THE PLANT CELL

Functional nuclear retention of pre-mRNA involving Cajal bodies during meiotic prophase in European larch (*Larix decidua*)

Magda Rudzka [ID] ,[1,2] Patrycja Wróblewska-Ankiewicz,[1,2] Karolina Majewska [ID] ,[1,2] Malwina Hyjek-Składanowska,[1,2,‡] Marcin Gołębiewski,[2,3] Marcin Sikora [ID] ,[2] Dariusz Jan Smoliński[1,2,*,†] and Agnieszka Kołowerzo-Lubnau[1,2,*,†]

1   Department of Cellular and Molecular Biology, Nicolaus Copernicus University, Torun 87-100, Poland
2   Centre for Modern Interdisciplinary Technologies, Nicolaus Copernicus University, Torun 87-100, Poland
3   Department of Plant Physiology and Biotechnology, Nicolaus Copernicus University, Torun 87-100, Poland

*Authors for correspondence: darsmol@umk.pl (D.J.S); agakol@umk.pl (A.K.L.)
†Senior authors
‡Present address: Laboratory of Protein Structure, International Institute of Molecular and Cell Biology, Warsaw, 02-109, Poland.
M.R. performed triple labeling of the RNA and proteins, PCR, confocal microscopy, and image analysis. K.M. and P.W.A. performed triple labeling of the RNA and proteins, qPCR, confocal microscopy, and image analysis. M.S.H. participated in experimental design, proteome analysis, performed cDNA library preparation, and participated in writing the MS. M.G. performed bioinformatical analysis and participated in discussion and analysis of results and writing. M.S. performed cDNA library preparation, PCR. D.J.S. and A.K.L. designed experiments, performed triple labeling of the RNA and proteins, qPCR, sequencing, confocal microscopy and image analysis, and participated in discussion and analyses of the results and writing the article.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plcell) is: Dariusz Jan Smoliński (darsmol@umk.pl).

Abstract

Gene regulation ensures that the appropriate genes are expressed at the proper time. Nuclear retention of incompletely spliced or mature mRNAs is emerging as a novel, previously underappreciated layer of posttranscriptional regulation. Studies on this phenomenon indicated that it exerts a significant influence on the regulation of gene expression by regulating export and translation delay, which allows the synthesis of specific proteins in response to a stimulus or at strictly controlled time points, for example, during cell differentiation or development. Here, we show that transcription in microsporocytes of European larch (*Larix decidua*) occurs in a pulsatile manner during prophase of the first meiotic division. Transcriptional activity was then silenced after each pulse. However, the transcripts synthesized were not exported immediately to the cytoplasm but were retained in the nucleoplasm and Cajal bodies (CBs). In contrast to the nucleoplasm, we did not detect mature transcripts in CBs, which only stored nonfully spliced transcripts with retained introns. Notably, the retained introns were spliced at precisely defined times, and fully mature mRNAs were released into the cytoplasm for translation. As similar processes have been observed during spermatogenesis in animals, our results illustrate an evolutionarily conserved mechanism of gene expression regulation during generative cells development in Eukaryota.

Open Access

IN A NUTSHELL

Background: Reports in both animals and plants have recently revealed that a part of protein-coding transcripts might be retained in the nucleus for most of their lifetime. To date, nuclear speckles were the only domain associated with the accumulation of mRNA in animal cells. We previously showed that transcription occurs in a pulsatile manner in larch microsporocytes, during prophase of the first meiotic division. After each pulse, transcription is turned off, but the newly synthesized transcripts are not exported immediately to the cytoplasm, but are instead retained in the nucleoplasm and Cajal bodies (CBs). CBs are self-organized structures that appear when there is a need for increased metabolism of certain RNA types. CBs are evolutionarily conserved in both animal and plant cells, indicating their fundamental role in eukaryotic cells.

Question: Are protein-coding transcripts retained in the nucleus and CBs in their mature or immature form? Does mRNA retention regulate protein synthesis? Does nuclear retention contribute to gene expression during development?

Findings: Here, we show that only non-fully spliced mRNAs with retained introns are stored in CBs. Retained introns are spliced at precisely defined times, and fully mature mRNAs are released into the cytoplasm for translation. Retained introns are removed posttranscriptionally at the developmental stage when the encoded protein is needed. Nuclear retention of intron-containing pre-mRNAs is a functional regulatory mechanism that prevents premature translation of proteins in transcriptionally silent meiotic cells.

Next steps: Future research will focus on the precise molecular mechanisms by which CBs are involved in mRNA retention. What is the role of cis-elements (sequences in pre-mRNA) and trans-elements in pre-mRNA retention in CBs? These questions should also help us understand posttranscriptional splicing after intron-retaining pre-mRNAs leave CBs. We believe that our work will serve as a starting point for many other studies that will expand our understanding of the involvement of CBs in nuclear retention of incompletely spliced mRNAs as a mechanism of posttranscriptional regulation of gene expression in plants.

Introduction

The life of an mRNA molecule begins with transcription by RNA polymerase II (RNA Pol II), followed by splicing and processing of its 5′- and 3′-ends, which generally take place at the nuclear transcription sites; the processed mRNA is then translocated to the cytoplasm, where it is translated and degraded. mRNAs are thought to predominantly reside in the cytoplasm for most of their lifetime (Halpern et al., 2015). However, in many cell types, about 30% of polyadenylated transcripts have been shown to be retained in the nucleus and thus undetected in the cytoplasm (Herman et al., 1976; Wegener and Müller-McNicoll, 2018). Some of these transcripts represent long noncoding RNAs (lncRNAs), mRNA-like molecules that are spliced and polyadenylated like true mRNAs and perform various regulatory functions (Inagaki et al., 2005). In recent years, reports in yeasts, animals, and plants have determined that protein-coding transcripts might be retained in the cell nucleus for most of their lifetime (e.g. Averbeck et al., 2005; Boothby et al. 2013; Halpern et al., 2015). Since mRNAs are transcribed and processed at the transcription sites and translated in the cytoplasm, this lengthy retention period points at the possible overlooked role of the cell nucleus (Halpern et al., 2015). Studies on nuclear retention of mRNAs have revealed the significant influence exerted on the regulation of gene expression by regulating mRNA export and delaying translation, which allows the synthesis of specific proteins in response to a stimulus (e.g. under stress conditions) or at strictly controlled times. Regulation of gene expression by nuclear retention of mRNAs has been demonstrated, among others, in mammalian metabolic tissues (Halpern et al., 2015), in generative cells (Averbeck et al., 2005; Boothby et al., 2013), or in cells under stress conditions (Prasanth et al., 2005; Ninomiya et al., 2011). Although the final result of such regulation is always the same (i.e. delayed translation), the transcripts can be retained in different forms. A regulatory mechanism of gene expression based on the retention of non-fully mature mRNAs was observed in generative cells. Recent studies on male cells of the fern hairy water clover (*Marsilea vestita*) showed that retention and subsequent removal of introns from pre-mRNAs regulates translation pattern during posttranscriptionally regulated spermatogenesis (Boothby et al., 2013). Sequencing analyses revealed that only one intron is usually left during mRNA splicing with a length generally not exceeding 100 nucleotides (nt). This intron is then removed posttranscriptionally when a cell is at the specific developmental stage that demands the encoded protein. This process is a functional mechanism that prevents premature translation of proteins required during specific stages of the development of transcriptionally silent gametes.

In mammalian metabolic tissues (beta cells, liver, and gut), a wide range of protein-coding genes was identified whose levels of spliced polyadenylated mature mRNAs were higher in the nucleus than in the cytoplasm (Halpern et al., 2015). These included genes such as the transcription factor

Carbohydrate response element binding protein (ChREBP), NLR Family Pyrin Domain Containing 6 (Nlrp6), glucokinase, and glucagon receptor. A single-molecule in situ technique used to quantify nuclear mRNA lifetimes revealed that such transcripts can spend hours in the nucleus before being exported to the cytoplasm. Mammalian genes are transcribed in bursts, leading to temporal fluctuations in cellular mRNA levels and variability among identical cells. Nuclear retention can effectively buffer these fluctuations, facilitating lower variability in cytoplasmic mRNA levels. A mechanism of mRNA nuclear retention was described in mouse (*Mus musculus*) liver cells, involving the protein-coding gene m̲ouse c̲ationic a̲mino acid t̲ransporter 2̲ (mCAT2) and the lncRNA *CAT2*-transcribed nuclear RNA (CTN-RNA; Prasanth et al., 2005). The mCAT2 protein is a cell-surface receptor involved in the cellular uptake of L-arginine, a precursor for the biosynthesis of nitric oxide (NO). The NO pathway is induced in cells experiencing various stress conditions, including viral infection and wound healing, as a part of the cellular defense mechanism. CTN-RNA is transcribed from the same locus as its protein-coding counterpart (*mCAT2*) but from an alternative promoter and transcription termination site, resulting in unique untranslated regions (UTRs) at both ends of the transcripts. Upon stress, the CTN-RNA is posttranscriptionally cleaved in its 3′-UTR, and the released mRNA, whose sequence is identical to the *mCAT2* mRNA, is transported to the cytoplasm and translated into mCAT2. This mechanism provides a rapid response, in the form of mCAT2 protein production, under stress conditions, to modulate L-arginine uptake for the NO pathway.

Nevertheless, there is little information about which mechanisms lead to retention of nuclear mRNAs and what causes their release to the cytoplasm. Splicing factors have been implicated in the retention of pre-mRNAs, especially those factors forming the early spliceosome complex (E), that is, U1 small nuclear RNA (snRNA) and U2AF (Takemura et al., 2011). The binding of these factors to mRNAs was shown to prevent export through nuclear pores (Fasken and Corbett, 2009). Another hypothesis proposes that nuclear retention factors bind the mRNAs, thus preventing them from binding the export factors. These factors are expected to anchor mRNAs in nuclear structures (Takemura et al., 2011).

Nuclear speckles (NSs) were thought to be the only cellular structure involved in the accumulation of polyadenylated (poly(A)$^+$) transcripts in animal cells. However, those transcripts accumulating in the speckles did not appear to be transported to the cytoplasm (Huang et al., 1994). In larch microsporocytes during diplotene, which lasts about 5 months in this species, large quantities of poly(A)$^+$ RNA are synthesized (Kołowerzo-Lubnau et al., 2015). We previously observed that poly(A)$^+$ RNAs accumulate in spherical structures within the nucleoplasm in mid-diplotene. We also demonstrated, using in situ hybridization at the ultrastructural level as well as double labeling of poly(A)$^+$ RNA and known markers of plant Cajal bodies (CBs), that CBs are

sites of poly(A)$^+$ RNA accumulation different than distributed localization of the poly (A)$^+$ RNA in the speckles (Kołowerzo et al., 2009). We also showed that the transcripts accumulating in CBs indeed encode proteins; that is, they were mRNAs. These mRNAs included transcripts for housekeeping genes, for example, mRNAs encoding pectin methyltransferase, ATP synthase, catalase, peroxidase, and α-tubulin. In addition to these transcripts, we also localized transcripts encoding proteins commonly detected in CBs, for example, splicing factors Sm antigen polypeptides SmD1, SmD2, and SmE, as well as mRNAs encoding the RNA polymerase II subunits RPB2 and RPB10 (Smoliński and Kołowerzo, 2012).

CBs are evolutionarily conserved structures present in both animal (Morris, 2008) and plant cells (Shaw and Brown, 2004), reflecting their fundamental role in eukaryotic cells. Intensive research, carried out for many years, has demonstrated the relationship between CBs and many processes associated with the metabolism of various types of RNA, mainly snRNA (Frey and Matera, 2001; Kiss, 2004; Stanek and Neugebauer, 2006) and small nucleolar RNA(Lamond and Spector, 2003). CBs are self-organized structures (Misteli, 2007) that form when a need arises for increased metabolism of certain types of RNA. Although assembly of spliceosomal subunits is possible in the nucleoplasm, the snRNPs (small nuclear ribonucleoprotein) assembly rate in CBs is around 10-fold greater than in the nucleoplasm (Klingauf et al., 2006; Novotný et al., 2011). In contrast, fibroblast cells, which have a low metabolism that does not require intensive assembly of snRNPs, do not have CBs. Here, the assembly of snRNPs takes place without CBs, and is therefore many times less efficient. A knockdown for *Coilin* (encoding a marker protein for these CBs) in cells with a high metabolism (generative cells or embryonic cells) resulted in the breakdown of CBs and led to cell death (Strzelecka et al., 2010).

Genes in larch microsporocytes are transcribed in bursts. Five such bursts were observed during diplotene. The diplotene stage is a period of intensive microsporocyte growth associated with the synthesis and accumulation of different RNAs and proteins. During this long period, microsporocytes increase their volume over three-fold. Larch microsporocytes display changes in chromatin morphology during this stage, alternating between transcriptionally silenced stages of chromatin condensation (contraction) and diffusion (relaxation) stages. We previously demonstrated that a large quantity of poly(A)$^+$ RNA in the nucleus of microsporocytes appears at the start of transcription. These transcripts were retained in the nucleus for a long time and were stored within the nucleoplasm and CBs. A pool of transcripts within the CBs gradually increased and reached a maximum at the stage preceding the release of large quantities of poly(A)$^+$ RNAs to the cytoplasm. CBs are the compartment in which poly(A)$^+$ transcripts are retained for a very long period (on the order of days). We observed that a nuclear pool of these transcripts is stored mainly in CBs during the stages that

end the flow of poly(A)$^+$ RNA cycle (Kołowerzo-Lubnau et al., 2015). CBs therefore appear to exert a significant effect on the nuclear retention of mRNAs, as well as on the subsequent export of these transcripts to the cytoplasm.

During the first three transcriptional bursts, mRNA is intensively synthesized. Large amounts of RNA Pol II and high levels of snRNPs accumulate in the nucleus (Kołowerzo-Lubnau et al., 2015). In the late diffusion stages, as during the fourth transcriptional burst and the poly(A)$^+$ cycle presented in this article, the synthesized mRNA is not directly subjected to translation but is instead retained in the nucleus, only later being transported to the cytoplasm for translation. In the fifth (and last) diffusion cycle, the levels of poly(A)$^+$ RNA increase slightly again, although splicing factors are still abundant. The mRNA synthesized in early stages is used during the diplotene stage and is not transmitted to dyad or tetrads. In contrast, mRNA synthesized in late diplotene stages that encode splicing factors and other essential enzymes and proteins, as well snRNAs and ribosomal ribonucleoprotein, accumulate and are most likely transmitted to the dyad and tetrads, and are used after intense transcription resumption. Based on our previous research and the results presented in this manuscript, we propose that pre-mRNA retention is used by meiotic cells for posttranscriptional regulation of expression via delayed mass translation at a certain point in the development of these generative cells. In addition, we show here that this mechanism is used to synthesize proteins "for later needs" necessary to develop four daughter cells (tetrads) when their genome is still inactive after meiotic divisions.

We selected transcripts encoding Sm proteins for detailed analyses, as they are retained in the nucleoplasm and CBs (Smoliński and Kołowerzo, 2012). To examine which mRNAs are transported to the cytoplasm and potentially translated after their nuclear retention, we performed a transcriptome analysis of cytoplasmic mRNAs. Among these transcripts, we identified a large group of mRNAs encoding proteins associated with transcription and posttranscriptional modifications, including mRNA encoding Sm proteins. We examined their distribution during a single cycle of cellular synthesis and turnover of poly(A)$^+$ RNA, and checked what form they are retained as in the nucleus. We also checked whether these transcripts are released into the cytoplasm after their retention period and whether the increase in these mRNAs in the cytoplasm correlates with the increase in the abundance of proteins they encode, which would suggest that nuclear retention is an element of regulation of expression of these mRNAs.

Results

Global pattern of poly (A)$^+$ RNA retention in the nucleus and CBs after pulse transcription activity in middle diplotene

We previously showed that larch microsporocytes undergo five bursts of de novo transcription during diplotene that corresponds to five cycles of increased levels of poly(A)$^+$

RNAs in the cell (Kołowerzo-Lubnau et al., 2015). Notably, the colocalization of nuclear poly(A)$^+$ RNAs with newly formed transcripts was relatively low in mid-diplotene (21% of the total poly(A)$^+$ RNAs pool), suggesting the occurrence of nuclear poly(A)$^+$ RNA retention during this period. The observation led us to perform a more detailed analysis of mRNA metabolism in mid-diplotene.

To explore RNA nuclear retention in these cells, we performed spatial and quantitative analysis of a single cycle of cellular synthesis and turnover of poly(A)$^+$ RNAs. The characteristic changes in poly(A)$^+$ RNA distribution allowed us to distinguish five stages in this cycle (Figure 1). Due to the extraordinarily long duration time in larch cells during diplotene and their high synchronicity within flower buds growing on the same branch (>90%; Majewska et al., 2021), it was possible to observe and analyze the sequence of stages of the poly(A)$^+$ RNA life cycle, which was essential for reverse transcription-quantitative PCR (RT-qPCR) and transcriptomic analysis. Previous studies showed that a single cycle lasts 10–11 days, with the first stage (Stage 1) lasting the longest (8–9 days). Stages 2–5 are much shorter and only last several hours each (Majewska et al., 2021). We used the U2 snRNA splicing factor (a CB marker) as an additional indicator to define the stages during the cycle (Figures 1, A and 2).

The abundance of U2 snRNA gradually increased in the nucleus area from Stages 1–5, especially in CBs, as shown in Figure 2. We previously showed that snRNAs are simultaneously synthesized with Sm proteins during the cycle (Hyjek et al., 2015). snRNAs are constituents of the spliceosome and are necessary together with Sm proteins for pre-mRNA splicing during the fifth transcription cycle. At the beginning of the cycle (that is during Stages 1 and 2), the levels of poly(A)$^+$ RNAs were significantly higher in the nucleus compared to the cytoplasm (Figure 1, A and B). The nuclear pool of poly(A)$^+$ RNAs was partially dispersed but also localized within numerous spherical structures corresponding to CBs (Figures 1, A and 2; Supplemental Figure S1). During Stage 1, the levels of poly(A)$^+$ RNAs were the highest relative to all other stages (Figure 1B). The levels of poly(A)$^+$ RNAs also correlated with a high signal intensity from the hyperphosphorylated form of RNA Pol II (Pol II O; Figure 1A), but saw their levels decrease in subsequent stages of the cycle (Supplemental Figure S2). This result indicated that the main burst of poly(A)$^+$ RNA synthesis takes place de novo during the first and longest stage. Starting at Stage 2, the nuclear pool of poly(A)$^+$ RNAs progressively decreased, with a concomitant increase in the cytoplasmic pool of poly(A)$^+$ RNAs (Figure 1, A and B). While the distribution pattern of poly(A)$^+$ RNA was similar to that seen in Stage 1 (Figure 1A; Supplemental Figure S3A); the proportion of poly(A)$^+$ RNAs in the cytoplasm and CBs changed. Indeed, the abundance of poly(A)$^+$ RNAs located within CBs increased constantly throughout the cycle (Figures 1, C and 2). Poly(A)$^+$ RNAs accumulated within the perinuclear region in a characteristic pattern, which we called endoplasmic reticulum-like (Figure 1A).
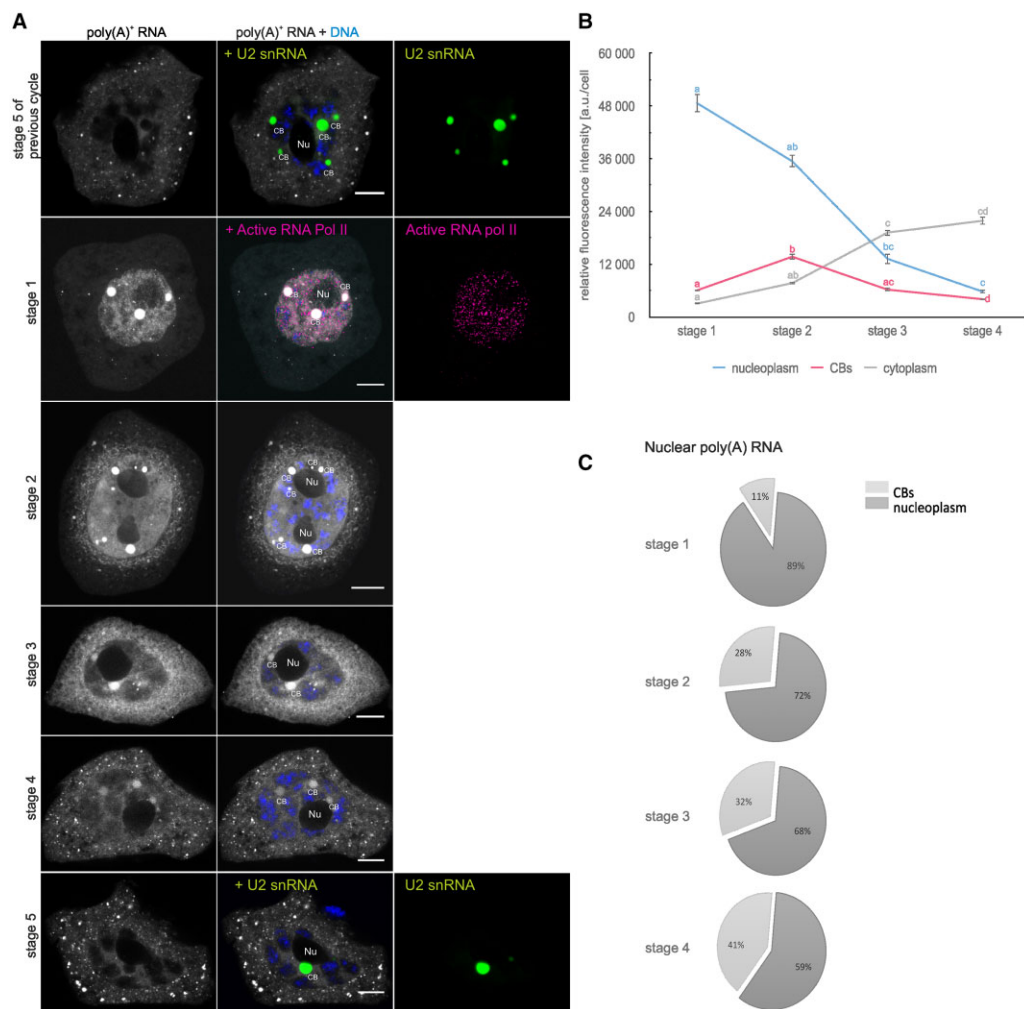
Figure 1 The cycle of cellular synthesis and turnover of poly(A) RNA. A, Detailed description is provided in the text; Labeling of U2 snRNAs was used to show CBs at Stage 5 when poly(A) RNAs do not accumulate in them. Labeling of RNA polymerase II in active form (phosphorylation of serine 2 in the CTD domain of the polymerase) was used to indicate the stage with the highest transcriptional activity. Nu, nucleolus. Bar = 10 μm. B, Levels of poly(A)$^+$ RNA in individual cell compartments (nucleoplasm, CBs, and cytoplasm) during a single cycle. Error bars indicate standard error of the mean (SEM). Different lowercase letters indicate Significant differences (at α = 0.05), according to Kruskal–Wallis test with Dunn's post hoc procedure using Benjamini–Hochberg FDR correction. n = 10 for each stage. C, Distribution of poly(A) + RNA nuclear fraction in the nucleoplasm or CBs fraction during a single cycle.

During Stage 3, we observed a dramatic drop in the nucleoplasmic pool of poly(A)$^+$ RNAs, which was accompanied by increased signal in the cytoplasm (Figure 1B) but as a more dispersed pattern compared to Stage 2 (Figure 1A; Supplemental Figure S3B). The signal intensity for poly (A)$^+$ RNAs in CBs decreased much more slowly (Figures 1, B and 2), as indicated by the ratio between the abundance of these transcripts within CBs over that of the total nuclear pool, which only exhibited up to a 5% increase compared to Stage 2, and up to a 22% increase relative to Stage 1 (Figure 1C). Microscopy observations revealed numerous and intensely stained CBs at Stage 3 (Figures 1A and 2). Stage 4 was characterized by a further decrease of the nucleoplasmic levels of poly(A)$^+$ RNAs (Figure 1B), whereas a significant portion (41% of the nuclear pool) of transcripts still accumulated in numerous CBs (Figures 1, A and C and 2; Supplemental Figure S4), which correlated with the higher

poly(A)$^+$ RNA levels within the cytoplasm (Figure 1B). We observed numerous bright foci, scattered over the entire cytoplasmic area (Figure 1A). Stage 5 was short-lived, being a quick transition between successive cycles, thus only allowing the observation of a few cells. During Stage 5, poly(A)$^+$ RNAs reached their lowest levels in the nucleoplasm of the entire cycle. These poly (A)$^+$ RNAs did not accumulate in CBs at all (Figure 1A; Supplemental Figure S4). Due to the very small number of cells at this stage, we did not perform a quantitative analysis.

The data presented above indicated that the temporal nuclear retention of RNA takes place in larch microsporocytes. The high levels of poly(A)$^+$ RNAs in the nucleus remained throughout the majority of the cycle, long after de novo transcription was turned off. These results confirm that a substantial portion of newly transcribed poly(A)$^+$ RNAs is not subject to immediate cytoplasmic export but is instead
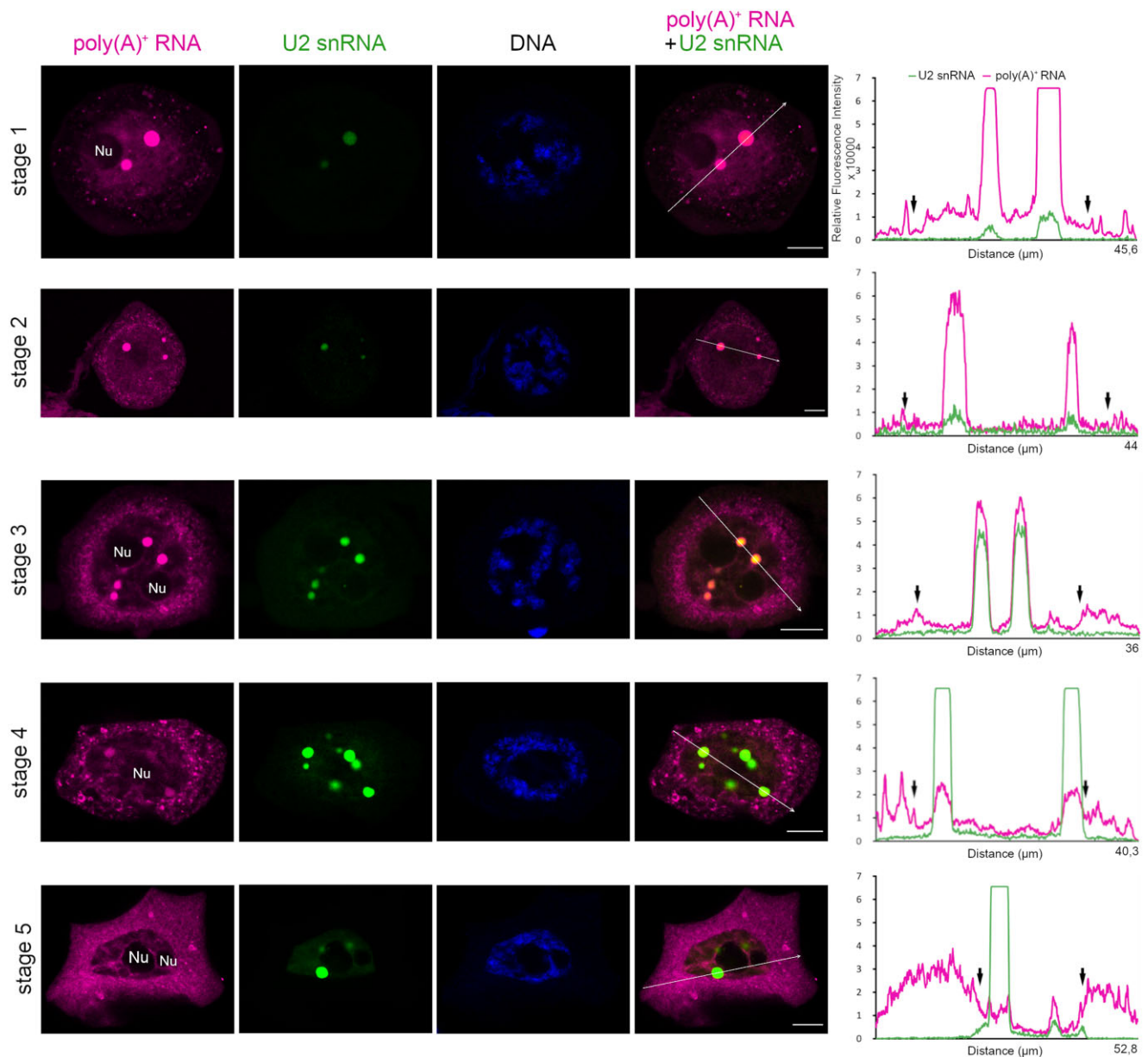
Figure 2 Multilabeling shows the levels and localization of U2 snRNA and poly(A)$^+$ RNA during the cycle. Left, Changes in the localization and amount of poly(A)$^+$ RNA and U2 snRNA, at each stage of the cycle. Right, fluorescence intensity for poly(A)$^+$ mRNA and U2 snRNA along the white line. Vertical black arrows indicate the edges of the nucleoplasm. Nu, nucleolus. Bar = 10 μm.

temporarily retained within the nucleus. Interestingly, the nuclear domains involved in this process were most likely CBs.

Although the nuclear pool of RNA gradually decreased from Stage 2 onward, the ratio of transcript abundance within CBs over that of the total nuclear pool increased, which pointed to CBs as the sites of RNA nuclear retention.

Cytoplasmic transcriptome analysis of microsporocytes in middle diplotene

The concentration of poly(A)$^+$ RNAs in CBs was 70–200 times higher than that in the rest of the nucleus. These small structures accumulated from 11% of the poly(A)$^+$ RNA pool in the nucleus at Stage 1 and up to 40% during

Stage 4 (Figures 1, C and 2), even though CBs only constituted 0.5%–1.5% of the volume of the nucleus. Several recent reports have focused on nuclear retention in both animal and plant cells and determined that transcripts are retained in various forms of maturity. The currently accepted rationale for the nuclear retention of mRNAs is as a means to control the time of mRNA export to the cytoplasm, thus delaying translation (Boothby et al., 2013; Halpern et al., 2015; Naro et al., 2017). To identify those mRNAs that are transported to the cytoplasm and potentially translated during a single cycle of cellular synthesis and turnover of poly(A)$^+$ RNA, we sequenced the cytoplasmic pool of transcripts of larch microsporocytes and determined their corresponding cytoplasmic proteome. We de novo

assembled the reads into Unigenes, which we then filtered for quality via PRINSEQ-LITE (Schmieder and Edwards, 2011). The reads from all libraries were pooled and assembled de novo using TRINITY version 2.8.2 with default settings. We then subjected the Unigenes to BLASTX (for potential mRNA sequences) and BLASTN (for other sequences; using the National Center for Biotechnology Information (NCBI) database for green plants), which resulted in the identification of 7,001 independent transcripts in the cytoplasm (Supplemental Figure S5A). We also conducted a functional annotation of these transcripts by gene ontology (GO) analysis, which returned 3,283 sequences (44%) assigned to 307,714 GO terms. The cytoplasmic transcripts detected in larch microsporocytes are listed in Supplemental Data Set 1. Raw reads generated during this project were deposited in the NCBI SRA database and are accessible through BioProject no. SAMN14564474.

Functional analysis showed that cytoplasmic transcripts include five major mRNA groups encoding proteins associated with: (1) transcription and posttranscriptional modifications; (2) ribosomes, translational and posttranslational modification of proteins; (3) mitochondria and energy transformations; (4) photosynthesis and plastids; and (5) the cytoskeleton and the cell wall (Supplemental Figure S5B). The most numerous group was transcripts encoding proteins associated with transcription and posttranscriptional modifications, among which were transcripts coding for the Sm proteins SmD1, SmD2, and SmG. Sm are nuclear proteins that are core components of spliceosomal ribonucleoproteins (UsnRNPs) and are thus involved in pre-mRNA splicing; notably, they localize to CBs in plants. We had previously shown that these specific mRNAs are retained in the nucleus and accumulate in CBs of diplotene microsporocytes in larch (Smoliński and Kołowerzo, 2012), prompting us to select these mRNAs for further analysis. We attempted to assess the form (mature or immature) in which the *Sm* transcripts were retained in the nucleus. Since we sequenced mature transcripts and because the larch genome is not fully sequenced, we resorted to designing primers flanking the locations of putative introns but anchored on exons, based on our larch RNA-seq dataset and the genomes of related species, mainly white spruce (*Picea glauca*) and Norway spruce (*Picea abies*). We sequenced the resulting PCR products and designed probes for fluorescent in situ hybridization (FISH).

FISH analysis showed that all analyzed *Sm* transcripts are retained in the microsporocytes nucleus in an immature form (Figure 3). In the case of the *SmD1* pre-mRNA, consisting of three exons and two introns, we detected strong signals in the nucleoplasm as well as in CBs for probes whose sequences were complementary to the first or second intron, respectively (Figure 3A). In contrast, for the pre-mRNA coding for the SmD2 and SmG proteins, which contained four exons and three introns, a single probe per transcript corresponding to one intron showed accumulation in the nucleoplasm and CBs (Figure 3, B and C). The signal derived from the probes specific for the other two introns was either very weak or limited only to a very small area in the nucleus, indicating that these introns may be quickly spliced out. These intron-containing transcripts, therefore, accumulate in the nucleus. The presence of unspliced introns, the so-called retained introns, determines the retention of such transcripts in the nucleus. Detailed analysis using multi-indirect FISH showed that the retention of immature forms of various *Sm* mRNAs coincides in time and place. CBs were the principal places of their concentration, especially the *SmG* mRNA (Figure 4; Supplemental Figure S6).

Participation of CBs in the retention of nonfully spliced pre-mRNA transcripts: Detailed analysis of *SmG* transcripts

In further studies, we conducted a more detailed analysis of the transcript encoding the SmG protein. The *SmG* pre-mRNA consisted of four exons and three introns, with the first intron being retained. Indeed, control polymerase chain reaction (PCR) and reverse transcriptase polymerase chain reaction (RT-PCR) reactions showed that the second intron likely undergoes co-transcriptional splicing, while only the first intron of *SmG* pre-mRNAs was retained, indicative of selective retention of particular introns in specific pre-mRNAs (Supplemental Figure S7). Using multi-FISH labeling, we analyzed the location and levels of *SmG* pre-mRNA retaining the first intron over a single cycle of cellular synthesis and turnover of poly(A)$^+$ RNA (Figure 5). We determined that these transcripts accumulate in both CBs and the nucleoplasm up to Stage 4 of the cycle (Figure 5A; Supplemental Figures S8–S12).

In Stage 5, the level of labeling was very low. We only detected small fluorescent foci in CBs and the nucleoplasm. Quantitative analysis revealed that the pre-mRNA levels gradually decrease in both CBs and the nucleoplasm from Stages 1–4 (Figure 5B). To confirm this result, we performed absolute quantification of *SmG* transcripts using RT-qPCR. To this end, we isolated microsporocytes from flower buds derived from one branch daily over 11 days during an entire cycle. We then selected four samples after validation of the stages by FISH on the collected material and isolated RNA for reverse transcription. These samples represented early Stage 1, late Stage 1, Stage 2/3 and Stage 4/5. We designed PCR primers such that the PCR product would include intron 1 and exon 2. Quantitative analysis showed that the number of copies of the immature form of the mRNA in Stage 1 increases slightly, which may indicate the ongoing synthesis of new mRNAs (Figure 5C). In the later stages, the copy number of this RNA gradually decreased. These data were consistent with the results obtained from in situ studies.

To check whether and when retained introns are spliced from their pre-mRNA, we designed a probe for the fully mature mRNA (thus after splicing of the first intron) and checked its location during a single cycle of cellular synthesis
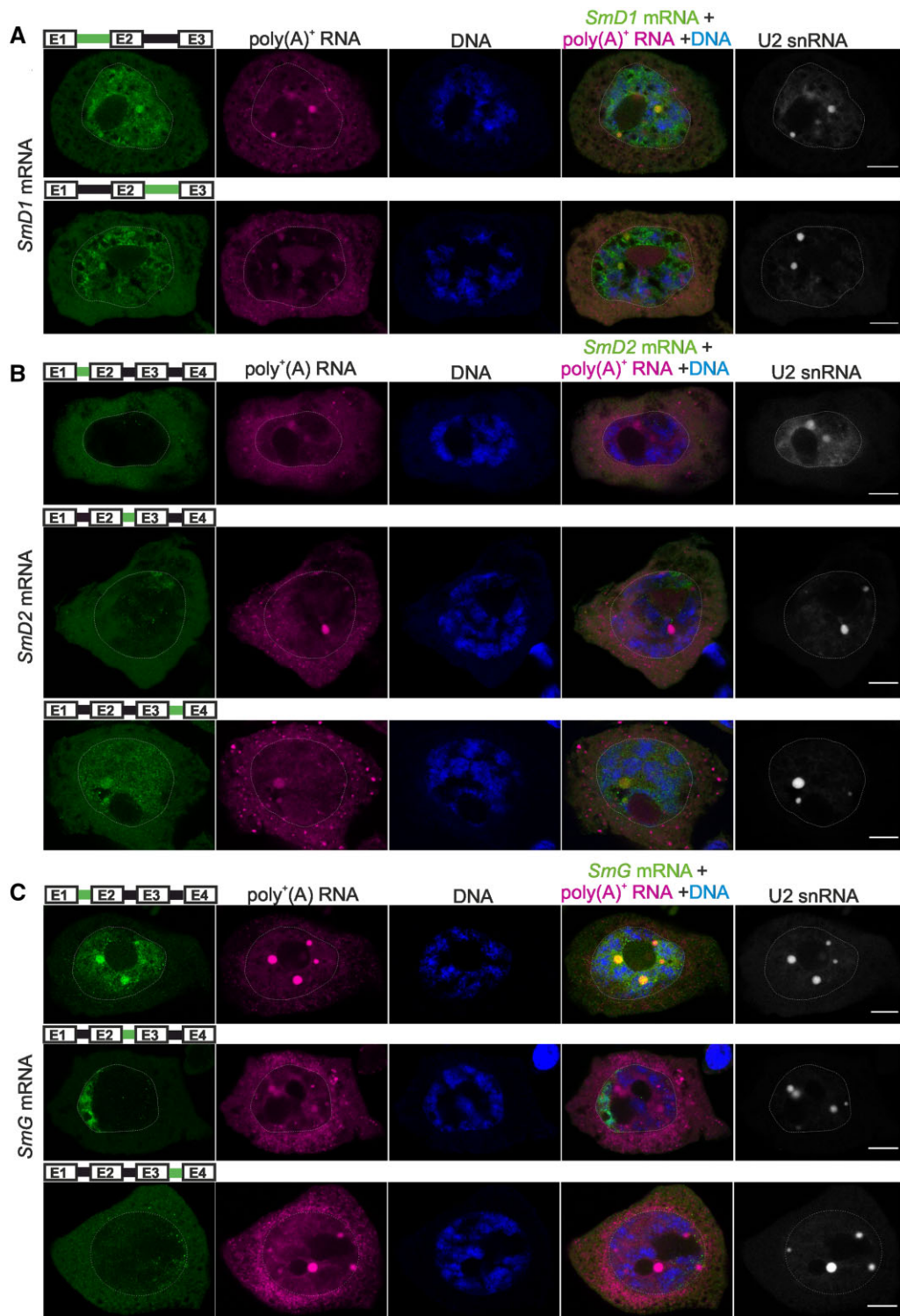
Figure 3 *Sm* mRNAs coding are retained in the nucleoplasm and CBs as pre-mRNAs containing defined intron(s). A, Localization of *SmD1* pre-mRNAs. Labeling with a probe complementary to intron 1 (top) or intron 2 (bottom), indicated in green in the transcript diagram, shows that both transcript types accumulate in the nucleoplasm and CBs. U2 snRNA labeling was used to identify CBs. B, Localization of *SmD2* pre-mRNAs. Labeling with a probe complementary to intron 1 (top), intron 2 (middle), or intron 3 (bottom), indicated in green in the transcript diagram, shows that only pre-mRNAs containing intron 3 accumulate in the nucleoplasm and CBs, while pre-mRNAs containing intron 1 or 2 are visible only in a few clusters. C, Localization of *SmG* pre-mRNAs. Labeling with a probe complementary to intron 1 (top), intron 2 (middle), or intron 3 (bottom), indicated in green in the transcript diagram, shows that only pre-mRNAs containing intron 1 accumulate in the nucleoplasm and CBs, while pre-mRNAs containing intron 2 or 3 are present only at the nucleus periphery. Bar = 10 µm.
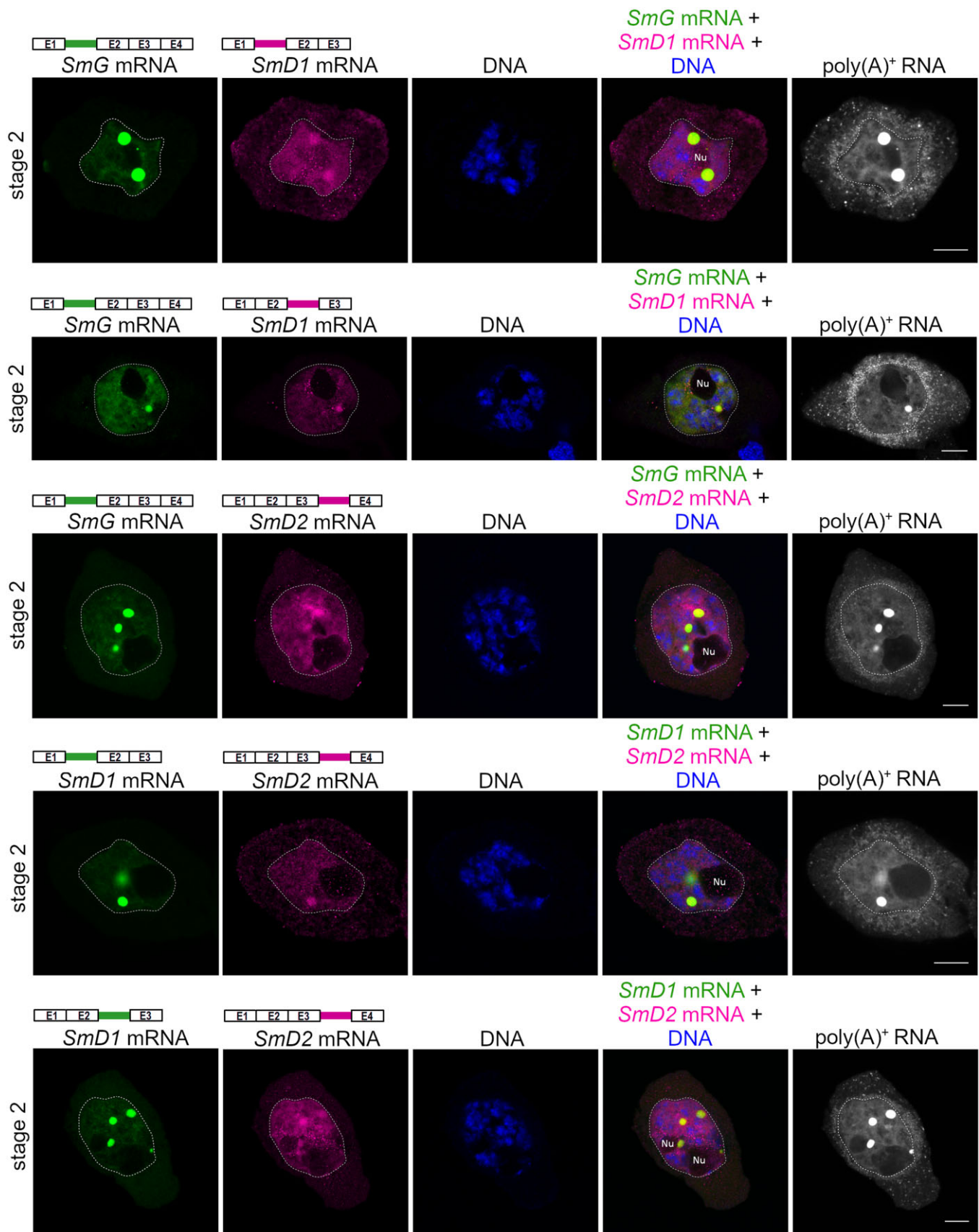
Figure 4 The retention of *Sm* pre-mRNAs co-occurs, largely in CBs. *SmG*, *SmD1*, and *SmD2* pre-mRNAs were labeled with a probe complementary to the intron indicated in the diagram above each panel. DNA was labeled with Hoechst 33342 and shown in blue. Poly(A)$^+$ RNA is shown in grayscale. Nu, nucleolus; Bar = 10 μm.
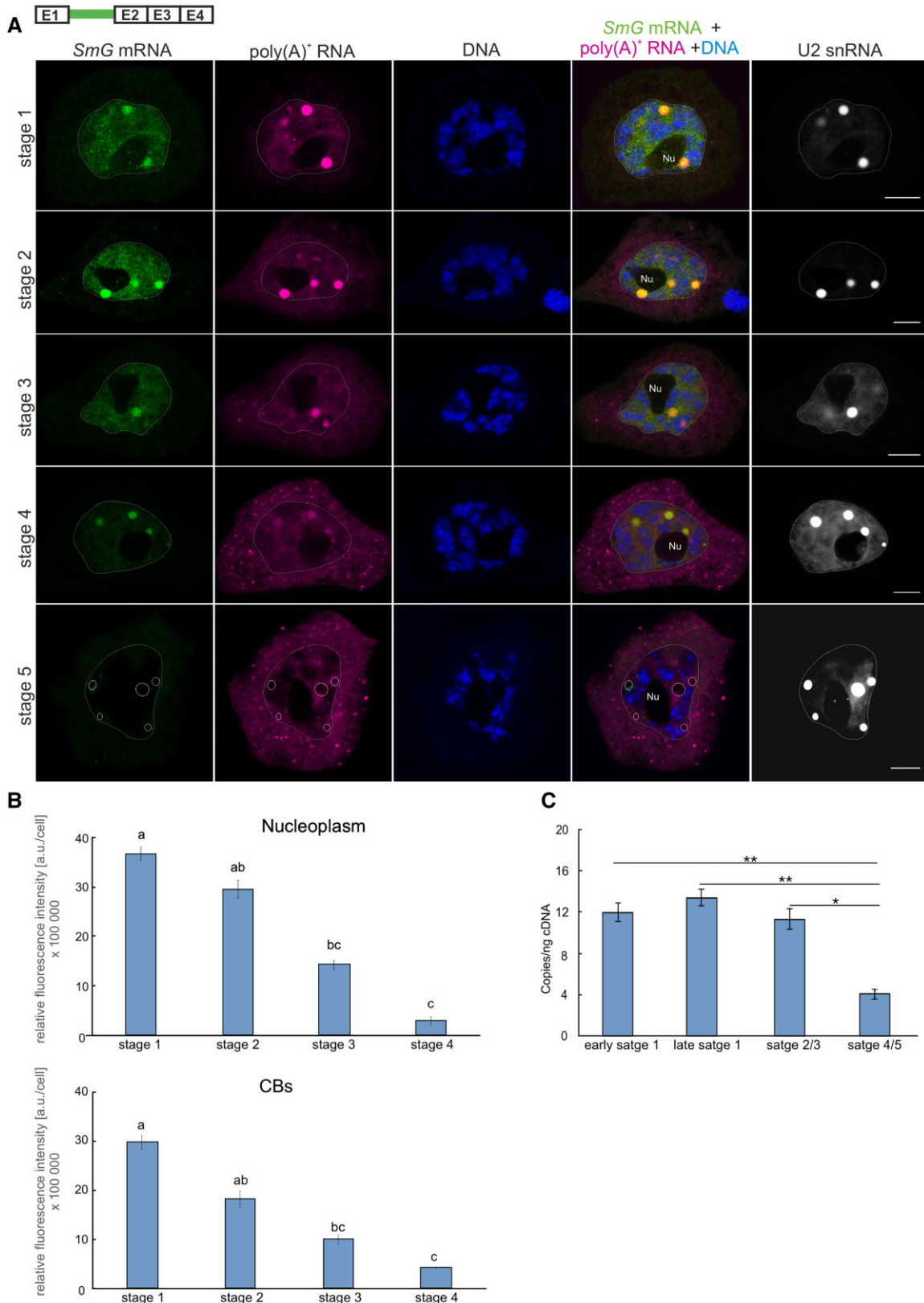
Figure 5 Detailed analysis of *SmG* pre-mRNAs that retain intron 1. A, Distribution of *SmG* pre-mRNAs retaining intron 1 against the background of the general pool of poly(A)$^+$ RNA distribution. U2 snRNA staining was used to show CBs. Details in the text. Nu, nucleolus; Bar = 10 μm. B, Levels of *SmG* pre-mRNA retaining intron 1 in the nucleoplasm (upper) or CBs (lower). Error bars indicate sem. Different lowercase letters indicate Significant differences (at α = 0.05), according to Kruskal–Wallis test with Dunn's post-hoc procedure using Benjamini–Hochberg FDR correction. n = 9 (Stage 1), 7 (Stage 2), 7 (Stage 3), and 8 (Stage 4). C, Amount of the immature forms of *SmG* pre-mRNA in successive stages of the cycle, as determined by RT-qPCR analysis. **$P < 0.01$; *$P < 0.05$ from ANOVA with Tukey's HSD.
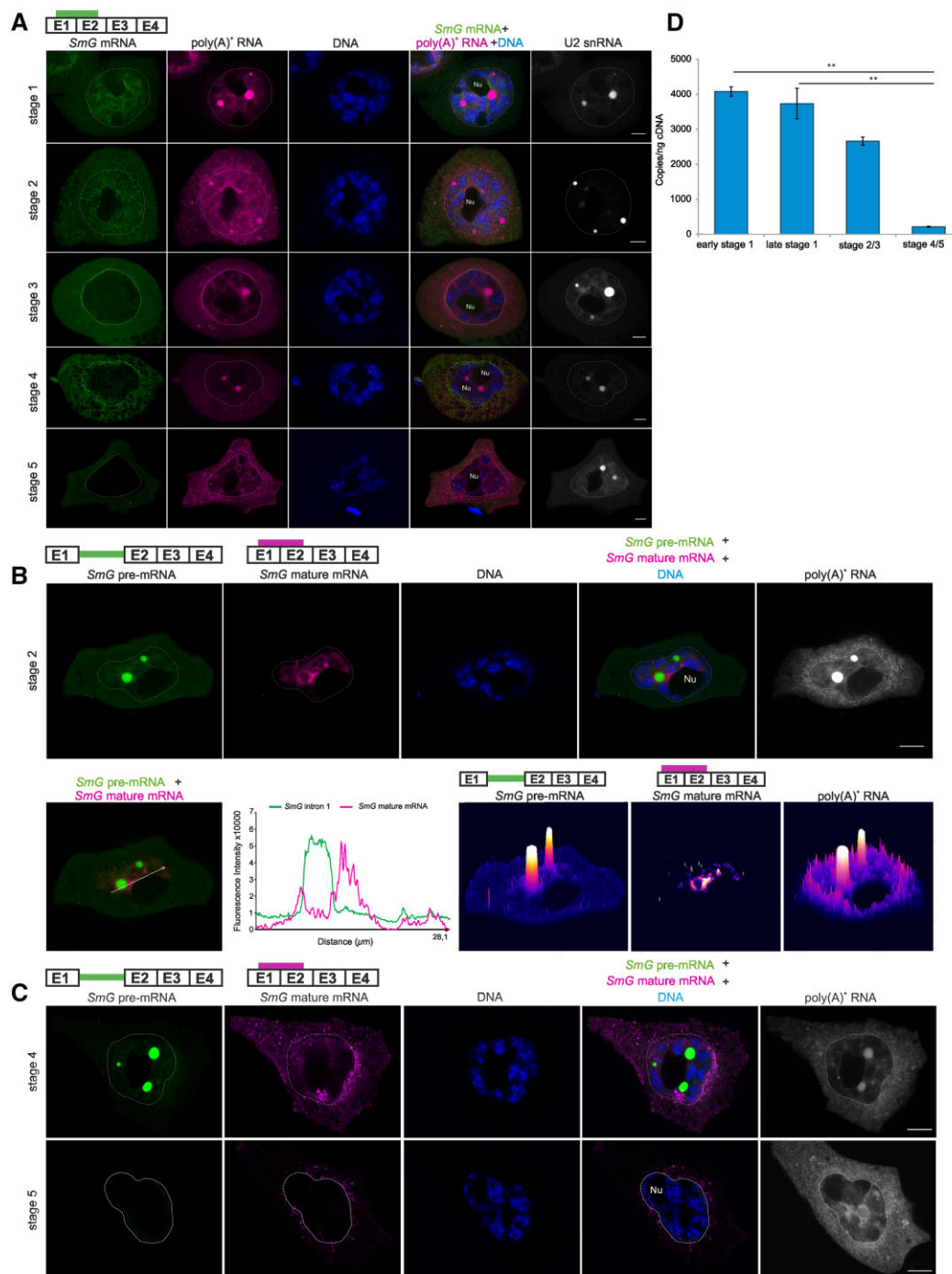
Figure 6 After removal of the retained intron, *SmG* mRNAs no longer accumulate in CBs. A, Mature *SmG* mRNA is absent from CBs at all stages, being detected in the nucleoplasm during Stages 1 and 2, concentrating along the nuclear envelope in Stage 3, and moving to the cytoplasm during Stages 4 and 5. Bar = 10 µm. B and C, Comparison of the distribution of immature and mature *SmG* mRNAs. B, Multi-labeling of the immature and mature form of *SmG* mRNA (upper). The fluorescence intensity of the *SmG* pre-mRNA and *SmG* mature mRNA along the white line is shown (lower). Right, 3D models of the general signal intensity of the analyzed cell. The signal intensity ranges from black (no signal) through blue, purple, red, orange (increasing signal intensity), ending at yellow and white (maximum foci). C, Multi-labeling of the immature and mature forms of *SmG* mRNA. D, Amount of the mature form of the *SmG* mRNA in successive cycle stages, by RT-qPCR analysis. **$P < 0.01$, as determined by a Kruskal–Wallis test with Dunn's post-hoc procedure with Benjamini–Hochberg FDR correction.

and turnover of poly(A)$^+$ RNA (Figure 6). We detected the fully spliced transcript as early as Stage 1 in the nucleoplasm. We did not observe its accumulation in CBs, in contrast to the nonfully spliced form present in CBs (Figure 6, B and C). We hypothesize that splicing of the retained intron(s) takes place immediately after the release

of these transcripts from CBs and before they are exported to the cytoplasm. In later stages, we detected fully spliced transcripts both in the nucleus and in the cytoplasm. For example, we observed and even nuclear and cytoplasmic labeling at Stage 2, followed by a drastic decrease in nuclear labeling intensity in favor of the cytoplasmic signal. Perhaps, the rate of export and further use of these transcripts vary depending on the stage of the cycle. During the first two stages, the significant accumulation of fully mature transcripts in the nucleus may suggest that their export rate to the cytoplasm is not very high. However, in later stages, the shift in the nuclear to cytosolic signal may reflect a rise in the mRNA export rate. We confirmed the results of in situ hybridization by absolute qPCR quantification analysis (Figure 6D).

Detailed analysis of the *SmG* transcripts showed significant involvement of CBs in their retention in the nucleus. Only the nonfully spliced form of transcripts with retained introns are present in CBs, raising a question as to whether they can be translated once released into the cytoplasm. To address this question, we analyzed a single cycle of expression of a specific protein-coding gene, with emphasis on the relationship between mRNA retention in CBs and protein synthesis in the cytoplasm. We selected the *Sm* gene family for this analysis, namely *SmD1*, *SmD2*, and *SmE*. We conducted a quantitative analysis of the corresponding mRNA levels during a single cycle in individual microsporocyte compartments: in the nucleoplasm, CBs, and the cytoplasm. We observed a particular correlation between the increase in mRNA signal in CBs and the appearance of mRNAs in the cytoplasm: mRNA levels in CBs peaked at Stage 3, at which point the cytoplasmic pool is very low. By the next stage, however, CB mRNA levels dropped while the cytoplasmic signal intensity increased (Supplemental Figure S13A), which may indicate intensive translation.

To assess whether the mRNAs exported to the cytoplasm area were translated, we also examined the abundance of Sm proteins with particular emphasis on the cytoplasm. We observed the accumulation of a new pool of Sm proteins in the cytosol. In fact, Sm protein levels in the cytoplasm increased rapidly in the last stages of the cycle, concomitantly with a rapid decrease in the mRNA pool located in CBs and a high level of mRNAs in the cytoplasm (Supplemental Figure S13B). Microscopy observations confirmed this result, as evidenced by in situ hybridization and immunolabeling. In Stage 5, where *Sm* mRNAs in microsporocytes are observed almost exclusively in the cytoplasm, we also detected intense protein labeling exclusively in the cytoplasm, likely reflecting de novo Sm protein synthesis (Supplemental Figures S13C and S14).

To confirm the hypothesis about the mass synthesis of Sm proteins at Stages 4 and 5, we performed a proteomics analysis from the cytoplasmic fraction of microsporocytes in the middle (Stages 2 and 3) and at the end of the cycle (Stages 4 and 5). We performed the experiment in triplicate for Stages 2 and 3 (Supplemental Data Set 2) and in quadruplicate for Stages 4 and 5 (Supplemental Data Set 3). We identified Sm proteins in all replicates of the cytoplasmic proteome for Stages 4 and 5 microsporocytes. However, Sm proteins were absent from the cytoplasmic proteome at Stages 2 and 3. Only once did the SmB protein appear in a small amount in one of the replicates of the analyzed proteome for the early stages of the cycle.

Therefore, we hypothesize that the retention of nonfully spliced forms of mRNAs in larch microsporocytes is functionally relevant, serving to regulate the time of export and thus control the synthesis time of individual proteins. The synthesis of Sm proteins at the end of the cycle would provide the splicing factors that will be necessary at the beginning of the next cycle, when transcriptional activity resumes (Supplemental Figure S13D).

Discussion

Nuclear retention of mRNA in larch microsporocytes controls mRNA export and translation time

As has been found in many other species, including mammals (Halpern et al., 2015), transcription during prophase in larch microsporocytes is a pulsating process (Kołowerzo-Lubnau et al., 2015). As our previous studies have shown, each transcription burst generates a poly(A)$^+$ RNA pool that is retained for some time in the nucleus after transcription has stopped as a consequence of chromatin rearrangement during meiotic prophase. In larch microsporocytes, during diplotene, there are four phases of chromatin contraction and five diffusion phases, when chromatin relaxation occurs (the so-called diplotene diffusion stage; Kołowerzo-Lubnau et al., 2015). Many studies have indicated that the retention of mRNAs in the cell nucleus results from mechanisms of quality control and detection of defective mRNAs and their subsequent degradation (Fasken and Corbett, 2009). In larch microsporocytes, the levels of poly(A)$^+$ RNAs retained within the nucleus are very high. Their nuclear pool significantly exceeds the cytoplasmic one throughout the entire diplotene (Kołowerzo-Lubnau et al., 2015). The scale of this phenomenon suggests that the retention of nuclear mRNAs in larch microsporocytes plays a new, underappreciated role in the regulation of certain genes, rather than just capturing defective transcripts.

Our results showed here that mRNAs retained at the beginning of the cycle of cellular synthesis and turnover of poly(A)$^+$ RNA are exported to the cytoplasm, initially in lower and then increasing quantities. We hypothesized that the export of mRNAs to the cytoplasm would be associated with their translation and the biosynthesis of functional proteins. To test this idea, we analyzed the cytoplasmic transcriptome of larch microsporocytes to identify mRNAs characterized by high export from the nucleus and high levels of mRNA in the cytoplasm. In the analyzed cytoplasmic transcriptome, the most numerous group was for transcripts encoding proteins associated with transcription and

posttranscriptional modifications, among which the *SmD1*, *SmD2*, and *SmG* transcripts. Furthermore, our previous research showed that Sm proteins and other splicing factors are intensely synthesized to process the products derived from the transcription burst occurring in diplotene (Hyjek et al., 2015). We therefore selected *Sm* mRNAs for further analysis. We asked whether the levels of exported *Sm* mRNAs correlate with the level of newly synthesized Sm proteins. We did observe newly synthesized Sm proteins at the last stage of the cycle, during which *Sm* mRNAs are almost exclusively in the cytoplasm, indicating that we are dealing with the synthesis and retention of functional transcripts that are exported and translated over time.

Moreover, our results point to the nuclear retention of *Sm* mRNAs as a means to introduce a new pool of Sm proteins at a particular moment in a cycle. Earlier studies on the biosynthesis of snRNPs complexes in larch microsporocytes have shown that Sm proteins and snRNAs are synthesized de novo in pulses (Smoliński et al., 2011; Hyjek et al., 2015). Thus, the assembly of functional snRNPs complexes occurs cyclically, as individual elements become available in a cell. We have shown that the moment of appearance of newly synthesized Sm proteins coincides with newly mass-synthesized RNA polymerase II complex and the start of a new transcription burst (Kołowerzo-Lubnau et al., 2015; Majewska et al., 2021). This observation indicates that the retention of *Sm* mRNAs in the nucleus delays the translation time of their encoding proteins to the next transcriptional round so that all necessary snRNP complexes are fully assembled for this next round of transcripts. The appearance of Sm proteins in the microsporocyte cytoplasm initiates the assembly of functional snRNPs complexes, which is a multistage process, including transcription of snRNAs, export of pre-snRNAs into the cytoplasm, Sm core assembly and exonucleolytic trimming of the snRNAs' 3′-ends, hypermethylation of 7-methylguanosine cap to form a 2,2,7-trimethylguanosine and re-import of snRNPs into the nucleus, where they are directed to CBs for nuclear maturation steps (Matera and Wang, 2014; Bohnsack and Sloan, 2018). The complexity of this process probably means that access to fully functional snRNPs complexes at the time of transcription burst in microsporocytes is likely to be limited. This deficit of fully functional splicing complexes, in addition to the mechanism of self-regulation of the levels of splicing proteins, such as Sm proteins, may complement the retention mechanism of mRNAs generated in a single transcriptional pulse. Insufficient accumulation of splicing factors can affect the selection of pre-mRNAs that will be spliced first and which pre-mRNAs might be retained for later processing. A specific feedback mechanism is created, depending on the availability of splicing factors. Naro et al. (2017) showed that the direct cause of intron retention (IR) is the low efficiency of their splicing, related to the high activity of RNA polymerase II. The high intensity of transcription causes competition between individual introns for the attachment of splicing factors, which leads to the retention of introns with weak splicing sites, a typical characteristic of retained introns. Decreasing the intensity of transcription by inhibiting the phosphorylation of RNA polymerase II has been shown to result in a significant improvement in the splicing efficiency of retained introns, which indicates that these introns bind spliceosomes less well, and therefore lose in the competition against other introns, and therefore remain unspliced (Naro et al., 2017).

Mechanisms of functional nuclear retention have been described, among others, in mammalian cells. A wide range of genes in highly metabolically active tissues has been shown to generate fully spliced, polyadenylated transcripts that are retained in the nucleus for periods that exceed their cytoplasmic lifetimes. These include genes such as those encoding the transcription factor ChREBP, Nlrp6, glucokinase, and the glucagon receptor (Halpern et al., 2015) and demonstrated that nuclear retention of mRNA can efficiently buffer cytoplasmic transcript levels from the noise that arises from transcriptional bursts.

An interesting example of functional retention is the formation of an mRNA reservoir in the nucleus of cells that will be transcriptionally silenced for an extended period. In the microspores of the fern *M. vestita*, despite an absence of transcriptional activity, proteins necessary for spermatogenesis are nevertheless synthesized at certain times in these cells (Boothby et al., 2013). Analysis of RNA-seq-derived transcriptomes revealed a large subset of intron-retaining transcripts that encode proteins essential for gamete development. These transcripts are retained within the cell nucleus, and then are posttranscriptionally spliced at the precise moment of gametic development, exported to the cytoplasm, and translated. The retention of introns is thus associated with an arrest of the translational events necessary for spermatogenesis in *M. vestita*. There is a widespread IR program in meiosis in male generative cell lines in the mouse, stabilizing highly expressed transcripts with strong relevance for spermatogenesis. Intron-retaining mRNAs accumulate in the nucleus of meiotic cells in fission yeast (*Schizosaccharomyces pombe*; Averbeck et al., 2005) and mouse (Naro et al., 2017) for hours or days. These intron-retaining mRNAs can be recruited to ribosomes for translation long after their synthesis. Nuclear retention of mRNAs can therefore also favor accumulation, storage, and the timely use of specific transcripts during highly organized cell differentiation programs, such as spermatogenesis.

Notably, our recent report suggested that cytoplasmic mRNP bodies (S-bodies) accumulate to high levels in larch microsporocytes in the fourth cycle (Hyjek-Skaładanowska et al., 2020). Thus, at the fourth cycle, many mRNAs are stored in the cytoplasm and probably not translated immediately, or even degraded. Therefore, we should consider that some of the export in the cytoplasm does not necessarily correlate with active translation. The participation of Sm proteins in regulating mRNA levels in the cytoplasm does not exclude their canonical role in splicing in the next fifth cycle of mRNA synthesis at the end of diplotene. In

addition, other mRNAs that undergo the same retention process as *Sm* mRNAs are translated, such as those encoding subunits of RNA polymerase II, as evidenced by the presence of the precursor nonphosphorylated form of RNA pol II in the cytoplasm in the fifth stage of the cycle (Majewska et al., 2021). This observation does not exclude a different fate of the fraction of mRNAs retained in the nucleus in addition to their later participation in protein synthesis (e.g. delayed translation or degradation). The data presented here indicate that a significant mRNA pool undergoing nuclear retention is eventually effectively translated, leading to efficient protein synthesis at the end of each cycle.

The positive role of mRNA nuclear retention has also been demonstrated in many cells under stress. Indeed, cells inhibit mRNA splicing in response to heat shock, which results in nuclear retention of some pre-mRNAs (Sadis et al., 1988; Saavedra et al., 1996; Shalgi et al., 2014). This mechanism may prevent the production of aberrant proteins. The presence of intron-retaining transcripts in the cell accumulating at substantial levels suggests that they are relatively stable in the nucleus and may have some function, perhaps serving as a reservoir of partially processed transcripts, which can later be fully spliced and exported to the cytoplasm, to rapidly produce proteins when stress conditions abate (Shalgi et al., 2014).

Sm mRNAs that are retained in larch microsporocytes are not fully spliced

Functional nuclear mRNA retention may relate to both fully mature transcripts (Halpern et al., 2015) as well as nonfully spliced transcripts (Boothby et al., 2013; Shalgi et al., 2014; Bergeron et al., 2015; Naro et al., 2017). Here we showed that *Sm* transcripts are retained in larch microsporocytes in the nucleoplasm and CBs as nonfully spliced pre-mRNAs, with one or more introns remaining. At precisely defined times, these retained introns are spliced, and fully mature mRNAs are then released into the cytoplasm, where the proteins are produced.

IR is a conserved form of alternative splicing (AS) characterized by the inclusion of intronic sequence in a mature transcript (intron retained transcripts [IRTs]). IR is an independent mechanism for controlling and enhancing the complexity of the transcriptome. IR facilitates rapid responses to biological stimuli, is involved in disease pathogenesis, and can generate novel protein isoforms (Vanichkina et al., 2018). AS is prevalent in land plants, and the generation of different mRNA variants from a single gene is tightly regulated. AS occurs frequently in response to altered light conditions, and also plays a crucial role in seedling photomorphogenesis. However, despite the frequency of this process and its essential contribution to plant development, the functional role of AS remains unknown for most splicing variants (Hartmann et al., 2018). Alternative transcripts often carry premature termination codons (PTCs). AS triggers the cytoplasmic RNA degradation pathway known as nonsense-mediated decay (NMD). However, the most prevalent AS events in plants often introduce PTCs that are insensitive to NMD. These NMD-resistant splice variants with IR localize to the nucleus of plant cells (Göhring et al., 2014), raising a question about the biological role of these unique RNA variants.

As shown by recent animal studies, the fate of intron-retaining transcripts can vary. IRTs can be exported to the cytoplasm to either generate distinct protein isoforms (Gontijo et al., 2011; Hossain et al., 2016) or be degraded by NMD (Jaillon et al., 2008; Peccarelli and Kebaara, 2014). However, some intron-containing mRNAs, as documented in yeast (Averbeck et al., 2005), plants (Boothby et al., 2013) and animal cells (Boutz et al., 2015; Naro et al., 2017) remain in the nucleus, leading to time-regulated protein production from these transcripts. In this case, IR does not lead to the generation of additional protein isoforms (Yap et al., 2012; Braunschweig et al., 2014; Boutz et al., 2015; Bresson et al., 2015; Liu et al., 2017).

In larch microsporocytes, intron-containing mRNAs are retained in the nucleus until splicing is completed. As illustrated here by the *SmG* mRNA retaining its first intron, the rate of posttranscriptional splicing of this intron regulates the export rate of this mRNA, thus regulating the time and levels of protein synthesis. This type of regulation is mainly observed in developing, specialized or stressed cells, often with lower or even inhibited transcriptional activity. Retained transcripts then act as a long-term pool that can be used to produce a given protein at a specific time.

Analysis of the transcriptome at various time intervals during the development of *M. vestita* microspores showed that a large proportion of the transcripts usually retain one relatively short intron (Boothby et al., 2013). RT-PCR confirmed that IRTs temporally precede their fully spliced isoforms during development. Studies using the spliceosome inhibitor spliceostatin A and the transcriptional inhibitor α-amanitin confirmed that the maturation of IRTs to mRNAs is spliceosome-dependent but independent of transcription. These studies showed that the retention of specific introns in the microspores of *M. vestita* is regulated during desiccation, a time when the spores undergo expansive translation of stored pre-mRNAs.

A similar mechanism of nuclear retention of intron-containing mRNAs was also described in mouse meiotic spermatocytes and postmeiotic spermatids (Naro et al., 2017). Most IR events were detected in meiotic spermatocytes, whereas these introns are spliced in postmeiotic cells, indicating that these transcripts are stable in the nucleus and can be maintained for days after their synthesis. IR regulation determines the timely usage of transcripts when transcription is repressed. These observations suggested that temporal regulation of IR plays an essential role in male germ cell differentiation.

Boutz et al. (2015) demonstrated the presence of intron-retaining mRNAs both in cell lines and liver tissue from adult animals. A common set of such transcripts was found among genes expressed in both embryonic stem cells and

adult livers, alongside sets of tissue-specific IRTs that are enriched for genes with particular functions related to the cell type in which they occur. Many of the functional gene categories in which IRTs are overrepresented encode products whose physiological concentrations must be tightly controlled. Regulation of intron-retaining mRNAs splicing kinetics may contribute an additional control point for fine tuning of gene expression in these cases.

Are CBs a checkpoint before exporting previously retained mRNAs into the cytoplasm?

In larch microsporocytes, a significant role can be attributed to CBs for nuclear retention and the subsequent release of mRNAs into the cytoplasm. From the very beginning of the cycle of cellular synthesis and turnover of poly(A)$^+$ RNA, mRNAs accumulate in CBs at a concentration of 70–200 times higher than in the surrounding nucleoplasm. We observed the highest levels of mRNAs in CBs just before the mass export of mRNAs to the cytoplasm, which indicates their relationship with switching the "retention program" to export (Figures 2, A and 6, A). This change from retention to export is probably related to the activation of mechanisms leading to the triggering of posttranscriptional splicing. In contrast to the nucleoplasm, we detected no mature transcripts in CBs. Only nonfully spliced transcripts with retained introns appeared to be stored in CBs, indicating that a type of switch must operate in CBs to modulate the release of retained transcripts and activate their posttranscriptional splicing.

To date, NSs have been considered to be nuclear domains associated with pre-mRNA storage. However, unspliced mRNAs (after spliceostatin A treatment), mRNAs with retained introns, and completely spliced mRNAs have recently been shown to be temporally sequestered within NSs (Girard et al., 2012; Carvalho et al., 2017). In larch microsporocytes, the accumulation of polyadenylated transcripts in NSs appears to be at much lower levels than in CBs (Kołowerzo et al., 2009). It remains unclear, however, why nonfully spliced mRNAs are targeted to CBs. As previously reported, one of the reasons for the retention of transcripts containing introns is the presence of spliceosome elements, such as U1 snRNA or U2AF, embedded in these sequences (Takemura et al., 2011). The retention of pre-mRNAs and associated spliceosomes in speckles is thus believed to be driven by the presumed presence of speckle localization signals in all splicing factors (Girard et al., 2012). Thus, spliceosomes, composed of many splicing factors, might exhibit a high affinity for speckles (Spector and Lamond, 2011). We hypothesize that the strong affinity of splicing factors associated with nonfully spliced mRNAs in microsporocytes determines their targeting to CBs. CBs are most commonly associated with the production of spliceosomal enzymatic backbones, known as small nuclear RNPs, that catalyze RNA splicing. Both during early and late stages of snRNPs biogenesis, CBs are enriched in their components, snRNA

precursors with sequences to be deleted at the 3′-end and fully assembled mature snRNPs after their reimport into the nucleus from the cytoplasm (Frey and Matera, 2001; Stanek and Neugebauer, 2006; Wang et al., 2016). Additionally, CBs perform quality control on snRNPs through the U4/U6 di-snRNP recycling factor Squamous cell carcinoma antigen recognized by T cells 3; (Novotný et al., 2015), explaining why CBs may serve as "boosters" to aid cells during spliceosome assembly to accommodate the additional load required in specific cells. A correlation has been noted between the availability of spliceosome elements and aberrant splicing. The reduction in the production of functional snRNPs complexes may lead to widespread AS, especially IR events (Zhang et al., 2008).

In larch microsporocytes, CBs appear to be an important element influencing the availability of splicing factors and the associated nuclear retention of mRNAs. They are also a retention site for nonfully spliced *Sm* mRNAs encoding Sm proteins, a component of snRNP complexes. CBs also appear to be responsible for activating the posttranscriptional splicing of these mRNAs, resulting in the formation of functional Sm proteins. At the same time, CBs participate in the synthesis and assembly of snRNP elements and assembly of spliceosomal subunits, all pointing to CBs as regulators of mRNA retention other than *Sm* mRNAs by regulating the rate of processing of individual elements of the snRNA complexes.

Materials and methods

Plant material and isolation of meiotic protoplasts

European larch (*Larix decidua* Mill.) anthers were collected from the same tree (Toruń, Poland, coordinates 53.021155 N, 18.570384 E) in successive meiotic prophase stages of diplotene (once a week from December 20, 2018 until January 20, 2019), to ensure constant experimental conditions. Anthers were fixed in 4% (w/v) paraformaldehyde in phosphate-buffered saline (PBS), pH 7.2, for 12 h and then squashed to release free microsporocytes. Meiotic protoplasts were isolated from these cells according to Kołowerzo et al. (2009) and were then subjected to immunofluorescence–FISH reactions.

Cytoplasmic fraction isolation and RNA purification for sequencing

To obtain the cytoplasmic fraction for sequencing libraries, the plant material was homogenized in liquid nitrogen, and then mixed in with ribonucleoprotein extraction buffer (RNP buffer; 200-mM Tris–HCl pH 9.0, 110-mM potassium acetate, 0.5% [v/v] Triton X-100, 0.1% [v/v] Tween-20, 2.5-mM DTT, Roche protease inhibitor) in the proportion of 10 mL per 1.5 g of material. After thawing on ice (~5 min), the material was filtered through one Miracloth layer (diameter 22–25 µm; Millipore Sigma, Burlington, MA, USA) and centrifuged in a swing out centrifuge at 1,500 g for 2 min at 4°C. The obtained supernatant was used as the cytoplasmic

fraction of the anther cells for RNA purification and cDNA library preparation.

Total RNA isolation from the cytoplasmic fraction (for sequencing libraries)

RNA isolation was performed using the Direct-zol TM RNA Microprep (Zymo Research Irvine, CA, USA) according to the manufacturer's instructions.

cDNA library preparation and sequencing

cDNA library generation and sequencing was performed using the TruSeq Stranded Total RNA with Ribo-Zero Plant kit (Illumina, San Diego, CA, USA) according to the protocol described in Hyjek-Składanowska et al. (2020). The libraries were sequenced on a MiSeq instrument (Illumina) with MiSeq Reagent Kit version 3 (150 cycles). The reads were filtered in terms of quality via PRINSEQ-LITE (Schmieder and Edwards, 2011). The sequence stretches with an average quality below 20 over a window of 20 nt were removed, and the reads were cut immediately before the first incidence of a degenerated base. Furthermore, trimmed reads shorter than 36 nt were removed. The reads from all libraries were pooled and assembled de novo using TRINITY version 2.8.2 with default settings (Grabherr et al., 2011).

Total RNA isolation from microsporocytes and reverse transcription (for PCR and RT-qPCR reaction)

Total RNA isolation was performed from the cytoplasmic fraction (for sequencing libraries) using the Direct-zol RNA Microprep kit (Zymo Research) according to the manufacturer's instructions. Isolation was performed from the suspension of microsporocytes (freshly collected anthers were squashed into PBS buffer, and then the suspension was filtered on 70-µm sterile cell sieves (BIOLOGIX Group, Jinan, China) to obtain a pure fraction of microsporocytes; the procedure was performed on ice using the TRI Reagent (Sigma-Aldrich Solutions, Burlington, MA, USA) according to manufacturer's instructions. The buds from which the microsporocytes were isolated came from one branch. Material for total RNA isolation was collected daily. The material corresponding to the first to fifth stages of a single cycle of cellular synthesis and turnover of poly(A) $^{+}$ RNA were selected for RT-qPCR.

Reverse transcription was performed using SuperScript IV VILO Master Mix with ezDNase Enzyme (Invitrogen, Waltham, MA, USA; Catalog number # 11766500) according to manufacturer's instructions. About 2,000 ng of total RNA were used for a single reaction.

PCR

A C1000 Touch Thermal Cycler (Bio-Rad, Hercules, CA, USA) was used for all PCR reactions. Genomic DNA was isolated from the same tree using a combination of bead-beating and flocculation according to the protocol of the PowerLyzer DNeasy kit (Qiagen, Hilden, Germany). The amplifications were performed in a final volume of 20 µL containing the following: 10-ng cDNA or 200–250 ng of genomic DNA, 1.5-mM $MgCl_2$ (Syngen), 0.2 mM of each dNTP (Invitrogen), standard Gold Taq Buffer with 1 unit of Gold Taq DNA polymerase (Syngen) and 5 µM of each primer. The PCR conditions for both primer pairs were as follows: denaturing at 95°C for 3 min, 34 cycles of 30 s at 95°C, 30 s at 56°C, 30 s at 72°C; final extension at 72°C for 5 min. The amplification products were analyzed using the HS DNA kit on an Agilent 2100 BioAnalyzer according to the manufacturer's instructions.

Primer sequences for amplification of the mRNA encoding SmG protein intron 1–exon 2 boundary (SmG I1-E2)—F: 5′-TCAGTTCAGGGTTTAAAGATCA-3′, R: 5′-AGTTTAGCCTTC TTAAACCAGA-3′; Primer sequences for amplification of the mRNA encoding SmG protein inside second intron (SmG I2F-I2R) F: 5′-GTTAAGCTTCTTGTCCATGTAC-3′, R: 5′-CTGTGCAGTTCCAAAATTATGA-3′.

qPCR

Absolute quantification based on SYBR-Green PCR assays was performed on a LightCycler 96 Instrument (Roche Life Science, Penzberg, Germany). For each 10-µL reaction, 2 µL of cDNA solution (1 ng of cDNA), 5-µL LightCycler 480 SYBR Green I Master (Roche Life Science), 1-µL forward primer (3 µM) and 1-µL reverse primer (3 µM) were mixed (primer sequences in Supplemental Figure S15). The PCR program was as follows: denaturation at 95°C for 15 s; annealing at 53°C for 15 s; synthesis 72°C for 30 s for 45 cycles. After each run, a melting curve analysis was performed by measuring fluorescence intensity in a temperature interval ranging from 65°C to 97°C. To assess primer–dimer and nonspecific products, blank controls containing water as a template were used for every primer pair in each experiment. Standard curves were prepared using purified PCR products as templates for the following numbers of molecules: 10, 100, 1×10^3, 1×10^4, 1×10^5, and 1×10^6 (Supplemental Table S1). Each reaction (both standards and cDNA samples) was run in five technical replicates. Threshold cycles (Ct) were determined by the second derivative method using the LighCycler software. The number of target molecules in each sample was calculated in Excel based on the obtained Ct values using the standard formula ($10^{(sampleCt - standard\ curve\ intercept)/standard\ curve\ slope}$).

Proteome analysis from the cytoplasmic fraction of microsporocytes

For proteomics analysis of the contents from the cytoplasmic fraction of cells, larch anthers were ground in liquid nitrogen and homogenized in protein extraction buffer (extraction buffer: 200-mM Tris–HCl pH 9.0, 110-mM potassium acetate, 0.5% [v/v] Triton X-100, 0.1% [v/v] Tween-20, 2.5-mM DTT, Roche Complete protease inhibitor; Cat # 04 693 116 001) in the proportion of 0.3 mL per 0.1 g of plant material. The mixture was centrifuged twice at 14,000g at 4°C, then centrifugated at 16,000g for 30 min at 4°C to obtain native cytosolic protein extract. LC–MS/MS analysis was performed by the Mass Spectrometry Laboratory (IBB PAS,

Warsaw, Poland) using an Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA). The proteins were identified with the MASCOT software version 2.4.1 (http://www.matrixscience.com/) and the TAIR10 database as reference proteome. The samples were analyzed with MScan software (http://proteom.ibb.waw.pl/mscan/index.html).

FISH in multiplex reactions

For hybridization, each probe was resuspended in hybridization buffer (30% [v/v] formamide, $4\times$ SSC, $5\times$ Denhardt's buffer, 1-mM EDTA, 50-mM phosphate buffer) at a concentration of 50 pmol/mL. Hybridization was performed overnight at 26°C in the dark. The antisense DNA oligonucleotides (Genomed, Warsaw, Poland; sequences of the oligonucleotides are listed in Supplemental Table S1) were resuspended in hybridization buffer at a concentration of 10 pmol/mL for each probe. Digoxygenine (DIG) probes were detected after hybridization using mouse anti-DIG (Roche, Basel, Switzerland; Cat # 11333062910) antibody in 0.05% (w/v) acetylated bovine serum albumin (acBSA) in PBS (diluted 1:200) in a humidified chamber overnight at 11°C and anti-mouse Alexa 488 (Invitrogen; Cat # A32723) antibody in 0.01% (w/v) acetylated BSA in PBS (diluted 1:1,000) in a humidified chamber for 1 h at 36°C.

Multiplex indirect FISH reactions

To show the simultaneous location of two different retained transcripts, multiplex indirect FISH reactions were performed. In the first step, one of the retained transcripts was hybridized with the DIG-labeled probe. In the next step, probes were detected using mouse anti-DIG (Roche) and rabbit anti-DIG (Invitrogen; Cat # 700772 diluted 1:100) primary antibodies. Incubation with the secondary anti-mouse Alexa 647 (Invitrogen; Cat # A-21235 diluted 1:100) and anti-rabbit Alexa 488 (Invitrogen; Cat # A32731 diluted 1:100) antibodies was carried out until full saturation. After this, hybridization with the DIG-labeled probe for the second retained transcript was performed. Probes were further detected by using rabbit anti-DIG primary antibodies. Incubation with the secondary anti-rabbit Alexa 488 (Invitrogen) and anti-rabbit Alexa 633 (Invitrogen) antibodies was carried out until full saturation.

Immunodetection of Sm proteins

For immunodetection, microsporocyte protoplasts were incubated in 0.1% (w/v) NaBH$_4$ for 2×5 min and washed with PBS, pH 7.2. Nonspecific antigens were blocked with PBS buffer containing 2% (w/v) acetylated BSA for 15 min. Sm proteins were detected by incubating with anti-Sm primary antibody (ANA Human reference serum, Centers for Disease Control and Prevention, Atlanta GA 30333) in 0.1% (w/v) acetylated BSA in PBS (1:300) in a humidified chamber overnight at 11°C and anti-human secondary antibody conjugated with Cy3, in 0.01% (w/v) acetylated BSA in PBS (1:100) or and anti-human secondary antibody conjugated with Alexa 633 (Invitrogen; Cat # A21091) in 0.01% (w/v) acetylated BSA in PBS (1:200) in a humidified chamber at 36°C for 1 h.

Design of multi-labeling reactions

Double-labeling immunofluorescence–FISH reactions (Sm proteins + mRNAs encoding Sm proteins; Sm proteins + pol RNA II phospho CTD Ser2 + scaRNA) were performed as described above. In the double-labeling immunofluorescence–FISH reactions, the immunocytochemical methods always preceded the in situ hybridization methods. When in situ hybridization was applied first, the subsequent levels of immunofluorescence signals were very low. After labelling, the slides were stained for DNA detection with Hoechst 33342 (1:1,000, Invitrogen; Cat # H3570) and mounted in ProLong Gold Antifade reagent (Invitrogen; Cat # P36930).

Control reactions

For in situ hybridization, incubations lacking the probe or lacking the primary antibody were performed. All control reactions produced negative results. For immunofluorescence, control incubations lacking the probe and primary antibody were performed. This control reaction produced negative results. Control incubations lacking only the probes (Supplemental Table S1) were also performed. This control reaction produced positive results.

Microscopy and quantitative measurements

Images were captured with a Leica SP8 confocal microscope with an optimized pinhole, long exposure time (200 kHZ) with a $63\times$ magnification (numerical aperture, 1.4) Plan Apochromat DIC H oil immersion lens. Images were collected sequentially in blue (Hoechst 33342), green (Alexa 488), red (Cy3), and far-red (Alexa 633, Cy5) channels. To minimize bleedthrough between fluorescence channels, the laser was set to low power (0.3% of maximum power) and sequential image captures were performed. For bleedthrough analysis and control experiments, the Leica SP8 software was used. Image acquisition was performed using constant parameters (laser power, detector gain, emission band, and resolution). At least six cells from a given stage were analyzed. Three-dimensional optical sections were acquired with a 0.5-μm step interval. The obtained data were corrected for background autofluorescence for all antigens and developmental stages, as determined by negative control signal intensities. ImageJ (Schneider et al., 2012) was used for image processing and analysis. Each reaction step was performed using the same temperatures, incubation times, and concentrations of probes and antibodies.

Statistical analyses

All statistical computations were performed in R version 4.0.3, and a significance level of 0.05 was used. qPCR results (numbers of amplicon molecules per ng of cDNA) were analyzed by means of Kruskal–Wallis test and Dunn's *post hoc* analysis using Benjamini–Hochberg's false discovery rate (FDR) as correction for multiple comparisons (R: dunn.test with method = "bh"; Supplemental Data Set 4).

Shapiro–Wilk's and Levene's tests (of the lawstat package) were used to check for assumption of normality and homogeneity of variance. As the assumptions did not hold, the Kruskal–Wallis test with Dunn's *post hoc* procedure was applied to test for differences between multiple groups, that is, signal levels at different stages. Spearman's rank correlation test (rcorr function of the Hmisc R package) was used to check if Sm proteins and their mRNA levels in the cytoplasm were correlated (Supplemental Data Set 4).

Accession numbers

Data supporting the findings of this work are provided in the main text and the supporting information files. Raw reads generated during this project were deposited in the NCBI SRA database and are accessible through BioProject no. SAMN14564474. The datasets for the transcriptomic analysis are available at https://www.ncbi.nlm.nih.gov/sra/?term=SAMN14564474 and at https://trace.ncbi.nlm.nih.gov/Traces/sra/?run=SRR11514884. The full dataset for the transcriptomic analysis (Trinity RNA-Seq de novo cytoplasmic transcriptome assembly) is available as Supplemental File S1. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2022) partner repository (http://www.proteomexchange.org/) with the dataset identifier PXD032074 (Project Name: Larch microsporocytes cytosplasmic fraction, Project DOI: 10.6019/PXD032074).

Supplemental data

The following materials are available in the online version of this article.

Supplemental Figure S1. Characteristic distribution pattern of poly(A)$^+$ RNA in Stage 1 of the cycle of cellular synthesis and turnover of poly(A)$^+$ RNAs.

Supplemental Figure S2. Distribution of the active form of polymerase RNA II during the single cycle of cellular synthesis and turnover of poly(A)$^+$ RNAs.

Supplemental Figure S3. Characteristic distribution pattern of poly(A)$^+$ RNAs in Stages 2 and 3 of the cycle of cellular synthesis and turnover of poly(A)$^+$ RNAs.

Supplemental Figure S4. Characteristic distribution pattern of poly(A)$^+$ RNAs in Stages 4 and 5 of the cycle of cellular synthesis and turnover of poly(A)$^+$ RNAs.

Supplemental Figure S5. Annotation of cytoplasmic mRNAs.

Supplemental Figure S6. Comparison of the distribution of different forms of *Sm* mRNAs.

Supplemental Figure S7. High Sensitivity DNA Bioanalyzer assay as a control for selective retention of particular introns in *Sm* transcripts.

Supplemental Figure S8. Characteristic distribution pattern of *SmG* pre-mRNA in Stage 1.

Supplemental Figure S9. Characteristic distribution pattern of *SmG* pre-mRNA in Stage 2.

Supplemental Figure S10. Characteristic distribution pattern of *SmG* pre-mRNA in Stage 3.

Supplemental Figure S11. Characteristic distribution pattern of *SmG* pre-mRNA in Stage 4.

Supplemental Figure S12. Characteristic distribution pattern of *SmG* pre-mRNA in Stage 5.

Supplemental Figure S13. Retained mRNAs are exported to the cytoplasm and translated.

Supplemental Figure S14. Distribution of Sm proteins during a single cycle of cellular synthesis and turnover of poly(A)$^+$ RNA.

Supplemental Figure S15. qPCR primers and standard curves.

Supplemental Figure S16. Control reactions.

Supplemental Table S1. Probes used in multi FISH labeling.

Supplemental Data Set 1. Cytoplasmic transcripts detected in larch microsporocytes.

Supplemental Data Set 2. Proteomics analysis from the cytoplasmic fraction of microsporocytes in the middle of the cycle (triplicate for Stages 2 and 3).

Supplemental Data Set 3. Proteomics analysis from the cytoplasmic fraction of microsporocytes at the end of the cycle (quadruplicate for Stages 4 and 5).

Supplemental Data Set 4. Summary of statistical analyses.

Supplemental File S1. Trinity RNA-seq de novo cytoplasmic transcriptome assembly. The full dataset for the transcriptomic analysis.

Funding

References

Averbeck N, Sunder S, Sample N, Wise JA, Leatherwood J (2005) Negative control contributes to an extensive program of meiotic splicing in fission yeast. Mol Cell **18**: 491–498

Bahar Halpern K, Caspi I, Lemze D, Levy M, Landen S, Elinav E, Ulitsky I, Itzkovitz S (2015) Nuclear retention of mRNA in mammalian tissues. Cell Rep **13**: 2653–2662

Bergeron D, Pal G, Beaulieu YB, Chabot B, Bachand F (2015) Regulated intron retention and nuclear pre-mRNA decay contribute to PABPN1 autoregulation. Mol Cell Biol **35**: 2503–2517

Boothby TC, Zipper RS, van der Weele CM, Wolniak SM (2013) Removal of retained introns regulates translation in the rapidly developing gametophyte of *Marsilea vestita*. Dev Cell **24**: 517–529

Boutz PL, Bhutkar A, Sharp PA (2015) Detained introns are a novel, widespread class of post-transcriptionally spliced introns. Genes Dev **29**: 63–80

Braunschweig U, Barbosa-Morais NL, Pan Q, Nachman EN, Alipanahi B, Gonatopoulos-Pournatzis T, Frey B, Irimia M, Blencowe BJ (2014) Widespread intron retention in mammals functionally tunes transcriptomes. Genome Res **24**: 1774–1786

Bresson SM, Hunter OV, Hunter AC, Conrad NK (2015) Canonical poly(A) polymerase activity promotes the decay of a wide variety of mammalian nuclear RNAs. PLoS Genet **11**: e1005610

Carvalho T, Martins S, Rino J, Marinho S, Carmo-Fonseca M (2017) Pharmacological inhibition of the spliceosome subunit SF3b triggers exon junction complex-independent nonsense-mediated decay. J Cell Sci **130**: 1519–1531

Fasken MB, Corbett AH (2009) Mechanisms of nuclear mRNA quality control. RNA Biol **6**: 237–241

Frey MR, Matera AG (2001) RNA-mediated interaction of Cajal bodies and U2 snRNA genes. J Cell Biol **154**: 499–509

Girard C, Will CL, Peng J, Makarov EM, Kastner B, Lemm I, Urlaub H, Hartmuth K, Lührmann R (2012) Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion. Nat Commun **3**: 994

Gontijo AM, Miguela V, Whiting MF, Woodruff RC, Dominguez M (2011) Intron retention in the Drosophila melanogaster Rieske Iron Sulphur Protein gene generated a new protein. Nat Commun **2**: 323

Göhring J, Jacak J, Barta A (2014) Imaging of endogenous messenger RNA splice variants in living cells reveals nuclear retention of transcripts inaccessible to nonsense-mediated decay in Arabidopsis. Plant Cell **26**: 754–764

Grabher MG, Haas BJ, Yassour M (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol **29**: 644–652

Halpern KB, Caspi I, Lemze D, Levy M, Landen S, Elinav E, Ulitsky I, Itzkovitz S (2015) Nuclear retention of mRNA in mammalian tissues. Cell Rep **13**: 2653–2662

Hartmann L, Wießner T, Wachter A (2018) Subcellular compartmentation of alternatively spliced transcripts defines SERINE/ARGININE-RICH PROTEIN30 expression. Plant Physiol **176**: 2886–2903

Herman RC, Williams JG, Penman S (1976) Message and non-message sequences adjacent to poly(A) in steady state heterogeneous nuclear RNA of HeLa cells. Cell **7**: 429–437

Hossain MA, Claggett JM, Edwards SR, Shi A, Pennebaker SL, Cheng MY, Hasty J, Johnson TL (2016) Posttranscriptional regulation of Gcr1 expression and activity is crucial for metabolic adjustment in response to glucose availability. Mol Cell **62**: 346–358

Huang S, Deerinck TJ, Ellisman MH, Spector DL (1994) In vivo analysis of the stability and transport of nuclear poly(A)+ RNA. J Cell Biol **126**: 877–899

Hyjek M, Wojciechowska N, Rudzka M, Kołowerzo-Lubnau A, Smoliński DJ (2015) Spatial regulation of cytoplasmic snRNP assembly at the cellular level. J Exp Bot **66**: 7019–7030

Hyjek-Składanowska M, Bajczyk M, Gołębiewski M, Nuc P, Kołowerzo-Lubnau A, Jarmołowski A, Smoliński DJ (2020) Core spliceosomal Sm proteins as constituents of cytoplasmic mRNPs in plants. Plant J **103**: 1155–1173

Inagaki S, Numata K, Kondo T, Tomita M, Yasuda K, Kanai A, Kageyama Y (2005) Identification and expression analysis of putative mRNA-like non-coding RNA in Drosophila. Genes Cell Devot Mol Cell Mech **10**: 1163–1173

Jaillon O, Bouhouche K, Gout JF, Aury JM, Noel B, Saudemont B, Nowacki M, Serrano V, Porcel BM, Ségurens B, et al. (2008) Translational control of intron splicing in eukaryotes. Nature **451**: 359–362

Kiss T (2004) Biogenesis of small nuclear RNPs. J Cell Sci **117**: 5949–5951

Klingauf M, Staněk D, Neugebauer KM (2006) Enhancement of U4/U6 small nuclear ribonucleoprotein particle association in Cajal bodies predicted by mathematical modeling. Mol Biol Cell **17**: 4972–4981

Kołowerzo A, Smoliński DJ, Bednarska E (2009) Poly(A) RNA a new component of Cajal bodies. Protoplasma **236**: 13–19

Kołowerzo-Lubnau A, Niedojadło J, Świdziński M, Bednarska-Kozakiewicz E, Smoliński DJ (2015) Transcriptional activity in

diplotene larch microsporocytes, with emphasis on the diffuse stage. PLoS One **10**: e0117337

Lamond AI, Spector DL (2003) Nuclear speckles: a model for nuclear organelles. Nat Rev Mol Cell Biol **4**: 605–612

Liu Y, Gonzàlez-Porta M, Santos S, Brazma A, Marioni JC, Aebersold R, Venkitaraman AR, Wickramasinghe VO (2017) Impact of alternative splicing on the human proteome. Cell Rep **20**: 1229–1241

Majewska K, Wróblewska-Ankiewicz P, Rudzka M, Hyjek-Składanowska M, Gołębiewski M, Smoliński DJ, Kołowerzo-Lubnau A (2021) Different patterns of mRNA nuclear retention during meiotic prophase in larch microsporocytes. Int J Mol Sci **22**: 8501 doi: 10.3390/ijms22168501.

Matera AG, Wang Z (2014) A day in the life of the spliceosome. Nat Rev Mol Cell Biol **15**: 108–121

Misteli T (2007) Beyond the sequence: cellular organization of genome function. Cell **128**: 787–800

Morris GE (2008) The Cajal body. Biochim Biophys Acta **1783**: 2108–2115

Naro C, Jolly A, Di Persio S, Bielli P, Setterblad N, Alberdi AJ, Vicini E, Geremia R, De la Grange P, Sette C (2017) An orchestrated intron retention program in meiosis controls timely usage of transcripts during germ cell differentiation. Dev Cell **41**: 82–93.e4

Ninomiya K, Kataoka N, Hagiwara M (2011) Stress-responsive maturation of Clk1/4 pre-mRNAs promotes phosphorylation of SR splicing factor. J Cell Biol **195**: 27–40

Novotný I, Blažíková M, Staněk D, Herman P, Malinsky J (2011) In vivo kinetics of U4/U6·U5 tri-snRNP formation in Cajal bodies. Mol Biol Cell **22**: 513–523

Novotný I, Stejskalová E, Matějů D, Klimešová K, Roithová A, Švéda M, Knejzlík Z, Staněk D (2015) SART3-Dependent Accumulation of Incomplete Spliceosomal snRNPs in Cajal Bodies. Cell Reports **10**: 429–440

Peccarelli M, Kebaara BW (2014) Regulation of natural mRNAs by the nonsense-mediated mRNA decay pathway. Eukaryotic Cell **13**: 1126–1135

Perez-Riverol Y, Bai J, Bandla C, Hewapathirana S, García-Seisdedos D, Kamatchinathan S, Kundu D, Prakash A, Fsize-Zipper A, Eisenacher M, et al. (2022) The PRIDE database resources in 2022: a Hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res **50**: D543–D552

Prasanth KV, Prasanth SG, Xuan Z, Hearn S, Freier SM, Bennett CF, Zhang MQ, Spector DL (2005) Regulating gene expression through RNA nuclear retention. Cell **123**: 249–263

Saavedra C, Tung KS, Amberg DC, Hopper AK, Cole CN (1996) Regulation of mRNA export in response to stress in Saccharomyces cerevisiae. Genes Dev **10**: 1608–1620

Sadis S, Hickey E, Weber LA (1988) Effect of heat shock on RNA metabolism in HeLa cells. J Cell Physiol **135**: 377–386

Schmieder R, Edwards R (2011) Quality control and preprocessing of metagenomic datasets. Bioinformatics **27**: 863–864

Schneider CA, Rasband WS, Eliceiri KW (2012) NIH image to imageJ: 25 years of image analysis. Nat Methods **9**: 671–675

Shalgi R, Hurt JA, Lindquist S, Burge CB (2014) Widespread inhibition of posttranscriptional splicing shapes the cellular transcriptome following heat shock. Cell Rep **7**: 1362–1370

Shaw PJ, Brown JWS (2004) Plant nuclear bodies. Curr Opin Plant Biol **7**: 614–620

Smoliński DJ, Kołowerzo A (2012) mRNA accumulation in the Cajal bodies of the diplotene larch microsporocyte. Chromosoma **121**: 37–48

Smoliński DJ, Wróbel B, Noble A, Zienkiewicz A, Górska-Brylass A (2011) Periodic expression of Sm proteins parallels formation of nuclear Cajal bodies and cytoplasmic snRNP-rich bodies. Histochem Cell Biolo **136**: 527–541

Spector DL, Lamond AI (2011) Nuclear speckles. Cold Spring Harbor Perspect Biol **3**: a000646

Stanek D, Neugebauer KM (2006) The Cajal body: a meeting place for spliceosomal snRNPs in the nuclear maze. Chromosoma **115**: 343–354

Strzelecka M, Oates AC, Neugebauer KM (2010) Dynamic control of Cajal body number during zebrafish embryogenesis. Nucleus (Austin, Texas), **1**: 96–108

Takemura R, Takeiwa T, Taniguchi I, McCloskey A, Ohno M (2011) Multiple factors in the early splicing complex are involved in the nuclear retention of pre-mRNAs in mammalian cells. Genes Cells Devot Mol Cell Mech **16**: 1035–1049

Vanichkina DP, Schmitz U, Wong JJL, Rasko JEJ (2018) Challenges in defining the role of intron retention in normal biology and disease. Semin Cell Dev Biol **75**: 40–49

Wang Q, Sawyer IA, Sung MH, Sturgill D, Shevtsov SP, Pegoraro G, Hakim O, Baek S, Hager GL, Dundr M (2016) Cajal bodies are linked to genome conformation. Nat Commun **7**: 10966

Wegener M, Müller-McNicoll M (2018) Nuclear retention of mRNAs - quality control, gene regulation and human disease. Semin Cell Dev Biol **79**: 131–142

Yap K, Lim ZQ, Khandelia P, Friedman B, Makeyev EV (2012) Coordinated regulation of neuronal mRNA steady-state levels through developmentally controlled intron retention. Genes Dev **26**: 1209–1223

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