Centrin diversity and basal body patterning across evolution: new insights from *Paramecium*

Anne Aubusson-Fleury¹⁺, Guillaume Balavoine², Michel Lemullos¹, Khaled Bouhouche³, Janine Beisson¹ and France Koll¹

**ABSTRACT**

First discovered in unicellular eukaryotes, centrins play crucial roles in basal body duplication and anchoring mechanisms. While the evolutionary status of the founding members of the family, Centrin2/ Vfl2 and Centrin3/cdc31 has long been investigated, the evolutionary origin of other members of the family has received less attention. Using a phylogeny of ciliate centrins, we identify two other centrin families, the ciliary centrins and the centrins present in the contractile filaments (ICL centrins). In this paper, we carry on the functional analysis of still not well-known centrins, the ICL1e subfamily identified in *Paramecium*, and show their requirement for correct basal body anchoring through interactions with Centrin2 and Centrin3. Using *Paramecium* as well as a eukaryote-wide sampling of centrins from completely sequenced genomes, we revisited the evolutionary story of centrins. Their phylogeny shows that the centrins associated with the ciliate contractile filaments are widespread in eukaryotic lineages and could be as ancient as Centrin2 and Centrin3.

**KEY WORDS:** Ciliary centrins, Basal body assembly, Basal body anchoring, Ciliated epithelia polarity, Centrin evolution

**INTRODUCTION**

The role of centrins in centriole/basal body/spindle pole body duplication (Salisbury et al., 1984; Wright et al., 1985; Baum et al., 1986) is a long-standing debate. It is now broadly agreed that centrins may be dispensable for centriole duplication, but are required for basal body biogenesis. This primary association of centrins with basal bodies is consistent with the generally admitted hypotheses that a cilium and associated basal body or basal body-centricrin diversity and function among eukaryotes.

**RESULTS**

Five functional centrin families can be identified in *Paramecium*

We revisited the phylogenetic relationships of centrin families in Ciliata in using the completely sequenced organisms: two *Paramecium* genomes, *P. tetraurelia* used in this study and *P. caudatum* which has undergone two less whole genome duplications resulting in less redundancy (McGrath et al., 2014);...
Tetrahymena thermophila, another species from the same taxonomic group (Adl et al., 2012); and an evolutionary more distant species, Oxytricha trifallax. 50 genes encoding centrin were found in P. tetraurelia, 22 in P. caudatum, 15 in T. thermophila, and 12 in O. trifallax. The resulting tree (Fig. 1A) shows a number of branches with sequences representative of the four species, likely representing ancestral ciliate families. Combining phylogenetic analysis and functional data, five main functional families can be defined (Fig. 1A). The first two families correspond to the classical basal body associated centrins, Centrin2 (Centrin2/VFL2) and Centrin3 (Centrin3/cdc31) (Ruiz et al., 2005). The phylogenetic position of four closely related paralogues (PtCen1, 5, 14 and 16) suggests that they could form a new family. PtCen1 and PtCen14 were identified in the ciliary proteome (Cildb, Arnaiz et al., 2014). GFP-tagged PtCen14 specifically localised to cilia (Fig. 1C) and its RNAi depletion resulted in altered swimming behaviour (Fig. S1). These centrins can thus be considered as ciliary centrins. The function of such a ciliary centrin has already been analysed in another protist, Trypanosoma (Wei et al., 2014). The fourth family contains the P. caudatum centrin required for the function of a ciliary voltage gated calcium channel (VGCC centrin, Gonda et al., 2004) and its P. tetraurelia ortholog, which is supposed to fulfill the same function. Finally, the ICL centrins, required for the assembly of the contractile network (Gogendeau et al., 2008) are grouped into a large functional fifth family (Fig. 1D). These sequences, quite divergent between ciliate species, do not form a common cluster. However, sequences representatives of P. caudatum and Tetrahymena are found in most branches corresponding to the previously defined ICL centrin subfamilies of P. tetraurelia, suggesting that they diverged before the emergence of these two genera.

Sequences representative of the four species are found in several well-supported branches: Centrin2, Centrin3, Ciliary and VGCC centrins. Strikingly, within the ICL centrins, only the ICL1e subfamily also comprises at least one sequence from all species including the more distantly related genus, O. trifallax. This suggests that ICL1e centrins could be ancient. It then seemed of interest to re-investigate the role of all members of this ICL1e subfamily in basal body biogenesis and organisation of the ICL.

Localisation of the ICL1e centrins

The ICL1e subfamily comprises seven genes coding for centrins corresponding to three ohnolog pairs, i.e. gene duplicates originating from whole genome duplication (ICL1e and ICL1g; Cen8 and Cen10; Cen12 and Cen15) and a single gene, Cen18. In order to ascertain their respective localisation and function, a representative from each ohnolog group (ICL1e, Cen10, Cen15, Cen18) was GFP-tagged. GFP-fusion constructions were introduced into two types of cells, the wild type (WT) and a mutant (BP) deleted for the Sfi-like centrin-binding protein and devoid of ICL. BP cells retain only a remnant of the infraciliary lattice closely associated to each basal body or basal body couplet, called ICL-organising centre (ICLOC) as it nucleates the assembly of the ICL under rescue conditions.

Localisation of ICL1e centrins

The ICL1e subfamily comprises seven genes coding for centrins corresponding to three ohnolog pairs, i.e. gene duplicates originating from whole genome duplication (ICL1e and ICL1g; Cen8 and Cen10; Cen12 and Cen15) and a single gene, Cen18. In order to ascertain their respective localisation and function, a representative from each ohnolog group (ICL1e, Cen10, Cen15, Cen18) was GFP-tagged. GFP-fusion constructions were introduced into two types of cells, the wild type (WT) and a mutant (BP) deleted for the Sfi-like centrin-binding protein and devoid of ICL. BP cells retain only a remnant of the infraciliary lattice closely associated to each basal body or basal body couplet, called ICL-organising centre (ICLOC) as it nucleates the assembly of the ICL under rescue conditions.

Localisation of the ICL1e centrins

The ICL1e subfamily comprises seven genes coding for centrins corresponding to three ohnolog pairs, i.e. gene duplicates originating from whole genome duplication (ICL1e and ICL1g; Cen8 and Cen10; Cen12 and Cen15) and a single gene, Cen18. In order to ascertain their respective localisation and function, a representative from each ohnolog group (ICL1e, Cen10, Cen15, Cen18) was GFP-tagged. GFP-fusion constructions were introduced into two types of cells, the wild type (WT) and a mutant (BP) deleted for the Sfi-like centrin-binding protein and devoid of ICL. BP cells retain only a remnant of the infraciliary lattice closely associated to each basal body or basal body couplet, called ICL-organising centre (ICLOC) as it nucleates the assembly of the ICL under rescue conditions.
ICL centrins are labelled by the monoclonal 1A9 which recognises the most acidic of the ICL centrins (Beisson et al., 2001). In WT cells (Fig. 2A), both GFP-ICL1e centrins and 1A9 label the whole network, but the two labels do not exactly overlap: the GFP-ICL1e labelling appears as discrete dots localised along a linear structure decorated by 1A9. In addition, all GFP-tagged members of the ICL1e subfamily localise as a dot near each single or paired basal body not recognised by the monoclonal 1A9. These results demonstrate that the 1A9 antibody does not recognise the ICL1e centrins. In BP cells, the ICLOC stands out as a bipartite structure, with a GFP-ICL1e centrin-labelled dot close to the basal body and not recognised by the monoclonal 1A9, and a small cluster of patches of less-defined shape, labelled by both 1A9 and the GFP-ICL1e centrins (Fig. 2B). The ICL1e dot, present in both WT and BP cells, thus appears as a constitutive basal body-associated structure. Because all the seven ICL1e centrins colocalise at this dot, we will call it the ICL1e complex.

To analyse the role of this ICL1e complex, we wanted to (1) ascertain the relative timing of duplication of basal body and ICL1e complex, (2) analyse the effect of ICL1e centrins depletion on basal body and ICL biogenesis and (3) investigate the possible functional relationships between the ICL1e complex and the basal body centrins, Centrin2 and Centrin3.

Timing of duplication of the ICL1e complex, the basal bodies and the ICL

All over the cell surface, basal bodies, organised as antero-posterior rows, are inserted in a submembranar cytoskeleton, the epiplam, segmented in longitudinal files of ‘cortical units’, within which one or two basal bodies are anchored (Iftode et al., 1989; Aubusson-Fleury et al., 2013). During division, the assembly of new basal bodies generates an alternance of parental and new basal bodies along each row. At the cell level, this duplication follows a precise spatio-temporal pattern which exactly restores the distribution of basal bodies in the two daughter cells (Fig. S2-1).

In dividing BP cells, the GFP-ICL1e complex duplicates first, while the 1A9-reactive spots will re-assemble later (Fig. 3A). This can also be observed in WT cells at the onset of division (Fig. S2-2). This stepwise assembly shows that the duplication of the ICLOC over cell divisions in BP cells, as well as of the ICL in WT cells, follows the duplication of the ICL1e complex.

We then compared the time of appearance of new ICL1e dots to that of SAS-6, the earliest known marker of nascent basal bodies in Paramecium (Jerka-Dziadosz et al., 2012). In cells expressing both RFP-SAS-6 and GFP-ICL1e, the SAS-6 and ICL1e foci co-localise at all new single basal bodies (Fig. 3B): the ICL1e complex that accompanies the new basal bodies as early as SAS-6 is detectable.
Therefore, the ICL1e complex duplicates along with the basal bodies before the assembly of the other components, to reconstitute the interphase ICL network in WT cells or the ICLOC in BP cells.

**Effects of ICL1e depletion on the cortical organisation of the basal bodies**

In order to analyse more precisely the role of the ICL1e complex, we depleted WT and BP cells by RNAi of representatives of each ohnolog group. Possible off-target effects were systematically screened for each construction (Fig. S3-2).

RNAi knockdown of ICL1e centrins in WT cells first causes a stripping of the ICL network (Fig. 4A), then scattered disorders in the organisation of basal body rows, and cell growth arrest after a few divisions. In BP cells, the earliest effect, soon after the first division under RNAi conditions, is the absence of ICLOC on newly assembled basal bodies (Fig. 4B) then followed by disorganisation of basal body rows after two or three divisions. In both BP and WT cells, patches of supernumerary mislocalised basal bodies appear. These erratic basal bodies are not anchored to the cell surface and may be found intra-cytoplasmically (Fig. 4C). At the ultrastructural level, the unattached basal bodies present a transition zone (Fig. 4D,E). Depletion of ICL1e thus does not impair duplication per se (since new basal bodies appear fully developed), but affects the insertion of newly formed basal bodies at the cell surface, a syndrome similar to the effect of inactivation of Centrin3 (Ruiz et al., 2005; Aubusson-Fleury et al., 2012).

The silencing of each member of the ICL1e subfamily yielded the same defective phenotype as induced by ICL1e centrin inactivation (Gogendeau et al., 2008). All the members of the ICL1e subfamily are essential to trigger the ICLOC/ICL assembly and the positioning of the new basal bodies. They do not complement each other and thus form a single functional complex.

**Topological relationships between ICL1e and the basal body Centrin2 and Centrin3, during division**

ICL1e, like Centrin2 and Centrin3 (Ruiz et al., 2005), are required for proper basal body positioning/anchoring in *Paramecium*. To ascertain the chronology of appearance of the three types of centrins, we compared it to that of the detection of basal body labelling at three successive stages of duplication of the cortical units (Fig. S2-1): at the onset of duplication just before detection of the new basal bodies within the cortical units, after duplication in newly individualized units with single basal bodies, and at the end of duplication after the formation of units with single and/or paired basal bodies.

In interphase cells, Centrin2 localises within the basal body cylinder, while Centrin3 localises anterior to single basal bodies, and between the two basal bodies when paired (Ruiz et al., 2005). At that stage, the localisations of ICL1e and Centrin3 are similar.

At the onset of unit duplication, before basal body duplication, a new spot of Centrin3 appears anterior to the anterior basal body of pairs (Fig. 5). Thus, at this stage, an anterior Centrin3 spot is associated to all paired and single basal bodies. Interestingly, the ICL1e labelling exactly follows the same pattern as Centrin3 (Fig. 5A).

A new spot of Centrin2 then appears just anterior to parental basal bodies (not shown). The new basal body is detected soon after by labelling with the ID5 antibody. The successive detection of a GFP-Centrin2 and of an ID5 spot suggests that Centrin2 is an early marker of basal body assembly, as already demonstrated for the centriole (Azimzadeh et al., 2009). As soon as the new basal body is detected, it is associated with an anterior spot of Centrin3 and ICL1e (Fig. 5B).

At the end of cell division, in some units, basal bodies duplicate to yield pairs: a GFP-Centrin2 spot first appears, then followed by the detection of a new basal body. Neither new GFP-Centrin3 nor GFP-ICL1e spots are detected in association with this new anterior basal body (Fig. 5C).

These observations show that, while Centrin2, localised inside the basal body, is an early marker of basal body assembly, the recruitment of Centrin3, localised outside the basal body (Ruiz et al., 2005) and of ICL1e at a new anterior paired basal body may be delayed until the next duplication: the full assembly and anchoring of a basal body do not need its association with Centrin3 and the ICL1e complex.

**Interactions between ICL1e and Centrin2 and Centrin3**

RNAi knockdown of either Centrin2 or Centrin3 disturbs basal body attachment to the cell surface, yielding intra-cytoplasmic basal bodies before the assembly of the other components, to reconstitute
bodies (Ruiz et al., 2005). Double-labelling of ICL1e and basal bodies shows that, upon inactivation of Centrin2, no GFP-ICL1e is found in association with intra-cytoplasmic basal bodies (Fig. 6A). Thus, although the two markers do not colocalise, Centrin2 is required for ICL1e localisation. In contrast, upon inactivation of Centrin3, GFP-ICL1e still localises at intra-cytoplasmic basal bodies (Fig. 6B). In the reverse situation, i.e. inactivation of ICL1e on GFP-Centrin3 expressing cells (Fig. 7), the localisation of GFP-Centrin3 is modified: the intensity of the Centrin3 spots is weaker on new basal bodies assembled after the onset of gene inactivation (Fig. 7A-D). Therefore, ICL1e is required for Centrin3 localisation.

To analyse further this decrease in GFP-Centrin3 fluorescence, a STED investigation was carried out to localise the GFP-Centrin3 spots more precisely. It appears that GFP-Centrin3 has a triple localisation: one spot sits at the tip of each basal body (Fig. 7B inset) and two spots at its base (Fig. 7E,G). These labels fit well previous ultrastructural observations, where the double proximal (Ruiz et al., 2005) and the tip localisations (Aubusson-Fleury et al., 2012) were observed. Upon ICL1e knockdown, at an early stage, the sole right basal spot (Fig. 7F,H) disappears, while the tip and basal left spots will disappear during the next duplications. The fact that all Centrin3 spots do not disappear concomitantly suggests that there

---

**Fig. 5. Topological relationships between Centrin2, Centrin3 and ICL1e during basal body duplication.** Confocal images. Three successive stages of basal body duplication are illustrated in cells expressing GFP-Centrin2 (A1-C1), GFP-Centrin3 (A2-C2) or GFP-ICL1e (A3-C3) and labelled with the ID5 antibody. (A) Pre-duplication, (B) duplication 1 leading to one basal body per cortical units, and (C) duplication 2, resulting in paired basal bodies in some units. The scheme to the right summarises the specific behaviour of each of the three centrins at successive stages. Circles: basal bodies; colour intensities represent successive basal bodies generations (0, parental basal bodies; 1, basal bodies from duplication 1; 2, basal bodies from duplication 2); blue, Centrin2; green, Centrin3; yellow, ICL1e. While the Centrin2 labelling (A and blue) is superimposed to the basal bodies at all three stages, Centrin3 (2 and green) and ICL1e (3 and yellow) all localise anterior of all basal bodies during the pre-duplication stage (A), and anterior of the single and of the sole posterior basal body in paired basal bodies at other stages (B and C). Each image represents a 10×10 µm area.
highly variable, up to 22 in the Ciliate genes. In other species, the number of centrin genes identified is single centrin in *C. rheinardii* was detected in sequenced genomes (Table S1). As expected, only one centrin gene cdc31families were discovered, we exclusively screened completely 770 indication that the story of centrins in the course of eukaryote even 24 in the Parabasalia to Gogendeau et al., 2008) and integrates genomes from all major (Ruiz et al., 2005; Azimzadeh and Bornens, 2007; Shi et al., 2008; The present analysis (Fig. 8) updates previous phylogenetic analyses Centrin diversification during eukaryotic evolution shows branches leading to individual genes of extremely variable some branches and gene losses in others. diversification has been complex, with multiple duplication events in several pathways for Centrin3 recruitment at the basal body. The ICL1e centrin is required for the assembly of the right Centrin3 spot, while it is not for the left one, which identifies the site of assembly of the anterior left filament (ALF), a Centrin3-dependent transient appendage which guides the positioning of the daughter basal body (Jerka-Dziadosz et al., 2012). Centrin2 and Centrin3 are sequentially required for the maturation and positioning of the basal bodies (Ruiz et al., 2005): the present observations add one actor along the functional sequence, Cen2-ICL1e-Cen3. This interdependence between the three basal body centrins suggests the existence of physical connections, likely ensured by centrin-binding proteins of the Sfi1 type, homologous to the Sfi13 described in *T. thermophila* and shown to affect basal body duplication and anchoring (Stemm-Wolf et al., 2013; Heydeck, et al., 2016).

In view of this essential function of the ICL1e centrins in basal body anchoring and organisation of the ciliary rows and of their intricate role with the other ubiquitous centrins, it seemed of interest to ascertain their phylogenetic position in relation to the other centrin families. **Centrin diversification during eukaryotic evolution**

The present analysis (Fig. 8) updates previous phylogenetic analyses (Ruiz et al., 2005; Azimzadeh and Bornens, 2007; Shi et al., 2008; Gogendeau et al., 2008) and integrates genomes from all major eukaryotic groups, as recently described (Adl et al., 2012). In addition to *Chlamydomonas rheinardii* and *Saccharomyces cerevisiae* in which the founding members of the Centrin2/xfl2 and Centrin3/ cdc31 families were discovered, we exclusively screened completely sequenced genomes (Table S1). As expected, only one centrin gene was detected in *S. cerevisiae*, but contrary to the prevalent idea of a single centrin in *C. rheinardii*, we found two additional centrin genes. In other species, the number of centrin genes identified is highly variable, up to 22 in the Ciliate *Paramecium caudatum* and even 24 in the Parabasalia *Trichomonas vaginalis*. This is a first indication that the story of centrins in the course of eukaryote
Fig. 8. Maximum-likelihood tree of centrins. The colours of the gene labels indicate the lineage of origin (red, SAR; brown, Excavates; green, Archaeplasts; blue, Amoebozoans; purple, Opisthokonts). The ICL subfamilies of Paramecium are noted in blue numbers. ND, not determined. Species abbreviations: Ath, Arabidopsis thaliana; Bna, Bigelowiella natans; Cpa, Cyanophora paradoxa; Cre, Chlamydomonas reinhardtii; Ddi, Dictyostelium discoideum; Ehu, Emiliana huxleyi; Fli, Reticyclomyxa filosa; Gin, Giardia intestinalis; Gpr, Gonapodya prolifera; Gsu, Gardinia sulphuraria; Gth, Guillardia theta; Hsa, Homo sapiens; Mbr, Monosiga brevicollis; Mve, Marsilea vestita; Ngr, Naegleria gruberi; Otr, Oxytricha trifallax; Pca, Paramecium caudatum; Pdu, Platynereis dumerii; Pfa, Plasmodium falciparum 3D7; Ppu, Porphyridium purpureum; Sce, Saccharomyces cerevisiae; Ska, Symbiodinium kawagutii; Smo, Selaginella moellendorffii; Tbr, Trypanosoma brucei; Tps, Thalassiosira pseudonana; Tth, Tetrahymena thermophila; Tva, Trichomonas vaginalis. Stippled black boxes indicate well supported phylogenetic groups (red and orange nodes): Centrin3, Ciliate ciliary centrins, and ICL centrins corresponding to the P. tetraurelia ICL1, ICL3/10/11 and ICL1e centrins. Potential extension of these well supported groups with reduced tree support (blue nodes) are indicated in colours (blue for ICL/ICL-like and green for ciliary centrins). The group of Centrin2 sequences is only weakly supported and some sequences jump out of this group. This is for example the case in Arabidopsis in which the two centrin genes, not considered as Centrin2 by some authors (Hodges et al., 2011), nevertheless display some of its functional properties (Molinier et al., 2004; Nishi et al., 2013).
length, not only across lineages but also in the same species. This is an indication that evolutionary constraints on primary structures have been highly variable between different gene types, maybe due to the number and flexibility of the molecular contacts that centrins make in the cell.

To ascertain the robustness of groupings, we compared the complete taxa sampling tree (Fig. S4-1) with a smaller tree, comprising only those genes that have evolved more slowly (see Materials and Methods section). The tree (Fig. 8) supports unequivocally only a limited number of groupings containing genes from several lineages, thus likely to represent ancestral eukaryotic genes. The monophyletic group comprising all Centrin3/cdc31 genes is particularly obvious, one of the two additional *Chlamydomonas* centrin genes unambiguously clustered in this group, confirming our previous analysis (Ruiz et al., 2005). A group containing short branch Centrin2/VFL2 sequences from a number of lineages is only weakly supported and some genes, interpreted as orthologous to Centrin2 for functional reasons, do not cluster with them. This may be interpreted as follows: an ancestral Centrin2 gene (with a high constraint on sequence evolution) is a certainty, owing to the diversity of lineages found in the weakly supported group. However, many other derived centrins (the long-branch genes) may have originated by duplication from this ancestral gene, thus making the evolutionarily conserved Centrin2 a paraphyletic stem group. The derived centrins do not cluster with Centrin2 in this tree because of their exuberantly high evolutionary speed. They instead branch deeper or group together due to long-branch attraction.

In favour of this scenario no other well-supported monophyletic group, with centrins coming from multiple lineages, appears in this tree. All other groupings mostly comprise long-branched genes. This is for example the case for one additional node which regroups several ciliate genes, including *Paramaecium* ciliary centrin genes, and a long branched cercozoan gene. In some cases, such as the *Trichomonas* centrins, all sequences of the node come from the same species, suggesting multiple gene duplications as already demonstrated (Carlton et al., 2007) and a rapid divergence of a paralogous gene sub-family. In other cases, when several eukaryotic lineages are represented, the statistical support of the node is low and the grouping dubious.

### Origin of the ICL/ICL-like centrin genes

Does our tree suggest an evolutionary origin of the ICL genes? Most, but not all, ICL sequences appear as long-branch genes. This indicates that they evolved rapidly with relaxed constraints as compared to the highly constrained Centrin2 and Centrin3 proteins. Retracing their evolutionary history is therefore difficult. The ICL sequences do not form a monophyletic group and may have originated from independent events of gene duplications from an ancestral centrin.

However, an important grouping, statistically supported and robust to resampling, emerges (ICL in Fig. 8). This branch groups not only many ciliate genes but also some genes from other species of the Stramenopile/Alveolates/Rhizaria (SAR) lineage (Adl et al., 2012). By contrast, a lone green algal sequence (*Chlamydomonas 184498*) jumps out of this group in the short branch tree, casting doubt on its affinities with ICL genes; but, even if no other plant sequence matches this additional branch, a sequence highly similar to this *Chlamydomonas* sequence is found in Volvox (Prochnik et al., 2010), another Chlorophyte able to divide in its flagellated stage, suggesting that this property could put a constraint to retain ICL-related genes.

This ICL grouping may in fact be even larger as a broader group emerges. In addition to the *Chlamydomonas* and Dynophyta sequences, this larger ICL-like group contains genes coming from another SAR lineage, as well as from potentially SAR-related lineages, Cryptophytes and Haptophytes. Next to this group is another branch with again a formally identified ICL gene from *Paramaecium* (ICL9), suggesting that this branch could also correspond to ICL/ICL-like genes. Although the group does not show a strong statistical significance, it is striking to find in this branch one sequence of another Ciliata (*Tetrahymena*). These groupings indicate that at least one ICL gene may have appeared by duplication of an ancestral centrin gene (either Centrin2 or Centrin3) in an ancestor of the SAR lineage or an even larger lineage. Finally, the occurrence of sequences from organisms of other lineages (Cryptophyta, Dictyostelias, Apicomplexa), and more unexpectedly from one Metazoan (*Platyneireis*) and the Choanozoan (*Monosiga*), show that some divergent centrins also exist in these organisms.

Within the ICL-like groups, few groupings emerge robustly. One is the ICL1e subfamily of ciliates. Interestingly, these particular genes have relatively short branches, indicating strong evolutionary constraints on their primary structures. Two similar short-branch genes from the dinoflagellate *Sympodium* and from the apicomplexan *Plasmodium* cluster with these ciliate centrins, although with a low significance. This suggests that a distinct ICL1e gene may have already existed in the ancestors of Alveolates. Besides ICL1e, two other well supported groupings are recognised within the ICL, one comprising only ciliate sequences, the other one comprising ciliate sequences and one foraminiferan ICL sequence, suggesting further that some diversification by gene duplication of the ICL had already taken place in the ancestor of the SAR lineage.

### DISCUSSION

In *P. tetraurelia*, the functional and molecular dissection of the ICL had led to the characterisation of a large number of centrins, phylogenetically distinct from the major Centrin2 and Centrin3 families (Gogendeau et al., 2008). We demonstrate here that the ICL1e centrins, one subfamily of the ICL centrins, contribute to the correct insertion and ordering of the basal bodies at the cell surface through an interaction with the ubiquitous centrins, Centrin2/VFL2 and Centrin3/cdc31. In addition, these proteins, which form a complex associated with the basal bodies, constitute the organising centre of the ICL.

In view of this essential role in basal body duplication and in maintenance of ciliary planar polarity, we revisited their evolutionary status. The existence of three ancient families previously identified (Ruiz et al., 2005; Azimzadeh and Bornens, 2007) is confirmed with a broader sampling: the centrins 1, 2 and 3 are unambiguously ancestral eukaryotic proteins. ICL centrins have a more recent origin but are still very ancient, as they are shared by the whole SAR lineage and even possibly a broader clade of eukaryotes. Among ICL centrins, the ICL1e sub-family emerges as particularly ancient, conserved and constrained.

### The essential functions of the ICL1e centrins

We show here that a basal body-associated complex, the ICL1e complex duplicates along with the basal bodies and that the sequential recruitment of the Centrin2, ICL1e and Centrin3 is required for correct duplication and positioning of the new basal body. The effects of ICL1e deficiency resemble those resulting of Centrin3 deficiency (Ruiz et al., 2005), as in both cases, the unanchored basal bodies appear well developed.

The genus *Paramaecium* has undergone several whole genome duplications (WGD) in the course of evolution (Aury et al., 2006),
and the comparison between different species allows us to trace the history of the ICL1e subfamily. In *P. caudatum*, which diverged from *P. tetraurelia* before the last two WGDs, only two ICL1e centrins were identified. Although not functionally studied, they can be assumed to have the same role as in *P. tetraurelia*. One ICL1e gene was also detected in two other Ciliates, *Tetrahymena* and *Oxytricha*. In *Tetrahymena*, this gene corresponds to the *Ttcen4* gene encoding a centrin which also localises at basal bodies (Stemm-Wolf et al., 2005). No functional analysis is available in this genus, but the similar localisation of this ICL1e suggests the same function in the two Ciliates. No functional data are available for the ICL1e in *Oxytricha*, but data obtained in another related species (Lemullois et al., 2004) suggest similar functions for centrins in this lineage.

Surprisingly, orthologs of ICL1e are also present in the two other groups of Alveolata, but not in any other group. The common cytological particularity of Alveolata is the presence of alveoles, which form a layer beneath the plasma membrane (Stelly et al., 1991). Basal bodies and extrusomes are anchored to the plasma membrane between two adjacent alveoles. Thus, to anchor, the new basal body has to be positioned very precisely not only with respect to its parent, but also to the superficial surrounding architecture. Our observations strongly suggest that, in addition to Centrin3 required to build the basal body (Vigue et al., 1984), as well as pro-basal bodies (Furness and Butler, 1985), are inserted. Although in such species, the function of this network has not yet been ascertained, its development, organisation and behaviour suggest that it would be required not only for contractility, but also for the transmission of the basal body pattern during the cell cycle. The fact that such a complexity of centrins is retained in *Paramecium* for a network dispensable for main cell functions could correspond to remnants of its evolutionary history. Alternatively, the network could be required under specific environmental conditions.

**Centrin diversification and contractility**

It is quite surprising that a large part of the centrin diversity in *P. tetraurelia* is represented by ICL centrins, required for the assembly of the infraciliary lattice. In *P. tetraurelia*, this diversity represents up to 38/50 genes, and the expression of all is required for the assembly of this network. The fact that this network seems to be dispensable for cell survival, at least under laboratory culture conditions, raises the question of how and why such a diversity has been retained in the course of evolution.

In *P. caudatum*, the ICL centrins represent 16/21 centrin genes, corresponding to the 10 subfamilies defined in *P. tetraurelia*. In our broad centrin phylogeny, two genes, each representing one subfamily (ICL8 and ICL9), exhibit long branches suggesting a rapid evolution. In addition, four subfamilies (ICL3a, ICL3b, ICL10 and ICL11) appear closely related, suggesting a more recent divergence. Interestingly, in *Tetrahymena thermophila*, which displays a general organisation similar to that of *Paramecium* but a centrin network reduced to an apical band (Jerka-Dziadosz et al., 2012), ICL1e is necessary to recruit Centrin3 in the proximity of the striated fibre where the pro-basal body is assembled during duplication (Ifode and Fleury-Aubusson, 2003). One hypothesis is that this additional system may generate an additional fine tuning during basal body positioning.

**Which centrins for which organisms?**

Based upon an extensive sampling of eukaryotic lineages, the present phylogenetic analysis confirms the diversity of centrins beyond the most conserved families, Centrin2/Vfl2 and Centrin3/cdc31, currently recognised in eukaryotes (Azimzadeh and Bornens, 2007). It leads to the following general observations:

1. Only two centrin genes, Centrin2 and Centrin3, were undoubtedly already present in the last common ancestor as suggested by previous phylogenies (Azimzadeh and Bornens, 2007). Our phylogeny confirms that Centrin2, considered as a core component of the centriole (Hodges et al., 2010), is maintained in organisms with basal bodies and lost or present but highly divergent in the other ones. In the few groups which have lost Centrin3, at least another divergent centrin is found in addition to the canonical Centrin2, suggesting that representatives of several centrin families are necessary to ensure the continuity of basal bodies (Fig. S4-2).

2. Additional centrins have appeared in other lineages. In some lineages, one or few additional centrins are detected but in two lineages (Ciliata and Parabasalia), whole new families comprising many genes have evolved. The tree presented here suggests that either Centrin2 or Centrin3 could have duplicated.

Among these divergent centrins, the ICL proteins are particularly remarkable. A large proportion of ciliate ICL centrins emerges as a single group comprising molecules found in the SAR lineage (or even from a larger SAR/Chromalveolate grouping). The tree thus suggests that at least one ICL-like molecule appeared early in eukaryotic evolution. The particular conservation and functional specialisation of one particular sub-type, ICL1e, suggests a scenario for the functional diversification of these ICL-like molecules. We show that ICL1e centrins play a crucial role in basal body anchoring in ciliates. The occurrence of closely related ICL1e-like molecules in other SAR groups indicates that ICL1e evolved by gene duplication in the SAR ancestor and could have retained its ancestral function in basal body maintenance. Later on, additional duplications of this ancestral ICL type took place but the new paralogs diverged further, evolved more rapidly and were recruited for building evolutionary novelties, like ciliate networks of ciliates. Because of its conserved function at the basal body, the ICL1e centrin has remained more constrained than the other paralogous ICL proteins. This scenario remains tentative as the function of potential ICL1e orthologs in groups other than ciliates remains to be explored.

Finally, besides the SAR group, other organisms display divergent centrins which could have other functions. In *Trypanosoma brucei*, the phylogenetically divergent centrin detected here has been shown to interact with one basal body centrin (Wang et al., 2012), like ICL1e in *Paramecium*, and to coordinate cell and nuclear division (Shi et al., 2008). In *Chlamydomonas*, where we detected one ICL-divergent gene in
addition to Centrin2 and Centrin3, centrin is also an important component of the basal body-nucleus connector (Geimer and Melkonian, 2005; Dutcher and O’Toole, 2016). Finally, and unexpectedly, one divergent centrin is also found in two Holozoa, Monosiga and Platynereis. It is not surprising that Monosiga displays one additional centrin: in most groups where cells are able to divide in their ciliated/flagellated stage, at least three different centrins are detected. The presence of a divergent centrin in a metazoan, Platynereis is more surprising. Indeed, the ability of ciliated cells to divide has not yet been detected in this lineage, where multiciliated cells correspond to a terminal differentiation (Brooks and Wallingford, 2014). To what extent the presence of a divergent centrin in Platynereis could be related to the existence of specific pathways of ciliation of its larval stage remains to be explored.

MATERIAL AND METHODS
Strains and culture procedures

Three strains were used: the wild-type reference strain d4-2 of Paramecium tetraurelia, the mutant nd7-1 (Skouri and Cohen, 1997) and the strain Δ-CenBP1/nd7-1 (Sehring et al., 2009). Cells were grown at 27°C in a wheat grass infusion (BHB, L’arbé de vie, Luçay Le Male, France) bacterised with Klebsiella pneumoniae and supplemented with 0.4 µg/ml β-sitosterol (Sonneborn, 1970; Beisson et al., 2010).

Gene cloning

Genes coding for centrin were amplified by PCR (using the oligonucleotides listed in Fig. S3-1) and introduced in appropriate restriction sites of the silencing plasmid or Paramecium expression vectors. Accession numbers: PtICL1: GSPATG00007681001; PtICL1g: GSPATG00021508001; PtCen8: GSPATG00000885001; PtCen10: GSPATG00000338001; PtCen12: GSPATG00000011001; PtCen15: GSPATG00001075001; PtCen18: GS-PATG00038985001. Amplifications were performed with high fidelity Phusion DNA polymerase (Thermo Scientific) according to the manufacture’s protocol.

GFP constructs

GFP coding sequence was 5’-inserted in each centrin gene and cloned in the KpnI site of the expression vector pPXV-GFP (Gogendeau et al., 2005). GFP-fusion constructs were placed under the control of the Paramecium calmodulin regulatory sequences with 4-6 nucleotides coding for glycine amino acid introduced between the GFP coding sequence and the centrin genes.

Obtaining transformed cell lines

ND7 cells were transformed by microinjection (Beisson et al., 2010). For each gene, at least 10 transformed clones were analysed.

RNAI by feeding

T7Pol-driven dsRNA were produced from at least 75% of the genomic sequence (Fig. S3-2) from each centrin gene amplified by PCR and cloned into the L4440 feeding vector (Timmons and Fire, 1998). Feeding experiments were performed as described (Beisson et al., 2010). Possible RNAI-off target effects were searched for each construct. As off-target sequences were detected using tools available at ParameciumDB (Arnatz et al., 2007) (Fig. S3-2), controls of feeding efficiency and/or specificity were performed by analysis of GFP-fluorescence decrease after feeding of clones expressing the GFP-ICL1e centrin with constructs used for inactivation of ICL1e (Gogendeau et al., 2008), PtCen8/PtCen15, PtCen12/PtCen10, and Cen18: significative GFP decrease was observed only after ICL1e inactivation (not shown), thus assessing the specificity of the RNAI. At least five independent experiments were carried out for each gene analysed.

Immunostaining

Cells were immunostained as previously described (Beisson et al., 2010). The antibodies were used at the following dilutions: anti-ICL centrin 1A9 culture supernatant (Beisson et al., 2001) 1:200, monoclonal anti-tubulin antibody culture supernatant 1D5 (Wehland and Weber, 1987) 1:100, the polyclonal PolyE anti-tubulin serum (Janke and Bulinski, 2011) 1:500, anti-GFP antibody from Interchim (Montluçon, France) 1:1000 and secondary antibodies labelled with AF488 or AF568 (Invitrogen-Molecular Probes, Eugene, OR, USA), or Cy5 (Life Technologies) diluted 1:200-1:500. Each experiment was done on samples of 50-200 cells.

Electron microscopy

Cells were prepared as previously described for morphological observations (Aubusson-Fleury et al., 2012). Ultrathin sections were examined with a Jeol transmission electron microscope at 120 kVolt.

Phylogenetic analysis

Screenings for centrin-like sequences was conducted on databanks of predicted peptides from whole genome sequencing (Table S1). In a few cases, this genome screen was complemented by an assembled transcriptome screen. Blast searches were conducted either online or with NgKlast v4.0 (Korilog, France). The sequence probe used for all blast searches was a concatenation of human centrin 1, 2 and 3. All hits above an E value threshold set at 1E-10 were reciprocally blasted against the human proteome and all those whose first human hit was a centrin, even with a low E value, were retained for further phylogenetic analyses. Five calmodulin-like sequences from different eukaryotic species were added as outgroups. A total of 149 sequences were thus retained. Sequences were aligned with the Muscle algorithm of Seaview (Gouy et al., 2010) and only the unambiguously aligned parts of the molecules were used for tree making. Alignments are available upon request to G.B. Tree making was performed with the Maximum Likelihood algorithm, as implemented in Seaview, using the LG model of amino acid substitution, no invariable sites, six categories of rate variations and aLRT scores for node robustness.

The multiplicity of the sequences, the relative shortness of the protein (only 158 as positions were retained for tree making) and the large variations in evolutionary rates result in a tree in which many nodes, especially deep ones, have a relatively low statistical value, as indicated by aLRT scores (Anisimova and Gascuel, 2006). To evaluate further node robustness, a selection of fast-evolving sequences was made. To this end, a table of Poisson distances between sequence pairs was calculated using the Distance Methods option of Seaview and the average distance of each centrin-like molecule to the calmodulin outgroup sequences (Table S2). We set arbitrarily a threshold at a value of 1.1 above which sequences were considered fast-evolving and prone to artefactual. This threshold resulted in the elimination of 43 sequences (29%) and another tree was built with the resulting set and the same options as above.

Acknowledgements

We are very grateful to Anne-Marie Tassin for support and interest and to Dr. Ulf Schwartz (Leica, Mannheim) for making the STED imaging and sharing his deep expertise in STED experimentation. We thank Carsten Janke (Curie Institute, Orsay, France) for the gift of PolyE serum.
Furness, N. D. and Butler, R. D. (1985). The cytology of the sheep rumen Ciliates. III. Ultrastructure of the genus Entodinium (Stein). J. Protozool. 32, 699-707.

Garreau de Loubresse, N., Keryer, G., Viguès, B. and Beisson, J. (1988). A contractile ciliateskeletal network of Paramecium: the infralittoral lacy. J. Cell Sci. 90, 329-336.

Garreau de Loubresse, N., Klotz, C., Viguès, B., Rutin, B. and Beisson, J. (1991). Ca2+-binding proteins and contractility of the infralittoral lacy in Paramecium. Biol. Cell 71, 217-225.

Geimer, S. and Melkonian, M. (2005). Centrin scaffold in Chlamydomonas reinhardtii revealed by immunoelectron microscopy. Eur. J. Cell Biol. 84, 1253-1263.

Gogendeau, D., Keller, A.-M., Yanagi, A., Cohen, J. and Koll, F. (2005). Nd6p, a novel protein with RCC1-like domains involved in exocytosis in Paramecium tetraurelia. Eur. J. Cell Biol. 84, 2129-2139.

Gogendeau, D., Beisson, J., Garreau de Loubresse, N., Ca:Jer, P.-R., Ruiz, F., Conde, J. and Sperling, L. Klotz, C. (2007). A Sflp-like centrin-binding protein mediates centric based Ca2+-dependent contractility in Paramecium. Eur. J. Cell Biol. 86, 1992-2000.

Gogendeau, D., Klotz, C., Arnaiz, O., Malinowska, A., Dadlez, M., Garreau de Loubresse, N., Ruiz, F., Koll, F. and Beisson, J. (2008). Functional diversification of centriols and cell morphology complexity. J. Cell Sci. 121, 65-74.

Gouda, K., Yoshida, A., Oomi, K. and Takahashi, M. (2004). Centrin is essential for the activity of the ciliary reversal-coupled voltage-gated Ca2+ channels. Biochem. Biophys. Res. Commun. 323, 891-897.

Gouy, M., Guidouin, S. and Gascuel, O. (2010). Sea-view version 4: a multiplatform graphical user interface for sequence alignments and phylogenetics tree building. Mol. Biol. Evol. 27, 221-224.

Heydecker, W., Stern-Woll, A. J., Knoep, J., Poh, C. C. and Winey, M. (2016). Sfr1, a Tetrahymena thermophila Sflp-repeat protein, mediates the production of cortical hair basal bodies. mSphere 10, e00275-00216.

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.

Hodges, M. E., Wickstead, B., Gull, K. and Langdale, J. A. (2011). Conservation of ciliary proteins in plants with no cilia. BMC Plant Biol. 11, 185-193.

Iftode, F. and Fleury-Aubusson, A. (2003). Structural inheritance in Paramecium: ultrastructural evidence for basal body and associated rootlets polarity transmission through binary fission. Biol. Cell 95, 39-51.

Iftode, F., Cohen, J., Ruiz, F., Rueda, A. T., Chen Shan, L., Adoutte, A. and Beisson, J. (1989). Development of surface pattern during division in Paramecium. I. Mapping of duplication and reorganization of cortical cytoskeletal structures in the wild type. Development 105, 191-211.

Janke, C. and Bulinski, J. C. (2011). Post-translational regulation of the microtubule ciliopathies: mechanisms and functions. Mol. Cell Biol. 12, 773-786.

Janke, J. and Bornens, M. (2002). Origin and evolution of the centrosome. Adv. Exp. Med. Biol. 507, 119-129.

Janke, J., Herteg, P., Delouvée, A., Euteneuer, U., Formstecher, E., Khodjakov, A. and Bornens, M. (2009). Hpcos is a centrin-binding protein required for assembly of full-length centrioles. J. Cell Biol. 185, 101-114.

Baum, P., Furlong, C. and Byers, B. (1986). Yeast gene required for spindle pole body duplication: homology of its product with Ca2+-binding proteins. Proc. Natl. Acad. Sci. USA 83, 5512-5516.

Beisson, J., Claer, J., Fleury-Aubusson, A., Garreau de Loubresse, N., Ruiz, F. and Klotz, C. (2001). Basal body-associated nucleation center for the centrin-based cortical cytoskeletal network in Paramecium. Protot 152, 339-354.

Beisson, J., Bétermin, M., Bré, M. H., Cohen, J., Duharcourt, S., Duret, L., Kung, C., Malinsky, S., Meyer, E., Preer, J. R., Jr et al. (2010). Paramecium tetraurelia: the renaissance of an early unicellular model. Cold Spring Harb. Protot.

Brooks, E. and Wallingford, J. B. (2014). Multiciliated cells. Methods Cell Biol. 127, 131-159.

Carlton, J. M., Hirt, R. P., Silva, J. C., Delcher, A. L., Schatz, M., Zhao, Q., Worton, R. J., Bidwell, S. L., Alsmark, U. C., Besteiro, S. et al. (2007). Draft genome sequence of the sexually transmitted pathogen Trichomonas vaginalis. Science 315, 207-212.

Carvalho-Santos, Z., Azimzadeh, J., Pereira-Leal, J. B. and Bettencourt-Dias, M. (2011). Evolution: Tracing the origins of centrioles, cilia, and flagella. J. Cell Biol. 194, 165-173.

Dantas, T. J., Daji, O. M. and Morrison, G. C. (2012). Such small hands: the roles of centrin/caltractins in the centriole and in genome maintenance. Cell. Mot. Life Sci. 69, 2979-2997.

Dutcher, S. K. and O'Toole, E. T. (2016). The basal bodies of Chlamydomonas reinhardtii. Cilia 5, 18.

Fleury-Aubusson, A., Iftode, F. and Adoutte, A. (1992). Molecular phylogeny of Ciliates: what does it tell us about the Evolution of the Cytoskeleton and of developmental strategies. Dev. Genet. 13, 247-254.
Nagamune, K., Moreno, S. N., Chini, E. N. and Sibley, L. D. (2008). Calcium regulation and signaling in apicomplexan parasites. *Subcell. Biochem.* 47, 70-81.

Nishi, R., Sakai, W., Tone, D., Hanaoka, F. and Sugasawa, K. (2013). Structure-function analysis of the EF-hand protein centrin-2 for its intracellular localization and nucleotide excision repair. *Nucleic Acid. Res.* 41, 6917-6929.

Pollard, T. D. (2016). Actin and actin-binding filaments. *Cold Spring Harb. Perspect Biol.* 1, 8.

Prochnik, S. E., Umen, J., Nedelcu, A. M., Hallmann, A., Miller, S. M., Nishii, I., Ferris, P., Kuo, A., Mitros, T., Fritz-Laylin, L. K. et al. (2010). Genomic analysis of organismal complexity in the multicellular green alga *Volvox carteri*. *Science* 329, 223-226.

Ruiz, F., Garreau de Loubresse, N., Klotz, C., Beisson, J. and Koll, F. (2005). Centrin deficiency in *Paramecium* affects the geometry of basal-body duplication. *Curr. Biol.* 15, 2097-2106.

Salisbury, J. L., Baron, A., Surek, B. and Melkonian, M. (1984). Striated flagellar roots: isolation and partial characterization of a calcium-modulated contractile organelle. *J. Cell Biol.* 99, 962-970.

Sehring, I. M., Klotz, C., Beisson, J. and Plattner, H. (2009). Rapid downregulation of the Ca2+-signal after exocytosis stimulation in *Paramecium* cells: essential role of a centrin-rich filamentous cortical network, the infraciliary lattice. *Cell Calcium* 45, 89-97.

Shi, J., Franklin, J. B., Yelinek, J. T., Ebersberger, I., Warren, G. and He, C. Y. (2008). Centrin4 coordinates cell and nuclear division in *T. brucei*. *J. Cell Sci.* 121, 3062-3070.

Skouri, F. and Cohen, J. (1997). Genetic approach to regulated exocytosis using functional complementation in *Paramecium*: identification of the ND7 gene required for membrane fusion. *Mol. Biol. Cell* 8, 1063-1071.

Sonneborn, T. M. (1970). Methods in paramecium research. *Method. Cell Physiol.* 4, 241-339.

Stelly, N., Mauger, J. P., Claret, M. and Adoutte, A. (1991). Cortical alveoli of *Paramecium*: a vast submembranar calcium storage compartment. *J. Cell Biol.* 113, 103-112.

Stemm-Wolf, A. J., Morgan, G., Giddings, T. H., Jr, White, E. A., Marchion, R., McDonald, H. B. and Winey, M. (2005). Basal body duplication and maintenance require one member of the *Tetrahymena thermophila* centrin gene family. *Mol. Biol. Cell* 16, 3606-3619.

Stemm-Wolf, A. J., Meehl, J. B. and Winey, M. (2013). Sfr13, a member of a large family of asymmetrically localized Sfi1-repeat proteins, is important for basal body separation and stability in *Tetrahymena thermophila*. *J. Cell Sci.* 126, 1659-1671.

Timmons, L. and Fire, A. (1998). Specific interference by ingested dsRNA. *Nature* 395, 854.

Vigués, B., Méténier, G. and Grolière, C. A. (1984). Biochemical and immunological characterization of the microfibrillar ecto-endoplasmic boundary in the Ciliate *Isotricha prostoma*. *Biol. Cell* 51, 67-78.

Wang, M., Gheiratmand, L. and He, C. Y. (2012). An interplay between Centrin2 and Centrin4 on the bi-lobed structure in *Trypanosoma brucei*. *Mol. Microbiol.* 83, 1153-1161.

Wehland, J. and Weber, K. (1987). Turnover of the carboxy-terminal tyrosine of alpha-tubulin and means of reaching elevated levels of de tyrosination in living cells. *J. Cell Sci.* 88, 185-203.

Wei, Y., Hu, H., Lun, Z. R. and Li, Z. (2014). Centrin3 in trypanosomes maintains the stability of a flagellar inner-arm dynein for cell motility. *Nat. Commun.* 5, 4060.

Wright, R. L., Salisbury, J. and Jarvik, J. W. (1985). A nucleus-basal body connector in *Chlamydomonas reinhardtii* that may function in basal body localization or segregation. *J. Cell Biol.* 101, 1903-1912.