Time-resolved quantitative proteomics implicates the core snRNP protein, SmB, together with the Survival of Motor Neuron protein, in neural trafficking.

Alan R Prescott[2], Alexandra Bales[1], John James[2], Laura Trinkle-Mulcahy[3] and Judith E. Sleeman[1].

[1]School of Biology, University of St Andrews, BMS, North Haugh, St Andrews, Fife, KY16 9ST, UK. [2]College of Life Sciences, University of Dundee, DD1 5EH, UK. [3] Department of Cellular and Molecular Medicine and Ottawa Institute of Systems Biology, University of Ottawa, 451 Smyth Road, Ottawa, ON K1H 8M5, Canada.

Corresponding Author: Judith E Sleeman, jes14@st-andrews.ac.uk, Tel:+44 (0)1334 463524, Fax:+44 (0)1334 462595

Running Title: A novel role for SmB in neural trafficking

Word count: 7133

Summary

The biogenesis of splicing snRNPs (small nuclear ribonucleoproteins) is a complex process, beginning and ending in the nucleus of the cell but including key stages that take place in the cytoplasm. In particular, the SMN (Survival Motor Neurons) protein complex is required for addition of the core Sm proteins to the snRNP. Insufficiency of SMN results in the inherited neurodegenerative condition, Spinal Muscular Atrophy (SMA). Details of the physical organization of the cytoplasmic stages of snRNP biogenesis are unknown. We have used time-resolved quantitative proteomics to identify proteins that associate preferentially with either newly assembled or mature splicing snRNPs. These data have allowed us to identify highly mobile SmB protein trafficking vesicles in neural cells. These vesicles are dependent on the cellular levels of SMN and SmB for their morphology and mobility. We propose that these represent a family of related vesicles, some of which play a role in snRNP biogenesis and some of which may play more diverse roles in cellular RNA metabolism.

Key Words: SILAC protemics/snRNPs/Spinal Muscular Atrophy/Vesicles/SMN

Introduction

The five uridine-rich splicosomal snRNPs, U1, U2, U4, U5 and U6 are essential components of the spliceosome required for removal by splicing of introns from pre-messenger RNA. With the exception of U6 snRNP, splicing U snRNPs have a complex pathway of biogenesis involving cytoplasmic and nuclear stages (reviewed in Fischer et al, 2011). Following their transcription by RNA polymerase II in the nucleus, the small nuclear RNAs (snRNAs) are exported into the cytoplasm where early events in snRNP maturation occur including hypermethylation of the 5' cap of the snRNA and trimming of the 3' end. Additionally, a heptameric ring of core, Sm, proteins is assembled onto the Sm binding site of the snRNA. Assembly of the Sm core onto the snRNA requires the SMN (survival motor neurons) complex (reviewed in Battle et al, 2006). This process has been the subject of considerable attention as lowered levels of SMN result in the inherited neurodegenerative condition, Spinal Muscular Atrophy (SMA) (Lefebvre et al, 1995; Lefebvre et al, 1997). It is currently unclear whether the specific pathological defects in SMA, affecting predominantly spinal motor neurons, are caused by a decreased capacity for splicing snRNP maturation or by defects in other, as yet uncharacterized, roles for SMN in neural cells. Although numerous reports have been made of accumulations of Sm proteins and/or SMN in the cytoplasm, primarily in cells over-expressing SMN or mutant SMN (Liu & Gall, 2007; Pellizzoni et al, 1998; Sleeman et al, 2003) the cellular location and organization of the addition of Sm proteins to the snRNA during snRNP assembly is unknown. Furthermore, the identity and role of SMN-rich cytoplasmic accumulations has been a subject of debate. In Drosophila, U bodies, containing snRNAs and SMN (Liu & Gall, 2007) have been proposed to play a role in snRNP assembly, while SMN-rich structures in mammalian neural cells, reported in some cases to lack Sm proteins, have been implicated in the localization of mRNAs (Fallini et al, 2010; Fallini et al, 2011; Fan & Simard, 2002; Hubers et al, 2011; Jablonka et al, 2001; Peter et al, 2011; Sharma et al, 2005; Todd et al, 2010a; Todd et al, 2010b; Zhang et al, 2006), with a proposed role in the local translation of mRNAs required for motor neuron function. Following the assembly of the Sm protein core, the partially assembled snRNP is imported into the nucleus. Once imported, snRNPs accumulate transiently in the Cajal body (CB) where the core snRNA is modified. Further maturation of snRNPs involves the addition of a number of additional proteins specific to each class of snRNP (for example, U1 snRNP contains the proteins U1A, U170K and U1C). Again, it is not known exactly where or how these stages are carried out, but it is assumed that these proteins are added after the snRNP

has been re-imported into the nucleus. Mature snRNPs are believed to be stored in splicing speckles from which they are recruited to active spliceosomes when required (reviewed in Spector & Lamond, 2011).

In order to dissect out the molecular mechanisms and sub-cellular organization of different stages of splicing snRNP biogenesis, in particular cytoplasmic stages of potential relevance to SMA, we employed a time-resolved quantitative proteomic approach. Differential labeling of cells containing FP-tagged newly assembled snRNPs (transiently transfected with FP-SmB) versus FP-tagged mature snRNPs (lines stably expressing FP-SmB) was achieved by culturing cells in tissue culture media containing different isotopes of the essential amino acids arginine and lysine. This technique, known as Stable Isotope Labeling by Amino Acids in Culture (SILAC) (Ong et al, 2002), was then combined with affinity purification/ mass spectrometry to facilitate comparison of the interactomes of snRNPs at these two distinct stages. Investigations informed by these data revealed a family of novel cytoplasmic vesicles, enriched in SmB and SMN. Unexpectedly, these vesicles appear to be heterogeneous in their constituents other than SMN and SmB, suggesting that they may be involved in several aspects of cellular RNA metabolism. The SmB/SMN vesicles are upregulated during neural differentiation, disrupted by over-expression of SMN or reduced expression of SmB and are highly mobile in neurites, implicating them in the pathology of SMA.

RESULTS

1) Expression of the core snRNP protein SmB enables the preferential labeling of mature and newly made snRNPs in a time-dependent and concentration-independent manner.

Our previous analysis of the splicing snRNP maturation pathway using FP-tagged Sm proteins revealed that newly-expressed Sm proteins are first detected diffusely in the cytoplasm, before accumulating transiently in Cajal bodies prior to forming the steady-state localization of mature splicing snRNPs in splicing speckles (Sleeman & Lamond, 1999). Since the Sm proteins are assembled into snRNPs prior to reaching speckles (Sleeman et al 2001, Sleeman 2007, supplementary Fig. 1), this defines a pathway of maturation for splicing snRNPs. This pathway does not influence the localization of pre-existing snRNPs within the cell. Hela cells constitutively expressing YFP-SmB can be fused with cells constitutively expressing CFP-SmB, and the import into the nucleus of newly-assembled snRNPs monitored. This results in differently localized mature and newly-assembled snRNPs (Fig. 1A). In cells stably expressing FP-SmB, the majority of FP-tagged snRNPs are mature, with only a small amount of newly assembled snRNPs present. Importantly for our experimental approach, the use of fluorescently tagged Sm proteins enables the maturity of tagged snRNPs to be assessed using live cell microscopy immediately prior to processing the cells for proteomic analysis.

2) Quantitative time-resolved proteomics identifies differences in the interactomes of newly assembled and mature splicing snRNPs.

To gain information about complexes involved in different stages of snRNP biogenesis, we adopted an affinity purification/ mass spectrometry (AP/MS) strategy involving differential stable isotope labeling, which allows a quantitative comparison of the relative distribution of interacting proteins between mature and newly assembled snRNPs and the ready elimination of non-specific contaminants. Mature snRNPs were isolated from a cell line constitutively expressing a YFP-tagged version of the core snRNP protein, SmB (Sleeman et al, 2003), while newly assembled snRNPs were isolated from cells transiently transfected with YFP-SmB and harvested after 24 hours, before the snRNPs had accumulated in nuclear speckles. The sub-cellular distribution of the YFP-SmB protein was assessed by fluorescence microscopy immediately prior to harvesting the cells (supplementary Fig. 1). The isolation

was carried out using the high-affinity GFP-Trap® reagent conjugated to agarose beads (Chromotek) (Fig. 1B).

This time-resolved screen identified a total of 344 proteins, 152 of which are deemed likely contaminants due to their presence in the 'bead proteome' previously established using the same reagents in HeLa cells (Trinkle-Mulcahy et al, 2008). In addition to identifying the remaining six members of the core Sm complex, we identified six members of the Lsm (like Sm) protein family, five of which form the core of splicing snRNP U6 and one of which, Lsm11, is present with SmB in the core of the low-abundance U7 snRNP which plays a role in histone RNA processing. Detection of Lsm11 confirms the sensitivity of the screen. Other key groups of proteins identified include non-Sm snRNP proteins and other splicing factors; SMN and several of the gemin proteins (members of the SMN complex); and members of the PRMT5 cytoplasmic methylation complex (table I).

Direct comparison of the degree of enrichment of detected proteins above background with mature (log H:L) vs. newly-assembled (log M:L) snRNPs clearly identifies interaction partners, both known and novel, and highlights key differences between the two pools (Fig. 2A). For example, although Sm proteins (yellow) and several other snRNP proteins (orange) are clearly enriched in both samples, non-snRNP splicing factors (green) are more enriched with mature snRNPs, while a larger number of transport factors (purple) co-purify with newly-assembled snRNPs. The log H:M ratios directly measure relative enrichment with mature (H) versus newly assembled (M) snRNPs, and information about relative abundance can be added by plotting these ratios vs. the summed intensities of all peptides identified for that particular protein (normalized by molecular weight) (Fig. 2B). As expected, the most highly enriched protein was the YFP-SmB bait protein. Although slightly more was pulled down from transiently transfected cells (newly-assembled snRNPs), it still falls below the chosen threshold (>1 log above or below the median log H:M) and is therefore not significant. The most highly enriched and abundant mature snRNP interactors are the Sm and other snRNP proteins. Importantly, a large number of non-snRNP splicing factors are also present in the Sm protein interactome, demonstrating conclusively for the first time that FP-tagged Sm proteins are not only assembled into splicing snRNPs, but are also recruited into the spliceosome. Most of the proteins known to be involved in snRNP maturation including the SMN complex and the PRMT5 complex involved in methylation of Sm

proteins plot near or slightly below the axis, indicating that they show a slight enrichment in the newly assembled sample. This is consistent with the suggestion that snRNP maturation is highly active in HeLa cells, as evidenced by their prominent Cajal bodies (Kaiser et al, 2008; Matera & Shpargel, 2006; Sleeman et al, 2001). Proteins enriched more highly in the newly assembled snRNP sample also include Crm1, which we have previously demonstrated to have a role in the targeting of new snRNPs to Cajal bodies (Sleeman, 2007) and a number of proteins potentially involved in the trafficking of new snRNPs in the cytoplasm. Since little is known about the physical organization and dynamics of the early stages of snRNP assembly in the cytoplasm, the identification of dynein heavy chain: a microtubule associated motor protein, and coatomer proteins beta and gamma, which are usually associated with endoplasmic reticulum trafficking vesicles was of particular interest. Early stages of snRNP assembly are of potential importance in understanding the molecular pathology of SMA. Dynein light intermediate chain (DYNCILI2) was also identified in the screen, falling just outside the threshold defining proteins preferentially interacting with newly-assembled snRNPs.

3) FP-tagged SmB forms highly mobile punctate structures in the neurites of human neuroblastoma cell line SH-SY5Y

In HeLa cells, cytoplasmic accumulations containing Sm proteins and SMN have been reported in cells over-expressing SMN (Sleeman et al, 2003), while U-bodies containing snRNPs, SMN and members of the SMN complex have been implicated in splicing snRNP maturation in Drosophila egg chambers (Cauchi et al, 2010; Liu & Gall, 2007). SMN-rich granules have also been detected in the cytoplasm of neural cells (Fallini et al, 2010; Jablonka et al, 2001; Peter et al, 2011; Sharma et al, 2005; Todd et al, 2010a; Todd et al, 2010b; Zhang et al, 2006). These have been widely reported to lack Sm proteins, leading to the suggestion that they are involved in alternative neural-specific roles for SMN. The potential importance of cytoplasmic stages of snRNP maturation in neural cells and its significance for SMA led us to investigate the cytoplasmic distribution of mCherry-tagged SmB in the human neuroblastoma cell line, SH-SY5Y. Time-lapse fluorescence microscopy of the SHY5mCherrySmB cell line, stably expressing mCherry-SmB in an SH-SY5Y human neuroblastoma background (Clelland et al, 2012), revealed small, rapidly moving, punctate structures within the cytoplasm (Fig. 3). Cytoplasmic structures were not observed in cells

stably expressing the U1-snRNP specific protein, U170K or the non-snRNP splicing factor ASF/SF2 fused to mCherry (Clelland et al, 2012) and other unpublished observations). Differentiation of SH-SY5Y cells using retinoic acid (RA) followed by brain derived neurotrophic factor (BDNF) leads to the production of extensive neurites (Clelland et al, 2009; Encinas et al, 2000). Differentiated cells from line SHY5mCherrySmB showed prominent SmB-containing structures moving rapidly in a bi-directional manner within the neurites (figure 3 and supplementary movie S1). The increased prominence and mobility of cytoplasmic structures containing SmB in differentiated cells suggests that they may play an important role in mature neural cells.

4) Correlative confocal fluorescence and electron microscopy identifies distinctive vesicles containing mCherry-SmB.

To gain detailed structural information about the cytoplasmic distribution of mCherry-SmB in the neurites of SH-SY5Y cells, we carried out correlative confocal and electron microscopy. Cells from line SHY5mCherrySmB were imaged by confocal microscopy, with selected cells processed for subsequent transmission electron microscopy. Careful spatial alignment of the resulting confocal and electron microscope images revealed that the mCherry-SmB signal is found in vesicles of around 50nm diameter with a slightly irregular outline (Fig. 4). In some cases, these vesicles are found adjacent to granular structures of similar size.

5) mCherry-SmB interacts with Dynein and the coatomer complex in cytoplasmic vesicles.

Our initial SILAC screen identified Dynein heavy chain (DYNHCI), part of a molecular motor responsible for transport of cargo along microtubules, and two sub-units of the coatomer complex γCOP and βCOP, associated with COPI trafficking vesicles, as proteins interacting preferentially with newly expressed SmB (Fig. 2 and table I). To confirm the interactions between SmB and these two protein complexes, immunoprecipitations were carried out. Whole cell lysates from line SHY5mCherry-SmB were incubated with RFP-TRAP® conjugated agarose beads (Chromotek) or unconjugated agarose beads and the isolated proteins used for western blot analysis with antibodies to DYNHCI and γCOP. Both of these endogenous proteins co-immunoprecipitated with mCherry-SmB (Fig. 5A). Since the coatomer complex is associated with COPI vesicles, normally linked to retrograde

transport within the golgi network (reviewed in Szul & Sztul, 2011), we next addressed the localization of mCherry-SmB with respect to COPI vesicles within the cytoplasm. We first established an SH-SY5Y cell line, SHY1GFPεCOP, stably expressing GFP-tagged εCOP (a gift from Presley and Lippencott-Schwartz) (Presley et al, 2002). Immunoprecipitation using GFP-TRAP® followed by mass spectrometric analysis confirmed that the GFP-εCOP associated with other members of the COPI complex (unpublished observations). mCherrySmB was then transiently expressed in this cell line and the two fluorescent proteins imaged in living cells. The resulting time-lapse sequences indicate that the mCherry-SmB structures represent a subset of GFP-εCOP-tagged vesicles (Fig. 5B and supplementary movie S2). Further evidence that the mCherry-SmB structures are a type of membrane vesicle was obtained by incubating differentiated cells from line SHY5mCherry-SmB with the fluorescent lipid dye, BODIPY488 (4μg/ml 16hrs). Time-lapse imaging showed co-localization between the mCherry-SmB vesicles and small lipid-rich structures stained with BODIPY488 (Fig. 5C,D, supplementary movies S3 and S4).

6) SmB vesicles move on microtubules

The identification of dynein, a microtubule motor, in our time-resolved proteomic screen suggested that newly-expressed Sm proteins interact with microtubules. In order to investigate the trafficking of the newly-identified Sm protein vesicles, differentiated cells from line SHY5mCherrySmB were imaged using time-lapse microscopy before and after treatment with inhibitors of cytoskeletal dynamics. In control cells, the vesicles showed saltatory movement characteristic of microtubule trafficking with maximum velocities of up to 2.5μ/s (Fig. 6A). Disruption of microtubules with nocodazole (30μM for three hours) resulted in the formation of larger, irregular SmB vesicles (Fig. 6B) and dramatically reduced their mobility (Fig. 6A). Inhibition of actin microfilaments with cytochalasin-D (2μg/ml for 2.5hrs) did not alter the appearance or mobility of the SmB vesicles (Fig. 6A). To confirm that the cytochalasin D treatment had been effective, cells were stained with phalloidin (Fig. 6C).

7) SmB vesicles contain SMN, members of the SMN complex and other Sm proteins but not other splicing factors or the transport factor CRM1.

A number of recent studies have identified mobile granules in neural cells that contain the survival motor neurons (SMN) protein. These granules showed partial association with other

members of the SMN complex (Fallini et al, 2010; Todd et al, 2010a; Todd et al, 2010b; Zhang et al, 2006) and with the COPI complex (Peter et al, 2011) but were not demonstrated to contain Sm proteins. To investigate the contents of the SmB vesicles, differentiated cells from line SHY5mCherrySmB were used for immunodetection of SMN, gemin2 and Sm proteins. However, although the SmB vesicles were clearly visible in fixed cells, they were disrupted on permeabilization of the cells using detergents or solvents (supplementary Fig. 2). This reinforces the conclusion that they are lipid-based structures. Furthermore, we also observed that different primary antibodies used to detect SMN appear to stain different cytoplasmic granules and show only partial co-localization when used together in the same cells (supplementary Fig. 3).

To circumvent these technical difficulties, cells from line SHY5mCherrySmB were transfected with plasmid constructs to express GFP-tagged versions of representative proteins from the classes identified in the SILAC screen. SMN, gemin2, and SmD1 represent proteins associated with snRNP maturation events; U5snRNP 100K represents non-Sm snRNP proteins; ASF/SF2 represents other splicing factors and CRM1 represents transport factors. We also included GFP-DCP1A (decapping mRNA 1) to determine whether the SmB-rich structures co-localized with P-bodies. Transfected cells were imaged either live, or fixed with paraformaldehyde but not permeabilized (Fig. 7, supplementary movies S5 and S6). GFP-SMN co-localized with mCherry-SmB in vesicles near the cell body and in neurites Fig. 7A,B). GFP-Gemin2 (Fig. 7C and supplementary movie S5) and GFP-SmD1 (Fig. 7D, supplementary movie S6) showed co-localization with a sub-set of the mCherry-SmB vesicles, but the co-localization was less complete than that seen between SmB and SMN. In contrast, GFP-U5 snRNP 100K (Fig. 7E) and GFP-ASF/SF2 (Fig. 7F) did not co-localize with the mCherry-SmB cytoplasmic vesicles (Fig. 7E). Both GFP-CRM1 (Fig. 7G) and GFP-DCP1A (Fig. 7H) showed cytoplasmic accumulations, but these did not co-localize with the SmB vesicles. This suggests that SmB and SMN form the core of the mobile vesicles, with Gemin2 and SmD1 present in a proportion of the vesicles, which are, therefore, implicated as the cytoplasmic sites of snRNP assembly. The roles for the rest of the SmB vesicles are not clear. The absence of U5snRNP 100K and ASF/SF2 is in agreement with current models of later stages of snRNP maturation and spliceosome assembly taking place within the nucleus. CRM1 does not appear to be important in the SmB vesicles, while the lack of co-localisation with DCP1A, together with the absence of P-body associated factors from the SILAC screen,

indicates that these structures are not P-bodies so are not likely to be involved in mRNA degradation.

8) Altered levels of SMN affect both the morphology and the dynamics of the SmB vesicles

In the experiments described above, we observed that cells expressing high levels of GFP-SMN showed a significant disruption of the mCherry-SmB cytoplasmic vesicles. The vesicles in these cells were larger and less mobile than vesicles in control cells (Fig. 8). Furthermore, they also contained higher levels of mCherry-SmB protein than seen in untransfected cells (0.23% s.e.m+/-0.05 of total cellular mCherrySmB signal is found in vesicles in control cells; 3.79% s.e.m+/-1.9 in cells transfected with GFP-SMN; Student's T-test $P<0.05$, $n=11$), suggesting that the SMN protein is instrumental in the formation and dynamics of the vesicles and in the regulation of their composition. Reduction of the amount of SMN present in cells from line SHY5mCherrySmB, using siRNA sequences that result in approximately 70% depletion of SMN (Clelland et al, 2012) did not, however, result in the loss of SmB cytoplasmic vesicles (Fig. 8D). Furthermore, depletion of PRMT5 by approximately 75% (Clelland et al, 2007) did not abolish the vesicles. Methylation of Sm proteins by PRMT5 is required for their addition to snRNAs during early snRNP maturation. To further investigate the requirements for SmB vesicle formation, we used siRNA to reduce the amount of SmB in an SH-SY5Y cell line constitutively expressing GFP-SMN (Clelland et al 2009) (Fig. 8E). This cell line expresses relatively high levels of GFP-SMN protein, resulting in the accumulation of GFP-SMN in a small number of cytoplasmic structures in addition to nuclear CB/gems. Reduction of SMN results in the loss of both nuclear and cytoplasmic GFP-SMN. Reduction of SmB by approximately 60% in these cells resulted in the complete loss of GFP-SMN accumulations from the cytoplasm together with a marked increase in nuclear bodies containing GFP-SMN. Due to the striking nature of this result, we investigated whether it was a direct result of SmB depletion or an indirect consequence of reduced snRNP assembly by inhibiting snRNP assembly with leptomycin B in the same cell line. LMB treatment, in contrast to SmB depletion, resulted in the appearance of a large number of GFP-SMN-rich structures in the cytoplasm, but no increase in nuclear GFP-SMN accumulations. This unexpected result strongly suggests that SmB has a key role in the formation and maintenance of SMN-rich cytoplasmic structures.

DISCUSSION

Time-resolved quantitative proteomics can be used to dissect complex biological pathways

Using the core snRNP protein SmB tagged with YFP, we have established a technique to identify proteins involved in different stages of the complex biogenesis pathway of splicing snRNPs. Our SILAC screen showed a clear enrichment of non-snRNP splicing factors and snRNP-specific proteins in the interactome of mature snRNPs when compared to that of newly expressed SmB. In contrast, an enrichment of the majority of snRNP processing factors, including SMN and a number of the gemin proteins, was seen in the newly-expressed sample. Importantly, the nuclear export factor CRM1, which has a role in targeting newly imported snRNPs to the Cajal body, was enriched in the newly-expressed interactome, confirming that differences observed in the interactomes reflect different stages of snRNP biogenesis and are not simply a consequence of higher amounts of cytoplasmic YFP-SmB in the newly-expressed sample. Confirmation of the interaction of SmB with both γCOP and DYNHCI in cell lines stably expressing SmB demonstrates that interactions identified by transiently expressing SmB are biologically relevant and not a result of a transient over-expression. The identification of a number of non-snRNP splicing factors in the YFP-SmB interactomes demonstrates that FP-tagged Sm proteins are incorporated into active spliceosomes, while the grouping of different classes of protein seen by plotting log H:L against log M:L has the potential to reveal information about the relationships between proteins identified in this and similar screens. Of particular interest in this respect is the position of the minor U11-U12 snRNP associated protein, U1SNRNPBP (SNRNP35), which is clearly separate from the major spliceosmal snRNP proteins in the top left section of the graph.

SmB colocalizes with COPI-associated vesicles on microtubules in neural cells

Among the proteins found to associate preferentially with newly-expressed SmB was the molecular motor protein, dynein heavy chain 1. Examination of the cytoplasmic distribution and dynamics of mCherry-SmB in the neuroblastoma cell line, SH-SY5Y, revealed highly mobile structures. Importantly, these structures were more prominent in cells differentiated using retinoic acid followed by BDNF. This suggests they may be of particular importance in neural cells. Analysis of the movement of these structures showed the saltatory motion associated with microtubule trafficking. The cessation of their movement in cells treated with nocodazole confirmed that they are transported along microtubules. The identification of

γCOP, a member of the COPI complex usually associated with vesicle trafficking in the endoplasmic reticulum, in the interactome of newly-expressed SmB was unexpected and led us to characterize the cytoplasmic SmB structures further. Expression of mCherry-SmB in an SH-SY5Y cell line constitutively expressing εCOP, another member of the coatomer complex, demonstrated that these two proteins co-localize in the mobile structures, while staining of cells expressing mCherry-SmB with the lipophilic dye BODIPY-488 demonstrates the SmB-containing structures to be lipid-rich. Together with their extreme sensitivity to extraction with detergents or alcohol when fixed with paraformaldehye for immunofluorescence, this strongly suggests that the SmB-containing structures are a novel class of lipid-rich vesicle associated with the coatomer complex. Correlative confocal fluorescence and transmission electron microscopy reveals that the SmB-containing structures have morphology consistent with lipid-rich vesicles. Interestingly, some of the vesicles appear to be immediately adjacent to granular structures of a similar size. This is consistent with recent electron microscopic analysis of SMN-containing structures associated with the Golgi complex in mouse motor neuron-like NSC34 cells (Ting et al, 2012).

How do SmB vesicles relate to other cytoplasmic structures implicated in snRNP processing and SMA?

A number of other cytoplasmic structures containing snRNP components and processing factors, including SMN, have been documented. In Drosophila, 'U' bodies, containing snRNAs, Sm proteins and the SMN protein, have been identified as a potential location for cytoplasmic stages of snRNP maturation (Liu & Gall, 2007). In contrast, several reports have identified structures in various neural cells containing the SMN protein and other members of the SMN-complex in the apparent absence of any snRNP components (Fallini et al, 2010; Jablonka et al, 2001; Peter et al, 2011; Sharma et al, 2005; Todd et al, 2010a; Todd et al, 2010b; Zhang et al, 2006). One of these reports (Peter et al, 2011) also demonstrated an association of SMN with the coatomer complex and with β-actin mRNA, suggesting that the SMN-containing structures are involved in trafficking messenger RNAs to growth cones and represent a nerve cell-specific role for SMN beyond its role in splicing snRNP assembly. The SmB vesicles that we have detected contain both SMN and the coatomer complex. They also show some co-localization with Gemin2 and SmD1, although Gemin2 and SmD1 appear to localize to only a sub-set of the SmB/SMN vesicles. Previous work from several groups has demonstrated that Gemin2 and other members of the snRNP assembly machinery co-localize

with a sub-set of SMN-positive structures in a number of neural cell types (reviewed in Fallini et al, 2012). Combined with our new data, this is suggestive of a family of related vesicles, with SMN and SmB as core components. The vesicles containing Gemin2, known to be responsible for the addition of Sm proteins to the snRNA core (Zhang et al, 2011), and other Sm proteins are likely to be involved in splicing snRNP assembly. Others may have different roles, for example in messenger RNA trafficking or signal recognition particle (SRP) assembly (Piazzon et al, 2013). We propose that SmB and SMN are core components of a family of cytoplasmic vesicles involved in ribonucleoprotein assembly, RNA trafficking and potentially other areas of RNA metabolism.

An expanding range of functions for Sm proteins

SmB, together with SmD3, has been implicated in the developmentally important localization of *oskar* mRNA in Drosophila (Gonsalvez et al, 2010) while Sm proteins are also involved in the maintenance of germ granule integrity in C.elegans (Barbee et al, 2002). A sub-set of Sm proteins including SmB have been demonstrated to interact with telomerase RNA and scaRNAs in mammalian cells (Fu & Collins, 2006; Fu & Collins, 2007), while in yeast, Sm proteins have recently been implicated in telomerase RNA processing (Tang et al, 2012). Taken together, these data demonstrate an expanding number of roles for SmB in the processing and regulation of RNAs beyond its role in pre-messenger RNA splicing. The bacterial Sm-like protein Hfq, the likely evolutionary origin of both Sm and LSm proteins, has been implicated as a general co-factor for chromosomally encoded regulatory RNAs (reviewed in Valentin-Hansen et al, 2004). We have now demonstrated that SmB is vital for the formation of SMN-rich cytoplasmic structures in neural cells, implicating it in the pathology of SMA whether or not that pathology is linked to snRNP assembly.

Cytoplasmic trafficking by dynein in neurodegeneration

The importance of dynein-based microtubule trafficking in neurons is well established (reviewed in Hirokawa et al, 2010), with dynein implicated in the pathologies of a number of neurodegenerative conditions (reviewed in Banks & Fisher, 2008; Eschbach & Dupuis, 2011). We have demonstrated that SmB interacts with DYNCH1 and forms vesicles that are trafficked on microtubules. Clearly, further analysis of these vesicles is required to fully understand their function or functions within neurons. It is, however, clear that both SMN and SmB are key components of these vesicles. There is increasing evidence for a role for

SMN in the transport and local translation of axonal mRNAs (reviewed in Fallini et al, 2012). The effect of over-expression of GFP-SMN in cell lines constitutively expressing mCherry-SmB is striking, with large, almost immobile structures forming. These abnormal structures also recruit a substantial amount of SmB. This is highly reminiscent of the effect of over-expression of SMN on the structure of the nucleus in SH-SY5Y cells (Clelland et al, 2009), where it results in the formation of a Cajal body complex containing SMN, coilin and Sm proteins in place of the separate CBs and gems usually seen in this cell line. This further implicates SMN in the formation of structures, both nuclear and cytoplasmic, involved in splicing snRNP metabolism. Suprisingly, however, many of the SmB/SMN vesicles did not co-localize with other members of the Sm protein family or of the SMN complex, making it unlikely that all of the SmB/SMN vesicles are involved in snRNP maturation. Importantly, our data suggest that the relative levels of expression of SmB and SMN determine the formation and mobility of a family of cytoplasmic vesicles implicated both in snRNP maturation, analogous to Drosophila U bodies, and in the trafficking of messenger RNAs, analogous to SMN-containing structures identified in neural cells. Interestingly, previous work has shown that GFP-tagged human SMN protein lacking exon 7, the most common mutation seen in SMA, is unable to localize to cytoplasmic structures in chick neurons (Zhang et al, 2003). We have shown that SMN and SmB are key to the integrity of a family of cytoplasmic trafficking vesicles, providing a mechanistic link between the SMN and SmB proteins and defects in RNA localization and trafficking within axons proposed in models of SMA (Fallini et al, 2012).

The Implications for SMA pathology: a splicing or a trafficking defect?

The SMN complex has a well-established role in assembling the Sm protein ring onto snRNAs during splicing snRNP maturation, with alterations both in the cellular snRNP repertoire and in pre-mRNA splicing patterns documented in several models of SMN (Baumer et al, 2009; Boulisfane et al, 2011; Campion et al, 2010; Gabanella et al, 2007; Zhang et al, 2008). Despite this, it is difficult to reconcile a defect in a process as ubiquitous as snRNP assembly with the motor neuron-specific symptoms seen in SMA, making a role for SMN in mRNA trafficking in neural cells an attractive alternative hypothesis. The involvement of SmB in SMN trafficking vesicles in neurites highlights the closely interlinked nature of mRNA processing and trafficking pathways. Taken together with the recent documentation of mis-splicing of Stasimon, proposed to have a role in vesicular cargo

transport, in a Drosophila model of SMA (Imlach et al, 2012; Lotti et al, 2012), our data suggest that the defects caused in cells by lowered levels of SMN converge on pathways of RNP transport that are of particular importance in cells, such as motor neurons, where local translation of mRNAs, at some distance from their site of transcription within the nucleus, is of key importance.

Materials and Methods

Cell lines and cell culture

HeLa and SH-SY5Y cells were from ATCC. Cells were cultured in DMEM with 10% FBS at 37ºC, 5% CO_2. Transfections were carried out using Effectene (Qiagen) according to the manufacturers instructions. pYFP-SmB, pCFP-SmB, pGFP-ASF, stable HeLa cell lines expressing YFP- and CFP-SmB and stable SH-SY5Y cell lines expressing mCherry-SmB and GFP-SMN have been described previously (Clelland et al, 2009; (Sleeman et al., 1998); Sleeman et al, 2003). pGFP-εCOP (Presley et al, 2002) was a gift from J. Presley and J. Lippencott-Schwartz. pGFP-DCP1A (Aizer et al., 2008) was a gift from Y. Shav-Tal. pGFP-U5snRNP 100K was a gift from A.I. Lamond. SH-SY5Y cell lines stably expressing GFP-εCOP were derived by selection with 200μg/ml G418 (Roche) following transfection. Differentiation of SH-SY5Y cells using retinoic acid and BDNF was carried out as described previously (Clelland et al, 2009), as were heterokaryon assays (Sleeman et al, 2001). BODIPY488 (Life Technologies) was added to culture medium at 2μg/ml overnight and leptomycin B (LMB) at 100ng/ml for 16 hours.

Microscopy and Image Analysis

For live-cell microscopy, cells were grown on 40 mm diameter glass coverslips and transfected as appropriate. Coverslips were transferred to an open chamber (Zeiss) within an environmental incubator (Solent Scientific, Segensworth, UK) on an Olympus DeltaVision RT microscope (Applied Precision), maintained at 37°C with 5% CO_2. Time-lapse images were taken at approximately 1s intervals using single z-sections or three z-sections at 0.5μm spacing for deconvolution, 100X 1.35 NA objective. Exposure times were between 0.2s and 1.0s, giving maximum pixel intensities of between 500 and 750. For fixed cell microscopy, z-sections of 0.2μm spacing were taken through the entire sample with exposures giving maximum pixel intensities of between 3000 and 4500. Deconvolution was carried out and movies generated using Volocity software (Perkin Elmer). Vesicles were tracked manually in Volocity. mCherry-SmB signals in whole cells and cytoplasmic vesicles were quantified using Volocity by identifying objects by their intensity and size. Line profiles were generated using SoftWorx (Applied Precision). Montages and overlays of images were made using Adobe Photoshop CS4. Bar charts and statistical analyses were generated using Prism4 (GraphPad).

Correlative Confocal and Electron Microscopy

Cells on CELLocate coverslips (Eppendorf) were fixed in 4% paraformaldehyde in PBS and imaged on an LSM700 confocal microscope using the 40X Plan-Neofluar NA1.3 and 100X alpha-Apochromat NA1.46 objectives. Z stacks of fluorescent cells were collected and located on the CELLocate grid using DIC optics. Cells were re-fixed in 2.5% gluteraldehyde in 0.1M cacodylate buffer pH7.2 and post-fixed in 1% OsO4 with1.5% Na ferrocyanide in cacodylate buffer. Coverslips were dehydrated in alcohol and embedded in Durcupan resin (Sigma). Coverslips were removed from the resin stubs leaving the CELLocate grid impression on the resin, and the region containing the required cells was trimmed and serially sectioned, collecting every section through the cells. Sections were stained with lead citrate only. Fluorescent cells were identified from their relationship to other cells in the DIC image and TEM images were taken from all the sections collected until the whole of the cell had been imaged. Stacks of images were imported into Volocity (Perkin-Elmer) and cross-correlated with the TEM images

Preparation of Lysates, IPs and Immunoblotting

Lysates were prepared as described previously (Sleeman et al, 2003). To isolate complexes containing mCherry-SmB, total cell lysates were cleared overnight with agarose beads then incubated for 2 hours with RFP-TRAP-A (Chromotek) or agarose beads. Following five washes with RIPA buffer (50mM Tris-HCl pH7.5; 150mM NaCl; 1% v/v NP40, 0.5% w/v NaDOC and protease inhibitors) samples were separated by SDS-PAGE (Novex, Invitrogen) and transferred to nitrocellulose membranes (Hybond-C+, GE Healthcare) for immunoblotting. Antibodies used were rat mAb anti-RFP (Chromotek, 1:500); goat polyclonal anti-γCOP (Santa Cruz, 1:250) and polyclonal rabbit anti-DYNHCI (Santa Cruz, 1:200). Secondary antibodies were horseradish peroxidase (HRP)-conjugated (Pierce, 1:20,000). Detection was carried out with ECL Plus (GE Healthcare) and imaged using a Fujifilm LAS-3000 imaging system.

siRNA knock-downs

Reduction of protein expression was achieved by transfecting the appropriate cell lines with siRNAs (Thermo Scientific) using viromer green (Lipocalyx GmbH) according to the manufacturer's instructions. Cells were lysed for assay by Western blot or fixed with paraformaldehyde for fluorescence microscopy 48 hours after transfection. Sequences used

were: SMN: CAGUGGAAAGUUGGGGACA, negative control targeting luciferase: UAAGGCUAUGAAGAGAUAC, positive control targeting cyclophilin B: GGAAAGACUGUUCCAAAAA, PRMT5: a mixture of CAACAGAGAUCCUAUGAUU, CCAAGUGACCGUAGUCUCA, UGGCACAACUUCCGGACUU, CGAAAUAGCUGACACACUA, SmB: a mixture of CCCACAAGGAAGAGGUACU, GCAUAUUGAUUACAGGAUG, CCGUAAGGCUGUACAUAGU, CAAUGACAGUAGAGGGACC.

Quantitative SILAC IP

Quantitative SILAC-based affinity purification/mass spectrometry experiments were carried out as previously described (Trinkle-Mulcahy et al, 2008). HeLa cells stably expressing free GFP were encoded by passaging in medium containing "light" L-arginine and L-lysine (R0K0; Sigma-Aldrich). HeLa cells stably expressing YFP-SmB were encoded by passaging in medium containing "heavy" [^{13}C/^{15}N]L-arginine and [^{13}C/^{15}N]L-lysine (R10K8; Cambridge Isotope Laboratories, Andover, MA). Parental cells were encoded by passaging for 7 days in medium containing "medium" [^{13}C]L-arginine and [4,4,5,5-D4]L-lysine (R6K4) and then transiently transfected with YFP-SmB for 24 hrs. The sub-cellular distribution of YFP-SmB in these samples was determined, by fluorescence microscopy, to be diffuse with some accumulation in CBs immediately prior to harvesting. Cells were harvested and extracts prepared by sonicating in RIPA buffer and centrifuging at 2800 g for 10 min at 4°C. Total protein concentrations were measured using a Biophotometer (Eppendorf).

For affinity purification of free GFP and YFP-tagged SmB, equivalent total protein amounts of each extract were incubated with GFP-Trap_A™ (Chromotek, Martinsried, Germany) for 1 hr at 4ºC. Extracts were then removed and the beads washed four times with ice-cold RIPA buffer. Proteins were eluted, separated by 1D SDS-PAGE, and trypsin-digested for MS analysis as described previously (Trinkle-Mulcahy et al, 2008).

Mass Spectrometry and Data Analysis

An aliquot of the tryptic digest (prepared in 5% acetonitrile/0.1% trifluoroacetic acid in water) was analyzed by LC-MS on an LTQ-Orbitrap mass spectrometer system (ThermoElectron, Rockford, IL) as previously described (Trinkle-Mulcahy et al, 2008). The Orbitrap was set to analyze the survey scans at 60,000 resolution and the top five ions in each

duty cycle selected for MS/MS in the LTQ linear ion trap. Database searching (against the human IPI database v3.68; 87,061 entries) and quantitation were performed using MaxQuant software v1.1.1.14 and the Andromeda search engine (Cox et al, 2009; Cox et al, 2011). The following criteria were used: peptide tolerance = 10 ppm, trypsin as the enzyme (2 missed cleavages allowed), carboxyamidomethylation of cysteine as a fixed modification. Variable modifications were oxidation of methionine and N-terminal acetylation. Medium SILAC labels were Arg6 and Lys4, heavy SILAC labels were Arg10 and Lys8. Minimum ratio count was 2 and quantitation was based on razor and unique peptides. Peptide and protein FDR was 0.01. The full dataset (minus common environmental contaminants as per http://maxquant.org and proteins identified via the decoy database) is provided as Supplemental Data.

Acknowledgements

The authors thank Lawrence Puente, Ottawa Hospital Research Institute Proteomics Core Facility for technical support; J. Lippencott_Schwartz (NIH, Bethesda) and J. Presley (McGill University, Quebec) for pεCOP-GFP; Y. Shav-Tal (Bar-Ilan University, Israel) for pGFP-DCP1A; A.I.Lamond (University of Dundee, UK) for pU5snRNP 100K; M. Fuszard and C. Botting, BSRC MS and Proteomics Facility, University of St Andrews for MS analysis of εCOP-GFP cell lines; G. Morris, RJAH Orthopaedic Hospital, Oswestry, UK for MANSMA1 antibody. LTM holds a CIHR New Investigator Salary Support Award. This work was supported in part by the Royal Society.

Author Contribution

ARP acquired and interpreted the correlative light/electron microscopy images and helped to edit the final manuscript. AB assisted with acquiring and analyzing data and with drafting the article. JJ assisted with acquiring and analyzing data. LTM analyzed the AP/MS data, generated the related tables and figures and helped to edit the final manuscript. JES came up with the original concept, designed all experiments, carried out imaging, affinity purification and Western blot experiments and generated the related figures, analyzed data and prepared the manuscript.

Conflict of Interest

The authors declare no conflict of interests.

REFERENCES

Aizer, A., Brody, Y., Ler, L. W., Sonenberg, N., Singer, R. H. and Shav-Tal, Y. (2008). The dynamics of mammalian P body transport, assembly, and disassembly in vivo. *Mol Biol Cell* **19**, 4154-4166.

Banks, G. T. and Fisher, E. M. (2008). Cytoplasmic dynein could be key to understanding neurodegeneration. *Genome Biol* **9**, 214.

Barbee, S. A., Lublin, A. L. and Evans, T. C. (2002). A novel function for the Sm proteins in germ granule localization during C. elegans embryogenesis. *Curr Biol* **12**, 1502-1506.

Battle, D. J., Kasim, M., Yong, J., Lotti, F., Lau, C. K., Mouaikel, J., Zhang, Z., Han, K., Wan, L. and Dreyfuss, G. (2006). The SMN complex: an assembly machine for RNPs. *Cold Spring Harb Symp Quant Biol* **71**, 313-320.

Baumer, D., Lee, S., Nicholson, G., Davies, J. L., Parkinson, N. J., Murray, L. M., Gillingwater, T. H., Ansorge, O., Davies, K. E. and Talbot, K. (2009). Alternative splicing events are a late feature of pathology in a mouse model of spinal muscular atrophy. *PLoS Genet* **5**, e1000773.

Boulisfane, N., Choleza, M., Rage, F., Neel, H., Soret, J. and Bordonne, R. (2011). Impaired minor tri-snRNP assembly generates differential splicing defects of U12-type introns in lymphoblasts derived from a type I SMA patient. *Hum Mol Genet* **20**, 641-648.

Campion, Y., Neel, H., Gostan, T., Soret, J. and Bordonne, R. (2010). Specific splicing defects in S. pombe carrying a degron allele of the Survival of Motor Neuron gene. *Embo J* **29**, 1817-1829.

Cauchi, R. J., Sanchez-Pulido, L. and Liu, J. L. (2010). Drosophila SMN complex proteins Gemin2, Gemin3, and Gemin5 are components of U bodies. *Exp Cell Res* **316**, 2354-2364.

Clelland, A. K., Bales, A. B. and Sleeman, J. E. (2012). Changes in intranuclear mobility of mature snRNPs provide a mechanism for splicing defects in spinal muscular atrophy. *J Cell Sci* **125**, 2626-2637.

Clelland, A. K., Kinnear, N. P., Oram, L., Burza, J. and Sleeman, J. E. (2009). The SMN protein is a key regulator of nuclear architecture in differentiating neuroblastoma cells. *Traffic* **10**, 1585-1598.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V. and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* **10**, 1794-1805.

Cox, J., Matic, I., Hilger, M., Nagaraj, N., Selbach, M., Olsen, J. V. and Mann, M. (2009). A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat Protoc* **4**, 698-705.

Encinas, M., Iglesias, M., Liu, Y., Wang, H., Muhaisen, A., Cena, V., Gallego, C. and Comella, J. X. (2000). Sequential treatment of SH-SY5Y cells with retinoic acid and brain-derived neurotrophic factor gives rise to fully differentiated, neurotrophic factor-dependent, human neuron-like cells. *J Neurochem* **75**, 991-1003.

Eschbach, J. and Dupuis, L. (2011). Cytoplasmic dynein in neurodegeneration. *Pharmacol Ther* **130**, 348-363.

Fallini, C., Bassell, G. J. and Rossoll, W. (2010). High-efficiency transfection of cultured primary motor neurons to study protein localization, trafficking, and function. *Mol Neurodegener* **5**, 17.

Fallini, C., Bassell, G. J. and Rossoll, W. (2012). Spinal muscular atrophy: the role of SMN in axonal mRNA regulation. *Brain Res* **1462**, 81-92.

Fallini, C., Zhang, H., Su, Y., Silani, V., Singer, R. H., Rossoll, W. and Bassell, G. J. (2011). The survival of motor neuron (SMN) protein interacts with the mRNA-binding

protein HuD and regulates localization of poly(A) mRNA in primary motor neuron axons. *J Neurosci* **31**, 3914-3925.

Fan, L. and Simard, L. R. (2002). Survival motor neuron (SMN) protein: role in neurite outgrowth and neuromuscular maturation during neuronal differentiation and development. *Hum Mol Genet* **11**, 1605-1614.

Fischer, U., Englbrecht, C. and Chari, A. (2011). Biogenesis of spliceosomal small nuclear ribonucleoproteins. *Wiley Interdiscip Rev RNA* **2**, 718-731.

Fu, D. and Collins, K. (2006). Human telomerase and Cajal body ribonucleoproteins share a unique specificity of Sm protein association. *Genes Dev* **20**, 531-536.

Fu, D. and Collins, K. (2007). Purification of human telomerase complexes identifies factors involved in telomerase biogenesis and telomere length regulation. *Mol Cell* **28**, 773-785.

Gabanella, F., Butchbach, M. E., Saieva, L., Carissimi, C., Burghes, A. H. and Pellizzoni, L. (2007). Ribonucleoprotein assembly defects correlate with spinal muscular atrophy severity and preferentially affect a subset of spliceosomal snRNPs. *PLoS One* **2**, e921.

Gonsalvez, G. B., Rajendra, T. K., Wen, Y., Praveen, K. and Matera, A. G. (2010). Sm proteins specify germ cell fate by facilitating oskar mRNA localization. *Development* **137**, 2341-2351.

Hirokawa, N., Niwa, S. and Tanaka, Y. (2010). Molecular motors in neurons: transport mechanisms and roles in brain function, development, and disease. *Neuron* **68**, 610-638.

Hubers, L., Valderrama-Carvajal, H., Laframboise, J., Timbers, J., Sanchez, G. and Cote, J. (2011). HuD interacts with survival motor neuron protein and can rescue spinal muscular atrophy-like neuronal defects. *Hum Mol Genet* **20**, 553-579.

Imlach, W. L., Beck, E. S., Choi, B. J., Lotti, F., Pellizzoni, L. and McCabe, B. D. (2012). SMN is required for sensory-motor circuit function in Drosophila. *Cell* **151**, 427-439.

Jablonka, S., Bandilla, M., Wiese, S., Buhler, D., Wirth, B., Sendtner, M. and Fischer, U. (2001). Co-regulation of survival of motor neuron (SMN) protein and its interactor SIP1 during development and in spinal muscular atrophy. *Hum Mol Genet* **10**, 497-505.

Kaiser, T. E., Intine, R. V. and Dundr, M. (2008). De novo formation of a subnuclear body. *Science* **322**, 1713-1717.

Lefebvre, S., Burlet, P., Liu, Q., Bertrandy, S., Clermont, O., Munnich, A., Dreyfuss, G. and Melki, J. (1997). Correlation between severity and SMN protein level in spinal muscular atrophy. *Nat Genet* **16**, 265-269.

Lefebvre, S., Burglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., Benichou, B., Cruaud, C., Millasseau, P., Zeviani, M. et al. (1995). Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* **80**, 155-165.

Liu, J. L. and Gall, J. G. (2007). U bodies are cytoplasmic structures that contain uridine-rich small nuclear ribonucleoproteins and associate with P bodies. *Proc Natl Acad Sci U S A* **104**, 11655-11659.

Lotti, F., Imlach, W. L., Saieva, L., Beck, E. S., Hao le, T., Li, D. K., Jiao, W., Mentis, G. Z., Beattie, C. E., McCabe, B. D. et al. (2012). An SMN-dependent U12 splicing event essential for motor circuit function. *Cell* **151**, 440-454.

Matera, A. G. and Shpargel, K. B. (2006). Pumping RNA: nuclear bodybuilding along the RNP pipeline. *Curr Opin Cell Biol* **18**, 317-324.

Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A. and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* **1**, 376-386.

Pellizzoni, L., Kataoka, N., Charroux, B. and Dreyfuss, G. (1998). A novel function for SMN, the spinal muscular atrophy disease gene product, in pre-mRNA splicing. *Cell* **95**, 615-624.

Peter, C. J., Evans, M., Thayanithy, V., Taniguchi-Ishigaki, N., Bach, I., Kolpak, A., Bassell, G. J., Rossoll, W., Lorson, C. L., Bao, Z. Z. et al. (2011). The COPI vesicle complex binds and moves with survival motor neuron within axons. *Hum Mol Genet* **20**, 1701-1711.

Piazzon, N., Schlotter, F., Lefebvre, S., Dodre, M., Mereau, A., Soret, J., Besse, A., Barkats, M., Bordonne, R., Branlant, C. et al. (2013). Implication of the SMN complex in the biogenesis and steady state level of the signal recognition particle. *Nucleic Acids Res* **41**, 1255-1272.

Presley, J. F., Ward, T. H., Pfeifer, A. C., Siggia, E. D., Phair, R. D. and Lippincott-Schwartz, J. (2002). Dissection of COPI and Arf1 dynamics in vivo and role in Golgi membrane transport. *Nature* **417**, 187-193.

Sharma, A., Lambrechts, A., Hao le, T., Le, T. T., Sewry, C. A., Ampe, C., Burghes, A. H. and Morris, G. E. (2005). A role for complexes of survival of motor neurons (SMN) protein with gemins and profilin in neurite-like cytoplasmic extensions of cultured nerve cells. *Exp Cell Res* **309**, 185-197.

Sleeman, J. (2007). A regulatory role for CRM1 in the multi-directional trafficking of splicing snRNPs in the mammalian nucleus. *J Cell Sci* **120**, 1540-1550.

Sleeman, J., Lyon, C. E., Platani, M., Kreivi, J. P. and Lamond, A. I. (1998). Dynamic interactions between splicing snRNPs, coiled bodies and nucleoli revealed using snRNP protein fusions to the green fluorescent protein. *Exp Cell Res* **243**, 290-304.

Sleeman, J. E. and Lamond, A. I. (1999). Newly assembled snRNPs associate with coiled bodies before speckles, suggesting a nuclear snRNP maturation pathway. *Curr Biol* **9**, 1065-1074.

Sleeman, J. E., Ajuh, P. and Lamond, A. I. (2001). snRNP protein expression enhances the formation of Cajal bodies containing p80-coilin and SMN. *J Cell Sci* **114**, 4407-4419.

Sleeman, J. E., Trinkle-Mulcahy, L., Prescott, A. R., Ogg, S. C. and Lamond, A. I. (2003). Cajal body proteins SMN and Coilin show differential dynamic behaviour in vivo. *J Cell Sci* **116**, 2039-2050.

Spector, D. L. and Lamond, A. I. (2011). Nuclear speckles. *Cold Spring Harb Perspect Biol* **3**.

Szul, T. and Sztul, E. (2011). COPII and COPI traffic at the ER-Golgi interface. *Physiology (Bethesda)* **26**, 348-364.

Tang, W., Kannan, R., Blanchette, M. and Baumann, P. (2012). Telomerase RNA biogenesis involves sequential binding by Sm and Lsm complexes. *Nature* **484**, 260-264.

Ting, C. H., Wen, H. L., Liu, H. C., Hsieh-Li, H. M., Li, H. and Lin-Chao, S. (2012). The Spinal Muscular Atrophy Disease Protein SMN Is Linked to the Golgi Network. *PLoS One* **7**, e51826.

Todd, A. G., Morse, R., Shaw, D. J., Stebbings, H. and Young, P. J. (2010a). Analysis of SMN-neurite granules: Core Cajal body components are absent from SMN-cytoplasmic complexes. *Biochem Biophys Res Commun* **397**, 479-485.

Todd, A. G., Morse, R., Shaw, D. J., McGinley, S., Stebbings, H. and Young, P. J. (2010b). SMN, Gemin2 and Gemin3 associate with beta-actin mRNA in the cytoplasm of neuronal cells in vitro. *J Mol Biol* **401**, 681-689.

Trinkle-Mulcahy, L., Boulon, S., Lam, Y. W., Urcia, R., Boisvert, F. M., Vandermoere, F., Morrice, N. A., Swift, S., Rothbauer, U., Leonhardt, H. et al. (2008). Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *J Cell Biol* **183**, 223-239.

Valentin-Hansen, P., Eriksen, M. and Udesen, C. (2004). The bacterial Sm-like protein Hfq: a key player in RNA transactions. *Mol Microbiol* **51**, 1525-1533.

Zhang, H., Xing, L., Rossoll, W., Wichterle, H., Singer, R. H. and Bassell, G. J. (2006). Multiprotein complexes of the survival of motor neuron protein SMN with Gemins traffic to neuronal processes and growth cones of motor neurons. *J Neurosci* **26**, 8622-8632.

Zhang, H. L., Pan, F., Hong, D., Shenoy, S. M., Singer, R. H. and Bassell, G. J. (2003). Active transport of the survival motor neuron protein and the role of exon-7 in cytoplasmic localization. *J Neurosci* **23**, 6627-6637.

Zhang, R., So, B. R., Li, P., Yong, J., Glisovic, T., Wan, L. and Dreyfuss, G. (2011). Structure of a key intermediate of the SMN complex reveals Gemin2's crucial function in snRNP assembly. *Cell* **146**, 384-395.

Zhang, Z., Lotti, F., Dittmar, K., Younis, I., Wan, L., Kasim, M. and Dreyfuss, G. (2008). SMN deficiency causes tissue-specific perturbations in the repertoire of snRNAs and widespread defects in splicing. *Cell* **133**, 585-600.

Figure Legends

Figure 1

A) *Splicing snRNPs can be labeled in a time-dependent manner to identify newly-assembled versus mature complexes*. HeLa cell lines constitutively expressing YFP-SmB (i) and CFP-SmB (ii) fused using polyethylene glycol and fixed 1 hour after fusion. In the cell constitutively expressing CFP-SmB, newly imported snRNPs tagged with YFP-SmB localize to the Cajal body (arrow). The identity of the Cajal body is confirmed using antibodies to coilin and SMN (white signal in iii). (iv) is an enlarged view of the fused cells seen in A,B and C indicating points X and Y between which an intensity profile of YFP-SmB and CFP-SmB is shown in (v). The intensity profile demonstrates that newly imported (YFP-tagged) snRNPs are confined to the CB (green asterisk), despite the presence of steady-state snRNPs tagged with CFP-SmB in speckles (cyan asterisk) as well as CBs. Maximum intensity projections of 0.2μm sections through the sample. Bar=10μm.

B) *Design of the quantitative affinity isolation experiments comparing the interactomes of newly assembled (YFP-SmB, transient) and mature (YFP-SmB, stable) splicing snRNPs.*

Figure 2) *Quantitative comparison of the interactomes of mature vs. newly-made snRNPs.*

A) *Plot of log H:L ratio (enrichment above background with stably-expressed YFP-SmB) versus log M:L ratio (enrichment above background with transiently-expressed YFP-SmB).* Potential contaminants cluster around a log ratio 0 (SILAC ratio 1:1) whereas proteins enriched equally with both fall in the upper right quadrant. Arrows indicate enrichment of proteins preferentially with either mature or newly-made snRNPs. Specific protein classes are highlighted, as indicated in the legend.

B) Plot of log H:M ratio (relative enrichment with stably- vs. transiently-expressed YFP-SmB) versus relative abundance (summed peptide intensity normalized for protein MW). Median log H:M ratio is indicated by the red line. Dashed red lines are thresholds, indicating proteins enriched >1 log from the median. Values for the highlighted proteins, and a summary of their relative enrichment, are presented in Table I.

Table I) *Distribution of interactors between mature and newly-made snRNPs.* Annotated dataset highlighting specific classes of proteins identified as co-purifying with YFP-SmB in our quantitative affinity purification/mass spectrometry experiment. Log SILAC ratios demonstrating enrichment above background are shown for YFP-SmB pulled down

following either stable (log H:L) or transient (log M:L) expression. Direct comparison of relative enrichment with mature and newly-made snRNPs is determined by their log H:M ratios. In this experiment the median log H:M was -0.873. Proteins were considered to be enriched more with a particular class of snRNPs if their log H:M was >1 log away from the median. Also listed here are Uniprot accession numbers, Gene and Protein names, total number of peptides detected and number of unique peptides detected. The full dataset is provided as Supplemental Table S1.

Figure 3) *Highly mobile SmB-containing structures are seen in human neuroblastoma cells*. Time-lapse images of cells from line SHY5mCherrySmB undifferentiated and differentiated with retinoic acid followed by BDNF. The cells show cytoplasmic accumulations of SmB (arrows) that are more distinct in the differentiated cells. The blue tracks in the bottom panels show the trajectories followed by the arrowed structures in the 2-minute duration of the time-lapse sequence. Maximum intensity projections of 3 deconvolved 0.5μm z-sections. Bar=10μm. See also supplementary movie S1.

Figure 4) *Correlative fluorescence and electron microscopy identifies distinctive vesicles containing mCherry-SmB*. Confocal images showing mCherry-SmB vesicles in differentiated SH-SY5Y cells (A to C). Selected fluorescent structures adjacent to clear structural points of reference within the cell (arrows and arrowheads) were chosen for detailed ultrastructural analysis by TEM. The mCherry-SmB structures are approximately 50nm in diameter and have a vesicular appearance (arrows in D and E, arrowhead in F). In some instances, they are found adjacent to granular structures of similar size (asterisk in F). Bar=10μm in A to C; 200nm in D to F.

Figure 5) *mCherry-SmB interacts with Dynein and the coatomer complex in cytoplasmic vesicles*.

A) *Endogenous DYNHCI and γCOP co-immunoprecipitate with mCherry-SmB*. The left hand panel demonstrates the efficiency of pull-down of mCherry-SmB using the RFP-TRAP reagent with a single band of the appropriate size (59KDa) seen in the 'input' lane. A small amount of mCherry-SmB remains in the unbound fraction, with none detected in the beads control lane and a strong enrichment seen in the RFP-TRAP sample. Probing of duplicate blots with antibodies to DYNCHI and γCOP shows bands of the expected sizes (~450 and

98KDa respectively) in the RFP-TRAP sample, but not in the beads control.

B) *mCherry-SmB co-localizes with a sub-set of COPI vesicles*. Transient expression of mCherry-SmB (left hand panels and magenta on overlay) in SH-SY5Y cells constitutively expressing GFP-εCOP (center panels and green on overlay) demonstrates co-localization of the two proteins in cytoplasmic structures (arrows). The right hand panel shows an enlarged view of the section outlined in the top left panel. This co-localization persists throughout the imaging period of 2 minutes confirming that these are genuine mobile structures.

C) and D) *mCherry-SmB vesicles stain with the lipophilic dye BODIPY 488*. Undifferentiated (C) and differentiated (D) cells constitutively expressing mCherry-SmB (left hand panels and magenta on overlay) stained with the lipid dye Bodipy488 (centre panels and green on overlay). Both undifferentiated and differentiated cells show partial co-localization of the two signals (arrows). In the differentiated cells, it is apparent that the SmB-rich structures represent a sub-set of the structures stained with Bodipy488 with some stained structures not containing mCherry-SmB (arrowheads). B) and D) single planes of focus, C) maximum intensity projection of 3x0.5μm z-sections. Bar=10μm. See also supplementary movies S2-S4.

Figure 6) *SmB vesicles move on microtubules*.

A) *Nocodazole, but not Cytochalasin D, disrupts the mobility of mCherry-SmB vesicles*. Graphical representations of movements of mCherrySmB vesicles in representative groups of cells treated with cytochalisin D, nocodazole and untreated control cells. The velocity of the vesicles between time points is plotted against time, each coloured line represents the track taken by a single vesicle. Mean and maximum track velocities, plotted as mean +/-s.e.m, show no alteration with cytochalasin D treatment, but a significant decrease following nocodazole treatment (two way Anova test with Tukey post-test, n=75).

B) *mCherrySmB vesicles are enlarged and clustered in nocodazole treated cells*. Maximum intensity projection of 3x0.5 z-sections from a representative time-lapse sequence shows enlarged clustered vesicles (arrows). Bar=10μm.

C) *Cytochalasin D treatment disrupts the actin cytoskeleton in SH-SY5Y cells*. Staining with Phalloidin (green) confirms that the actin cytoskeleton is disrupted by the duration of cytochalasin D treatment used for the time-lapse studies (arrows). Nuclei are counterstained with DAPI (blue). Maximum intensity projection of deconvolved 0.2μm z-sections through the sample. Bar=10μm.

Figure 7) *mCherry-SmB vesicles contain SMN, gemin2 and SmD1 but not U5snRNP 100K, ASF/SF2, CRM1 or DCP1A*. Expression of GFP-SMN in cells from line SHY5-mCherry SmB demonstrates a high degree of co-localization between SmB (top and magenta in overlay) and SMN (centre and green in overlay) in vesicles (arrows) both near the cell body (A) and in neurites (B). The SMN complex protein, gemin2 (C) and the Sm protein SmD1 (D) expressed as GFP fusions also show some co-localization with SmB vesicles (arrows). GFP-U5snRNP 100K (E) and GFP-ASF/SF2 (F) do not accumulate in cytoplasmic vesicles, despite the presence of SmB vesicles in the cells (arrows). Both GFP-CRM1 (G) and GFP-DCP1A (H) show cytoplasmic accumulations (arrowheads). These do not, however, co-localise with the SmB vesicles (arrows). Single z-sections. Bar=10μm. See also supplementary figure S2 and supplementary movies S5 and S6. Percent of mCherry SmB vesicles also containing SMN is 98.48% s.e.m.+/-1.52; percent of mCherrySmB vesicles also containing GFP-SmD1 is 12.78% s.e.m.+/-3.94; percent of mCherrySmB vesicles also containing GFP-Gemin2 is 49.9% s.e.m.+/-9.79. These differences are statistically significant (Anova $P<0.05$, n=30).

Figure 8) *Over-expression of GFP-SMN results in enlarged vesicles with low mobility, while reduction of endogenous SmB results in loss of SMN-rich cytoplasmic structures*.

A) *Undifferentiated cell from line SHY5-mCherrySmB transiently expressing GFP-SMN*. mCherrySmB (i) and magenta in overlay is found in large cytoplasmic structures (arrows) co-localizing with GFP-SMN (ii and green on overlay). Single z-sections from a representative time-lapse sequence. Bar=10μm.

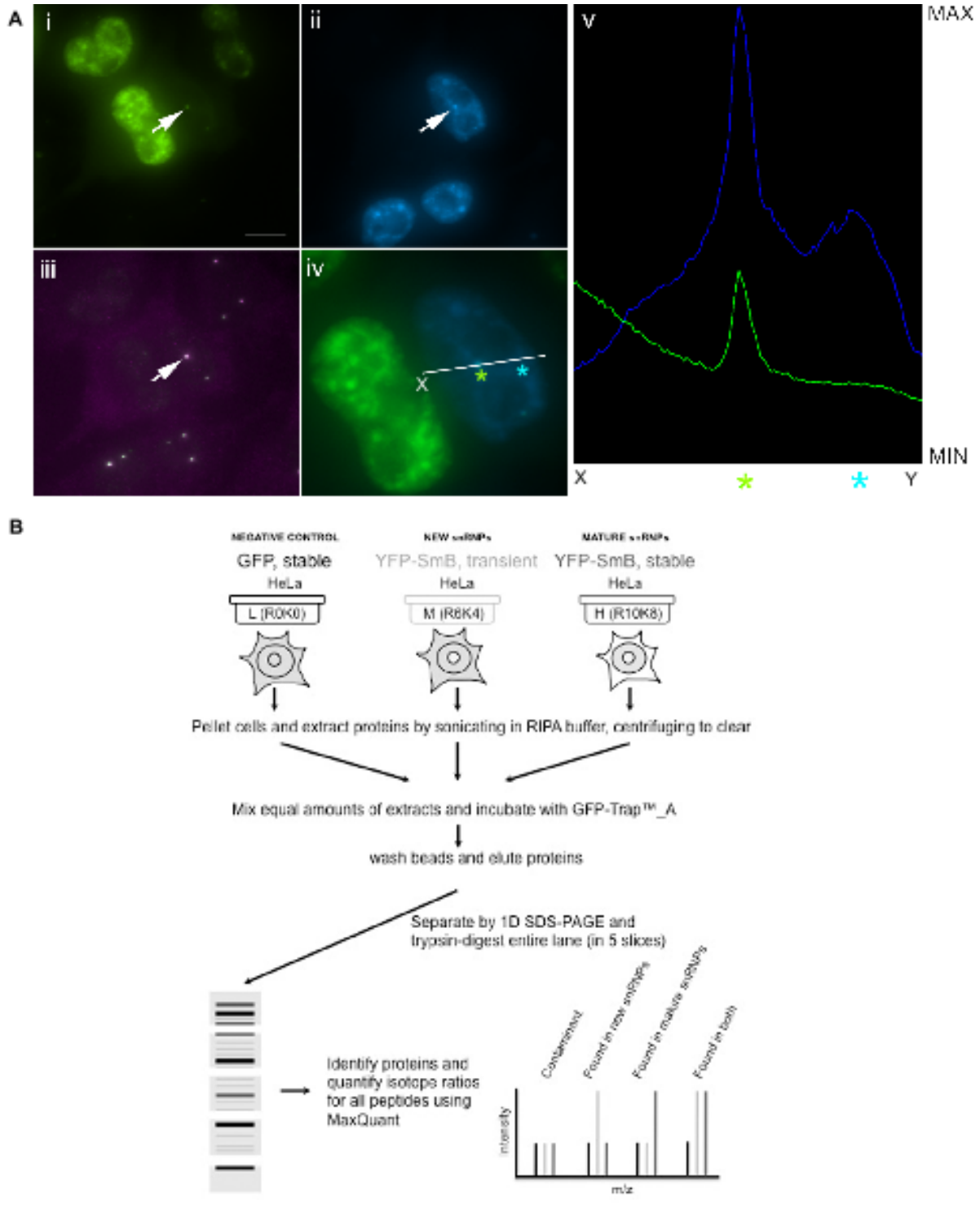
B) *Graphical representations of movements of mCherrySmB vesicles in cells expressing GFP-SMN and untransfected control cells*. The velocity of the vesicles between time points is plotted against time, each coloured line represents the track taken by a single vesicle. Total duration of the GFP-SMN-expressing time course is longer due to the time required for dual-color imaging at each time point.

C) *Both mean and maximum track velocities are significantly reduced by over-expression of GFPSMN*. Mean and maximum track velocities, plotted as mean+/- s.e.m, are significantly decreased by expression of GFP-SMN (unpaired Student's T test, n=89).

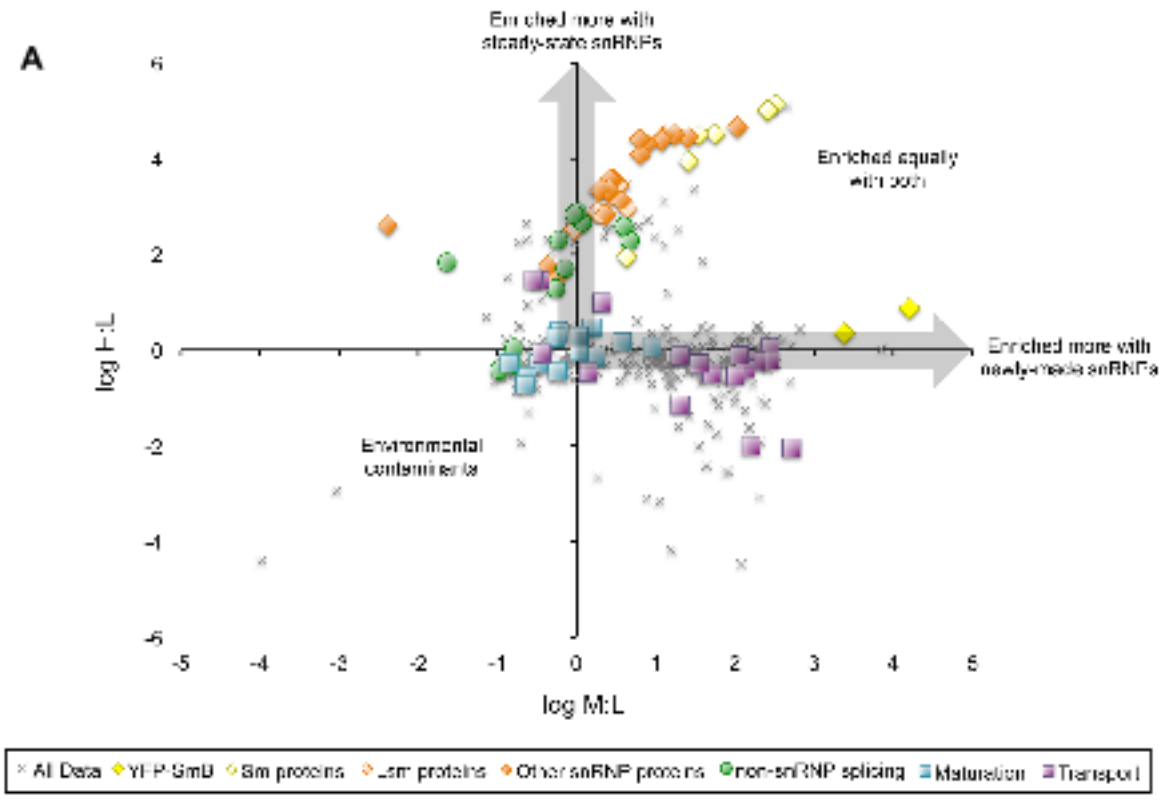
D) SmB vesicles persist in cells with reduced levels of SMN or PRMT5. Cells from line SHY5mCherry SmB transfected with siRNA sequences to reduce SMN (siSMN) or PRMT5

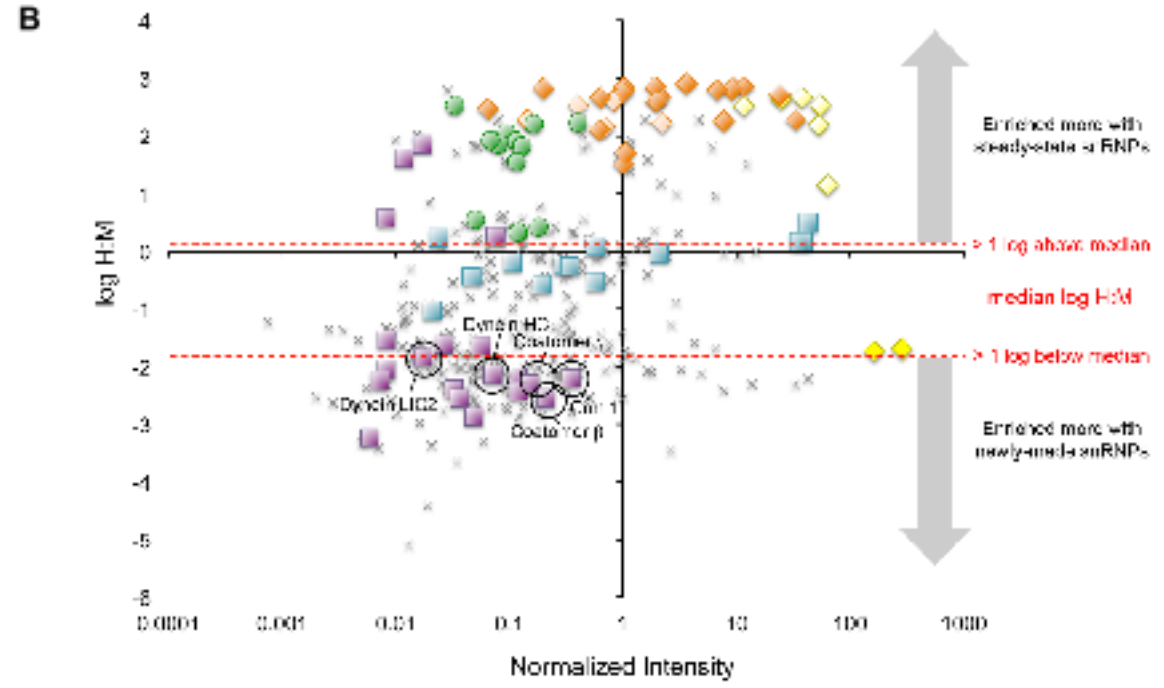
(siPRMT5) or non-targeting duplexes (-ve control) contain detectable SmB vesicles (arrows). The efficiency of transfection was extremely high (>90% of cells) as evidenced by the use of siGLO-labelled duplexes targeting cyclophillin D (siGLO, arrows).

E) Depletion of SmB results in loss of SMN-rich cytoplasmic structures. Cells from line SHY10 GFP-SMN transfected with siRNA sequences to reduce SmB expression (siSmB) contain large numbers of nuclear accumulations of GFP-SMN (arrows) but no cytoplasmic accumulations. Cells transfected with siRNA sequences targeting luciferase (-ve control) show the expected pattern of a small number of nuclear accumulations (arrows) and some cytoplasmic accumulations (arrowhead). Cells transfected with sequences to reduce SMN (siSMN) show no nuclear or cytoplasmic accumulations of SMN. In contrast to the result with siSmB, cells from the same line treated with leptomycin B (16hrs LMB) show a large number of cytoplasmic accumulations of GFP-SMN (arrowheads) when compared to control, untreated cells (no LMB), which contain a small number of nuclear accumulations (arrows) in addition to cytoplasmic signal. Bar=10μm.

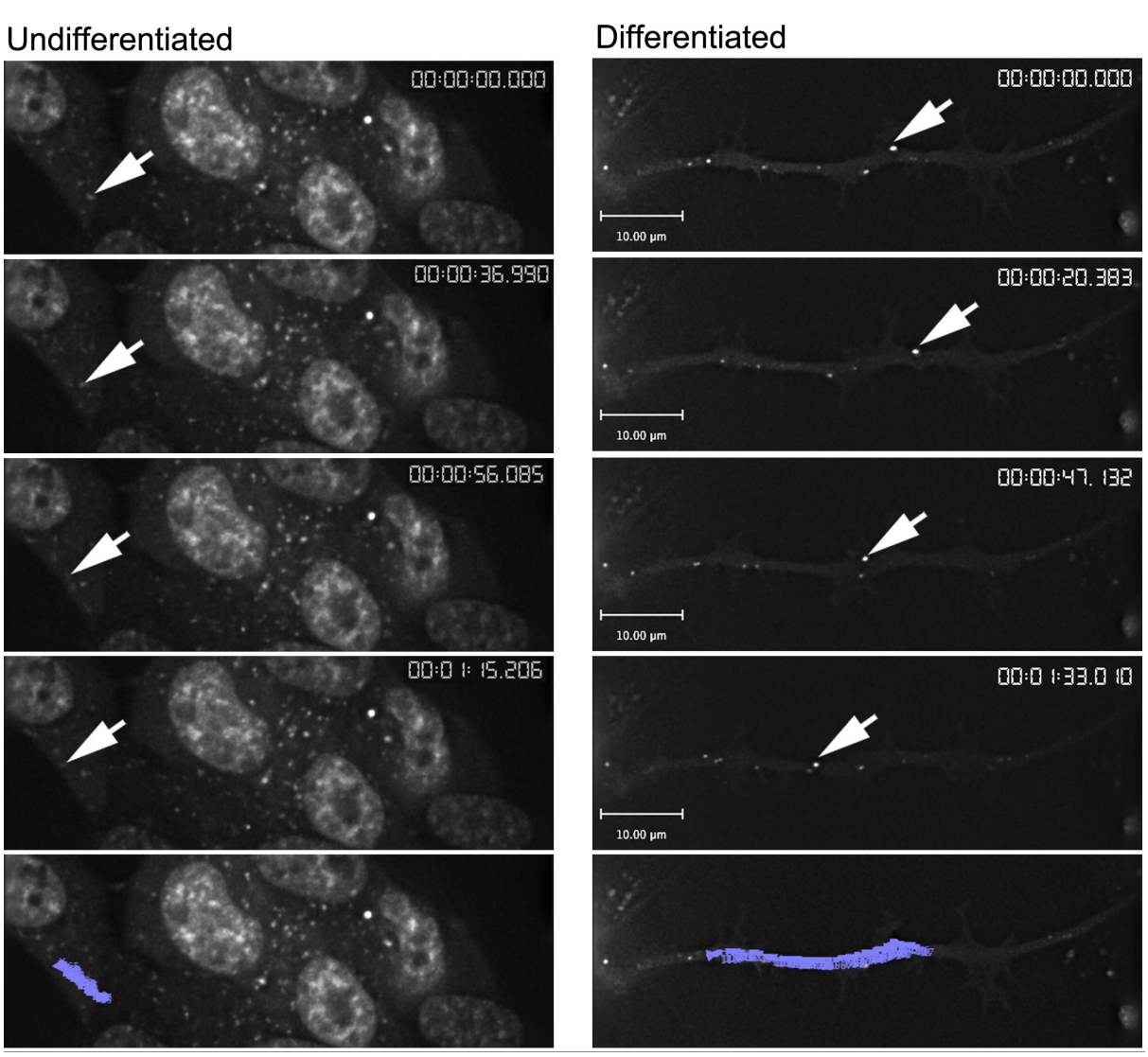

A
i
ii
iii
iv
v
MAX
MIN
X
Y
B
NEGATIVE CONTROL
GFP, stable
HeLa
L (R0K0)
NEW snRNPs
YFP-SmB, transient
HeLa
M (R6K4)
MATURE snRNPs
YFP-SmB, stable
HeLa
H (R10K8)
Pellet cells and extract proteins by sonicating in RIPA buffer, centrifuging to clear
Mix equal amounts of extracts and incubate with GFP-Trap™_A
wash beads and elute proteins
Separate by 1D SDS-PAGE and
trypsin-digest entire lane (in 5 slices)
Identify proteins and
quantify isotops ratios
for all peptides using
MaxQuant
Contaminant
Found in new snRNPs
Found in mature snRNPs
Found in both
intensity
m/z

1

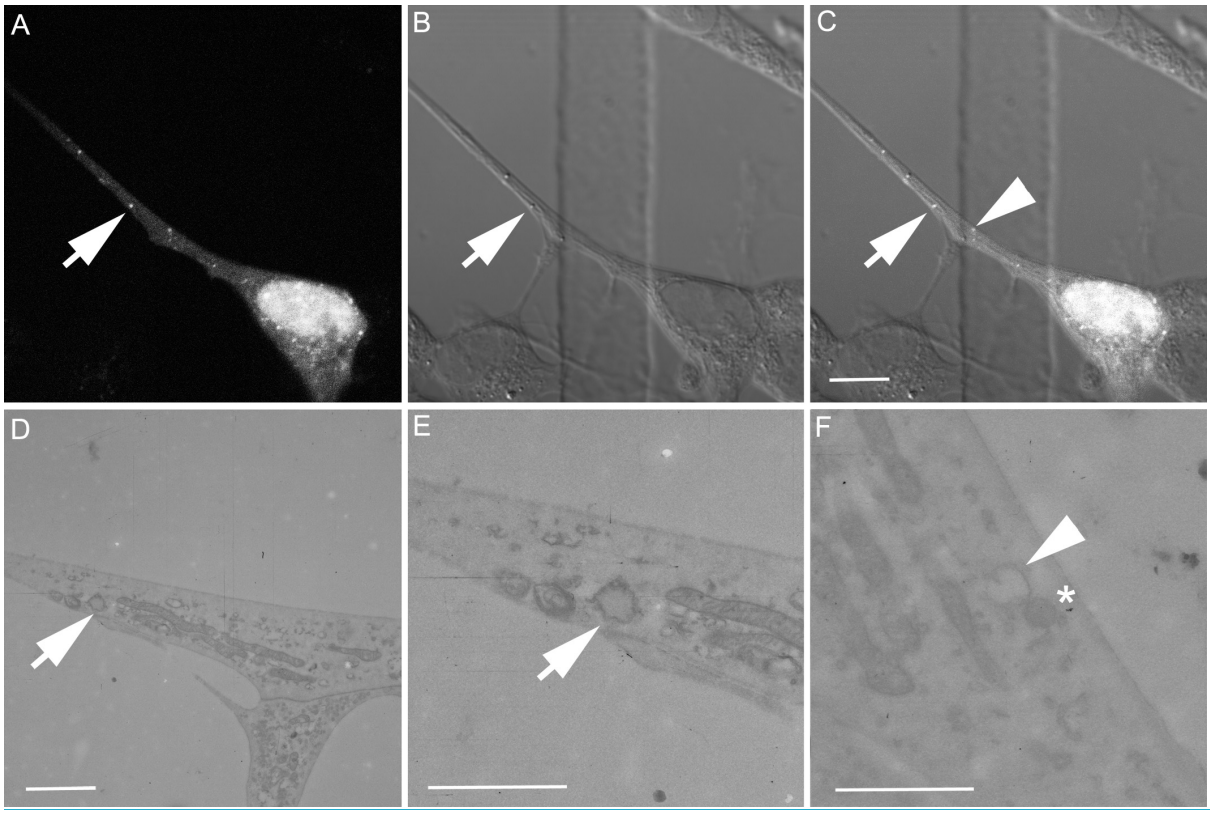

A
Enriched more with steady-state snRNPs
Enriched equally with both
Enriched more with newly-made snRNPs
Environmental contaminants
log H:L
log M:L
All Data
YF-P-SmB
Sm proteins
Lsm proteins
Other snRNP proteins
non-snRNP splicing
Maturation
Transport


B
Enriched more with steady-state snRNPs
> 1 log above median
median log H:M
> 1 log below median
Enriched more with newly-made snRNPs
Dynein HC
Dynein LIC2
Clathrin
Coatomer β
log H:M
Normalized Intensity

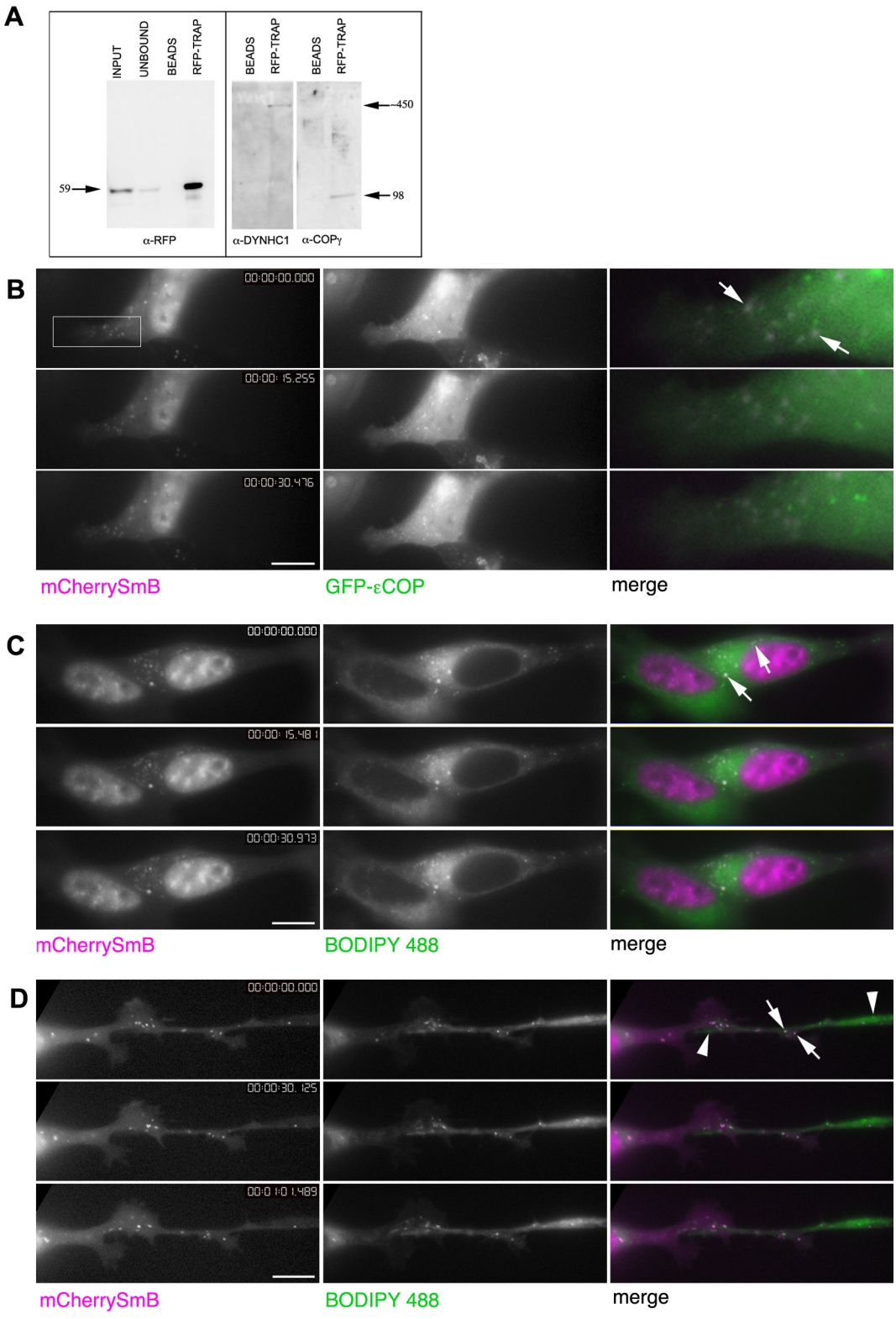
1


Undifferentiated
Differentiated
00:00:00.000
00:00:36.990
00:00:56.085
00:01:15.206
00:00:00.000
10.00 µm
00:00:20.383
10.00 µm
00:00:47.132
10.00 µm
00:01:33.010
10.00 µm

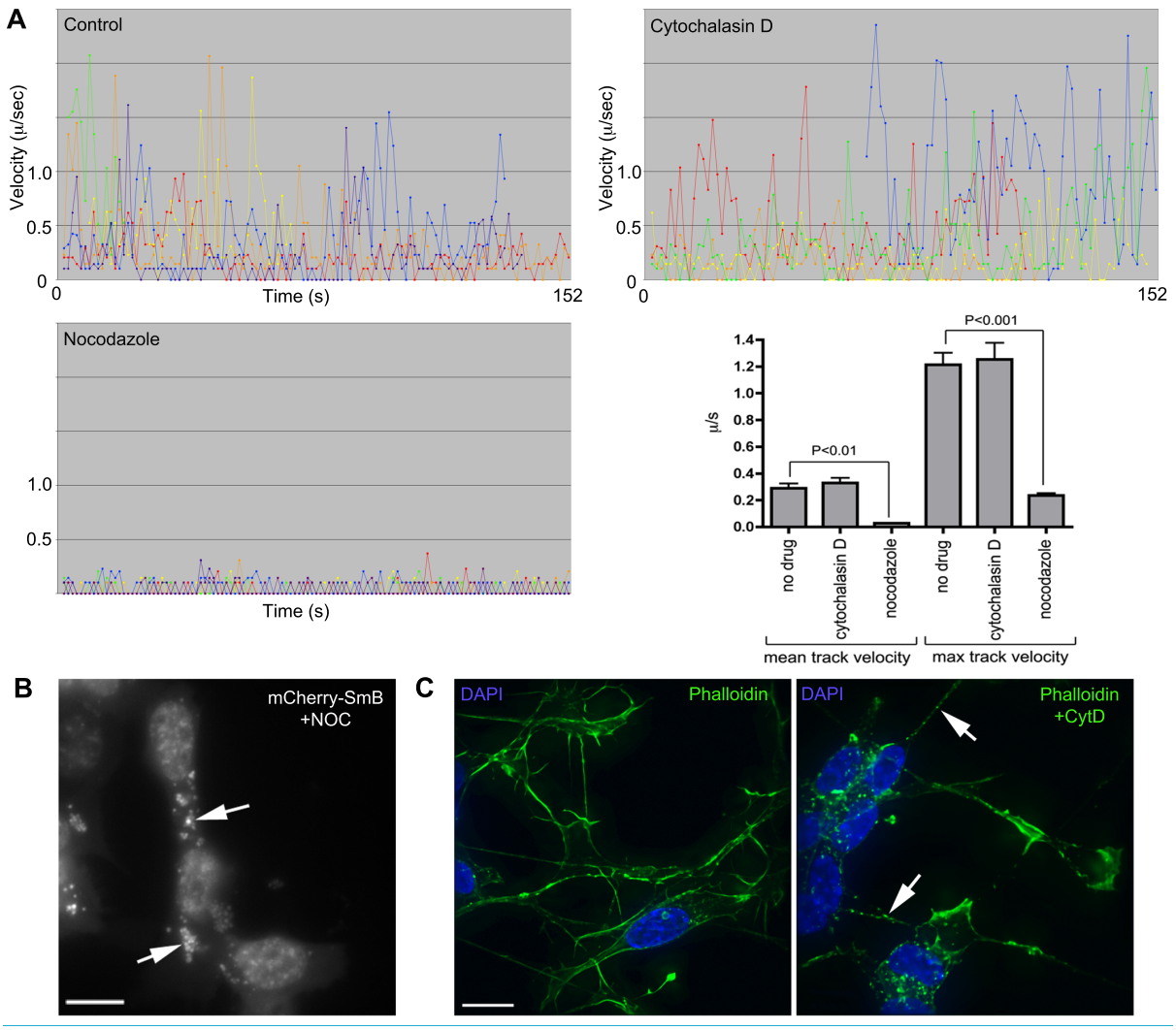
1



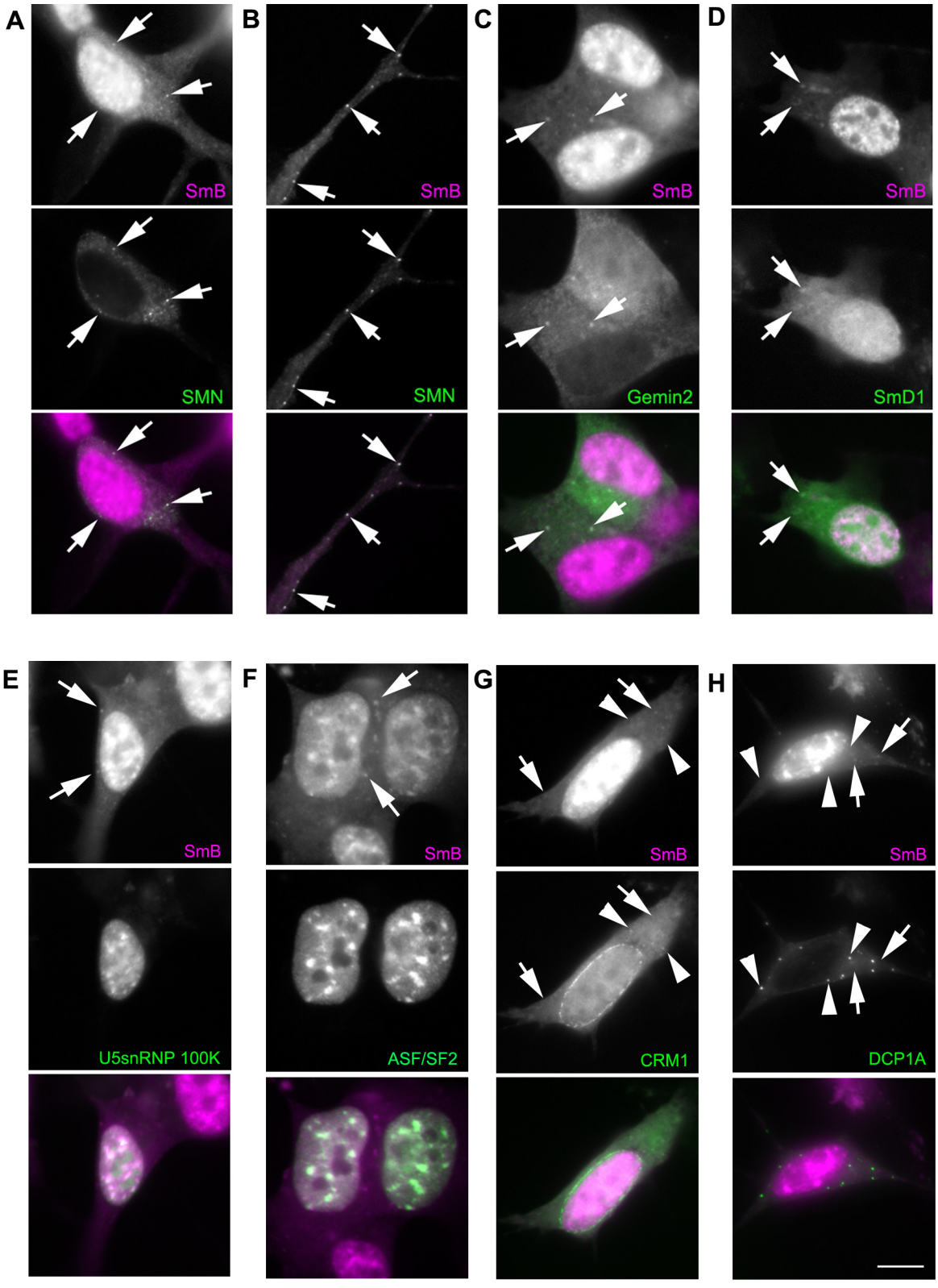
1


A
INPUT
UNBOUND
BEADS
RFP-TRAP
BEADS
RFP-TRAP
BEADS
RFP-TRAP
~450
98
59
α-RFP
α-DYNHC1
α-COPγ
B
00:00:00.000
00:00:15.255
00:00:30.476
mCherrySmB
GFP-εCOP
merge
C
00:00:00.000
00:00:15.481
00:00:30.973
mCherrySmB
BODIPY 488
merge
D
00:00:00.000
00:00:30.125
00:01:01.489
mCherrySmB
BODIPY 488
merge

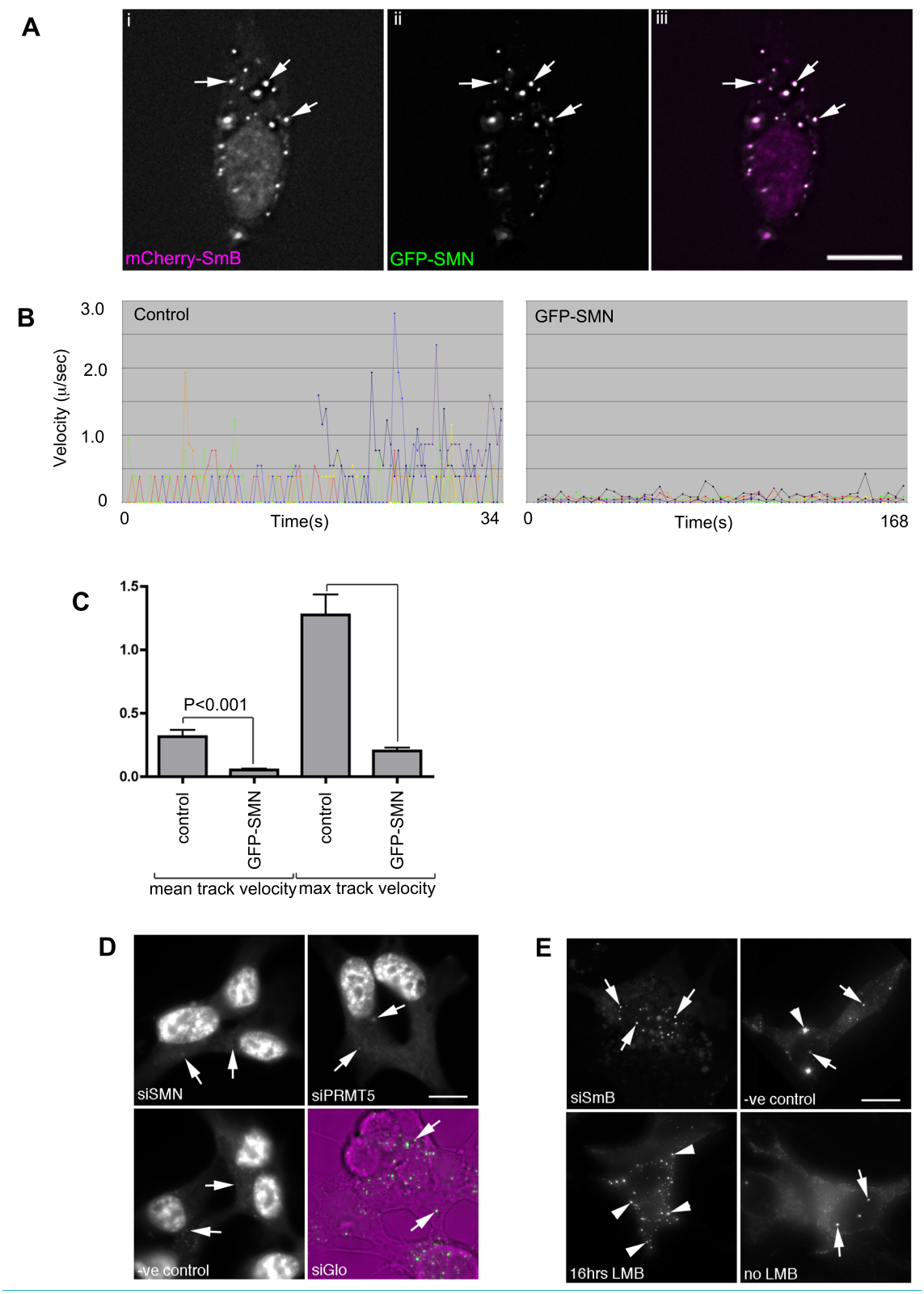
1


A
Control
Velocity (μ/sec)
1.0
0.5
0
0
Time (s)
152
Cytochalasin D
Velocity (μ/sec)
1.0
0.5
0
0
152
Nocodazole
1.0
0.5
Time (s)
P<0.001
P<0.01
μ/s
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
no drug
cytochalasin D
nocodazole
no drug
cytochalasin D
nocodazole
mean track velocity
max track velocity
B
mCherry-SmB
+NOC
C
DAPI
Phalloidin
DAPI
Phalloidin
+CytD

1


A
SmB
SMN
B
SmB
SMN
C
SmB
Gemin2
D
SmB
SmD1
E
SmB
U5snRNP 100K
F
SmB
ASF/SF2
G
SmB
CRM1
H
SmB
DCP1A

1


A
i
ii
iii
mCherry-SmB
GFP-SMN
B
Control
GFP-SMN
Velocity (μ/sec)
3.0
2.0
1.0
0
0
Time(s)
34
0
Time(s)
168
C
1.5
1.0
0.5
0.0
P<0.001
control
GFP-SMN
control
GFP-SMN
mean track velocity
max track velocity
D
siSMN
siPRMT5
-ve control
siGlo
E
siSmB
-ve control
16hrs LMB
no LMB