Connecting cis-elements and trans-factors with mechanisms of developmental regulation of mRNA translation in meiotic and haploid mammalian spermatogenic cells

Kenneth C Kleene

Department of Biology, University of Massachusetts Boston, 100 Morrissey Boulevard, Boston, Massachusetts 02125-3393, USA

Correspondence should be addressed to K C Kleene; Email: kenneth.kleene@umb.edu

Abstract

mRNA-specific regulation of translational activity plays major roles in directing the development of meiotic and haploid spermatogenic cells in mammals. Although many RNA-binding proteins (RBPs) have been implicated in normal translational control and sperm development, little is known about the keystone of the mechanisms: the interactions of RBPs and microRNAs with cis-elements in mRNA targets. The problems in connecting factors and elements with translational control originate in the enormous complexity of post-transcriptional regulation in mammalian cells. This creates confusion as to whether factors have direct or indirect and large or small effects on the translation of specific mRNAs. This review argues that gene knockouts, heterologous systems, and overexpression of factors cannot provide convincing answers to these questions. As a result, the mechanisms involving well-studied mRNAs (Ddx4/Mvh, Prm1, Prm2, and Sycp3) and factors (DICER1, CPEB1, DAZL, DDX4/MVH, DDX25/GRTH, translin, and ELAV1/HuR) are incompletely understood. By comparison, mutations in elements can be used to define the importance of specific pathways in regulating individual mRNAs. However, few elements have been studied, because the only reliable system to analyze mutations in elements, transgenic mice, is considered impractical. This review describes advances that may facilitate identification of the direct targets of RBPs and analysis of mutations in cis-elements. The importance of upstream reading frames in the developmental regulation of mRNA translation in spermatogenic cells is also documented.

Introduction

Spermatogenesis, the process of cell renewal and differentiation that produces male gametes, is divided into three main phases, each with distinct cell type and function, shown in Fig. 1 (Russell et al. 1990). The first is the mitotic phase in which diploid spermatogonia proliferate in the mitotic cycle, greatly amplifying the number of cells. This is followed by the meiotic phase in which two copies of the diploid chromosome set in spermatocytes are shuffled and reduced to a haploid set. During the final phase, spermiogenesis, haploid spermatids differentiate into highly polar cells specialized for motility and fertilization, spermatozoa. Spermatogenesis is extremely complex and occurs in many developmental stages (Fig. 1). Spermatocytes pass through six stages, distinguished by the morphology of the meiotic chromosomes, and mouse spermatids pass through 16 stages, known as steps, distinguished by the morphology of the developing acrosome, nucleus, and sperm tail (Russell et al. 1990).

Translational regulation plays an important role in controlling the development of spermatocytes and spermatids (reviewed in Paronetto & Sette (2010), Nguyen-Chi & Morello (2011) and Idler & Yan (2012)). The most familiar example originates in chromatin remodeling in which histones are replaced by transition proteins and protamines during late spermiogenesis (reviewed in Meistrich et al. (2003)). Round spermatids in steps 1–8 have round transcriptionally active nuclei. The nuclei of elongating spermatids in steps 9 through early 12 change from round to elongated accompanied by a drastic reduction in transcriptional activity (Kierszenbaum & Tres 1975). As a result, elongated spermatids in steps 12–16 differentiate without new mRNA transcripts. Many mRNAs are transcribed in spermatocytes and round spermatids and stored in a translationally inactive state in the absence of the encoded protein for several days to more than 2 weeks, before the protein is first detected in elongating and elongated spermatids (Fig. 1). In addition, translational control directs the beginning and end of the first meiotic prophase, and all mRNAs in pachytene spermatocytes and round spermatids are partially repressed by an unknown global mechanism (discussed below).

mRNA-specific translational regulation in eukaryotic cells is usually mediated by cis-elements that bind...
trans-factors, RNA-binding proteins (RBPs), or small non-coding RNAs, sncRNAs (Jackson et al. 2010). By binding elements, factors promote or block the recruitment of other factors and ribosomes to mRNAs, thereby decreasing, negative control, or stimulating, positive control, the translation of mRNA-coding sequences into protein. However, one type of cis-element discussed below, upstream reading frames, does not necessarily involve the binding of factors to elements.

The importance of translational regulation in spermatogenesis is further demonstrated by mutations in a large number of genes, mostly knockout models, which cause abnormal sperm development, reductions in male fertility, and failure to express proteins at the normal developmental stage (Table 1). Although many reviews discussing these and related subjects cited here affirm translational control as an important problem in spermatogenesis, little is known about the basic features of the mechanisms: the interactions of factors with elements and the magnitude of the change in translational activity mediated by factors and elements. This review focuses on systems in which elements and factors have and have not been connected, strategies that may facilitate the identification and validation of elements, and important controversies to be resolved in future work.

### Elucidating mechanisms of translational control with gene knockouts, overexpression of factors, heterologous systems, and mutations in elements

The challenges in connecting mRNA targets with regulatory factors can be traced to a major insight: post-transcriptional regulation of gene expression in...
eukaryotic cells is much more complex than anyone could have predicted 10 years ago (reviewed in Bartel (2009), Kishore et al. (2010) and Morris et al. (2010)). Post-transcriptional gene regulation in mammalian cells is realized at the levels of alternative splicing, polyadenylation site selection, mRNA export, translation, stability, and localization, and individual mRNAs are often regulated at multiple levels. These processes are mediated by the activities of RBPs, and snRNAs, the best studied of which are microRNAs.

Table 1 Factors implicated in translational control in spermatocytes and spermatids.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Knockout phenotype</th>
<th>mRNA targets (evidence)</th>
<th>Positive/ negative</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADAR1/TENR (dsRBD)</td>
<td>Infertile, oligospermia, spermiation, nuclear shape, flagellum, mitoity throughpout, azoosperma, total block step 6, acrosome</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Connolly et al. (2005)</td>
</tr>
<tr>
<td>Boll/Boule (ssRNA, RRM)</td>
<td>Infertile, azoosperma, total block step 6, acrosome</td>
<td>Unknown</td>
<td>Pos</td>
<td>VanGompel &amp; Xu (2010)</td>
</tr>
<tr>
<td>CELF3/BRUNOL1 (ssRNA, RRM)</td>
<td>Infertile, mild oligosperma, no morphological defects</td>
<td>Unknown</td>
<td>?</td>
<td>Dev et al. (2007)</td>
</tr>
<tr>
<td>CELF1/CUGBP1 (ssRNA, RRM)</td>
<td>Normal to intertile, step 7 block, apoptotic round spermatids</td>
<td>Unknown</td>
<td>?</td>
<td>Kress et al. (2007)</td>
</tr>
<tr>
<td>DDX4/MVH (ssRNA, DEAD-box RNA helicase)</td>
<td>Infertile, azoosperma, total block near leptotene–zygot o gene</td>
<td>Pmp1, Pmp2, Trnp1, and Tnp2 histone H4, many others (IG, IP, MP)</td>
<td>Pos</td>
<td>Tsai-Morris et al. (2004, 2012)</td>
</tr>
<tr>
<td>DDX25/GRT1H (DEAD-box RNA helicase)</td>
<td>Infertile, azoosperma, massive meiotic metaphase I apoptosis, total block step 8, nuclear elongation</td>
<td>Hsp2a (IG)</td>
<td>Pos</td>
<td>Sun et al. (2010)</td>
</tr>
<tr>
<td>EIF4G3 (binds EIF4E and PABPC1)</td>
<td>Infertile, azoosperma, total block before metaphase I</td>
<td>Hsp2a (IG, many)</td>
<td>Pos</td>
<td>Chi et al. (2011)</td>
</tr>
<tr>
<td>ELAVL1/HuR (ssRNA, RRM)</td>
<td>Infertile, azoosperma, apoptosis at or before meiotic division, nuclear elongation</td>
<td>Trp2, Prm2 (IG in 3T3 cells)</td>
<td>Neg</td>
<td>Dai et al. (2011)</td>
</tr>
<tr>
<td>miRNA469 (microRNA)</td>
<td>Not studied</td>
<td>Unknown</td>
<td>Neg</td>
<td>Yang et al. (2007)</td>
</tr>
<tr>
<td>YBX2/MSY2 (ssRNA, CSD)</td>
<td>Infertile, azoosperma, minor pachyten e and massive step 8 apoptosis, nuclear elongation and condensation, flagella, spermiation, mRNA degradation</td>
<td>Unknown</td>
<td>Neg</td>
<td>Lu et al. (2006)</td>
</tr>
<tr>
<td>YBX3/MSY4/MSY3L (ssRNA, CSD)</td>
<td>Infertile or fertile (line dependent), pachyten e apoptosis, no morphological defects</td>
<td>All? Pmp1, Pmp2, Trnp1 (IG)</td>
<td>Neg and Pos</td>
<td>Yanagiya et al. (2010)</td>
</tr>
<tr>
<td>PABPC1 (ssRNA, RRM)</td>
<td>Not studied</td>
<td>All? Pmp1, Pmp2, Trnp1 (IG)</td>
<td>Pos and Neg</td>
<td>Yanagiya et al. (2010)</td>
</tr>
<tr>
<td>PAIP2a (binds PABPC1)</td>
<td>Infertile, severe oligosperma, round spermatid apoptosis, nuclear elongation flagella, spermiation</td>
<td>Spag16, Neet1, Spedys (IG, IP)</td>
<td>Pos</td>
<td>Paronetto et al. (2009)</td>
</tr>
<tr>
<td>KHDRBS1/SAM68 (ssRNA, KH)</td>
<td>Infertile, azoosperma or severe oligosperma, pachyten e and round spermatid apoptosis, flagella, nuclei</td>
<td>Unknown</td>
<td>?</td>
<td>Pires-daSilva et al. (2001) Zhong et al. (1999)</td>
</tr>
<tr>
<td>TARBP2/PRBP (dsRNA, dsRBP)</td>
<td>Infertile, severe oligosperma, round spermatid apoptosis, spermiation, nuclear elongation and condensation</td>
<td>Pmp1, Pmp2 (IG, indirect?)</td>
<td>Pos</td>
<td>Kashiiwabara et al. (2002)</td>
</tr>
<tr>
<td>PAPI/TTPAP (unknown)</td>
<td>Infertile, azoosperma, total block 8 block pachyten e apoptosis, step 7 block</td>
<td>Unknown</td>
<td>Pos?</td>
<td>Chennathuzukuzhi et al. (2003)</td>
</tr>
<tr>
<td>TSN/Translin (ssRNA)</td>
<td>Fertile, mild oligospermia, pachyten e apoptosis, no morphological defects</td>
<td>Pmp2, Gapdhst, many? (RRL)</td>
<td>Neg?</td>
<td>Kashiiwabara et al. (2002)</td>
</tr>
</tbody>
</table>

*The official name of the protein is according to the Mouse Genome Informatics website and common synonyms. The information in parentheses lists the preference for binding single-stranded or double-stranded RNA and the type of RNA-binding domain. The RNA-binding domains include RRM, RNA recognition motif; CSD, cold shock domain; KH, KH-domain; dsRBP, double-stranded RNA-binding domain; ZFD, zinc-finger domain. It is difficult to summarize the phenotypes of the knockouts accurately within the space limitations of a table, since the phenotypes are influenced by genetic background and exhibit different effects on the differentiation of various organelles and biases due to differences in the depth of analysis. The descriptions of the phenotypes in the primary literature have been reproduced above, but question marks have been inserted when the developmental abnormalities have not been described precisely. For knockouts that cause partial blocks, the organelles or developmental processes that exhibit abnormalities are listed. The mRNAs listed are proposed to be direct targets of the knocked-out factor. The information in parentheses describes the evidence: IG, in vivo genetic effects based on gene knockouts; u.v., u.v. cross-linking assays; IP, co-immunoprecipitation; IPPM, immunoprecipitation of polysomes and microarray analysis; RRL, rabbit reticulocyte. References of the articles containing supporting information.*
miRNAs. It is estimated that the mammalian genome encodes about 400 miRNAs and 600 RBPs recognized by various types of RNA-binding domains (RBDs) (Bartel 2009, Ascano et al. 2012). However, the number of RBPs is potentially much greater because the genome encodes an additional ∼1400 uncharacterized proteins with domains that can bind RNA and/or DNA (Morris et al. 2010, Ascano et al. 2012). In addition, polyadenylated mRNAs in somatic cells bind several hundred proteins that lack recognizable RBDs (Baltz et al. 2012, Castello et al. 2012).

The current paradigm is that individual RBPs and miRNAs interact with many, sometimes thousands of, mRNAs and that the expression of individual mRNAs is regulated cooperatively by multiple RBPs and miRNAs (Bartel 2009, Kishore et al. 2010, Morris et al. 2010). However, the functions of RBPs and miRNAs in post-transcriptional regulation in somatic mammalian cells differ significantly. miRNAs base-pair with target mRNAs and guide mRNA degradation, translational repression, or both, but the decrease in the expression of individual proteins is usually slight, ‘fine-tuning’ (Bartel 2009, Fabian & Sonenberg 2012). In contrast, RBPs increase or decrease mRNA degradation and translation, by magnitudes ranging from small to large.

Another striking difference between RBPs and miRNAs is that RBPs often coordinate the expression of proteins with related functions accompanying the differentiation of specific cell types or changes in the physiological state (Kishore et al. 2010, Morris et al. 2010). By comparison, miRNAs usually fine-tune the expression of large numbers of proteins with unrelated functions (Bartel 2009, Fabian & Sonenberg 2012) and more rarely the expression of proteins with closely related functions (Mestdagh et al. 2010).

About 20 factors have been implicated in translational regulation by gene knockouts, most of which have deleterious effects on sperm development (see Table 1). Pinpointing the defects in post-transcriptional regulation that result from knockouts in spermatogenic cells is remarkably difficult, because of problems in distinguishing between mRNA targets that are regulated directly and indirectly by a factor. Specifically, many factors regulate groups of target mRNAs directly, but some of the direct mRNA targets encode factors that regulate secondary targets, creating confusion as to whether the effect of the knockout is direct or indirect. This problem is magnified by the interactions of factors with very large numbers of mRNAs and functions in multiple post-transcriptional processes.

The problems in relating phenotypic effects to direct and indirect targets are illustrated by two RBPs with important functions in spermatogenesis, YBX2/MSY2 and ELAVL1/HuR (Yang et al. 2007, Chi et al. 2011). YBX2/MSY2 belongs to the family of Y-box proteins that interacts with thousands of, and potentially all, mRNAs in somatic cells and functions in transcription, splicing, positive and negative regulation of translation, stability, and mRNP structure (Skabkin et al. 2006). Similarly, ELAVL1/HuR binds U-rich elements in the 3’ UTRs and introns of thousands of mRNAs in cultured mammalian cells, enhancing stability and translation by mechanisms that are sometimes coupled to pre-mRNA processing (Lebedeva et al. 2011, Mukherjee et al. 2011). As some of the most responsive mRNAs to siRNA knockdown of ELAVL1/HuR encode proteins with functions in regulating gene expression, ELAVL1/HuR has been described as ‘a regulator of regulators’ (Mukherjee et al. 2011).

Distinguishing between the direct and indirect effects of knockouts is especially difficult in spermatogenic cells, because RBPs that are necessary for spermatogenesis are usually expressed for more than a week, a period during which the cells undergo profound developmental changes (Fig. 1). In fact, the earliest phenotype of knockouts is often increased apoptosis, followed days later by morphological abnormalities (examples are shown in Table 1). These compound phenotypes could be produced by defective regulation of two sets of genes, early and late, with independent or dependent functions. Obviously, a knockout that blocks sperm development before the stage at which an mRNA is normally activated or repressed in wild-type mice precludes inferring that the knocked-out factor directly regulates that mRNA.

Knockouts often produce different abnormalities in different cells at the same stage in a single testis (Zhong et al. 1999, Dass et al. 2007, Reynolds et al. 2007) or different phenotypes on different genetic backgrounds (Connolly et al. 2005, Reynolds & Cooke 2005). These phenotypes are the hallmarks of incomplete and variable expressivity and penetrance, indicative of modifying factors that compensate for or intensify the phenotype created by the absence of a factor.

Studies of the overexpression of factors are also difficult to interpret because elevated levels of a factor can produce regulatory effects that differ from the effects of normal levels. This is clearly illustrated by findings that high levels of PABPC1 repress mRNA translation in the Krebs cell-free translation lysate and round spermatids, in contrast to the usual function of lower levels of PABPC1 in stimulating translation (Yanagiya et al. 2010). Similarly, depletion of ELAV1/HuR with a conditional knockout decreases Hspa2 mRNA translation in late spermatocytes (Chi et al. 2011), in agreement with a well-established function of this RBP in translational activation in somatic cells (Lebedeva et al. 2011). In contrast, the overexpression of ELAV1/HuR in spermatogenic cells in transgenic mice also decreases Hspa2 mRNA translation, consistent with direct or indirect functions in translational repression (Chi et al. 2011). In addition, the overexpression of Y-box protein YBX3/MSY4 in round and elongated spermatids produces deformed spermatozoa (Giorgini et al. 2002), whereas the Ybx3/MSy4 knockout does not (Lu et al. 2006).
The functions of factors in translational regulation in spermatogenic cells have also been inferred from studies in cultured somatic cells, *Xenopus* oocytes and the rabbit reticulocyte cell-free lysate (RRL). Heterologous systems can produce interpretable results when the interaction between a factor and an element is sufficient to regulate translation, sometimes with the assistance of factors conveniently present in the heterologous cells (Ling et al. 2002, Collier et al. 2005). However, this approach is particularly risky with spermatogenic cells, which express many factors that differ qualitatively and quantitatively from those in somatic cells and oocytes (references in Kleene (2003), Nguyen-Chi & Morello (2011) and Idler & Yan (2012)). In some cases, the addition of a factor at a superphysiological level may alter translational activity only at excessive levels (see above) or the activity of the factor may depend on post-translational modifications that are not fulfilled in the heterologous milieu. There are several reports of misleading results in the RRL (Kleene & Smith 1994, Svitkin et al. 2009). A striking example is that TARBP2 is a non-specific translational repressor in the RRL, while the gene knockout implicates TARBP2 in a pathway that specifically activates *Prm1* and *Prm2* mRNA translation (Lee et al. 1996, Zhong et al. 1999). In general, heterologous systems cannot elucidate the magnitude of the regulatory effects of elements and factors on mRNA translation in spermatogenic cells.

Numerous elements in flies, worms, *Xenopus* oocytes and cultured mammalian cells have been identified with mutations that abrogate translational control in vivo. Mutations in elements eliminate uncertainties concerning the identity of the target mRNA and have often led to evidence that a short mRNA sequence is a translational control element (TCE) that binds a specific factor. Mutations have also provided evidence that elements mediate distinctive mechanisms involving the length of the poly(A) tail, recruitment of PABP without a poly(A) tail, a block to initiation of translation at the level of free mRNPs, subunit joining or reinitiation, regulation in a specific cell type or developmental stage, translation of upstream open reading frames (uORFs), or spatial relations between elements (Mueller & Hinnebusch 1986, McGrew et al. 1989, Shyu et al. 1989, Levy et al. 1991, Ostareck et al. 2001, Collier et al. 2005, Lyabin et al. 2011). As mutations in elements that inactivate individual regulatory pathways can usually be created, they can provide decisive evidence whether a specific factor has a large or small effect on the translational activity of a single mRNA and about the relative importance of multiple pathways in regulating a single mRNA and the importance of the position of elements (Shyu et al. 1989, Levy et al. 1991, Lyabin et al. 2011).

Studies of mutations in elements can also be problematic. They are less likely than knockouts to produce phenotypes. Deletions can produce misleading phenotypes by altering the positions of elements (see comments regarding the position of the *Prm1* Y-box recognition sequence (YRS) below) or by abrogating the binding of multiple factors. The huge number of factors involved in post-transcriptional regulation makes it likely that replacing a UTR with a heterologous UTR will introduce extraneous elements with unpredictable regulatory effects. In addition, elements may be present in multiple copies in individual UTRs, some or all of which must be mutated to understand their function, and the effects of mutating one element may only be realized in the presence of other elements in the same or a different UTR (Mehta et al. 2006, Bartel 2009). In general, mutations that produce strong phenotypes are easy to analyze experimentally, and mutations in small numbers of bases minimize the chance of abrogating the binding of multiple factors.

It is also desirable to identify mutations in elements that result in the total loss of negative or positive translational control of an individual mRNA at a specific stage of spermatogenesis. Mutations of positive control elements are yet to be studied.

Notably, mutations in negative control elements do not result in the total loss of translational repression in round spermatids. In somatic mammalian cells, >85% of fully active mRNAs sediment with polysomes in sucrose gradients (Schmidt et al. 1999, Mathews et al. 2007). In contrast, mutation or deletion of negative TCEs in the 3′ UTRs of two mRNAs that are strongly repressed in round spermatids, *Prm1* and *Smcp*, results in relatively low levels of polysomal mRNA, 35–55% (Schmidt et al. 1999, Zhong et al. 2001, Hawthorne et al. 2006, Bagarova et al. 2010). In addition, each of the more than 60 active mRNAs in pachytene spermatocytes and spermatids that have been analyzed with sucrose gradients exhibit 25–55% polysomal mRNA in total testsis (Kleene 2003, Paronetto et al. 2009, Kleene et al. 2010, Chi et al. 2011). Many of these partially repressed mRNAs, exemplified by the *Ldhc* mRNA, exhibit a similar polysomal loading in pachytene spermatocytes and round spermatids (Kleene 2003). Thus, the strong initial repression of developmentally regulated mRNAs appears to be superimposed on a ubiquitous background of partial global repression. The factors that are potentially involved in the partial global repression include YBX2/MSY2, PABPC1, EIF4E, EIF4BP1, and ribosomal protein SP6 (Schmidt et al. 1999, Messina et al. 2010, Yanagiya et al. 2010). At present, the contribution of each of these factors to the partial repression of all mRNAs and the strong initial repression of developmentally regulated mRNAs is unknown.

Unfortunately, few mutations in regulatory elements in spermatogenic cells have been analyzed, because cell culture and cell-free translation systems that reproduce the patterns of translational control in spermatogenic cells have not been authenticated. Mutations in transgenic mice have rarely been studied, probably because the approach is expensive, laborious, and prone
to negative results. Potential solutions to these problems are described below.

Tables 1 and 2 list factors and elements that have been implicated in translational regulation in mammalian spermatogenesis by studies of gene knockouts, heterologous systems, and transgenic mice. Significantly, most of these factors have not been connected with targets. The following sections scrutinize the small number of systems in which elements have been implicated in translational control in meiotic and haploid spermatogenic cells.

**Elements in the Prm1 3’ UTR: TCE and YRS**

The Prm1 mRNA is first detected in step 7 spermatids and is repressed in free mRNPs until the Prm1 protein is first detected in step 10 spermatids, see Fig. 1 (Kleene 1989, Mali et al. 1989, Meistrich et al. 2003). In classic studies using the Prm1 promoter and 5’ UTR and human GH (hGH) reporter, Braun et al. (1989) demonstrated that the timing of Prm1 mRNA translation is regulated by the 3’ UTR in transgenic mice. Subsequent studies of deletion and point mutations identified two elements in the Prm1 3’ UTR that repress translation in early spermatids. However, it is important to note that both Ybx3 and Msy4 fail to acknowledge that the Ybx3 gene is expressed as two alternatively spliced mRNAs (Mastrangelo & Kleene 2005) encoding isoforms of different sizes and similar abundance, detected by western blots of testicular proteins (Davies et al. 2000, Giorgini et al. 2002), which are referred to as YBX3L (long) and YBX3S (short). Both isoforms are conserved in multiple mammalian species, implying that each isoform has at least partially distinct functions.

Fajardo et al. (1997) studied an hGH–Prm1 mRNA in which 95 nts were deleted from the middle of the Prm1 3’ UTR, thereby moving the YRS from its natural position, 109 nts upstream of the poly(A) signal, to an abnormal position, 15 nts upstream of the poly(A) signal. This 3’ UTR, with the YRS in an abnormal position, delays hGH expression in transgenic mice, but the initial repression is incomplete because low levels of hGH are detected in round spermatids (Fajardo et al. 1997). Giorgini et al. (2001) subsequently showed that a mutation in this YRS, which abrogates the binding of Y-box proteins YBX3L/MSY4 and YBX2/MSY2 in vitro, results in translational repression in early spermatids. However, a transgene in which the TCEs are mutated in the full-length Prm1 3’ UTR abrogates translational repression in early spermatids.
spermatids, suggesting that the YRS in its natural position is insufficient to repress translation (Zhong et al. 2001).

The literature to date has revealed that Prm1 mRNA translation in round spermatids is repressed by an unidentified factor that binds the TCE and that YBX2/MSY2 and YBX3/MSY4 do not repress translation by binding the YRS in its natural position (Zhong et al. 2001). The Prm1 mRNA is also activated by pathways involving the DEAD-box RNA helicase, DDX25/GRTH, and the protein kinase/double-stranded RBP, TARBP2 (Zhong et al. 1999, Tsai-Morris et al. 2012). However, none of the pathways, direct or indirect, that repress and activate Prm1 mRNA translation are known.

**uORFs in the Smcp and c-Mos 5′ UTRs**

The 5′ UTRs of 30–50% of human mRNAs contain uORFs, characterized as reading frames that begin with an AUG codon in the 5′ UTRs and are discontinuous with the authentic coding sequence (Hinnebusch 2005). The translation of uORFs promotes the release of the ternary complex (eIF2–GTP–tRNAiMet) from the 40S ribosomal subunit, which prevents ribosomes from re-initiating translation at the authentic ORF downstream. Since uORFs are translated, repression by uORFs produces small polysomes instead of free mRNPs.

The repression of mRNA translation by uORFs in various human mRNAs varies from ~30% (fine-tuning) to ~100% (total repression) and is often accompanied by small decreases in mRNA levels (Calvo et al. 2009). The strength of repression is determined by many parameters such as secondary structure, rare codons, context of the uORF initiation codon, sequence of the encoded peptide, length and position of the uORF, and trans-acting factors (Sachs & Geballe 2006, Medenbach et al. 2011). In general, parameters that prolong the translation of uORFs promote the release of the ternary complex and decrease the translation of the authentic ORF. Conversely, parameters that prolong the period between termination of the uORF and initiation at the authentic initiation codon promote rebinding of the ternary complex to the scanning 40S subunit and increase the translation of the authentic ORF. Although the parameters that determine the strength of repression by uORFs are too complex to predict by examining the primary sequence of 5′ UTRs, the functions of uORFs in translational repression can easily be demonstrated by mutating the uORF AUG initiation codon to a non-start codon. The uORFs of two mRNAs that are expressed in spermatids have been demonstrated to developmentally regulate gene expression in transgenic mice.

The SMCP protein is first detected in step 11 spermatids 6 days after the mRNA is detected in step 3 spermatids, see Fig. (Shih & Kleene 1992, Cataldo et al. 1996). Translational control of the Smcp mRNA has been studied in transgenic mice using the Smcp promoter, the Gfp-coding region, and the permutations of wild-type and mutant Smcp and Gfp UTRs. The translational repression of the Smcp mRNA in early spermatids is primarily mediated by the 3′ UTR, which produces free mRNPs and delays GFP expression by ~4 days. By comparison, the 5′ UTR produces a one- to two-day delay and small polysomes (Hawthorne et al. 2006, Bagarova et al. 2010). The initial repression by the Smcp 5′ UTR is mediated by two short uORFs, because mutating the uORF AUG codons to AGG eliminates the delay in GFP expression (Bagarova et al. 2010). The uORFs are also involved in the developmental control of Smcp mRNA translation, because the Smcp 3′ UTR exerts a positive control that overcomes the small polysomes mediated by the uORFs in late spermatids (Bagarova et al. 2010, TA Chordhury and KC Kleene, unpublished observations).

Steel et al. (1996) studied the spermatid-specific 5′ UTR of the Mos/c-Mos mRNA using the β-galactosidase reporter under the control of the constitutive cytomegalovirus IE promoter. A transgene containing the wild-type Mos 5′ UTR, which contains four short uORFs, restricts β-galactosidase expression to spermatids, whereas a transgene in which the upstream AUG initiation codons are mutated to AAG restricts β-galactosidase expression to the periphery of seminiferous tubules, presumably spermatogonia and spermatocytes. It is unclear whether the striking effect of the Mos uORFs on the stage of β-galactosidase expression is mediated by differences in mRNA translation or stability.

Students of translational control in spermatogenic cells have paid little attention to uORFs even though uORFs are a common, versatile, and easily recognized TCE. The Akap4, Gapdh, Spata18, Smcp, Spata19, and primate HILS1 5′ UTRs contain uORFs that are conserved in diverse mammalian species (Chowdhury & Kleene 2012, KC Kleene, unpublished observations). Interestingly, the uORFs of five of these mRNAs have features that correlate with strong translational repression: the termination codons of the Gapdh and Spata19 uORFs are just upstream or downstream of the authentic AUG codon, the Akap4 uORFs are especially long, and the Spata18 and Gapdh uORFs are predicted to bind Y-box proteins and translain respectively (Yang et al. 2003, Chowdhury & Kleene 2012). Mutations have not been studied to determine whether the uORFs have large or small effects on the developmental expression of these mRNAs in spermatids.

**DAZL, the deleted in azoospermia-like RBP**

The deleted in azoospermia, DAZ, family of proteins is named for a cluster of four genes on the Y-chromosome that has been implicated in 10–15% of cases of human male infertility (reviewed in Reynolds & Cooke (2005) and VanGompel & Xu (2011)). DAZ genes are only present in primates, including old-world monkeys and humans, but other vertebrates have two autosomal paralogs, deleted in azoospermia-like (DAZL)
and BOLL/BOULE, and all invertebrates likely have BOLL orthologs.

On a mixed background, the homozygous Dazl knockout produces a final block to the development of spermatocytes marked by the failure of chromosome synapsis during the leptotene–zygote transition (Saunders et al. 2003). The mRNAs encoding a major protein in the lateral element of the synaptonemal complex, SYCP3, and the DEAD-box RNA helicase, mouse Vasa homolog (DDX4/MVH), have been identified as targets of DAZL (Venables et al. 2001, Reynolds et al. 2005). The Sycep3 and Ddx4/Mvh 3’ UTRs contain U-rich DAZL-binding elements (discussed below). The Sycep3 and Ddx4/Mvh knockouts also produce blocks at or near the leptotene–zygote transition respectively similar to the stage of final arrest in Dazl-null mice (Tanaka et al. 2000, Yuan et al. 2000).

DAZL activates Sycep3 and Ddx4/Mvh mRNA translation, because the Dazl knockout reduces DDX4/MVH and SYCP3 protein expression in spermatocytes, without altering the levels of either mRNA (Reynolds et al. 2005, 2007). In addition, injection of Xenopus laevis oocytes with mouse Dazl mRNA enhances the translation of luciferase reporter mRNAs containing the Ddx4/Mvh and Sycep3 3’ UTRs, and the stimulation of translation is abrogated by mutation of the DAZL-binding sites (Reynolds et al. 2005, 2007).

Collier et al. (2005) analyzed the mechanism by which DAZL stimulates mRNA translation with tethering assays in Xenopus oocytes. In these tethering assays, synthetic mRNAs encoding fusions of five DAZ family proteins with the bacteriophage MS2 coat protein were injected into Xenopus oocytes followed by injections of luciferase reporter mRNA containing 3’ UTR MS2 binding sites. Tethering simplifies the assay of translational activity of reporter mRNAs by circumventing the process by which DAZL normally binds its target mRNAs. Xenopus DAZL, mouse DAZL, human DAZL, human DAZ, and human Boule stimulate luciferase activity without significant effects on mRNA stability. Collier et al. (2005) further demonstrated that all five DAZ family proteins bind the C-terminal domain of Xenopus poly(A)-binding protein, PABP (the ortholog of the mammalian cytoplasmic PABP known as PABPC1), and mapped the segment of DAZL that binds the PABP C-terminal domain with deletion mutations using pull-down and yeast two-hybrid assays. The analysis of DAZL deletions in tethering assays demonstrated that binding of DAZL to PABP is necessary for the stimulation of translation. Thus, DAZL activates translation by binding a 3’ UTR element and recruiting initiation factor PABP independent of the poly(A) tail.

As summarized in Table 2, DAZL binds a loosely defined set of elements in various mRNAs that are rich in U and contain the trinucleotide GUU. Indeed, the identification of the DAZL-binding elements is the most convincing of those surveyed here because it is based on consistent findings with a variety of assays including in vivo u.v. cross-linking and mutations that reduce binding of the DAZL protein and stimulation of translation in tethering assays (Venables et al. 2001, Reynolds et al. 2005). Surprisingly, quantitative binding assays demonstrate that the affinity of DAZL for GUU is independent of the surrounding bases (Jenkins et al. 2011). The brevity of this element complicates identifying the targets of DAZL, because GUU appears in many mRNAs, most of which presumably are not strongly upregulated by DAZL. This conundrum might be resolved by evidence that multiple DAZL-binding elements in a single 3’ UTR augment translational stimulation (Collier et al. 2005). Alternatively, the stimulation by DAZL may be modulated by additional factors.

**CPEB1: the cytoplasmic polyadenylation element-binding protein**

The cytoplasmic polyadenylation element (CPE), usually UUUAUUAUUU, and the CPE-binding protein (CPEB1) form the core of mechanisms that repress and activate mRNA translation in Xenopus oocytes (reviewed in Radford et al. (2008)). Translationally regulated mRNAs are synthesized in the nucleus with long poly(A) tails. Following export to the cytoplasm, the poly(A) tails are initially shortened to ~20 As, resulting in translational repression, and later lengthened to ~150 As, resulting in translational activation. CPEB1, CPE, and proximity of the CPE and AUAAUUAAU polyadenylation signal are required for both the poly(A) shortening that represses translation by limiting PABP recruitment and the poly(A) lengthening that activates translation by recruiting PABP.

The *Cpeb1* mRNA is expressed in pachytene spermatocytes and round spermatids, but the developmental expression of the protein has not been analyzed (Tay & Richter 2001). The homozygous *Cpeb1* knockout reduces male fertility about 16-fold, accompanied by increased apoptosis and chromatin fragmentation and large decreases in the number of post-meiotic cells and levels of two major synaptonemal complex proteins, SYCP1 and SYCP3. The final stage at which spermatocytes develop normally in *Cpeb1*-null mice has not been determined.

The *Sycep3* mRNA appears to be regulated directly by CPEB1 in wild-type oocytes because its CPE, UUUUAUAAU, efficiently u.v.-cross-links to *Xenopus* CPEB1, and translational activation is accompanied by poly(A) lengthening. The evidence that CPEB1 activates the *Sycep1* mRNA in oocytes is much weaker because its CPE, UUUUAUAAAUUU, is a poor match for the CPE consensus sequence and u.v.-cross-links inefficiently to *Xenopus* CPEB1. More importantly, there is little difference in the length of the *Sycep1* poly(A)s in the ovaries of wild-type and *Cpeb1* knockout mice.

Unfortunately, the poly(A) tails on the *Sycep1* and *Sycep3* mRNAs were not measured in *Cpeb1*-null and...
DAZL, CPEB1, and DDX4/MVH probably control the initiation of meiosis by different mechanisms

Gene knockouts reveal that a trio of well-established activators of mRNA translation, CPEB1, DAZL, and DDX4/MVH, are necessary to initiate meiosis in male mice. The Dazl knockout blocks the transition from leptotene to zygotene (Saunders et al. 2003). The Ddx4 and Cpeb1 knockouts also block early meiosis, but the morphology of the chromosomes that distinguish the normal and abnormal stages has not been defined (Tanaka et al. 2000, Tay & Richter 2001).

DAZL and CPEB1 are thought to stimulate translation by different mechanisms. DAZL stimulates translation by recruiting PABPC to the body of the 3’ UTR, while CPEB1 stimulates translation by poly(A) lengthening recruiting PABPC (Radford et al. 2008, Brook et al. 2009). The mouse Sycp3 mRNA contains CPE- and DAZL-RNA-binding elements. Interestingly, the Ddx4 3’ UTRs in many mammalian species including mouse contain consensus CPEs ~21–24 nt upstream of the poly(A) signal (Kleene 2013 unpublished). Thus, translation of the Sycp3 and Ddx4 mRNAs is potentially stimulated by CPEB1 as well as DAZL.

Assigning a mechanism to DDX4 is difficult because DDX4 has multiple functions. DDX4 belongs to the family of DEAD-box RNA helicases that melts secondary structure and often stimulates the translation of large groups of mRNAs (reviewed in Jankowsky 2011 and Parsyan et al. 2011). In addition, DDX4 functions in the assembly of germ-cell RNP granules and blocks the proliferation of retrotransposons, processes with potential indirect effects on post-transcriptional gene expression (Kirino et al. 2010, Kuramochi-Miyagawa et al. 2010).

Interestingly, the Drosophila ortholog of DDX4, Vasa, is an atypical RNA helicase because it stimulates the translation of the mei-P26 mRNA by binding a U-rich 3’ UTR element and recruiting the initiation factor Eif5b (Liu et al. 2009, Parsyan et al. 2011). The sequence-specific binding of Vasa is thought to be mediated by an RNA-binding motif, RGG, which appears eight times in the ~200-amino acid-long N-terminal segments of Vasa, outside of the RNA helicase domain (Liu et al. 2009). Four RGG motifs are present in the N-terminal segment of mouse and human DDX4, and the IUPred algorithm (http://iupred.enzim.hu) predicts that the N-terminal segments of Vasa and DDX4 are unstructured (Kleene 2013 unpublished), a feature shared by many RBPs (Castello et al. 2012). Thus, DDX4 and Vasa could activate mRNA translation by two mechanisms, by binding specific sequences in mRNAs and by melting the secondary structure. A related question is whether the two isoforms of DDX4 and Vasa with RGG motifs in different contexts differ functionally.

In adult testis, 858 potential mRNA targets of DDX4 have been identified by immunoprecipitation of native mRNPs with DDX4 antibody and microarray analysis (Nagamori et al. 2011), a technique known as RIP-chip. The identification of direct RNA targets of DEAD-box helicases is considered difficult (Jankowsky 2011) and may be complicated by the fact that DDX4 interacts with many RBPs. Most of these putative targets are expressed in pachytene spermatocytes and spermatids. Since the Ddx4 knockout blocks early meiosis (Tanaka et al. 2000), it cannot be used to determine whether DDX4 is necessary for translational regulation in pachytene spermatocytes and spermatids. mRNAs encoding proteins that initiate meiosis have not been identified, probably because they are present at low levels in adult testis. Perhaps, early meiotic targets of DDX4 could be identified in pure populations of early spermatocytes using in vivo u.v. crosslinking and immunoprecipitation discussed below, which appears to be able to capture mRNAs that are in direct contact with RNA helicases in vivo (Baltz et al. 2012, Castello et al. 2012).

Defining the roles of DDX4, CPEB1, and DAZL in the expression of the other two factors and meiotic proteins such as SYCP1 and SYCP3 would greatly increase our understanding of the mechanisms that control the initiation of meiosis in spermatocytes. The function of CPEB1 in activating the translation of the Sycp1, Sycp3, and Ddx4 mRNAs in early meiosis could easily be determined by measuring poly(A) lengths with ePAT (Janicke et al. 2012) and polysomal loading in leptotene and zygotene spermatocytes in wild-type and Cpeb1-null mice. However, determining the magnitude of stimulation by each factor will probably require in vivo genetic studies of mutations in elements specific for DDX4, CPEB1, and DAZL.

The testis–brain RBP, translin, and TSNAX/TRAX

The testis–brain RBP (TB-RBP) was initially proposed to repress Prm2 mRNA translation in round spermatids by specifically binding two elements named Y and H in a 40-nt segment of the Prm2 3’ UTR referred to as transcript c: 5’ CUGAGCCCCUGAGCCUGCAAGGAGCCACGAGACUGAC (Kwon & Hecht 1991, 1993). The Y- and H-elements are single and double underlined respectively. Subsequent work has demonstrated that TB-RBP is the DNA- and RBP translin, TSN (Wu et al. 1997), and that TSN is usually associated with a partner, translin-associated factor X (TSNAX/TRAX), which lacks nucleic acid-binding activity but possesses RNase activity (reviewed in Jaendling & McFarlane 2010). TSN has been implicated in several forms of post-transcriptional gene expression of many mRNAs in wild-type mouse testes. This is an important gap because the factors that regulate meiosis often differ in oocytes and spermatocytes (Cohen et al. 2006) and Sycp3 mRNA translation is also thought to be stimulated by DAZL (Brook et al. 2009).
spermatogenic cells: translational repression in round spermatids, export of mRNA to the cytoplasm in spermatocytes, and sharing of post-meiotic transcripts by transport of mRNA through cytoplasmic bridges. The functions of TSN, TSNAX, and the Y- and H-elements are confused by controversy and the absence of in vivo genetic evidence.

One controversy concerns the proportions of TSN homomer and TSN-TSNAX heteromer in spermatogenic cells. Finkenstadt et al. (2002) reported that virtually all TSN is associated with TSNAX in testis and cerebellum, while Cho et al. (2004) reported that the proportion of TSN homomer progressively increases fourfold to fivefold in pachytene spermatocytes, round spermatids, and elongated spermatids. This disparity is important because the affinity of the TSN homomer for transcript c is much greater and exhibits greater sequence specificity than those of the TSN-TSNAX heteromer (Chennathuzukuzhi et al. 2001, Lluis et al. 2010).

Another controversy concerns the sequence of bases to which TSN binds. Li & Baraban (2004) demonstrated that the recombinant and natural TSN-TSNAX heteromer in cerebellum extracts binds interspersed G residues instead of the specific bases of the Y- and H-elements. At present, none of the 68 Y- and H-elements in 42 mRNAs predicted computationally by similarity to the Prm2 Y- and H-elements are confirmed by RNA binding studies of natural and mutated elements (Han et al. 1995).

The critical question here is whether TSN represses mRNA translation in round spermatids. Evidence that partially purified TSN from testis extracts and recombinant TSN homomer repress translation in the RRL (Kwon & Hecht 1993, Yang et al. 2003) is subject to artifacts of heterologous systems and disagreement about the proportions of TSN homomer and TSN-TSNAX heteromer in vivo discussed above.

A gene trap mutation that abolishes the binding of TSN to mRNA has been created (Chennathuzukuzhi et al. 2003) and could be used to analyze the functions of TSN. However, the phenotype of the Tsn knockout is in disagreement with the expected consequences of depleting the putative repressor of many mRNAs in round spermatids. Premature translation of the Prm1 and Tnp2 mRNAs in transgenic mice produces deformed sperm and reduced male fertility (Lee et al. 1995, Tseden et al. 2007). The Tsn knockout might, therefore, be expected to have drastic effects on spermatid differentiation, because TSN is postulated to repress the translation of the Gapdh, Prm1, Prm2, Tnp1, Tnp2, and many other mRNAs. In contrast, the Tsn knockout reduces sperm number and litter size by approximately twofold, increases apoptosis in spermatocytes and results in the degradation of TSNAX, but does not increase the frequency of apoptotic round spermatids and abnormal sperm (Chennathuzukuzhi et al. 2003). The phenotype of the Celf3/Brunol1 knockout is very similar to that of the Tsn knockout, and it was interpreted as evidence that CELF3/BRUNOL1 ‘might be non-essential for spermatogenesis’ (Dev et al. 2007). The possibility that TSN does not repress translation in vivo cannot be dismissed because there are no reports of determinations whether the Tsn knockout results in a precocious activation of repressed mRNAs by the analysis of polysomal loading with sucrose gradients or protein expression in the testes of prepubertal mice.

Recent work has shown that the TSN-TSNAX heteromer, sometimes referred to as C3PO, functions in the RNA interference pathway by increasing the rate of degradation of the passenger RNA by the RNA interference silencing complex (Tian et al. 2011, Ye et al. 2011). This is probably a minor or redundant function because knockouts of the Tsn and/or Tsnax genes in fission yeast, flies, and mammals are viable and fertile (Jaendling & McFarlane 2010).

**Approaches for connecting elements and factors**

The challenges in connecting mRNA targets with regulatory factors originate in the number of different factors and the variety of parameters that influence the binding of RBPs to RNA. RBPs in eukaryotic cells are classified by RBDs, which exhibit different modes of interaction with RNA and functions (reviewed in Kishore et al. (2010), Morris et al. (2010) and Elliott & Ladomery (2011)). Table 1 lists a diverse assortment of RBPs that have been implicated in translational regulation in spermatogenic cells. These include RBDs that bind specific sequences of bases in a single-stranded RNA (ssRNA), DAZL and CPEB1, or a specific sequence of bases in all mRNAs, PABPC1. Other RBPs, such as TARBP2 and ADAR1, contain the dsRBD that binds the ribose sugars and phosphates of the backbones of double-stranded RNA segments instead of the bases. Some dsRBPs bind dsRNA segments in mRNAs nonspecifically, while others bind mRNA-specific configurations of multiple dsRNA segments. Argonaute proteins form complexes with miRNAs that are guided to specific mRNAs by complementary base-pairing (Bartel 2009, Fabian & Sonenberg 2012). Two factors in Table 1, EIF4G3 and PAIP2a, are not known to bind mRNA directly, but instead interact with initiation factors that bind mRNA. The versatility of interactions of RBPs with mRNA is documented in the general reviews cited above as well as specialized reviews of RBP families (Skabkin et al. 2006, Brook et al. 2009, Jankowsky 2011).

The binding of snRNAs and RBPs to mRNAs is also influenced by the accessibility to mRNAs (Kishore et al. 2010, Morris et al. 2010). Access is determined by mRNA secondary structure, RBPs, and snRNAs, which promote or block binding to specific elements, and RNP remodeling proteins such as Y-box proteins and DEAD-box RNA helicases, which dissociate and rearrange dsRNA segments and RBPs (Skabkin et al. 2006, Jankowsky 2011).
The specificity of interactions of RBPs with RNAs can also be defined by intracellular localization. Certain DEAD-box RNA helicases interact non-specifically with RNA in vitro, but remodel specific RNAs in vivo because the helicases function in a multiprotein complex containing hundreds of proteins and snRNAs that regulates specific classes of RNAs and RNP (Jankowsky 2011). This may be a property of the two DEAD-box helicases, DDX4/MVH and DDX25/GRTH, which are major constituents of the cytoplasmic germ-cell RNP granules known as the intermitochondrial cement in spermatocytes and the chromatoïd body in spermatids (Meikar et al. 2011).

The elements and mRNAs to which factors bind can be identified by biochemical studies of molecular interactions, computational searches for elements shared by co-regulated RNAs, or a combination of both, e.g. co-immunoprecipitation and microarray analysis or deep sequencing (reviewed in Kishore et al. (2010), Morris et al. (2010) and Ascano et al. (2012)). Each approach is distinguished by tradeoffs in simplicity, cost, and reliability in distinguishing bona fide targets and false positives, and every approach is capable of producing misleading information. Not surprisingly, the sequences of elements obtained using different techniques differ markedly (Table 2). Some of these differences arise because techniques that define elements precisely reveal that RBPs bind sequences that are often shorter, 3–6 nts, and more degenerate than were thought previously (Ray et al. 2009, Jenkins et al. 2011, Ascano et al. 2012).

The complexity of the parameters that determine the interactions of RBPs with RNA provides a compelling argument for identifying RBP–mRNA interactions in the natural environment of living cells. This has been achieved by variations of a procedure known as CLIP, based on the principle that u.v. irradiation at 254 nm cross-links bases and amino acids that are in direct contact in living cells and does not cross-link proteins (Ule et al. 2005). After cross-linking, mRNP fragments containing short segments of mRNA surrounding an element are purified with limited RNase digestion, immunoprecipitation to selectively concentrate the RNA–protein complexes, and SDS–PAGE to separate complexes from non-cross-linked RNAs and proteins.

A recent modification of CLIP, photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP), utilizes the incorporation of photoactivatable nucleosides and u.v. cross-linking at 365 nm in vivo (Hafner et al. 2010, Ascano et al. 2012). PAR ribonucleosides, such as 4-thiouridine and 6-thioguanosine, are incorporated into RNA by cells and cross-linked to protein at a wavelength, 365 nm, that minimizes damage to the RNA. Opinions differ whether cross-linking of PAR bases at 365 nm is 1.3- or >100-fold more efficient than that of natural bases at 254 nm (Kishore et al. 2011, Ascano et al. 2012). Importantly, cross-linking of amino acids to 4-thiouracil and 6-thioguanine in mRNA produces T to C and G to A transitions in cDNA produces respectively, providing criteria to identify the precise binding site and distinguish cross-linked copies from non-cross-linked contaminants. High-throughput sequencing of CLIP cDNAs, HITS-CLIP, has also been shown to detect single-nucleotide deletions at bases that are cross-linked to amino acids (Zhang & Darnell 2011). Thus, deep sequencing of mRNA fragments purified with PAR-CLIP and CLIP is capable of defining elements by the affinity of the RBP for all permutations of bases within and surrounding the element and the context of the element in relation to the local secondary structure and binding of other RBPs and snRNAs.

The invention of PAR-CLIP has stimulated transcriptome-wide analyses of RBP–mRNA interactions. PAR-CLIP and CLIP, in combination with oligo-dT purification of mRNPs and quantitative proteomics, identify ~800–860 proteins that bind polyadenylated mRNA in cultured mammalian cells (Baltz et al. 2012, Castello et al. 2012). Using deep sequencing, Baltz et al. (2012) also created a transcriptome-wide inventory of nucleotide sites that are bound by RBPs. PAR-CLIP identifies ~5000–7500 mRNAs bound by ELAVL1/HuR using 4-thiouridine, a judicious PAR base because this RBP binds U-rich elements (Kishore et al. 2011, Lebedeva et al. 2011, Mukherjee et al. 2011).

These studies have also shed light on the advantages and disadvantages of various methods for identifying RBP–mRNA interactions that are important in post-transcriptional regulation. The identification of RBP–RNA interactions is influenced by many technical factors including the quality of the antibody, the specificity and levels of the RNase used to purify short RNA–RBP complexes, and the depth of sequencing (Morris et al. 2010, Kishore et al. 2011, Lebedeva et al. 2011, Mukherjee et al. 2011).

It is clear that a single procedure does not identify all RBP–mRNA interactions. Irradiation at 254 nm preferentially cross-links pyrimidines with cysteine, lysine, phenylalanine, tyrosine, and tryptophan (Ule et al. 2005), and the potential for bias with PAR-CLIP is greater because cross-linking is restricted to elements containing the PAR base. Nevertheless, most RBPs are cross-linked to polyadenylated mRNA with both CLIP and PAR-CLIP, although 10–20% of the RBPs are identified with either procedure, but not both (Castello et al. 2012).

PAR-CLIP identifies about fourfold more ELAV1/Hur targets than does immunoprecipitation of native mRNPs without cross-linking, RIP-chip (Mukherjee et al. 2011). The greater number of mRNAs identified with PAR-CLIP was explained by the idea that PAR-CLIP detects both stable and transient mRNA–RBP interactions, while RIP-chip only detects stable interactions. This hypothesis is consistent with the observation that mRNA targets identified with RIP-chip exhibit slightly greater decreases...
in mRNA levels in the ELAV1/HUR siRNA knockdown than those identified with PAR-CLIP. Many targets identified with RIP-chip were not confirmed with PAR-CLIP (Mukherjee et al. 2011). The cause of this difference is unclear, but a potential explanation is that ELAV1/HUR binds some mRNAs indirectly through complexes with other RBPs. The identification of mRNA targets with CLIP, PAR-CLIP, or RIP-chip does not indicate that an interaction has major regulatory effects. Indeed, siRNA knockdown of ELAV1/HUR reveals that many putative targets undergo very small decreases in mRNA levels and that the most responsive mRNAs have multiple 3' UTR ELAV1/HUR-binding sites or binding sites in both introns and 3' UTRs (Lebedeva et al. 2011, Mukherjee et al. 2011).

The combination of CLIP, proteomics, and deep sequencing has the potential to enormously increase our understanding of the interactions of mRNAs and RBPs in spermatogenic cells. Fortunately, the costs of deep sequencing are decreasing rapidly, and PAR bases can probably be incorporated easily in seminiferous tubule culture (Klee 1993). Importantly, the ability of CLIP and PAR-CLIP to detect transient RNA–RBP interactions enables the identification of mRNA targets of RBPs that usually do not recognize specific sequences of bases in mRNA, e.g. argonautes, dsRBPs, and RNA helicases (Kishore et al. 2011, Zhang & Darnell 2011, Baltz et al. 2012, Castello et al. 2012). This will facilitate the identification of the direct targets of miRNAs and RBPs in spermatogenic cells that are presently unknown, such as ADAR, TARBP2, DDX4/MVH, and DDX25/GRT. Despite the successes of CLIP and PAR-CLIP, Castello et al. (2012) caution that individual RBP–mRNA interactions should be validated before inferring mechanisms and consequences of post-transcriptional regulation. In fact, CLIP was one of the several techniques that were used to identify the elements to which DAZL binds, and these elements were later validated by studies of mutations in tethering assays (Reynolds et al. 2005, 2007).

miRNAs and endo-siRNAs

The role of sncRNAs in spermatogenesis has been studied with knockouts of the Dicer1 gene encoding a Class III double-stranded RNA endonuclease that functions in the biosynthesis of two classes of sncRNAs, miRNAs and endogenous small interfering RNAs, endo-siRNAs (Song et al. 2011). Mature miRNAs and endo-siRNAs bind argonaute proteins and guide complexes to mRNA targets by base-pairing (Bartel 2009, Fabian & Sonenberg 2012). Endo-siRNAs appear to promote mRNA degradation and regulate transcription (Song et al. 2011), while miRNAs usually repress mRNA translation, decrease stability, or both (Bartel 2009, Fabian & Sonenberg 2012). A prolonged controversy concerning the mechanism by which miRNAs repress translation in somatic cells has been resolved recently with evidence that miRNAs function in a pathway that inhibits binding of the small ribosomal subunit to the 5' cap, usually followed by deadenylation (reviewed in Fabian & Sonenberg (2012) and Izaurralde (2012)).

The function of DICER1 has been studied using conditional CRE–Lox knockouts at specific stages of spermatogenesis, because the germ-line deletion of Dicer1 results in embryonic lethality (Bernstein et al. 2003). Dicer1 knockouts in pre-natal male germ cells and post-natal spermatogonia produce male infertility accompanied by abnormalities in round and elongated spermatids and severe decreases in sperm number (Korhonen et al. 2011, Romero et al. 2011). However, the pre-natal germ-cell Dicer1 knockout produces defects in early spermatocytes that are not observed with the spermatogonia knockout.

The phenotypes of male conditional Dicer1 knockouts will be particularly difficult to connect with specific mRNAs for reasons that have been cogently discussed by Bartel (2009). One reason is that DICER1 functions in the biosynthesis of hundreds of miRNAs and endo-siRNAs. The large number of sncRNAs that are depleted by Dicer1 knockouts greatly magnifies the problems in identifying direct targets: each miRNA interacts with tens to thousands of mRNAs, and some of these mRNAs presumably encode RBPs that regulate the expression of other genes at the transcriptional and post-transcriptional levels (Bartel 2009). Another reason is that knockouts of individual miRNAs often fine-tune the expression of hundreds of mRNAs, which cumulatively produce pronounced phenotypes (Bartel 2009). However, subtle changes in the expression of individual mRNAs are difficult to detect experimentally.

An important challenge of future work will be to determine whether individual miRNAs and endo-siRNAs have large or small effects in regulating the translation and stability of specific mRNAs in spermatocytes and spermatids.

Identification of cis-elements and factors that coordinate the translation of mRNAs

Deep sequencing identification of the cis-elements bound by RBPs will probably lead to insights into the factors that co-regulate mRNAs with similar patterns of developmental regulation (Kishore et al. 2010, Morris et al. 2010). However, current knowledge of the developmental changes in mRNA translation in meiotic and haploid spermatogenic cells is arguably too imprecise to identify co-regulated mRNAs. The intricate differentiation of spermatids involves chromatin remodeling and the ordered construction of the acrosome, outer dense fibers, fibrous sheath, mitochondrion sheath, and other subcellular structures unique to the spermatozoan.
Translational control in spermatogenic cells

(Russell et al. 1990, Meistrich et al. 2003). Clustering co-regulated mRNAs according to the stages of translational repression and activation and association with specific organelles is likely to be more informative than the usual Ingenuity-based classifications based on functions in somatic cells. The temporal translational activity of relatively few mRNAs has been precisely described and is often confused by disagreements in reports from different laboratories and multiple mRNA variants encoded by a single gene. For example, the Odf2 mRNA, which is studied as a typical mRNA that is transcribed in round spermatids and translated in elongated spermatids, is expressed as two variants with different 3′ UTRs, 103- and 1085-nt long. The effect of the alternative Odf2 3′ UTRs on post-transcriptional regulation is unknown.

Iguchi et al. (2006) used microarrays and sucrose gradients in a transcriptome-wide analysis of polysomal loading of prepubertal testes in which the most advanced cells were pachytene spermatocytes (17 days), round spermatids (22 days), and elongated spermatids (adult). Unfortunately, the inferences of developmental changes in polysomal loading are confused by high levels of polysomal mRNA at stages at which the corresponding protein cannot be detected and presentation of the findings in a form (absent or present) that does not indicate the proportions of mRNA in the free mRNP and polysomal fractions (discussed in Kleene et al. (2010)). Arava et al. (2003) provided a state-of-the-art prototype for this kind of study.

Iguchi et al. (2006) also proposed that the Pgk2 and Dbil5 mRNAs exemplify a novel mode of developmental translational control in which mRNAs are stored in free mRNPs in pachytene spermatocytes and translationally activated in round spermatids. Unfortunately, this idea is not supported by either their data or the literature. RT-PCR and northern blots concur that the Pgk2 mRNA is expressed from preleptotene spermatocytes to elongated spermatids (McCarrey et al. 1992, Yoshioka et al. 2007), while immunocytochemistry in four laboratories reveals that the PGK2 protein is first detected in step 9 (Vandeberg et al. 1981) or step 12 spermatids (Kramer 1981, Bluthmann et al. 1982, Danshina et al. 2010). Similarly, in situ hybridization and immunohistochemistry demonstrate that the Dbil5 mRNA and protein are respectively first detected in step 6–8 and step 16 spermatids (Kolmer et al. 1997, Pusch et al. 2000). The developmental expression of the Dbil5 and Pgk2 mRNAs and proteins is diagrammed in Fig. 1. The microarray and RT-qPCR data of Iguchi et al. (2006) also show that the polysomal loading of the Pgk2 and Dbil5 mRNAs increases sharply between round spermatids and elongated spermatids. Thus, all of these findings concur with the generally accepted view that the developmental regulation of Pgk2 and Dbil5 mRNA translation in spermatids circumvents the transcriptional inactivity imposed by chromatin remodeling in elongating and elongated spermatids.

I know of no mRNAs that have been demonstrated by in situ hybridization, immunohistochemistry, and sucrose gradients to undergo significant increases in translational activity in round spermatids. Further work will be required to determine whether RBPs that have been implicated in the development of round spermatids are required to activate the translation of individual mRNAs at specific stages or to sustain a constant rate of mRNA translation (Paronetto et al. 2009, VanGompel & Xu 2010).

The task of deducing the functions of factors and elements identified in high-throughput studies of mRNA translation would be facilitated by a carefully annotated compilation of the developmental expression of mRNA variants and proteins in spermatogenic cells. At present, this list does not exist and would require a painstaking analysis of the literature to assess the quality of the evidence and resolve disagreements between reports.

Validating cis-elements with studies of mutations in vivo

Studies of gene expression in spermatogenic cells at every level, from transcription to translation and stability, have been thwarted by the absence of a suitable technique for validating cis-elements with mutations. In vivo genetic analyses of mutations in elements have been limited to transgenic mice, which are considered too costly and risky by many. However, it is debatable whether the inconvenience of transgenic mice is worse than the question of whether information based on studies of mutations in heterologous systems might be invalid. Such studies would also benefit from quantification to determine whether a specific element has a major or minor effect on translational activity (Kleene et al. 2010). This impasse can potentially be solved by combining several of the strategies described below.

The risk of squandering resources on mutations in cis-elements that produce no phenotypes can be reduced by using comparative genomics to identify conserved sequences or RNA binding assays to identify elements. Both approaches have been used successfully by molecular biologists, with the caveat that elements often have negligible regulatory effects. Predicting binding sites of RBPs in spermatogenic cells that are likely to produce strong phenotypes will become easier as information becomes available from transcriptome-wide studies in somatic and spermatogenic cells. The studies cited above indicate that the critical parameters vary among RBPs and elements and include binding affinity, number of binding sites, binding sites in both introns and 3′ UTRs, positions within 3′ UTRs and 5′ UTRs, and distance from the binding sites of miRNAs and other RBPs.

Comparative genomics is appealing because it is fast and cheap. The public databases contain more-or-less
modified spermatogonial stem cells, are able to fertilize eggs through copulation or microinjection of sperm and produce normal offspring, the cultured stem cells appear to be free of epigenetic and genetic mutations (Kanatsu-Shinohara et al. 2005). Creating transgenic and knockout mice by transfection of spermatogonial stem cells and repopulation of the testes of sterile recipients is technically easier and more efficient than by conventional approaches using pronuclear injection and blastocyst chimeras. Thus, the use of spermatogonial stem cells has been advanced as a cost-effective method for creating knockout and transgenic mice, and the approach may become widely available in core facilities. Again there is a downside; the repopulation of seminiferous tubules is a prolonged process requiring 2–3 months for the injected stem cells to differentiate into spermatozoa. It is unclear whether this disadvantage offsets the advantages in decreased screening and time to generate transgenic lines.

The analysis of mutations in transgenic mice can also be expedited by studying the GFP reporter in seminiferous tubule squashes in transgenic founders. Each founder can be analyzed in several hours, eliminating delays and costs in establishing lines, and the levels of GFP fluorescence can be quantified with programs such as ImageJ (Ventelä et al. 2000a, Bagarova et al. 2010). One testis from an adult male can be analyzed biochemically and the other with seminiferous tubule squashes. Phase-contrast images of tubule squashes display the morphological details that distinguish every stage of meiotic and haploid spermatogenic cells more clearly than do photographs of paraffin sections (Kotaja et al. 2004). A reporter in which GFP is fused to acrosin localizes GFP in the acrosome, the critical organelle in stage round spermatids, and may prevent the release of soluble GFP from lysed cells (Ventelä et al. 2000b).

Importantly, the stage at which proteins encoded by the Smcp–Gfp and Prm1–hGH transgenes are first detected is independent of variations in mRNA levels originating in differences in the number and position of transgene insertions (Braun et al. 1989, Fajardo et al. 1997, Giorgini et al. 2001, Bagarova et al. 2010).

A study analyzing approximately ten mutations by whatever means would initiate a new era in studies of gene regulation in spermatogenic cells.

Conclusions and important questions to be answered

Table 1 lists factors that have been implicated in translational regulation in spermatogenic cells. The knockouts of most of these factors produce male infertility or subfertility and abnormal sperm, demonstrating the importance of post-transcriptional regulation in normal development of spermatozoa. However, the knockouts of the Tsn, Cell3/Bruno11, and Ybx3 genes do not produce abnormal sperm and therefore appear to have minor functions. Unfortunately, none of
the cis-elements mentioned in this review has been demonstrated to interact with a specific factor and regulate translation in its natural position in spermatogenic cells. The failure to connect factors and target mRNAs with translational control limits our understanding of the mechanisms, which in turn hinders formulating hypotheses that are likely to lead to important insights into the mechanisms by which abnormal post-transcriptional regulation decreases male fertility.

The combination of CLIP, deep sequencing, and transgenic mice provides approaches to rigorously identify the direct targets of specific RBPs and miRNAs and to determine their precise function in the regulation of specific mRNAs. These techniques should make it possible to identify individual miRNAs, endo-siRNAs, and mRNA targets that are major regulators of spermatogenesis.

Mutations in regulatory elements in transgenic mice also provide a direct approach to resolving controversies concerning the factors and elements that regulate the translation of the Prm2, Sycp3, and Hspa2 mRNAs. The translation of the Prm2 mRNA has been proposed to be repressed by the Y- and H-elements, a 3' UTR sequence similar to the Prm1 TCE and two coding region elements that bind the atypically long 26-nt mRNA 469 (Kwon & Hecht 1993, Zhong et al. 2001, Dai et al. 2011), while the Sycp3 and Ddx4/Mvh mRNAs have been proposed to be activated by the CPE- and DAZL-binding sites (Tay & Richter 2001, Brook et al. 2009, see above). If the activation of Hspa2 mRNA translation by EIF4G3 and ELAVL1/HuR can be connected with specific elements, it will be possible to determine which pathway is more important for the exit from meiosis (Sun et al. 2010, Chi et al. 2011). It should also be possible to identify the unknown factor that binds the Prm1 TCE with CLIP and immunoprecipitation with the appropriate antibody, regardless of whether the factor is an scRNA or an RBP. Bridging these gaps will increase our understanding of the mechanisms that regulate three important transitions in spermatogenesis: the initiation of meiosis, the completion of meiosis and the differentiation of late spermatids.

Declaration of interest

The author declares that there is no conflict of interest that could be perceived as prejudicing the impartiality of this review.

Funding

This research did not receive any specific grant from any funding agency in the public, commercial or not-for-profit sector.

Acknowledgements

I am grateful to the many scientists whose careful work provided insights into the complicated and fascinating domains of eukaryotic post-transcriptional gene regulation and spermatogenesis.

www.reproduction-online.org

References


Brook M, Smith JW & Gray NK 2009 The DAZL and PABP families: RNA-binding proteins with interrelated roles in translational control in oocytes. Reproduction 137 595–617. (doi:10.1530/REP-08-0524)

Calvo SE, Pagliarini DJ & Mootha VK 2009 Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. PNAS 106 7507–7512. (doi:10.1073/pnas.0810916106)


Chennathuzukuzhi VM, Kurihara Y, Bray JD & Hecht NB 2001 Trx (translin-associated factor X), a primarily cytoplasmic protein, inhibits the binding of TB-RBP (translin) to RNA. Journal of Biological Chemistry 276 13256–13263. (doi:10.1074/jbc.M009707200)


Cho YS, Chennathuzukuzhi VM, Handel MA, Eppig J & Hecht NB 2004 The relative levels of translasin-associated factor X (TRAX) and testis brain RNA-binding protein determine their nucleocytoplasmic distribution in male germ cells. Journal of Biological Chemistry 279 31514–31523. (doi:10.1074/jbc.M401422200)


Downloaded from Bioscientifica.com at 12/18/2018 01:05:59AM
via free access

Giorgini F, Davies HG & Braun RE

Finkenstadt PM, Jeon M & Baraban JM

Fabian MR & Sonenberg N


Iguchi N, Tobias JW & Hecht NB 2006 Expression profiling reveals meiotic male germ cell mRNAs that are translationally up- and down-regulated. PNAS 103 7712–7717. (doi:10.1073/pnas.0510999103)


Kirschner WA & Tres LL 1975 Structural and transcriptional features of the mouse spermatid genome. Journal of Cell Biology 65 258–270. (doi:10.1083/jcb.65.2.258)


Kwon YK & Hecht NB 1993 Binding of a phosphoprotein to the 3'-terminal stem-loop-binding protein enhances translation through a U-rich motif in its 3' UTR. Genes and Development 23 2742–2752. (doi:10.1101/gad.1820709)


Reynolds N, Collier B, Bingham V, Gray NK & Cooke HJ 2007 Translation of the synaptonemal complex component Scy73 is enhanced in vivo by the germ cell specific regulator Dazl. RNA 13 974–981. (doi:10.1261/ma.465502)
Shyu AB, Greenberg ME & Belasco JG 1989 The c-fos transcript is targeted for rapid decay by two distinct mRNA degradation pathways. Genes and Development 3 60–72. (doi:10.1101/gad.3.1.60)

VanDergoemp M & Xu EY 2011 The roles of the DAZ family in spermatogenesis: more than just translation? Spermatogenesis 1 36–46. (doi:10.4161/spmg.1.1.14659)


Wu XQ, Gu W, Meng X & Hecht NB 1997 The RNA-binding protein, TB-RBP, is the mouse homologue of translin, a recombination protein associated with chromosomal translocations. PNAS 94 5640–5645. (doi:10.1073/pnas.94.11.5640)


