*, 2009; Cassioli *et al.*, 2021), transcription factor translocation (Scott *et al.*, 2017; Horwitz and Birk, 2021), and DNA damage signalling (den Dulk *et al.*, 2008; Chaki *et al.*, 2012; Choi *et al.*, 2013), with the latter two drawing a direct connection between cilia and nucleus (also reviewed in (McClure-Begley and Klymkowsky, 2017)). It has indeed been shown that several IFT and BBSome proteins can localise to the nucleus (Gascue *et al.*, 2012; Scott *et al.*, 2017; Shi *et al.*, 2018; Ewerling *et al.*, 2023), and that they fulfil

potential roles in gene regulation (Gascue *et al.*, 2012; Scott *et al.*, 2017), although the reason as to why they do this and how common this is among eukaryotic clades remains unclear.

In this review, we aim to provide an insight in the proteomic makeup of LECA from a membrane coat point of view, how the membrane-bounded organelles arose, and how they had opportunity to interact during eukaryotic evolution. Ultimately, we establish a possible evolutionary trajectory for how ciliary proteins acquired nuclear functions, and if currently observed nuclear functions could be ancestrally conserved from LECA.

Proteomic makeup of LECA endomembrane sorting systems

The ability to deform membranes is not exclusive to eukaryotes and can be observed in multiple cases of prokaryotic life. Bending membranes is essential for cell division regardless of kingdom. Some prokaryotes even form intricate membrane invaginations that need to be stabilised by proteins to maintain their structure. In fact, bacterial organisms with stacked membrane invaginations were first considered as ancestors of eukaryotes, as these structures resemble endomembrane organelles (Neumann *et al.*, 2014) and, in some cases, even engulf parts of the genetic information (Fuerst and Sagulenko, 2011). These are however analogous structures that arose independently in prokaryotes, and do not represent ancestors of eukaryotic endomembrane systems (McInerney *et al.*, 2011). Although it remains a matter of debate which organelles arose first and by what means, phylogenetic reconstruction of LECA paints the picture of a surprisingly complex cell equipped with the hallmark organelles of extant eukaryotes, and whose proteomic diversity is potentially on par with modern eukaryotes (Koumandou *et al.*, 2013). To face the challenges that come from hosting multiple organelles, LECA had an established set of distinct trafficking pathways to and from organelles and the plasma membrane that provided specialised cargo delivery and directionality of transport (Fig. 2). Since there are already excellent studies and reviews encompassing the plethora of complexes that were likely established in LECA together with their regulatory protein families and tethering complexes (Field *et al.*, 2007; Dacks and Field, 2007; DeGrasse *et al.*, 2009; Elias *et al.*, 2012; Rout and Field, 2017; Field and Rout, 2019), we focus here on the evolution of the ones most closely related to the eukaryotic cilium: the type I, type II, nucleoporin, and IFT family of membrane coats.

Protocoatomer-derived vesicle coats

Membrane coats differentiate vesicles and proteins depending on their origin and destination and were an essential tool to sort the highly complex endomembrane machinery in LECA. By the time different eukaryotic lineages expanded, LECA most likely had a complex system of endomembrane organelles and subcellular compartments established that warranted sorting of cargo (Field *et al.*,

2007; Dacks and Field, 2007; Koumandou *et al.*, 2007, 2013; Elias *et al.*, 2012). Three families stand out by their connection through an ancestral protein complex: type I, type II, and IFT/BBSome coating proteins. They all trace back to a so-called 'proto-coatomer' protein complex that was likely established before LECA (Devos *et al.*, 2006; O'Reilly *et al.*, 2011; Sampathkumar *et al.*, 2013; van Dam *et al.*, 2013; Stuwe *et al.*, 2014). Although they diverge substantially in sequence, their α-solenoid/β-propeller structure is highly conserved and hints at their shared evolutionary history. This group of protocoatomer-derived proteins comprises of several subfamilies: the type I and type II coats, the nucleoporins (Nups) building the nuclear pore complex (NPC), and the ciliary trafficking proteins of IFT and the BBSome, all of which shape the identity of their respective compartments/organelles.

The type I family of coats consists of coatomer protein I (COPI), five adaptins (AP-1 – 5), and clathrin. COPI coats vesicles in retrograde Golgi-to-ER and intra-Golgi transport (Duden, 2003; Lee *et al.*, 2004; Beck *et al.*, 2009; Spang, 2009), with additional roles in endosomal transport, lipid homeostasis, mRNA transport, and NE breakdown (Guo *et al.*, 1994; Aniento *et al.*, 1996; Gu *et al.*, 1997; Daro *et al.*, 1997; Liu *et al.*, 2003; Trautwein *et al.*, 2004; Prunuske *et al.*, 2006; Gabriely *et al.*, 2007; Bi *et al.*, 2007; Beller *et al.*, 2008; Razi *et al.*, 2009; Soni *et al.*, 2009; Todd *et al.*, 2013; Thiam *et al.*, 2013; Wilfling *et al.*, 2014; Zabezhinsky *et al.*, 2016). Given their versatility in different trafficking pathways, it is likely they were present early in eukaryotic evolution and adapted to multiple functions. AP-1 and AP-2 in concert with clathrin build vesicle coats that destine their cargo from the plasma membrane or the trans-Golgi network to early/recycling endosomes. The remaining AP-3 – 5 are involved in sorting of vesicles for the late endosome/lysosome, or have currently undefined functions (Edeling *et al.*, 2006; Paczkowski *et al.*, 2015). Clathrin coats vesicles budding from the plasma membrane, and vesicles destined to the *trans*-Golgi network and endosomes (Vigers *et al.*, 1986).

Type II family coats are COPII and HOPS (homotypic fusion and vacuole protein sorting)/CORVET (class C core vacuole-endosome tethering) proteins (Koumandou *et al.*, 2007). COPII proteins facilitate the anterograde transport of vesicles from the ER towards Golgi, while the HOPS/CORVET coats are also involved in early to late endosomal sorting, and late endosome to lysosome/vacuole transport.

The Nups that build the NPC share structural features of and are related to both type I and type II coats (Devos *et al.*, 2004; Brohawn *et al.*, 2008), making the NPC a chimera and likely a late addition to the organellar landscape of the developing eukaryote. They form distinct complexes of eightfold symmetry in the nuclear envelope that allow selective diffusion of proteins and molecules through the NPC. Through duplication, Nups diversified before LECA, which led to formation of

subcomplexes within the NPC: the cytosolic face with filaments attached to a ring, which itself is connected and anchored via inner rings and membrane rings to the nucleosolic complex with nuclear rings and the nuclear basket (Makarov *et al.*, 2021).

The last group of protocoatomer-derived membrane coats present in LECA does not actually coat a membrane. The IFT/BBSome complexes share a common ancestor that is related to COPIε (van Dam *et al.*, 2013) and form multiprotein complexes that are needed for ciliary trafficking. They link motor proteins to their cargo and aid in the entry and exit of ciliary cargo. In contrast to other protocoatomer family proteins, IFT/BBSome proteins share considerable sequence identity across different eukaryotes, which reflects their highly specialised function in ciliary trafficking. Recent studies could show that they not only localise to the cilium or the ciliary base, but also to the nucleus of some species (Gascue *et al.*, 2012; Scott *et al.*, 2017; Shi *et al.*, 2018; Horwitz and Birk, 2021; Ewerling *et al.*, 2023). Nuclear functions could be shown for some of the IFT/BBSome proteins, but not all. This raises the question of how it could be that ciliary proteins are able to enter the nucleus, and whether this phenomenon can be traced back to the evolutionary history of nucleus and cilia. For this, we will highlight the possible order of emergence for each of the two organelles and infer potential co-evolution scenarios that could explain the duality of ciliary proteins functioning in the nucleus of eukaryotic cells.

Cilia

Cilia and flagella are an eukaryotic invention that was present in LECA (Koumandou *et al.*, 2013). Across the eukaryotic Tree of Life, cilia can be found in all major lineages but with frequent secondary losses (Pazour, 2004; Hodges *et al.*, 2010; Carvalho-Santos *et al.*, 2011). The purpose of cilia is largely homologous, serving both sensory and locomotive functions (Satir *et al.*, 2008; Bloodgood, 2010). Cilia are anchored to the cell via the basal body, a barrel-like structure made from nine symmetrically arranged microtubule (MT) triplets that serves as a nucleation point for the ciliary axoneme (Fig. 1). The basal body itself is highly conserved as is evident from phylogenetic reconstruction of its components (Hodges *et al.*, 2010). The axoneme that is the centre scaffold of the extending cilium is itself a structure made of MT doublets with an additional central pair in the case of motile cilia. As the structure and function of cilia and the basal body are conserved, so are the proteomes, with crucial components essentially identical in most ciliated species (Wickstead and Gull, 2007; Hodges *et al.*, 2011). The A- and B-tubules of each axonemal doublet are the ‘highway’ for all ciliary trafficking that is needed to assemble, disassemble, and maintain the cilium in its architectural and functional integrity. On these tracks (specifically the B-tubule), highly conserved, specialised intraflagellar transport (IFT) complexes IFT-A, IFT-B, and the BBSome, are moved by kinesin-2 family (anterograde) or dynein 1b/cytoplasmic dynein 2 family (retrograde) motor

proteins that connect to ciliary cargo proteins (e.g. transmembrane receptors) via the BBSome functioning as an adaptor complex (reviewed in Taschner and Lorentzen, 2016; Prevo *et al.*, 2017). This highly conserved machinery is likely derived from a diversification of the 'protocoatomer' during the shaping of the organellar landscape in the transition from first to last eukaryotic common ancestor (van Dam *et al.*, 2013). How did the building blocks for cilia arise before LECA? Their high degree of conservation calls for an early adoption of the flagellar machinery that is directly linked to the possible nuclear functions of ciliary proteins that can be appreciated today.

Origin of eukaryotic cilia

As for other eukaryotic organelles, there are a number of theories about how cilia came to be. Three major schools of thought have emerged since the 1960s, but only two of them persist to this day. One of the first hypotheses was formulated by Lynn Margulis (later Sagan), alongside her theory of sequential endosymbiosis that tried to explain the origin of other organelles like mitochondria. She hypothesised that the cilium is (like the mitochondrion) the result of an endosymbiotic event between an archaebacterium and, in this case, a bacterial spirochaete (Sagan, 1967). The theory is however poorly supported by phylogenetic relations of eukaryotic ciliary proteins and spirochaete bacterial proteins, and was therefore abandoned. Another theory postulated a viral origin, where the virus contained information about precursors of tubulin, and where the capsid served as a template for the basal body/ciliary axoneme (Satir *et al.*, 2007). In this scenario, the cell would have supplied the motors and IFT precursors to the budding virus (its capsid serving as basal body), resulting in an elongating axoneme. However, there is also little supporting data for this theory. A third hypothesis was first introduced in 1974 and describes an autogenous origin of both cilia and basal body from a primitive microtubule organising centre (MTOC) (Pickett-Heaps, 1974). This hypothesis is favoured by many researchers as it does not necessitate a fortuitous union of two prokaryotic cells where one of them would lose its outer membrane and complete genetic fingerprint (as would be the case in Sagan's theory) and also allows a gradual evolution of the cytoskeletal and MT network from prokaryotic precursors rather than a sudden genetic addition of a viral entity.

If we assume the theory of an autogenous origin for cilia and basal bodies to be true, there is a scenario where the primitive microtubular network co-evolved with ancestors of ciliary trafficking proteins and motor proteins. Initially, the IFT precursors and motors could have facilitated directed transport of receptors and vesicles to and from the plasma membrane (as was already hypothesised by Jékely and Arendt, 2006) and by paralogous expansion, begin to focus on one specific route, potentially with specific cargo affinity for receptors and receptor-based signalling components. The membrane at the destination would then be enriched in signalling molecules, providing significant advantage for sensory perception over random receptor distribution over the whole cell surface, up

to the point where receptors would inhibit each other sterically at the membrane patch subdomain (Berg and Purcell, 1977). Consequentially, once this membrane patch is oversaturated, selective pressure would favour a cell surface architecture that further allows enhancement of signal capture. Cilia provide this advantage: Mathematical modelling of ciliary ligand capture rates concludes that an immotile cilium has the same capture rate constant as a membrane patch 3.8 times the ciliary surface area (Hickey *et al.*, 2021). In shear flow, the cilia even enhance the capture rate sixfold. This theory provides a plausible evolutionary scenario for the arisal of a directed transport facilitated by emerging families of trafficking proteins and molecular motors. It could also help explain how the cilium improved the survival chances of early eukaryote ancestors, where even incremental elongations of the ciliary axoneme would provide an advantage compared to non-ciliated cells.

Given that there are many advantages for ciliated organisms from a biophysical point of view, it is necessary to trace back the components that make cilia possible to their evolutionary origins to substantiate the basis of the theory. Here, we concentrate on the main scaffolding protein of cilia, tubulin, and the IFT complexes IFT-A, IFT-B, and the BBSome, the latter three still without confident placement in their evolutionary context. However, recent advances in phylogenetic reconstructions of the complexes allows for a more refined view on the evolution of cilia.

Evolution of tubulin and MT-dependent trafficking

There is indeed evidence of a distant relative to tubulin in prokaryotes. The prokaryotic *FtsZ* gene codes for a protein needed for chromosome segregation (Löwe and Amos, 2009) that is still needed for plastid division in plants and algae (Kuroiwa *et al.*, 1998; Pyke, 1999). Its loose phylogenetic similarity led to it being hypothesised as ancestor of eukaryotic tubulin (Faguy and Doolittle, 1998). Physiologically, the structure of the protein and its dynamics are conserved: polymerisation of both FtsZ and tubulin results in longitudinal extension of fibres, with a dynamically assembling and disassembling. However, a crucial difference between the two is the ability to interact with other proteins: prokaryotic FtsZ does not show any interaction with accessory factors as does eukaryotic tubulin (Michie and Löwe, 2006). In eukaryotes, tubulin is subjected to many post-translational modifications and interacts with a plethora of other cytosolic proteins and cofactors (Erickson, 2007). With regards to cilia, one group of these clearly stands out: molecular motor proteins, such as kinesin and dynein that are needed for transport along the MT network within the cell, but also along the axoneme of cilia and flagella. Both motor protein families are essential for ciliary function and integrity as they associate with the IFT particles to facilitate cargo delivery to, from, and within the cilium. From studies on the evolutionary relationships of different dynein subclasses, the above theory of a polarised membrane patch prior to the establishment of the cilium is plausible as IFT-related dyneins were the first family to diverge from cytoplasmic dyneins and therefore

theoretically able to transport IFT precursors before the ciliary dyneins had evolved (Hartman and Smith, 2009). Cilia-specific dyneins further provide the means for motile cilia. We can infer from phylogenetic reconstructions of different subclasses of dyneins that LECA had multiple paralogous dynein subclasses, meaning that they were already expanding pre-LECA (Wickstead and Gull, 2007).

The establishment of a sophisticated ciliary transport system however not only requires motors and tracks but also a means to attach cargoes onto the motors. Eukaryotic cells implement different IFT complexes to facilitate ciliary cargo delivery to their destination. The evolutionary history of IFT complexes IFT-A, IFT-B and the BBSome does not only shed light on the FECA-to-LECA transition, but also how parts of these subcomplexes may have acquired additional functions in an increasingly complex pre-eukaryotic cell.

Evolution of IFT

Intraflagellar transport is the process of cargo delivery from the base of the cilium through a selective barrier along the axoneme to the tip, and back down and out of the cilium again. Eukaryotic cells established three main complexes to facilitate this movement: IFT-A, IFT-B and the BBSome (Rosenbaum and Witman, 2002; Taschner and Lorentzen, 2016; Prevo *et al.*, 2017; Nakayama and Katoh, 2018). IFT-A and IFT-B directly interact with kinesin-2 family (anterograde) and cytoplasmic dynein 2 family (retrograde) proteins for locomotion of pre-assembled IFT trains. These trains are assembled at the ciliary base and loaded with various cargo proteins, such as G-protein coupled receptors (Qin *et al.*, 2005; Huang *et al.*, 2007; Williams *et al.*, 2014), motility-related outer dynein arm proteins (Piperno *et al.*, 1996; Qin *et al.*, 2004; Ishikawa *et al.*, 2014), αβ-tubulin heterodimers (Hao *et al.*, 2011; Craft *et al.*, 2015) and the BBSome (Lechtreck *et al.*, 2009). The BBSome itself is an adaptor for membrane-anchored receptors that need to be transported into the cilium (Lechtreck *et al.*, 2013), but is also required for regulated export of retrograde cargo from the ciliary space (Liew *et al.*, 2014).

Despite their distinct functions in cilia and flagella, all three complexes most likely share a common ancestor that was present in LECA, the so-called ‘protocoatomer’ (van Dam *et al.*, 2013). Following the consensus, this protocoatomer was the ancestor of all eukaryotic membrane-coating proteins that facilitate inter-organellar vesicle trafficking, including endoplasmic reticulum, Golgi complex, endosomes, and the plasma membrane. The IFT complexes are also structurally related, however confidently placing the root between type I coats, type II coats, and the IFT/BBSome remains a challenge to this day. Unlike proteins from the other two families, IFT/BBSome proteins share astonishing sequence homology across vastly different eukaryotes together with the structural conservation that allow them to be confidently traced back to LECA. The high degree of sequence

conservation indicates however that the proteins were more functionally restricted and had less opportunity to diversify and expand as other coating complexes that retained structure, but not sequence, throughout their evolution.

For the complexes themselves, the most likely scenario of emergence was postulated some years ago (van Dam *et al.*, 2013). Here, the protocoatomer was the ancestral complex that gave rise to an IFT precursor related to COP I ε. The first subcomplex would have been IFT-B, parts of which expanded paralogously and gave rise to BBSome components BBS4 and BBS8. The latest addition in this model would be IFT-A. Interestingly, parts of each complex bear resemblance to different COP I subunits. Protein domains found in COP I α and β' can be found in IFT-A (IFT122, IFT140, WDR19, WDR35) and IFT-B (IFT80, IFT172), domains aligning with parts of COP I ε in IFT-A (TTC21), IFT-B (IFT88, TTC26, TTC30) and the BBSome (BBS4, BBS8) (van Dam *et al.*, 2013). This mingling of different COP I family protein domains could be a hint of a parallel expansion of both COP I and IFT/BBSome proteins, where duplications in COP I subunits could have led to their inclusion into the developing IFT complexes. It is noteworthy that like the BBSome, the COP I protein family is crucial for the transport of transmembrane proteins (such as receptors) However, for the COP I proteins to diversify, one would have to infer that there were proteins present before LECA that were not protocoatomer subunits, but already diversified COP I proteins (Fig. 3). If the COP I family had already been established, and we assume that the organellar paralogy model works, that means that the cilium would have been in development during the shaping of the organellar landscape of the increasingly complex pre-LECA cell. It is therefore likely that some form of directed transport along the MT network already took place while cilia were acquiring their distinctive trafficking complexes, fitting the previous hypothesis that the cilium began as a specialised membrane patch that would have benefited from an own, distinct coating for cargo.

While the exact order of emergence of ciliary trafficking complexes remains disputable, it is likely that the cilium was not the last feature to be established in LECA. In the previously stated hypothesis, the cilium emerged sometime after the acquisition of the mitochondrion (so that lipid bilayers could gradually replace archaeal terpenes as cellular boundaries) as it is established by complexes derived from membrane-coating protocoatomer-related proteins. The ciliary trafficking complexes share common architectural hallmarks with COP I proteins. COP I facilitates trafficking of vesicles within the Golgi apparatus (GA) and the retrograde transport of vesicles from GA to ER (Serafini *et al.*, 1991). In a less diversified coating system with less specific components to sort cargo meticulously and which might have been present in a pre-LECA cell, we can only infer that a directionality for cargo transport was already established by the time ciliary trafficking complexes emerged from the theory stated above. However, this vesicular transport must have originated from

a sort of proto-ER that allowed the integration of transmembrane proteins into vesicles that were then coated and sent to the plasma membrane. This raises an important question for the non-ciliary roles of the emerging ciliary trafficking proteins: Were non-ciliary functions ancestral or did ciliary proteins ‘slip’ into new functions by parallel inventions of diversifying eukaryotes and radiating gene families that established organellar identity? Our theory of simultaneous evolution of cilium and nucleus only depends on directed transport as a prerequisite for early ciliary evolution. The cilium could possibly even start to arise before the establishment of the nuclear membrane. To further investigate how ciliary proteins acquired nuclear functions, we need to dissect the architecture and evolutionary history of the nucleus, nuclear pore complexes, and nucleocytoplasmic transport adapters.

Eukaryotic nucleus

The nucleus is the namesake of the domain *Eukarya* (from Greek εὖ “well, good”, and κάρυον “nut, kernel”) and the container of genetic information in an eukaryotic cell. The nucleus is a membrane-bounded organelle adjacent to the ER and its membrane is, in fact, continuos with the ER’s membrane. It sequesters DNA from the cytoplasm, thereby restricting access of soluble molecules and allowing precise spatiotemporal regulation of gene activation and repression, and transcription and translation. Its invention separates the prokaryotes (bacteria and archaea) from eukaryotes and is considered one of the major transitions in life (Szathmáry and Smith, 1995; Smith and Szathmáry, 1997). The nucleoplasm is accessible via highly selective nuclear pore complexes (NPCs), multi-protein channels in the MDa range with an eigthfold symmetry that regulate import and export of mRNAs and proteins. Phylogenetic reconstruction is not simple as sequence similarity between species is generally low, but comparative genomics revealed that LECA most likely had a fully functional NPC (Neumann *et al.*, 2010).

Architecture and evolution of nuclear pore complexes

Despite its omnipresence in eukaryotes, the proteomic makeup and overall architecture of the nucleus and specifically of the NPCs is subjected to quite some variability (Bapteste *et al.*, 2005). Throughout the eukaryotic Tree of Life, it has undergone considerable lineage-specific adaptations (Fernandez-Martinez and Rout, 2021) that reflect the adaptability of the system to multiple biological niches. NPC architecture is the same between eukaryotes: The nucleoporins (Nups) forming the NPC build a core scaffold of eightfold-symmetrical linked outer and inner ring on the cytosolic and nucleosolic face of the pore; the nucleoplasmic face is directly integrated into the nuclear laminar. The NPC extends into both directions, on the cytosolic face with cytoplasmic filaments and on the nuclear face with proteins forming the nuclear basket. The central channel is decorated with phenylalanine-glycine (FG) repeat Nups that obstruct free diffusion by forming a

brush-like structure. While this structure is present in principle in all eukaryotes, the Nups themselves are only well conserved on a structural level (Rout *et al.*, 2000; Cronshaw *et al.*, 2002; Alber *et al.*, 2007; Iwamoto *et al.*, 2010; Tamura *et al.*, 2010; Ori *et al.*, 2013; Asakawa *et al.*, 2014; Eibauer *et al.*, 2015; Fischer *et al.*, 2015; Obado *et al.*, 2016; Kosinski *et al.*, 2016; Kim *et al.*, 2018; Mosalaganti *et al.*, 2018). They are all composed of the same α-solenoid/β-propeller structure found in descendants of the hypothetical protocoatomer. It is highly likely that Nups were present in LECA (Neumann *et al.*, 2010) and that they expanded paralogously before the radiation of eukaryotic lineages (Devos *et al.*, 2004; Alber *et al.*, 2007; Kosinski *et al.*, 2016; Kim *et al.*, 2018; Beck *et al.*, 2018). As they have little sequence conservation, evolutionary pressure was not on the exact amino acid composition of the Nups, but more on their ability to form the same complex in different ways. Nups are related to other membrane coating proteins from both COP I and COP II families (Kim *et al.*, 2018), but also adaptin-like proteins (Alber *et al.*, 2007; Knockenhauer and Schwartz, 2016; Kosinski *et al.*, 2016; Hoelz *et al.*, 2016), and have low levels of sequence conservation between lineages (Mans *et al.*, 2004; Bapteste *et al.*, 2005; DeGrasse *et al.*, 2009). This argues against a “nucleus-first” scenario in the evolution of eukaryotes, or at the least would mean that the nucleus co-evolved with ER and GA, as COP I and COP II coats are needed for transport between ER, GA and plasma membrane. Given that both families appear in the nucleus, primitive forms of the other organelles could have been present at the time NPCs were evolving.

The evolutionary origin of Nups (and other protocoatomer-derived proteins for that matter) likely lies in prokaryotic, archaeal α-solenoid/β-propeller domain proteins (Field and Rout, 2019). As previously discussed, the expansion of this protein family led to the diverse membrane coats that shape organellar identity and assemble into some of the most intricate multiprotein complexes of eukaryotic cells. Likewise, the association of early Nup precursor proteins with membranes could likely lead to primitive repetitive structures resembling proto-NPCs in the (possibly incomplete) NE. Through duplications, the NPC gained considerable diversity in composition that retained interaction sites with other NPC subunits.

Although the forming NE and NPC were initially likely to be ‘leaky’ in the sense that they did not fully sequester DNA and its interacting proteins from the cytosol, there are obvious advantages of an incomplete NE and NPC. Even purely physical obstacles like a non-selective membrane with rather large fenestrations would slow DNA-interacting molecules’ diffusion, limiting gene activation and repression and ultimately shaping signalling cascades and cell cycle stages. Permeability of these structures for gene regulatory components is an evolutionarily selectable feature, favouring more complex and functionally fine-tuned structures. Higher selectivity would pressure the evolving cell into co-evolution of proteins that are desired in close proximity to DNA

for gene regulatory (e.g. methylation, transcriptional activators or repressors) or stability (e.g. damage repair, histones) purposes, and their unhindered entry into the restricted compartment. After establishment of a full NE, the only point of entry for proteins into the nucleus would be the pores formed by proto-NPCs that need not be the sophisticated, highly selective complexes that were most likely present in LECA. Even non-selective pores formed by membrane-deforming proto-Nups would convey size selection simply by their architecture: Depending on the size of the molecule that wants to traverse the pore (where larger molecules diffuse more slowly than small ones, (Grathwohl, 2012)) and the diameter of the pore itself, the probability of a molecule not only encountering, but passing through a non-selective pore, becomes smaller as molecule size increases and pore size decreases (Purcell, 1977; Berg, 1993; Rout *et al.*, 2003; Zilman *et al.*, 2007). Extant NPCs do however implement further measures to ensure that no unwanted molecules enter the nuclear space. The actual pore itself is decorated with phenylalanine-glycine (FG) moieties (Rout *et al.*, 2000; Alber *et al.*, 2007) that stem from so-called FG-Nups. These FG-Nups are highly dynamic, intrinsically disordered proteins with large hydrodynamic radii and effectively decrease the size of the pore further (Denning *et al.*, 2003). They prevent random passage through the nuclear pore by inter- and intramolecular electrostatic interactions (Gehrke *et al.*, 1997; Kawai *et al.*, 2003) resulting in high entropy within the NPC. Establishing such an intricate structure as the NPC to prevent free nuclear localisation however comes with the need to also import and export proteins and molecules at the right times. Part of this regulation is conveyed by FG-Nups themselves; the selective import and export of nuclear-bound molecules is however also enhanced by adapter proteins, the karyopherins.

Karyopherins

Karyopherins (Kaps) are a large family of proteins found in nuclear import and export pathways of eukaryotic cells. Phylogenetic reconstruction showed that LECA probably had an extensive repertoire of Kaps (O'Reilly *et al.*, 2011), consistent with the presence of a rather complex NPC and therefore also a sophisticated nuclear import and export system (O'Reilly *et al.*, 2011). They also form mostly α-solenoid/β-propeller structures, with the dominant protein domains being HEAT and ARM repeats (Bayliss *et al.*, 2000, 2002; Stuwe *et al.*, 2014), and are related to Nups and other membrane coats via their likely proto-coatomer origin. Depending on the function they fulfil in their respective pathways, they are also conveniently called importins and exportins. Both protein families are cargo adapters for nuclear proteins. Through hydrophobic interactions of Kaps with the FG residues within the NPCs, the Kap-cargo-complex is able to overcome the high entropy that would normally impede free diffusion of the cargo proteins (Iovine *et al.*, 1995). Although diffusion in presence of FG repeats is possible, the presence of specialised import and export proteins that

compete for the same binding sites in NPCs effectively blocks unspecific passage through higher affinity (Zilman *et al.*, 2007; Jovanovic-Talisman *et al.*, 2009; Zilman, 2009).

Passage through the NPC is not only dependent on the presence of Kaps. The pathways use the RanGTP/RanGDP gradient across the NE to further guarantee correct direction of transport. Importins in their cargo-bound state translocate through the NPC (Izaurralde *et al.*, 1997) and release their cargo in the nucleus before binding nucleoplasmic RanGTP and being exported again. In the cytoplasm, hydrolysis of RanGTP to RanGDP releases importin from the export complex to be used in another cycle. For proteins to be exported from the nucleus, exportin associates with its cargo and nuclear RanGTP to form the export complex, only to be released in the cytoplasm through hydrolysis of RanGTP to RanGDP. This gradient is crucial for nucleocytoplasmic transport (Lui and Huang, 2009).

Cargo is recognised by nuclear localisation or export signals (NLS and NES, respectively) within their amino acid sequences and are bound by the respective importin or exportin. As Kaps have been established before radiation of eukaryotic phyla, so must have been the sequences marking proteins as nuclear cargo, i.e. NLS and NES. Canonical sequences seem to be functional across different eukaryotic lineages, so they most likely also originated before LECA. To understand how proteins can function in the nucleus, it is worth exploring how nuclear signal sequences came to be in the first place.

Evolution of nuclear signal sequences NLS and NES

Nuclear localisation and export signals (NLS and NES) are amino acids sequences that confer nuclear import or export capability to a protein. There is a range of sequences and biophysical properties that define classes of NLS and NES, and those sequence motifs are highly conserved throughout eukaryotes. Classical NLSs are either monopartite or bipartite targeting signals comprised of basic amino acids that associate with Kaps for nuclear import. Monopartite NLSs contain mostly lysine and/or arginine residues and loosely follow the consensus motif K-(K/R)-X-(K/R) first identified in the SV40 large T antigen (Kalderon *et al.*, 1984a; b; Dang and Lee, 1988; Boulikas, 1993). Bipartite NLSs are separated by a spacer of variable length, but otherwise share the same biochemical properties. The first description of bipartite NLSs was in the protein nucleoplasmin, with their consensus sequence being $(K/R)(K/R)X_{10–12}(K/R)_{3/5}$ (Dingwall and Laskey, 1991; Robbins *et al.*, 1991; Makkerh *et al.*, 1996; Lange *et al.*, 2007). Other motifs have been identified since, such as the monopartite acidic M9 motif (also known as PY-NLS, containing proline and tryptophan) from the mRNA-binding protein hnRNPA1 (Siomi and Dreyfuss, 1995; Görlich *et al.*, 1996; Weis *et al.*, 1996; Pollard *et al.*, 1996; Bonifaci *et al.*, 1997). Although all are either recognised directly by importin β or sequentially by importin α and then importin β, and

different motifs convey different transport rates (Ray *et al.*, 2015), all motifs are sufficient to facilitate nuclear import.

NESs function analogously to NLSs, but facilitate export from rather than import into the nucleus. They are comprised of hydrophobic stretches of amino acids usually rich in leucines, following the consensus (L/I/V/F/M)-$X_{2–3}$-(L/I/V/F/M)-$X_{2–3}$-(L/I/V/F/M)-X-(L/I/V/F/M) (Wen *et al.*, 1995; Bogerd *et al.*, 1996; Kim *et al.*, 1996; Kosugi *et al.*, 2008). Nuclear export is facilitated by association of the Kap exportin-1 (Xpo1; also called chromosome region maintenance 1, CRM1) (Adachi and Yanagida, 1989; Toda *et al.*, 1992; Fornerod *et al.*, 1997; Fukuda *et al.*, 1997), and no other major protein export pathways have been described to date.

The high degree of functional conservation of NLS and NES argues for an ancient origin. Interestingly, structural evidence for a prokaryotic origin comes from ribosome analyses. Prokaryotic ribosomal proteins contain water and magnesium ions that give the ribosome structural integrity. In human ribosomes, water and magnesium are replaced by moieties that maintain the same contacts to rRNA and functionally also serve as NLS (Melnikov *et al.*, 2015). As ribosomal proteins comprise about half of the most conserved proteins throughout all domains of life (Koonin, 2003), future proteomic analyses into deeply branching eukaryotes, archaea and bacteria could uncover more similarities and strengthen this theory. Recent work could point to a similar direction for the origin of NLS: NLS-like amino acid sequences can also be identified in the DNA-binding regions of prokaryotic proteins (Lisitsyna *et al.*, 2020). And indeed, in extant eukaryotes, NLSs often co-localise with DNA- or RNA-binding motifs (LaCasse and Lefebvre, 1995; Cokol *et al.*, 2000; Lisitsyna *et al.*, 2020). It would make sense that the first proteins to contain NLSs in the prokaryote-to-eukaryote transition would be the proteins already associated with DNA, as their DNA-associated functions most likely would also be needed in the evolving pre-LECA cell. Purifying selection could then have sorted out obsolete proteins. This would sufficiently explain how a primitive nuclear proteome could arise in the first place.

Ciliary proteins in the nucleus

How can the evolutionary insights into organelle acquisition and formation help us formulate a hypothesis for nuclear functions of ciliary proteins? There are clear links between the two cellular compartments: Both have likely been present since LECA (Jékely, 2003; Jékely and Arendt, 2006; Hodges *et al.*, 2010; van Dam *et al.*, 2013; Field *et al.*, 2014; Devos *et al.*, 2014; Rout and Field, 2017) and were major factors in what discerns eukaryotes from prokaryotes (Szathmáry and Smith, 1995; Smith and Szathmáry, 1997). The proteins that shape and maintain the respective structures have a shared common ancestor, the hypothetical proto-coatomer (Devos *et al.*, 2006; O'Reilly *et*

al., 2011; Sampathkumar *et al.*, 2013; van Dam *et al.*, 2013; Stuwe *et al.*, 2014). Response to extracellular clues often happens through the cilium that serves as a specialised region for receptors and signalling molecules, effectors of which often regulate gene expression on different levels (Nozawa *et al.*, 2013; Satir and Satir, 2019). Proteins from both compartments can be found in the other, e.g. Nups participating in a selective diffusion barrier at the ciliary base (Kee *et al.*, 2012; Kee and Verhey, 2013; Marquez *et al.*, 2021), and ciliary kinesins that shuttle between cilia and the nucleus (Morris *et al.*, 2004; Huang *et al.*, 2021). The ciliary kinesin KIF7 even mimicks a DNA double strand through homodimerisation to bind GLI transcription factors to the ciliary axoneme and sequestering it from the soluble phase of cilia (Haque *et al.*, 2022). Recent studies showed that proteins of the BBSome, a highly conserved ciliary trafficking complex, can not only localise to the nucleus of mammalian cells (Ewerling *et al.*, 2023) but play active roles in gene regulation (Gascue *et al.*, 2012; Scott *et al.*, 2017; Horwitz and Birk, 2021) and cell homeostasis (Marchese *et al.*, 2022). Other parts of IFT complexes have been found in the nucleus of the ciliate *Paramecium tetraurelia* (Shi *et al.*, 2018), although no clear functions could be deducted so far. The ciliary KAP3 (kinesin associated protein 3), a highly conserved ciliary protein, uses its ARM repeats for both ciliary and nuclear localisation, and a RanGTP gradient controls its translocation from the cytosol to either nucleus or cilia of *Chlamydomonas reinhardtii* (Huang *et al.*, 2021).

Shedding light onto the evolution of organelles also moonlights potential intersections and platforms that are needed for co-evolution of protein subsets of each organelle. Their shared evolutionary history allows to infer potential trajectories for co-evolution. Our hypothesis supports a ‘late nucleus’ model were the NE is only formed after or concomitantly with the split between ciliary and cytosolic coating complexes, and where the NE/NPC co-evolved with the cilium. We need only infer minimal prerequisites (polarised trafficking, established microtubule network) for this theory, so the proto-cilium (or even just a polarised membrane patch) could arise quite early during the FECA-to-LECA transition, possibly even before the nucleus. Several points speak for a co-evolution of nucleus and cilia and, by extension, for potentially ancestral nuclear roles of ciliary proteins. As mentioned above, many extant proteins have functions in both compartments. But the similarities do not end there. Both the nucleus and the cilium use RanGTP/RanGDP gradients and importins to facilitate translocation from the cytosol to the respective compartment, the latter albeit to a lesser extent and only for specific proteins (Dishinger *et al.*, 2010). In fact, a “ciliary pore complex” (CPC) was proposed some time ago that acts akin to a NPC and even might recognise ‘ciliary localisation sequences’ (CLS) similar to NLS (Kee *et al.*, 2012; Kee and Verhey, 2013). This CPC is structurally distinct from the NPC as it is not constituted of all canonical NPC functional subunits and is not anchored to the membrane via transmembrane Nups (Kee *et al.*, 2012), but might convey import selectivity and anchoring via association with other ciliary gating complexes

such as the NPHP/MKS complex (Garcia-Gonzalo and Reiter, 2012; Takao *et al.*, 2017). Nucleoporins have been identified at the ciliary base, orchestrating ciliary import in a manner akin to their role at nuclear pores. Nup98 (Endicott and Brueckner, 2018), Nup205 (Marquez *et al.*, 2021), Nup62 (Kee *et al.*, 2012), Nup93, Nup98, and Nup188 (Endicott *et al.*, 2015; del Viso *et al.*, 2016), as well as Nup85 (Endicott and Brueckner, 2018), are all situated at the base of cilia facilitating ciliary cargo sorting and playing pivotal roles in determining the left-right axis in mammals (Endicott *et al.*, 2015; Blasius *et al.*, 2019; Heinß *et al.*, 2020). While the NPC is tethered to the nuclear envelope through interactions between membrane ring Nups and other NPC constituents, cilia present just one surface with a lipid bilayer that could potentially serve as an anchor for the CPC. Nevertheless, the NPHP module responsible for cargo sorting occupies the membrane-facing ciliary space within the transition zone (Sang *et al.*, 2011; Wang *et al.*, 2022) and has demonstrated interactions with Nups (Blasius *et al.*, 2019). Whether Nups are also affixed to the axonemal microtubules, forming a complete NPC/CPC, or if they line the inner perimeter of the cilium in a "brush border" arrangement, remains currently undisclosed, necessitating further comprehensive interactomic investigations to fully elucidate the spatial arrangement of Nups at the ciliary base.

Concluding remarks

Taken together, the evolution of single organelles and their proteomes allows us to form a hypothesis how nuclear functions of ciliary trafficking proteins came to be: If cilia were born from a polarised membrane patch with signalling receptors, it is possible that effector molecules of these receptors regulated a change in gene expression in response to environmental cue, meaning the signalling was directly from proto-cilium to DNA. Low-complexity, non-expanded signalling pathways might have been in place before sequential/combinatorial regulation, meaning proteins from the proto-cilium also affected gene regulation directly and needed to associate with DNA. For this, ‘ciliary’ proteins would also have needed DNA-binding motifs (that were also present before cilia) that could eventually evolve into NLS, as has been proposed (LaCasse and Lefebvre, 1995; Cokol *et al.*, 2000; Lisitsyna *et al.*, 2020). Through this, ciliary proteins would have a preexisting propensity to enter the nucleus before the NE and NPC were fully in place. The obvious evolutionary advantage in pre-eukaryotes would be that no additional layers of regulation or *de novo* invention of nuclear effectors were needed due to recycling of existing mechanisms like nuclear import. This would also explain why some of the Nups could be re-purposed to function in ciliary import sorting. In this scenario, ‘ciliary’ proteins would have dual functions until cilia were fully established after NPCs were functionally selective barriers. With addition of layers for signalling cascade regulation, some of the ciliary proteins would likely have lost the need to

translocate to the nucleus directly and instead remained ciliary (Fig. 4). However, to fully answer the question in which order organelles and compartments evolved it is crucial to gain an understanding the relationship between the derivatives of the protocoatomer. The uncertain placement of acquisition of BBSome/IFT complexes relative to acquisition of other organelles makes it hard to untangle the possible scenarios of how ciliary proteins could have evolved nuclear functions. Phylogenetic reconstruction of earlier diverging eukaryotic clades will bring clarity to how NPCs looked and functioned in LECA, and will also help resolve the relationship between type I, type II and IFT/BBSome coats. This will help determining the exact sequence of organelle evolution, and by extension shed light onto eukaryogenesis itself by clarifying how basic biological functions arose *de novo* in the developing pre-eukaryotic cell.

Declaration of interests

The authors declare no competing interests.

Author contributions

Conceptualization, A.E. and H.L.M.-S.; Investigation, A.E.; Resources, H.L.M.-S.; Data curation, A.E.; Writing – Original Draft, A.E.; Writing – Review & Editing, A.E. and H.L.M.-S.; Visualization, A.E.; Supervision, H.L.M.-S.; Project administration, H.L.M.-S.; Funding Acquisition, H.L.M.-S.

References

Alber F, Dokudovskaya S, Veenhoff LM, Zhang W, Kipper J, Devos D, Suprapto A, Karni-Schmidt O, Williams R, Chait BT, Sali A, Rout MP (2007) The molecular architecture of the nuclear pore complex. *Nature* **450**: 695–701.

Alexander DL, Mital J, Ward GE, Bradley P, Boothroyd JC (2005) Identification of the Moving Junction Complex of Toxoplasma gondii: A Collaboration between Distinct Secretory Organelles. *PLOS Pathogens* **1**: e17.

Araç D, Dulubova I, Pei J, Huryeva I, Grishin NV, Rizo J (2005) Three-dimensional Structure of the rSly1 N-terminal Domain Reveals a Conformational Change Induced by Binding to Syntaxin 5. *Journal of Molecular Biology* **346**: 589–601.

Asakawa H, Yang H-J, Yamamoto TG, Ohtsuki C, Chikashige Y, Sakata-Sogawa K, Tokunaga M, Iwamoto M, Hiraoka Y, Haraguchi T (2014) Characterization of nuclear pore complex components in fission yeast Schizosaccharomyces pombe. *Nucleus* **5**: 149–162.

Bapteste E, Charlebois RL, MacLeod D, Brochier C (2005) The two tempos of nuclear pore complex evolution: highly adapting proteins in an ancient frozen structure. *Genome Biology* **6**: R85.

Bayliss R, Littlewood T, Stewart M (2000) Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. *Cell* **102**: 99–108.

Bayliss R, Littlewood T, Strawn LA, Wente SR, Stewart M (2002) GLFG and FxFG nucleoporins bind to overlapping sites on importin-beta. *The Journal of Biological Chemistry* **277**: 50597–50606.

Berg HC, Purcell EM (1977) Physics of chemoreception. *Biophysical Journal* **20**: 193–219.

Bloodgood RA (2010) Sensory reception is an attribute of both primary cilia and motile cilia. *Journal of Cell Science* **123**: 505–509.

Bogerd HP, Fridell RA, Benson RE, Hua J, Cullen BR (1996) Protein sequence requirements for function of the human T-cell leukemia virus type 1 Rex nuclear export signal delineated by a novel in vivo randomization-selection assay. *Molecular and Cellular Biology* **16**: 4207–4214.

Bonifaci N, Moroianu J, Radu A, Blobel G (1997) Karyopherin beta2 mediates nuclear import of a mRNA binding protein. *Proceedings of the National Academy of Sciences of the United States of America* **94**: 5055–5060.

Boulikas T (1993) Nuclear localization signals (NLS). *Critical Reviews in Eukaryotic Gene Expression* **3**: 193–227.

Bremer N, Tria FDK, Skejo J, Garg SG, Martin WF (2022) Ancestral State Reconstructions Trace Mitochondria But Not Phagocytosis to the Last Eukaryotic Common Ancestor. *Genome Biology and Evolution* **14**: evac079.

Brighouse A, Dacks JB, Field MC (2010) Rab protein evolution and the history of the eukaryotic endomembrane system. *Cellular and Molecular Life Sciences* **67**: 3449–3465.

Brohawn SG, Leksa NC, Spear ED, Rajashankar KR, Schwartz TU (2008) Structural Evidence for Common Ancestry of the Nuclear Pore Complex and Vesicle Coats. *Science* **322**: 1369–1373.

Carlton JG, Martin-Serrano J (2007) Parallels Between Cytokinesis and Retroviral Budding: A Role for the ESCRT Machinery. *Science* **316**: 1908–1912.

Carvalho-Santos Z, Azimzadeh J, Pereira-Leal JoséB, Bettencourt-Dias M (2011) Tracing the origins of centrioles, cilia, and flagella. *Journal of Cell Biology* **194**: 165–175.

Cassioli C, Onnis A, Finetti F, Capitani N, Brunetti J, Compeer EB, Niederlova V, Stepanek O, Dustin ML, Baldari CT (2021) The Bardet–Biedl syndrome complex component BBS1 controls T cell polarity during immune synapse assembly. *Journal of Cell Science* **134**: jcs258462.

Cavalier-Smith T (1987) The origin of eukaryotic and archaebacterial cells. *Annals of the New York Academy of Sciences* **503**: 17–54.

Cavalier-Smith T (2002) The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. *International Journal of Systematic and Evolutionary Microbiology* **52**: 297–354.

Chaki M, Airik R, Ghosh AK, Giles RH, Chen R, Slaats GG, Wang H, Hurd TW, Zhou W, Cluckey A, Gee HY, Ramaswami G, Hong C-J, Hamilton BA, Červenka I, Ganji RS, Bryja V, Arts HH, van Reeuwijk J, Oud MM, Letteboer SJF, Roepman R, Husson H, Ibraghimov-Beskrovnaya O, Yasunaga T, Walz G, Eley L, Sayer JA, Schermer B, Liebau MC, Benzing T, Le Corre S, Drummond I, Janssen S, Allen SJ, Natarajan S, O'Toole JF, Attanasio M, Saunier S, Antignac C, Koenekoop RK, Ren H, Lopez I, Nayir A, Stoetzel C, Dollfus H, Massoudi R, Gleeson JG, Andreoli SP, Doherty DG, Lindstrad A, Golzio C, Katsanis N, Pape L, Abboud EB, Al-Rajhi AA, Lewis RA, Omran H, Lee EY-HP, Wang S, Sekiguchi JM, Saunders R, Johnson CA, Garner E, Vanselow K, Andersen JS, Shlomai J, Nurnberg G, Nurnberg P, Levy S, Smogorzewska A, Otto EA, Hildebrandt F (2012) Exome Capture Reveals ZNF423 and CEP164 Mutations, Linking Renal Ciliopathies to DNA Damage Response Signaling. *Cell* **150**: 533–548.

Choi HJC, Lin J-R, Vannier J-B, Slaats GG, Kile AC, Paulsen RD, Manning DK, Beier DR, Giles RH, Boulton SJ, Cimprich KA (2013) NEK8 Links the ATR-Regulated Replication Stress Response and S Phase CDK Activity to Renal Ciliopathies. *Molecular Cell* **51**: 423–439.

Cokol M, Nair R, Rost B (2000) Finding nuclear localization signals. *EMBO reports* **1**: 411–415.

Craft JM, Harris JA, Hyman S, Kner P, Lechtreck KF (2015) Tubulin transport by IFT is upregulated during ciliary growth by a cilium-autonomous mechanism. *Journal of Cell Biology* **208**: 223–237.

Cronshaw JM, Krutchinsky AN, Zhang W, Chait BT, Matunis MJ (2002) Proteomic analysis of the mammalian nuclear pore complex. *The Journal of Cell Biology* **158**: 915–927.

Dacks JB, Doolittle WF (2002) Novel syntaxin gene sequences from Giardia, Trypanosoma and algae: implications for the ancient evolution of the eukaryotic endomembrane system. *Journal of Cell Science* **115**: 1635–1642.

Dacks JB, Field MC (2007) Evolution of the eukaryotic membrane-trafficking system: origin, tempo and mode. *Journal of Cell Science* **120**: 2977–2985.

Dam TJP van, Townsend MJ, Turk M, Schlessinger A, Sali A, Field MC, Huynen MA (2013) Evolution of modular intraflagellar transport from a coatomer-like progenitor. *Proceedings of the National Academy of Sciences* **110**: 6943–6948.

Delaval B, Bright A, Lawson ND, Doxsey S (2011) The cilia protein IFT88 is required for spindle orientation in mitosis. *Nature Cell Biology* **13**: 461–468.

Devos D, Dokudovskaya S, Alber F, Williams R, Chait BT, Sali A, Rout MP (2004) Components of Coated Vesicles and Nuclear Pore Complexes Share a Common Molecular Architecture. *PLOS Biology* **2**: e380.

Devos D, Dokudovskaya S, Williams R, Alber F, Eswar N, Chait BT, Rout MP, Sali A (2006) Simple fold composition and modular architecture of the nuclear pore complex. *Proceedings of the National Academy of Sciences* **103**: 2172–2177.

Dingwall C, Laskey RA (1991) Nuclear targeting sequences — a consensus? *Trends in Biochemical Sciences* **16**: 478–481.

Dishinger JF, Kee HL, Jenkins PM, Fan S, Hurd TW, Hammond JW, Truong YN-T, Margolis B, Martens JR, Verhey KJ (2010) Ciliary entry of the kinesin-2 motor KIF17 is regulated by importin-β2 and RanGTP. *Nature Cell Biology* **12**: 703–710.

Dulk B den, Eijk P van, Ruijter M de, Brandsma JA, Brouwer J (2008) The NER protein Rad33 shows functional homology to human Centrin2 and is involved in modification of Rad4. *DNA Repair* **7**: 858–868.

East MP, Kahn RA (2011) Models for the functions of Arf GAPs. *Seminars in Cell & Developmental Biology* **22**: 3–9.

Eibauer M, Pellanda M, Turgay Y, Dubrovsky A, Wild A, Medalia O (2015) Structure and gating of the nuclear pore complex. *Nature Communications* **6**: 7532.

Elde NC, Morgan G, Winey M, Sperling L, Turkewitz AP (2005) Elucidation of Clathrin-Mediated Endocytosis in Tetrahymena Reveals an Evolutionarily Convergent Recruitment of Dynamin. *PLOS Genetics* **1**: e52.

Elias M, Brighouse A, Gabernet-Castello C, Field MC, Dacks JB (2012) Sculpting the endomembrane system in deep time: high resolution phylogenetics of Rab GTPases. *Journal of Cell Science* **125**: 2500–2508.

Erickson HP (2007) Evolution of the cytoskeleton. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* **29**: 668–677.

Ewerling A, Maissl V, Wickstead B, May-Simera HL (2023) Neofunctionalization of ciliary BBS proteins to nuclear roles is likely a frequent innovation across eukaryotes. *iScience* **26**: 106410.

Faguy DM, Doolittle WF (1998) Cytoskeletal proteins: The evolution of cell division. *Current Biology* **8**: R338–R341.

Field MC, Gabernet-Castello C, Dacks JB (2007) Reconstructing the Evolution of the Endocytic System: Insights from Genomics and Molecular Cell Biology. In: *Eukaryotic Membranes and Cytoskeleton: Origins and Evolution*, Advances in Experimental Medicine and Biology. Springer, New York, NY, pp. 84–96.

Finetti F, Paccani SR, Riparbelli MG, Giacomello E, Perinetti G, Pazour GJ, Rosenbaum JL, Baldari CT (2009) Intraflagellar transport is required for polarized recycling of the TCR/CD3 complex to the immune synapse. *Nature Cell Biology* **11**: 1332–1339.

Fischer J, Teimer R, Amlacher S, Kunze R, Hurt E (2015) Linker Nups connect the nuclear pore complex inner ring with the outer ring and transport channel. *Nature Structural & Molecular Biology* **22**: 774–781.

Fukuda M, Asano S, Nakamura T, Adachi M, Yoshida M, Yanagida M, Nishida E (1997) CRM1 is responsible for intracellular transport mediated by the nuclear export signal. *Nature* **390**: 308–311.

Gascue C, Tan PL, Cardenas-Rodriguez M, Libisch G, Fernandez-Calero T, Liu YP, Astrada S, Robello C, Naya H, Katsanis N, Badano JL (2012) Direct role of Bardet–Biedl syndrome proteins in transcriptional regulation. *Journal of Cell Science* **125**: 362–375.

Gould SB, Garg SG, Martin WF (2016) Bacterial Vesicle Secretion and the Evolutionary Origin of the Eukaryotic Endomembrane System. *Trends in Microbiology* **24**: 525–534.

Hao K, Chen Y, Yan X, Zhu X (2021) Cilia locally synthesize proteins to sustain their ultrastructure and functions. *Nature Communications* **12**: 6971.

Hao L, Thein M, Brust-Mascher I, Civelekoglu-Scholey G, Lu Y, Acar S, Prevo B, Shaham S, Scholey JM (2011) Intraflagellar transport delivers tubulin isotypes to sensory cilium middle and distal segments. *Nature Cell Biology* **13**: 790–798.

Haque F, Freniere C, Ye Q, Mani N, Wilson-Kubalek EM, Ku P-I, Milligan RA, Subramanian R (2022) Cytoskeletal regulation of a transcription factor by DNA mimicry via coiled-coil interactions. *Nature Cell Biology* **24**: 1088–1098.

Hartman H, Smith TF (2009) The evolution of the cilium and the eukaryotic cell. *Cell Motility* **66**: 215–219.

Hickey D, Vilfan A, Golestanian R (2021) Ciliary chemosensitivity is enhanced by cilium geometry and motility. *eLife* **10**: e66322.

Hodges ME, Scheumann N, Wickstead B, Langdale JA, Gull K (2010) Reconstructing the evolutionary history of the centriole from protein components. *Journal of Cell Science* **123**: 1407–1413.

Hodges ME, Wickstead B, Gull K, Langdale JA (2011) Conservation of ciliary proteins in plants with no cilia. *BMC Plant Biology* **11**: 185.

Horwitz A, Birk R (2021) BBS4 Is Essential for Nuclear Transport of Transcription Factors Mediating Neuronal ER Stress Response. *Molecular Neurobiology* **58**: 78–91.

Hu Q, Milenkovic L, Jin H, Scott MP, Nachury MV, Spiliotis ET, Nelson WJ (2010) A Septin Diffusion Barrier at the Base of the Primary Cilium Maintains Ciliary Membrane Protein Distribution. *Science* **329**: 436–439.

Huang K, Diener DR, Mitchell A, Pazour GJ, Witman GB, Rosenbaum JL (2007) Function and dynamics of PKD2 in Chlamydomonas reinhardtii flagella. *Journal of Cell Biology* **179**: 501–514.

Huang S, Dougherty LL, Avasthi P (2021) Separable roles for RanGTP in nuclear and ciliary trafficking of a kinesin-2 subunit. *Journal of Biological Chemistry* **296**.

Hurley JH, Emr SD (2006) THE ESCRT COMPLEXES: Structure and Mechanism of a Membrane-Trafficking Network. *Annual Review of Biophysics and Biomolecular Structure* **35**: 277–298.

Hurley JH, Hanson PI (2010) Membrane budding and scission by the ESCRT machinery: it's all in the neck. *Nature Reviews Molecular Cell Biology* **11**: 556–566.

Iovine MK, Watkins JL, Wente SR (1995) The GLFG repetitive region of the nucleoporin Nup116p interacts with Kap95p, an essential yeast nuclear import factor. *The Journal of Cell Biology* **131**: 1699–1713.

Ishikawa H, Ide T, Yagi T, Jiang X, Hirono M, Sasaki H, Yanagisawa H, Wemmer KA, Stainier DY, Qin H, Kamiya R, Marshall WF (2014) TTC26/DYF13 is an intraflagellar transport protein required for transport of motility-related proteins into flagella. *eLife* **3**: e01566.

Iwamoto M, Asakawa H, Hiraoka Y, Haraguchi T (2010) Nucleoporin Nup98: a gatekeeper in the eukaryotic kingdoms. *Genes to Cells* **15**: 661–669.

Izaurralde E, Kutay U, Kobbe C von, Mattaj IW, Görlich D (1997) The asymmetric distribution of the constituents of the Ran system is essential for transport into and out of the nucleus. *The EMBO journal* **16**: 6535–6547.

Jékely G (2003) Small GTPases and the evolution of the eukaryotic cell. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* **25**: 1129–1138.

Jékely G, Arendt D (2006) Evolution of intraflagellar transport from coated vesicles and autogenous origin of the eukaryotic cilium. *BioEssays* **28**: 191–198.

Kee HL, Dishinger JF, Lynne Blasius T, Liu C-J, Margolis B, Verhey KJ (2012) A size-exclusion permeability barrier and nucleoporins characterize a ciliary pore complex that regulates transport into cilia. *Nature Cell Biology* **14**: 431–437.

Killick-Kendrick R, Peters W (1978) *Rodent malaria*. Academic Press, London, New York.

Kim FJ, Beeche AA, Hunter JJ, Chin DJ, Hope TJ (1996) Characterization of the nuclear export signal of human T-cell lymphotropic virus type 1 Rex reveals that nuclear export is mediated by position-variable hydrophobic interactions. *Molecular and Cellular Biology* **16**: 5147–5155.

Kim SJ, Fernandez-Martinez J, Nudelman I, Shi Y, Zhang W, Raveh B, Herricks T, Slaughter BD, Hogan JA, Upla P, Chemmama IE, Pellarin R, Echeverria I, Shivaraju M, Chaudhury AS, Wang J, Williams R, Unruh JR, Greenberg CH, Jacobs EY, Yu Z, Cruz MJ de la, Mironska R, Stokes DL, Aitchison JD, Jarrold MF, Gerton JL, Ludtke SJ, Akey CW, Chait BT, Sali A, Rout MP (2018) Integrative structure and functional anatomy of a nuclear pore complex. *Nature* **555**: 475–482.

Kim JC, Ou YY, Badano JL, Esmail MA, Leitch CC, Fiedrich E, Beales PL, Archibald JM, Katsanis N, Rattner JB, Leroux MR (2005) MKKS/BBS6, a divergent chaperonin-like protein linked to the obesity disorder Bardet-Biedl syndrome, is a novel centrosomal component required for cytokinesis. *Journal of Cell Science* **118**: 1007–1020.

Kontou A, Herman EK, Field MC, Dacks JB, Koumandou VL (2022) Evolution of factors shaping the endoplasmic reticulum. *Traffic (Copenhagen, Denmark)* **23**: 462–473.

Kosinski J, Mosalaganti S, Appen A von, Teimer R, DiGuilio AL, Wan W, Bui KH, Hagen WJH, Briggs JAG, Glavy JS, Hurt E, Beck M (2016) Molecular architecture of the inner ring scaffold of the human nuclear pore complex. *Science (New York, N.Y.)* **352**: 363–365.

Koumandou VL, Dacks JB, Coulson RM, Field MC (2007) Control systems for membrane fusion in the ancestral eukaryote; evolution of tethering complexes and SM proteins. *BMC Evolutionary Biology* **7**: 29.

Koumandou VL, Wickstead B, Ginger ML, Giezen M van der, Dacks JB, Field MC (2013) Molecular paleontology and complexity in the last eukaryotic common ancestor. *Critical Reviews in Biochemistry and Molecular Biology* **48**: 373–396.

Kuroiwa T, Kuroiwa H, Sakai A, Takahashi H, Toda K, Itoh R (1998) The Division Apparatus of Plastids and Mitochondria. In: Jeon KW (Ed.), *International Review of Cytology*, Academic Press, pp. 1–41.

LaCasse EC, Lefebvre YA (1995) Nuclear localization signals overlap DNA- or RNA-binding domains in nucleic acid-binding proteins. *Nucleic Acids Research* **23**: 1647–1656.

Lechtreck KF, Brown JM, Sampaio JL, Craft JM, Shevchenko A, Evans JE, Witman GB (2013) Cycling of the signaling protein phospholipase D through cilia requires the BBSome only for the export phase. *Journal of Cell Biology* **201**: 249–261.

Lechtreck K-F, Johnson EC, Sakai T, Cochran D, Ballif BA, Rush J, Pazour GJ, Ikebe M, Witman GB (2009) The Chlamydomonas reinhardtii BBSome is an IFT cargo required for export of specific signaling proteins from flagella. *Journal of Cell Biology* **187**: 1117–1132.

Leung KF, Dacks JB, Field MC (2008) Evolution of the multivesicular body ESCRT machinery; retention across the eukaryotic lineage. *Traffic (Copenhagen, Denmark)* **9**: 1698–1716.

Liew GM, Ye F, Nager AR, Murphy JP, Lee JS, Aguiar M, Breslow DK, Gygi SP, Nachury MV (2014) The intraflagellar transport protein IFT27 promotes BBSome exit from cilia through the GTPase ARL6/BBS3. *Developmental Cell* **31**: 265–278.

Lindemann CB (2022) The flagellar germ-line hypothesis: How flagellate and ciliate gametes significantly shaped the evolution of organismal complexity. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* **44**: e2100143.

Lisitsyna OM, Kurnaeva MA, Arifulin EA, Shubina MY, Musinova YR, Mironov AA, Sheval EV (2020) Origin of the nuclear proteome on the basis of pre-existing nuclear localization signals in prokaryotic proteins. *Biology Direct* **15**: 9.

Loidl J (2016) Conservation and Variability of Meiosis Across the Eukaryotes. *Annual Review of Genetics* **50**: 293–316.

López-García P, Moreira D (2006) Selective forces for the origin of the eukaryotic nucleus. *BioEssays* **28**: 525–533.

Löwe J, Amos LA (2009) Evolution of cytomotive filaments: The cytoskeleton from prokaryotes to eukaryotes. *The International Journal of Biochemistry & Cell Biology* **41**: 323–329.

Makkerh JPS, Dingwall C, Laskey RA (1996) Comparative mutagenesis of nuclear localization signals reveals the importance of neutral and acidic amino acids. *Current Biology* **6**: 1025–1027.

Mans BJ, Anantharaman V, Aravind L, Koonin EV (2004) Comparative genomics, evolution and origins of the nuclear envelope and nuclear pore complex. *Cell Cycle (Georgetown, Tex.)* **3**: 1612–1637.

Marchese E, Caterino M, Viggiano D, Cevenini A, Tolone S, Docimo L, Di Iorio V, Del Vecchio Blanco F, Fedele R, Simonelli F, Perna A, Nigro V, Capasso G, Ruoppolo M, Zacchia M (2022) Metabolomic fingerprinting of renal disease progression in Bardet-Biedl syndrome reveals mitochondrial dysfunction in kidney tubular cells. *iScience* **25**: 105230.

McClure-Begley TD, Klymkowsky MW (2017) Nuclear roles for cilia-associated proteins. *Cilia* **6**: 8.

Melnikov S, Ben-Shem A, Yusupova G, Yusupov M (2015) Insights into the origin of the nuclear localization signals in conserved ribosomal proteins. *Nature Communications* **6**: 7382.

Michie KA, Löwe J (2006) Dynamic Filaments of the Bacterial Cytoskeleton. *Annual Review of Biochemistry* **75**: 467–492.

Mitchell DR (2007) The Evolution of Eukaryotic Cilia and Flagella as Motile and Sensory Organelles. In: *Eukaryotic Membranes and Cytoskeleton: Origins and Evolution*, Advances in Experimental Medicine and Biology. Springer, New York, NY, pp. 130–140.

Morita E, Sandrin V, Chung H-Y, Morham SG, Gygi SP, Rodesch CK, Sundquist WI (2007) Human ESCRT and ALIX proteins interact with proteins of the midbody and function in cytokinesis. *The EMBO Journal* **26**: 4215–4227.

Mosalaganti S, Kosinski J, Albert S, Schaffer M, Strenkert D, Salomé PA, Merchant SS, Plitzko JM, Baumeister W, Engel BD, Beck M (2018) In situ architecture of the algal nuclear pore complex. *Nature Communications* **9**: 2361.

Nakayama K, Katoh Y (2018) Ciliary protein trafficking mediated by IFT and BBSome complexes with the aid of kinesin-2 and dynein-2 motors. *The Journal of Biochemistry* **163**: 155–164.

Neumann N, Lundin D, Poole AM (2010) Comparative Genomic Evidence for a Complete Nuclear Pore Complex in the Last Eukaryotic Common Ancestor. *PLOS ONE* **5**: e13241.

Obado SO, Brillantes M, Uryu K, Zhang W, Ketaren NE, Chait BT, Field MC, Rout MP (2016) Interactome Mapping Reveals the Evolutionary History of the Nuclear Pore Complex. *PLOS Biology* **14**: e1002365.

O'Reilly AJ, Dacks JB, Field MC (2011) Evolution of the Karyopherin-β Family of Nucleocytoplasmic Transport Factors; Ancient Origins and Continued Specialization. *PLOS ONE* **6**: e19308.

Ori A, Banterle N, Iskar M, Andrés-Pons A, Escher C, Khanh Bui H, Sparks L, Solis-Mezarino V, Rinner O, Bork P, Lemke EA, Beck M (2013) Cell type-specific nuclear pores: a case in point for context-dependent stoichiometry of molecular machines. *Molecular Systems Biology* **9**: 648.

Pickett-Heaps J (1974) The evolution of mitosis and the eukaryotic condition. *Biosystems* **6**: 37–48.

Piperno G, Mead K, Henderson S (1996) Inner dynein arms but not outer dynein arms require the activity of kinesin homologue protein KHP1(FLA10) to reach the distal part of flagella in Chlamydomonas. *The Journal of Cell Biology* **133**: 371–379.

Pollard VW, Michael WM, Nakielny S, Siomi MC, Wang F, Dreyfuss G (1996) A novel receptor-mediated nuclear protein import pathway. *Cell* **86**: 985–994.

Prevo B, Scholey JM, Peterman EJG (2017) Intraflagellar transport: mechanisms of motor action, cooperation, and cargo delivery. *The FEBS Journal* **284**: 2905–2931.

Pyke KA (1999) Plastid Division and Development. *The Plant Cell* **11**: 549–556.

Qin H, Burnette DT, Bae Y-K, Forscher P, Barr MM, Rosenbaum JL (2005) Intraflagellar transport is required for the vectorial movement of TRPV channels in the ciliary membrane. *Current biology: CB* **15**: 1695–1699.

Qin H, Diener DR, Geimer S, Cole DG, Rosenbaum JL (2004) Intraflagellar transport (IFT) cargo: IFT transports flagellar precursors to the tip and turnover products to the cell body. *Journal of Cell Biology* **164**: 255–266.

Raiborg C, Stenmark H (2009) The ESCRT machinery in endosomal sorting of ubiquitylated membrane proteins. *Nature* **458**: 445–452.

Ramakrishnan C, Maier S, Walker RA, Rehrauer H, Joekel DE, Winiger RR, Basso WU, Grigg ME, Hehl AB, Deplazes P, Smith NC (2019) An experimental genetically attenuated live vaccine to prevent transmission of Toxoplasma gondii by cats. *Scientific Reports* **9**: 1474.

Ray M, Tang R, Jiang Z, Rotello VM (2015) Quantitative tracking of protein trafficking to the nucleus using cytosolic protein delivery by nanoparticle-stabilized nanocapsules. *Bioconjugate chemistry* **26**: 1004–1007.

Robert A, Margall-Ducos G, Guidotti J-E, Brégerie O, Celati C, Bréchot C, Desdouets C (2007) The intraflagellar transport component IFT88/polaris is a centrosomal protein regulating G1-S transition in non-ciliated cells. *Journal of Cell Science* **120**: 628–637.

Rosenbaum JL, Witman GB (2002) Intraflagellar transport. *Nature Reviews Molecular Cell Biology* **3**: 813–825.

Rout MP, Aitchison JD, Suprapto A, Hjertaas K, Zhao Y, Chait BT (2000) The yeast nuclear pore complex: composition, architecture, and transport mechanism. *The Journal of Cell Biology* **148**: 635–651.

Sagan L (1967) On the origin of mitosing cells. *Journal of Theoretical Biology* **14**: 225-IN6.

Sampathkumar P, Kim SJ, Upla P, Rice WJ, Phillips J, Timney BL, Pieper U, Bonanno JB, Fernandez-Martinez J, Hakhverdyan Z, Ketaren NE, Matsui T, Weiss TM, Stokes DL, Sauder JM, Burley SK, Sali A, Rout MP, Almo SC (2013) Structure, Dynamics, Evolution, and Function of a Major Scaffold Component in the Nuclear Pore Complex. *Structure* **21**: 560–571.

Samson RY, Obita T, Freund SM, Williams RL, Bell SD (2008) A Role for the ESCRT System in Cell Division in Archaea. *Science* **322**: 1710–1713.

Sanderfoot A (2007) Increases in the Number of SNARE Genes Parallels the Rise of Multicellularity among the Green Plants. *Plant Physiology* **144**: 6–17.

Satir P, Guerra C, Bell AJ (2007) Evolution and persistence of the cilium. *Cell Motility* **64**: 906–913.

Satir P, Mitchell DR, Jékely G (2008) Chapter 3 How Did the Cilium Evolve? In: *Current Topics in Developmental Biology*, Ciliary Function in Mammalian Development. Academic Press, pp. 63–82.

Scott CA, Marsden AN, Rebagliati MR, Zhang Q, Chamling X, Searby CC, Baye LM, Sheffield VC, Slusarski DC (2017) Nuclear/cytoplasmic transport defects in BBS6 underlie congenital heart disease through perturbation of a chromatin remodeling protein. *PLOS Genetics* **13**: e1006936.

Shi L, Koll F, Arnaiz O, Cohen J (2018) The Ciliary Protein IFT57 in the Macronucleus of Paramecium. *Journal of Eukaryotic Microbiology* **65**: 12–27.

Smith JM, Szathmáry E (1997) *The Major Transitions in Evolution*. OUP Oxford.

Stenmark H (2009) Rab GTPases as coordinators of vesicle traffic. *Nature Reviews Molecular Cell Biology* **10**: 513–525.

Stuwe T, Lin DH, Collins LN, Hurt E, Hoelz A (2014) Evidence for an evolutionary relationship between the large adaptor nucleoporin Nup192 and karyopherins. *Proceedings of the National Academy of Sciences* **111**: 2530–2535.

Szathmáry E, Smith JM (1995) The major evolutionary transitions. *Nature* **374**: 227–232.

Tamura K, Fukao Y, Iwamoto M, Haraguchi T, Hara-Nishimura I (2010) Identification and Characterization of Nuclear Pore Complex Components in Arabidopsis thaliana[W][OA]. *The Plant Cell* **22**: 4084–4097.

Taschner M, Lorentzen E (2016) The Intraflagellar Transport Machinery. *Cold Spring Harbor Perspectives in Biology* **8**: a028092.

Tetenbaum-Novatt J, Rout MP (2010) The Mechanism of Nucleocytoplasmic Transport through the Nuclear Pore Complex. *Cold Spring Harbor Symposia on Quantitative Biology* **75**: 567–584.

Venard CM, Vasudevan KK, Stearns T (2020) Cilium axoneme internalization and degradation in chytrid fungi. *Cytoskeleton* **77**: 365–378.

Villeneuve AM, Hillers KJ (2001) Whence Meiosis? *Cell* **106**: 647–650.

Vitre B, Guesdon A, Delaval B (2020) Non-ciliary Roles of IFT Proteins in Cell Division and Polycystic Kidney Diseases. *Frontiers in Cell and Developmental Biology* **8**.

Wen W, Meinkoth JL, Tsien RY, Taylor SS (1995) Identification of a signal for rapid export of proteins from the nucleus. *Cell* **82**: 463–473.

Wickstead B, Gull K (2007) Dyneins Across Eukaryotes: A Comparative Genomic Analysis. *Traffic* **8**: 1708–1721.

Williams CL, McIntyre JC, Norris SR, Jenkins PM, Zhang L, Pei Q, Verhey K, Martens JR (2014) Direct evidence for BBSome-associated intraflagellar transport reveals distinct properties of native mammalian cilia. *Nature Communications* **5**: 5813.

Yoshizawa AC, Kawashima S, Okuda S, Fujita M, Itoh M, Moriya Y, Hattori M, Kanehisa M (2006) Extracting Sequence Motifs and the Phylogenetic Features of SNARE-Dependent Membrane Traffic. *Traffic* **7**: 1104–1118.

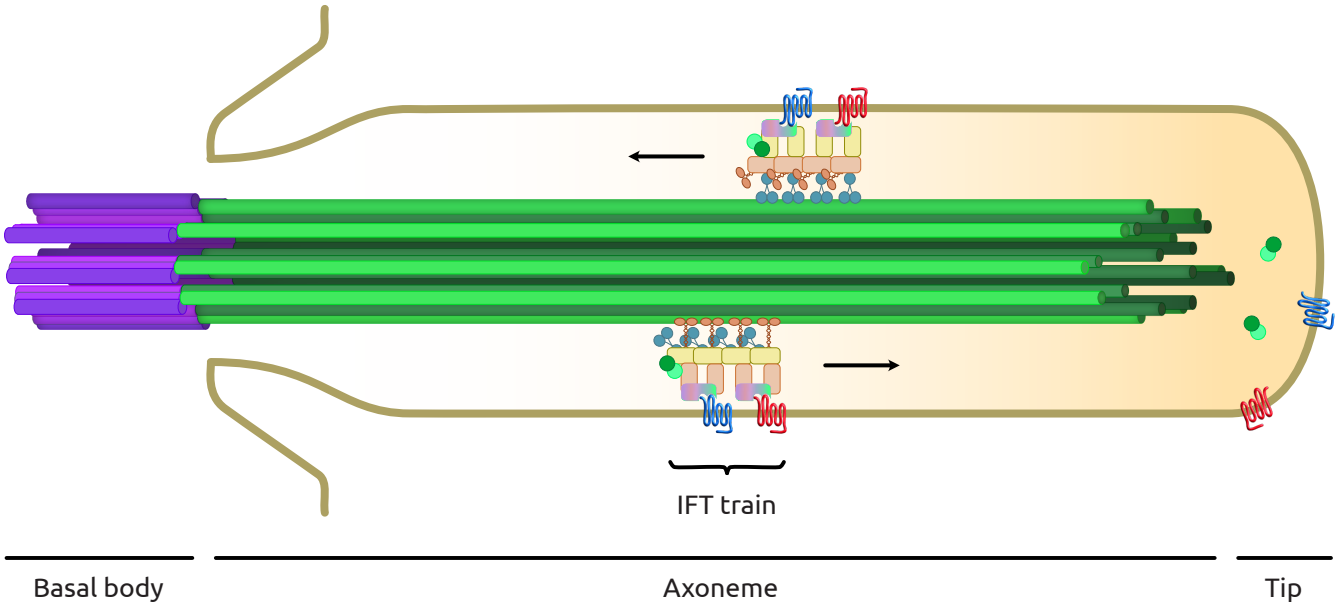
Figure legends

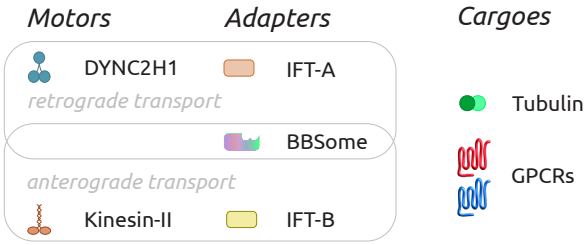
Figure 1: Overview of a eukaryotic cilium and intraflagellar transport machinery. Cargo is transported from the proximal basal body along the ciliary axoneme to the tip, where cargo is released, the IFT train re-structured and send back towards the basal body.

Figure 2: Overview of protocoatomer-derived vesicle coats and their field of action. While parts of the IFT/BBSome are derived from protocoatomer, they are not vesicle coats *sensu stricto*. They do however associate with membranes for cargo delivery.

Figure 3: Potential series of events that led to the current spectrum of membrane vesicle coats. α-solenoid and β-propellers are present in bacteria and archaea, but only separately. Lateral gene transfer and fusion events likely lead to the formation of the primordial α/β chimera that nucleated the protocoatomer.

Figure 4: Sequence of events leading to the ER, GA, cilium, and NE. In the post-FECA cell, the initial protocoatomer likely facilitated directed delivery from the plasma membrane to a proto-ER and *vice versa*. DNA was freely accessible due to a lack of NE structure. With the enrichment of receptors on a specific patch of the plasma membrane, the cell developed a primitive membrane protrusion, the proto-cilium, that further enhanced signalling capabilities. Pre-existing signalling cascades were inherited, novel pathways established, and regulated via intermediate regulatory levels. With the paralogous expansion of the protocoatomer, the cell was able to differentiate between inwards and outwards transport, and to establish an intermediate compartment for sorting, the proto-GA. The type I coats used for inwards transport could also be co-opted by the developing proto-cilium. After further diversification of type I and type II coats, the NE could form with primitive pores, limiting free diffusion of ‘ciliary’ signalling effector proteins into the forming NE. However, some of the ciliary proteins still could enter (either due to size or presence of advantageous protein structures), while others relied on intermediate messengers that could. At the time the NE with NPCs was fully formed, some signalling pathways must have been lost. Some pathways made use of other proteins that could enter, and others retained their nuclear functions fully.


IFT train
Basal body
Axoneme
Tip


Motors
Adapters
DYNC2H1
IFT-A
retrograde transport
BBSome
anterograde transport
Kinesin-II
IFT-B

Cargoes
Tubulin
GPCRs

Figure 1

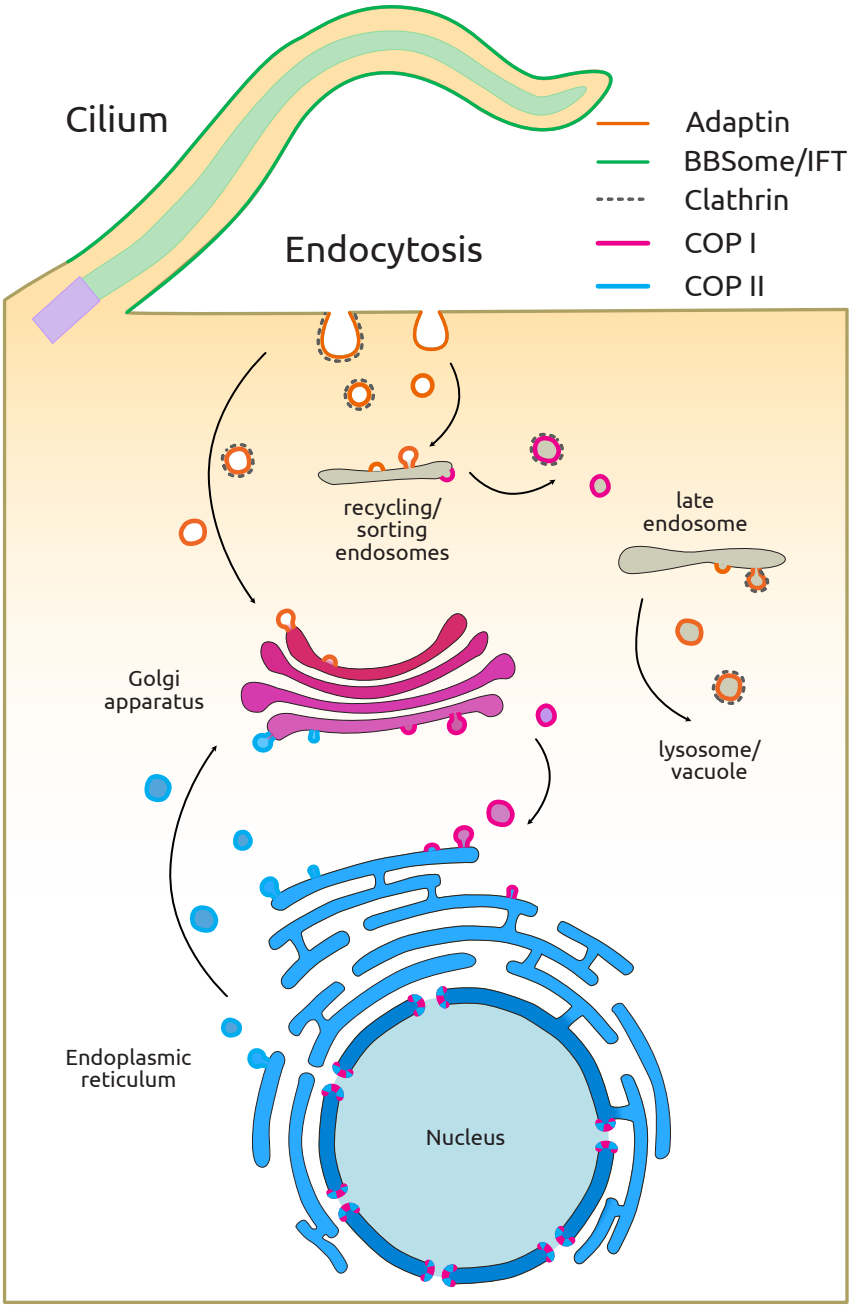

Cilium
Endocytosis
Adaptin
BBSome/IFT
Clathrin
COP I
COP II
recycling/
sorting
endosomes
late
endosome
Golgi
apparatus
lysosome/
vacuole
Endoplasmic
reticulum
Nucleus

Figure 2

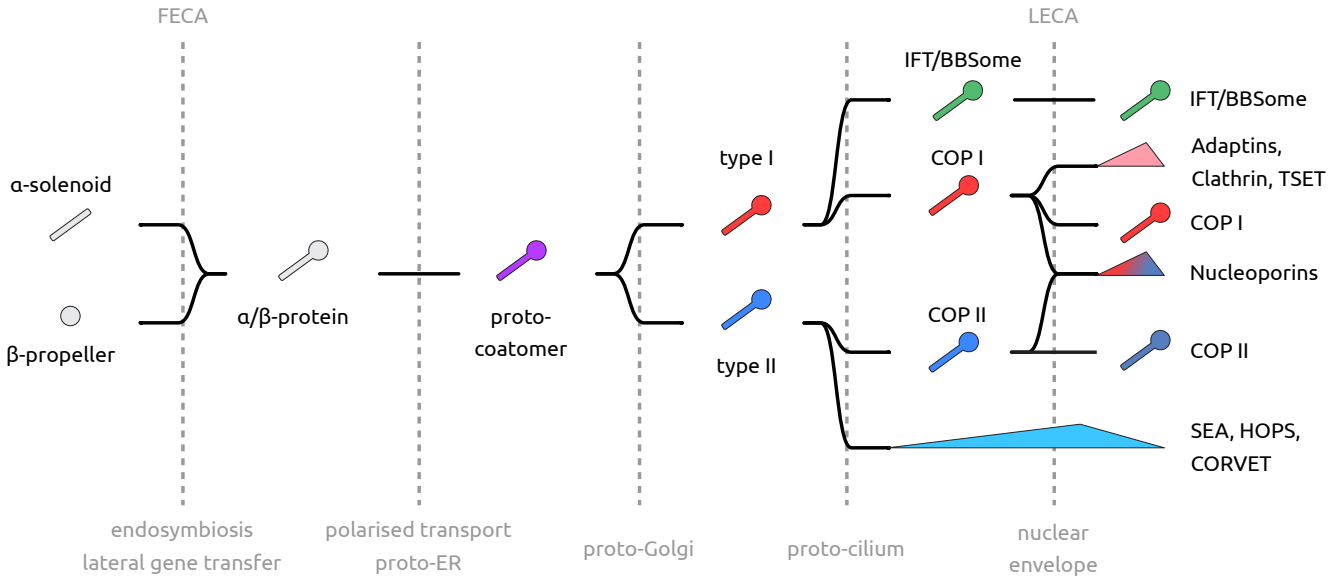

FECA
LECA
α-solenoid
β-propeller
α/β-protein
proto-
coatomer
type I
type II
IFT/BBSome
COP I
COP II
IFT/BBSome
Adaptins,
Clathrin, TSET
COP I
Nucleoporins
COP II
SEA, HOPS,
CORVET
endosymbiosis
lateral gene transfer
polarised transport
proto-ER
proto-Golgi
proto-cilium
nuclear
envelope

Figure 3

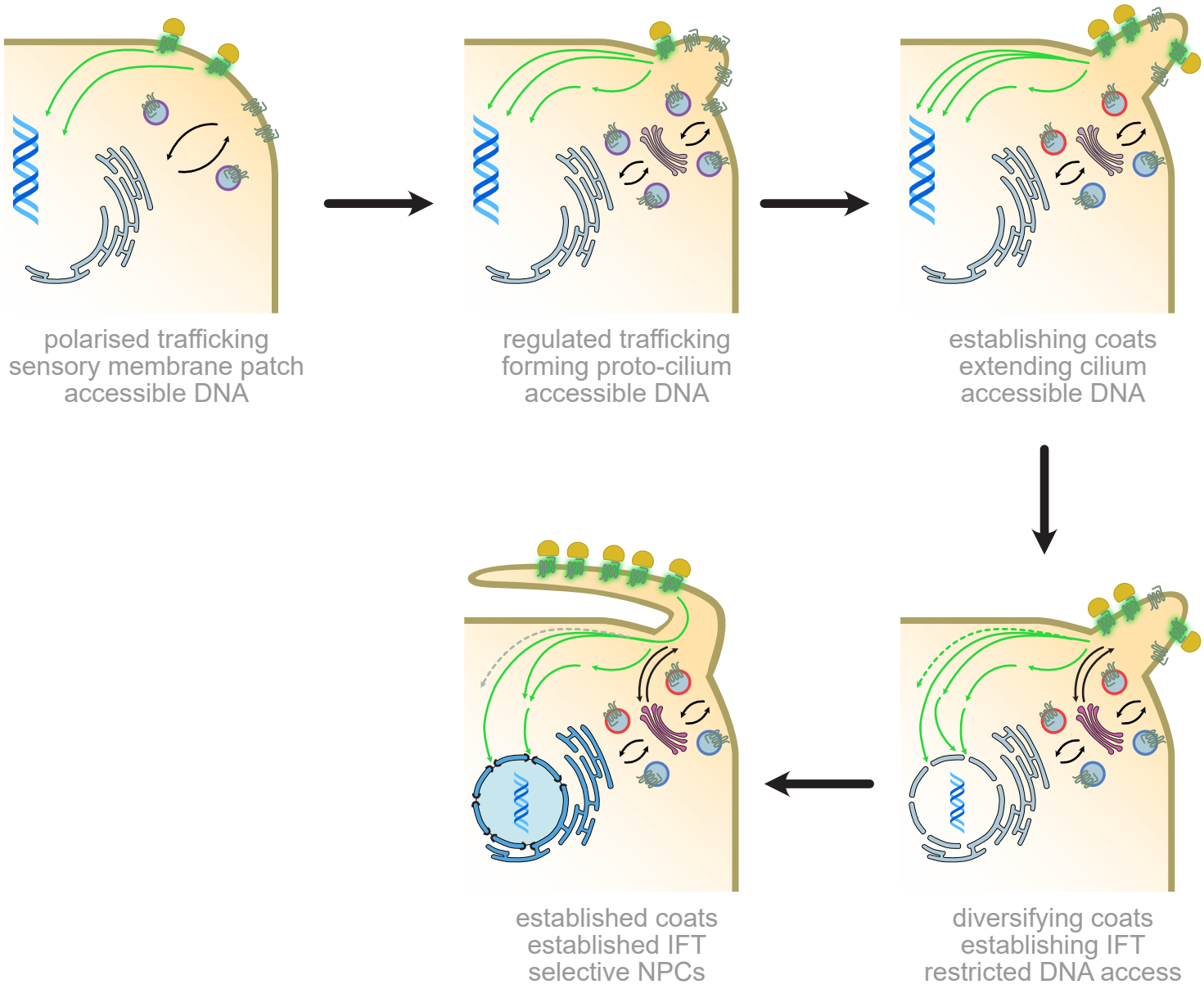

polarised trafficking
sensory membrane patch
accessible DNA
regulated trafficking
forming proto-cilium
accessible DNA
establishing coats
extending cilium
accessible DNA
diversifying coats
establishing IFT
restricted DNA access
established coats
established IFT
selective NPCs

Figure 4

5. DISCUSSION

Cilia play crucial roles in numerous eukaryotic cells. Their proper function is vital for both single-celled organisms and complex multicellular life forms, contributing to activities like movement, sensory perception, and growth. The newfound extraciliary functions of ciliary proteins in mammalian cells expand their influence beyond their original ciliary duties. This prompts the question of whether these functions remain consistent in other eukaryotes, potentially as ancestral traits, or if they were adopted independently through convergent evolution. To address this, we embarked on reconstructing the evolutionary trajectory of the BBSome and chaperonin-like BBS proteins. We identified orthologues and predicted their nuclear localisation using bioinformatic signal sequence prediction. Additionally, we harnessed the research potential of non-model organisms such as insects to illuminate the gene regulatory and nuclear roles of BBS proteins. By contextualising current research findings within the framework of eukaryotic evolution and the transition from prokaryotes to eukaryotes, we aimed to elucidate the reasons behind the presence of ciliary proteins in the nucleus.

5.1. CONSERVATION OF BBS PROTEINS IN EUKARYOTES

Our analysis of the distribution of BBSome components across different eukaryotic organisms validates the widespread occurrence of these proteins and their strong evolutionary connection with the presence of cilia (Ewerling et al., 2023b; Hodges et al., 2010). Additionally, our investigation indicates that the more limited chaperonin-like BBS proteins are not exclusive to Metazoa; they also appear in amoebae and early-branching organisms like *Malawimonas jakobiformis*. Contrary to earlier studies (Mukherjee and Brocchieri, 2013), our findings do not replicate the presence of BBS6-like proteins in oomycetes. This suggests that the chaperonin-like BBS proteins are unique to the Opimoda lineage of eukaryotes.

We observed a consistent presence of BBSome proteins as a set. This set encompasses at least one representative from each subset of BBSome proteins, specifically the BBS1-like, TPR-domain-containing, and PH-domain-containing families (Fig. 8). There are exceptions to this pattern, particularly in lineages where the most recently diverging species have undergone a gradual reduction in ciliary protein complexes, and eventually lost the cilium altogether. For instance, in *Phytophthora*, although *P. ramorum* lacks TPR proteins BBS4 and BBS8, as well as the BBS1-like protein BBS2, its closely related species *P. infestans* retains all three (van Dam et al., 2013). This reduction in the protein set might indicate a gradual loss of cilia in stramenopiles, exemplified by

Phaeodactylum's complete loss of cilia and associated proteins. A similar trend of reduction and loss is evident in other clades; for instance, *Batrachochytrium dendrobatidis*, an early branching fungal representative, still possesses a cilium but lacks the BBSome complex (Ewerling et al., 2023b). This loss of BBSome proteins comes with a non-canonical cytoplasmic ciliogenesis (Venard et al., 2020), a trend that aligns with cases like *Plasmodium* and *Toxoplasma*, two alveolates that assemble cytosolic cilia without needing a functional BBSome (Killick-Kendrick and Peters, 1978). Another example of BBSome reduction is *Drosophila melanogaster*, where all BBSome components are present but BBS2 and BBS7 (Ewerling et al., 2023a). A deeper look into other insect genomes, transcriptomes and proteomes however revealed that *Drosophila* is atypical in its BBSome composition. While no other queried insects possess homologues to chaperonin-like BBS proteins, it appears that the BBSome itself is highly conserved, apart from a BBS1 homologue in honeybees (Ewerling et al., 2023a). The impact of *D. melanogaster* as a model organism for developmental and neuronal biology remains undisputed (Jennings, 2011), however for cilia biology, the organism is less well suited as it lacks two of eight BBSome components. Other insects may serve as more representative examples to research ciliary proteins in ciliary and non-ciliary contexts.

Terrestrial fungi and *Cryptosporidium*, which no longer develop cilia, also lack BBSome protein counterparts, except for a notable match in BBS5 found in *Toxoplasma gondii*. RNAseq data validates the expression of this orthologue, primarily during a life-cycle stage known to be non-ciliated (Ramakrishnan et al., 2019). Given that certain BBSome components might be dispensable, and their loss precedes the gradual loss of the entire complex and subsequently cilia, a question arises regarding why *Toxoplasma* retains a single BBS protein. This situation could indicate a potential residual function for BBS5 in *Toxoplasma*'s cytosolic ciliary assembly or, which is more likely, suggests that the protein serves a non-ciliary role.

Further investigation of the identified orthologs through sequence analysis yielded insights into the distribution of nuclear localisation or export signals (NLS and NES, respectively). It became evident that predicted NLS signals manifest in a binary manner, showing an "all-or-nothing" trend, while NES signals exhibit a more dispersed pattern across various prediction scores. Among the BBS proteins, BBS7 exhibited the highest likelihood of receiving high NES scores followed by BBS1 and BBS8, whereas BBS4 and BBS9 showed the highest probability of obtaining high NLS scores (Ewerling et al., 2023b). To determine whether there is a preference for stronger nuclear signal sequences during open or closed mitosis, a permutation test was conducted. The results showed no significant increase in the likelihood of NLS or NES presence in BBS proteins during

either form of mitosis. This implies that the nuclear localisation of BBS proteins likely did not co-evolve with the transition to a specific form of mitosis; instead, it seems to have emerged as a recent adaptation in particular lineages. Combined with the potential co-evolution of ciliary proteins and the NE, it would be fitting that the nuclear localisation of BBS proteins was established before any form of closed mitosis (due to the lack of a functional NE).

When overexpressed in human cells, BBS1, BBS2, BBS7, BBS9, BBS4, BBS8, and BBS5 were also detected in the nucleus. This nuclear localisation likely arises from an active transport mechanism between the cytosol and nucleus. Notably, the nuclear pore's size-exclusion barrier restricts the free diffusion of molecules with a mass exceeding 40 kDa, a criterion met by all nuclear-localising BBS proteins except for BBS5. Interestingly, despite lacking classical NES, BBS5 is detectable in the nuclear fraction and is also evident in immunocytochemical staining. The presence of two pleckstrin-homology-like domains in BBS5 could theoretically facilitate interaction with nuclear phosphoinositides, which have been implicated in chromatin remodeling and transcriptional regulation (Keune et al., 2011), further pointing towards gene regulatory functions for BBS proteins in the nucleus.

5.2. NUCLEAR FUNCTIONS OF BBS PROTEINS

Regulation of differential gene expression during development of tissues is a novel role for BBS proteins that has only been described recently (Gascue et al., 2012; Scott et al., 2017). But BBS proteins are also involved in short-term responses to more immediate environmental cues. Recent studies have shown that BBS proteins are recruited for translocation of transcription factors after application of cytotoxic stress: BBS4, which is normally localised at centrosomes in cycling cells or the basal body in quiescent cells, translocates from cytosol to nucleus after induction of ER stress (Horwitz and Birk, 2021). Here, BBS4 associates with transcription factors that initiate the unfolded protein response (UPR). This is reminiscent of a similar mechanism seen in BBS6 that is disturbed in McKusick-Kaufman Syndrome (MKKS), where a mutation disrupts the shuttling of BBS6 in and out of the nucleus, enriching SMARCC1 in the cytoplasm. Unpublished work from our group also showed that induction of cyto- and genotoxic stress induces a translocation of BBS4, BBS6, and BBS8 from cytoplasm to nucleus (Patnaik et al., in prep.). Therefore it seems like BBS proteins have a pivotal role in both developmental, long-term regulation of genes, and short-term responses to environmental cues by associating with transcription factors and effector proteins and ensuring cytonucleoplasmic translocation.

Individual BBS proteins can enter the nucleus – at least in transiently transfected human cells (Ewerling et al., 2023b). Not all BBS proteins localise to the nucleus to the same extent or in the same manner (Ewerling et al., 2023b), indicating that their entry and exit from the nucleus is differentially regulated, probably because their nuclear funtions differ and are independent of each other. The study confirmed previous work showing BBS7 localisation to the nucleus in mammalian cell lines (Gascue et al., 2012). There, BBS7 was implicated in transcriptional regulation by interaction with E3 ubiquitin ligase RING2 (RNF2). RNF2 itself regulates histone-monoubiquitylation on genes involved throughout development (Cao et al., 2005). BBS7 therefore might not only regulate developmental processes via its function in the BBSome and the signalling hub that is the primary cilium, but also by regulating chromatin-associated proteins. In agreement with this, target genes of RNF2 were misregulated upon depletion of BBS7 (Gascue et al., 2012) and in our data, BBS7 was also mostly localised at the nuclear membrane (Ewerling et al., 2023b). The nucleoplasmic face of nuclear pores is known as a hotspot for transcriptional regulation (Casolari et al., 2004; Ishii et al., 2002), concomitant with chromatin remodelling that is typically seen after histone modifications. The predominant localisation of BBS7 and BB8 could be an indication that they are involved in either chromatin-remodelling complexes such as the RNF2-containing polycomb repressive complex 1 (PRC1) or a direct interaction with transcriptional regulators that are typically found at sites of active transcription, such as transcription factors or RNA polymerases.

Another process linked to the inner nuclear membrane and the nuclear pore complexes in particular is the export of nascent mRNAs. During transcription, mRNAs with bound proteins (mRNPs) are exported through NPCs. This is facilitated (among others) by the TREX2-complex (Stewart, 2019, 2007) located at the nucleoplasmic site of NPCs, the nuclear basket (Umlauf et al., 2013). As BBS8 seems to localise to the nuclear periphery, it is possible that it interacts with parts of the TREX2 machinery. TREX2 is well conserved from yeast to mammals, so a regulatory role across eukaryotes is not unlikely, given the conserved character of BBS proteins themselves. BBS proteins' NLS and NES could act in two ways: on one hand, their nuclear signal sequences could aid in importing parts of regulatory complexes into the nucleus, while on the other hand, having an NES could recruit exportins to mRNP complexes to facilitate cytoplasmic translocation. Different patterns in nuclear staining for BBS proteins (Ewerling et al., 2023b) indicate that not all BBS proteins take part in the same functions or complexes. This might be a crucial point of spatiotemporal regulation targeting different mRNPs at different points throughout the cell cycle

and organismal development. Many targets of RNF2 require a precisely regulated on-off-switch: RNF2 regulates X chromosome inactivation in mammals (Almeida et al., 2017; de Napoles et al., 2004), differentiation of germ layers (Morey et al., 2012), and modulation of HOX gene expression through histone modifications (Buchwald et al., 2006; Tavares et al., 2012).

Our data consolidates that BBS proteins can have roles outside of ciliary trafficking. However, as not all BBSome proteins localise to the nucleus and those that do have distinct patterns (Ewerling et al., 2023b), it is likely that BBS protein involvement in nuclear processes is independent of their function as a complex. This nuclear localisation does not appear to be the result of the BBSome (or individual proteins) being coincidentally incorporated into the nucleus due to evolution towards open mitosis, but is equally likely in all eukaryotic species possessing BBS proteins. This leaves two options for the evolution of BBS proteins' nuclear function: Given that BBS proteins are found across all eukaryotic taxa, even on occasion in species without classical ciliogenesis, this might place the root of nuclear functions of BBS proteins early during the radiation of eukaryotes, potentially even at or before LECA. This scenario is likely for BBS7, as we could show through ancestral state reconstruction that the hypothetical LECA orthologue bears a strong NES, likely making it nuclear (Ewerling et al., 2023b). Another explanation is an independent convergence scenario where conserved BBS proteins filled an ecological niche for gene regulation on numerous occasions. This is the case for BBS4 and BBS9, where reconstructed sequences at the base of Choanozoa and in *Thecamonas trahens* in Opimoda, and *Monocercomonoides* in Diphoda have NLSs, but without direct connections to each other. Either theory can be elucidated by investigation of their nuclear interactome and gene regulatory functions in different eukaryotic species to conduct comparative interactomic studies. This information is crucial to understand the evolution of the cilium, nucleus and the origins of regulatory networks in early eukaryotes that have potentially persisted until today.

5.3. Non-model organisms for non-ciliary functions

Cilia biology has always used many different model organisms to elucidate the structure and function of eukaryotic cilia. Earliest studies have been conducted in the single-celled green alga *Chlamydomonas reinhardtii* in the mid of the 20th century, and have led to the discovery of universal ciliary properties such as IFT, ciliary assembly and disassembly, and signalling functions (Bernstein, 1995; Dutcher, 1995; Harris, 2009; Lechtreck et al., 2009; Lechtreck and Geimer, 2000; Pröschold et al., 2005). Scientists have since used a smörgåsbord of ciliated species to answer the

specific research question at hand, and did so to great avail (Vincensini et al., 2011), revealing lineage-specific intricacies to cilia biology. However, researching non-ciliary functions of ciliary proteins puts researchers in a quandary: using ciliated species such as *Chlamydomonas*, *Tetrahymena*, or Choanoflagellates comes with the advantage of guaranteed presence of ciliary proteins; the downside being that these proteins will mainly be localised at the cilium or BB. Studying ciliary proteins outside their ciliary context requires novel approaches to proteomic and interactomic analyses. While existing model system can be manipulated in order to obtain pure organelle-specific protein samples through subcellular fractionation, the process must be adjusted to fit the biochemical properties of each organism. Existing model organisms can be used in such a way, but broader phylogenetic representation is needed to make confident statements about inherited phenotypes and functions. In order to find novel organisms for researching non-ciliary functions of ciliary BBS proteins, we chose insects as a potential clade.

Insects provide the advantage of expressing cilia only in two tissues, neurons and sperm cells (Keil, 2012). This restriction allows for separation of ciliated and non-ciliated tissue, and simultaneously guarantees the presence of ciliary genes in the organism. Earlier studies (Ewerling et al., 2023b; Hodges et al., 2010) have used the fruit fly *Drosophila melanogaster* and the honeybee *Apis mellifera* as well-studied representatives of the insect clade (Grozinger and Zayed, 2020; Jennings, 2011; Trapp et al., 2017; Weinstock et al., 2006). However, both species lack BBS proteins conserved in other species, and therefore offer limited transferability of interactomic studies to other eukaryotes. Broader sampling and phylogenetic reconstruction of BBS homologues in insects has revealed that both *D. melanogaster* and *A. mellifera* are unique in the way that they assemble “incomplete” BBSomes (Ewerling et al., 2023a, 2023b; Hodges et al., 2010), and that other early-branching insects do indeed possess homologues to all BBS proteins.

BBS proteins have been implicated in the establishment of a social parasitism phenotype in Cape honeybees (Wallberg et al., 2016) and could play a role in differential gene regulation, resulting in distinct phenotypes. Honeybees establish social hierarchy and labour distribution within a caste via differential gene expression. While we could not identify a BBS1 homologue, we found BBS4, 5, 7 and 8 being most differentially regulated between queens (female) and drones (male) (Ewerling et al., 2023a), all of which have been implied to potentially fulfil nuclear roles (Ewerling et al., 2023b; Gascue et al., 2012; Horwitz and Birk, 2021). Single proteins of the BBSome complex are higher expressed in the brain, hinting at an independent role from the BBSome. Further subsets show that BBS8 is significantly enriched in drone brains compared to queen brains, highlighting its potential

role in gene regulation. Within females, BBS4 shows highest expression in antennae and glands. Antennae are enriched with neuronal olfactory receptors and are using in intraspecies communication across insects (Sharma et al., 2015), meaning signal transduction likely happens through cilium-mediated signalling cascades. The presence of one, but not all BBSome proteins implies an elevated role for this protein in signal transduction in particular. At the same time, BBS4 is also higher expressed in the hypopharyngeal, mandibular, sting, and Nasonov glands. Apart from the sting gland, they serve communication functions between individuals, and actively shape the establishment of social castes in and behaviour of the colony (Ahmad et al., 2021; Kamakura, 2011). Clearly, single BBS proteins perform crucial functions in honeybees, either through their functions in ciliary maintenance, or as single proteins to a currently unknown extent. Given that there is an extensive toolkit for molecular manipulation of insect cells (Ikonomou et al., 2003) and that insects are excellent models for social and behavioural studies, the thought of “adopting” insects as a model for non-ciliary functions of ciliary proteins is certainly worth entertaining.

5.4. A co-evolution scenario for nuclear cilia proteins

One of the major transitions of life has been the development of a sophisticated, compartmentalised endomembrane system that probably contributed to the success of eukaryotes (Smith and Szathmáry, 1997). While the “original” nuclear membrane was likely freely permeable, the nuclear envelope has since evolved into a tightly regulated, selective transport barrier (Field et al., 2014; Hoelz et al., 2011; Wilson and Dawson, 2011). During interphase, proteins and nucleic acids must pass the NPCs by association with importin and exportin-class proteins. In some organisms (including green plants and animals), the nuclear envelope completely breaks down during mitosis, while in others (such as yeasts), the envelope remains intact. The mitotic NE breakdown is however not binary. Many eukaryotes partially break down the NE, allowing microtubules to enter the nuclear space through fenestrated openings, the extent of which is dependent on the species (Fig. 9). The breakdown of the nuclear envelope may serve as an entry point into a compartment that is otherwise inaccessible without a NLS. On the other hand, in an organism with a closed form of mitosis, all nuclear proteins must be actively transported into the nucleus. We reasoned that if nuclear roles for BBS proteins are conserved, there might be a higher need for NLS in BBS orthologues in species with a closed form of mitosis and, conversely, a potential increase of NES in BBS orthologues in species with an open mitosis. Our permutation analysis (Ewerling et al., 2023b) showed no statistical bias towards NLS or NES signals in association with mitotic mode. Hence the appearance of nuclear signal sequences is not coincidental with the evolution of a specific mitotic

pattern. However, while we predicted mostly NES signals for human BBS proteins, we still find nuclear localisation via both subcellular fractionation and immunofluorescent staining which seems counter-intuitive at first glance. The discrepancy between predicted nuclear localisation (and mainly nuclear export) comes from the bioinformatic models used to make the predictions. Our analysis was based on prediction of classical, basic NLS signals (Nguyen Ba et al., 2009). Computational NLS predictors only recognise a relatively low percentage of true NLSs (37% for NLStradamus against a dataset of non-yeast sequences) with good confidence. This means however that many NLSs are not recognised, and non-classical (e.g. PY-NLS) or protein-specific NLSs are not covered at all. Computational analysis alone is therefore not sufficient to fully answer the evolutionary question of nuclear BBS functions and highlights the need for molecular biological confirmation in species of interest. To trace back the origin of obviously functional nuclear signal sequences in "ciliary" BBS proteins, it is worth highlighting the similarities between nuclear and ciliary import systems and to look at their evolutionary history.

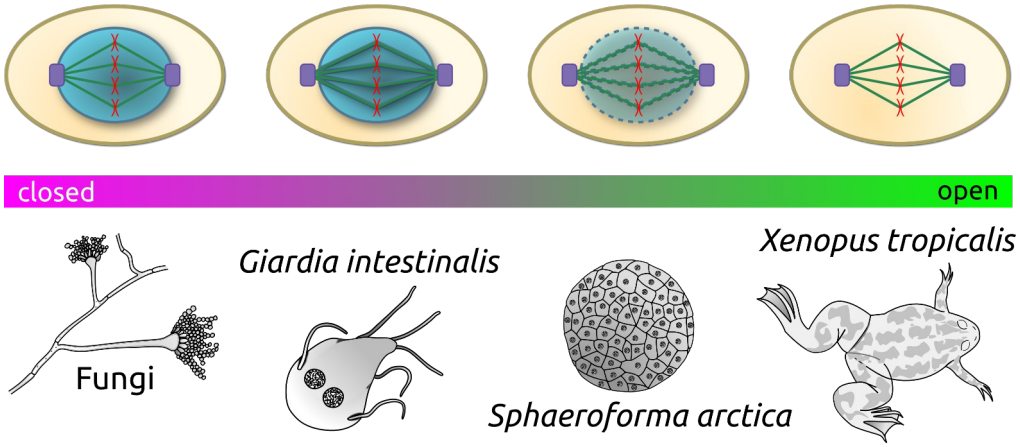


Figure 9: Spectrum of nuclear envelope breakdown during mitosis with representative species. The nuclear envelope (NE) behaves dynamically during mitosis in different organisms: in closed mitosis, the NE remains intact, and spindle poles are anchored within the NE, as seen e.g. in fungi. Other species have small or large fenestrations in the NE for microtubules to enter the nucleoplasm. Complete NE breakdown during mitosis can be observed e.g. in animals and plants. Adapted from Ewerling et al, 2023b.

5.4.1. THE CILIARY PORE – A FOSSIL OF CO-EVOLUTION?

Nuclear pore complexes and intra-organellar transport vesicle coats, including components of the IFT and BBSome are likely derived from a hypothetical protocoatomer complex that must have emerged early in eukaryotic evolution (Obado et al., 2016). These ancient coating complexes first aided in bending and stabilising primitive membrane invaginations and, through duplications, diverged to serve the specialised functions we see today (Field et al., 2011; Field and Rout, 2019). As components from both BBSome and IFT complexes share a common ancestor with type I family

coat proteins (van Dam et al., 2013) and the latter are found in endomembrane organelles, IFT must have at least emerged concomitantly with the establishment of a proto-ER and/or proto-GA. The proteins of NPCs are structurally related to both COPI- and COPII-family proteins (Devos et al., 2006, 2004), and are therefore also most likely a rather late addition to the “pre-karyotic” cell, i.e. after diversification of membrane coats from the protocoatomer. However, nucleoporins can be found at the base of the cilium, regulating ciliary import in a similar fashion as they do at nuclear pores (Kee et al., 2012). Nup98 (Endicott and Brueckner, 2018), Nup205 (Marquez et al., 2021), Nup62 (Kee et al., 2012), Nup93, Nup98 and Nup188 (del Viso et al., 2016; Endicott et al., 2015) as well as Nup85 (Endicott and Brueckner, 2018) can all be found at the ciliary base of cilia, interacting with the NPHP module for ciliary cargo sorting and playing crucial roles for mammalian left-right-axis determination (Blasius et al., 2019; Endicott et al., 2015; Heinß et al., 2020). This led to the proposition of a “ciliary pore complex” (CPC) akin to the nuclear pore complex, albeit with a distinct architecture (Kee and Verhey, 2013; Takao and Verhey, 2016) (Fig. 10). While the NPC is anchored to the NE via interaction between membrane ring Nups and other NPC components, cilia only provide one surface with a lipid bilayer that could potentially anchor the CPC to a membrane. However, the cargo sorting NPHP module occupies the membrane-facing periciliary space in the transition zone (Sang et al., 2011; Wang et al., 2022) and was found to interact with Nups (Blasius et al., 2019). Whether Nups are also anchored to the axonemal MTs and form a “full” NPC/CPC, or if they decorate the inner perimeter of the cilium, forming a “brush border”, is curently unknown and will require further interactomic studies and high-resolution imaging to fully unravel the topology of Nups at the ciliary base.

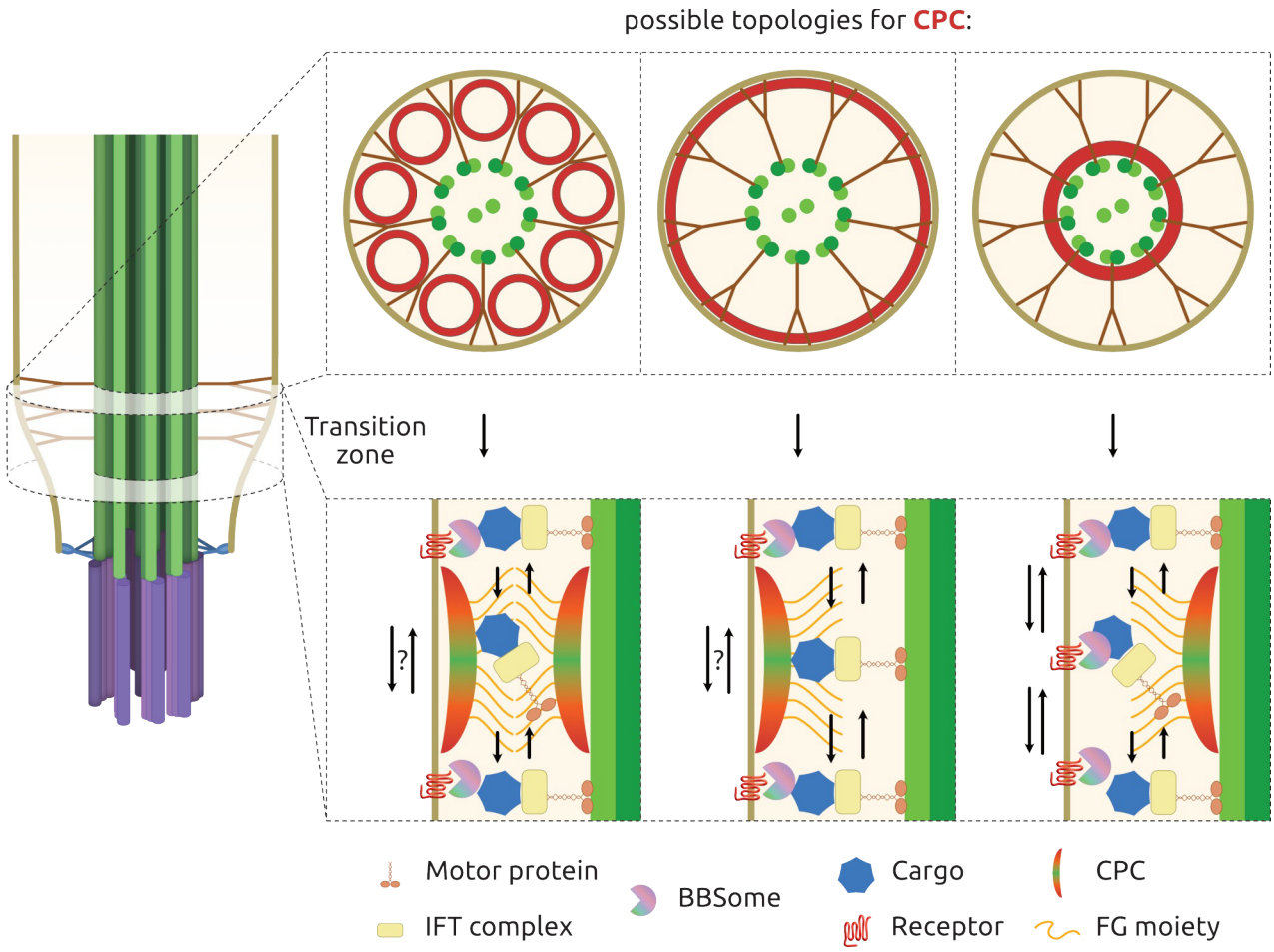


Figure 10: The ciliary pore complex. Although nucleoporins localise to the base of eukaryotic cilia, and they convey selectivity similar to their function in NPCs, the actual structure of the ciliary pore complex (CPC) is not yet resolved. Interactions suggest that the CPC sits in the transition zone and coordinates selective passage of proteins through interaction with other transition zone modules, but three constellations are theoretically possible. Adapted from Kee and Verhey, 2013, and Takao and Verhey, 2016.

This leaves three possible options for the appearance of ciliary proteins at and in the nucleus: Either I) the cilium developed before the NE, making the ciliary functions ancestral and the nuclear functions co-opted; II) the NE emerged first, and the close association of the cilium-nucleating centriole led to nuclear proteins “hitch-hiking” to the cilium and adopting ciliary roles; or III) there was a substantial period of time where cilium and nucleus co-evolved, with both partitions being rather “leaky” before diffusion barriers (transition zone in cilia, NPCs ind nuclei) were completely in place. For several ciliary cargoes, including crucial anterograde IFT protein KIF17 (Dishinger et al., 2010), the “nuclear” translocation system with importins, RanGDP/GTP gradient and NLS could be confirmed, but others seem to bypass the need for some parts of the system. Interestingly, primary cilia do not have any nucleoporins at the base (Breslow et al., 2013). While there are differences in nuclear and ciliary import and export, it is still possible that both systems once shared

a common ancestral machinery. Given the close phylogenetic relationship between membrane coats and the IFT/BBSome proteins that derived from the protocoatomer, the latter co-evolution scenario would explain best how predominantly ciliary proteins could withstand the evolutionary pressure of a selective NPC without acquiring nuclear functions *de novo*. However, as the proteins it is possible for them to gain nuclear functions, as seems to be the case for BBS4 and BBS9 (Ewerling et al., 2023b). Future analysis comparing the divergence of ciliary vesicle and cargo coat complexes such as the BBSome from nucleoporins and the ancestral protocoatomer needs to be conducted to untangle their shared evolutionary history.

6. Impact of this work and outlook

Cilia have fascinated researchers ever since their discovery by Antoni van Leeuwenhoek in 1675. They were the first organelle to be described, and their structure and functions are still actively being researched today. Their function as cellular motility and signalling apparatus remains conserved from single-celled protists to multicellular organisms. It was only recently that they became relevant to clinicians and cell biologists alike as crucial organelles for homeostasis and development in a pathological context, and the interest in these organelles has never been higher. The proteins underlying human diseases known as ciliopathies are conserved, and studying them outside their pathological context has brought insights into cell biology and evolutionary biology. Recent studies could show that ciliary proteins are however not restricted to their ciliary compartment, and that they play roles in various other cell biological contexts. These advances in ciliary research play a crucial role in understanding the disease phenotypes of ciliopathies as they can help to explain the phenotypic plasticity that can be observed in patients. Gaining insights into the interactomes of ciliary proteins outside of cilia opens the way for novel treatment strategies that may help alleviate the symptoms that are not directly linked to dysfunctional cilia, but rather their gene regulatory roles.

With this work, we employ state-of-the-art bioinformatic analyses to identify BBSome components as part if the ciliary IFT machinery across a broadly sampled set of eukaryotic species. We confirm previous studies and additionally identify single BBS proteins that must serve non-ciliary functions. Our in-depth analysis of insects as potential models for research of non-ciliary functions of BBS proteins revealed them as a promising clade for future behavioural and interactomic approaches. We found that BBS proteins are more conserved in insects than previously thought. Using molecular biology tools, we consolidated the nuclear presence of BBS proteins in human cells, showing that

BBS proteins are far more often nuclear than was expected from literature. Finally, we construct a conceivable scenario of early eukaryotic evolution that allowed “ciliary” proteins to co-evolve with the nucleus, allowing proteins of both organelles to adapt to dual functions, and leading to the gene regulatory roles of BBS proteins we see today.

Current model organisms have provided and will continue to provide a valuable resource for ciliary biology. However, we could show that in some respects, not all models are as suitable for specific research questions, e.g. nuclear functions of ciliary proteins. With our work, we could identify several other organisms that could prove to be valuable: insects could also be used as a suitable model for non-ciliary, gene regulatory functions of ciliary proteins, as their behaviour and social hierarchy are heavily influenced by differential gene expression. Our work shows that BBS genes are significantly differentially expressed between sexes and certain tissues, and that they could influence sociality in honeybees specifically. We also identified Discicristates as a potential model for non-ciliary BBS functions as their BBS orthologues contain above-average counts of nuclear signal sequences. Discicristates encompass many eukaryotic pathogens such as Trypanosomatida, and elucidating non-ciliary functions of BBS proteins in these organisms also helps advance the research into treatments against other disease-causing agents.

Our bioinformatic work and its corroboration in human cell lines have unearthed highly interesting patterns for nuclear localisation of BBS proteins. By ancestral state reconstruction, we were not only able to infer which of the BBS proteins are expected to show nuclear phenotypes can be expected in other eukaryotes, but we could also pinpoint which proteins likely had a recent gain of nuclear function, and which protein has an ancestrally conserved nuclear phenotype. The findings can be experimentally validated by expressing the reconstructed sequences in basal eukaryotes and assessing the localisation biochemically, and could also be refined by including more species and expanding the orthologue searches.

This work provides the basis for further studies of the nuclear functions of BBS genes in a variety of eukaryotic organisms. We could show that the nuclear phenotype of BBS proteins is more prevalent than previously anticipated, and that single BBS proteins stand out in terms of conservation. However, as the localisation pattern differs between proteins (and potentially between species), the crucial next step is now to analyse the nuclear interactome of BBS proteins. Different BBS proteins have been shown to regulate different gene expression pathways, some of which in ways that are also highly conserved in eukaryotes. It is now of interest to compare specifically the

nuclear interactomes of different BBS proteins in different species to see if not only the localisation, but also the gene regulatory functions are conserved across eukaryotes, and if BBS proteins play roles in regulating the response to similar external influences, e.g. in DNA damage response or by influencing DNA methylation patterns.

With sequencing technology becoming cheaper and more reliable, transcriptomics and proteomics becoming more and more sensitive, and increasingly interdisciplinary research, the field of cilia biology is on the brink of a revolution. The acknowledgement of ciliary proteins outside of cilia and their apparent impact on a plethora of biological mechanisms from cargo transport to cell cycle regulation and gene regulation has transformed cilia research from either a molecular/developmental biology or evolutionary biology discipline to a highly dynamic, integrative science incorporating methods and insights from both worlds, resulting in a new, syncretic approach to pressing research questions from both fields.

References

Adl, S.M., Bass, D., Lane, C.E., Lukeš, J., Schoch, C.L., Smirnov, A., Agatha, S., Berney, C., Brown, M.W., Burki, F., Cárdenas, P., Čepička, I., Chistyakova, L., del Campo, J., Dunthorn, M., Edvardsen, B., Eglit, Y., Guillou, L., Hampl, V., Heiss, A.A., Hoppenrath, M., James, T.Y., Karnkowska, A., Karpov, S., Kim, E., Kolisko, M., Kudryavtsev, A., Lahr, D.J.G., Lara, E., Le Gall, L., Lynn, D.H., Mann, D.G., Massana, R., Mitchell, E.A.D., Morrow, C., Park, J.S., Pawlowski, J.W., Powell, M.J., Richter, D.J., Rueckert, S., Shadwick, L., Shimano, S., Spiegel, F.W., Torruella, G., Youssef, N., Zlatogursky, V., Zhang, Q., 2019. Revisions to the Classification, Nomenclature, and Diversity of Eukaryotes. Journal of Eukaryotic Microbiology 66, 4–119. https://doi.org/10.1111/jeu.12691

Adl, S.M., Simpson, A.G.B., Farmer, M.A., Andersen, R.A., Anderson, O.R., Barta, J.R., Bowser, S.S., Brugerolle, G., Fensome, R.A., Fredericq, S., James, T.Y., Karpov, S., Kugrens, P., Krug, J., Lane, C.E., Lewis, L.A., Lodge, J., Lynn, D.H., Mann, D.G., Mccourt, R.M., Mendoza, L., Moestrup, Ø., Mozley-Standridge, S.E., Nerad, T.A., Shearer, C.A., Smirnov, A.V., Spiegel, F.W., Taylor, M.F.J.R., 2005. The New Higher Level Classification of Eukaryotes with Emphasis on the Taxonomy of Protists. Journal of Eukaryotic Microbiology 52, 399–451. https://doi.org/10.1111/j.1550-7408.2005.00053.x

Ahmad, S., Khan, S.A., Khan, K.A., Li, J., 2021. Novel Insight Into the Development and Function of Hypopharyngeal Glands in Honey Bees. Frontiers in Physiology 11.

Akella, J.S., Carter, S.P., Nguyen, K., Tsiropoulou, S., Moran, A.L., Silva, M., Rizvi, F., Kennedy, B.N., Hall, D.H., Barr, M.M., Blacque, O.E., 2020. Ciliary Rab28 and the BBSome negatively regulate extracellular vesicle shedding. eLife 9, e50580. https://doi.org/10.7554/eLife.50580

Alber, F., Dokudovskaya, S., Veenhoff, L.M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B.T., Sali, A., Rout, M.P., 2007. The molecular architecture of the nuclear pore complex. Nature 450, 695–701. https://doi.org/10.1038/nature06405

Almeida, M., Pintacuda, G., Masui, O., Koseki, Y., Gdula, M., Cerase, A., Brown, D., Mould, A., Innocent, C., Nakayama, M., Schermelleh, L., Nesterova, T.B., Koseki, H., Brockdorff, N., 2017. PCGF3/5-PRC1 initiates Polycomb recruitment in X chromosome inactivation. Science 356, 1081–1084. https://doi.org/10.1126/science.aal2512

Andrade, M.A., Perez-Iratxeta, C., Ponting, C.P., 2001. Protein Repeats: Structures, Functions, and Evolution. Journal of Structural Biology 134, 117–131. https://doi.org/10.1006/jsbi.2001.4392

Aniento, F., Gu, F., Parton, R.G., Gruenberg, J., 1996. An endosomal beta COP is involved in the pH-dependent formation of transport vesicles destined for late endosomes. Journal of Cell Biology 133, 29–41. https://doi.org/10.1083/jcb.133.1.29

Anvarian, Z., Mykytyn, K., Mukhopadhyay, S., Pedersen, L.B., Christensen, S.T., 2019. Cellular signalling by primary cilia in development, organ function and disease. Nature Reviews Nephrology 15, 199–219. https://doi.org/10.1038/s41581-019-0116-9

Asakawa, H., Yang, H.-J., Yamamoto, T.G., Ohtsuki, C., Chikashige, Y., Sakata-Sogawa, K., Tokunaga, M., Iwamoto, M., Hiraoka, Y., Haraguchi, T., 2014. Characterization of nuclear pore complex components in fission yeast Schizosaccharomyces pombe. Nucleus 5, 149–162. https://doi.org/10.4161/nucl.28487

Asiedu, M., Wu, D., Matsumura, F., Wei, Q., 2009. Centrosome/Spindle Pole–associated Protein Regulates Cytokinesis via Promoting the Recruitment of MyoGEF to the Central Spindle. MBoC 20, 1428–1440. https://doi.org/10.1091/mbc.e08-01-0001

Bachmann-Gagescu, R., Phelps, I.G., Stearns, G., Link, B.A., Brockerhoff, S.E., Moens, C.B., Doherty, D., 2011. The ciliopathy gene cc2d2a controls zebrafish photoreceptor outer segment development through a role in Rab8-dependent vesicle trafficking. Human Molecular Genetics 20, 4041–4055. https://doi.org/10.1093/hmg/ddr332

Badano, J.L., Mitsuma, N., Beales, P.L., Katsanis, N., 2006. The Ciliopathies: An Emerging Class of Human Genetic Disorders. Annual Review of Genomics and Human Genetics 7, 125–148. https://doi.org/10.1146/annurev.genom.7.080505.115610

Balczon, R., Bao, L., Zimmer, W., 1994. PCM-1, A 228-kD centrosome autoantigen with a distinct cell cycle distribution. Journal of Cell Biology 124, 783–793. https://doi.org/10.1083/jcb.124.5.783

Balczon, R., Simerly, C., Takahashi, D., Schatten, G., 2002. Arrest of cell cycle progression during first interphase in murine zygotes microinjected with anti-PCM-1 antibodies. Cell Motility 52, 183–192. https://doi.org/10.1002/cm.10043

Baldauf, S.L., 2003. The Deep Roots of Eukaryotes. Science 300, 1703–1706. https://doi.org/10.1126/science.1085544

Bapteste, E., Charlebois, R.L., MacLeod, D., Brochier, C., 2005. The two tempos of nuclear pore complex evolution: highly adapting proteins in an ancient frozen structure. Genome Biol 6, R85. https://doi.org/10.1186/gb-2005-6-10-r85

Bardet, G., 1920. Sur un syndrome d'obesite congenitale avec polydactylie et retinite pigmentaire (contribution a l'etude des former cliniques de l'obesite hypophysaire). Amedee Le Grand 470.

Bayliss, R., Littlewood, T., Stewart, M., 2000. Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. Cell 102, 99–108. https://doi.org/10.1016/s0092-8674(00)00014-3

Bayliss, R., Littlewood, T., Strawn, L.A., Wente, S.R., Stewart, M., 2002. GLFG and FxFG nucleoporins bind to overlapping sites on importin-beta. J Biol Chem 277, 50597–50606. https://doi.org/10.1074/jbc.M209037200

Beales, P.L., 2005. Lifting the lid on Pandora's box: the Bardet-Biedl syndrome. Current Opinion in Genetics & Development, Genetics of disease 15, 315–323. https://doi.org/10.1016/j.gde.2005.04.006

Beck, M., Mosalaganti, S., Kosinski, J., 2018. From the resolution revolution to evolution: structural insights into the evolutionary relationships between vesicle coats and the nuclear pore. Curr Opin Struct Biol 52, 32–40. https://doi.org/10.1016/j.sbi.2018.07.012

Beck, R., Adolf, F., Weimer, C., Bruegger, B., Wieland, F.T., 2009. ArfGAP1 Activity and COPI Vesicle Biogenesis. Traffic 10, 307–315. https://doi.org/10.1111/j.1600-0854.2008.00865.x

Beller, M., Sztalryd, C., Southall, N., Bell, M., Jäckle, H., Auld, D.S., Oliver, B., 2008. COPI Complex Is a Regulator of Lipid Homeostasis. PLOS Biology 6, e292. https://doi.org/10.1371/journal.pbio.0060292

Benaim, G., Bermudez, R., Urbina, J.A., 1990. Ca2+ transport in isolated mitochondrial vesicles from Leishmania braziliensis promastigotes. Molecular and Biochemical Parasitology 39, 61–68. https://doi.org/10.1016/0166-6851(90)90008-A

Bentley, M.L., Corn, J.E., Dong, K.C., Phung, Q., Cheung, T.K., Cochran, A.G., 2011. Recognition of UbcH5c and the nucleosome by the Bmi1/Ring1b ubiquitin ligase complex. EMBO J 30, 3285–3297. https://doi.org/10.1038/emboj.2011.243

Bernstein, M., 1995. Flagellar kinesins: New moves with an old beat. Cell Motility 32, 125–128. https://doi.org/10.1002/cm.970320211

Betts, H.C., Puttick, M.N., Clark, J.W., Williams, T.A., Donoghue, P.C.J., Pisani, D., 2018. Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. Nat Ecol Evol 2, 1556–1562. https://doi.org/10.1038/s41559-018-0644-x

Bi, J., Tsai, N.-P., Lu, H.-Y., Loh, H.H., Wei, L.-N., 2007. Copb1-facilitated axonal transport and translation of κ opioid-receptor mRNA. Proceedings of the National Academy of Sciences 104, 13810–13815. https://doi.org/10.1073/pnas.0703805104

Biedl, A., 1922. Ein Geschwister mit adiposogenitaler Dystrophie. Deutsche medicinische Wochenschrift 48, 1630.

Blasius, T.L., Takao, D., Verhey, K.J., 2019. NPHP proteins are binding partners of nucleoporins at the base of the primary cilium. PLOS ONE 14, e0222924. https://doi.org/10.1371/journal.pone.0222924

Blitzer, A.L., Panagis, L., Gusella, G.L., Danias, J., Mlodzik, M., Iomini, C., 2011. Primary cilia dynamics instruct tissue patterning and repair of corneal endothelium. PNAS 108, 2819–2824. https://doi.org/10.1073/pnas.1016702108

Bloodgood, R.A., 2010. Sensory reception is an attribute of both primary cilia and motile cilia. Journal of Cell Science 123, 505–509. https://doi.org/10.1242/jcs.066308

Braschi, E., Goyon, V., Zunino, R., Mohanty, A., Xu, L., McBride, H.M., 2010. Vps35 Mediates Vesicle Transport between the Mitochondria and Peroxisomes. Current Biology 20, 1310–1315. https://doi.org/10.1016/j.cub.2010.05.066

Braun, D.A., Hildebrandt, F., 2017. Ciliopathies. Cold Spring Harb Perspect Biol 9, a028191. https://doi.org/10.1101/cshperspect.a028191

Bremer, N., Tria, F.D.K., Skejo, J., Garg, S.G., Martin, W.F., 2022. Ancestral State Reconstructions Trace Mitochondria But Not Phagocytosis to the Last Eukaryotic Common Ancestor. Genome Biology and Evolution 14, evac079. https://doi.org/10.1093/gbe/evac079

Bremer, N., Tria, F.D.K., Skejo, J., Martin, W.F., 2023. The Ancestral Mitotic State: Closed Orthomitosis With Intranuclear Spindles in the Syncytial Last Eukaryotic Common Ancestor. Genome Biology and Evolution 15, evad016. https://doi.org/10.1093/gbe/evad016

Breslow, D.K., Koslover, E.F., Seydel, F., Spakowitz, A.J., Nachury, M.V., 2013. An in vitro assay for entry into cilia reveals unique properties of the soluble diffusion barrier. Journal of Cell Biology 203, 129–147. https://doi.org/10.1083/jcb.201212024

Buchwald, G., van der Stoop, P., Weichenrieder, O., Perrakis, A., van Lohuizen, M., Sixma, T.K., 2006. Structure and E3-ligase activity of the Ring-Ring complex of polycomb proteins Bmi1 and Ring1b. EMBO J 25, 2465–2474. https://doi.org/10.1038/sj.emboj.7601144

Burki, F., Roger, A.J., Brown, M.W., Simpson, A.G.B., 2020. The New Tree of Eukaryotes. Trends in Ecology & Evolution 35, 43–55. https://doi.org/10.1016/j.tree.2019.08.008

Cao, R., Tsukada, Y.-I., Zhang, Y., 2005. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Mol Cell 20, 845–854. https://doi.org/10.1016/j.molcel.2005.12.002

Carvalho-Santos, Z., Azimzadeh, J., Pereira-Leal, José.B., Bettencourt-Dias, M., 2011. Tracing the origins of centrioles, cilia, and flagella. Journal of Cell Biology 194, 165–175. https://doi.org/10.1083/jcb.201011152

Casolari, J.M., Brown, C.R., Komili, S., West, J., Hieronymus, H., Silver, P.A., 2004. Genome-Wide Localization of the Nuclear Transport Machinery Couples Transcriptional

Status and Nuclear Organization. Cell 117, 427–439. https://doi.org/10.1016/S0092-8674(04)00448-9

Cassioli, C., Onnis, A., Finetti, F., Capitani, N., Brunetti, J., Compeer, E.B., Niederlova, V., Stepanek, O., Dustin, M.L., Baldari, C.T., 2021. The Bardet–Biedl syndrome complex component BBS1 controls T cell polarity during immune synapse assembly. Journal of Cell Science 134, jcs258462. https://doi.org/10.1242/jcs.258462

Cavalier-Smith, T., 2002. The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. International Journal of Systematic and Evolutionary Microbiology 52, 297–354. https://doi.org/10.1099/00207713-52-2-297

Chaki, M., Airik, R., Ghosh, A.K., Giles, R.H., Chen, R., Slaats, G.G., Wang, H., Hurd, T.W., Zhou, W., Cluckey, A., Gee, H.Y., Ramaswami, G., Hong, C.-J., Hamilton, B.A., Červenka, I., Ganji, R.S., Bryja, V., Arts, H.H., van Reeuwijk, J., Oud, M.M., Letteboer, S.J.F., Roepman, R., Husson, H., Ibraghimov-Beskrovnaya, O., Yasunaga, T., Walz, G., Eley, L., Sayer, J.A., Schermer, B., Liebau, M.C., Benzing, T., Le Corre, S., Drummond, I., Janssen, S., Allen, S.J., Natarajan, S., O'Toole, J.F., Attanasio, M., Saunier, S., Antignac, C., Koenekoop, R.K., Ren, H., Lopez, I., Nayir, A., Stoetzel, C., Dollfus, H., Massoudi, R., Gleeson, J.G., Andreoli, S.P., Doherty, D.G., Lindstrad, A., Golzio, C., Katsanis, N., Pape, L., Abboud, E.B., Al-Rajhi, A.A., Lewis, R.A., Omran, H., Lee, E.Y.-H.P., Wang, S., Sekiguchi, J.M., Saunders, R., Johnson, C.A., Garner, E., Vanselow, K., Andersen, J.S., Shlomai, J., Nurnberg, G., Nurnberg, P., Levy, S., Smogorzewska, A., Otto, E.A., Hildebrandt, F., 2012. Exome Capture Reveals ZNF423 and CEP164 Mutations, Linking Renal Ciliopathies to DNA Damage Response Signaling. Cell 150, 533–548. https://doi.org/10.1016/j.cell.2012.06.028

Chaudhuri, I., Söding, J., Lupas, A.N., 2008. Evolution of the beta-propeller fold. Proteins 71, 795–803. https://doi.org/10.1002/prot.21764

Choi, H.J.C., Lin, J.-R., Vannier, J.-B., Slaats, G.G., Kile, A.C., Paulsen, R.D., Manning, D.K., Beier, D.R., Giles, R.H., Boulton, S.J., Cimprich, K.A., 2013. NEK8 Links the ATR-Regulated Replication Stress Response and S Phase CDK Activity to Renal Ciliopathies. Molecular Cell 51, 423–439. https://doi.org/10.1016/j.molcel.2013.08.006

Cingolani, G., Petosa, C., Weis, K., Müller, C.W., 1999. Structure of importin-β bound to the IBB domain of importin-α. Nature 399, 221–229. https://doi.org/10.1038/20367

Cook, K.L., Soto-Pantoja, D.R., Abu-Asab, M., Clarke, P.A., Roberts, D.D., Clarke, R., 2014. Mitochondria directly donate their membrane to form autophagosomes during a novel mechanism of parkin-associated mitophagy. Cell Biosci 4, 16. https://doi.org/10.1186/2045-3701-4-16

Cronshaw, J.M., Krutchinsky, A.N., Zhang, W., Chait, B.T., Matunis, M.J., 2002. Proteomic analysis of the mammalian nuclear pore complex. J Cell Biol 158, 915–927. https://doi.org/10.1083/jcb.200206106

Dacks, J.B., Field, M.C., 2007. Evolution of the eukaryotic membrane-trafficking system: origin, tempo and mode. Journal of Cell Science 120, 2977–2985. https://doi.org/10.1242/jcs.013250

Dacks, J.B., Robinson, M.S., 2017. Outerwear through the ages: evolutionary cell biology of vesicle coats. Current Opinion in Cell Biology, Cell Organelles 47, 108–116. https://doi.org/10.1016/j.ceb.2017.04.001

Daro, E., Sheff, D., Gomez, M., Kreis, T., Mellman, I., 1997. Inhibition of Endosome Function in CHO Cells Bearing a Temperature-sensitive Defect in the Coatomer (COPI) Component ε-COP. Journal of Cell Biology 139, 1747–1759. https://doi.org/10.1083/jcb.139.7.1747

de Duve, C., 2007. The origin of eukaryotes: a reappraisal. Nat Rev Genet 8, 395–403. https://doi.org/10.1038/nrg2071

De Duve, C., 1969. Evolution of the peroxisome. Ann N Y Acad Sci 168, 369–381. https://doi.org/10.1111/j.1749-6632.1969.tb43124.x

de Napoles, M., Mermoud, J.E., Wakao, R., Tang, Y.A., Endoh, M., Appanah, R., Nesterova, T.B., Silva, J., Otte, A.P., Vidal, M., Koseki, H., Brockdorff, N., 2004. Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. Dev Cell 7, 663–676. https://doi.org/10.1016/j.devcel.2004.10.005

Deatherage, B.L., Cookson, B.T., 2012. Membrane Vesicle Release in Bacteria, Eukaryotes, and Archaea: a Conserved yet Underappreciated Aspect of Microbial Life. Infection and Immunity 80, 1948–1957. https://doi.org/10.1128/iai.06014-11

DeGrasse, J.A., DuBois, K.N., Devos, D., Siegel, T.N., Sali, A., Field, M.C., Rout, M.P., Chait, B.T., 2009. Evidence for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic ancestor. Mol Cell Proteomics 8, 2119–2130. https://doi.org/10.1074/mcp.M900038-MCP200

del Viso, F., Huang, F., Myers, J., Chalfant, M., Zhang, Y., Reza, N., Bewersdorf, J., Lusk, C.P., Khokha, M.K., 2016. Congenital Heart Disease Genetics Uncovers Context-Dependent Organization and Function of Nucleoporins at Cilia. Developmental Cell 38, 478–492. https://doi.org/10.1016/j.devcel.2016.08.002

den Dulk, B., van Eijk, P., de Ruijter, M., Brandsma, J.A., Brouwer, J., 2008. The NER protein Rad33 shows functional homology to human Centrin2 and is involved in modification of Rad4. DNA Repair 7, 858–868. https://doi.org/10.1016/j.dnarep.2008.02.004

Denning, D.P., Patel, S.S., Uversky, V., Fink, A.L., Rexach, M., 2003. Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded. Proc Natl Acad Sci U S A 100, 2450–2455. https://doi.org/10.1073/pnas.0437902100

Devos, D., Dokudovskaya, S., Alber, F., Williams, R., Chait, B.T., Sali, A., Rout, M.P., 2004. Components of Coated Vesicles and Nuclear Pore Complexes Share a Common Molecular Architecture. PLOS Biology 2, e380. https://doi.org/10.1371/journal.pbio.0020380

Devos, D., Dokudovskaya, S., Williams, R., Alber, F., Eswar, N., Chait, B.T., Rout, M.P., Sali, A., 2006. Simple fold composition and modular architecture of the nuclear pore complex. Proceedings of the National Academy of Sciences 103, 2172–2177. https://doi.org/10.1073/pnas.0506345103

Devos, D.P., Reynaud, E.G., 2010. Evolution. Intermediate steps. Science 330, 1187–1188. https://doi.org/10.1126/science.1196720

Dishinger, J.F., Kee, H.L., Jenkins, P.M., Fan, S., Hurd, T.W., Hammond, J.W., Truong, Y.N.-T., Margolis, B., Martens, J.R., Verhey, K.J., 2010. Ciliary entry of the kinesin-2 motor KIF17 is regulated by importin-β2 and RanGTP. Nat Cell Biol 12, 703–710. https://doi.org/10.1038/ncb2073

Dokudovskaya, S., Waharte, F., Schlessinger, A., Pieper, U., Devos, D.P., Cristea, I.M., Williams, R., Salamero, J., Chait, B.T., Sali, A., Field, M.C., Rout, M.P., Dargemont, C., 2011. A Conserved Coatomer-related Complex Containing Sec13 and Seh1 Dynamically Associates With the Vacuole in Saccharomyces cerevisiae*. Molecular & Cellular Proteomics 10. https://doi.org/10.1074/mcp.M110.006478

Doolittle, W.F., 1998. You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. Trends Genet 14, 307–311. https://doi.org/10.1016/s0168-9525(98)01494-2

Duden, R., 2003. ER-to-Golgi transport: COP I and COP II function (Review). Molecular Membrane Biology 20, 197–207. https://doi.org/10.1080/0968768031000122548

Dutcher, S.K., 1995. Flagellar assembly in two hundred and fifty easy-to-follow steps. Trends in Genetics 11, 398–404. https://doi.org/10.1016/S0168-9525(00)89123-4

Edeling, M.A., Smith, C., Owen, D., 2006. Life of a clathrin coat: insights from clathrin and AP structures. Nat Rev Mol Cell Biol 7, 32–44. https://doi.org/10.1038/nrm1786

Eguether, T., San Agustin, J.T., Keady, B.T., Jonassen, J.A., Liang, Y., Francis, R., Tobita, K., Johnson, C.A., Abdelhamed, Z.A., Lo, C.W., Pazour, G.J., 2014. IFT27 Links the BBSome to IFT for Maintenance of the Ciliary Signaling Compartment. Developmental Cell 31, 279–290. https://doi.org/10.1016/j.devcel.2014.09.011

Eibauer, M., Pellanda, M., Turgay, Y., Dubrovsky, A., Wild, A., Medalia, O., 2015. Structure and gating of the nuclear pore complex. Nat Commun 6, 7532. https://doi.org/10.1038/ncomms8532

Elias, M., Brighouse, A., Gabernet-Castello, C., Field, M.C., Dacks, J.B., 2012. Sculpting the endomembrane system in deep time: high resolution phylogenetics of Rab GTPases. J Cell Sci 125, 2500–2508. https://doi.org/10.1242/jcs.101378

Endicott, S.J., Basu, B., Khokha, M., Brueckner, M., 2015. The NIMA-like kinase Nek2 is a key switch balancing cilia biogenesis and resorption in the development of left-right asymmetry. Development 142, 4068–4079. https://doi.org/10.1242/dev.126953

Endicott, S.J., Brueckner, M., 2018. NUP98 Sets the Size-Exclusion Diffusion Limit through the Ciliary Base. Current Biology 28, 1643-1650.e3. https://doi.org/10.1016/j.cub.2018.04.014

Ewerling, A., Graebling, I., Wierczeiko, A., Kotzurek, E., Foitzik, S., Gerber, S., Colgan, T.J., May-Simera, H., 2023a. Analysis of human BBS protein homologues in insects support alternative non-ciliary functions. https://doi.org/10.1101/2023.07.28.550953

Ewerling, A., Maissl, V., Wickstead, B., May-Simera, H.L., 2023b. Neofunctionalization of ciliary BBS proteins to nuclear roles is likely a frequent innovation across eukaryotes. iScience 26, 106410. https://doi.org/10.1016/j.isci.2023.106410

Fernandez-Martinez, J., Rout, M.P., 2021. One Ring to Rule them All? Structural and Functional Diversity in the Nuclear Pore Complex. Trends in Biochemical Sciences 46, 595–607. https://doi.org/10.1016/j.tibs.2021.01.003

Field, M.C., Dacks, J.B., 2009. First and last ancestors: reconstructing evolution of the endomembrane system with ESCRTs, vesicle coat proteins, and nuclear pore complexes. Current Opinion in Cell Biology, Cell structure and dynamics 21, 4–13. https://doi.org/10.1016/j.ceb.2008.12.004

Field, M.C., Gabernet-Castello, C., Dacks, J.B., 2007. Reconstructing the Evolution of the Endocytic System: Insights from Genomics and Molecular Cell Biology, in: Eukaryotic Membranes and Cytoskeleton: Origins and Evolution, Advances in Experimental Medicine and Biology. Springer, New York, NY, pp. 84–96. https://doi.org/10.1007/978-0-387-74021-8_7

Field, M.C., Koreny, L., Rout, M.P., 2014. Enriching the Pore: Splendid Complexity from Humble Origins. Traffic 15, 141–156. https://doi.org/10.1111/tra.12141

Field, M.C., Rout, M.P., 2019. Pore timing: the evolutionary origins of the nucleus and nuclear pore complex. F1000Res 8, F1000 Faculty Rev-369. https://doi.org/10.12688/f1000research.16402.1

Field, M.C., Sali, A., Rout, M.P., 2011. On a bender—BARs, ESCRTs, COPs, and finally getting your coat. Journal of Cell Biology 193, 963–972. https://doi.org/10.1083/jcb.201102042

Fischer, J., Teimer, R., Amlacher, S., Kunze, R., Hurt, E., 2015. Linker Nups connect the nuclear pore complex inner ring with the outer ring and transport channel. Nat Struct Mol Biol 22, 774–781. https://doi.org/10.1038/nsmb.3084

Florea, L., Caba, L., Gorduza, E.V., 2021. Bardet-Biedl Syndrome-Multiple Kaleidoscope Images: Insight into Mechanisms of Genotype-Phenotype Correlations. Genes (Basel) 12, 1353. https://doi.org/10.3390/genes12091353

Forsythe, E., Beales, P.L., 2013. Bardet–Biedl syndrome. European Journal of Human Genetics 21, 8–13. https://doi.org/10.1038/ejhg.2012.115

Forterre, P., 2013. The Common Ancestor of Archaea and Eukarya Was Not an Archaeon. Archaea 2013, e372396. https://doi.org/10.1155/2013/372396

Forterre, P., 2011. A new fusion hypothesis for the origin of Eukarya: better than previous ones, but probably also wrong. Research in Microbiology, Archaea and the Tree of Life 162, 77–91. https://doi.org/10.1016/j.resmic.2010.10.005

Forterre, P., 2002. The origin of DNA genomes and DNA replication proteins. Curr Opin Microbiol 5, 525–532. https://doi.org/10.1016/s1369-5274(02)00360-0

Gabriely, G., Kama, R., Gerst, J.E., 2007. Involvement of Specific COPI Subunits in Protein Sorting from the Late Endosome to the Vacuole in Yeast. Molecular and Cellular Biology 27, 526–540. https://doi.org/10.1128/MCB.00577-06

Gao, Z., Lee, P., Stafford, J.M., von Schimmelmann, M., Schaefer, A., Reinberg, D., 2014. An AUTS2-Polycomb complex activates gene expression in the CNS. Nature 516, 349–354. https://doi.org/10.1038/nature13921

Gascue, C., Tan, P.L., Cardenas-Rodriguez, M., Libisch, G., Fernandez-Calero, T., Liu, Y.P., Astrada, S., Robello, C., Naya, H., Katsanis, N., Badano, J.L., 2012. Direct role of Bardet–Biedl syndrome proteins in transcriptional regulation. J Cell Sci 125, 362–375. https://doi.org/10.1242/jcs.089375

Gehrke, S.H., Fisher, J.P., Palasis, M., Lund, M.E., 1997. Factors determining hydrogel permeability. Ann N Y Acad Sci 831, 179–207. https://doi.org/10.1111/j.1749-6632.1997.tb52194.x

Gould, S.B., Garg, S.G., Martin, W.F., 2016. Bacterial Vesicle Secretion and the Evolutionary Origin of the Eukaryotic Endomembrane System. Trends in Microbiology 24, 525–534. https://doi.org/10.1016/j.tim.2016.03.005

Gribaldo, S., Poole, A.M., Daubin, V., Forterre, P., Brochier-Armanet, C., 2010. The origin of eukaryotes and their relationship with the Archaea: are we at a phylogenomic impasse? Nat Rev Microbiol 8, 743–752. https://doi.org/10.1038/nrmicro2426

Grimm, R., Singh, H., Rachel, R., Typke, D., Zillig, W., Baumeister, W., 1998. Electron Tomography of Ice-Embedded Prokaryotic Cells. Biophysical Journal 74, 1031–1042. https://doi.org/10.1016/S0006-3495(98)74028-7

Grozinger, C.M., Zayed, A., 2020. Improving bee health through genomics. Nat Rev Genet 21, 277–291. https://doi.org/10.1038/s41576-020-0216-1

Gu, F., Aniento, F., Parton, R.G., Gruenberg, J., 1997. Functional Dissection of COP-I Subunits in the Biogenesis of Multivesicular Endosomes. Journal of Cell Biology 139, 1183–1195. https://doi.org/10.1083/jcb.139.5.1183

Guo, Q., Vasile, E., Krieger, M., 1994. Disruptions in Golgi structure and membrane traffic in a conditional lethal mammalian cell mutant are corrected by epsilon-COP. Journal of Cell Biology 125, 1213–1224. https://doi.org/10.1083/jcb.125.6.1213

Haeckel, E.H.P.A., 1866. Generelle Morphologie der Organismen : allgemeine Grundzüge der organischen Formen-Wissenschaft, mechanisch begründet durch die von Charles Darwin reformirte Descendenz-Theorie. Berlin : G. Reimer.

Harris, E.H., 2009. The Chlamydomonas Sourcebook: Introduction to Chlamydomonas and Its Laboratory Use: Volume 1. Academic Press.

Heinß, N., Sushkin, M., Yu, M., Lemke, E.A., 2020. Multifunctionality of F-rich nucleoporins. Biochemical Society Transactions 48, 2603–2614. https://doi.org/10.1042/BST20200357

Hodges, M.E., Scheumann, N., Wickstead, B., Langdale, J.A., Gull, K., 2010. Reconstructing the evolutionary history of the centriole from protein components. J Cell Sci 123, 1407–1413. https://doi.org/10.1242/jcs.064873

Hodges, M.E., Wickstead, B., Gull, K., Langdale, J.A., 2011. Conservation of ciliary proteins in plants with no cilia. BMC Plant Biology 11, 185. https://doi.org/10.1186/1471-2229-11-185

Hoelz, A., Debler, E.W., Blobel, G., 2011. The Structure of the Nuclear Pore Complex. Annu. Rev. Biochem. 80, 613–643. https://doi.org/10.1146/annurev-biochem-060109-151030

Hoelz, A., Glavy, J.S., Beck, M., 2016. Toward the atomic structure of the nuclear pore complex: when top down meets bottom up. Nat Struct Mol Biol 23, 624–630. https://doi.org/10.1038/nsmb.3244

Horwitz, A., Birk, R., 2021. BBS4 Is Essential for Nuclear Transport of Transcription Factors Mediating Neuronal ER Stress Response. Mol Neurobiol 58, 78–91. https://doi.org/10.1007/s12035-020-02104-z

Hsiao, Y.-C., Tong, Z.J., Westfall, J.E., Ault, J.G., Page-McCaw, P.S., Ferland, R.J., 2009. Ahi1, whose human ortholog is mutated in Joubert syndrome, is required for Rab8a localization, ciliogenesis and vesicle trafficking. Human Molecular Genetics 18, 3926–3941. https://doi.org/10.1093/hmg/ddp335

Hu, Q., Milenkovic, L., Jin, H., Scott, M.P., Nachury, M.V., Spiliotis, E.T., Nelson, W.J., 2010. A Septin Diffusion Barrier at the Base of the Primary Cilium Maintains Ciliary Membrane Protein Distribution. Science 329, 436–439. https://doi.org/10.1126/science.1191054

Huang, S., Dougherty, L.L., Avasthi, P., 2021. Separable roles for RanGTP in nuclear and ciliary trafficking of a kinesin-2 subunit. Journal of Biological Chemistry 296. https://doi.org/10.1074/jbc.RA119.010936

Ikonomou, L., Schneider, Y.-J., Agathos, S.N., 2003. Insect cell culture for industrial production of recombinant proteins. Appl Microbiol Biotechnol 62, 1–20. https://doi.org/10.1007/s00253-003-1223-9

Iovine, M.K., Watkins, J.L., Wente, S.R., 1995. The GLFG repetitive region of the nucleoporin Nup116p interacts with Kap95p, an essential yeast nuclear import factor. J Cell Biol 131, 1699–1713. https://doi.org/10.1083/jcb.131.6.1699

Ishii, K., Arib, G., Lin, C., Van Houwe, G., Laemmli, U.K., 2002. Chromatin Boundaries in Budding Yeast: The Nuclear Pore Connection. Cell 109, 551–562. https://doi.org/10.1016/S0092-8674(02)00756-0

Iwamoto, M., Asakawa, H., Hiraoka, Y., Haraguchi, T., 2010. Nucleoporin Nup98: a gatekeeper in the eukaryotic kingdoms. Genes to Cells 15, 661–669. https://doi.org/10.1111/j.1365-2443.2010.01415.x

Izaurralde, E., Kutay, U., von Kobbe, C., Mattaj, I.W., Görlich, D., 1997. The asymmetric distribution of the constituents of the Ran system is essential for transport into and out of the nucleus. EMBO J 16, 6535–6547. https://doi.org/10.1093/emboj/16.21.6535

Jékely, G., Arendt, D., 2006. Evolution of intraflagellar transport from coated vesicles and autogenous origin of the eukaryotic cilium. BioEssays 28, 191–198. https://doi.org/10.1002/bies.20369

Jennings, B.H., 2011. Drosophila – a versatile model in biology & medicine. Materials Today 14, 190–195. https://doi.org/10.1016/S1369-7021(11)70113-4

Jin, H., White, S.R., Shida, T., Schulz, S., Aguiar, M., Gygi, S.P., Bazan, J.F., Nachury, M.V., 2010. The Conserved Bardet-Biedl Syndrome Proteins Assemble a Coat that Traffics Membrane Proteins to Cilia. Cell 141, 1208–1219. https://doi.org/10.1016/j.cell.2010.05.015

John, P., Whatley, F.R., 1975. Paracoccus denitrificans and the evolutionary origin of the mitochondrion. Nature 254, 495–498. https://doi.org/10.1038/254495a0

Jovanovic-Talisman, T., Tetenbaum-Novatt, J., McKenney, A.S., Zilman, A., Peters, R., Rout, M.P., Chait, B.T., 2009. Artificial nanopores that mimic the transport selectivity of the nuclear pore complex. Nature 457, 1023–1027. https://doi.org/10.1038/nature07600

Kamakura, M., 2011. Royalactin induces queen differentiation in honeybees. Nature 473, 478–483. https://doi.org/10.1038/nature10093

Karpenahalli, M.R., Lupas, A.N., Söding, J., 2007. TPRpred: a tool for prediction of TPR-, PPR- and SEL1-like repeats from protein sequences. BMC Bioinformatics 8, 2. https://doi.org/10.1186/1471-2105-8-2

Kawai, T., Saito, K., Lee, W., 2003. Protein binding to polymer brush, based on ion-exchange, hydrophobic, and affinity interactions. J Chromatogr B Analyt Technol Biomed Life Sci 790, 131–142. https://doi.org/10.1016/s1570-0232(03)00090-4

Kee, H.L., Dishinger, J.F., Lynne Blasius, T., Liu, C.-J., Margolis, B., Verhey, K.J., 2012. A size-exclusion permeability barrier and nucleoporins characterize a ciliary pore complex that regulates transport into cilia. Nat Cell Biol 14, 431–437. https://doi.org/10.1038/ncb2450

Kee, H.L., Verhey, K.J., 2013. Molecular connections between nuclear and ciliary import processes. Cilia 2, 11. https://doi.org/10.1186/2046-2530-2-11

Keeling, P.J., 2010. The endosymbiotic origin, diversification and fate of plastids. Philos Trans R Soc Lond B Biol Sci 365, 729–748. https://doi.org/10.1098/rstb.2009.0103

Keeling, P.J., Burger, G., Durnford, D.G., Lang, B.F., Lee, R.W., Pearlman, R.E., Roger, A.J., Gray, M.W., 2005. The tree of eukaryotes. Trends in Ecology & Evolution 20, 670–676. https://doi.org/10.1016/j.tree.2005.09.005

Keil, T.A., 2012. Sensory cilia in arthropods. Arthropod Structure & Development 41, 515–534. https://doi.org/10.1016/j.asd.2012.07.001

Kenny, T.D., Beales, P.L., 2013. Ciliopathies: A Reference for Clinicians. OUP Oxford.

Keune, W. jan, Bultsma, Y., Sommer, L., Jones, D., Divecha, N., 2011. Phosphoinositide signalling in the nucleus. Advances in Enzyme Regulation 51, 91–99. https://doi.org/10.1016/j.advenzreg.2010.09.009

Killick-Kendrick, R., Peters, W., 1978. Rodent malaria. Academic Press, London, New York.

Kim, J.C., Badano, J.L., Sibold, S., Esmail, M.A., Hill, J., Hoskins, B.E., Leitch, C.C., Venner, K., Ansley, S.J., Ross, A.J., Leroux, M.R., Katsanis, N., Beales, P.L., 2004. The Bardet-Biedl protein BBS4 targets cargo to the pericentriolar region and is required for microtubule anchoring and cell cycle progression. Nature Genetics 36, 462. https://doi.org/10.1038/ng1352

Kim, J.C., Ou, Y.Y., Badano, J.L., Esmail, M.A., Leitch, C.C., Fiedrich, E., Beales, P.L., Archibald, J.M., Katsanis, N., Rattner, J.B., Leroux, M.R., 2005. MKKS/BBS6, a

divergent chaperonin-like protein linked to the obesity disorder Bardet-Biedl syndrome, is a novel centrosomal component required for cytokinesis. Journal of Cell Science 118, 1007–1020. https://doi.org/10.1242/jcs.01676

Kim, S.J., Fernandez-Martinez, J., Nudelman, I., Shi, Y., Zhang, W., Raveh, B., Herricks, T., Slaughter, B.D., Hogan, J.A., Upla, P., Chemmama, I.E., Pellarin, R., Echeverria, I., Shivaraju, M., Chaudhury, A.S., Wang, J., Williams, R., Unruh, J.R., Greenberg, C.H., Jacobs, E.Y., Yu, Z., de la Cruz, M.J., Mironska, R., Stokes, D.L., Aitchison, J.D., Jarrold, M.F., Gerton, J.L., Ludtke, S.J., Akey, C.W., Chait, B.T., Sali, A., Rout, M.P., 2018. Integrative structure and functional anatomy of a nuclear pore complex. Nature 555, 475–482. https://doi.org/10.1038/nature26003

Klinger, C.M., Spang, A., Dacks, J.B., Ettema, T.J.G., 2016. Tracing the Archaeal Origins of Eukaryotic Membrane-Trafficking System Building Blocks. Molecular Biology and Evolution 33, 1528–1541. https://doi.org/10.1093/molbev/msw034

Klink, B.U., Gatsogiannis, C., Hofnagel, O., Wittinghofer, A., Raunser, S., 2020. Structure of the human BBSome core complex. eLife 9, e53910. https://doi.org/10.7554/eLife.53910

Knockenhauer, K.E., Schwartz, T.U., 2016. The Nuclear Pore Complex as a Flexible and Dynamic Gate. Cell 164, 1162–1171. https://doi.org/10.1016/j.cell.2016.01.034

Kobe, B., Kajava, A.V., 2000. When protein folding is simplified to protein coiling: the continuum of solenoid protein structures. Trends in Biochemical Sciences 25, 509–515. https://doi.org/10.1016/S0968-0004(00)01667-4

Kontou, A., Herman, E.K., Field, M.C., Dacks, J.B., Koumandou, V.L., 2022. Evolution of factors shaping the endoplasmic reticulum. Traffic 23, 462–473. https://doi.org/10.1111/tra.12863

Kosinski, J., Mosalaganti, S., von Appen, A., Teimer, R., DiGuilio, A.L., Wan, W., Bui, K.H., Hagen, W.J.H., Briggs, J.A.G., Glavy, J.S., Hurt, E., Beck, M., 2016. Molecular architecture of the inner ring scaffold of the human nuclear pore complex. Science 352, 363–365. https://doi.org/10.1126/science.aaf0643

Koumandou, V.L., Dacks, J.B., Coulson, R.M., Field, M.C., 2007. Control systems for membrane fusion in the ancestral eukaryote; evolution of tethering complexes and SM proteins. BMC Evol Biol 7, 29. https://doi.org/10.1186/1471-2148-7-29

Koumandou, V.L., Wickstead, B., Ginger, M.L., Giezen, M. van der, Dacks, J.B., Field, M.C., 2013. Molecular paleontology and complexity in the last eukaryotic common ancestor. Critical Reviews in Biochemistry and Molecular Biology 48, 373–396. https://doi.org/10.3109/10409238.2013.821444

Kubo, A., Tsukita, S., 2003. Non-membranous granular organelle consisting of PCM-1: subcellular distribution and cell-cycle-dependent assembly/disassembly. Journal of Cell Science 116, 919–928. https://doi.org/10.1242/jcs.00282

Kurland, C.G., Collins, L.J., Penny, D., 2006. Genomics and the irreducible nature of eukaryote cells. Science 312, 1011–1014. https://doi.org/10.1126/science.1121674

Lechtreck, K.F., Brown, J.M., Sampaio, J.L., Craft, J.M., Shevchenko, A., Evans, J.E., Witman, G.B., 2013. Cycling of the signaling protein phospholipase D through cilia requires the BBSome only for the export phase. Journal of Cell Biology 201, 249–261. https://doi.org/10.1083/jcb.201207139

Lechtreck, K.-F., Geimer, S., 2000. Distribution of polyglutamylated tubulin in the flagellar apparatus of green flagellates. Cell Motility 47, 219–235. https://doi.org/10.1002/1097-0169(200011)47:3<219::AID-CM5>3.0.CO;2-Q

Lechtreck, K.-F., Johnson, E.C., Sakai, T., Cochran, D., Ballif, B.A., Rush, J., Pazour, G.J., Ikebe, M., Witman, G.B., 2009. The Chlamydomonas reinhardtii BBSome is an IFT cargo required for export of specific signaling proteins from flagella. Journal of Cell Biology 187, 1117–1132. https://doi.org/10.1083/jcb.200909183

Lee, E.-Y., Choi, D.-Y., Kim, D.-K., Kim, J.-W., Park, J.O., Kim, S., Kim, S.-H., Desiderio, D.M., Kim, Y.-K., Kim, K.-P., Gho, Y.S., 2009. Gram-positive bacteria produce membrane vesicles: Proteomics-based characterization of Staphylococcus aureus-derived membrane vesicles. PROTEOMICS 9, 5425–5436. https://doi.org/10.1002/pmic.200900338

Lee, M.C.S., Miller, E.A., Goldberg, J., Orci, L., Schekman, R., 2004. Bi-Directional Protein Transport Between the Er and Golgi. Annual Review of Cell and Developmental Biology 20, 87–123. https://doi.org/10.1146/annurev.cellbio.20.010403.105307

Leung, K.F., Dacks, J.B., Field, M.C., 2008. Evolution of the multivesicular body ESCRT machinery; retention across the eukaryotic lineage. Traffic 9, 1698–1716. https://doi.org/10.1111/j.1600-0854.2008.00797.x

Liew, G.M., Ye, F., Nager, A.R., Murphy, J.P., Lee, J.S., Aguiar, M., Breslow, D.K., Gygi, S.P., Nachury, M.V., 2014. The intraflagellar transport protein IFT27 promotes BBSome exit from cilia through the GTPase ARL6/BBS3. Dev Cell 31, 265–278. https://doi.org/10.1016/j.devcel.2014.09.004

Liu, J., Prunuske, A.J., Fager, A.M., Ullman, K.S., 2003. The COPI Complex Functions in Nuclear Envelope Breakdown and Is Recruited by the Nucleoporin Nup153. Developmental Cell 5, 487–498. https://doi.org/10.1016/S1534-5807(03)00262-4

Lui, K., Huang, Y., 2009. RanGTPase: A Key Regulator of Nucleocytoplasmic Trafficking. Mol Cell Pharmacol 1, 148–156. https://doi.org/10.4255/mcpharmacol.09.19

Luo, N., Conwell, M.D., Chen, X., Kettenhofen, C.I., Westlake, C.J., Cantor, L.B., Wells, C.D., Weinreb, R.N., Corson, T.W., Spandau, D.F., Joos, K.M., Iomini, C., Obukhov, A.G., Sun, Y., 2014. Primary cilia signaling mediates intraocular pressure sensation. PNAS 111, 12871–12876. https://doi.org/10.1073/pnas.1323292111

Luo, X., Schoch, K., Jangam, S.V., Bhavana, V.H., Graves, H.K., Kansagra, S., Jasien, J.M., Stong, N., Keren, B., Mignot, C., Ravelli, C., Undiagnosed Diseases Network, Bellen, H.J., Wangler, M.F., Shashi, V., Yamamoto, S., 2021. Rare deleterious de novo missense variants in Rnf2/Ring2 are associated with a neurodevelopmental disorder with unique clinical features. Hum Mol Genet 30, 1283–1292. https://doi.org/10.1093/hmg/ddab110

Malik, H.S., Eickbush, T.H., Goldfarb, D.S., 1997. Evolutionary specialization of the nuclear targeting apparatus. Proceedings of the National Academy of Sciences 94, 13738–13742. https://doi.org/10.1073/pnas.94.25.13738

Mans, B.J., Anantharaman, V., Aravind, L., Koonin, E.V., 2004. Comparative genomics, evolution and origins of the nuclear envelope and nuclear pore complex. Cell Cycle 3, 1612–1637. https://doi.org/10.4161/cc.3.12.1316

Marchese, E., Caterino, M., Viggiano, D., Cevenini, A., Tolone, S., Docimo, L., Di Iorio, V., Del Vecchio Blanco, F., Fedele, R., Simonelli, F., Perna, A., Nigro, V., Capasso, G., Ruoppolo, M., Zacchia, M., 2022. Metabolomic fingerprinting of renal disease progression in Bardet-Biedl syndrome reveals mitochondrial dysfunction in kidney tubular cells. iScience 25, 105230. https://doi.org/10.1016/j.isci.2022.105230

Marcotte, E.M., Pellegrini, M., Yeates, T.O., Eisenberg, D., 1999. A census of protein repeats11Edited by J. M. Thornton. Journal of Molecular Biology 293, 151–160. https://doi.org/10.1006/jmbi.1999.3136

Margulis, L., 1970. Origin of eukaryotic cells: evidence and research implications for a theory of the origin and evolution of microbial, plant, and animal cells on the Precambrian earth. Yale University Press, New Haven.

Marquez, J., Bhattacharya, D., Lusk, C.P., Khokha, M.K., 2021. Nucleoporin NUP205 plays a critical role in cilia and congenital disease. Developmental Biology 469, 46–53. https://doi.org/10.1016/j.ydbio.2020.10.001

Martin, W., 2017. Symbiogenesis, gradualism, and mitochondrial energy in eukaryote origin. Period. Biol. 119, 141–158. https://doi.org/10.18054/pb.v119i3.5694

Martin, W., 1999. A briefly argued case that mitochondria and plastids are descendants of endosymbionts, but that the nuclear compartment is not. Proceedings of the Royal Society of London. Series B: Biological Sciences 266, 1387–1395. https://doi.org/10.1098/rspb.1999.0792

Martin, W., Koonin, E.V., 2006. Introns and the origin of nucleus-cytosol compartmentalization. Nature 440, 41–45. https://doi.org/10.1038/nature04531

Martin, W., Müller, M., 1998. The hydrogen hypothesis for the first eukaryote. Nature 392, 37–41. https://doi.org/10.1038/32096

Martin, W.F., Tielens, A.G.M., Mentel, M., Garg, S.G., Gould, S.B., 2017. The Physiology of Phagocytosis in the Context of Mitochondrial Origin. Microbiol Mol Biol Rev 81, e00008-17. https://doi.org/10.1128/MMBR.00008-17

May-Simera, H.L., Ross, A., Rix, S., Forge, A., Beales, P.L., Jagger, D.J., 2009. Patterns of expression of Bardet-Biedl syndrome proteins in the mammalian cochlea suggest noncentrosomal functions. Journal of Comparative Neurology 514, 174–188. https://doi.org/10.1002/cne.22001

McGinty, R.K., Henrici, R.C., Tan, S., 2014. Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. Nature 514, 591–596. https://doi.org/10.1038/nature13890

McInerney, J.O., Martin, W.F., Koonin, E.V., Allen, J.F., Galperin, M.Y., Lane, N., Archibald, J.M., Embley, T.M., 2011. Planctomycetes and eukaryotes: A case of analogy not homology. BioEssays 33, 810–817. https://doi.org/10.1002/bies.201100045

Mereschkowsky, C., 1905. Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. Biol Centralbl 25, 593.

Mills, D.B., Boyle, R.A., Daines, S.J., Sperling, E.A., Pisani, D., Donoghue, P.C.J., Lenton, T.M., 2022. Eukaryogenesis and oxygen in Earth history. Nat Ecol Evol 6, 520–532. https://doi.org/10.1038/s41559-022-01733-y

Mitchell, D.R., 2007. The Evolution of Eukaryotic Cilia and Flagella as Motile and Sensory Organelles, in: Eukaryotic Membranes and Cytoskeleton: Origins and Evolution, Advances in Experimental Medicine and Biology. Springer, New York, NY, pp. 130–140. https://doi.org/10.1007/978-0-387-74021-8_11

Morey, L., Pascual, G., Cozzuto, L., Roma, G., Wutz, A., Benitah, S.A., Di Croce, L., 2012. Nonoverlapping functions of the Polycomb group Cbx family of proteins in embryonic stem cells. Cell Stem Cell 10, 47–62. https://doi.org/10.1016/j.stem.2011.12.006

Mosalaganti, S., Kosinski, J., Albert, S., Schaffer, M., Strenkert, D., Salomé, P.A., Merchant, S.S., Plitzko, J.M., Baumeister, W., Engel, B.D., Beck, M., 2018. In situ architecture of the algal nuclear pore complex. Nat Commun 9, 2361. https://doi.org/10.1038/s41467-018-04739-y

Mukherjee, K., Brocchieri, L., 2013. Ancient Origin of Chaperonin Gene Paralogs Involved in Ciliopathies. J Phylogenetics Evol Biol 1.

Müller, M., Mentel, M., van Hellemond, J.J., Henze, K., Woehle, C., Gould, S.B., Yu, R.-Y., van der Giezen, M., Tielens, A.G.M., Martin, W.F., 2012. Biochemistry and Evolution of Anaerobic Energy Metabolism in Eukaryotes. Microbiology and Molecular Biology Reviews 76, 444–495. https://doi.org/10.1128/mmbr.05024-11

Nachury, M.V., Loktev, A.V., Zhang, Q., Westlake, C.J., Peränen, J., Merdes, A., Slusarski, D.C., Scheller, R.H., Bazan, J.F., Sheffield, V.C., Jackson, P.K., 2007. A Core Complex of BBS Proteins Cooperates with the GTPase Rab8 to Promote Ciliary Membrane Biogenesis. Cell 129, 1201–1213. https://doi.org/10.1016/j.cell.2007.03.053

Naor, A., Gophna, U., 2013. Cell fusion and hybrids in Archaea: prospects for genome shuffling and accelerated strain development for biotechnology. Bioengineered 4, 126–129. https://doi.org/10.4161/bioe.22649

Neumann, N., Lundin, D., Poole, A.M., 2010. Comparative Genomic Evidence for a Complete Nuclear Pore Complex in the Last Eukaryotic Common Ancestor. PLOS ONE 5, e13241. https://doi.org/10.1371/journal.pone.0013241

Nguyen Ba, A.N., Pogoutse, A., Provart, N., Moses, A.M., 2009. NLStradamus: a simple Hidden Markov Model for nuclear localization signal prediction. BMC Bioinformatics 10, 202. https://doi.org/10.1186/1471-2105-10-202

Obado, S.O., Brillantes, M., Uryu, K., Zhang, W., Ketaren, N.E., Chait, B.T., Field, M.C., Rout, M.P., 2016. Interactome Mapping Reveals the Evolutionary History of the Nuclear Pore Complex. PLOS Biology 14, e1002365. https://doi.org/10.1371/journal.pbio.1002365

O'Malley, M.A., Leger, M.M., Wideman, J.G., Ruiz-Trillo, I., 2019. Concepts of the last eukaryotic common ancestor. Nat Ecol Evol 3, 338–344. https://doi.org/10.1038/s41559-019-0796-3

O'Reilly, A.J., Dacks, J.B., Field, M.C., 2011. Evolution of the Karyopherin-β Family of Nucleocytoplasmic Transport Factors; Ancient Origins and Continued Specialization. PLOS ONE 6, e19308. https://doi.org/10.1371/journal.pone.0019308

Ori, A., Banterle, N., Iskar, M., Andrés-Pons, A., Escher, C., Khanh Bui, H., Sparks, L., Solis-Mezarino, V., Rinner, O., Bork, P., Lemke, E.A., Beck, M., 2013. Cell type-specific nuclear pores: a case in point for context-dependent stoichiometry of molecular machines. Mol Syst Biol 9, 648. https://doi.org/10.1038/msb.2013.4

Ou, G., Blacque, O.E., Snow, J.J., Leroux, M.R., Scholey, J.M., 2005. Functional coordination of intraflagellar transport motors. Nature 436, 583. https://doi.org/10.1038/nature03818

Paczkowski, J.E., Richardson, B.C., Fromme, J.C., 2015. Cargo adaptors: structures illuminate mechanisms regulating vesicle biogenesis. Trends in Cell Biology 25, 408–416. https://doi.org/10.1016/j.tcb.2015.02.005

Poole, A.M., Gribaldo, S., 2014. Eukaryotic origins: How and when was the mitochondrion acquired? Cold Spring Harb Perspect Biol 6, a015990. https://doi.org/10.1101/cshperspect.a015990

Poole, A.M., Phillips, M.J., Penny, D., 2003. Prokaryote and eukaryote evolvability. Biosystems 69, 163–185. https://doi.org/10.1016/s0303-2647(02)00131-4

Prevo, B., Scholey, J.M., Peterman, E.J.G., 2017. Intraflagellar transport: mechanisms of motor action, cooperation, and cargo delivery. The FEBS Journal 284, 2905–2931. https://doi.org/10.1111/febs.14068

Pröschold, T., Harris, E.H., Coleman, A.W., 2005. Portrait of a Species. Genetics 170, 1601–1610. https://doi.org/10.1534/genetics.105.044503

Prunuske, A.J., Liu, J., Elgort, S., Joseph, J., Dasso, M., Ullman, K.S., 2006. Nuclear Envelope Breakdown Is Coordinated by Both Nup358/RanBP2 and Nup153, Two Nucleoporins with Zinc Finger Modules. MBoC 17, 760–769. https://doi.org/10.1091/mbc.e05-06-0485

Ramakrishnan, C., Maier, S., Walker, R.A., Rehrauer, H., Joekel, D.E., Winiger, R.R., Basso, W.U., Grigg, M.E., Hehl, A.B., Deplazes, P., Smith, N.C., 2019. An experimental genetically attenuated live vaccine to prevent transmission of Toxoplasma gondii by cats. Sci Rep 9, 1474. https://doi.org/10.1038/s41598-018-37671-8

Raven, P.H., 1970. A Multiple Origin for Plastids and Mitochondria. Science 169, 641–646. https://doi.org/10.1126/science.169.3946.641

Razi, M., Chan, E.Y.W., Tooze, S.A., 2009. Early endosomes and endosomal coatomer are required for autophagy. Journal of Cell Biology 185, 305–321. https://doi.org/10.1083/jcb.200810098

Robert, A., Margall-Ducos, G., Guidotti, J.-E., Brégerie, O., Celati, C., Bréchot, C., Desdouets, C., 2007. The intraflagellar transport component IFT88/polaris is a centrosomal protein regulating G1-S transition in non-ciliated cells. J Cell Sci 120, 628–637. https://doi.org/10.1242/jcs.03366

Roger, A.J., Muñoz-Gómez, S.A., Kamikawa, R., 2017. The Origin and Diversification of Mitochondria. Curr Biol 27, R1177–R1192. https://doi.org/10.1016/j.cub.2017.09.015

Roger, A.J., Simpson, A.G.B., 2009. Evolution: Revisiting the Root of the Eukaryote Tree. Current Biology 19, R165–R167. https://doi.org/10.1016/j.cub.2008.12.032

Rout, M.P., Aitchison, J.D., Suprapto, A., Hjertaas, K., Zhao, Y., Chait, B.T., 2000. The yeast nuclear pore complex: composition, architecture, and transport mechanism. J Cell Biol 148, 635–651. https://doi.org/10.1083/jcb.148.4.635

Rout, M.P., Field, M.C., 2017. The Evolution of Organellar Coat Complexes and Organization of the Eukaryotic Cell. Annual Review of Biochemistry 86, 637–657.

https://doi.org/10.1146/annurev-biochem-061516-044643

Sagan, L., 1967. On the origin of mitosing cells. Journal of Theoretical Biology 14, 225-IN6. https://doi.org/10.1016/0022-5193(67)90079-3

Samson, R.Y., Obita, T., Freund, S.M., Williams, R.L., Bell, S.D., 2008. A Role for the ESCRT System in Cell Division in Archaea. Science 322, 1710–1713. https://doi.org/10.1126/science.1165322

Sandman, K., Reeve, J.N., 2006. Archaeal histones and the origin of the histone fold. Current Opinion in Microbiology, Antimicrobials/Genomics 9, 520–525. https://doi.org/10.1016/j.mib.2006.08.003

Sandman, K., Reeve, J.N., 1998. Origin of the eukaryotic nucleus. Science 280, 501, 503. https://doi.org/10.1126/science.280.5363.499d

Sang, L., Miller, J.J., Corbit, K.C., Giles, R.H., Brauer, M.J., Otto, E.A., Baye, L.M., Wen, X., Scales, S.J., Kwong, M., Huntzicker, E.G., Sfakianos, M.K., Sandoval, W., Bazan, J.F., Kulkarni, P., Garcia-Gonzalo, F.R., Seol, A.D., O'Toole, J.F., Held, S., Reutter, H.M., Lane, W.S., Rafiq, M.A., Noor, A., Ansar, M., Devi, A.R.R., Sheffield, V.C., Slusarski, D.C., Vincent, J.B., Doherty, D.A., Hildebrandt, F., Reiter, J.F., Jackson, P.K., 2011. Mapping the NPHP-JBTS-MKS Protein Network Reveals Ciliopathy Disease Genes and Pathways. Cell 145, 513–528. https://doi.org/10.1016/j.cell.2011.04.019

Satir, P., Mitchell, D.R., Jékely, G., 2008. Chapter 3 How Did the Cilium Evolve?, in: Current Topics in Developmental Biology, Ciliary Function in Mammalian Development. Academic Press, pp. 63–82. https://doi.org/10.1016/S0070-2153(08)00803-X

Schlacht, A.D.W., 2016. Evolution of the vesicle formation machinery (Doctor of Philosophy). University of Alberta, Department of Cell Biology. https://doi.org/10.7939/R38K75473

Schwechheimer, C., Kuehn, M.J., 2015. Outer-membrane vesicles from Gram-negative bacteria: biogenesis and functions. Nat Rev Microbiol 13, 605–619. https://doi.org/10.1038/nrmicro3525

Scott, C.A., Marsden, A.N., Rebagliati, M.R., Zhang, Q., Chamling, X., Searby, C.C., Baye, L.M., Sheffield, V.C., Slusarski, D.C., 2017. Nuclear/cytoplasmic transport defects in BBS6 underlie congenital heart disease through perturbation of a chromatin remodeling protein. PLOS Genetics 13, e1006936. https://doi.org/10.1371/journal.pgen.1006936

Seitz, K.W., Lazar, C.S., Hinrichs, K.-U., Teske, A.P., Baker, B.J., 2016. Genomic reconstruction of a novel, deeply branched sediment archaeal phylum with pathways for acetogenesis and sulfur reduction. ISME J 10, 1696–1705. https://doi.org/10.1038/ismej.2015.233

Seo, S., Baye, L.M., Schulz, N.P., Beck, J.S., Zhang, Q., Slusarski, D.C., Sheffield, V.C., 2010. BBS6, BBS10, and BBS12 form a complex with CCT/TRiC family chaperonins and mediate BBSome assembly. PNAS 107, 1488–1493. https://doi.org/10.1073/pnas.0910268107

Sharma, K.R., Enzmann, B.L., Schmidt, Y., Moore, D., Jones, G.R., Parker, J., Berger, S.L., Reinberg, D., Zwiebel, L.J., Breit, B., Liebig, J., Ray, A., 2015. Cuticular Hydrocarbon Pheromones for Social Behavior and Their Coding in the Ant Antenna. Cell Reports 12, 1261–1271. https://doi.org/10.1016/j.celrep.2015.07.031

Shi, L., Koll, F., Arnaiz, O., Cohen, J., 2018. The Ciliary Protein IFT57 in the Macronucleus of Paramecium. Journal of Eukaryotic Microbiology 65, 12–27. https://doi.org/10.1111/jeu.12423

Simpson, A.G.B., Roger, A.J., 2004. The real 'kingdoms' of eukaryotes. Current Biology 14, R693–R696. https://doi.org/10.1016/j.cub.2004.08.038

Simpson, A.G.B., Roger, A.J., 2002. Eukaryotic Evolution: Getting to the Root of the Problem. Current Biology 12, R691–R693. https://doi.org/10.1016/S0960-9822(02)01207-1

Singh, S.K., Gui, M., Koh, F., Yip, M.C., Brown, A., 2020. Structure and activation mechanism of the BBSome membrane protein trafficking complex. eLife 9, e53322. https://doi.org/10.7554/eLife.53322

Skejo, J., Garg, S.G., Gould, S.B., Hendriksen, M., Tria, F.D.K., Bremer, N., Franjević, D., Blackstone, N.W., Martin, W.F., 2021. Evidence for a Syncytial Origin of Eukaryotes from Ancestral State Reconstruction. Genome Biology and Evolution 13, evab096. https://doi.org/10.1093/gbe/evab096

Smith, J.M., Szathmáry, E., 1997. The Major Transitions in Evolution. OUP Oxford.

Smith, T.F., 2013. Diversity of WD-repeat Proteins, in: Madame Curie Bioscience Database [Internet]. Landes Bioscience.

Smith, T.F., Gaitatzes, C., Saxena, K., Neer, E.J., 1999. The WD repeat: a common architecture for diverse functions. Trends in Biochemical Sciences 24, 181–185. https://doi.org/10.1016/S0968-0004(99)01384-5

Soni, K.G., Mardones, G.A., Sougrat, R., Smirnova, E., Jackson, C.L., Bonifacino, J.S., 2009. Coatomer-dependent protein delivery to lipid droplets. Journal of Cell Science 122, 1834–1841. https://doi.org/10.1242/jcs.045849

Soubannier, V., McLelland, G.-L., Zunino, R., Braschi, E., Rippstein, P., Fon, E.A., McBride, H.M., 2012. A Vesicular Transport Pathway Shuttles Cargo from Mitochondria to Lysosomes. Current Biology 22, 135–141. https://doi.org/10.1016/j.cub.2011.11.057

Spang, A., 2009. On vesicle formation and tethering in the ER–Golgi shuttle. Current Opinion in Cell Biology 21, 531–536. https://doi.org/10.1016/j.ceb.2009.03.003

Spang, A., Caceres, E.F., Ettema, T.J.G., 2017. Genomic exploration of the diversity, ecology, and evolution of the archaeal domain of life. Science 357, eaaf3883. https://doi.org/10.1126/science.aaf3883

Spang, A., Saw, J.H., Jørgensen, S.L., Zaremba-Niedzwiedzka, K., Martijn, J., Lind, A.E., van Eijk, R., Schleper, C., Guy, L., Ettema, T.J.G., 2015. Complex archaea that bridge the gap between prokaryotes and eukaryotes. Nature 521, 173–179. https://doi.org/10.1038/nature14447

Stewart, M., 2019. Structure and Function of the TREX-2 Complex. Subcell Biochem 93, 461–470. https://doi.org/10.1007/978-3-030-28151-9_15

Stewart, M., 2007. Molecular mechanism of the nuclear protein import cycle. Nat Rev Mol Cell Biol 8, 195–208. https://doi.org/10.1038/nrm2114

Strassert, J.F.H., Irisarri, I., Williams, T.A., Burki, F., 2021. A molecular timescale for eukaryote evolution with implications for the origin of red algal-derived plastids. Nat Commun 12, 1879. https://doi.org/10.1038/s41467-021-22044-z

Szathmáry, E., Smith, J.M., 1995. The major evolutionary transitions. Nature 374, 227–232. https://doi.org/10.1038/374227a0

Taherbhoy, A.M., Huang, O.W., Cochran, A.G., 2015. BMI1-RING1B is an autoinhibited RING E3 ubiquitin ligase. Nat Commun 6, 7621. https://doi.org/10.1038/ncomms8621

Takao, D., Verhey, K.J., 2016. Gated entry into the ciliary compartment. Cell. Mol. Life Sci. 73, 119–127. https://doi.org/10.1007/s00018-015-2058-0

Tamura, K., Fukao, Y., Iwamoto, M., Haraguchi, T., Hara-Nishimura, I., 2010. Identification and Characterization of Nuclear Pore Complex Components in Arabidopsis thaliana[W][OA]. Plant Cell 22, 4084–4097. https://doi.org/10.1105/tpc.110.079947

Taschner, M., Lorentzen, E., 2016. The Intraflagellar Transport Machinery. Cold Spring Harb Perspect Biol 8, a028092. https://doi.org/10.1101/cshperspect.a028092

Tavares, L., Dimitrova, E., Oxley, D., Webster, J., Poot, R., Demmers, J., Bezstarosti, K., Taylor, S., Ura, H., Koide, H., Wutz, A., Vidal, M., Elderkin, S., Brockdorff, N., 2012. RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. Cell 148, 664–678. https://doi.org/10.1016/j.cell.2011.12.029

Thiam, A.R., Antonny, B., Wang, J., Delacotte, J., Wilfling, F., Walther, T.C., Beck, R., Rothman, J.E., Pincet, F., 2013. COPI buds 60-nm lipid droplets from reconstituted water–phospholipid–triacylglyceride interfaces, suggesting a tension clamp function. Proceedings of the National Academy of Sciences 110, 13244–13249. https://doi.org/10.1073/pnas.1307685110

Tian, X., Zhao, H., Zhou, J., 2023. Organization, functions, and mechanisms of the BBSome in development, ciliopathies, and beyond. eLife 12, e87623. https://doi.org/10.7554/eLife.87623

Todd, A.G., Lin, H., Ebert, A.D., Liu, Y., Androphy, E.J., 2013. COPI transport complexes bind to specific RNAs in neuronal cells. Human Molecular Genetics 22, 729–736. https://doi.org/10.1093/hmg/dds480

Trapp, J., McAfee, A., Foster, L.J., 2017. Genomics, transcriptomics and proteomics: enabling insights into social evolution and disease challenges for managed and wild bees. Mol Ecol 26, 718–739. https://doi.org/10.1111/mec.13986

Trautwein, M., Dengjel, J., Schirle, M., Spang, A., 2004. Arf1p Provides an Unexpected Link between COPI Vesicles and mRNA in Saccharomyces cerevisiae. MBoC 15, 5021–5037. https://doi.org/10.1091/mbc.e04-05-0411

van Dam, T.J.P., Townsend, M.J., Turk, M., Schlessinger, A., Sali, A., Field, M.C., Huynen, M.A., 2013. Evolution of modular intraflagellar transport from a coatomer-like progenitor. PNAS 110, 6943–6948. https://doi.org/10.1073/pnas.1221011110

Van Valen, L.M., Maiorana, V.C., 1980. The Archaebacteria and eukaryotic origins. Nature 287, 248–250. https://doi.org/10.1038/287248a0

Veland, I.R., Awan, A., Pedersen, L.B., Yoder, B.K., Christensen, S.T., 2009. Primary Cilia and Signaling Pathways in Mammalian Development, Health and Disease. NEP 111, p39–p53. https://doi.org/10.1159/000208212

Venard, C.M., Vasudevan, K.K., Stearns, T., 2020. Cilium axoneme internalization and degradation in chytrid fungi. Cytoskeleton 77, 365–378. https://doi.org/10.1002/cm.21637

Vigers, G.P., Crowther, R.A., Pearse, B.M., 1986. Three-dimensional structure of clathrin cages in ice. EMBO J 5, 529–534. https://doi.org/10.1002/j.1460-2075.1986.tb04242.x

Vincensini, L., Blisnick, T., Bastin, P., 2011. 1001 model organisms to study cilia and flagella. Biology of the Cell 103, 109–130. https://doi.org/10.1042/BC20100104

Wallberg, A., Pirk, C.W., Allsopp, M.H., Webster, M.T., 2016. Identification of Multiple Loci Associated with Social Parasitism in Honeybees. PLOS Genetics 12, e1006097. https://doi.org/10.1371/journal.pgen.1006097

Wallin, I.E., 1923. The Mitochondria Problem. The American Naturalist 57, 255–261.

Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., Zhang, Y., 2004. Role of histone H2A ubiquitination in Polycomb silencing. Nature 431, 873–878. https://doi.org/10.1038/nature02985

Wang, L., Wen, X., Wang, Z., Lin, Z., Li, C., Zhou, H., Yu, H., Li, Y., Cheng, Y., Chen, Y., Lou, G., Pan, J., Cao, M., 2022. Ciliary transition zone proteins coordinate ciliary

protein composition and ectosome shedding. Nat Commun 13, 3997. https://doi.org/10.1038/s41467-022-31751-0

Weinstock, G.M., Robinson, G.E., Gibbs, R.A., Weinstock, G.M., Weinstock, G.M., Robinson, G.E., Worley, K.C., Evans, J.D., Maleszka, R., Robertson, H.M., Weaver, D.B., Beye, M., Bork, P., Elsik, C.G., Evans, J.D., Hartfelder, K., Hunt, G.J., Robertson, H.M., Robinson, G.E., Maleszka, R., Weinstock, G.M., Worley, K.C., Zdobnov, E.M., Hartfelder, K., Amdam, G.V., Bitondi, M.M.G., Collins, A.M., Cristino, A.S., Evans, J.D., Michael, H., Lattorff, G., Lobo, C.H., Moritz, R.F.A., Nunes, F.M.F., Page, R.E., Simões, Z.L.P., Wheeler, Diana, Carninci, P., Fukuda, S., Hayashizaki, Y., Kai, C., Kawai, J., Sakazume, N., Sasaki, D., Tagami, M., Maleszka, R., Amdam, G.V., Albert, S., Baggerman, G., Beggs, K.T., Bloch, G., Cazzamali, G., Cohen, M., Drapeau, M.D., Eisenhardt, D., Emore, C., Ewing, M.A., Fahrbach, S.E., Forêt, S., Grimmelikhuijzen, C.J.P., Hauser, F., Hummon, A.B., Hunt, G.J., Huybrechts, J., Jones, A.K., Kadowaki, T., Kaplan, N., Kucharski, R., Leboulle, G., Linial, M., Littleton, J.T., Mercer, A.R., Page, R.E., Robertson, H.M., Robinson, G.E., Richmond, T.A., RodriguezZas, S.L., Rubin, E.B., Sattelle, D.B., Schlipalius, D., Schoofs, L., Shemesh, Y., Sweedler, J.V., Velarde, R., Verleyen, P., Vierstraete, E., Williamson, M.R., Beye, M., Ament, S.A., Brown, S.J., Corona, M., Dearden, P.K., Dunn, W.A., Elekonich, M.M., Elsik, C.G., Forêt, S., Fujiyuki, T., Gattermeier, I., Gempe, T., Hasselmann, M., Kadowaki, T., Kage, E., Kamikouchi, A., Kubo, T., Kucharski, R., Kunieda, T., Lorenzen, M., Maleszka, R., Milshina, N.V., Morioka, M., Ohashi, K., Overbeek, R., Page, R.E., Robertson, H.M., Robinson, G.E., Ross, C.A., Schioett, M., Shippy, T., Takeuchi, H., Toth, A.L., Willis, J.H., Wilson, M.J., Robertson, H.M., Zdobnov, E.M., Bork, P., Elsik, C.G., Gordon, K.H.J., Letunic, I., Hackett, K., Peterson, J., Felsenfeld, A., Guyer, M., Solignac, M., Agarwala, R., Cornuet, J.M., Elsik, C.G., Emore, C., Hunt, G.J., Monnerot, M., Mougel, F., Reese, J.T., Schlipalius, D., Vautrin, D., Weaver, D.B., Gillespie, J.J., Cannone, J.J., Gutell, R.R., Johnston, J.S., Elsik, C.G., Cazzamali, G., Eisen, M.B., Grimmelikhuijzen, C.J.P., Hauser, F., Hummon, A.B., Iyer, V.N., Iyer, V., Kosarev, P., Mackey, A.J., Maleszka, R., Reese, J.T., Richmond, T.A., Robertson, H.M., Solovyev, V., Souvorov, A., Sweedler, J.V., Weinstock, G.M., Williamson, M.R., Zdobnov, E.M., Evans, J.D., Aronstein, K.A., Bilikova, K., Chen, Y.P., Clark, A.G., Decanini, L.I., Gelbart, W.M., Hetru, C., Hultmark, D., Imler, J.-L., Jiang, Haobo, Kanost, M., Kimura, K., Lazzaro, B.P., Lopez, D.L., Simuth, J., Thompson, G.J., Zou, Z., De Jong, P., Sodergren, E., Csűrös, M., Milosavljevic, A., Johnston, J.S., Osoegawa, K., Richards, S., Shu, C.-L., Weinstock, G.M., Elsik, C.G., Duret, L., Elhaik, E., Graur, D., Reese, J.T., Robertson, H.M., Robertson, H.M., Elsik, C.G., Maleszka, R., Weaver, D.B., Amdam, G.V., Anzola, J.M., Campbell, K.S., Childs, K.L., Collinge, D., Crosby, M.A., Dickens, C.M., Elsik, C.G., Gordon, K.H.J., Gramates, L.S., Grozinger, C.M., Jones, P.L., Jorda, M., Ling, X., Matthews, B.B., Miller, J., Milshina, N.V., Mizzen, C., Peinado, M.A., Reese, J.T., Reid, J.G., Robertson, H.M., Robinson, G.E., Russo, S.M., Schroeder, A.J., St Pierre, S.E., Wang, Y., Zhou, P., Robertson, H.M., Agarwala, R., Elsik, C.G., Milshina, N.V., Reese, J.T., Weaver, D.B., Worley, K.C., Childs, K.L., Dickens, C.M., Elsik, C.G., Gelbart, W.M., Jiang, Huaiyang, Kitts, P., Milshina, N.V., Reese, J.T., Ruef, B., Russo, S.M., Venkatraman, A., Weinstock, G.M., Zhang, L., Zhou, P., Johnston, J.S., Aquino-Perez, G., Cornuet, J.M., Monnerot, M., Solignac, M., Vautrin, D., Whitfield, C.W., Behura, S.K., Berlocher, S.H., Clark, A.G., Gibbs, R.A., Johnston, J.S., Sheppard, W.S., Smith, D.R., Suarez, A.V., Tsutsui, N.D., Weaver, D.B., Wei, X., Wheeler, David, Weinstock, G.M., Worley, K.C., Havlak, P., The Honeybee Genome Sequencing Consortium, Overall project leadership:, Principal investigators:, Community coordination:, Annotation section leaders:, Caste development and reproduction:, EST sequencing:, Brain and behaviour:, Development and metabolism:, Comparative and evolutionary analysis:, Funding agency management:, Physical and genetic mapping:, Ribosomal RNA genes and related retrotransposable elements:, Gene prediction and consensus gene set:, Honeybee disease and immunity:, BAC/fosmid library construction and analysis:, G+C content:, Transposable elements:, Gene regulation including miRNA and RNAi:, Superscaffold assembly:, Data management:, Chromosome structure:, Population genetics and SNPs:, Genome assembly:, 2006. Insights into social insects from the genome of the honeybee Apis mellifera. Nature 443, 931–949. https://doi.org/10.1038/nature05260

Whittaker, R.H., 1969. New Concepts of Kingdoms of Organisms. Science 163, 150–160. https://doi.org/10.1126/science.163.3863.150

Wickstead, B., Gull, K., 2007. Dyneins Across Eukaryotes: A Comparative Genomic Analysis. Traffic 8, 1708–1721. https://doi.org/10.1111/j.1600-0854.2007.00646.x

Wilfling, F., Thiam, A.R., Olarte, M.-J., Wang, J., Beck, R., Gould, T.J., Allgeyer, E.S., Pincet, F., Bewersdorf, J., Farese, R.V., Walther, T.C., 2014. Arf1/COPI machinery acts directly on lipid droplets and enables their connection to the ER for protein targeting. Elife 3, e01607. https://doi.org/10.7554/eLife.01607

Wilson, K.L., Dawson, S.C., 2011. Functional evolution of nuclear structure. Journal of Cell Biology 195, 171–181. https://doi.org/10.1083/jcb.201103171

Woese, C.R., Fox, G.E., 1977. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. Proceedings of the National Academy of Sciences 74, 5088–5090. https://doi.org/10.1073/pnas.74.11.5088

Zabezhinsky, D., Slobodin, B., Rapaport, D., Gerst, J.E., 2016. An Essential Role for COPI in mRNA Localization to Mitochondria and Mitochondrial Function. Cell Reports 15, 540–549. https://doi.org/10.1016/j.celrep.2016.03.053

Zaremba-Niedzwiedzka, K., Caceres, E.F., Saw, J.H., Bäckström, D., Juzokaite, L., Vancaester, E., Seitz, K.W., Anantharaman, K., Starnawski, P., Kjeldsen, K.U., Stott, M.B., Nunoura, T., Banfield, J.F., Schramm, A., Baker, B.J., Spang, A., Ettema, T.J.G., 2017. Asgard archaea illuminate the origin of eukaryotic cellular complexity. Nature 541, 353–358. https://doi.org/10.1038/nature21031

Zhang, Q., Yu, D., Seo, S., Stone, E.M., Sheffield, V.C., 2012. Intrinsic Protein-Protein Interaction-mediated and Chaperonin-assisted Sequential Assembly of Stable Bardet-Biedl Syndrome Protein Complex, the BBSome. J. Biol. Chem. 287, 20625–20635. https://doi.org/10.1074/jbc.M112.341487

Zhu, L., Wang, Z., Wang, W., Wang, C., Hua, S., Su, Z., Brako, L., Garcia-Barrio, M., Ye, M., Wei, X., Zou, H., Ding, X., Liu, L., Liu, X., Yao, X., 2015. Mitotic Protein CSPP1 Interacts with CENP-H Protein to Coordinate Accurate Chromosome Oscillation in Mitosis. J. Biol. Chem. 290, 27053–27066. https://doi.org/10.1074/jbc.M115.658534

Zilman, A., 2009. Effects of multiple occupancy and interparticle interactions on selective transport through narrow channels: theory versus experiment. Biophys J 96, 1235–1248. https://doi.org/10.1016/j.bpj.2008.09.058

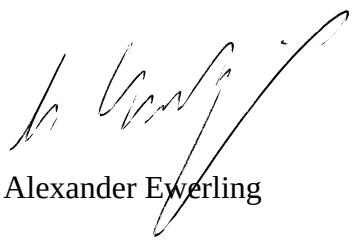
Zilman, A., Di Talia, S., Chait, B.T., Rout, M.P., Magnasco, M.O., 2007. Efficiency, selectivity, and robustness of nucleocytoplasmic transport. PLoS Comput Biol 3, e125. https://doi.org/10.1371/journal.pcbi.0030125

STATUTORY DECLARATION

I hereby declare that I have done this work independently and without the use of any aids other than those specified. All parts which are taken literally or analogously from publications or other sources are clearly marked as such. The work has not yet been published in the same or similar form and has not yet been submitted to an examination authority.

I am aware that the violation of this regulation will lead to failure of the thesis.

Mainz, August 2023



Alexander Ewerling

ACKNOWLEDGEMENTS

Firstly, I want to thank Helen May-Simera for supervising me since my Master's up until now. It has been an incredibly eventful journey. Kudos to you, not only for treading into unfamiliar territory with this project, but also for believing in me to be the right one for the job. I was surely the odd one out with the prospective zoo we wanted to start running in the lab, but your enthusiasm for the project kept the spirits up. Your guidance was immeasurably valuable, and I think some of your optimistic spirit seems to have rubbed off on me after all. Thank you.

My second thanks goes out to Bill Wickstead. Neither of us probably were expecting that you would be dragged into this project as much as you were, but I sure am glad it turned out this way. Learning from you has not only taught me hands-on bioinformatics, but also approaching problems in ways I had never imagined before, and that has come in handy more than once. I always aspired to be an evolutionary biologist, and working with you has brought me one step closer to this goal. It was a pleasure working with you. Cheers.

A big thanks also to Martin Kaltenpoth and Susanne Gerber, for 'holding our horses' when Helen and I got a bit too excited and overshot the goal with what we wanted to implement into the project. Meetings and discussions with you were always fruitful and provided us with a better, streamlined train of thought in the end. Thanks a lot.

I want to thank all of the people from the lovely May-Simera group, past, present and future, who kept me company throughout the last few years, in dire times and seasons of joy. Doing a PhD without you would not have been the same, and you bunch know how to make life worthwhile. Special thanks to Vanessa, for putting up with me in your early days and supporting my work from the beginning. Thank you.

Rike – I know you do not want to hear this and it is going to be incredibly awkward for you. You have had an enormous impact on my life. I did not imagine things would go the way they did, but I am happy that they did so. Your continuing support has made hard times bearable and kept me sane, and I can only hope to support you in the same way in times of need. Thank you. And also thanks to Linda, of course.

The last, but surely not least, thank you goes out to my friends and family, who kept me running through this wild time in more ways than one. With you, there is possibly nothing I could not face, and for that I am eternally grateful.

Eines zu seyn mit Allem, was lebt, in seeliger
Selbstvergessenheit wiederzukehren in's All der Natur,
das ist der Gipfel der Gedanken und Freuden,
das ist die heilige Bergeshöhe, der Ort der ewigen Ruhe,
wo der Mittag seine Schwüle und der Donner
seine Stimme verliert und das kochende Meer
der Wooge des Kornfeldes gleicht.

Auf dieser Höh' steh ich oft, mein Bellarmin!
Aber ein Moment des Besinnens wirft mich herab.
Ich denke nach und finde mich, wie ich zuvor war,
allein, mit allen Schmerzen der Sterblichkeit,
und meines Herzens Asyl, die ewig einige Welt, ist hin;
die Natur verschließt die Arme, und ich stehe,
wie ein Fremdling, vor ihr, und verstehe sie nicht.

– Friedrich Hölderlin: Hyperion oder der Eremit in Griechenland (1797/99)