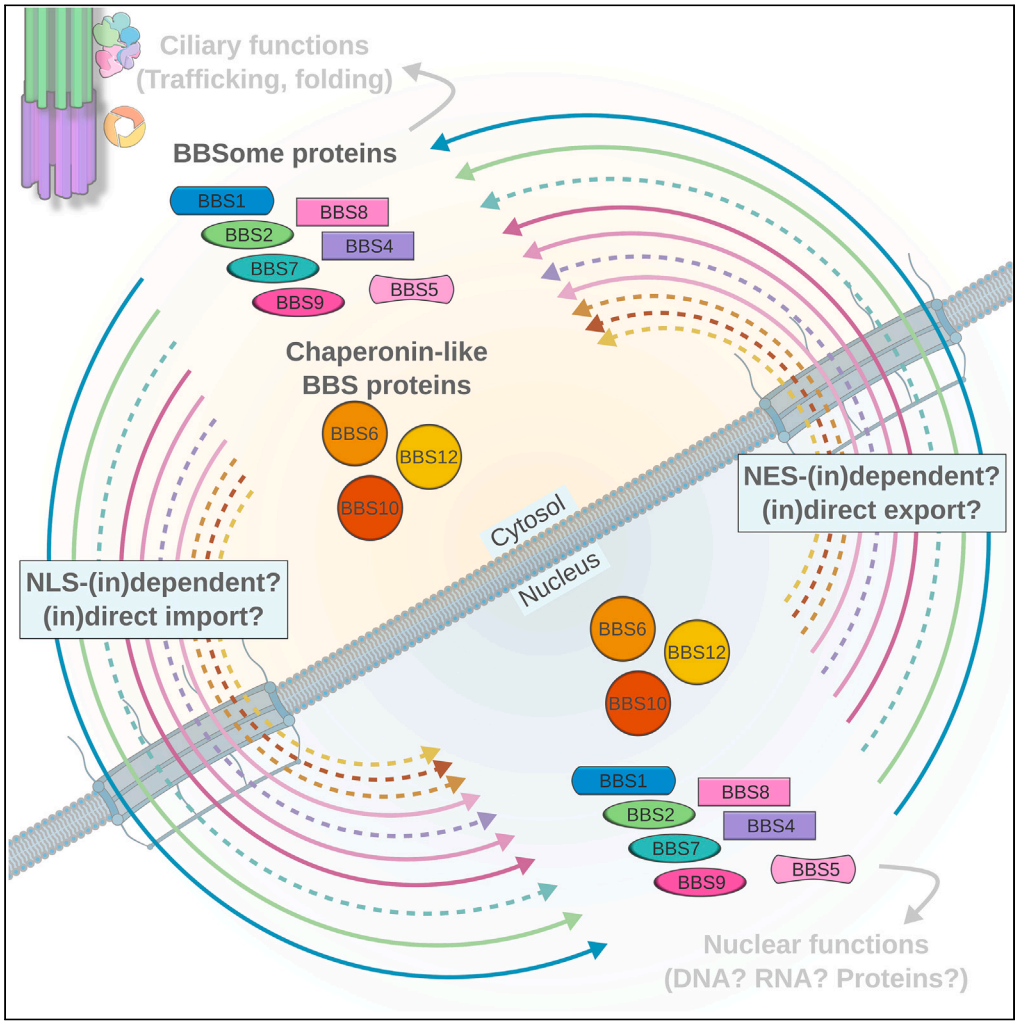
iScience

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

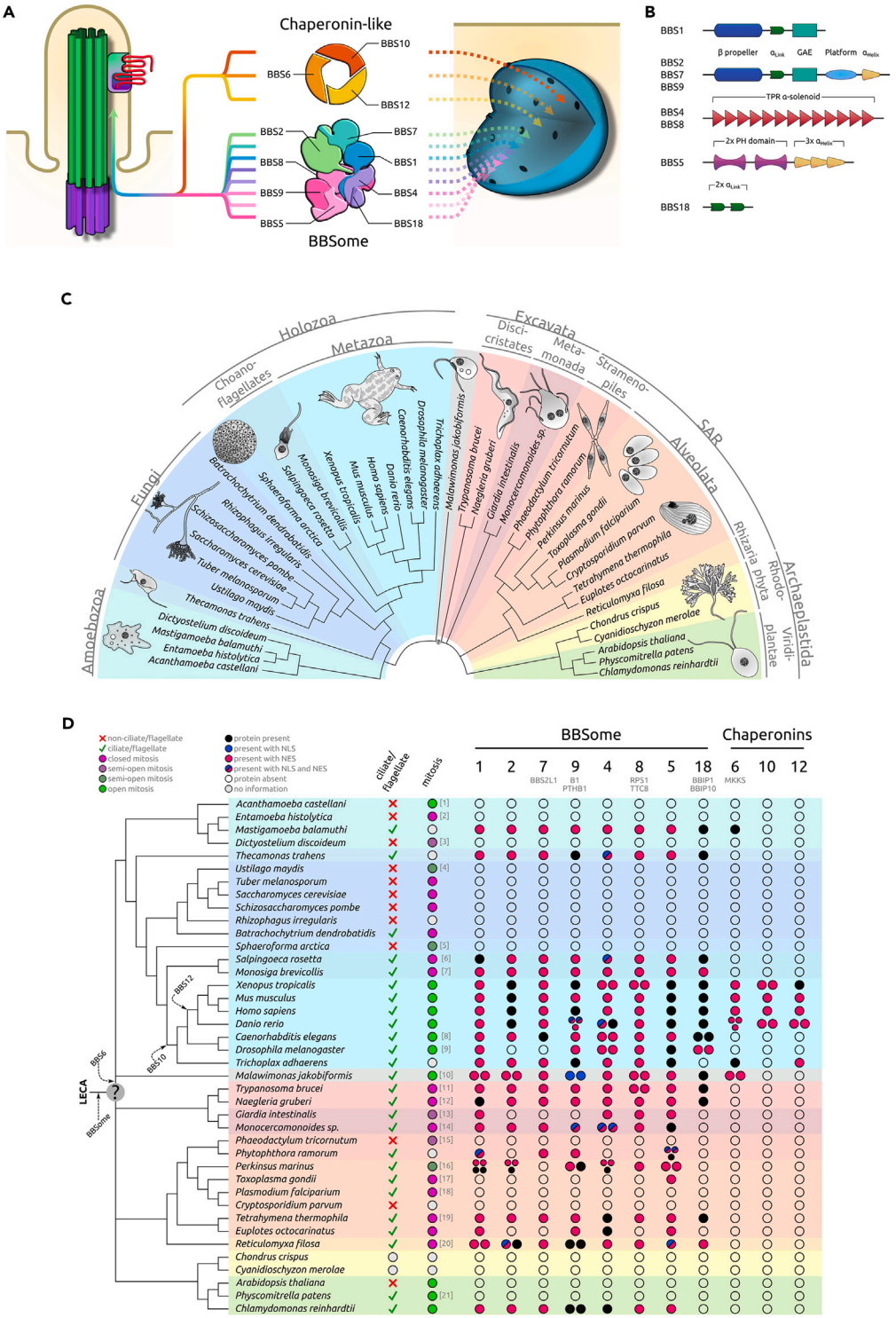


Figure 1. BBS proteins in the eukaryotic Tree of Life

(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.

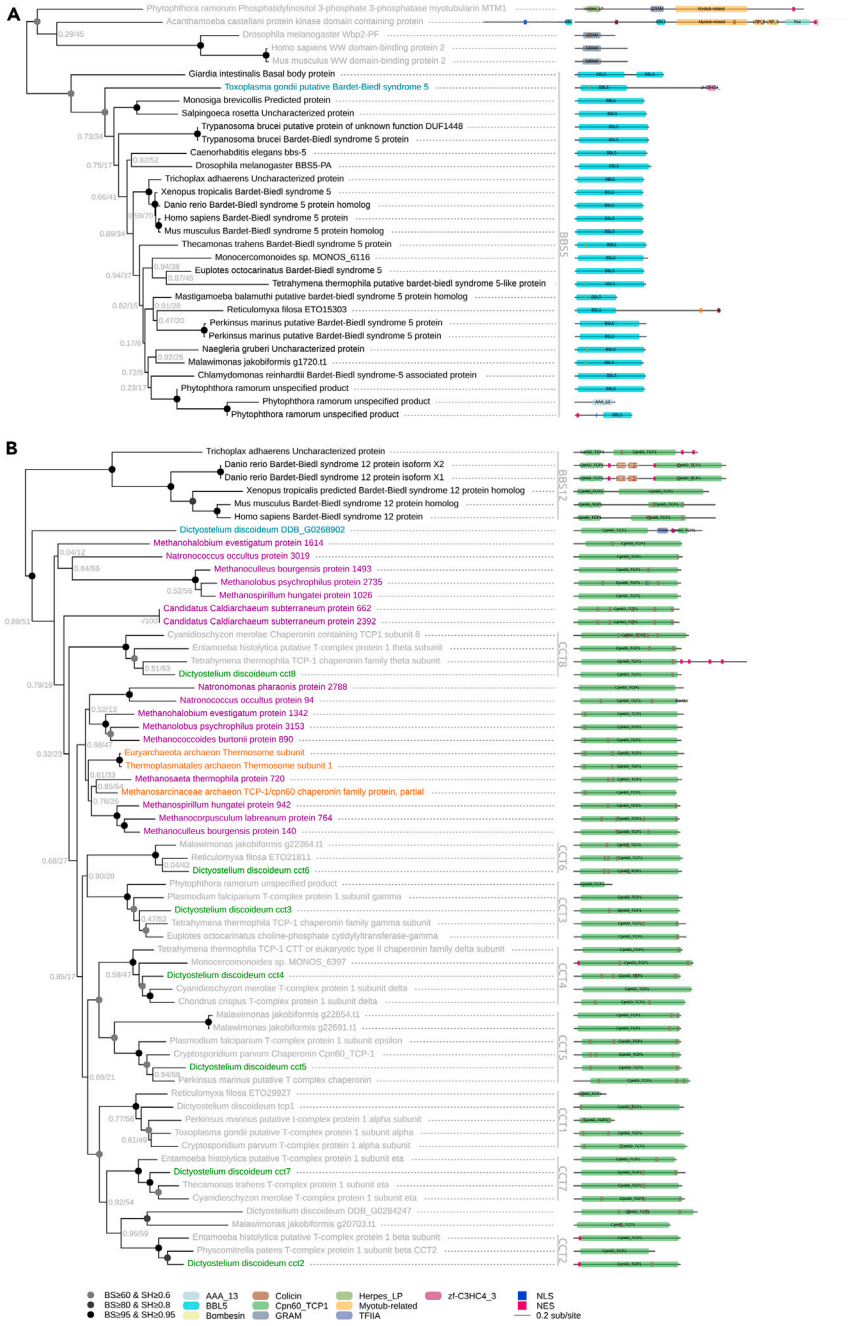


Figure 2. Maximum-likelihood (ML) phylogenetic relationships of BBS5 and BBS12

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)

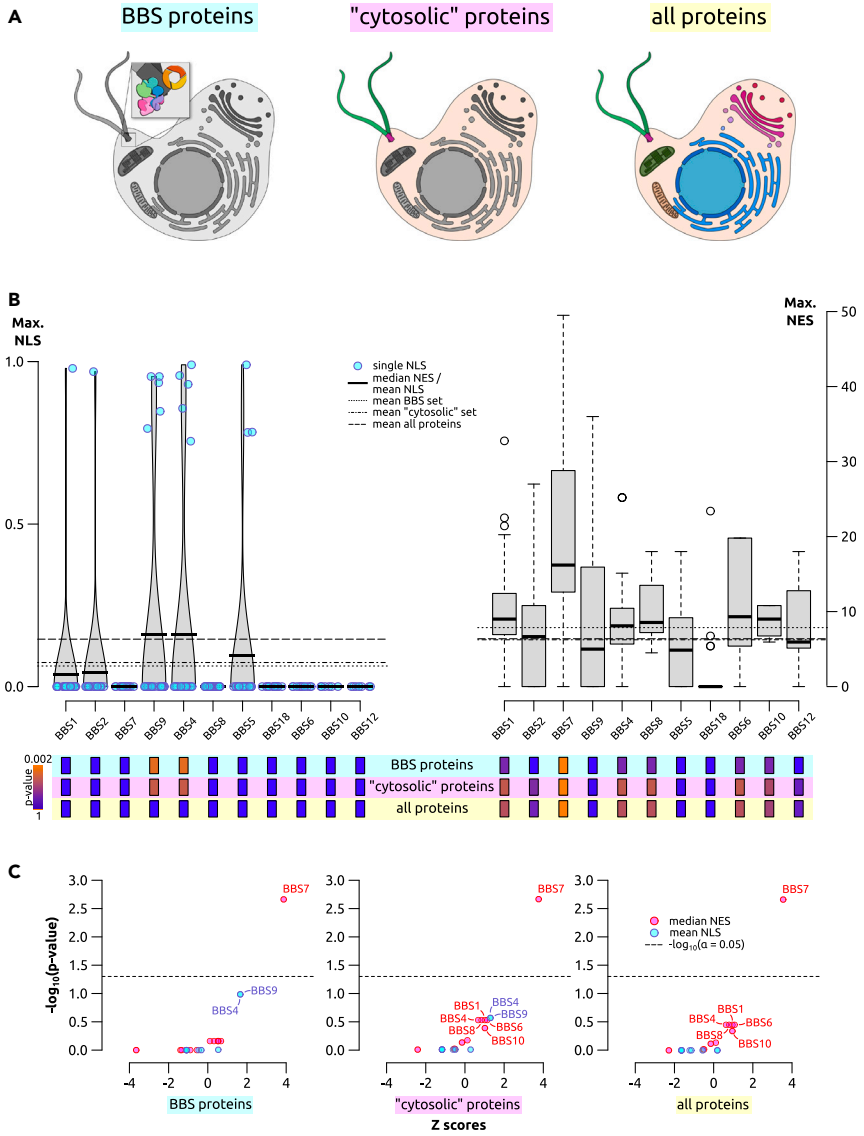


Figure 3. Distribution of predicted nuclear signal sequences in BBS orthologues

(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

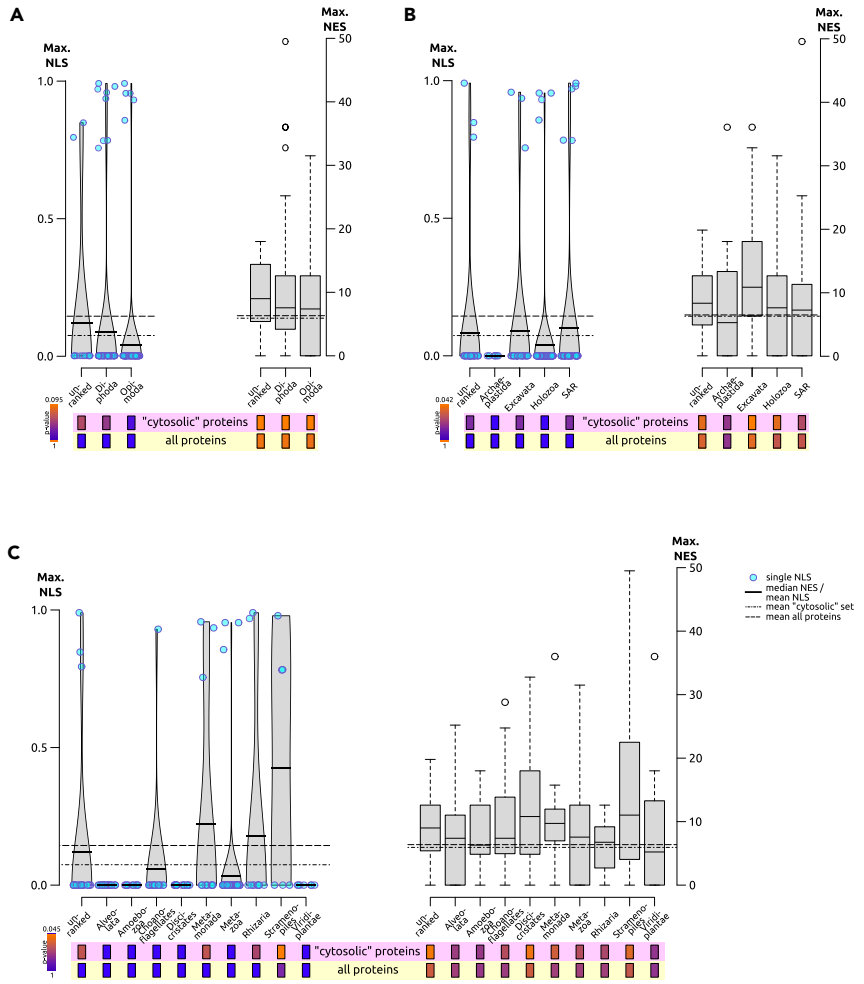


Figure 4. Comparison of BBS protein nuclear signal sequence scores to other proteins of the same taxonomic unit

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

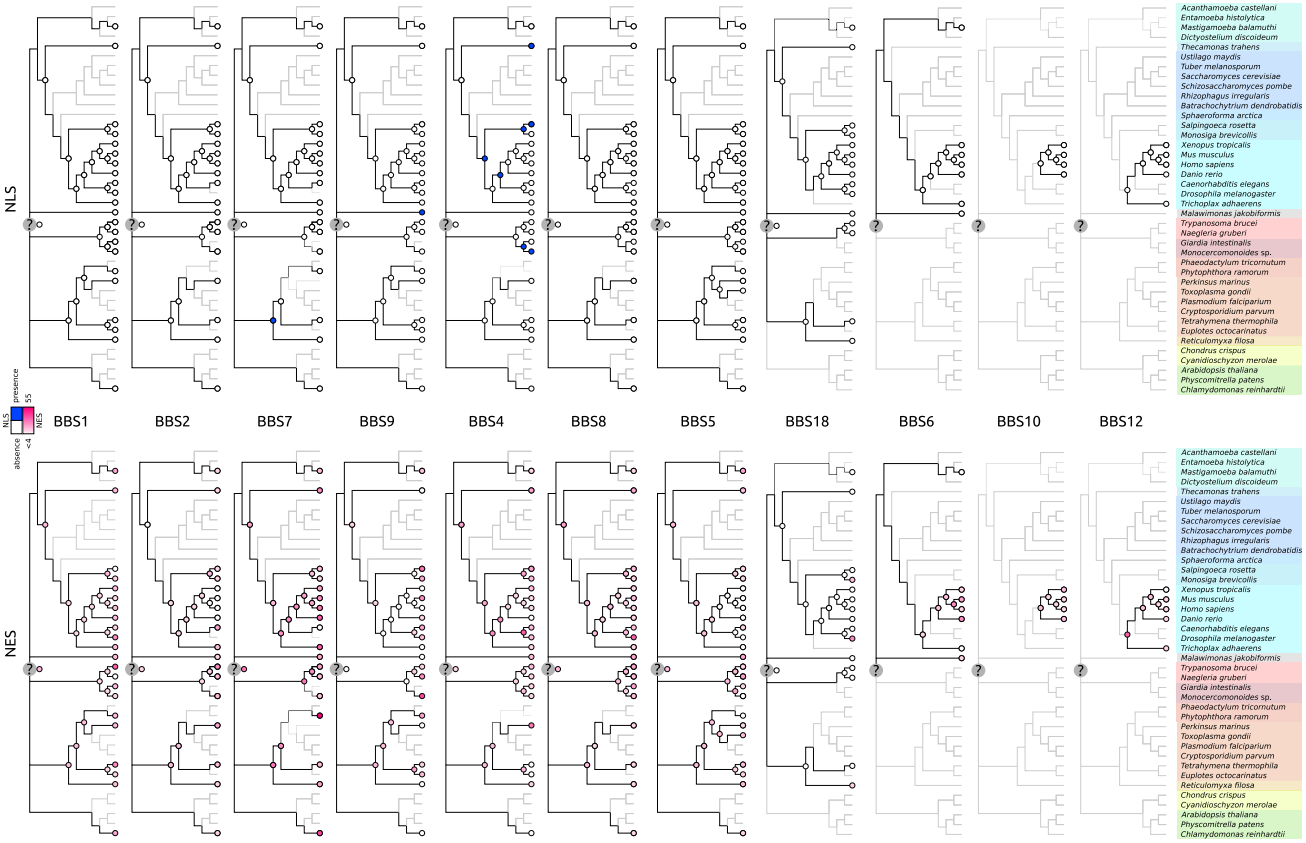


Figure 5. Ancestral sequence reconstruction of BBS proteins

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

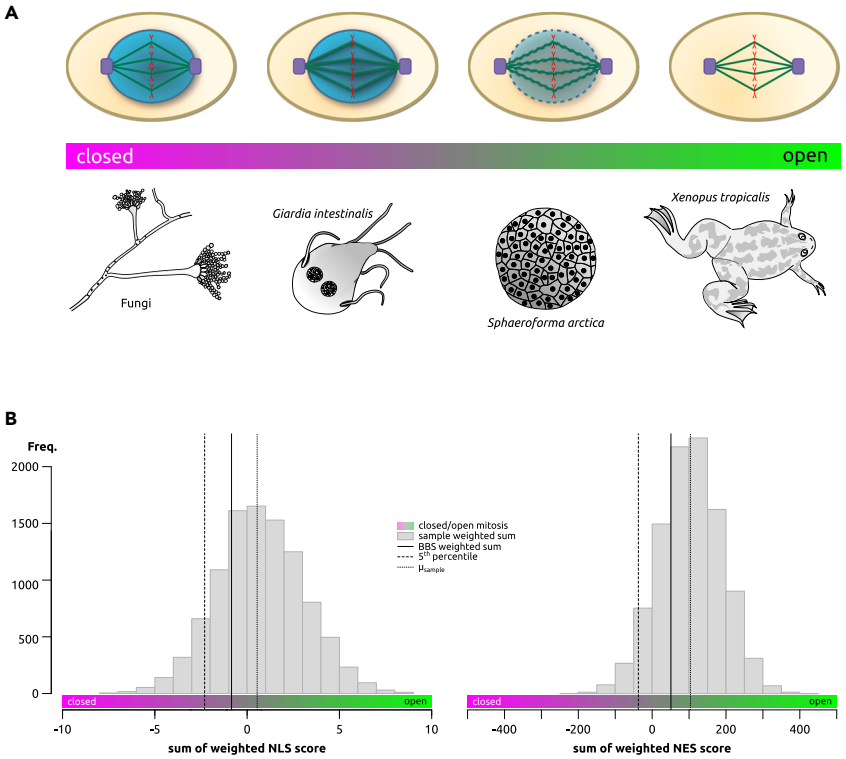


Figure 6. Mode of mitosis and presence of predicted nuclear signal sequences are not correlated

(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]

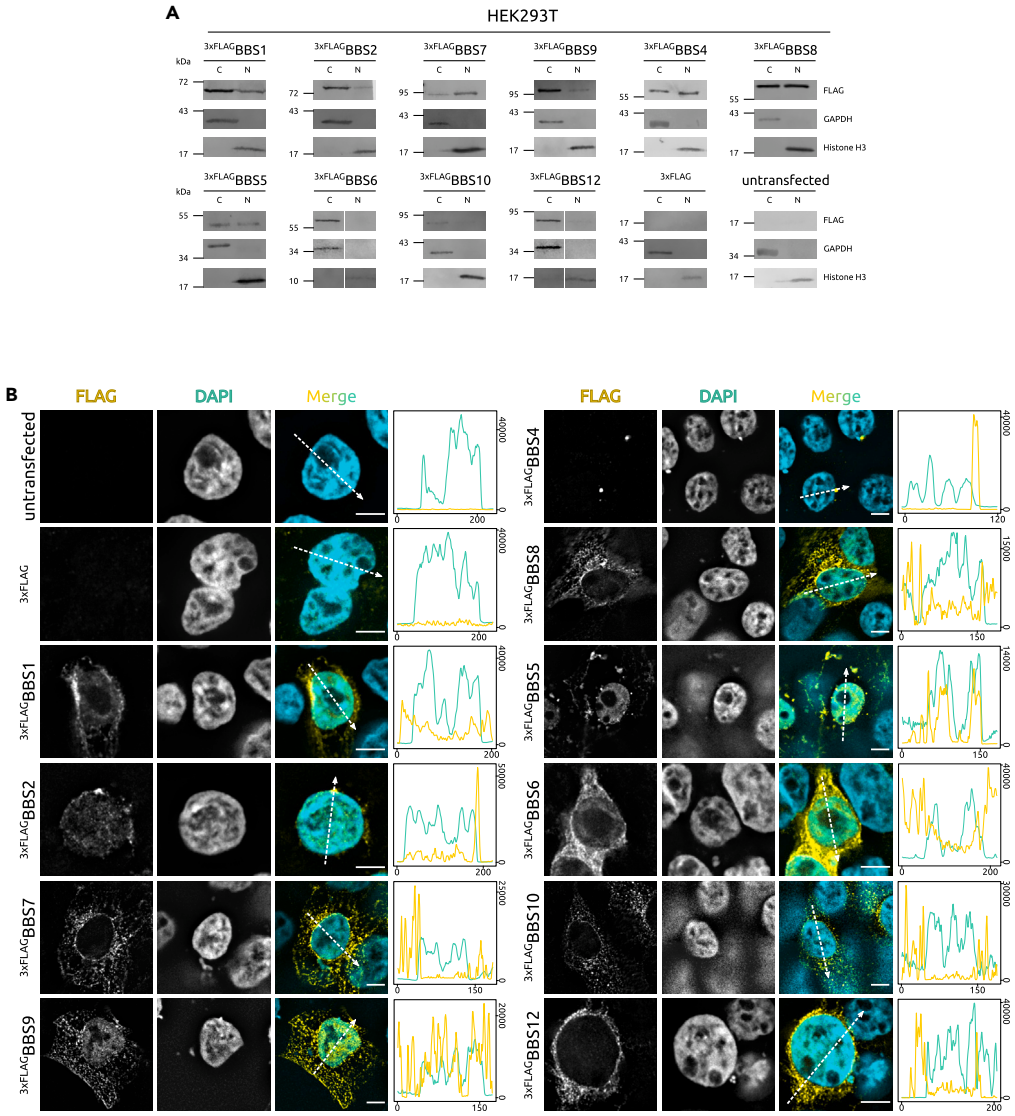


Figure 7. Nuclear localization of different BBS proteins

(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

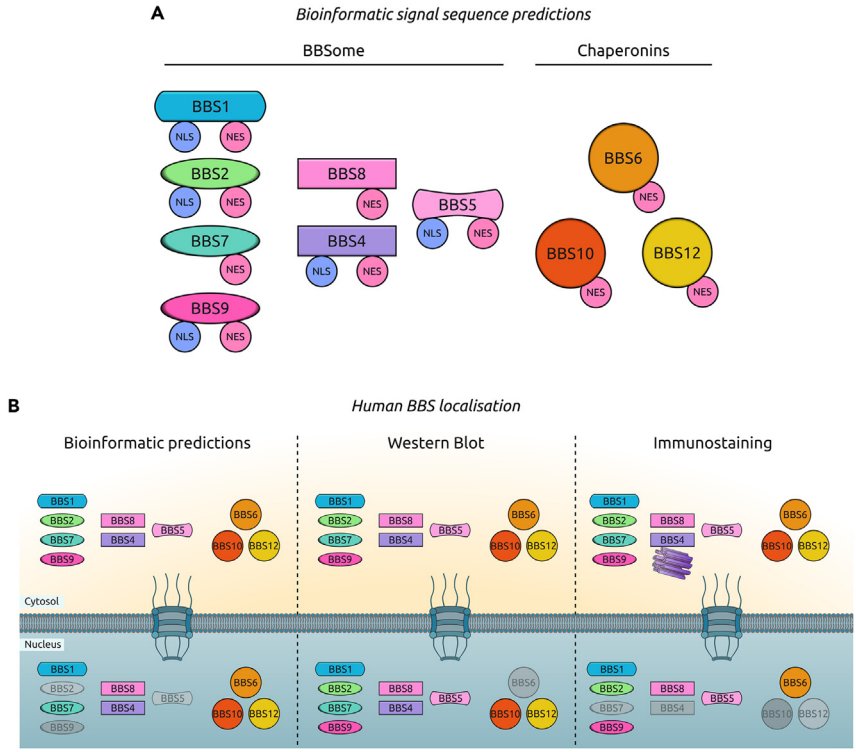


Figure 8. Schematic summary of bioinformatic, molecular and immunological experiments

(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, $0.05 \leq p$; *: $p < 0.05$.