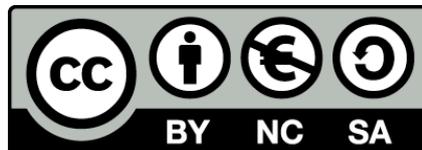




UNIVERSITAT<sub>DE</sub>  
BARCELONA

**Combinatorial perspective  
on the gene expression circuits established  
by the CPEB-family of RNA binding proteins**

Berta Duran Arqué



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement- NoComercial – CompartirIgual 4.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento - NoComercial – CompartirIgual 4.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution-NonCommercial-ShareAlike 4.0. Spain License.**

**Combinatorial perspective on  
the gene expression circuits  
established by the CPEB-family  
of RNA binding proteins**

**Berta Duran Arqué**





UNIVERSITAT DE  
BARCELONA

INSTITUTE  
FOR RESEARCH  
IN BIOMEDICINE



UNIVERSITAT DE BARCELONA

FACULTAT DE FARMÀCIA

PROGRAMA DE DOCTORAT EN BIOMEDICINA

# Combinatorial perspective on the gene expression circuits established by the CPEB-family of proteins

Memòria presentada per Berta Duran Arqué per optar al títol de doctor per la  
Universitat de Barcelona.

Aquesta tesi ha estat realitzada sota la direcció del Dr. Raúl Méndez de la Iglesia a  
l'Institut de Recerca Biomèdica (IRB).

Dr. Raúl Méndez  
de la Iglesia, director

Dr. Antonio Zorzano  
Olarte, tutor

Berta Duran Arqué,  
doctoranda

Barcelona, 2020



*“Everybody’s looking for their tiny piece of meaning.  
Some fleeting, perfect thing that might make them more alive.”*

Kate Tempest in *The Bricks That Built the Houses*



# Acknowledgements

Five years ago, I moved to Barcelona to start this doctoral thesis. Since, life and thesis have been co-existing in a fragile equilibrium, occasionally merging into one, only to split again and go back to their tense co-existence. Thesis and life have nurtured and shaped one another. So, if you ever stepped through my life, bits of you may now be scattered through this thesis. For that, I wish to thank you.

First and foremost, I would like to thank Raúl. Raúl, it has been a pleasure to join the team. Thank you for betting on me and for giving me space to wander, discover and express myself. I would also like to thank Fátima Gebauer, Jens Lüders and Antonio Zorzano for the advice, ideas and direction provided in the advisory committees. I am also thankful to the IRB community and IRB staff, specially to Gian and Marina from the Proteomics facility, Óscar, Adrià and Camille from the Bioinformatics and Biostatistics facility, Nacho and David from the Genomics facility as well as Anna, Lidia and Nikos from the Microscopy facility. Thank you for your support and sincere involvement.

Undoubtedly, most thank-yous and without-you-it-would-not-have-been-possibles go to the Rmendez lab members, both past and present. Because, truly, I cannot imagine spending 5 years in better company. Thank you all for the support, laughs, criticism and cakes, pies, muffins, cookies and palmeras de Morata.

Specifically, thank you, Carlos and Jordina, you set a really high standard for the rest of Rmendez PhD students. Thank you, Vitto, for being so critical of yours and others' work. Thank you, Gonzalo, for being an example of positivity and motivation. Thank you, Eulàlia, for welcoming us to the *Xenopus* crew and guiding us when we were PhD toddlers. Thank you, Chiara, for always being willing to lend a hand and for being an example of hard work. Thank you, Vero, for being such an all-rounder and for your fresh pragmatic take on life. Thank you, Judit, for being straightforward, supportive and great company. Thank you, Alba, for being a reference of good organization and good practices. Thank you, Annarita, for your generosity and kindness.

And, of course, thank you, Pinchis, it has been wild to share struggles, buffers, music and learn plenty of new words with you. Thank you, Anna, I admire your calm positive attitude. Thank you, Anna F., I am impressed by your capacity to stay concentrated and calm in the face a big project. Thank you, Marcos, for being someone one can talk to about anything. To the senior-pinchis, Irene, Rosa and Clara, thank you three for

showing how one can rock a PhD and everything else in life, I am still in awe. And last, thank you, Manu, I could not have envisioned sharing the PhD with someone so funny, creative and open.

To my friends, old and new, who have shared this past five years in Barcelona with me, thank you all. Laia, no saps com m'alegro d'haver-te conegut. Gràcies per ser companya d'entrenaments i de pis, per acollir-me sense fer preguntes. Thank you, Nina, I am glad the rollercoaster ride is over although it was amazing to be by your side. Albert, Edu, Edu, Edu, Gerard, Saúl, Matteo, Rebecca, Sara, Suguru... it's been a blast! Thank you for trusting me and inspiring me to grow, I have learned so much alongside you all. I a la Troja plus, gràcies per ser-hi i per totes les aventures.

Another big thank you goes to my family, who might still not know what this thesis is about, whether I have a job or not and why I have put myself through this, but who are unconditionally by my side. Gràcies, mare i pare, entre moltes altres, per l'ajuda en les múltiples mudances. Gràcies aviets, a tots quatre, de cadascú de vosaltres he après i segueixo aprenent què és i com prendre'm això de la vida. Gràcies Marta i Josep, tietes, tiets, cosines i cosins. Also, to those who once were family, I am immensely grateful to have been accompanied by you, Lukas and Cova and, from afar, Anja and all the Fertigs.

Por último, Carlos, gracias por co-fundar este equipo. Gracias por ayudarme a relativizar los problemas y reírte de mi conmigo. Gracias por las aventuras, los "chistes" y por la banda sonora.

A tots, gràcies per tant.

# Contents

## **Acknowledgements**

## **Abbreviations** **5**

## **Abstract** **11**

## **Introduction** **13**

The flows of genetic information . . . . . 15

The life cycle of a canonical mRNA . . . . . 16

Oogenesis and anatomy of the growing immature oocyte . . . . . 21

Meiotic maturation and fertilization . . . . . 24

Spatiotemporal control of gene expression in meiotic maturation. . . . . 28

The CPEB protein family across tissues and organisms. . . . . 32

Higher-order mRNP organization . . . . . 36

## **Aims** **39**

## **Results** **43**

The CPEBs are co-expressed in *X. laevis* oocytes. . . . . 45

CPEB2 and CPEB3 are phosphorylated in meiotic maturation. . . . . 47

CPEB3 is phosphorylated at proline-directed sites clustered on the N-terminal domain. . . . . 49

Bulk CPEB2 phosphorylation also occurs at proline-directed sites. . . . . 53

ERK2 and Cdk1:CycB phosphorylate CPEB3. . . . . 56

CPEB2 and CPEB3 multiple phosphorylations impact cellular distribution . . . . . 58

CPEB2 and CPEB3 are more dynamic than CPEB1 . . . . . 61

The CPEB1-3 co-localize in the same compartments . . . . . 65

The CPEBs share a core proximome of mRNA processing, storage and repression proteins . . . . .	66
The CPEBs regulate an overlapping set of mRNAs . . . . .	75
Comparative analysis of CPEB1-4 targetomes. . . . .	83
<b>Discussion</b>	<b>87</b>
Functions of the CPEB-family in meiotic maturation . . . . .	90
A matter of protein concentration? . . . . .	91
LLPS in meiotic maturation. . . . .	93
<b>Conclusions</b>	<b>95</b>
<b>Materials and methods</b>	<b>99</b>
Buffers . . . . .	101
<i>X. laevis</i> oocytes preparation . . . . .	102
Oocyte lysis . . . . .	103
Subcloning and mutagenesis. . . . .	103
<i>In vitro</i> transcription (ivt) and polyadenylation . . . . .	104
SDS-PAGE and WB. . . . .	104
B-isox precipitation. . . . .	105
Protein IP. . . . .	105
Lambda Protein Phosphatase ( $\lambda$ -PPase) assay . . . . .	107
CPEB3 p-site mapping by MS. . . . .	107
SVI <i>X. laevis</i> BioID. . . . .	109
RIP-Seq . . . . .	112
Protein overexpression and purification from <i>E. coli</i> . . . . .	115
<i>In vitro</i> phosphorylation assays with oocyte lysates . . . . .	116

<i>In vitro</i> phosphorylation assays with oocyte lysates and kinase inhibitors. . . . .	116
<i>In vitro</i> phosphorylation assays with recombinant ERK2 and Cdk1:CycB. . . . .	117
Cell culture and transfection. . . . .	117
CPEB2 and CPEB3 distribution in transfected U-2 OS cells . . . . .	118
CPEB2 and CPEB3 co-localization with CPEB1. . . . .	119
Live cell imaging and FRAP. . . . .	120
<b>Appendices</b>	<b>121</b>
<b>References</b>	<b>139</b>



# Abbreviations



<b>l-PPase</b>	Lambda Protein Phosphatase
<b>A, Ala</b>	Alanine
<b>APA</b>	Alternative polyadenylation
<b>ARE</b>	AU-rich element
<b>B-isox</b>	Biotinylated-isoxazole
<b>C-ter</b>	C-terminal
<b>CBB</b>	Coomassie Brilliant Blue
<b>CDS</b>	Coding sequence
<b>CPE</b>	Cytoplasmic polyadenylation element
<b>D, Asp</b>	Aspartic acid
<b>DMEM</b>	Dulbecco Modified Eagle Medium
<b>DTT</b>	Dithiothreitol
<b>E, Glu</b>	Glutamic acid
<b>e.g.</b>	<i>Exempli gratia</i>
<b>eIF</b>	Eukaryotic initiation factor
<b>Elu</b>	Elate or elution
<b>et al.</b>	<i>et alter</i>
<b>FBS</b>	Fetal Bovine Serum
<b>FC</b>	Fold change
<b>FRAP</b>	Fluorescence recovery after photobleaching
<b>GFP</b>	Green Fluorescent Protein
<b>GSEA</b>	Gene set enrichment analysis
<b>GO</b>	Gene Ontology
<b>GV</b>	Germinal vesicle
<b>GVBD</b>	GV breakdown
<b>H, His</b>	Histidine
<b>H1</b>	Histone H1
<b>H1K</b>	Histone H1 kinase
<b>HA</b>	Hemagglutinine epitope
<b>Hex</b>	Hexanucleotide

<b>i.e.</b>	<i>ie est</i>
<b>iBAQ</b>	Intensity-based absolute quantification
<b>IDR/IDP</b>	Intrinsically disordered region/protein
<b>In</b>	Input
<b>IP</b>	Immunoprecipitation
<b>IVT</b>	In vitro transcription/transcribed
<b>K, Lys</b>	Lysine
<b>KD</b>	Knockdown
<b>KO</b>	Knockout
<b>L</b>	Ladder
<b>LB</b>	Luria Broth
<b>LC-MS/MS</b>	Liquid chromatography followed by tandem mass spectrometry
<b>LCR</b>	Low complexity region
<b>LLPS</b>	Liquid-liquid phase separation
<b>LTM</b>	Long-term memory
<b>MBE</b>	Musashi-binding element
<b>MBS</b>	Modified Barth's Solution
<b>MBT</b>	Mid-blastula transition
<b>Met-I/II</b>	Metaphase I or II
<b>MI</b>	First meiotic division
<b>MII</b>	Second meiotic division
<b>MIP</b>	Max intensity projection
<b>MPF</b>	M-phase promoting factor
<b>mRNA</b>	Messenger RNA
<b>MS</b>	Mass Spectrometry
<b>MW</b>	Molecular weight
<b>N-ter</b>	N-terminal
<b>NEM</b>	N-Ethylmaleimide
<b>NES</b>	Normalized enrichment score
<b>NI</b>	Not-injected

<b>p-adj</b>	Adjusted p-value
<b>P-body</b>	Processing body
<b>p-mim</b>	Phosphorylation mimetic
<b>p-null</b>	Phosphorylation null
<b>p-site</b>	Phosphorylation site
<b>P, Pro</b>	Proline
<b>P</b>	Progesterone
<b>PAGE</b>	Polyacrylamide gel electrophoresis
<b>PAS</b>	Polyadenylation signal
<b>PBD</b>	Proximity-dependent biotinylation
<b>PBE</b>	Pumilio binding element
<b>PBS</b>	Phosphate-buffered saline
<b>PC</b>	Principal component
<b>PCA</b>	PC analysis
<b>PGC</b>	Primordial germ cell
<b>PIC</b>	Pre-initiation complex
<b>Pol</b>	Polymerase
<b>Pro-I</b>	Prophase of the first meiotic division
<b>PSM</b>	Peptide spectrum matches
<b>PTM</b>	Post-translational modification
<b>R, Arg</b>	Arginine
<b>RBP</b>	RNA-binding protein
<b>Rep</b>	Replicate
<b>RIP-Seq</b>	RNA-immunoprecipitation Sequencing
<b>RNP</b>	Ribonucleoprotein particle
<b>ROI</b>	Region of interest
<b>RPKM</b>	Reads per kilobase per million mapped reads
<b>rpm</b>	Revolutions per minute
<b>RRM</b>	RNA recognition motif
<b>RT</b>	Room temperature

<b>S, Ser</b>	Serine
<b>SDS</b>	Sodium dodecyl sulfate
<b>SG</b>	Stress granule
<b>SII</b>	Stage II
<b>SVI</b>	Stage VI
<b>t-half</b>	Half time of recovery
<b>T, Thr</b>	Threonine
<b>TBS</b>	Tris-buffered saline
<b>UTR</b>	Untranslated region
<b>WB</b>	Western blot
<b>wms</b>	White maturation spot
<b>wt</b>	Wild type
<b>ZGA</b>	Zygotic genome activation

# Abstract

The complex changes that take place in the mature *Xenopus* oocyte and early embryo are orchestrated in the absence of transcription. Until zygotic transcription starts, after the mid-blastula transition, cells rely on tight spatiotemporal translational regulation. Some maternal mRNAs accumulate during oocyte growth and are stored, translationally silent. Upon stimuli, stored, silenced mRNAs become cytoplasmically polyadenylated and, subsequently, engage in translation. The timing and extent of translational activation are dictated by a complex code of 3'UTR motifs recognized by RNA-binding proteins.

In meiotic maturation, at least three sequential polyadenylation waves occur. First, in response to progesterone, a single Aurora kinase A phosphorylation triggers CPEB1-directed cytoplasmic polyadenylation of mRNAs that are required for Cdk1 activation and meiotic progression. Second, activated Cdk1 targets CPEB1 for degradation, triggering a second polyadenylation surge that is necessary for the MI-MII transition. Last, CPEB4, synthesized from the first wave and activated by Cdk1 and ERK2 upon meiotic progression, drives a third wave during the second meiotic division that is required for the metaphase-II arrest.

Unlike the well-studied roles of CPEB1 and CPEB4, the roles of the remaining family members, CPEB2 and CPEB3, remain uncharacterized. In this thesis we have performed a systematic investigation of the CPEB-family of RBPs in meiotic maturation in order to elucidate their combinatorial contribution to gene expression regulation. We have determined that CPEB1 and the CPEB2-4 subfamily differ in their expression dynamics, concentration and regulation. Like CPEB4, CPEB2 and CPEB3 are regulated by N-terminal hyperphosphorylation that causes dissolution of the CPEB-condensates. Furthermore, we have found that all CPEBs co-localize and are proximal to mRNA repression and storage proteins, probably reflecting their inclusion within large repressive mRNPs in the oocyte. We have also found that all CPEBs bind a highly overlapping subset of mRNAs, although CPEB1 and CPEB2-4 could differentially regulate a small subset of targets.

All in all, we have contributed to understanding how the multiple CPEBs co-exist and how their activities are coordinated in the cell to dictate complex expression patterns.



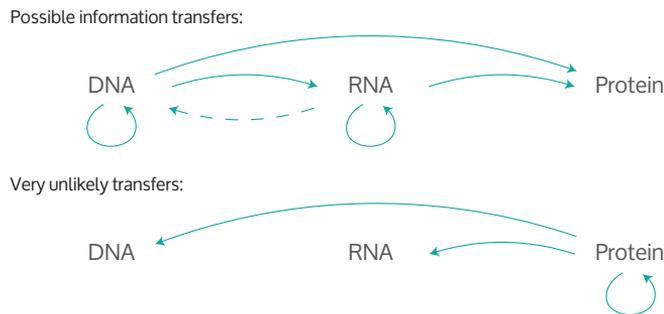
# Introduction



## The flows of genetic information

With a few exceptions, all cells that constitute a multicellular organism host the same DNA material. However, the output of this material differs, giving rise to cells as different as a neuron, a cardiomyocyte, a leukocyte or a fibroblast. How the genetic output is controlled is a fundamental biological question to which we keep adding layers of complexity.

In a seminal lecture back in 1957, with a lot of the pieces of the puzzle missing, Francis Crick proposed the Central Dogma - or, arguably, basic assumption (Cobb 2017) - of Molecular Biology and the Sequence Hypothesis. Crick addressed the general rules of sequence information transfer between three polymer molecules: DNA, RNA and protein. Considering all possible information transfers, the sequence hypothesis stated that overall transfer from nucleic acid to protein did exist, arguing that “the main function of the genetic material is to control (not necessarily directly) the synthesis of proteins.” cit. (Crick 1970). The Central Dogma was a negative statement saying that transfers from protein did not exist: “once information has passed into protein it *cannot get out again* [...]” cit. (Crick 1958) (**Figure 1**).



**Figure 1. The flow of genetic information.** The Sequence Hypothesis and Central Dogma proposed by Crick, from an unpublished note dating to 1956.

In 1961, two publications appeared back-to-back in *Nature*, demonstrating the existence of the short-lived RNA species that serves as an intermediate between genes and proteins: messenger RNA (mRNA) (Brenner et al. 1961, Gros et al. 1961). Soon after, still in 1961, breakthrough work from Nirenberg and Matthaei cracked the “genetic code”, deciphering its first “word”: a poly-uracyl homopolymer codes for a chain of phenylalanines (Nirenberg & Matthaei 1961). Subsequent work in Nirenberg’s lab exploited the cell-free system they had set up - where they added synthetic RNAs and determined

which aminoacid chain was being synthesized - to complete the correspondence between nucleotide triplets and aminoacids.

In parallel, Jacob and Monod were studying the synthesis of enzymes in bacteria in response to lactose addition to the culture media. They proposed the operon model, whereby the product of a regulator gene would directly regulate synthesis of a group of functionally-related structural genes, revealing a “co-ordinated program of protein synthesis and the means of controlling its execution” (Jacob & Monod 1961). The *lac* operon system is widely recognized as the first description of transcriptional gene regulation.

In the nearly sixty years since these pioneering findings and hypotheses, our understanding of gene expression has continued to expand. Crick’s hypotheses are still cited because the paradigm is that proteins constitute the functional end-product of gene expression, and these end-products are produced from DNA through an intermediate RNA message. However, we also know, that in order for a protein to be present in a given place at a given time, DNA elements that do not encode for RNA are needed and RNA elements that do not encode for protein are also needed. That is, DNA, RNA and proteins contain more information than what was deciphered in those early days, more than the aminoacid sequence. Many of the ways information is contained in DNA and RNA are not fully understood. Thus, additional codes remain to be cracked.

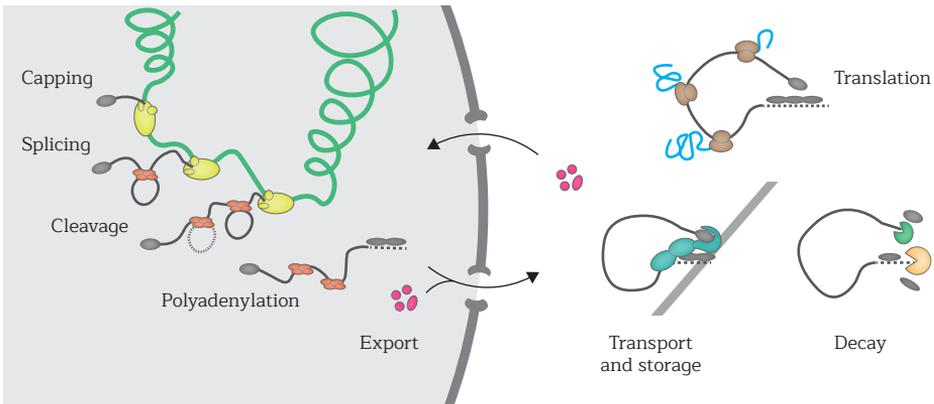
In the lab, we are interested in the CPEB family of RNA-binding proteins (RBPs) and two information codes: the messenger ribonucleoprotein particle (mRNP) code and the 3’ untranslated region (3’ UTR) code. Because of our focus in post-transcriptional gene expression and regulation, first, the life cycle of an RNA polymerase (pol) II transcript will be described. Next, because the experimental model used in this thesis are *Xenopus laevis* oocytes, their biology and the events that unfold in meiotic maturation will be laid out. Then, the specific mechanisms of gene expression regulation that operate in *X. laevis* meiotic maturation will be discussed. For completeness, the roles of the CPEB family of proteins beyond meiotic maturation will be introduced. Finally, I will briefly review the state-of-the-art concerning phase transitions of RBPs and mRNP granules.

## The life cycle of a canonical mRNA

From synthesis until degradation, mRNAs go through their life cycle accompanied by proteins, non-coding RNAs, small metabolites and ions in the form of dynamic mRNPs. The mRNP identity at any given time determines its properties, fate and function. This

has been termed the mRNP code. The mRNP is born with the mRNA and is constantly and coordinately remodeled through its life (for reviews see Gehring et al. 2017, Singh et al 2015).

The first step in the mRNA - and mRNP - life cycle is co-transcriptional nuclear processing, which in turn consists of three sequential events coordinated with RNA pol II (Herzel et al. 2017, Martinez-Rucobo et al. 2015): 5' end capping, splicing and 3' end cleavage and polyadenylation. Thereafter, mRNPs are exported to the cytoplasm where they can be transported to specific subcellular locations, stored, translated and eventually degraded (**Figure 2**).



**Figure 2. The mRNA life cycle.** Stages, as described in the text, of the life cycle of a capped and polyadenylated RNA polymerase II transcript. Redrawn with inspiration from Silver (Hieronymus & Silver 2004).

### Nuclear pre-mRNA processing and export

The first processing step - 5' end capping - consists on the addition of an N7-methyl-guanosine ( $m^7G$ ) moiety linked by a 5'-5' triphosphate chain to the first transcribed nucleotide. It is a three-step reaction, initiated as soon as the nascent transcript emerges from the RNA polymerase exit tunnel - at 20-40 nucleotides long (Saldit-Georgieff et al. 1980). Capping is important to ensure RNA protection from 5'-exonucleases and is required for the subsequent processing steps, as well as export and translation. Throughout the mRNA lifetime the cap is bound by different cap-binding complexes that constitute an important protein interaction hub.

Splicing is performed by the spliceosome, a protein-directed metalloribozyme that assembles on specific mRNA sequences and catalyzes two trans-esterification reactions

that lead to intron excision. The spliceosome is a large dynamic molecular machine - consisting of 90 core proteins in yeast and 175 in humans (Fabrizio et al. 2009) - and its assembly and disassembly occur as many times as there are introns present in the pre-mRNA. The developmentally and cell-type-regulated usage of splice sites that enables various splicing outcomes is termed alternative splicing. It has been estimated that over 90% of multi-exon human genes undergo alternative splicing (Pan et al. 2008, Wang et al. 2008).

The last step in pre-mRNA processing is 3' end cleavage and polyadenylation. Cleavage defines the 3' end of the transcript and it is directed by a multipartite motif in the pre-mRNA, consisting of: a polyadenylation signal (PAS) - degenerate A-rich hexanucleotide (hex), AAUAAA and AUUAAA being the most common variants - within 40 nucleotides upstream the cleavage site, U/GU-rich sequences immediately downstream the cleavage site plus additional distal elements (reviewed in Tian & Graber 2012). Cleavage is succeeded by polyadenylation of the exposed 3' end. In higher eukaryotes, a tail of circa 250 adenines is added that, similarly to the m<sup>7</sup>G cap at the other end, serves as an interaction platform and confers protection from 3'-exonucleases. Cleavage and polyadenylation are executed by the cleavage and polyadenylation specificity factor complex (CPSF) and cleavage stimulating factor complex (CSTF). These complexes bind to the PAS and downstream sites, respectively, and recruit other factors - such as the scaffold Symplekin (SYMPK) and the poly(A) polymerase - that altogether constitute the active mRNA 3' processing complex (Shi & Manley 2015). Like alternative splicing, alternative selection of cleavage and polyadenylation sites (APA) is a widespread regulated process that contributes to transcript diversity (Flavell et al. 2008, Sandberg et al. 2008).

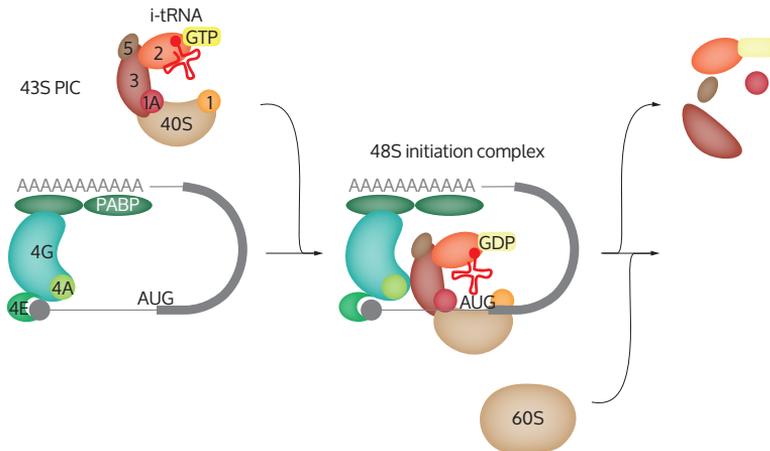
Pre-mRNA processing is coupled to transcription termination and cytoplasmic export of the mature mRNP (Perales & Bentley 2009, Xie & Ren 2019). Cytoplasmic export through the nuclear pores is unidirectional and requires participation of many export factors.

### **Life in the cytoplasm**

Once in the cytoplasm, the fate of mRNPs loses the linearity it had until this point. Upon export, mRNPs may immediately enter the translationally active pool or may be held quiescent until some signal is sensed. This regulated onset of translation is often coupled to transport to specific subcellular regions. Because spatiotemporal regulation of translation has been most studied in highly polarized cells - such as oocytes or neurons - it will be discussed in the context of *X. laevis* oocytes in the section 'Spatiotempo-

ral control of gene expression in meiotic maturation’.

Translation is another process executed by a large molecular machine, the ribosome.

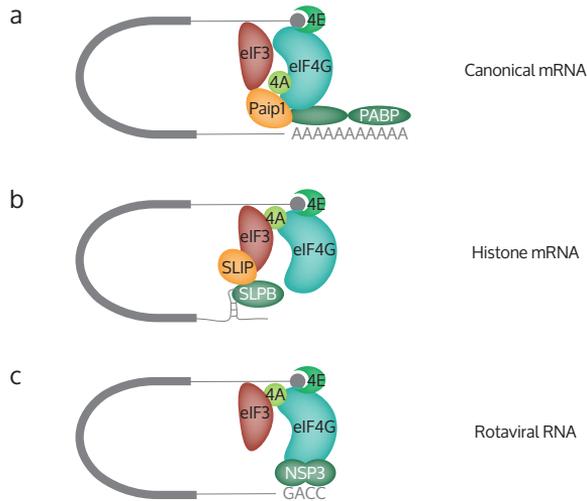


**Figure 3. Cap-dependent translation initiation.** The 43S PIC (comprising a 40S subunit, eIF2-GTP, initiator tRNA, eIF3, eIF1, eIF1A and eIF5, drawn in brown shades) is recruited to the mRNA by the eIF4F cap-binding complex (comprising eIF4E, eIF4G and eIF4A drawn in green shades). The 43S PIC scans the 5'UTR until it recognizes an initiation codon and commits to it, forming the 48S initiation complex. Further remodeling upon base-pairing with the start codon leads to release of eIFs and joining of the 60S ribosomal subunit.

The first step, initiation, requires participation of a myriad of factors and is viewed as the most regulated rate-limiting step of translation. Initiation requires recruitment of the 40S small ribosomal subunit to the mRNA cap to enable scanning through the 5'UTR. The interaction between the ribosome and the cap is not direct, it is mediated by the cap-binding complex - physically associated to the cap - and the eukaryotic initiation factor (eIF) 3 complex - physically associated to the small ribosomal subunit and other factors that altogether constitute the 43S pre-initiation complex (PIC) (Ryu & Kim 2017). Once recruited to the cap, the PIC scans the 5'UTR, 5' to 3', until the initiator tRNA recognizes the start codon. Then, the scanning mRNP is remodeled such that the 60S large ribosomal subunit is recruited, initiation factors disassemble and elongation ensues (**Figure 3**).

The elongating ribosome travels the mRNA, decoding and adding new aminoacids to the growing peptide chain until it encounters a stop codon. Stop codons are recognized by release factors, rather than aminoacyl-tRNAs, that trigger hydrolysis of the

peptidyl-tRNA bond, releasing the newly synthesized polypeptide. Dissociation of the ribosome, assisted by recycling and initiation factors, follows suit. Recycling yields ribosomal subunits ready for translation initiation or re-initiation (Gerovac & Tampé 2019).



**Figure 4. Modes of 5-to-3' communication in translation initiation.** a) In canonical capped and polyadenylated mRNAs eIF4G interacts simultaneously with eIF4E (4E) and PABP. Paip1 may also establish end-to-end communication and participate in translation initiation by interacting with eIF4A (4A), eIF3 and PABP. b) Histone mRNAs are not polyadenylated but possess a 3' stemloop (SL) that is recognized by the SL-binding protein (SLBP). SLBP-interacting protein (SLIP) binds eIF3 to promote translation initiation. c) The rotaviral RNA GACC 3'-terminal sequence recruits NSP3 dimers that interact with eIF4G to initiate translation. Adapted from Fakim & Fabian 2019.

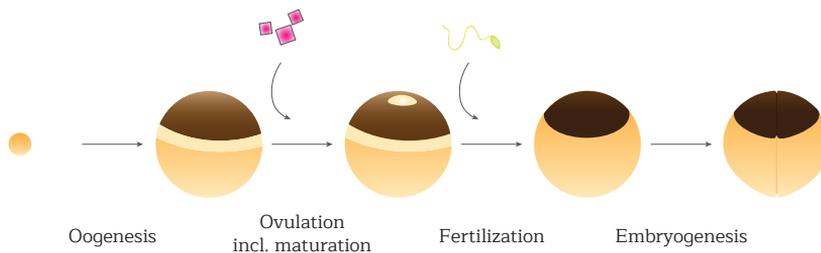
5' to 3' end communication synergistically enhances translation (Gallie 1991) and protects transcripts from degradation. In canonical capped and polyadenylated transcripts, this synergy is due to interaction between the cytoplasmic cap-binding complex, eIF4F - composed of eIF4E, eIF4G and eIF4A -, and the cytoplasmic poly(A)-binding protein, PABPC1. Specifically, between eIF4G and PABPC1 (Imataka 1998, Wells et al. 1998). This constitutes the canonical 5' to 3' communication and is often referred to as the closed-loop model. However, 5' to 3' communication is not limited to canonical mRNAs: replication-dependent histone mRNAs or rotaviral RNAs, both of which are capped but not polyadenylated, also rely on 5' to 3' communication for efficient translation, as illustrated on **Figure 4** (reviewed in Fakim & Fabian 2019). Importantly, 5' to 3' communication is also critical for translational repression and decay.

The last stage of an mRNA's life is inevitably decay. Decay requires destabilization of the structures that confer protection from exonucleases - that is, the cap and the poly(A) tail - or internal cleavage by endonucleases. Decay mechanisms operate at all stages of the mRNA life cycle to ensure transcript fidelity and fine control of gene expression. In eukaryotes, bulk mRNA decay is triggered by reversible poly(A) tail shortening. Deadenylation is followed by degradation from either transcript end: from the 3' end by the exosome complex or from the 5' end, after decapping, by XRN1. Deadenylation and decapping-dependent decay constitutes another example where communication between transcript ends must exist for the effector function.

Several deadenylases have been identified in eukaryotes - e.g. PAN2-PAN3, PARN, Nocturnin, Angell/2 - however, the main cytoplasmic deadenylase activity is the CCR4-NOT complex - which in mammals contains two of four possible catalytic subunits: CNOT6/CNOT6L (CCR4) and CNOT7/CNOT8 (CAF1). In addition to coupling deadenylation to decapping and decay, the CCR4-NOT multi-subunit complex also coordinates translational repression, both by deadenylation-dependent and independent mechanisms (Mayya & Duchaine 2019).

## Oogenesis and anatomy of the growing immature oocyte

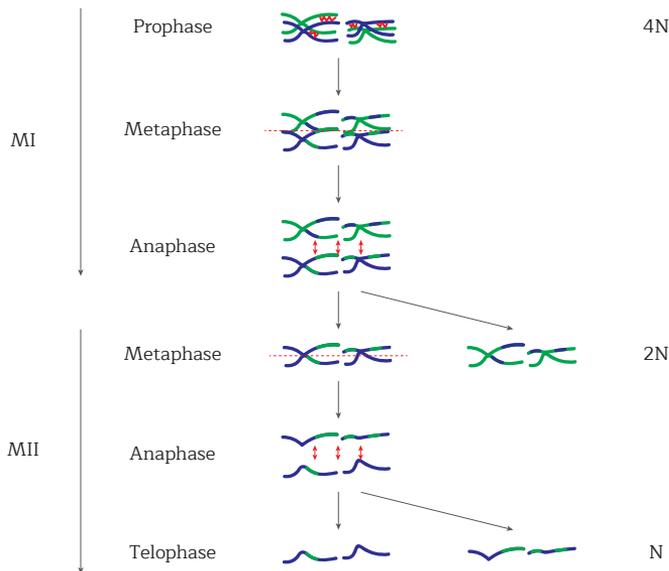
The journey of female cells from diploid stem cells to functional haploid reproductive cells can be broken down into three different processes. It starts in the embryonic ovary with oogenesis - whereby germ cells commit to differentiation - and subsequent oocyte growth during the first meiotic arrest - which yields immature oocytes. It is followed, in the sexually-mature organism, by hormone-induced ovulation, which includes oocyte maturation to fertilization-competency and release from the ovaries into the oviduct.



**Figure 5. Stages in female reproductive cell formation.** Schematic representation of the stages of oocyte and embryo formation in *X. laevis* in accordance to the definitions given in the text. Progesterone is represented by pink rectangles. Redrawn from Sato & Tokmakov 2020.

It culminates with fertilization, which is the sperm-induced transition from eggs to zygotes and thus the start of zygotic and then embryonic development (Sato & Tokmakov 2020) (Figure 5).

Primordial germ cells (PGCs) are a pluripotent population specified in early embryogenesis. PGCs migrate to the site of the developing gonad, the genital ridge. Their arrival to the genital ridge coincides with genital ridge differentiation into a primitive ovary. Concurrently, PGCs initiate differentiation into female gametes – which in the pre-meiotic stages receive the name of oogonia – and expand mitotically. Upon expansion, oogonia enter a special evolutionary-conserved cell division, whereby synchronous mitoses characterized by incomplete cytokinesis result in the formation of clonal cell clusters, called nests or cysts. Meiosis begins once the cysts have mitotically expanded - in *Xenopus*, germline cysts contain 16 ( $2^4$ ) cells or cystocytes.

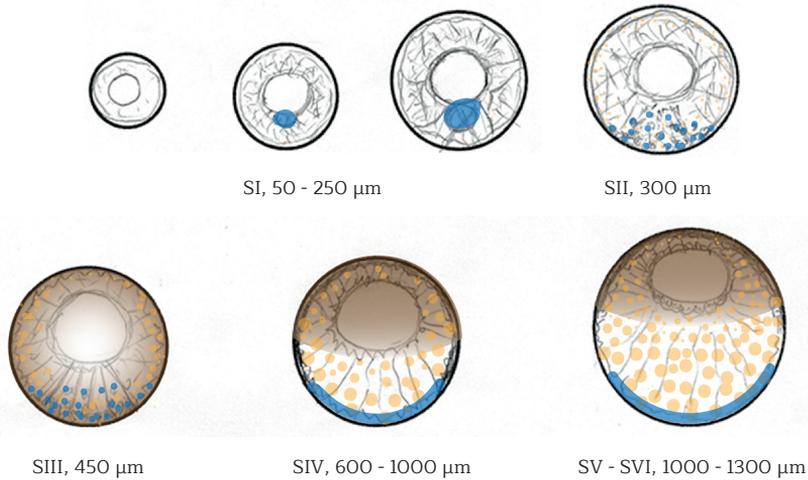


**Figure 6. Schematic representation of meiosis.** Redrawn with inspiration from Chaigne 2014.

Meiosis is the specialized cell division reserved to germinal cells. It consists of two tandem divisions without intervening DNA replication, hence, it produces haploid cells. The first division is called Meiosis I (MI) and the second Meiosis II (MII), all the steps within each Meiosis get the I or II appended to specify which division they belong to. MI is atypical in that genetic recombination takes place between homologous chromosomes. Recombination brings genetic diversity and increases the duration and complexity of prophase I (Pro-I), which is further divided into five stages: leptotene, zygotene,

pachytene, diplotene and diakinesis. MI is also atypical in that, at anaphase, homologous chromosomes segregate instead of sister chromatids (**Figure 6**).

In *X. laevis*, all cystocytes within cysts coordinately enter Pro-I (Kloc et al. 2004), becoming oocytes. Oocytes synchronously progress through Pro-I until, at late pachytene, invading prefollicular cells break down the cysts and oocytes lose their synchrony. As oocytes asynchronously reach diplotene, they become meiotically arrested, a phase that can last up to several years, until the organism's sexual maturity.



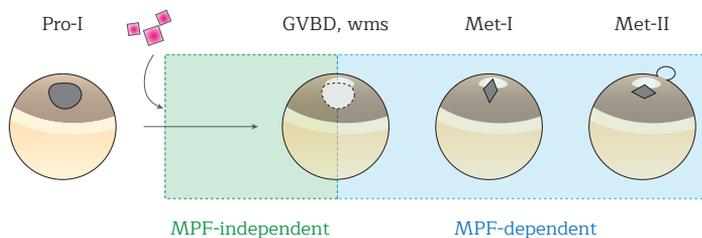
**Figure 7. Staging of oocytes growth as proposed by Dumont (Dumont 1972).** Schematic drawing of *X. laevis* oocytes as they progress from stages I to VI. The Balbiani body and vegetally-localized RNPs are painted blue, yolk platelets are yellow and pigment granules brown. Adapted from Carotenuto and Tussellino 2018.

During this prolonged cell cycle arrest, also called dictyate, oocytes experience enormous growth – from 50-100 to 1300  $\mu\text{m}$  in diameter – as well as other transformations that affect both the oocyte and the surrounding follicular cells (reviewed in Rasar & Hammes 2006). The classification defined by Dumont (Dumont 1972), discretizes the continuous oocyte growth into six morphological stages (**Figure 7**). Stage I oocytes are transparent and ideal for live imaging studies (Powrie et al. 2016). As they progress through stage II (SII), they become progressively translucent. At stage III, pigment synthesis and vitellogenesis begin. During vitellogenesis, which spans until stage V, oocytes uptake liver substrates and accumulate them in yolk platelets. These reservoirs will sustain embryo development until the swimming tadpole stage, in the event of fertilization. Post-vitellogenic stage VI (SVI) oocytes are fully-grown and ready for ovulation.

SVI oocytes are approximately 1  $\mu\text{L}$  in volume, contain up to 4.3  $\mu\text{g}$  of total RNA and more than 500  $\mu\text{g}$  of protein, 80% of which allocated in yolk platelets (Rasar & Hammes 2006). Oocytes are highly asymmetric cells: their big nucleus, or germinal vesicle (GV), is eccentrically positioned and defines the animal pole. The vegetal pole is defined as the opposite side, it has larger and more densely packed yolk platelets. The pigment-containing granules - melanosomes - also sit asymmetrically, as a result, the animal pole is light to dark brown whereas the vegetal pole is pale yellow. These and other asymmetries ensure the asymmetry of the meiotic division, which produces one egg and two small polar bodies (Verlhac & Wingman Lee 2010). In addition, oocyte polarity is important for developmental axes specification (Gerhart et al. 1989).

## Meiotic maturation and fertilization

Once *X. laevis* females attain reproductive maturity, pituitary and ovarian hormone stimulation trigger ovulation. In a broad sense, ovulation can be defined as the ensemble of oocytic and stromal responses that produce mature, follicle-free, fertilization-competent eggs (Sato & Tokmakov 2020). Following this definition, the main processes of ovulation are follicular rupture - that releases the oocytes to the oviduct - and oocyte maturation, which in turn includes maturation of the coat and meiotic maturation. Maturation of the coat is the modification of the glycoprotein-rich extracellular matrix that envelops the oocytes - the vitelline membrane - and its further coating with jelly layers. It happens in the oviduct with participation of epithelial cells (Sato 2014). Meiotic maturation is the oocyte's progression from the primary arrest at Pro-I until a secondary arrest at metaphase II (Met-II) (**Figure 8**).



**Figure 8. Meiotic maturation.** Meiotic maturation is the progesterone-triggered (in pink) process from SVI fully-grown immature oocytes to cytoplasmically mature oocytes. Upon meiotic resumption, de novo protein synthesis leads to MPF activation, which marks MI-entry and drives meiotic progression. GV breakdown (GVBD) and white maturation spot (wms) formation are hallmarks of MI-entry.

Isolated follicle-free *X. laevis* oocytes undergo meiotic maturation upon stimulation with progesterone or other steroids (Goascogne et al. 1985). Because of the uncoupling to maturation of the coat, *ex vivo* meiotic maturation produces cytoplasmically-mature oocytes, albeit not fertilization-competent. Despite this, it constitutes an invaluable model for cell cycle study. It is in this *ex vivo* setting that the biochemistry governing the Pro-I arrest and subsequent release, transition from MI to MII and the metaphase Met-II arrest have been elucidated.

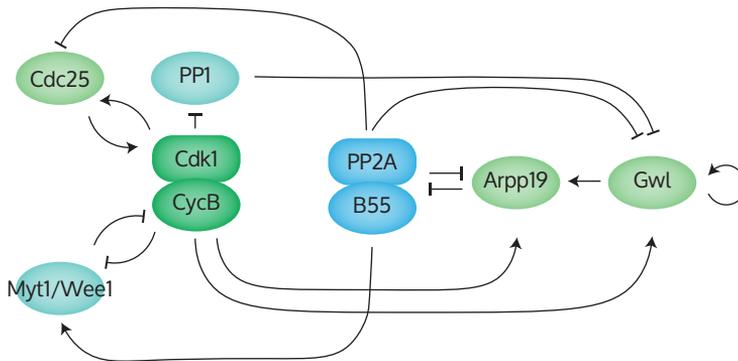
The key activity for meiotic progression is MPF, which initially stood for Maturation Promoting Factor, although it was later renamed M-phase Promoting Factor, owing to the conservation of this activity not only across species but also between meiosis and mitosis. Historically, MPF was defined on a functional basis as the transferable cytoplasmic activity that could trigger meiotic re-entry of an untreated immature recipient. Although not faithful to the historical definition, nowadays most literature uses the term MPF interchangeably with Cdk1:Cyclin-B (CycB) (reviewed in Kishimoto 2018). The Cdk1:CycB dimer is a serine/threonine kinase, where Cdk1 possesses the kinase activity and CycB is a regulatory subunit.

At Pro-I arrest, Protein Kinase A (PKA) keeps the tightly controlled Cdk1:CycB activity at bay. Although both CycB and Cdk1 are present in the oocyte cytoplasm, the complex is subject to inhibitory phosphorylations. In addition, the Cdk1:CycB antagonist phosphatase PP2A:B55 is active. Progesterone signaling triggers meiotic resumption, which is marked by PKA inactivation and subsequent *de novo* mRNA translation. At least three *de novo* synthesized proteins, participate redundantly in Cdk1:CycB activation: Mos, CycB and Ringo/Speedy (see recent review by Meneau et al. 2020). Cdk1:CycB activation, within 3 to 5 hours from meiotic resumption, marks MI entry and meiotic progression.

Besides its role in Cdk1:CycB activation, the germline-specific kinase Mos activates the mitogen-activated protein kinase (MAPK, also known as ERK) signaling pathway. Despite the early synthesis of Mos, MAPK axis activation occurs concurrently with Cdk1:CycB activation at the onset of meiotic progression. The Mos-MAPK axis has essential conserved functions in formation of the meiotic spindles, inhibition of DNA replication in interkinesis and sustaining the MII arrest.

Following initial activation, Cdk1:CycB can sustain its own activation in the absence of the initial trigger. The Cdk1:CycB autoregulatory mechanism has two branches. On the one hand, Cdk1:CycB phosphorylates its direct regulators Cdc25 and Myt1, promot-

ing its active form. On the other hand, Cdk1:CycB inhibits the antagonist phosphatase PP2A:B55 via activation of the Gwl-Arpp19/ENSA axis (Dupre et al. 2013, Mochida et al. 2016). Alternatively, the Cdk1:CycB network can be viewed as two groups of antagonistic proteins, M-phase-promoting (Cdk1:CycB, Cdc25, Gwl-Arpp19/ENSA) versus interphase-promoting (Myt1/Wee1, PP2A:B55, PP1) (Domingo-Sananes et al. 2011) (**Figure 9**). These complex interactions confer MPF its bistable switch characteristic.



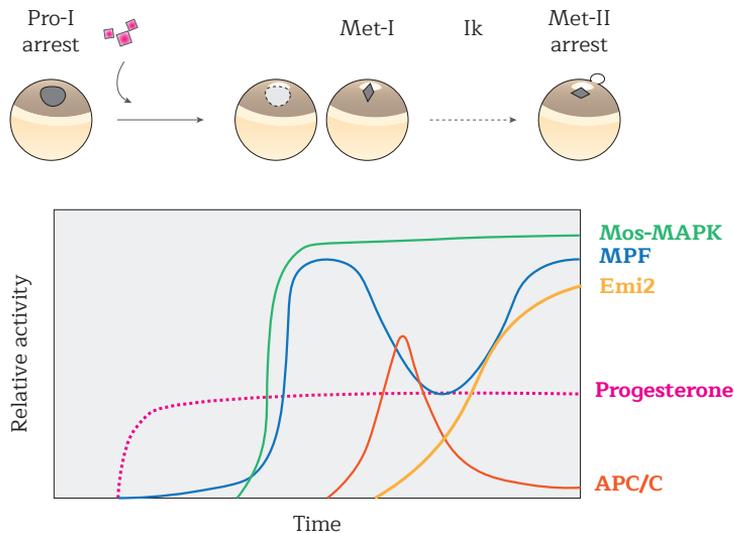
**Figure 9. Players of the MPF bistable switch.** Activating and inhibitory relationships between the kinases and phosphatases implicated in MPF activity. M-phase-promoting proteins represented in green shades, interphase-promoting in blue. Adapted from Kishimoto 2018.

In addition to sustaining its own activation, Cdk1:CycB phosphorylates a myriad of substrates that direct meiotic events. Starting with chromosome condensation, followed by GVBD, spindle formation, spindle relocation to the cortex and alignment of the homologous chromosomes at the metaphase plate. GVBD and spindle relocation to the cortex cause formation of a wms on the oocyte's animal cortex, which is used as a surrogate marker of meiotic progression.

In the same way that the Cdk1:CycB circuitry is self-activating, it is also self-terminating. Cdk1:CycB activates the Anaphase-Promoting Complex/Cyclosome (APC/C) and targets CycB for degradation via APC/C-mediated proteolysis. APC/C is the key activity for anaphase progression, polar body emission and MI exit. In turn, the consequent partial Cdk1 inactivation is necessary for MI exit and interkinesis - the MI to MII transition.

In contrast to a mitotic interphase, interkinesis is brief: the paired sister chromatids remain condensed, the GV is not reformed and the S-phase is omitted. During interkinesis, Cdk1:CycB is reactivated due to multiple actions coordinated by the Mos-MAPK

pathway. Cdk1:CycB reactivation and APC/C inhibition are necessary for MII entry and maintenance of these activities is also crucial for the Met-II arrest that follows. At these stages, the main effector of APC/C inhibition is Emi2, whose activity and stability are controlled by Cdk1:CycB and Mos-MAPK (Tang et al. 2008). The relative activity of the main meiotic kinases is represented in **Figure 10**.



**Figure 10. Key kinase activities of meiotic maturation.** Time-courses of various kinase activities in response to a progesterone pulse. Ik: interkinesis.

The oocytes only overcome the MII arrest and terminate the meiotic division upon fertilization. Fertilization triggers a transient increase in cytoplasmic  $\text{Ca}^{2+}$  concentration that releases APC/C inhibition, allowing it to target cyclins, securins and others for degradation, promoting anaphase, second polar body emission and MII exit. Mechanistically, Calmodulin-dependent Kinase II (CaMKII) senses the calcium rise and phosphorylates Emi2, enhancing Emi2 phosphorylation by yet another kinase, Polo-like Kinase 1 (Plx1). Ultimately, Plx1 phosphorylation targets Emi2 - the key APC/C inhibitor - for degradation by  $\text{SCF}^{\text{FTRCP}}$  (Rauh et al. 2005). In parallel, the calcium wave also activates the phosphatase Calcineurin (CaN), that contributes to APC/C activation by negatively regulating Emi2 and positively regulating APC/C (Heim et al. 2018, Mochida & Hunt 2007, Nishiyama et al. 2007).

Many events occur upon the sperm-induced cell cycle resumption: the contents of cortical granules are released onto the perivitelline space and modify the vitelline mem-

brane, preventing further sperm binding and fusion, MII is completed and the second polar body is emitted, the male and female pronuclei form, chromatin decondenses and the two pronuclei fuse forming a diploid genome. The diploid genome duplicates and proceeds to mitosis, initiating embryonic development.

## Spatiotemporal control of gene expression in meiotic maturation

A unique aspect of meiotic maturation is that it is orchestrated post-transcriptionally. Fully-grown immature oocytes are transcriptionally silent and this global transcriptional silencing is maintained until early development. Specifically, in *Xenopus*, ubiquitous zygotic genome activation (ZGA) is attained at the mid-blastula transition (MBT), after the first 12 mitotic divisions. Thus, all the events that take place in meiotic maturation until MBT are driven by the maternal cytoplasm with the transcripts and proteins accumulated during oocyte growth. Because of the uncoupling of transcription from the gene expression pathway, oocytes provide a privileged scenario to study post-transcriptional gene expression.

Temporally and spatially-regulated translation of subsets of transcripts is essential for meiotic maturation and early development. Some transcripts - which mostly function in mitochondrial and ribosomal biogenesis - are actively translated in prophase-arrested oocytes but are switched off upon meiotic re-entry. Concurrently, other transcripts - associated with cell cycle and embryo development - follow the opposite trend, when produced in the growing oocyte they are stored translationally silent, or masked, in the cytoplasm (Luong et al. 2020, F. Yang et al. 2020). At specific time points upon meiotic resumption masked transcripts are mobilized and engage in translation.

In the oocyte, the translation dynamics just described are often - albeit not exclusively - coupled to cytoplasmic changes in poly(A) tail length. Specifically, repressed mRNPs have shortened poly(A) tails whereas translational activation is preceded by poly(A) tail lengthening (Colot & Rosbash 1982, Dworkin & Dworkin-Rastl 1985, Eichhorn et al. 2016, F. Yang et al. 2020).

### Cis-acting signals and trans-acting factors of cytoplasmic polyadenylation

Historically, two cis-acting signals in the mRNA 3'UTR were found to be key for cytoplasmic polyadenylation-dependent translation activation upon meiotic re-entry:

the ubiquitous PAS hexanucleotide - which is implicated in nuclear polyadenylation and was introduced in “The life cycle of a canonical mRNA” on page 16 - and a U-rich motif termed cytoplasmic polyadenylation element (CPE), consensus UUUUUAU and UUUUAAU (Fox et al. 1989, McGrew et al. 1989).

Following identification of these cis-acting signals, the respective trans-acting factors were characterized. On the one hand, the PAS is bound by a cytoplasmic form of the CPSF complex (Dickson et al. 1999). On the other hand, the CPE is recognized by the CPE-binding protein (CPEB) (Hake & Richter 1994), later renamed CPEB1. Together, these two cis-acting signals and their trans-acting factors define the subset of mRNAs that are masked at Pro-I and become activated early in meiotic maturation.

However, first of all, not all transcripts containing a CPE and PAS conform to this norm and, second, not all transcripts whose translational activation is dependent on the CPE and PAS are activated simultaneously. Transcripts are activated in sequential bursts at distinct times, referred to as polyadenylation waves.

### **A combinatorial code of cis-acting signals**

The current established model is that the number, relative position and combination of cis-acting signals specify translation patterns. These signals can be linear motifs, secondary and tertiary RNA structures or a combination of both and are most often found on the 3' UTR. There have been several attempts, using different approaches, to crack this combinatorial code of cis-acting signals.

Looking at *cyclin B1-B5* 3'UTRs, Piqué et al. (Piqué et al. 2008) formulated the following rules of CPE and PAS-dictated timing and extent of polyadenylation in meiotic maturation. (1) Translational repression is mediated by two or more CPEs, at an optimal distance of 10 to 12 nucleotides regardless of their distance to the PAS. (2) Translational activation requires at least one CPE, whose distance from the PAS is determinant to the extent of activation. The optimal is 25 nucleotides and it needs to be no further than 100 nucleotides and not overlapping it. (3) Another cis-acting signal, Pumilio-binding element (PBE), bound by the RBP Pumilio (PUM2), can also drive activation in combination with a non-consensus CPE. (4) CPEs and PBEs have an additive effect on the extent of activation. (5) Early activation requires CPE(s) that do not overlap the PAS, instead late activation occurs with at least two CPEs, of which one overlaps the PAS. Six modes of regulation derived from these rules were used to make genome-wide predictions and a random sample of predicted behaviours were validated experimentally.

Subsequent work from Belloc et al. (Belloc & Méndez 2008) addressed and added the contribution of AU-rich elements (AREs) to the proposed CPE-PAS combinatorial code. Upon meiotic resumption, C3H-4 - an ARE-binding protein - is synthesized. C3H-4 recruits CCR4-NOT and mediates deadenylation of its targets, hence, in mRNAs with CPEs and AREs these two opposing forces affect poly(A) tail length, resulting in yet finer temporal control of translational activation. In addition, Weill et al. (Weill et al. 2017) elucidated the combinatorial contribution of CPEs and Musashi-binding elements (MBEs) and, thus, the functional interaction between CPEB1 and the RBP Musashi-1. In *mos* 3'UTR, MBEs modulate the preferential use of certain CPEs by producing changes in RNA secondary structure that increase CPE accessibility.

Recently, a number of studies have looked at meiotic maturation and early development in an unbiased way, collecting simultaneous transcriptome and translome data and generalizing some of the early findings derived from candidate-based approaches. For example, Luong et al. (Luong et al. 2020) employed a combination of RiboTag, RNA Sequencing and reporter assays and demonstrated a global switch in translation pattern at Pro-I exit in mouse oocytes. Yang et al. (F. Yang et al. 2020) monitored the transcriptome, translome (by polyribosome analysis) and poly(A) tail length (by Tail-Seq) in *X. laevis* meiotic maturation. In agreement with Luong et al. they found that CPEs are important to escape the general deadenylation imposed upon meiotic resumption and that the density of U-rich CPE-like elements within 100 nucleotides from the PAS correlates with translation activation.

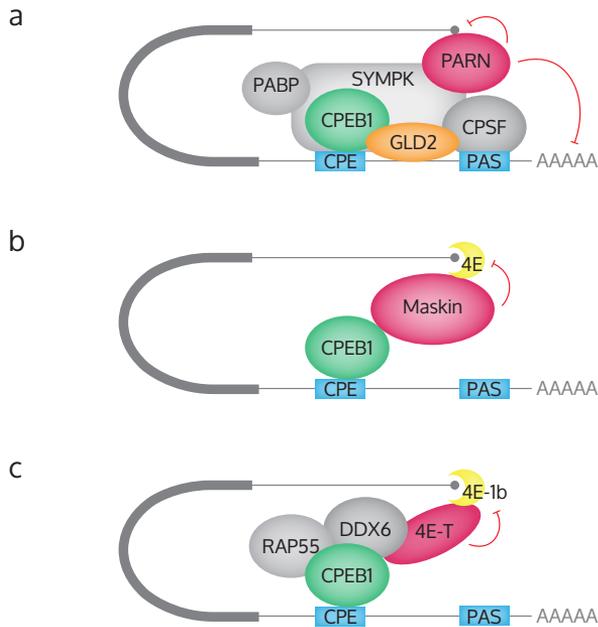
An additional recent investigation by Dai et al. (Dai et al. 2019) looked at *cpeb1*, *cnot6l* and *btg4* 3'UTRs in mouse oocytes and described the capacity of PASs other than the one closest to the 3' end to mediate cytoplasmic polyadenylation. Although some of their findings disagree with previous work, this work raises the possibility that several PAS can simultaneously influence the poly(A) tail length.

### **CPEB1 mRNP composition, architecture and remodeling**

Cis-acting signals in the mRNA are bound by trans-acting factors. All the directly and indirectly-bound proteins determine mRNP identity and ultimately mRNP fate. Thus, understanding the trans-acting factors in translational repression and activation has long been a research focus of many groups.

It is generally accepted that repressive mRNPs maintain a low translation rate by (1) skewing the adenylation/deadenylation activities toward the latter and (2) interfering

with eIF4F assembly, either by binding the m<sup>7</sup>G cap or the cap-binding protein eIF4E. Precluding eIF4F assembly would prevent translation initiation whereas the effect of shortening the poly(A) tail is less understood. Over time, several - often incompatible - models of CPEB1-mediated repression have arisen. The models are summarized in **Figure 11**.



**Figure 11. Mutually exclusive models of CPEB1-mediated repression.** Schematic representation of the different CPEB1-mediated repression models discussed in the text.

The repressive activity has been attributed to PARN, a deadenylase protein whose recruitment by CPEB1 to the mRNP outcompetes the activity of the poly(A) polymerase Gld2, yielding mRNAs with a short poly(A) tail (Barnard et al. 2004, Kim & Richter 2006). In addition, through its cap-binding activity, PARN also prevents eIF4F assembly (Kim & Richter 2006) (**Figure 11a**). Maskin, an eIF4E and CPEB1-binding protein has also been described to be the effector of CPEB1-mediated repression (Stebbins-Boaz et al. 1999) (**Figure 11b**). Alternatively, another eIF4E-binding protein has been involved in CPEB1-mediated repression: eIF4E-Transporter (4E-T) (Minshall et al. 2007). In this last model (**Figure 11c**), CPEB1 interacts - in large mRNPs reminiscent of processing bodies (P-bodies) - with the RNA helicase Xp54 (*X. laevis* homolog of DDX6), Pat100, Rap55 and an oocyte-specific eIF4E, eIF4E1b. This oocyte-specific cap-binding protein is a weak cap binder that prefers 4E-T over eIF4G.

Thus, PARN, Maskin and 4E-T have been proposed to preclude eIF4F assembly in CPEB1 repression complexes. In contrast, PARN is the sole deadenylase claimed to be in these complexes. However, CPEB1 was found in the interactome of the CCR4-NOT catalytic subunit CAF1 (CNOT7/CNOT8). In line, a non-CPEB1-specific RBP-mediated repression mechanism was proposed, where repression is mediated by recruitment of the CAF1-CNOT1-DDX6-4E-T axis (Waghray et al. 2015). This repression mechanism would be eIF4E-independent and highly resemble miRNA-mediated repression (Kamenska et al. 2014, 2016, Meijer et al. 2013, Waghray et al. 2015). In somatic cells, CPEB1-mediated Caf1 recruitment via Tob has been demonstrated to mediate mRNA deadenylation (Ogami et al. 2014). Whether these different mRNPs reflect the simultaneous contribution of independent co-existing compartments or whether they are snapshots of the same mRNP at different maturation stages remains to be addressed.

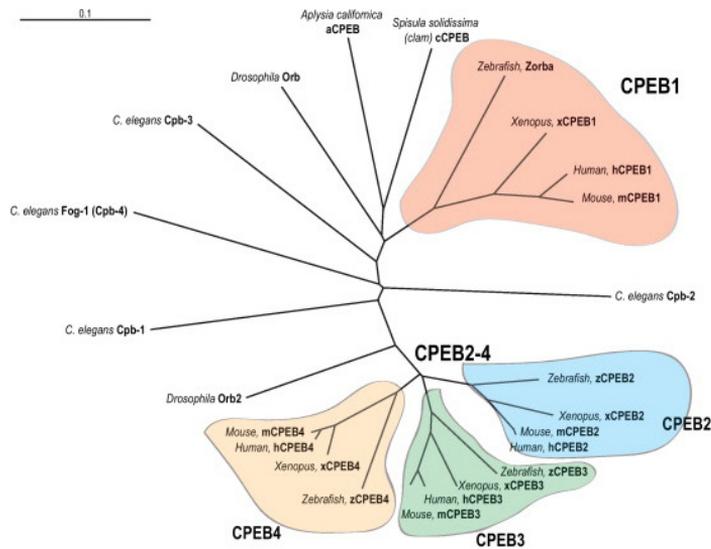
Conversely, so far, the only poly(A) polymerase associated to CPEB1-mediated translational activation is Gld2. Gld2 is already present in repression complexes, however the adenylation/deadenylation balance is skewed toward deadenylation due to PARN's dominance. Upon meiotic re-entry, CPEB1-phosphorylation-induced mRNP remodeling causes PARN to be excluded, allowing Gld2 to polyadenylate by default (Kim & Richter 2006).

## The CPEB protein family across tissues and organisms

The CPEB protein family is represented by a sole member in the surf clam *Spisula solidissima* and sea slug *Aplysia californica*, in the Mollusca phylum (Si et al. 2003, Walker et al. 1999). *Drosophila melanogaster*, in the Arthropoda phylum, possesses two CPEB paralogs, Orb and Orb2. Vertebrates, a subphylum within the Chordata phylum, have four CPEB protein family members, CPEB1-4. Phylogenetic analyses indicate that CPEB genes can be divided in two subfamilies, one represented by Orb, CPEB1 and paralogs, and a second integrated by Orb2, CPEB2-4 and other paralogs (Theis et al. 2003, Wang & Cooper 2010) (**Figure 12**).

Conservation among CPEB paralogs is mostly observed in their C-terminal region. Specifically, sequence identity in this domain is at least 95% in any pairwise comparison among human CPEB2-4, and at most 46% in comparisons to human CPEB1. The C-terminal region consists of two RNA recognition motifs (RRMs) followed by a zinc-binding domain of the ZZ-type (Merkel et al. 2013) and confers the RNA-binding function. The CPEB1 and CPEB2-4 subfamilies regulate overlapping mRNA populations

(Igea & Méndez 2010, Novoa et al. 2010, Ortiz-Zapater et al. 2012, Stepien et al. 2016) since they both recognize linear CPE elements and related U-rich motifs (Stepien et al. 2016). However, the CPEB2-4 subfamily might possess broader substrate specificity including recognition of TTTTGT (CPE-G) motifs (Stepien et al. 2016, Suñer et al. manuscript in preparation).



**Figure 12. The CPEB family of RBPs.** Unrooted phylogenetic tree of representative CPEB proteins across species. CPEB1 vertebrate orthologs (in red) are the most distant members of the family. Reproduced, with permission, from Fernández-Miranda & Méndez 2012.

In contrast to the C-termini, the N-termini of the CPEBs are highly variable in both length and amino acid composition. In this case, sequence identity is, at most, 40% in pairwise comparisons among CPEB2-4 and 15% in comparisons with CPEB1. Despite lacking canonical sequence conservation and not containing any known protein domains, CPEB2-4 N-termini are intrinsically disordered and display conserved low complexity regions (LCRs), such as polyQ, tracts that might indicate functional conservation. Moreover, the N-termini harbor all regulatory cues characterized to date. CPEB1 and CPEB4 are regulated by phosphorylation in their N-terminal domains in meiotic maturation (Guillén-Boixet et al. 2016; Mendez, Murthy et al. 2000; Setoyama et al. 2007). Neuronal CPEB3 is activated by mono-ubiquitylation in the N-terminal domain, although precise mapping of the modification site has not been reported (Pavlopoulos et al. 2011).

## The CPEBs in meiotic maturation

Many studies of CPEB function have focused on oocytes, where CPEB1 was initially identified and characterized. Indeed, the importance of CPEB1/Orb at various stages of female gametogenesis and early development is well established in mice and flies through loss-of-function animal models (Christerson & McKearin 1994, Lantz et al. 1994, Racki & Richter 2006, Tay et al. 2003, Tay & Richter 2001).

In the oocyte, CPEB1 directs polyadenylation of at least two cytoplasmic polyadenylation bursts, or waves. First, in response to progesterone signaling, CPEB1 is phosphorylated at S174. This single phosphorylation is sufficient for mRNP remodeling and translation of 'early' transcripts, necessary for Cdk1:CycB activation and Pro-I exit (Mendez et al., 2000). Second, at MI, CPEB1 is further phosphorylated by Cdk1:CycB and Polo-like kinase I (Plx1). These phosphorylations target CPEB1 for degradation (Setoyama et al., 2007) and the resulting decrease of CPEB1 amount is correlated with translation of 'late' transcripts. Partial CPEB1 destruction and late polyadenylation are required for the MI-MII transition (Mendez et al., 2002).

A third polyadenylation wave is required for MII entry and MII arrest, this one is driven by another member of the CPEB family of proteins, CPEB4. CPEB4 is synthesized as a result of CPEB1-driven early polyadenylation. In addition, CPEB4 function as a translational activator requires multiple phosphorylations by Cdk1:CycB and ERK2 (Guillén-Boixet et al., 2016; Igea & Méndez, 2010). Thus, CPEB4 is produced in a CPEB1-dependent manner and is activated by the kinase that triggers CPEB1 destruction.

Whether CPEB2 and CPEB3 also play roles in cytoplasmic polyadenylation in meiotic maturation and what the functional interactions among all the family members are remain to be investigated.

## In the brain and beyond

The roles of the CPEBs are not limited to the oocyte. In fact, the CPEBs are expressed and have functions across many tissues, such as, but not limited to, sperm differentiation (Tay & Richter 2001, Xu et al. 2012), mammary gland development (Pascual et al. 2020) or liver homeostasis (Maillo et al. 2017).

Importantly, if any tissue comes second to the oocyte in terms of CPEB studies that

is the central nervous system. The four CPEB-paralogues are co-expressed in hippocampal neurons (Theis et al. 2003, Wu et al. 1998), where they participate in localized protein synthesis (Fioriti et al. 2015, Huang et al. 2002, 2003, Udagawa et al. 2012). Localized protein synthesis at synapses produces the necessary morphological and functional changes that allow long-term memory (LTM) formation and its underlying synaptic plasticity (reviewed in Costa-Mattioli et al. 2009). The functional implication of CPEB1 and CPEB3 in LTM has been demonstrated in multiple knockout (KO) mouse models (Alarcon et al. 2004, Berger-Sweeney et al. 2006, Chao et al. 2013, Fioriti et al. 2015).

Perhaps the most interesting insights derived from brain studies concern the molecular mechanisms of CPEB-mediated translational control. In mammalian neurons, CPEB1 mediates CPE-dependent translational repression together with the eIF4E-binding protein Neuroguidin, the poly(A) polymerase Gld2 and the deadenylase PARN (Jung et al. 2006, Udagawa et al. 2012). Synaptic stimulation results in Aurka or CamKII-catalyzed CPEB1 phosphorylation, necessary for polyadenylation-induced translation (Atkins 2004, Huang et al. 2002). Thus, the molecular mechanism of CPEB1 in neurons proposed to date is conserved with the one in the oocyte.

In contrast, CPEB3 has been proposed to exert its translation activator function via the regulated formation of a functional amyloid (Driscaldi et al. 2015, Fioriti et al. 2015; Hervas et al. 2020, Pavlopoulos et al. 2011, Stephan et al. 2015), a mechanism that would be conserved with Orb2 (Majumdar et al. 2012, White-Grindley et al. 2014) and *Aplysia* CPEB (ApCPEB) (Si et al. 2003, 2010). In basal conditions, CPEB3 is sumoylated and resides in P-body-like RNPs, where it represses target mRNAs (Driscaldi et al. 2015, Ford et al. 2019). Upon synaptic stimulation, CPEB3 mono-ubiquitylation triggers amyloid fibrillization, which correlates to polyadenylation-dependent translational activation of target mRNAs (Fioriti et al. 2015, Hervas et al. 2020, Pavlopoulos et al. 2011). Thus, in this model, mono-ubiquitylation switches CPEB3 activity from repressor to activator by promoting a change in molecular state. Finer mechanistic detail on how amyloid fibers activate translation, whether this mode-of-action is applicable to the other CPEB2-4 subfamily members and, if so, what distinguishes brain from oocytic CPEB4 behaviour are some of the open questions in the field.

An alternative model is that CPEB3-target mRNA translation is due to suppression of inhibitory CPEB3 activity, via synaptic-stimulation-induced Calpain-2-catalyzed cleavage (Huang et al. 2006, Wang & Huang 2012). In this case, CPEB3 would not have dual repressor and activator roles.

In non-neuronal somatic cells, CPEB3 and CPEB4-mediated mRNA deadenylation has been shown to rely on interaction with Tob and subsequent Caf1 recruitment (Hosoda et al. 2011). Instead, CPEB2 - which arguably remains the most enigmatic CPEB - has been demonstrated to downregulate translation at the elongation step, via direct interaction with the elongation factor (eEF) 2 (Chen & Huang 2012).

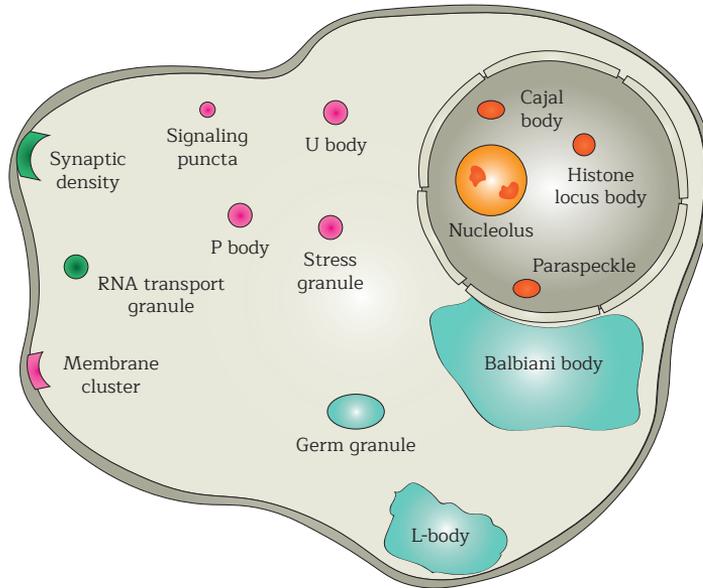
## Higher-order mRNP organization

In recent years, the number of nucleic-acid binding proteins that participate in membrane-less compartments has exploded. In fact, membrane-less compartments have emerged as a common organizing principle in the cell and we now understand some of the biophysical principles behind their regulated formation. **Figure 13** summarizes some of the known membrane-less compartments in eukaryotic cells.

Membrane-less compartments are micron-scale structures not delimited by membranes that concentrate protein and nucleic acids at discrete sites. These condensates are thought to assemble through a type of phase transition, specifically spontaneous separation of a homogenous solution into two distinct liquid phases, hence liquid-liquid phase separation (LLPS). Macromolecules are not found homogeneously in both phases, but are rather depleted in one phase and enriched in the other (see reviews by (Alberti 2017, Alberti et al. 2019, Banani et al. 2017)). Compartments formed by LLPS, such as nucleoli or *Caenorhabditis elegans* P granules, exhibit liquid-like properties; they are round, they can deform and are able to fuse with one another, plus, their protein components are mobile and exchange rapidly with the surrounding cytoplasm (Brangwynne et al. 2009, 2011). However, the degree of “liquidity” of membrane-less compartments is a continuum that ultimately depends on their specific composition, environmental conditions as well as their maturation process - with time, liquid compartments can further transition into gel, glass or solid-like states (Alberti et al. 2019).

In fact, maintaining condensate fluidity is an active ATP-dependent process that involves chaperones and disaggregases. Improper maintenance of fluidity - dubbed molecular aging - that causes droplet hardening has been linked to neurodegenerative diseases (reviewed in Li et al. 2013, Ramaswami et al. 2013). However, in other cases, condensates with solid-like properties are functional. Examples of solid-like condensates include: Balbiani bodies, involved in preserving dormancy of vertebrate oocytes (Boke et al. 2016); localization-bodies (L-bodies), involved in vegetal transport and localization in *X. laevis* oocytes (Neil et al. 2020); yeast Rim4 aggregates, involved in

translational repression in gametogenesis (Berchowitz et al. 2015); ApCPEB, Orb2 and vertebrate CPEB3 aggregates, involved in memory formation (reviewed in Si & Kandel 2016). Among the cited examples, several have been characterized as amyloid-like, which is a specific cross- $\beta$  fibrillar conformation.



**Figure 13. Subcellular structures formed by phase-separation.** Schematic representation of cellular compartments proposed to form by phase-separation. Some are cell-type specific, such as the Balbiani body, germ granules and L-bodies (in blue) which are germ-cell-specific or synaptic densities and RNA transport granules which are neuronal-specific (in green). Adapted from Banani (Banani et al. 2017).

The physical basis for LLPS of RNA-protein complexes is multivalency, both in the form of protein-protein and protein-RNA interactions (Li et al. 2012). At the protein level, these interactions are established by either multiple modular domains or regions of intrinsic disorder. Intrinsically-disordered proteins/regions (IDPs/IDRs) exhibit conformational flexibility, that is, they transition between multiple low-energy conformations. IDRs are by definition compositionally-biased, they are deficient in hydrophobic residues and enriched in polar and charged residues. In addition, IDRs often contain LCRs, which are stretches of even stronger compositional bias, such as prion-like domains or RGG boxes (Chong et al. 2018, Franzmann & Alberti 2019).

Beyond their common molecular grammar, membrane-less compartments are func-

tionally diverse. They have been implicated in ribosome biogenesis, splicing, translational repression and activation, signal transduction, nucleation of cytoskeletal structures, organelle storage and memory formation among others. The proposed advantages and implications of phase separation in these processes are multifarious (reviewed in Alberti et al. 2019). Briefly, phase separation is viewed as a fast-sensing mechanism that can buffer cellular concentrations and achieve high local concentrations - which in turn can both favor and inhibit reactions. In addition, condensates can have emergent properties, which are properties that arise in the condensate that its constituents do not possess on their own. For example, condensates could generate mechanical forces or act as mechanical filters.

There is evidence that the CPEBs can form or are found in phase-separated compartments. Endogenous CPEB1 is found in multi-phasic mRNPs associated to vegetal mRNA localization, termed L-bodies, in early stage *X. laevis* oocytes (Neil et al. 2020). In HeLa cells, CPEB1 is also found in cytoplasmic hubs, where it co-localizes with the P-body proteins Dcp1 and GW182 and with the stress granule (SG) marker eIF2 upon stress (Wilczynska et al. 2005). Full-length CPEB4 as well as its N-terminal domain can phase separate in vitro and form condensates in an ectopic overexpression system (Guillén-Boixet et al. 2016). Neuronal CPEB3, Orb2 and ApCPEB form amyloid fibrils upon synaptic stimulation (Fioriti et al. 2015, Hervas et al. 2020, Si et al. 2010). Last but not least, in Zebrafish oocytes, the CPEB1 targets *mos* and *cyclin B1* mRNAs are found in different granules that are disassembled shortly after induction of oocyte maturation (Horie & Kotani 2016; Kotani et al. 2013). Whether the CPEBs are clients or scaffolds of phase-separated compartments, which specific compartments they are found in, what is their composition and internal organization and what is the relevance of phase separation for CPEB function are important questions yet to be addressed.

# Aims



The aim of this thesis was to unveil the contribution of the different members of the CPEB-family of RNA binding proteins to gene expression regulation during cell cycle. We focused on expression and regulation of the least characterized family members in our model system - CPEB2 and CPEB3 - and we subsequently addressed the contribution of the whole CPEB-family to gene expression regulation. The specific questions we pursued are the following:

1. Determine whether CPEB2 and CPEB3 are expressed in meiotic maturation.
2. Characterize the expression dynamics of CPEB2 and CPEB3.
3. Characterize the regulatory mechanisms that affect CPEB2 and CPEB3.
4. Identify the enzymes responsible for CPEB2 and CPEB3 post-translational modification.
5. Investigate the effect of post-translational modifications on CPEB2 and CPEB3 function.
6. Define a subset of high-confidence CPEB1-4 proximal proteins.
7. Define and compare the target mRNAs bound by CPEB1-4.



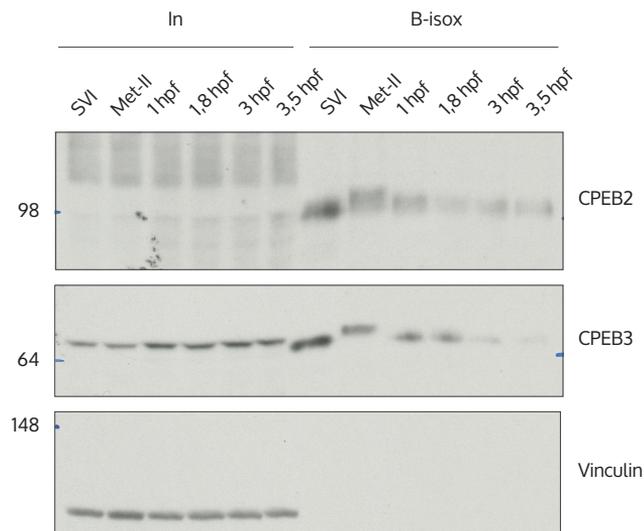
# Results



## The CPEBs are co-expressed in *X. laevis* oocytes

In order to understand how the CPEB family of RNA-binding proteins as a whole regulates gene expression, we chose meiotic maturation as our scenario. Meiotic maturation is a cell-cycle synchronized model whereby the sequential roles of CPEB1 and CPEB4 are well-established. Thus, we started by asking whether the remaining family members, CPEB2 and CPEB3, were also expressed and if so, which were their expression dynamics.

We performed biotinylated isoxazole (b-isox) precipitation followed by tandem mass spectrometry (MS/MS) and, complementarily, b-isox precipitation followed by polyacrylamide gel electrophoresis (PAGE) and Western blot (WB) of oocyte lysates at three maturation stages: SVI, MI-entry (marked by GVBD, in turn scored by wms formation) and Met-II-arrest (at least three hours after wms) (**Table 1** and **Figure 14**). B-isox is a small molecule that has the ability to precipitate RNP granule proteins, presumably because b-isox crystals can capture extended  $\beta$ -strand conformations that IDPs/IDRs easily adopt (Han et al. 2012, Kato et al. 2012). Therefore, we reasoned that we would be able to enrich the CPEBs with this compound since they are RNP granule proteins with IDRs.



**Figure 14. Endogenous CPEB2 and CPEB3 expression by b-isox precipitation followed by PAGE-immunoblot in meiotic maturation and early development time-courses.** Representative immunoblots, CPEB2 n = 1, CPEB3 n = 4. Vinculin is used as a loading control and precipitation negative control. Hpf: hours post-fertilization.

As shown in **Table 1**, we were able to detect all four CPEBs with at least one PSM in at least one time-point. We detected CPEB1 at all time-points, with its highest expression at SVI and decreasing thereafter, consistent with its partial degradation. We detected unique peptides of CPEB2-4 and also shared peptides that were impossible to assign, owing to the high degree of sequence identity in the C-terminal domain. As shown in **Figure 14**, we detected both CPEB2 and CPEB3 following PAGE-WB of b-isox precipitates with specific antibodies. Both proteins were expressed at SVI and throughout the maturation and embryonic time-courses with varying mobilities, most likely, caused by PTMs. Specifically, the CPEB3 mobility shift was observed at MI-entry (**Figure S1** on page 123), coinciding with the time-point of CPEB4 activation by ERK2 and Cdk1 (Guillén-Boixet et al. 2016, Igea & Méndez 2010).

**Table 1. Endogenous CPEBs expression detected by MS/MS in b-isox-precipitated meiotic maturation lysates, expressed as PSMs, n = 1.**

Protein	SVI	GVBD	Met-II
CPEB2			1
CPEB3	12		4
CPEB4	3	3	2
CPEB2; CPEB3	3	1	2
CPEB2; CPEB3; CPEB4	15	9	14
CPEB1	839	167	27

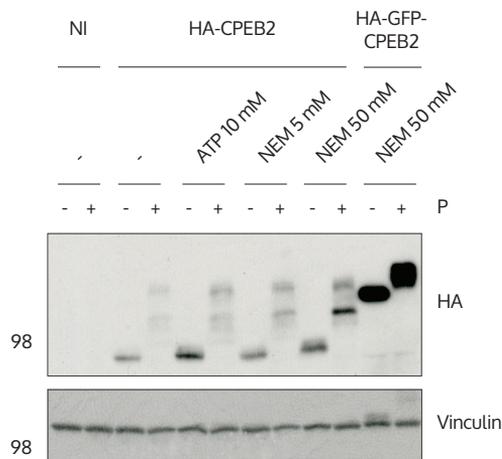
Independent high-throughput interrogations of the *X. laevis* oocyte proteome have also detected multiple CPEBs. Specifically, a phosphoproteomics study by Peuchen et al. identified CPEB1 and CPEB3 phospho-peptides at all time-points between SVI oocytes and 2-cell embryos (Peuchen et al. 2017). Deep genome-free proteomics of the egg by Wuhr et al. that identified and estimated the concentration of 11000 proteins detected CPEB1 and CPEB2 and estimated their concentrations in 43 and 2 nM, respectively (Wuhr et al. 2014). Why the whole CPEB2-4 subfamily has not been identified in these studies is probably because of their low expression in combination with the high sequence identity of the CPEB2-4 C-termini – since, usually, only uniquely-mapping peptides are considered. Transcripts per million (TPM) calculated from RNA-Seq datasets also lend support to the idea that CPEB2-4 are less abundant than CPEB1 in the oocyte (Session et al. 2016).

Taking together our findings about CPEB2 and CPEB3, added to prior knowledge about CPEB1 and CPEB4, we found that all CPEBs are co-expressed at SVI and modified during the meiotic cell cycle. We determined that CPEB3 is modified at meiotic progression, like CPEB4 and unlike CPEB1, which is targeted for degradation at meiotic

progression. Thus, at the expression level, CPEB1 and CPEB2-4 differ in their expression dynamics, concentration and time-point of regulation in oocytes.

## CPEB2 and CPEB3 are phosphorylated in meiotic maturation

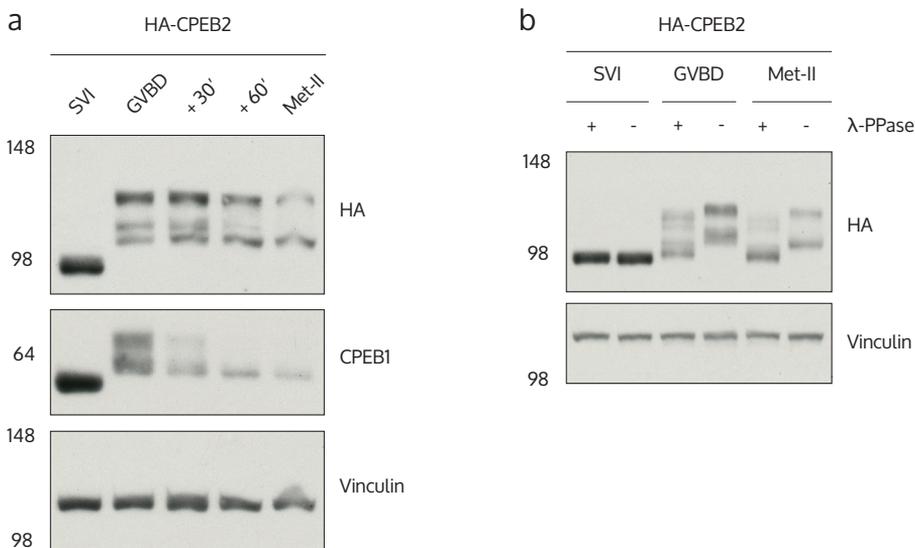
Based on the prior knowledge that both CPEB1 and CPEB4 are regulated by phosphorylation - albeit with opposite consequences for their function (Guillén-Boixet et al. 2016; Mendez et al. 2002; Mendez, Hake et al. 2000; Mendez, Murthy et al. 2000; Setoyama et al. 2007) - and in light of the changing mobility of CPEB2 and CPEB3 in PAGE-WB (**Figure 14**), our second question was whether CPEB2 and CPEB3 are also regulated by phosphorylation. To this aim we performed time-course analyses in oocytes overexpressing the proteins of interest.



**Figure 15. Overexpressed CPEB2 detection differs with lysis conditions.** Representative immunoblot (n = 2) showing different lysis conditions and an N-terminal GFP fusion. NI: not-injected. P: progesterone. MW, in kDa, indicated on the left side of the film scans. Vinculin immunoblots serve as loading controls.

Detection of overexpressed HA-CPEB2 proved to be more difficult than detection of the other CPEBs. To optimize it, we tested different lysis protocols, as well as different CPEB2 constructs (**Figure 15**). Lysis with addition of 10 mM ATP - millimolar ATP concentrations solubilize hydrophobic proteins (Patel et al. 2017) - provided a modest improvement in CPEB2 recovery, most noticeable on its fastest-migrating least-modified form. Lysis with addition of the irreversible cysteine-peptidase inhibitor N-ethylma-

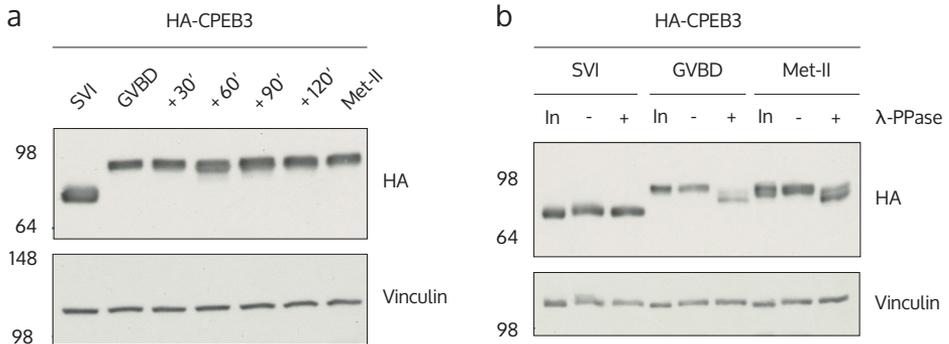
leimide (NEM) - an alkene reactive towards thiol groups - had a concentration-dependent effect on CPEB2 recovery in all its bands, higher NEM concentration resulted on better recovery. We also tested two reciprocal Green Fluorescent Protein (GFP) fusion constructs. We found that an N-terminal GFP fusion improved CPEB2 detection. Instead, the reciprocal C-terminal fusion provided no improvement (data not shown). In light of these results, we used 100 mM NEM in the lysis buffer of subsequent time-course experiments, including those involving CPEB3, in order to set a comparable scenario.



**Figure 16. CPEB2 is phosphorylated throughout meiotic maturation.** a) Representative immunoblot ( $n = 2$ ) of polyadenylated HA-CPEB2 mRNA expression in a maturation time-course. b) Representative immunoblot ( $n = 2$ ) of a lambda phosphatase assay of polyadenylated-HA-CPEB2-microinjected oocytes at different maturation time-points. In both experiments, oocyte lysis was performed in the presence of 0.1 M NEM. MW, in kDa, indicated on the left side of the film scans. Vinculin immunoblots serve as loading controls.

Maturation time-course experiments of HA-CPEB2-overexpressing oocytes revealed an increase in apparent molecular weight (MW) at GVBD and all subsequent time-points (**Figure 16a**). MI-entry produced at least three discrete slower migrating bands. Interestingly, the overall protein intensity decreased over time. When we incubated stimulated lysates with lambda Protein Phosphatase ( $\lambda$ -PPase, **Figure 16b**), the mobility shift was partially abrogated. Hence, the shifts were due, at least partially, to phosphorylation.

Similarly, when we overexpressed HA-CPEB3, we observed a shift in apparent MW at GVBD that was sustained until the Met-II arrest (**Figure 17a**). In this case, the amount of protein remained constant throughout the time-courses. In addition, the mobility shift was partially lost upon incubation with  $\lambda$ -PPase (**Figure 17b**), indicating that CPEB3 was also phosphorylated.

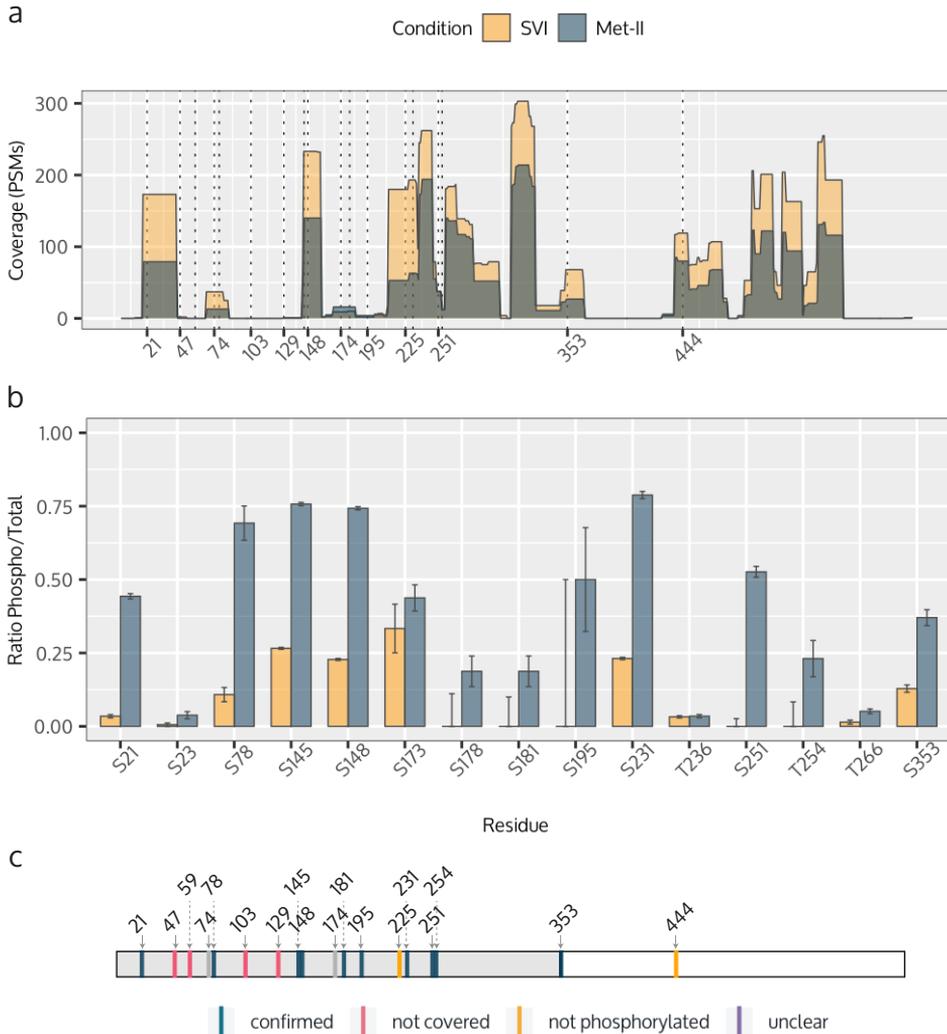


**Figure 17. CPEB3 is phosphorylated throughout meiotic maturation.** a) Representative immunoblot ( $n = 2$ ) of polyadenylated HA-CPEB3 mRNA expression in a maturation time-course. b) Representative immunoblot ( $n = 2$ ) of a Lambda phosphatase assay of polyadenylated-HA-CPEB3-microinjected oocytes at different maturation time-points. In both a) and b), oocyte lysis was performed in the presence of 0.1 M NEM. In: input. P: progesterone. MW, in kDa, indicated on the left side of the film scans. Vinculin immunoblots serve as loading controls. CPEB1 immunoblot is included as a loading and maturation control.

From these results we conclude that – like CPEB1 and CPEB4 - CPEB2 and CPEB3 are regulated by phosphorylation, implying that the entire CPEB family of RBPs is regulated by phosphorylation in meiotic maturation. CPEB2 and CPEB3 are phosphorylated at meiotic resumption, like CPEB4, however, CPEB2 levels decrease throughout maturation whereas CPEB3 expression is sustained.

## CPEB3 is phosphorylated at proline-directed sites clustered on the N-terminal domain

We next asked which were the phosphorylated residues in CPEB2 and CPEB3. We followed two complementary approaches for phosphorylation-site (p-site) mapping: an unbiased screening by protein immunoprecipitation (IP) followed by MS and generation of phosphorylation-mimetic and phosphorylation-null mutants by targeted mutagenesis of candidate sites.



**Figure 18. CPEB3 p-site mapping by MS.** a) Coverage, in counts, of each position throughout the protein sequence. SP-sites are indicated with dotted lines. b) Phospho-to-total occurrences ratio determined by MS for the indicated polypeptide positions on SVI and Met-II oocytes. The ratios are calculated from the pool of 4 biological replicates. Only positions where there is a relative gain of phosphorylation are displayed. Error bars represent  $\pm$  ratio error. c) Summary of the phosphorylation status of all proline-directed sites according to the results displayed on a) and b).

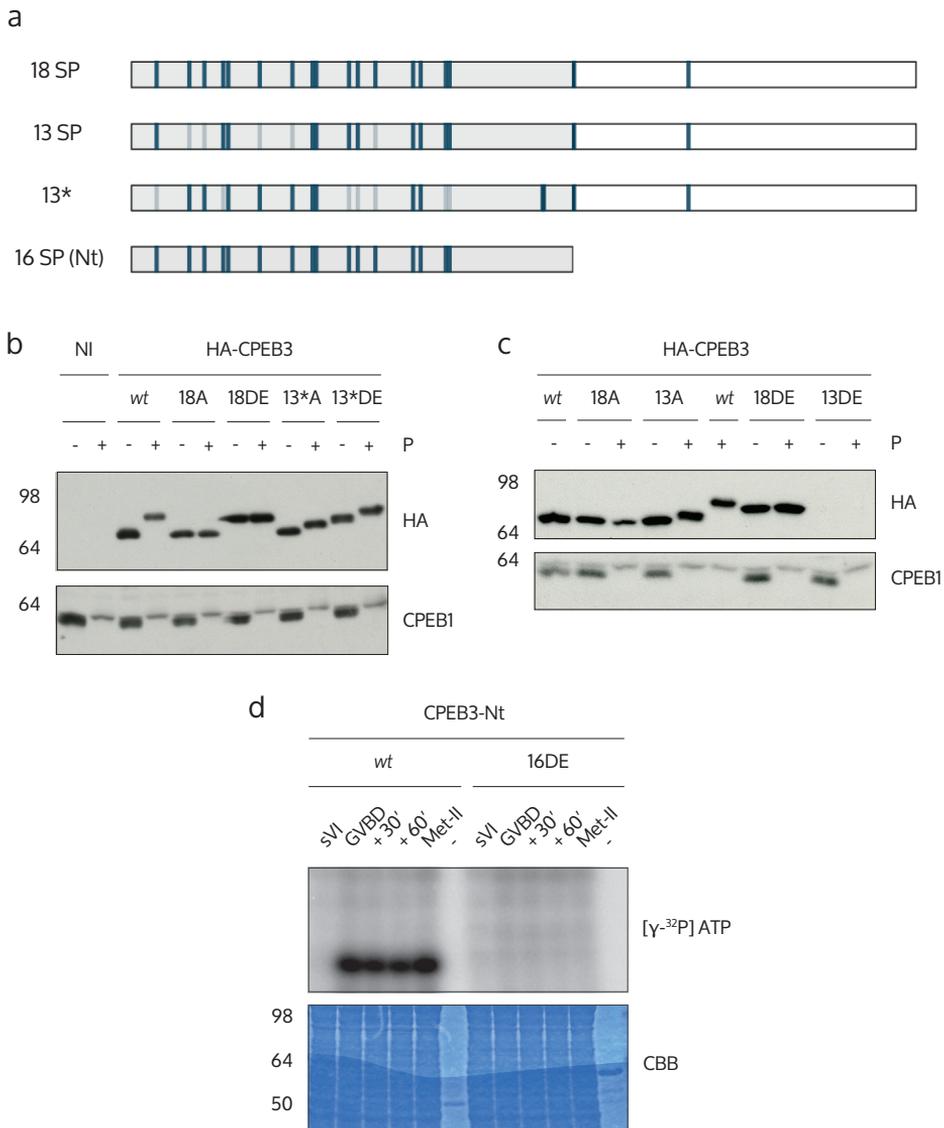
The unbiased screening by MS revealed that multiple residues in CPEB3 were phosphorylated upon progesterone stimulation. **Figure 18** panel a shows the coverage of CPEB3 in the pool of MS mapping experiments. In **Figure 18b**, only the residues in CPEB3 for which we detected a relative gain of phosphorylation upon progesterone

stimulation are plotted. The values from which **Figure 18b** was constructed can be found on “Appendices”, **Table S1** on page 123. Briefly, we detected a relative gain of phosphorylation on eleven sites, namely S21, S78, S145, S148, S178, S181, S195, S231, S251, T254 and S353. Except for S178, all these sites were consensus proline-directed sites. Overall, these proline-directed sites represented ten out of the eighteen proline-directed sites on CPEB3.

Regarding the eight proline-directed sites for which we did not detect a gain of phosphorylation, four - S47, S59, T103 and S129 - were not captured in any peptide in our experiments, as reflected by the coverage (**Figure 18a**). Missing data is a well-known limitation of label-free MS that can have multiple sources (O’Brien et al. 2018). For example, the endopeptidases cut sites or the resulting peptides’ ionization efficiency. Therefore, the number of observed peptides can usually be increased by using multiple endopeptidases for fragmentation. For this reason, we used trypsin and chymotrypsin in our assay, and still did not capture any peptides containing the aforementioned positions. For another two of the candidate proline-directed sites - S225 and S444 -, the evidence indicated that they were not phosphorylated upon progesterone stimulation. Last, for two sites - T74 and S174 -, although the ratios suggested no gain of phosphorylation, they were found in phosphorylated peptides, whereby the phosphorylation was assigned to another residue - S78 and S173 or S178, respectively. Fine mapping of the phosphorylation sites in these multiply phosphorylated peptides would require MS mapping with mutant versions of CPEB3.

To summarize the MS mapping - as illustrated on **Figure 18c** - ten out of eighteen proline-directed sites were phosphorylated upon progesterone stimulation, four were not observed with this experimental setup, two were unclear and two were not phosphorylated. Interestingly, all phosphorylations captured occurred on the N-terminal domain of CPEB3.

Complementarily, we performed targeted mutagenesis of candidate phosphorylation sites. On the one hand, the candidate serine or threonine residues were replaced by alanine, resulting in non-phosphorylatable constitutively-unphosphorylated versions of the protein (referred to as phosphorylation-null, p-null or by indicating the number of substituted residues followed by A). On the other hand, the candidate serine or threonine sites were substituted by aspartic or glutamic acid, respectively, resulting in non-phosphorylatable constitutively-phosphorylated mimic versions (referred to as phosphorylation-mimetic, p-mim or indicating the number of mutated residues followed by DE). With both non-phosphorylatable mutants, we expected to abrogate the



**Figure 19. CPEB3 p-site mapping by targeted mutagenesis.** a) Summary of mutants. Vertical lines indicate the relative position of mutated residues. b) Representative immunoblot [18(SP),  $n = 6$ ; 13\*,  $n = 3$ ] of the apparent MW shift of 18(SP) and 13\* mutants upon progesterone stimulation relative to *wt* HA-CPEB3. c) Representative immunoblot [13(SP),  $n = 3$ ] of the apparent MW shift of 18 and 13(SP) mutants upon progesterone stimulation relative to *wt*. NI: not-injected. P: progesterone. CPEB1 included as a loading and maturation control. d) Representative autoradiography ( $n = 2$ ) of radiolabeled ATP incorporation by *wt* or 16DE CPEB3-Nt upon incubation with oocyte lysates at the indicated maturation time-points. CBB-stained gel shown as loading control. Nt: N-terminal. CBB: Coomassie brilliant blue. [ $\gamma$ -<sup>32</sup>P] ATP: Phosphorous-32-labelled ATP. MW, in kDa, on the left side of the film scans.

mobility retardation induced by progesterone stimulation on SDS-PAGE.

We tested three mutants by SDS-PAGE, represented on **Figure 19a**: a mutant of all proline-directed sites (18 SP), a mutant of 13 proline-directed sites (13 SP), where the sites left *wild type* (*wt*) correspond to the sites with little or no coverage in the MS mapping (i.e. S47, S59, T103, S129 and S195, **Figure 19c**) and a mutant of 13 serines and threonines including some that are not proline-directed (13\*). In the 18 SP mutants - both phosphorylation-null and mimetic - the band retardation induced by progesterone was fully abrogated (**Figure 19**, b and c). Instead, in both 13\* and 13 SP mutants there was still a shift upon stimulation (**Figure 19b**). Importantly, the 13 SP mutants provided information that complemented the MS results: although its mimetic version was not expressed, the mobility shift of the null indicated that at least some of the 5 sites not captured by MS were getting phosphorylated (**Figure 19c**). So, in short, mutating 18 proline-directed sites blocked the mobility shift while mutating any less still produced a progesterone-induced mobility shift. Moreover, we captured phosphorylation at the residues not observable by MS.

A more direct and quantitative approach to measure phosphorylation are *in vitro* phosphorylation assays, whereby phosphorylation of a recombinant protein substrate is measured by radiolabeled ATP incorporation. We did *in vitro* phosphorylation assays of bacterially-produced *wt* and non-phosphorylatable CPEB3 N-termini to further support our findings (**Figure 19a**). We found, in line with previous results, that *wt* CPEB3 N-terminus incorporated radiolabeled ATP at all time-points from GVBD until Met-II. In contrast, the non-phosphorylatable mutant did not, indicating that all possible phosphorylation sites were blocked.

Taken together, the MS and targeted mutagenesis mapping of phosphorylation sites revealed that CPEB3 was phosphorylated in as many as 16 proline-directed sites that lied in its N-terminal domain.

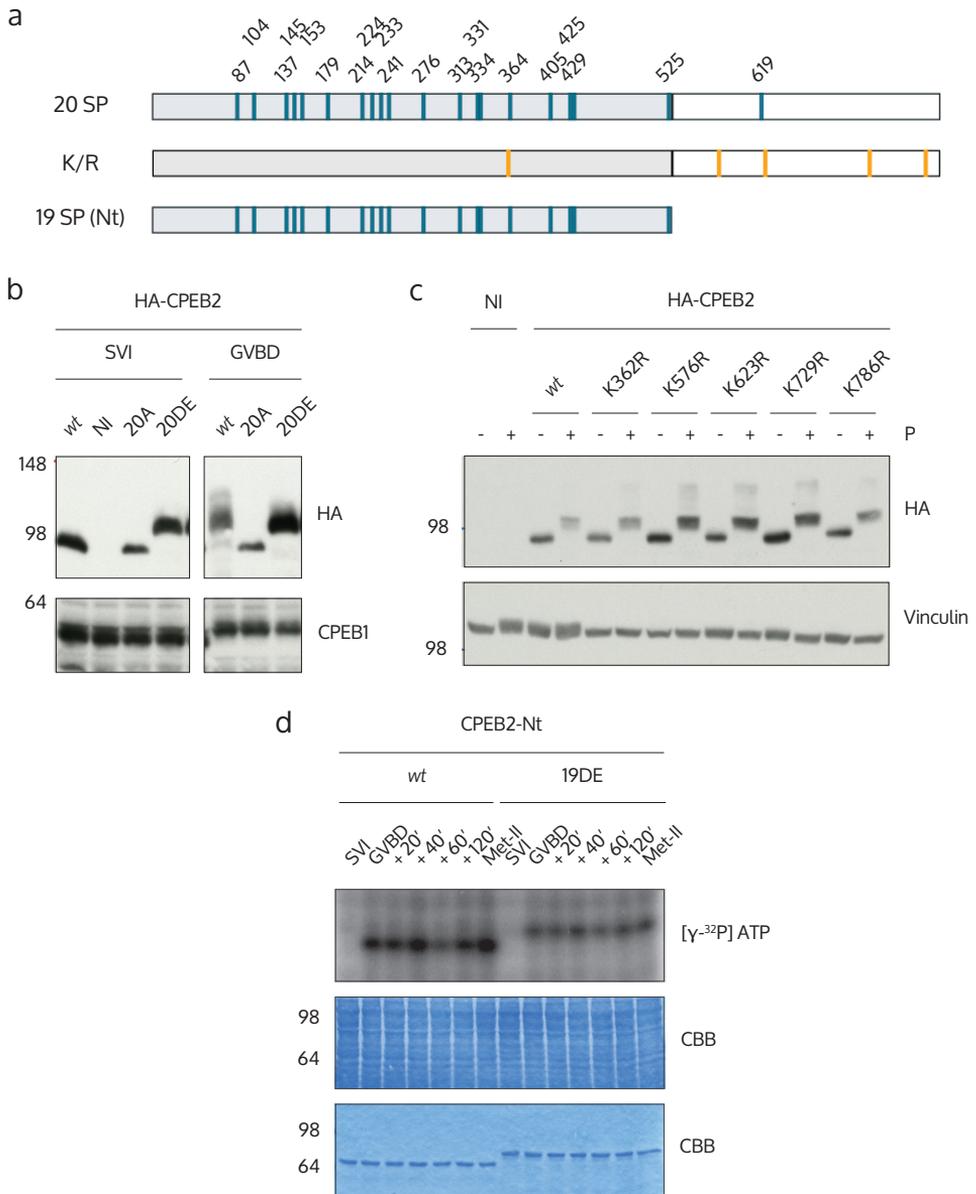
## Bulk CPEB2 phosphorylation also occurs at proline-directed sites

To map the regulatory sites on CPEB2 we followed the targeted mutagenesis approach because of the difficulty of purifying sufficient CPEB2 for MS assays. Our starting hypothesis was that CPEB2 would also be regulated at proline-directed sites given that the mobility shift is concurrent with the activation of proline-directed kinases.

In consequence, we generated the phosphorylation-null and phosphorylation-mimetic non-phosphorylatable mutants of the 20 proline-directed sites in the protein (20 SP: 20A and 20DE, respectively, **Figure 20a**). We observed that, analogously to CPEB3, blocking all these sites fully abrogated the mobility shift induced by progesterone (**Figure 20b**). Intriguingly though, the 20 SP mutants did not mimic the complex multiple-band shift of *wt* CPEB2. We speculated that the slower migrating forms observed in the *wt* could correspond to additional PTMs.

We reasoned that the additional PTM might be mono-ubiquitylation or sumoylation given that, in mouse neurons, CPEB3 activity is regulated by these modifications (Drisaldi et al. 2015, Pavlopoulos et al. 2011). We addressed whether CPEB2 was being sumoylated at five out of thirty possible lysines in the protein based on three predictors of sumoylation. GPS-SUMO (Zhao et al. 2014) predicted a non-consensus sumoylation site at K786, JASSA (Beauclair et al. 2015) returned K729, and SUMOgo (Chang et al. 2018) K623 and K576. These four predicted sumoylation sites are on the C-terminal region and would therefore be conserved between CPEB2-4. In addition, we tested K362, because it is the only consensus phosphorylation-dependent sumoylation site. We separately mutated each of these lysines to non-sumoylatable arginines and micro-injected the corresponding mutants to oocytes (lysine mapping represented on **Figure 20a**, result on **Figure 20c**). All five mutants exhibited the same migration at SVI and in response to progesterone than the *wt* protein, indicating that neither of these residues harbored the PTMs responsible for the MW shift observed.

The additional PTMs could also be phosphorylations. In fact, in the high-throughput phosphoproteomics investigation by Peuchen and colleagues (Peuchen et al. 2017), it was proposed that proline-directed kinases, responsible for bulk phosphorylation in meiotic maturation, could work in concert with broad-specificity kinases CK2, CK1 and GSK3 to generate hyperphosphorylated clusters. When we performed *in vitro* phosphorylation assays with bacterially-produced CPEB2 N-terminus and monitored radiolabeled ATP incorporation, we observed that *wt* CPEB2 incorporated radiolabeled ATP from GVBD until MII (**Figure 20d**). The mutant of all proline-directed sites (19DE) incorporated less signal than the *wt*, but incorporated signal nonetheless. This information supported the hypothesis that there were additional phosphorylation sites on CPEB2. Importantly, despite not having mapped all phosphorylation sites on CPEB4, Guillén-Boixet et al. demonstrated that 10 out of 12 phosphorylations mapped were sufficient for full activation of CPEB4-mediated polyadenylation (Guillén-Boixet et al., 2016). They proposed that CPEB4 could operate as an ultrasensitive switch, a system whereby only about half of the sites are required for activation and extra sites contrib-



**Figure 20. CPEB2 p-site mapping by targeted mutagenesis.** a) Summary of mutants. Candidate S/T indicated in blue, K in pink. b) Representative immunoblot (n = 3) of the apparent MW shift of the 20(SP) mutants upon progesterone stimulation relative to *wt* HA-CPEB2. c) Representative immunoblot (n = 2) of the apparent MW shift of the candidate-lysine mutants upon progesterone stimulation relative to *wt*. In both b) and c), oocyte lysis was performed with addition of 0.1 M NEM. Vinculin used as loading control. d) Representative autoradiography (n = 2) of radiolabeled ATP incorporation by *wt* or 19DE CPEB2-Nt upon incubation with oocyte lysates at the indicated maturation time-points. CBB-stained gels are shown as loading controls. Nt: N-terminal. CBB: Coomassie brilliant blue. [ $\gamma$ -<sup>32</sup>P] ATP: Phosphorous-32-labelled ATP.

ute to increasing ultrasensitivity (Ferrell & Ha, 2014a, 2014b, 2014c; Guillén-Boixet et al., 2016).

An additional piece of data we extracted from **Figure 20d** was that the phosphorylation kinetics of *wt* and mutant CPEB2 N-termini paralleled the kinetics of Cdk1:CycB activity, peaking at MI and MII and decreasing at interkinesis (**Figure 10**, page 27).

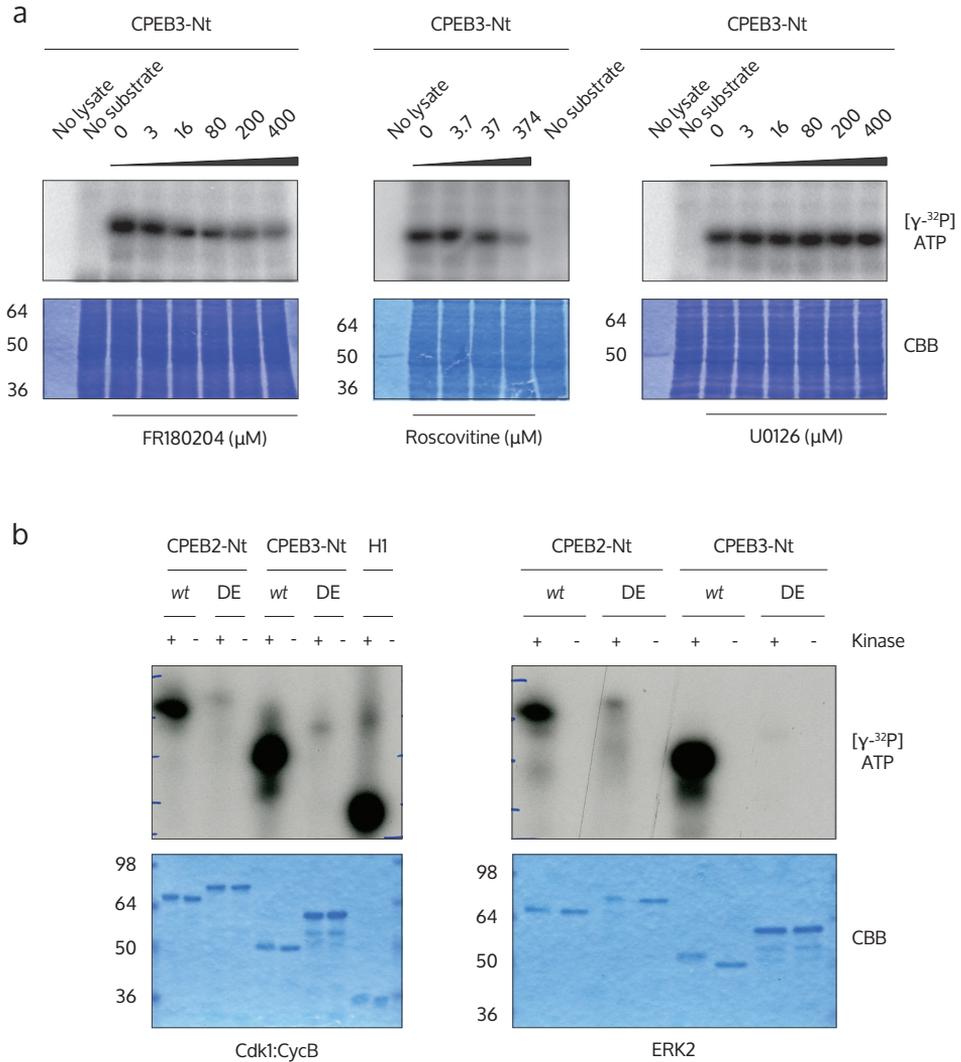
To summarize, CPEB2 proline-directed sites clustered on the N-terminal domain are phosphorylated upon meiotic progression in parallel to Cdk1:CycB activity. However, there might be additional PTMs - possibly phosphorylations - on CPEB2.

## ERK2 and Cdk1:CycB phosphorylate CPEB3

We then asked about the kinases responsible for CPEB3 phosphorylation. We have found that CPEB3 was phosphorylated at proline-directed sites at MI-entry. Moreover, three proline-directed kinases reportedly account for bulk phosphorylation during oocyte maturation and the first mitotic cycle (Peuchen et al. 2017): the cyclin-dependent kinase-cyclin pair Cdk1:CycB, MEK1 and p42MAPK. For clarification, MEK1 and p42MAPK are part of the Mos-MAPK kinase axis presented in the introduction, which more specifically is Mos-MEK1-p42MAPK-p90Rsk. Note that p42MAPK is nowadays mostly referred to as ERK2, which is the nomenclature we will use. We therefore performed a targeted kinase screening using the following inhibitors: an ERK2 and the closely related ERK1 (also known as p44MAPK) inhibitor, FR180204; a MEK1 and closely related MEK2 inhibitor, U0126 and a Cdk inhibitor, Roscovitine.

To test the inhibitors, we incubated Met-II lysates, recombinant CPEB3 N-termini and kinase inhibitors, at different inhibitor concentrations (**Figure 21a**). We observed a dose-dependent inhibition of phosphorylation in lysates treated with FR180204 or Roscovitine. However, we observed no decrease in radioactivity incorporation when treating with U0126. These results indicate that CPEB3 is phosphorylated by ERK2 and Cdk1:CycB in meiotic maturation.

To validate this, we did *in vitro* phosphorylation assays of recombinant CPEB3 N-termini with recombinant Cdk1:CycB or ERK2. *Wt* CPEB3 was strongly phosphorylated by both Cdk1:CycB and ERK2, whereas the proline-directed sites mutant was not (**Figure 21b**), thus strengthening the findings that we mapped all phosphorylation sites on CPEB3 and that CPEB3 was a substrate to Cdk1:CycB and ERK2 kinases.



**Figure 21. Cdk1:CycB and ERK2 phosphorylate CPEB3.** a) Representative autoradiographies ( $n = 2$ ) of recombinant CPEB3-Nt *in vitro* phosphorylation with Met-II oocyte lysates at the indicated kinase inhibitor concentrations. b) Representative autoradiographies ( $n = 2$ ) of recombinant CPEB3-Nt and CPEB2-Nt *in vitro* phosphorylation with recombinant Cdk1:CycB and ERK2. Histone H1 (H1) is a positive control in the Cdk1:CycB condition. MW, in kDa, indicated on the left. CBB-stained gels are shown as loading controls. Nt: N-terminal. CBB: Coomassie brilliant blue. [ $\gamma$ - $^{32}$ P] ATP: Phosphorous-32-labelled ATP.

Although CPEB2 was not included in the kinase screening, we tested whether it was a substrate to Cdk1:CycB and ERK2 *in vitro*. *Wt* CPEB2 was significantly more phosphorylated than the non-phosphorylatable mutant by both kinases (Figure 21b). However,

the mutant did incorporate radioactive signal upon incubation with ERK2, again indicating that there might be additional ERK2 phosphorylation sites.

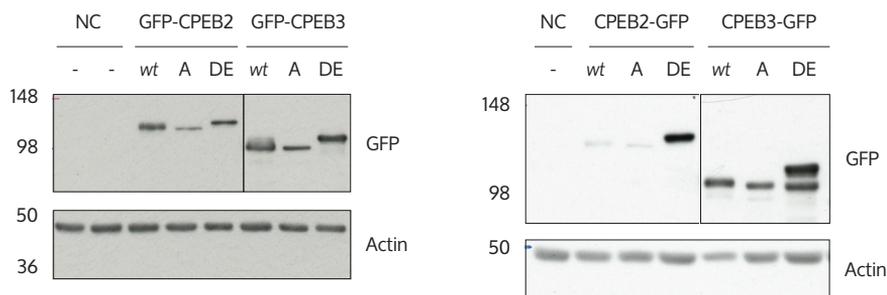
To recapitulate, CPEB3 was phosphorylated by the meiotic proline-directed kinases ERK2 and Cdk1:CycB, abiding to the same regulatory mechanism demonstrated for CPEB4 (Guillén-Boixet et al. 2016). CPEB2 was also potentially regulated by Cdk1:CycB given its phosphorylation dynamics, as highlighted earlier on **Figure 20d**. In addition, we obtained further evidence supporting that CPEB2 was phosphorylated beyond consensus proline-directed sites. In the future, CPEB2 kinase screening assays and phospho-site mapping could help clarify this point.

## CPEB2 and CPEB3 multiple phosphorylations impact cellular distribution

There are now numerous examples of RBPs regulated by PTMs, whereby PTMs promote or inhibit the ability of the RBP to form condensates. PTMs that affect key condensate proteins are often found in their IDRs and modulate the interactions with other condensate components (reviewed in Bah & Forman-Kay 2016, Söding et al. 2020). In the case of CPEB4, extensive N-terminal phosphorylation was demonstrated to impact its biomaterial properties and not to affect its protein interactors. When expressed in U-2 OS cells, CPEB4 had a punctuate distribution while a non-phosphorylatable mimetic mutant was diffuse. Furthermore, endogenous CPEB4 eluted at denser sucrose fractions at interkinesis than at M-phase, coinciding with its phosphorylation by proline-directed cell cycle kinases (Guillén-Boixet et al. 2016). Given the multiple phosphorylations on CPEB2 and CPEB3 N-terminal region, an IDR, we asked whether these had an effect on the ability of CPEB2 and CPEB3 to form condensates.

To approach the question, we quantified CPEB2 and CPEB3 cellular distribution in the human osteosarcoma cell line U-2 OS. We opted for U-2 OS because of their ease to use in imaging studies. Briefly, we transiently transfected U-2 OS with GFP-fused CPEBs and performed imaging after the cells were fixed. We blindly classified the images according to the fluorescence distribution as either aggregated or diffuse. **Figure 22** shows representative immunoblots of the expression of the different CPEB2 and CPEB3 fusions. Although we saw different expression of the different constructs, these effects were at the population level and did not interfere with our subsequent analysis, since we selected single cells of comparable total fluorescence intensity. These immunoblots also provide a control of cleavage, which is commonly observed with fusion

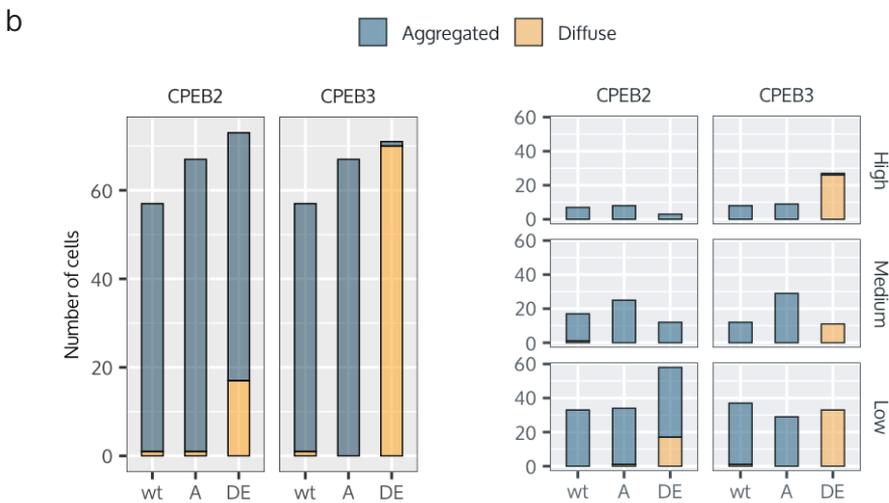
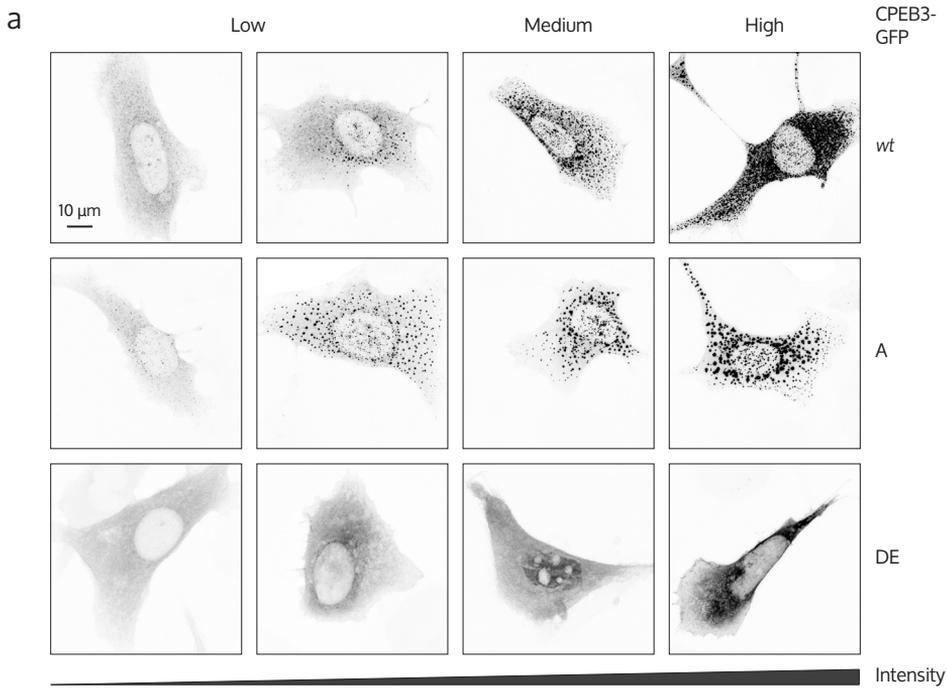
proteins. We observed cleavage in the case of the CPEB3-GFP non-phosphorylatable mimetic but considered that the relative contribution of the cleaved protein was well below that of the full length and, thus, that its influence would be negligible. **Figure 23b** shows representative maximum intensity projections of cells overexpressing CPEB3-GFP – *wt* and non-phosphorylatable mutants - arranged by increasing mean fluorescence intensity to illustrate the methodology.



**Figure 22. CPEB2 and CPEB3 phosphorylation-mutants transfection in U-2 OS cells.** Representative immunoblots ( $n = 2$ ) of CPEB2 and CPEB3 transfections in U-2 OS cells, of both GFP fusions. NC: mock-transfected negative control. A: phosphorylation-null. DE: phosphorylation-mimetic. Actin used as loading control. MW, in kDa, indicated on the left.

The results of the binary aggregated/diffuse classification are presented in **Figure 23b** left panel. We found that the constitutively non-phosphorylated mutants and the *wt* displayed a similar behaviour and they both formed cytoplasmic aggregates in the vast majority of cells. This was indeed expected since the cells were imaged at interphase when ERK2 and Cdk1 are inactive. In contrast, the phosphorylation mimetic mutants of CPEB2 and CPEB3 adopted a more diffuse distribution in cells. In the case of CPEB3, we observed near-total “solubilization” of the phosphorylation mimetic mutant relative to the *wt* or phosphorylation null. Instead, in the case of CPEB2, the “solubilization” of the phosphorylation mimetic mutant was partial, only 23% (17/73) of the cells counted were diffuse.

When we stratified the counted cells in three bins of mean total fluorescence – high, medium and low - (right panel of **Figure 23b**) we found that all the diffuse CPEB2-expressing cells fell under the low intensity bin. Thus, considering that fluorescence intensity can be a proxy for protein concentration, this observation is consistent with a concentration-dependent condensate formation mechanism, whereby the critical concentration of the CPEB2-20DE mutant would fall within the concentration range captured in the experiment. On the contrary, the critical concentration of CPEB3 18DE would be above the concentrations captured.

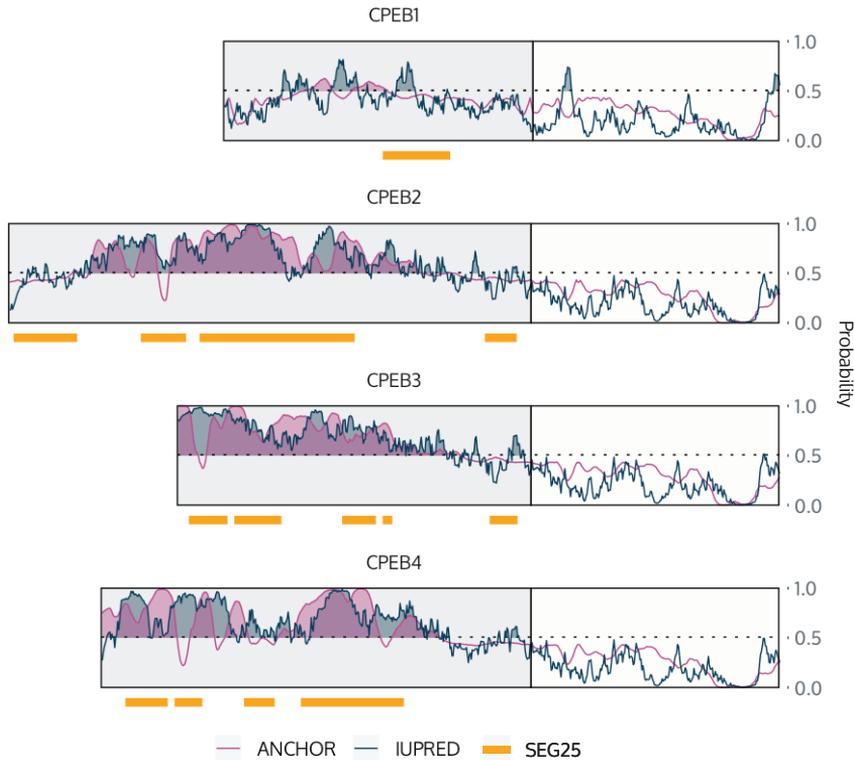


**Figure 23. CPEB2 and CPEB3 phosphorylation-mimetic mutants adopt a more diffuse distribution than their wt counterparts in U-2 OS cells.** a) Examples of the maximum intensity projections of CPEB3-GFP-transfected U-2 OS cells from which the aggregated/diffuse pattern was counted. Mean fluorescence intensity category indicated on top. b) Total cells counted with an aggregated or diffuse distribution pattern in different transfections ( $n = 4$ ). The right panel represents the same counts stratified by mean fluorescence intensity. *Wt*: wild type. *A*: phosphorylation-null. *DE*: phosphorylation-mimetic.

From these results we concluded that CPEB2 and CPEB3 N-terminal hyperphosphorylation impacts their ability to form biomolecular condensates. Our results suggest that phosphorylation increases the critical concentration at which these proteins undergo phase separation. Consequently, upon phosphorylation, the condensates are dissolved.

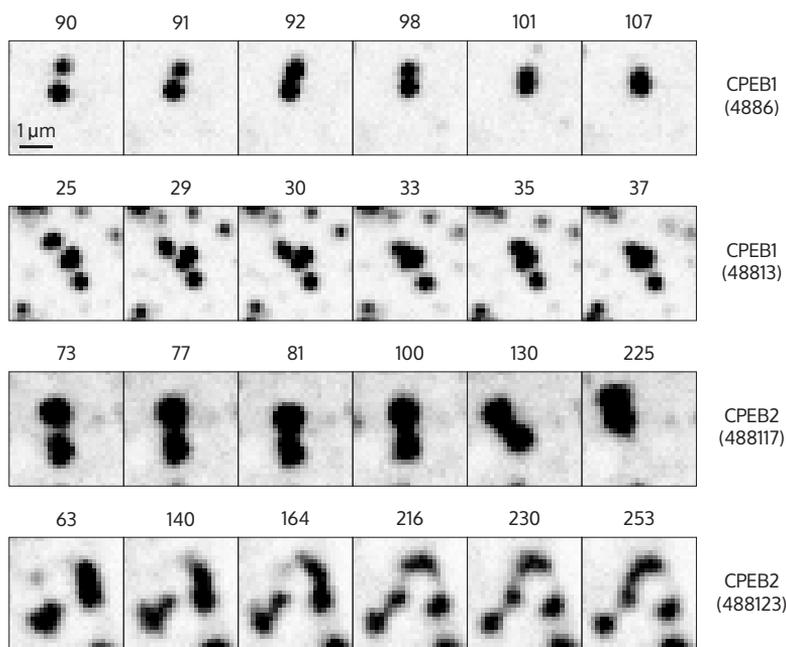
## CPEB2 and CPEB3 are more dynamic than CPEB1

Having found that CPEB2 and CPEB3 formed condensates in U-2 OS cells, we sought to further characterize these assemblies, with a special interest in understanding whether different CPEBs assembled in different or similar structures.



**Figure 24. CPEB2-4 are predicted to be disordered and contain LCRs within their N-terminal.** IUPred2A and ANCHOR2 score profiles of CPEB1-4. The stretches of predicted disorder (IUPred2A) and disordered-to-order upon binding (ANCHOR2) are shaded. Yellow stretches represent SEG predictions of LCRs (with parameters: window 25, lowcut 3.3, highcut 3.0). Proteins' N- and C-termini are delimited by boxes.

As mentioned in the introduction, CPEBs have a canonical RBP architecture, consisting of a divergent unstructured N-terminus and a highly-conserved C-terminus that contains the known RNA-binding domains. **Figure 24** shows several sequence features of the CPEBs. Specifically, ANCHOR scores, which measure the probability of disordered segments to undergo folding-upon-binding (Dosztanyi et al. 2009), IUPred scores, which measure context-dependent disorder probability (Mészáros et al. 2018) and, last, SEG, which is a predictor of compositionally-biased regions, i.e. LCRs (Wootton & Federhen 1993). The CPEB2-4 subfamily was predicted to be more disordered and contain more LCRs than CPEB1.



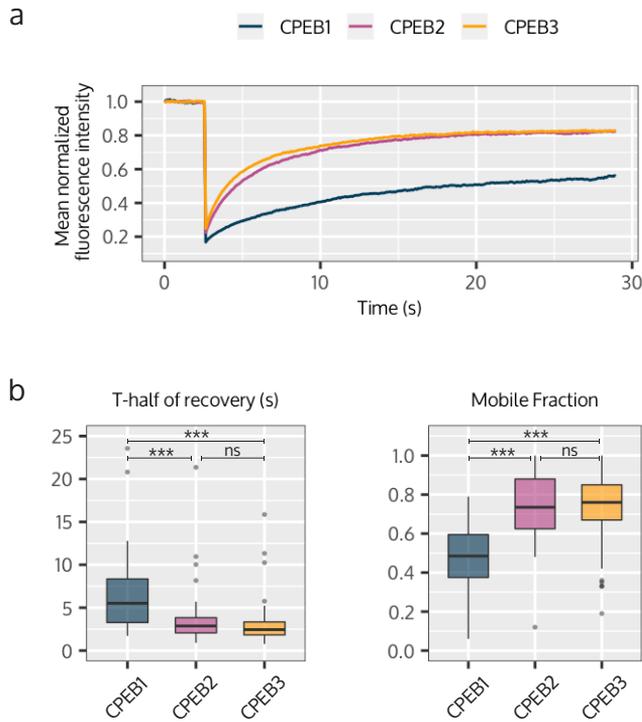
**Figure 25. CPEBs exhibit a dynamic behavior in vivo.** Selected frames from live imaging experiments displaying fusion events. The video frames indicated above; videos taken at 11.36 frames per second. The condition and video identifier are indicated on the right.

We investigated the dynamicity of the condensates formed by CPEB1, CPEB2 and CPEB3 by live imaging and fluorescence recovery after photobleaching (FRAP) in transiently-transfected U-2 OS cells. Although in previous experiments (**Figure 22**) we did not observe a different behaviour of the N- and C-terminal fluorescent protein fusions, we did live imaging and FRAP of the C-terminal fusions only, so as to avoid interfering with the N-terminal IDRs (Alberti et al. 2018).

Live imaging, at 24-hours post-transfection, showed CPEB1, CPEB2 and CPEB3 assemblies that were oftentimes round and mobile. A few fusion and fission events were reported in the case of CPEB1 and CPEB2, shown in **Figure 25**. These fusion and fission events happened in variable timescales, up to tens of seconds (e.g. the frames shown on **Figure 25** panel 3 represent 13.4 seconds). Sometimes, there was a delay between the first contact and fusion into a larger sphere. For reference, the timescales of fusion and the delay-upon-contact of *X. laevis* nucleoli experimentally determined by Brangwynne et al. were in the order of minutes whereas the timescales of fusion of *C. elegans* P granules were in the order of seconds (Brangwynne et al. 2009, 2011). The dynamics we observed are thus slow if we consider that for liquids, fusion and fission timescales increase linearly with condensate size, and that our condensates are smaller than nucleoli or P granules. The slow dynamics together with the anecdotal number of fusions and fissions could suggest molecular aging of the condensates into solid-like states, which is not uncommon, especially in overexpression systems (Alberti 2017).

Therefore, our results suggest CPEB1, CPEB2 and CPEB3 form condensates by LLPS but that at 24-hours post-transfection these structures have experienced ageing. Live imaging at earlier time-points, e.g. 12-hours, or live imaging endogenously-tagged protein, e.g. eGFP knock-in cell lines, would help clarify whether this artifact is occurring.

FRAP is a live imaging technique that consists on partially or totally bleaching a fluorescently-labelled hub and monitoring fluorescence recovery. Modelling the recovery provides information of the underlying molecular dynamics including, but not limited to, exchange between the dilute and dense phases of condensates (Alberti et al. 2019, McSwiggen et al. 2019). The mean curves of CPEB1-3 full-FRAP time-courses are presented on **Figure 26a** for illustration. The bleaching event took place at frame 30 (corresponding to 2.64 seconds) and the recovery was monitored for additional 300 frames (26.4 seconds). The full-FRAP recoveries were fitted to a single exponential model and two descriptors were obtained: the half time of recovery ( $t_{\text{half}}$ , **Figure 26b**, left panel), the time at which 50% of the recovery was achieved, and the mobile fraction (**Figure 26b**, right panel), the maximum final fluorescence recovered relative to the initial value. As shown on **Figure 26** and **Table 2**, CPEB1 had a significantly greater half time of recovery and smaller mobile fraction than CPEB2 and CPEB3, meaning it recovered slower and to a lesser extent from the bleaching event. CPEB2 and CPEB3 displayed comparable dynamics. The recoveries here determined for CPEB1-3 are not comparable to the CPEB4 dynamics previously determined by Guillén-Boixet et al. (Guillén-Boixet et al. 2016) because, although we used the same experimental setup, the CPEB4 GFP fusion was N- instead of C-terminal.



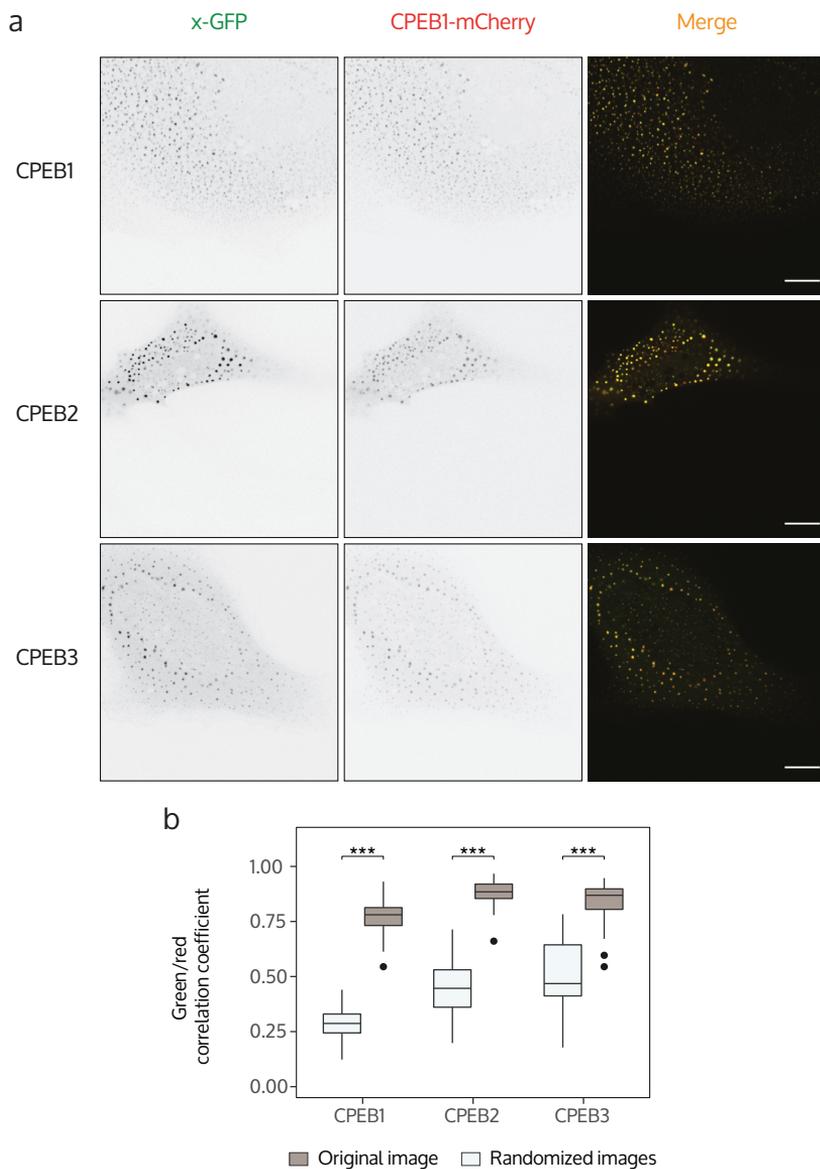
**Figure 26. CPEB2 and CPEB3 are more dynamic than CPEB1 as determined by FRAP.** a) Average mean normalized fluorescence intensity curves of all cells included in each condition in the FRAP experiments. b) Half-time of recovery (in seconds) and mobile fraction distributions determined by FRAP. Number of photobleaching events counted per condition: CPEB1, 48; CPEB2, 78; CPEB3, 84. These correspond to three biological replicates. Comparison between groups is done by Kruskal-Wallis rank sum test (significance level of 5%) and post-hoc Dunn's test with Bonferroni correction (\*\*\*: adjusted p-value  $\leq 0.001$ ; ns: not significant).

**Table 2. Summary of CPEB1-3 FRAP model variables.** St. dev.: standard deviation.

Condition	t-half (s)		Mobile fraction		Events
	Mean	St. dev.	Mean	St. dev.	
CPEB1	6.35	4.41	0.48	0.17	48
CPEB2	3.38	2.69	0.75	0.18	78
CPEB3	2.90	2.17	0.75	0.16	84

All in all, live imaging experiments revealed that the CPEBs form condensates whose properties are compatible with a mechanism of formation by LLPS. Importantly, CPEB1 forms assemblies that are less dynamic than those formed by CPEB2 or CPEB3.

## The CPEB1-3 co-localize in the same compartments



**Figure 27. The CPEBs co-localize in the same compartments.** a) Representative images of CPEB1-mCherry co-localization with CPEB1-3-GFP. b) Distribution of the Pearson correlation coefficients of the red and green channels in the acquired images or the randomized images. Data from two independent experiments, CPEB1-GFP  $n = 44$ , CPEB2-GFP  $n = 38$  and CPEB3-GFP  $n = 37$ . Paired samples were compared with Wilcoxon signed-rank test with a Bonferroni correction to account for multiple testing (\*\*\*: adjusted  $p$ -value  $\leq 0.001$ ). Scale bar = 10  $\mu\text{m}$ .

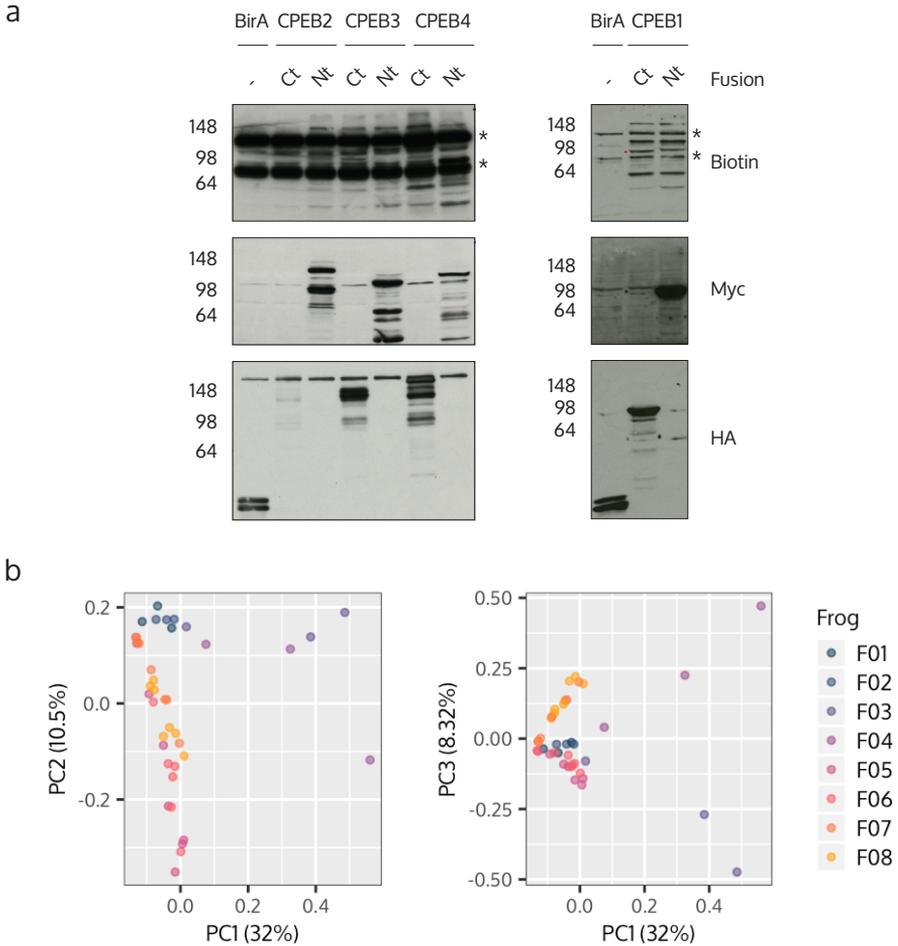
We also asked, whether when co-expressed in U-2 OS, CPEB1, CPEB2 and CPEB3 would co-localize in the same condensates. In order to address it, in collaboration with Manuel Cañete (Cañete 2020), we co-expressed all pairwise combinations of CPEB1-3-GFP with CPEB1-mCherry and imaged fixed cells with an aggregated pattern 24-hours post-transfection.

We found that CPEB1, CPEB2 and CPEB3 had a near-perfect co-localization in condensates (**Figure 27**). Thus, despite their different biomaterial properties (**Figure 26** and **Table 2**), the CPEBs resided in the same compartments. Whether they are in different subcompartments within these structures and what happens to these structures when the CPEBs are modified remain open questions.

## The CPEBs share a core proximate of mRNA processing, storage and repression proteins

In order to address the functional nature of the condensates and the co-existence of the different CPEBs in the same compartment we decided to approach the subject from a proteomic perspective. We chose a proximity-dependent biotinylation (PDB) coupled to MS method, specifically, BioID (Roux et al. 2012). In BioID, PDB relies on a modified version of *Escherichia coli*'s biotin ligase – BirA - that is fused in frame to the protein of interest. The BirA-fusion protein catalyzes conversion of biotin to activated biotin. In turn, activated biotin, which is reactive to primary amines, covalently labels proximal proteins in the native cellular environment. Last, biotinylated proteins are affinity-purified with streptavidin beads and identified by MS.

The BioID methodology was originally developed for a human cell line (Roux et al. 2012). Since, it has been adapted to a variety of cell lines, tissues and organisms (Sears et al. 2019). To address our biological questions, we employed the BioID methodology adapted to *X. laevis* oocytes by Manuel Cañete (Cañete 2020). In contrast to cultured cells, our scenario required 40-hour-incubation at 18°C in 20  $\mu$ M biotin-supplemented media for efficient labeling. The lengthened incubation time, relative to cultured cells protocols, was possibly required due to the low incubation temperature. Kim and colleagues (Kim et al. 2016) observed that, 16 and 25°C-incubation resulted in a decrease of 90 and 85% BirA activity, respectively, relative to 37°C. Additional details on the BioID setup for *X. laevis* oocytes can be found on “SVI *X. laevis* BioID” on page 109.



**Figure 28. CPEBs BioID expression and biotinylation.** a) Representative loading (myc and HA tags) and biotinylation immunoblots ( $n = 4$ ) of the CPEBs BioID. The top row indicates the position of the BirA moiety (Nt: N-terminal; Ct: C-terminal). The asterisk marks constitutively biotinylated proteins. MW, in kDa, indicated on the left side of the film scans. NC: negative control. b) Principal component analysis of the eight BioID replicates, colored by replicate (i.e. frog). Components one through three are shown, alongside their contribution to total variance (%). The variables have been scaled and centered.

We performed BioID experiments in quadruplicate where we tagged CPEB1-4 with BirA at both the N- and C-terminal ends. We employed the BirA enzyme alone as negative control of endogenously biotinylated proteins and unspecific bead binders and endogenously biotinylated proteins. **Figure 28a** shows representative immunoblots of loading and biotinylation of all constructs used, where expression and biotinylation by all constructs was observed. The bands marked with asterisks correspond to constitutively biotinylated carboxylases present in the oocyte (Tong 2013), that we identified

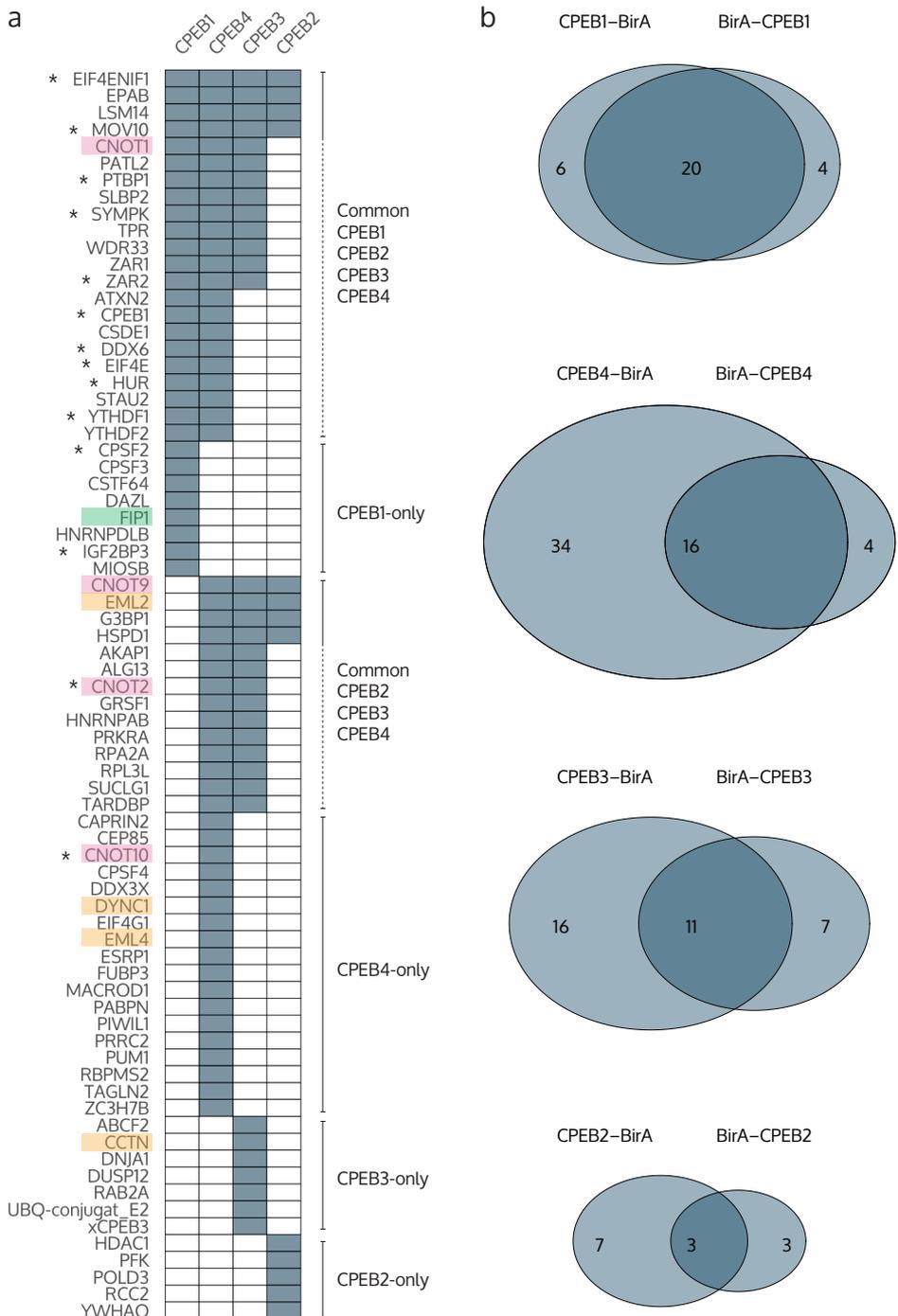
by MS (data not shown). The anti-myc and anti-HA immunoblots should be interpreted as a binary readout of expression; the cleavage observed is an artifact of the lysis conditions (data not shown).

In total, eight biological replicates constituted the experiment, with frogs 1 through 4 used in the CPEB1 BioID and frogs 5 through 8 for CPEB2-4. The frogs used were neither inbred nor littermates, consequently, we dealt with large variability in our data, which is a double-edged sword of low statistical power but high biological relevance. In **Figure 28b**, a principal component analysis (PCA) of the data is shown, with the three main components and their contribution to total variance. The samples are colored by frog to illustrate that the biological replicates and the experimental run (i.e. frogs 1-4 versus Frogs 5-8) were a greater source of variability than the conditions (i.e. BirA, CPEB1, CPEB2, CPEB3 or CPEB4).

	Condition				Control					
	F01	F02	F03	F04	F01	F02	F03	F04		
Protein A	6	8	6	7	4	3	3	4	Imputed	Complete cases
Protein B	5	6	NA	NA	2	1	2	2		
Protein C	3	NA	2	3	2	NA	NA	4	Discarded	Incomplete cases
Protein D	1	NA	NA	NA	1	2	1	2		
Protein E	7	7	7	8	NA	NA	1	NA	Recovered	
Protein F	NA	6	5	5	NA	NA	NA	NA		

**Figure 29. Graphical summary of the in-house hits-selection methodology.** Paired condition-control iBAQ values of six imaginary proteins across four replicates, indicating all possible outcomes depending on the number and distribution of missing values (NA: not available).

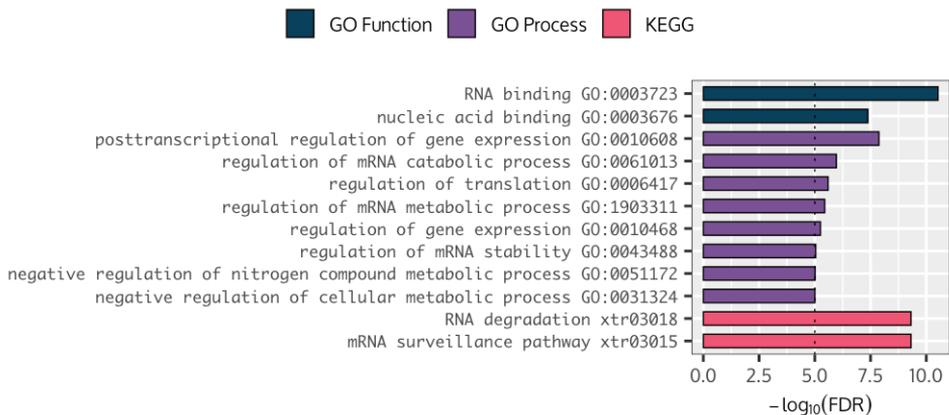
In our label-free bottom-up-based quantitation, we estimated protein abundance from the iBAQ values, which is the sum of peptide intensities divided by the number of theoretical observable tryptic peptides (Krey et al. 2014). Missing observation can constitute more than 50% of bottom-up MS datasets (O’Brien et al. 2018) and pose a challenge for differential enrichment analyses. In our case, we took a conservative approach to deal with missing data. Briefly, we performed single value imputation at the protein level in the cases where as few as one or two observations were missing. For those bait-prey pairs, we computed the fold-change and p-value of differential enrichment using linear models. The remainder non-complete cases were discarded except for extreme cases where missing observations clustered in the negative control and not in the condition. Those cases could in fact be perfect hits and, as such, undetectable in the control condition. Thus, we included them although neither the magnitude nor



**Figure 30. SVI CPEBs proximomes.** a) Summary of SVI BioID results showing the shared and own hits of each CPEB, considering the union of both BirA fusions for each bait (n = 4). b) Venn diagrams showing the hits overlap between reciprocal fusions.



To obtain a functional view on the proximome, we used STRING (Szklarczyk et al. 2019), which is a database of known and predicted protein association. When entering the CPEBs 74-protein proximome plus the baits themselves we obtained a high-confidence network with 43 nodes - excluding disconnected nodes - and 130 edges (**Figure 31**). The connectivity was greater than expected in a random set of proteins of the same size, implying biological relatedness. The network was organized in two main clusters, one constituted by pre-mRNA processing machinery (left) and another by CCR4-NOT components and others (right), associated to general mRNA deadenylation and deadenylation-dependent decay. Additionally, there were three pairs of isolated proteins among which POLD3 and RPA2 revealed a link to the nuclear compartment and the EML2 and EML4 pair a link to the microtubule network. Importantly, with the interactions characterized in this BioID experiment we added 35 nodes to this tightly connected network.



**Figure 32. Enriched functionalities in the CPEBs proximomes.** GSEA with GO and KEGG terms done with STRING.

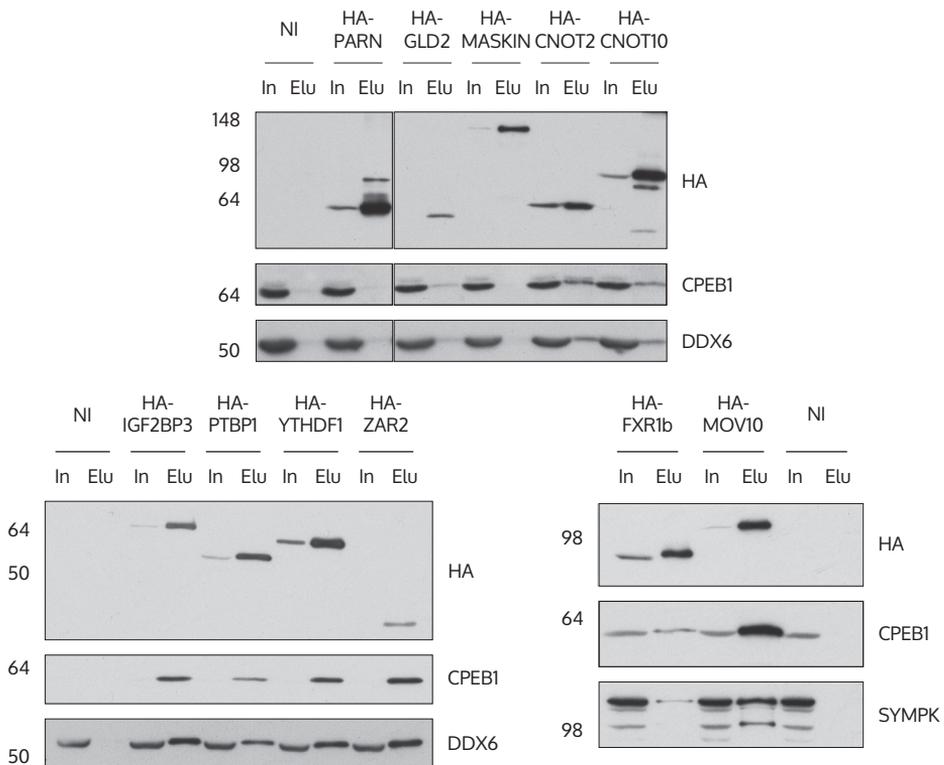
In addition, we asked which functional categories were over-represented in the CPEBs proximome. For this purpose, we employed gene sets from two commonly used resources of gene set annotation: Gene Ontology (GO, Carbon et al. 2019) and KEGG (Kanehisa 2019), and did a gene set enrichment analysis (GSEA). Setting a false discovery rate (FDR) threshold of 0.00001 a total of 12 terms were enriched (**Figure 29**). The top enriched gene sets were GO molecular function ‘RNA binding’, GO biological process ‘posttranscriptional regulation of gene expression’ and KEGG ‘RNA degradation’ and ‘mRNA surveillance pathway’. All these functions are consistent with the known roles of the CPEBs, and by extension CPEB-interacting proteins, and thus provided confirmation that our proximome dataset is biologically relevant.

We also interrogated the dataset about specific components related to known roles of the CPEBs in translation regulation. In the proximome there were neither proteins with poly(A) polymerase nor deadenylase activity. There were, however, proteins associated to these activities (**Figure 30a**, highlighted in green and pink). Specifically, FIP1 is an RBP and CPSF-associated factor that mediates the interaction with poly(A) polymerase (Kaufmann et al. 2004). While FIP1 has an established role in nuclear pre-mRNA processing and APA (Lackford et al. 2014), its role in cytoplasmic polyadenylation has not been demonstrated. Regarding deadenylases, we found subunits of the CCR4-NOT complex - CNOT1, CNOT2, CNOT9 and CNOT10 - albeit none of the catalytic subunits. We did not detect Tob either, which in somatic cells bridges CPEB1, CPEB3 and CPEB4 to the deadenylase Caf1 (Hosoda et al. 2011, Ogami et al. 2014) We also identified several cytoskeletal and motor proteins (**Figure 30a**, highlighted in orange), concretely, the microtubule-associated proteins EML2 and EML4 in the CPEB2-4 proximomes, as well as Dynein C1 in the CPEB4 proximome and the actin-binding protein Cortactin in the CPEB3 proximome. Among these, so far, only dynein has been implicated with mRNP function. Specifically, dynein directs vegetal mRNP transport and can be visualized decorating L-bodies (Gagnon et al. 2013, Neil et al. 2020). The association of CPEB2-4 to microtubules and microfilaments opens an exciting avenue that remains to be further explored.

Interestingly, the proximomes also revealed differences between CPEBs. Most notably, CPSF components, CSTF and FIP1, all part of the polyadenylation machinery (left cluster in **Figure 31**), appear as CPEB1-specific components. Contrastingly, SYMPK, the polyadenylation machinery scaffold, appears in the proximome of both CPEB1 and CPEB2-4 families. These could suggest a differential function of CPEB1 and CPEB2-4 in cytoplasmic polyadenylation and translational activation that remains to be explored in detail. Another CPEB1-specific component is the germline-specific-RBP DAZL. DAZL is also a dual function RBP that participates in translational repression at SVI and can participate both repression or activation in meiotic maturation, depending on the specific target 3'UTR context (C. R. Yang et al. 2020). DAZL and CPEB1 have been described to interact and co-regulate mRNA translation but mechanistic insight about how they achieve this is lacking (Sousa Martins et al. 2016, C. R. Yang et al. 2020). According to our data, DAZL interacts with CPEB1 but not with CPEB2-4.

We validated the BioID by testing some of the candidates with an independent complementary technique, co-immunoprecipitation (co-IP). We performed co-IP of 14 proteins found in the BioID dataset (highlighted with asterisks in **Figure 30a**) and three additional proteins that have historically being linked to CPEB1 function in *X. laevis*

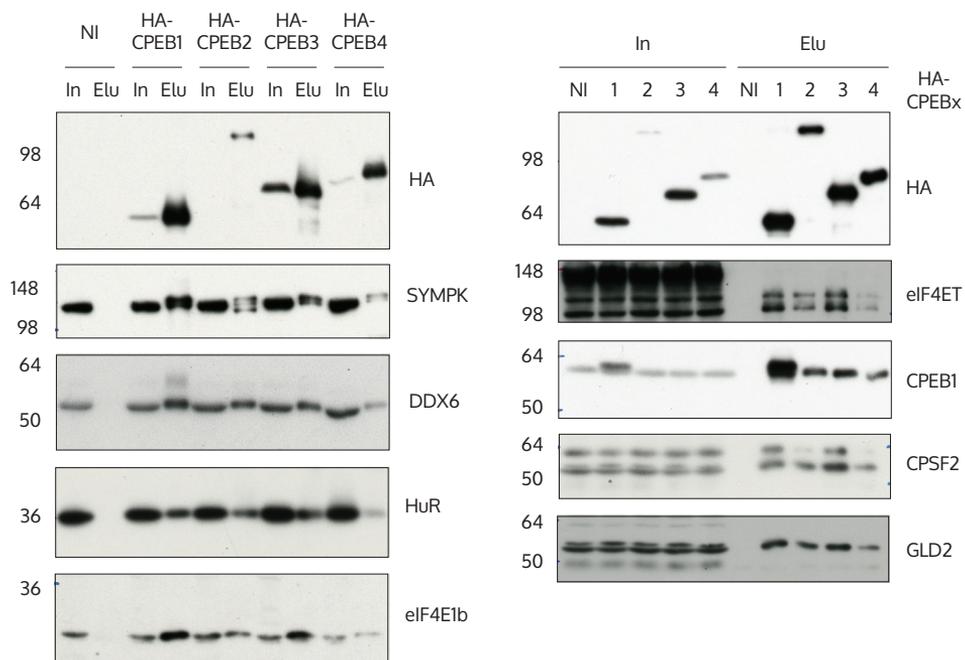
oocytes, namely, PARN, Maskin and Gld2 (**Figure 33**, **Figure 34**). By co-IP we detected the poly(A) polymerase Gld2 with all CPEBs. In contrast, we did not find PARN nor Maskin. The validation confirmed all the proximome candidates tested, thus providing robustness to the BioID methodology. Among the candidates differentially enriched in the CPEB1 or CPEB2-4 proximomes, we only tested CPSF2, which we pulled-down with the four CPEB constructs, thus, disagreeing with the BioID observation. This discrepancy could be methodological, since pull-downs are not quantitative and lysis and incubation can alter interactions. Ideally, we would further validate these candidates with an independent technique, such as size-exclusion chromatography, immunofluorescence or an additional PDB-method, such as TurboID.



**Figure 33. BioID proximomes are validated by co-IP, part 1.** Representative co-IP immunoblots of HA-tagged candidates IPs and endogenous CPEB1 ( $n = 2$ ). DDX6 and Symplekin are shown as input loading controls. MW, in kDa, indicated on the left side of the film scans.

To sum up, we identified interactions between the CPEBs and 74 proteins associated to different biological processes and compartments, mostly related to pre-mRNA processing, repression and deadenylation. Among the CPEB1 complexes described in the

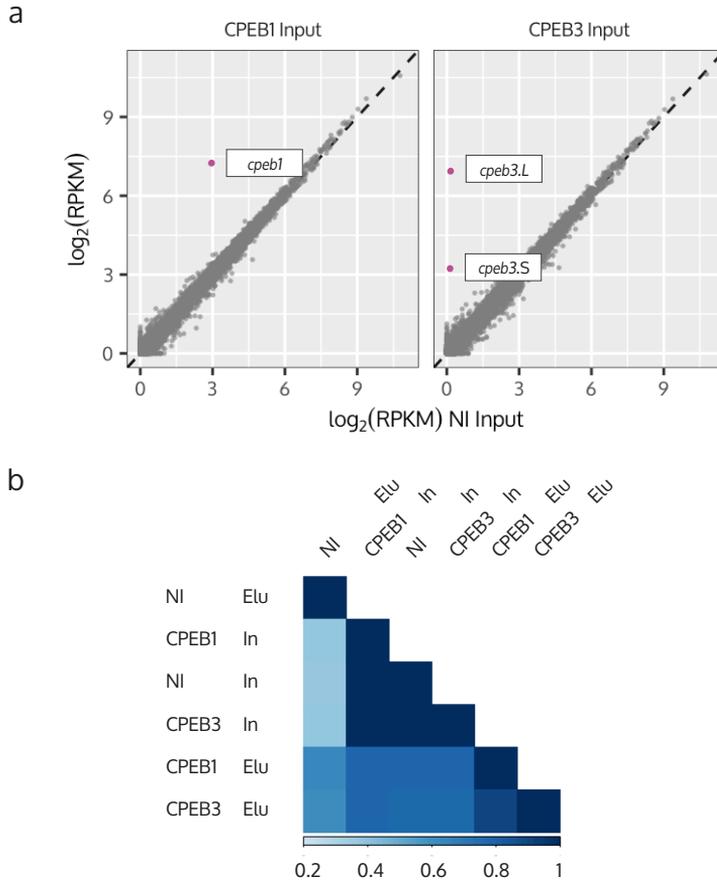
literature and presented in the introduction, our dataset is most consistent with the model from Minshall et al. (Minshall et al. 2007) (**Figure 11c**), in which CPEB1 mediates mRNA repression in P-body-like compartments together with DDX6, PATL, LSM14, eIF4E-T and 4E1b. We find these and additional P-body and SG components in the CPEB1 and also CPEB2-4 proximomes, indicating that the CPEBs reside in P-body-like compartments and that they all participate in mRNA repression. Simultaneously, the CPEB1 proximome is consistent with the dual repression and activation model proposed by Barnard et al. and Méndez et al. (repression represented **Figure 11a**) (Barnard et al. 2004; Mendez, Murthy et al. 2000). Both repression and cytoplasmic polyadenylation components are present in SVI CPEB1-mRNPs, although we did not find the catalytic subunits PARN nor Gld2 and we identified CCR4-NOT and FIP1 instead. Importantly, the cytoplasmic polyadenylation machinery was not so evidently present in the CPEB2-4 proximomes, pointing to potential differences in the activation mechanism between CPEBs.



**Figure 34. BioID proximomes are validated by co-IP, part 2.** Representative co-IP immunoblots of HA-CPEBx IPs and endogenous candidates (n = 3). MW, in kDa, indicated on the left side of the film scans.

# The CPEBs regulate an overlapping set of mRNAs

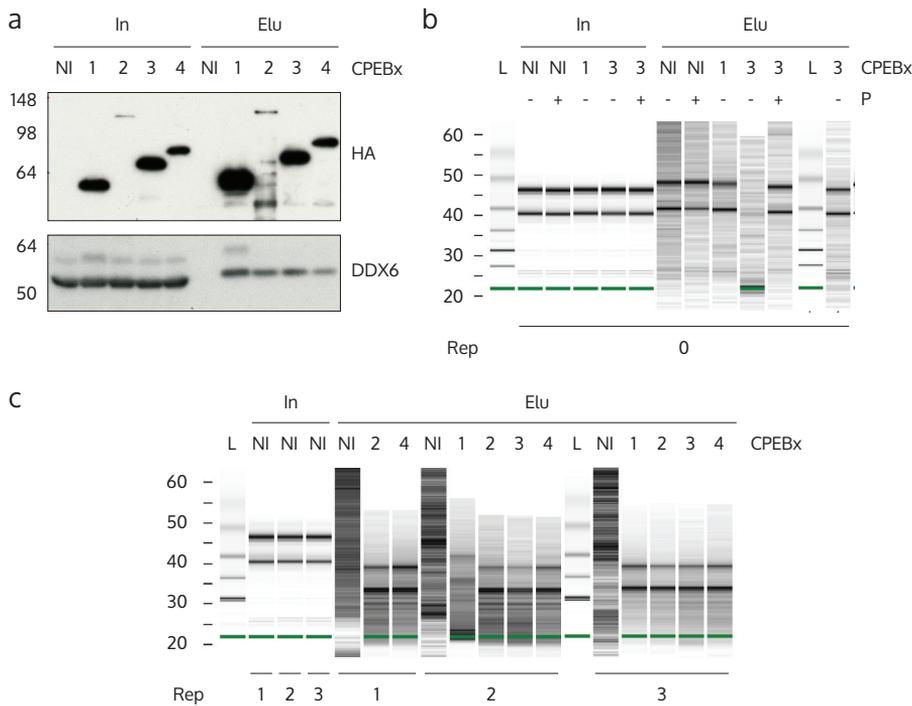
Motivated to understand the differences and similarities between CPEBs, another major aim of this thesis was to determine the targets of each CPEB. To do so, we performed RNA-immunoprecipitation Sequencing (RIP-Seq) of the four CPEBs in SVI oocytes.



**Figure 35. mRNA microinjection does not have global transcriptomic effects.** a) Correlation in  $\log_2(\text{RPKM})$  between the CPEB1 and CPEB3-microinjected inputs and the NI input from the RIP-seq pilot experiment. Enriched genes are labeled and highlighted in pink. b) Correlation matrix of the samples in the RIP-Seq pilot experiment. NI: not-injected. Elu: eluate. In: input.

We performed HA RIP of HA-tagged microinjected CPEBs, in order to have the most comparable setup. We chose to use the N-terminally GFP tagged form of CPEB2 to ob-

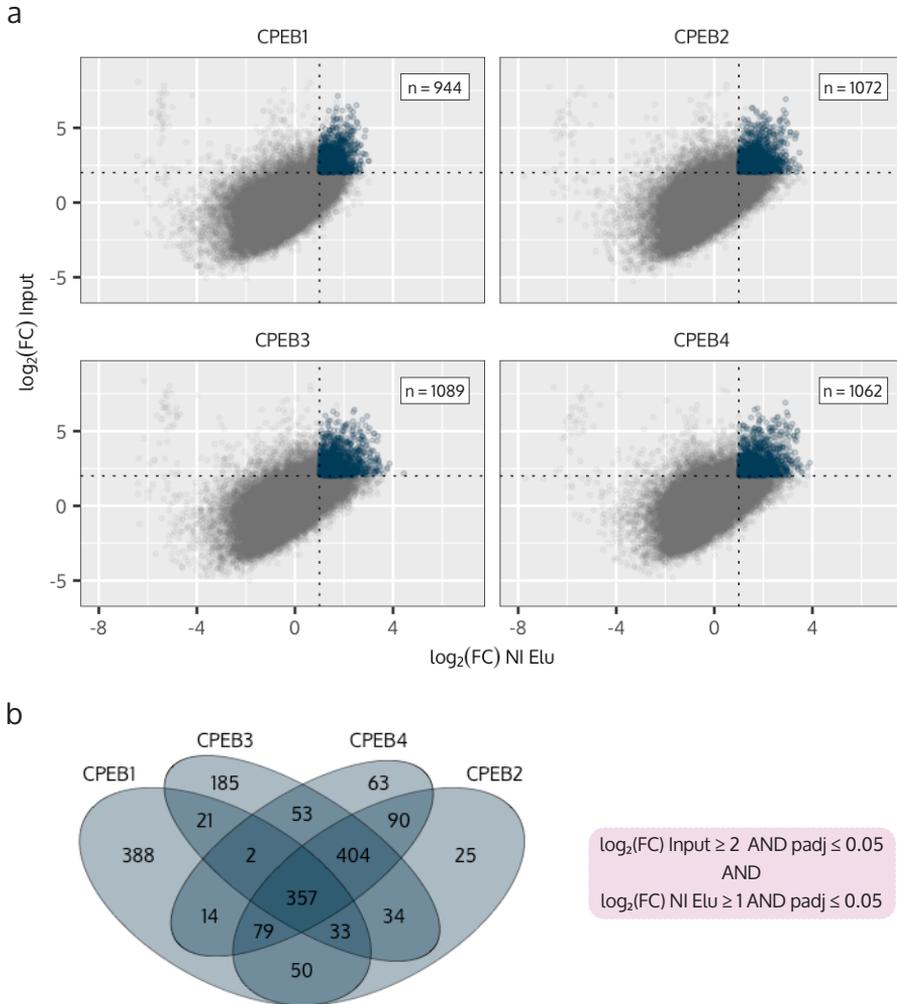
tain higher expression (**Figure 15**) more comparable to CPEB1, CPEB3 and CPEB4. In a pilot experiment we determined that the RNA-microinjected inputs were very similar to the not-injected (NI) inputs. They exhibited near-perfect correlation (**Figure 35b**) and only the genes corresponding to the microinjected mRNAs were upregulated (**Figure 35a**). Hence, as expected, microinjection and bait overexpression did not impact the global SVI transcriptome. Consequently, we included a single common NI input control per set of conditions within the same biological replicate, reducing the number of libraries to sequence. As a tradeoff, we could not calculate a meaningful enrichment of the overexpressed mRNAs, therefore we excluded the CPEBs mRNAs from the analysis and we did not address whether each CPEB regulated its own mRNA.



**Figure 36. RIP-Seq quality controls.** a) Representative immunoblot of HA-IPs performed for RIP-Seq (n = 4). The input loading is equivalent to 1 oocyte while the eluate loading is 4. b) Bioanalyzer electrophoresis of the immunoprecipitated RNA in the pilot experiment (replicate 0). The CPEB3 - P eluate was re-run on a separate gel (last 2 lanes). c) Bioanalyzer electrophoresis of the immunoprecipitated RNA in biological replicates 1 to 3. L: ladder. NI: not-injected. Elu: eluate. In: input. P: progesterone. Rep: replicate. MW, in kDa in the case of immunoblots and seconds in the case of RNA gels, indicated on the left side of panels.

Briefly, to determine the targets of each CPEB, we sequenced twenty libraries coming from four biological replicates, corresponding to three biological replicates of each

CPEB and two negative controls per replicate: the input and an HA eluate of NI oocytes, to control for unspecific purification (NI IP or blank IP). An example of the protein fraction recovered in a typical RIP-Seq experiment is presented on **Figure 36a**. The complementary RNA fraction is shown on panels b and c (featuring some of the samples from the pilot experiment, replicate O, that were not included in the final analysis).



**Figure 37. Overview and overlap of CPEBs hits subsets.** a) Principal component analysis of the twenty RIP-Seq samples. Components one through three are shown, alongside their contribution to total variance (%). The variables have been scaled and centered. b) Fold-change (in  $\log_2$  scale) relative to the input and not-injected elution of all genes detected in the RIP-Seq experiment. Pink dashed lines mark the fold-change thresholds set for the analysis [ $\log_2(\text{FC}) \text{ Input} \geq 2$ ,  $\log_2(\text{FC}) \text{ NI Elu} \geq 1$ ]. Within the upper right quadrants, genes in blue have an adjusted p-value  $\leq 0.05$  for both normalizers. The count of genes making the four cuts is indicated in each panel. c) Venn diagram displaying the overlap of hits between CPEBs.

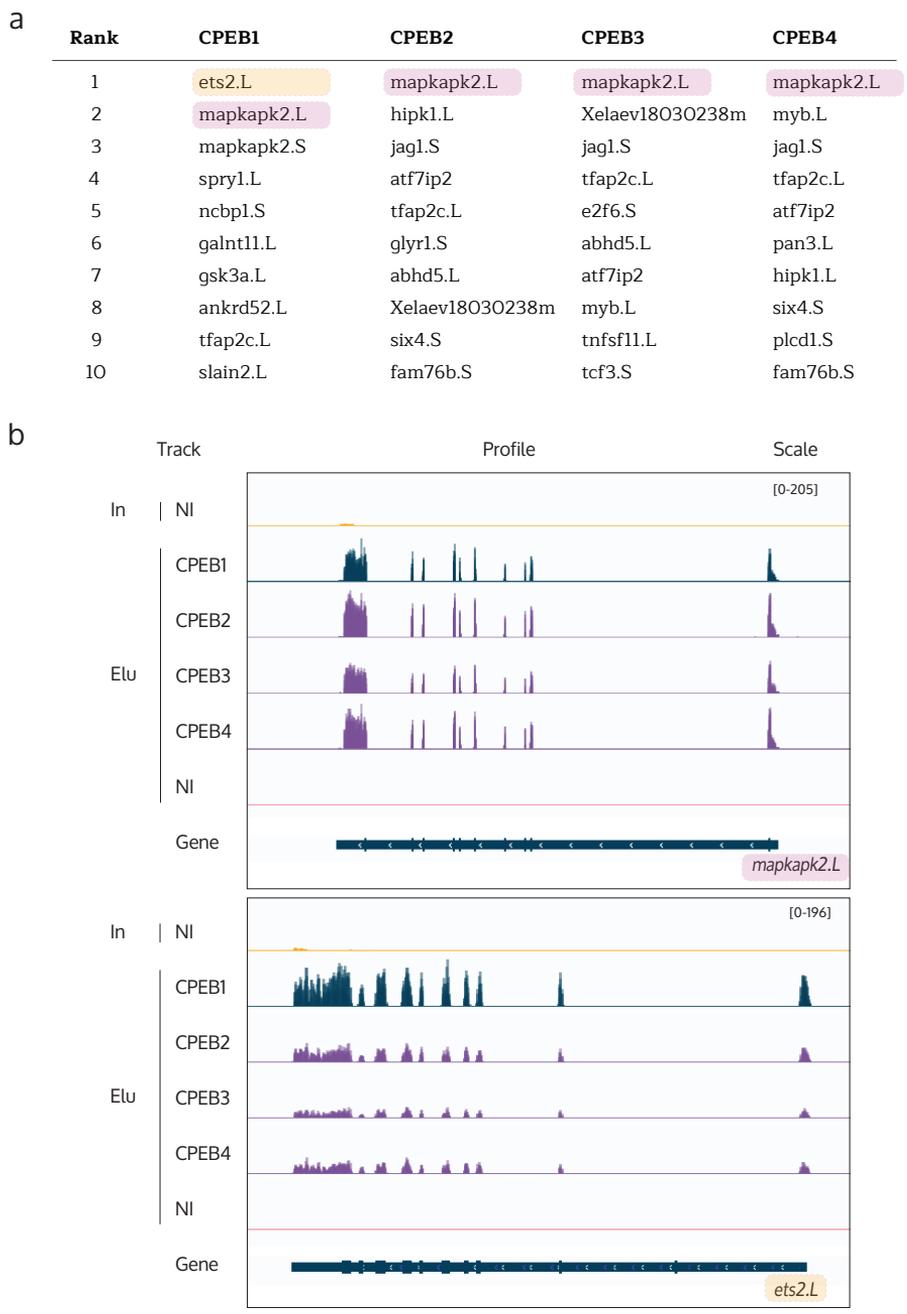
We defined a subset of high-confidence targets for each CPEB based on their differential enrichment to the input and to blank IP: we set a threshold to the input of 4-fold and to the blank IP of 2-fold, and we demanded the adjusted p-value of these contrasts be smaller than or equal to 0.05.

With these requirements, each CPEB displayed a subset of roughly one thousand hits (**Figure 37a**). The relationships between the four subsets are depicted in the Venn diagram in **Figure 37b**. With the given definition, the union of all CPEB targets consisted on 1798 mRNAs and 357 mRNAs formed the intersection. CPEB2-4 shared a large percentage of their hits, with an overlap of 761 targets. On the opposite end, CPEB1 had a big percentage of hits, 388 or 41%, that it did not share with any other CPEB.

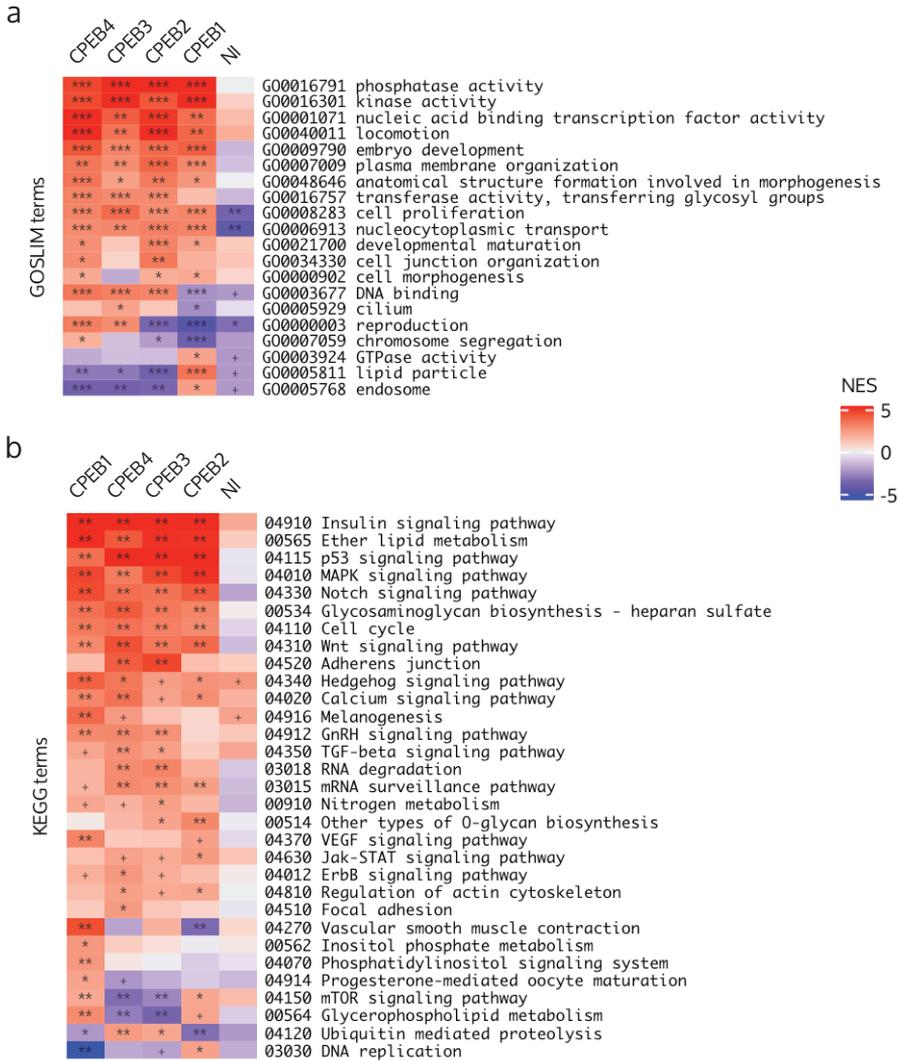
For illustration, a list of the top 10 ranked hits for each CPEB are presented in **Figure 38a**, along with two examples of the read mapping in panel b. Interestingly, although the CPEBs bind regulatory sequences in the mRNAs 3'UTRs, since we did not perform RNA fragmentation and RNA degradation is limited in oocytes, we obtained reads throughout the entire transcripts length.

To obtain a global functional picture on the RIP-Seq dataset, we performed GSEA of GO and KEGG gene sets with ROAST (D. Wu et al. 2010). Relevant GO slim (simplified GO containing a subset of GO terms) and KEGG terms across conditions are shown in **Figure 39**. Terms that were shared by all CPEBs included, but were not limited to, GO 'kinase and phosphatase activity', 'embryo development', 'nucleocytoplasmic transport', 'cell proliferation' and KEGG 'cell cycle' plus a number of signaling pathways, namely 'Wnt', 'MAPK', 'Notch', 'Insulin' and 'p53'. These were consistent with the known functions of the CPEBs and our biological context.

In addition, we wondered if there were any distinctive features of CPEBs' targets over the whole transcriptome, especially at the 3'UTR level. We specifically asked about 3'UTR length and arrangement of CPEs in the 3' UTR. We hypothesized that CPEB-targets would have longer 3' UTRs and would be enriched in arrangements of CPEs that predict regulation by CPEBs, according to the CPE-code described by Piqué et al. (Piqué et al. 2008). To address these questions, we used a *X. laevis* SVI oocyte 3'UTR dataset recently published by Yang et al. (F. Yang et al. 2020), where 3'UTRs were annotated using experimental 3' end determination data. Working with this dataset, as opposed to using the UCSC dataset, had the advantage that all annotated 3'UTRs were from our biological context. It also had the advantage of providing relative abundance of all 3'UTR isoforms per gene.



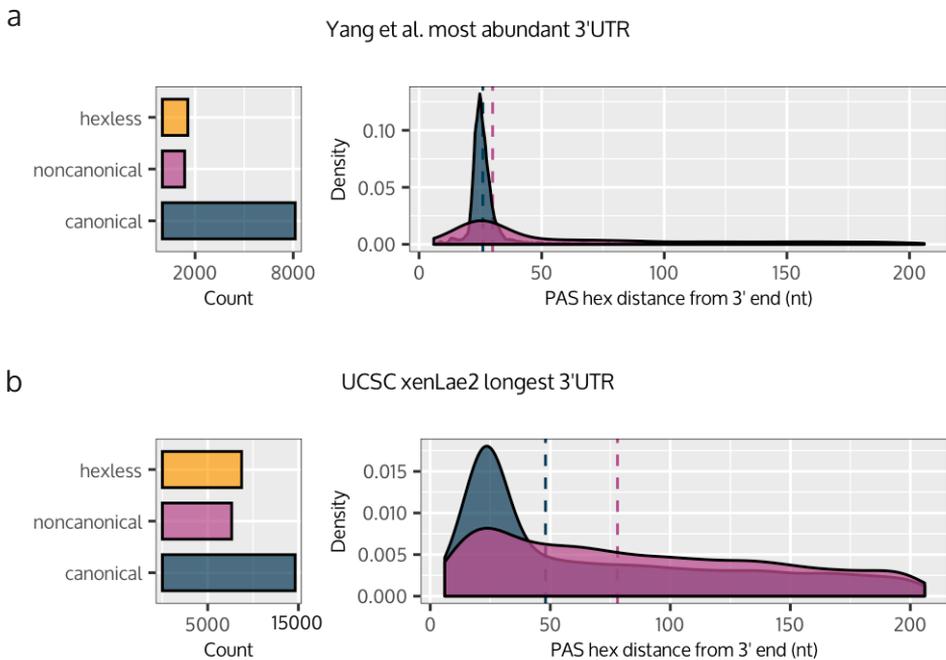
**Figure 38. Top-10 ranked hits.** a) Top 10 ranked hits of each CPEB from our RIP-Seq data. The rank is calculated using the module of the fold-changes (relative to the input and to the not-injected elution). b) Read mapping of two top candidates, *mapkapk2.L* and *ets2.L*, across conditions. Each track contains the overlay of all replicates. The read scale – specified on the top right - is the same for all samples.



**Figure 39. ROAST GSEA reveals shared and own functions of each CPEB.** a) GOSLIM terms enrichment across conditions relative to the input. b) KEGG terms enrichment across conditions relative to the input. Only terms with a positive significant (adjusted p-value  $\leq 0.05$ ) score for at least one CPEB are included. Moreover, terms with a positive significant score for the NI elution negative control are excluded. NES: normalized enrichment score. +: adjusted p-value  $\leq 0.1$ ; \*: adjusted p-value  $\leq 0.05$ ; \*\*: adjusted p-value  $\leq 0.01$ ; \*\*\*: adjusted p-value  $\leq 0.001$ .

To evaluate the robustness of this 3'UTR dataset, we compared the presence/absence and, in the case of presence, distance to the transcript 3' end of canonical and non-canonical PAS hexanucleotides between UCSC xenLae2 and Yang et al. 3'UTR datasets (Figure 40). In Yang et al.'s dataset, selecting the most abundant 3'UTR per

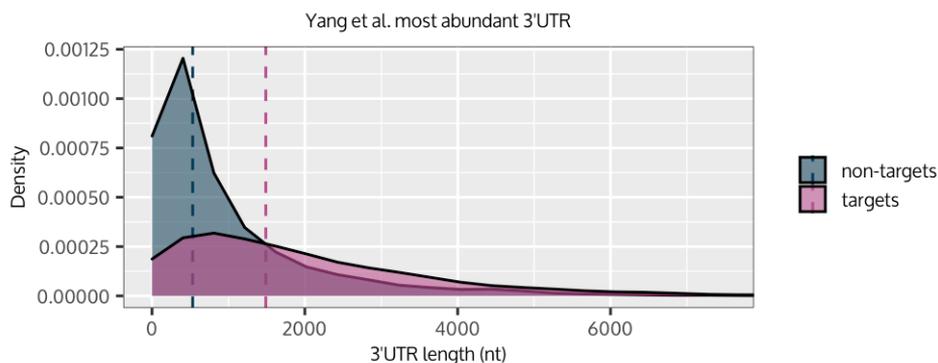
gene, 86% had a canonical or non-canonical PAS hex within 200 nucleotides from the transcript end and only 14% of transcripts did not, referred to as ‘hexless’ (**Figure 40a**, left). The median distances of the PAS hex to the transcript end were 26 and 30 nucleotides, respectively, for canonical and non-canonical PAS hex and in 95% of the cases the PAS hex was within 122 nucleotides from the transcript end (**Figure 40a**, right), as reported for this dataset and as observed with human mRNAs (X. Wu & Bartel 2017, F. Yang et al. 2020). In contrast, in the longest 3’ UTR per gene from the UCSC xenLae2 dataset, the percentage of hexless rose to 28% and the distance distribution was a lot more spread, with medians of 48 and 78 and 75% of the population falling within 122 nucleotides from the transcript end (**Figure 40b**). In light of these results, subsequent 3’ UTR level analyses were performed with Yang et al.’s dataset.



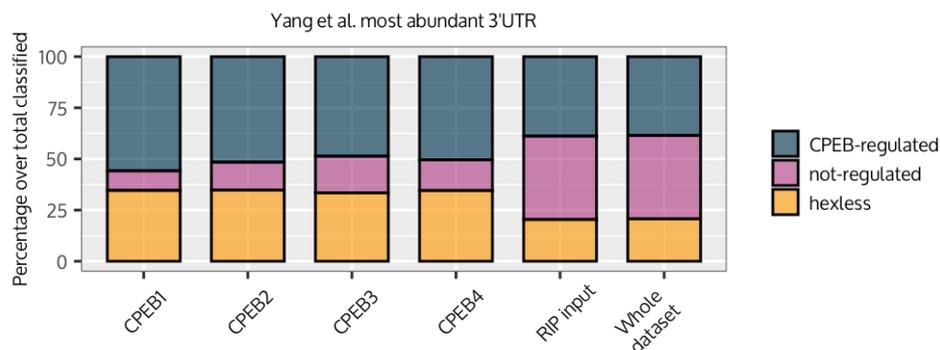
**Figure 40. Presence/absence and distance of PAS hex to the transcript 3’ end in two independent *X. laevis* 3’UTR datasets.** Left: number of transcripts in each dataset that harbor a PAS hex (canonical or non-canonical) within 200 nucleotides from the transcript end. Right: distance distribution of the PAS hex in the 3’UTR. Median distances are represented with a dashed line.

When we checked the 3’ UTR length of targets of any CPEB relative to non-targets, we found that targets were significantly longer than non-targets. The median 3’UTR length of targets was 1489 nucleotides versus 532 for non-targets and the length dis-

tributions were significantly different (**Figure 41**; non-parametric Wilcoxon rank-sum test p-value  $< 2.2 \cdot 10^{-16}$ ). If we excluded the 14% of hexless transcripts from the dataset, the differences between targets and non-targets were maintained. This agrees with previously published results from numerous systems that altogether indicate that highly-regulated localized transcripts have longer 3'UTRs (Courel et al. 2019, Khong et al. 2017, Stepien et al. 2016, Tushev et al. 2018).



**Figure 41. CPEB-targets have longer 3'UTRs than non-targets.** 3'UTR length distribution of targets of any CPEB and non-targets. Medians indicated with a dashed line.



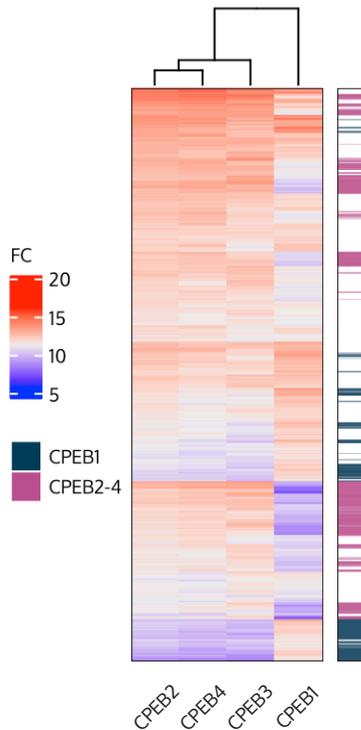
**Figure 42. RIP-Seq-determined CPEB-targets are enriched in motif architectures that predict regulation by CPEBs.**

We also found that targets of all CPEBs were enriched in arrangements of CPEs that predict regulation by CPEBs (**Figure 42**). For this classification, in view of our previous results (**Figure 40**), we used the combinatorial code defined by Piqué et al. but included non-canonical PAS hex and relaxed the maximal distance allowed from the PAS hex to the transcript end to 60 nucleotides. We determined that 55% of CPEB1 RIP-Seq

targets would be predicted to be CPEB-regulated by the combinatorial code, in contrast to 39% in the RIP-Seq input. We also noticed that the PAS hex was underrepresented in the CPEB-target groups relative to the input but the implications of this observation remain to be explored.

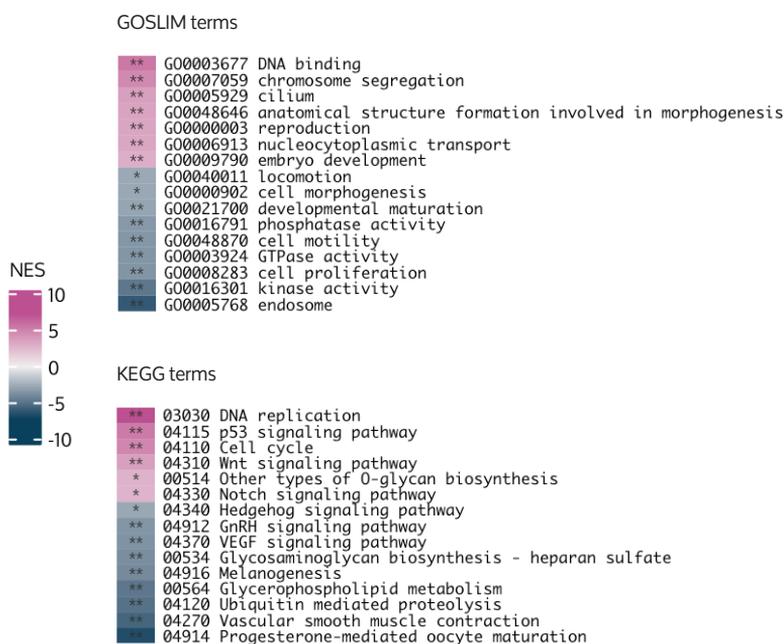
In short, we defined high-confidence target subsets for each CPEB. The mRNA targets of all four CPEBs were largely overlapping, with as much as 87% overlap and as little as 38% (in pairwise comparisons, relative to the largest set). Moreover, we found that the CPEB targets had significantly longer 3'UTRs than non-targets and were enriched in the motif architectures described to predict CPEB-regulation.

## Comparative analysis of CPEB1-4 targetomes



**Figure 43. CPEB1 and CPEB2-4 regulate different target subsets.** Complete-linkage hierarchical clustering of all genes that are targets of at least one CPEB based on the module of their fold-changes. The right hand heatmap shows genes enriched in the CPEB1 versus CPEB2-4 conditions ( $\log_2(\text{FC}) \geq 1$  and adjusted p-value  $\leq 0.05$ ).

We next sought to investigate whether we could delineate the differences between the targets bound by each CPEB. To this end, we first performed unsupervised clustering of the differential enrichment of the IPs versus controls for genes defined as targets of at least one CPEB. The clustering revealed that CPEB1 and CPEB2-4 formed two well-differentiated groups (**Figure 43**), hence we performed differential enrichment analysis of CPEB1 versus CPEB2-4 and defined differentially regulated targets by CPEB1 or CPEB2-4 as those that were at least 2-fold enriched in either group (adjusted p-value < 0.05). Within the subset of 1798 targets of any CPEB, 234 mRNAs were targets of CPEB1 and preferentially CPEB1-regulated, 414 were targets of CPEB2, CPEB3 or CPEB4 and preferentially CPEB2-4-regulated and 1148 were not differentially regulated in either group (**Figure 43**).



**Figure 44. ROAST GSEA of CPEB1 versus CPEB2-4 (union) RIP-Seq datasets with GOSLIM and KEGG terms.** All terms shown are also enriched over the Input background (NES > 0 and adjusted p-value ≤ 0.05). \*: adjusted p-value ≤ 0.05, \*\*: 0.001 ≤ adjusted p-value ≤ 0.01.

In order to find functional differences between the genes regulated by one or the other subfamily, we performed ROAST GSEA of CPEB2-4 versus CPEB1 (**Figure 44**). Overall, we found that processes related to DNA and the nucleus were enriched in the CPEB2-4 RIP-Seqs and, in contrast, terms related to the secretory pathway were enriched in the CPEB1 dataset.

We also wondered whether these preferentially-regulated subsets differed in their 3'UTR motif composition. Accordingly, we performed *de novo* motif enrichment analysis with two independent software, HOMER and MEME-Suite (Bailey et al. 2009, Benner et al. 2017). *De novo* motif enrichment revealed that targets of any CPEB, CPEB1-preferentially-regulated targets and CPEB2-4-preferentially regulated targets were all enriched in U-rich motifs (Table 3, Table 5, Table S10, Table S11). Specifically, CPEB1-preferentially-regulated targets were enriched in 'UGUUUUUAA', according to both software (Table 3, Table S12), which could correspond to a canonical CPE preceded by a poly-U stretch and a G. This motif was also significantly enriched when we compared the CPEB1-preferential targets to CPEB2-4-preferential targets (Table 4). The reciprocal comparison told us that CPEB2-4-preferential targets were instead enriched in a different U-rich motif 'UUUUGUA' both relative to the input and relative to CPEB1-preferential targets (Table 5, Table 6). This motif matches the definition of CPE-G, thus we found that CPEB2-4 targets are enriched in CPE-G.

**Table 3. HOMER differential motif enrichment in the 3'UTRs of CPEB1-preferentially regulated targets relative to input, as defined from RIP-Seq data.** Total target sequences = 147, total background sequences = 6407.

Rank	Motif	P-value	% of Targets	% of Background
1		1E-16	49.66%	18.19%
2		1E-13	38.10%	13.03%

**Table 4. HOMER differential motif enrichment in the 3'UTRs of CPEB1-preferentially regulated targets relative to CPEB2-4-preferentially regulated targets, as defined from RIP-Seq data.** Total target sequences = 147, total background sequences = 185.

Rank	Motif	P-value	% of Targets	% of Background
1		1E-16	67.35%	22.60%

**Table 5. HOMER differential motif enrichment in the 3'UTRs of CPEB2-4-preferentially regulated targets relative to input, as defined from RIP-Seq data.** Total target sequences = 210, total background sequences = 6805.

Rank	Motif	P-value	% of Targets	% of Background
1		1E-12	48.57%	24.87%

**Table 6. HOMER differential motif enrichment in the 3'UTRs of CPEB2-4-preferentially regulated targets relative to input, as defined from RIP-Seq data.** Total target sequences = 210, total background sequences = 148.

Rank	Motif	P-value	% of Targets	% of Background
1		1E-13	38.57%	5.89%

In summary, our RIP-Seq data has allowed us to define and explore two small subsets of targets differentially-regulated by either CPEB1 or the CPEB2-4 subfamily. Both subsets are enriched in U-rich motifs, although these motifs are different: canonical CPEs are enriched in the CPEB1 subset, whereas CPE-G are consistently enriched in CPEB2-4 subfamily preferential targets. Functionally, the CPEB1-preferential targets are enriched in genes involved in the secretory pathway, whereas nuclear processes are overrepresented among CPEB2-4 targets.

# Discussion



Meiotic maturation and early development are fascinating scenarios that have long been explored to address multifarious biological questions. Mature fully-grown oocytes, arrested at Pro-I, are transcriptionally silent. Thus, in the meiotic division and early embryonic development, gene expression relies on the timely mobilization of large stocks of RNAs accumulated during oocyte growth.

During meiotic maturation, the CPEB-family members CPEB1 and CPEB4 act sequentially to promote phase-specific activation of CPE-containing mRNAs that ensure meiotic progression and cell cycle arrest at MII, respectively (Guillén-Boixet et al. 2016, Igea & Méndez 2010, Stebbins-Boaz et al. 1996). In this work, we have added new information about CPEB2 and CPEB3 expression and regulation in the meiotic cell cycle. We have also comparatively investigated the CPEB family of RBPs proximosomes, mRNA targets and biomaterial properties. All in a quest to understand gene expression regulation by the whole CPEB family of RBPs.

Briefly, we have found that all CPEBs are co-expressed in oocytes and modified in the meiotic cell cycle (**Figure 14, Table 1**). CPEB1 is subject to a single activating phosphorylation at meiotic resumption (Mendez, Hake et al. 2000), while subsequently multiple Cdk1-phosphorylations at the onset of meiotic progression target it for degradation (Mendez et al. 2002, Setoyama et al. 2007). Instead, the CPEB2-4 subfamily members are subject to multiple Cdk1- and ERK2-phosphorylations in their N-terminal domain at the onset of meiotic progression (**Figure 16-Figure 21**, Guillén-Boixet et al. 2016). These multiple phosphorylations do not target them for degradation but, instead, change the proteins' properties making them more soluble and potentially triggering condensate dissolution (**Figure 23**, Guillén-Boixet et al. 2016).

We have determined that all CPEBs are capable of assembling into condensates with features compatible with formation by LLPS (**Figure 25, Figure 26, Table 2** and Guillén-Boixet et al. 2016). The condensates formed by the CPEBs differ in their dynamics, CPEB1 leaning to the more “solid” end of the spectrum than CPEB2 and CPEB3. Despite their different regulation and biomaterial properties, when co-expressed at interkinesis, all CPEBs co-exist in the same compartments (**Figure 27** and additional data about CPEB4 not shown). In line with this, as determined by BioID, in SVI-arrested oocytes, all CPEBs are near a core of proteins implicated in 3'end processing and translational repression. However, components differentially detected in the proximosomes indicate architectural and functional differences between the CPEBs (**Figure 30-Figure 34**).

Importantly, and again, despite their differential regulation and biomaterial properties, we have found that the four CPEBs bind a largely overlapping subset of mRNAs. CPEB-target mRNAs have long 3'UTRs, are enriched in motif architectures predictive of regulation by CPEBs and have a greater density of CPEs and CPE-like hexamers (**Figure 41**, **Figure 42**, **Table S10** and **Table S11**). Our data supports that all CPEBs bind linear CPE-like motifs and that the CPEB2-4 subfamily can, in addition, bind the CPE-G motif, thus controlling a subset of mRNAs not bound by CPEB1 (**Table 5** and **Table 6**). And vice-versa, CPEB1 also binds differentially a small subset of target mRNAs relative to CPEB2-4 (**Figure 43**).

With this work we have raised a number of hypotheses and questions that will be discussed and contextualized in the following sections. First and foremost, what can we infer from the function of the CPEBs, especially CPEB2 and CPEB3, in meiotic maturation? Second, how could non-physiological protein concentrations affect interpretation of some of our key experiments? And, last, what limitations does this work have and what are some of the future directions to take.

## Functions of the CPEB-family in meiotic maturation

CPEB1 and CPEB4 have sequential functions in *X. laevis* meiotic maturation. CPEB1 is required for meiotic resumption and progression whereas CPEB4 for the Met-II arrest. These functions have been demonstrated in knock-down (KD) oocytes generated by microinjection of either antisense oligonucleotides or neutralizing antibodies (Guillén-Boixet et al. 2016, Igea & Méndez 2010, Stebbins-Boaz et al. 1996).

Complementary, mouse models have provided additional information on the roles of the CPEBs in meiotic maturation and antecedent or preceding biological processes, i.e. germ cell formation and embryonic development. Specifically, CPEB1 knock-out (KO) female mice were born without ovaries or small vestigial ones, due to the requirement of CPEB1 for synaptonemal complex formation during Pro-I (Tay & Richter 2001). Females where CPEB1 was knocked-down specifically in growing oocytes presented numerous oocyte and ovarian abnormalities as well as compromised fertility (Racki & Richter 2006). Contrastingly, CPEB4-KO mice were born in mendelian ratios and displayed no overt phenotype at young ages under unchallenged conditions (Maillo et al. 2017). Females were fertile and had normal ovarian histology (Maillo et al. 2017). Similarly, CPEB2-KO and CPEB3-KO mouse models generated in our lab were also fertile and had normal ovarian histology (Pascual 2018; Martín, Suñer, Pascual et al. unpublished).

work).

This seeming discrepancy in CPEB4 function between mouse and *Xenopus* could have different explanations. First of all, there could be a differential requirement of CPEB4 in the Met-II arrest in *Xenopus* and mouse. Second, the phenotype of misaligned chromosomes observed in *Xenopus* Met-II-arrested-oocytes need not necessarily result in infertility, especially given that the penetrance is at most 50% (Guillén-Boix et al. 2016). Third, it is possible that CPEB2-4 single KO mice display some degree of subfertility, detectable only in thorough fertility studies. Last, considering our characterization of CPEB2 and CPEB3 in meiotic maturation, it is plausible to propose that the CPEB2-4 subfamily functions redundantly in maintaining the Met-II arrest. Hence, in KO mouse models, compensatory mechanisms between CPEB2-4 could be at play, masking the phenotype (El-Brolosy et al. 2019).

Functional redundancy and compensatory mechanisms within the CPEB2-4 subfamily have been previously proposed in other systems. Specifically, in the lactating mammary gland, CPEB4 transcriptional upregulation was observed in a constitutive CPEB2-KO model and proposed to functionally compensate for CPEB2 loss (Pascual 2018). In mouse embryonic fibroblasts, CPEB2 and CPEB4 have been proposed to function redundantly in Vinexin recruitment to SGs (Chang & Huang 2014). Moreover, constitutive double CPEB2/CPEB4-KO mice were embryonic lethal, unlike any other pairwise combination of CPEB2-4, for which one possible explanation is functional compensation between CPEB2 and CPEB4 (Martín, Suñer, Pascual et al. unpublished work).

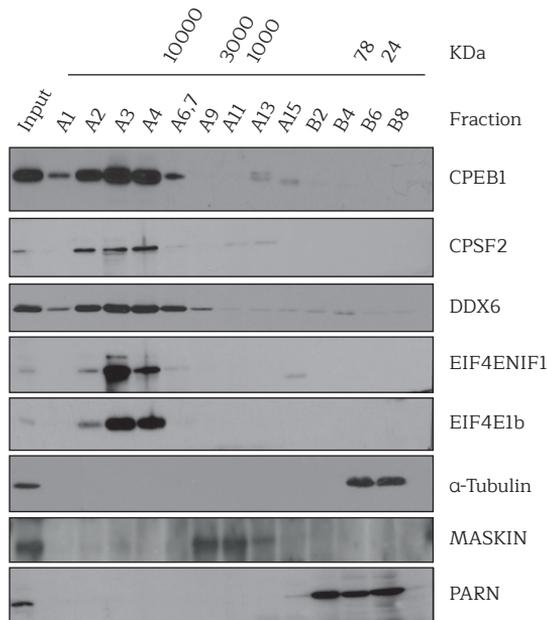
## A matter of protein concentration?

Although we have hereby presented that all CPEBs are expressed in meiotic maturation, we do not ignore that the amounts detected of CPEB2, CPEB3 and CPEB4 suggest very low cellular concentrations, which are mirrored by their mRNA concentrations and consistent with the proteomics studies of others (Peuchen et al. 2017, Session et al. 2016, Wühr et al. 2014). This has implications for the interpretation of some of the experiments presented in this thesis that have been done with overexpression.

LLPS has been suggested to be a universal property of proteins and nucleic acids under specific conditions, many of which may never be met physiologically (Alberti et al. 2019). Given the dependency of LLPS on protein concentration, the concentrations

of CPEB2-4 in meiotic maturation could be insufficient for them to drive condensate formation. However, even at low concentrations, CPEB2-4 could be clients in condensates nucleated by other proteins or their target mRNAs.

There is evidence that CPEB1 is found in large heterotypic mRNPs in oocytes. Recently, Neil and colleagues imaged CPEB1 in 5-10  $\mu\text{m}$  cytoplasmic gel-like mRNPs associated to vegetal mRNA transport in SII oocytes (Neil et al. 2020). L-bodies, as these large condensates were termed, eluted in the void volume of size-exclusion columns. Similarly, in SVI oocytes - where imaging is not so amenable - CPEB1 chromatographed in the void volume corresponding to complexes of 10 MDa and larger, along with other components of the repression and cytoplasmic polyadenylation machinery (**Figure 45**). Either imaging endogenous CPEB2-4 in previtellogenic oocytes or detection in size-exclusion chromatography fractions or sucrose gradients would help clarify this point.



**Figure 45. CPEB1 is found in large particles together with polyadenylation and repression machinery.** Size-exclusion chromatography of *X. laevis* oocytes Murray extracts, performed by Maria Piqué.

Interpretation of our RIP-Seq also warrants consideration of endogenous protein amounts. The RIP-Seq presented demonstrates that all CPEBs can bind largely the same mRNAs, that there are differences between CPEB1 versus CPEB2-4 binding and,

within the CPEB2-4 subfamily, between CPEB3 versus CPEB2/4. In vivo, with high CPEB1 and low CPEB2-4 concentration, CPEB2-4 control may be restricted to the subset of mRNAs defined as CPEB2-4-preferential - a subset of 415 genes that were recovered at least two-fold more efficiently for CPEB2, CPEB3 or CPEB4 than for CPEB1. The CPEB2-4-preferentially-regulated subset was related to G1-to-S cell cycle control and positive regulation of RNA pol II transcription and is henceforth in agreement with the CPEB2-4 function proposed earlier.

## LLPS in meiotic maturation

In this thesis we have demonstrated that CPEB1, CPEB2 and CPEB4 localize in cellular condensates whose properties are compatible with formation by LLPS. The CPEBs colocalize among them and also co-localize to other granule proteins such as DDX6 or CNOT2 (preliminary data, not shown). Upon phosphorylation, CPEB2-4 adopt a diffuse cellular distribution rather than granular, and CPEB1 is degraded.

With ours and other's data already presented, we propose a model whereby in SVI oocytes, the CPEBs localize with repressed mRNAs in heterotypic cytoplasmic condensates. Condensates might be the best way to store mRNAs for a long time in a protected environment.

Moreover, we hypothesize that upon Cdk1 activation and MI-entry, widespread phosphorylation triggers condensate disassembly and translational activation. In fact, disassembly of *mos* and cyclin B1 RNA granules during oocyte maturation has been reported in Zebrafish oocytes (Horie & Kotani 2016, Kotani et al. 2013). In neurons, Neuronal granules disassembly.

We envision a great regulatory complexity in condensate disassembly, involving multiple kinases and post-translational modifiers, as well as modifications exerted on multiple condensate proteins. The effects of these modifications on condensate interactions will ultimately all combine to give rise to distinct disassembly dynamics. Thus, it would be of great interest, in order to understand CPEB-mediated regulation in meiotic maturation and, more broadly, the mechanisms of granule disassembly, to map modifications that occur on all condensate proteins upon a given stimuli.

The regulatory complexity in condensate disassembly and translational repression and activation also involves mRNAs, not as passive cargo controlled by proteins, but

as active players in molecular interactions that are also susceptible to regulation. We identified the methylation readers YTHDF1 and YTHDF2 in the CPEBs proximalomes and RNA-methylation has previously been linked to both repression and mRNP condensation. Thus, we think that the effect of RNA-modifications in granule dynamics and translational activation and repression should be further investigated.

# Conclusions



In this thesis we have conducted a detailed characterization of CPEB2 and CPEB3 in meiotic maturation and investigated the relationships established between the four CPEB-family members. The conclusions we have obtained are the following:

1. The four CPEBs are co-expressed in *X. laevis* meiotic maturation.
2. CPEB3 is extensively phosphorylated at meiotic progression.
3. CPEB3 phosphorylation occurs at proline-directed sites concentrated on its N-terminal domain, specifically at sites: S21, S78, S145, S148, S178, S181, S195, S231, S251, T254 and S353 and at least one of S47, S59, T103, S129 and S195.
4. CPEB3 is phosphorylated by the proline-directed kinases Cdk1 and ERK2.
5. CPEB2 is extensively phosphorylated at consensus proline-directed sites clustered on its N-terminal domain and, possibly, at additional non-consensus sites.
6. CPEB1, CPEB2 and CPEB3 form dynamic cytoplasmic condensates that display properties compatible with LLPS.
7. CPEB2 and CPEB3 extensive N-terminal phosphorylation has a solubilizing effect on the cytoplasmic condensates they form.
8. CPEB2 and CPEB3 condensates are more liquid than CPEB1 condensates.
9. When co-expressed, CPEB1, CPEB2 and CPEB3 co-localize in the same compartments.
10. CPEB1-4 share a core proximome of mRNA processing, storage and repression proteins, that includes pre-mRNA 3'end processing machinery, components of the CCR4-NOT machinery, microRNA-induced silencing machinery and RBPs involved in RNA transport and translational control. However, beyond this common core, each CPEB also displays unique interactors.
11. The CPEBs regulate a largely overlapping set of mRNAs, characterized by having longer 3'UTRs than non-regulated transcripts.
12. The CPEB1 subfamily and CPEB2-4 subfamily each differentially regulate a subset of mRNAs.



# Materials and methods



# Buffers

## Modified Barth's Solution (MBS)

88 mM NaCl, 1 mM KCl, 1 mM MgSO<sub>4</sub>, 2.5 mM NaHCO<sub>3</sub>, 0.7 mM CaCl<sub>2</sub>, pH 7.8.

## H1K 0.4% NP40 lysis solution

80 mM sodium β-glycerophosphate pH 7.4, 20 mM EGTA, 15 mM MgCl<sub>2</sub>, 0.5 mM Na<sub>3</sub>VaO<sub>4</sub> (sodium orthovanadate), 0.4% NP40, 1x Complete EDTA-free Protease Inhibitors (Roche).

## Laemmli sample buffer

180 mM Tris-HCl pH 6.8, 30% glycerol, 6% SDS, 0.03% bromophenol blue, 180 mM DTT.

## IP lysis buffer

20 mM Tris-HCl pH 8, 100 mM NaCl, 0.4% NP40, 1 mM EDTA, 1 mM MgCl<sub>2</sub>, 1x Complete EDTA-free protease inhibitors (Roche).

## H1K protease inhibitors

80 mM sodium β-glycerophosphate, 0.5 mM sodium orthovanadate.

## BioID lysis buffer

50 mM Tris-HCl pH 7.4, 100 mM NaCl, 1x Complete EDTA-free protease inhibitors (Roche), 1 mM PMSF.

## Proteinase K buffer

200 mM Tris-HCl pH 7.5, 100 mM NaCl, 10 mM EDTA, 1% SDS, 5 μL/mL Ribolock (40 U/μL, Thermo Fisher Scientific)

### **Protein Purification lysis buffer**

25 mM Tris-HCl pH 8, 1 M NaCl, 5 mM MgCl<sub>2</sub>, 1% NP40, 5% glycerol, 6 M urea, 10 mM imidazole, 1x Complete EDTA-free Protease Inhibitors (Roche), 1 mM PMSF.

### **Protein Purification wash buffer**

25 mM Tris-HCl pH 8, 0.5 M NaCl, 5 mM MgCl<sub>2</sub>, 5% glycerol, 6 M urea, 20 mM imidazole, 1x Complete EDTA-free Protease Inhibitors (Roche), 1 mM PMSF.

### **Protein Purification elution buffer**

25 mM Tris-HCl pH 8, 0.3 M NaCl, 5 mM MgCl<sub>2</sub>, 5% glycerol, 6 M urea, 300 mM imidazole, 1x Complete EDTA-free Protease Inhibitors (Roche), 1 mM PMSF.

### **Protein Purification dialysis buffer**

25 mM Tris-HCl pH 8, 0.1 M NaCl, 5 mM MgCl<sub>2</sub>, 10% glycerol, 10 mM imidazole, 1x Complete EDTA-free Protease Inhibitors (Roche), 1 mM PMSF.

### **U-2 OS lysis buffer**

50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% NP40, 5 mM EGTA, 5 mM EDTA, 20 mM NaF, 10 mM β-glycerophosphate, 1 mM DTT, 1x Complete EDTA-free protease inhibitors (Roche), 1 mM PMSF, Phosphatase Inhibitors Cocktail 2 (P5726, Sigma Aldrich) and Phosphatase Inhibitors Cocktail 3 (POO44, Sigma Aldrich).

## ***X. laevis* oocytes preparation**

Ovarian lobes were obtained by laparoscopy of *X. laevis* females. The ovarian lobes were kept in MBS and digested by addition of 0.8 mg/mL Collagenase (C9891, Sigma-Aldrich) and 0.48 mg/mL Dispase II (D4693, Sigma-Aldrich) for 2 hours at 22°C with gentle rocking. After digestion, oocytes were thoroughly washed in MBS and kept in MBS throughout selection, microinjection and maturation.

SVI oocytes were selected manually using a stereomicroscope (Leica MZ9.5). Mi-

croinjection was performed under the microscope with Nanoject II Auto-Nanoliter Injector (Drummond) and 3.5'' glass capillaries (3-000-203-G/X, Drummond). Oocytes were microinjected with 46 nL of 50 ng/ $\mu$ L in vitro transcribed RNA, unless otherwise specified. For protocols requiring overnight incubation, oocytes were kept at 18°C in the incubator.

Oocyte maturation was induced by addition of 10  $\mu$ M progesterone (P8783, Merck). Maturation was scored by the appearance of a wms.

MII-arrested oocyte activation was induced with 4  $\mu$ M calcium ionophore A23187 (C7522, Merck). Activation was scored by observation of cortical contraction of the animal pole pigmentation.

## Oocyte lysis

Oocytes were collected in microcentrifuge tubes, the MBS carefully removed and, unless otherwise specified, lysed by pipetting in 10  $\mu$ L/oocyte cold H1K 0.4% NP40 lysis solution. Homogenates were spun 10 minutes at 12000 x g 4°C, after which the clear phase was recovered and used for downstream applications.

## Subcloning and mutagenesis

Inserts were produced by PCR with custom-designed DNA oligonucleotides (ordered from Merck). Restriction sites were introduced with the primers. Inserts and plasmids were digested with the corresponding restriction enzymes (both from NEB and Thermo Fisher Scientific) and purified with the Nucleospin Gel and PCR Clean-up kit (Macherey-Nagel). Ligation was performed at 1:3 vector:insert molar ratio of with 400 U T4 DNA Ligase (NEB) for 2 hours at RT. 50  $\mu$ L Subcloning Efficiency DH5a Competent Cells (Thermo Fisher Scientific) were transformed with 2  $\mu$ L of ligation product following the manufacturer's instructions. A sample of the resulting colonies was screened by colony-PCR or by sequencing (EZ-Seq service, Macrogen) after plasmid DNA miniprep (with NucleoSpin Plasmid Miniprep kit, Macherey Nagel). The positive clones were used for larger-scale DNA preparations (with NucleoBond Xtra Maxim Kit, Macherey Nagel). All large-scale DNA preps were confirmed by sequencing.

When suitable restriction sites could not be introduced, subcloning was performed

by Gibson assembly (with a commercial master mix, from NEB).

All plasmids for expression in mammalian cells were produced subcloning with In-Fusion (Takara Bio) following manufacturer's instructions. The linearized pPEU vectors were provided by the Protein Expression Facility of IRB.

Last, some plasmids were generated by mutagenesis rather than DNA copy-paste. Mutagenesis reactions were done with QuikChange Lightning Multi Site-Directed Mutagenesis Kit (Agilent).

The accessions or sequences of the CPEB protein used throughout this thesis are listed on **Table S14**. The oligonucleotides used for CPEB2 and CPEB3 mutagenesis are listed on **Table S15** and **Table S16**. All plasmids used throughout this thesis are listed on **Table S17**, **Table S18**, **Table S19** and **Table S20**.

## ***In vitro* transcription (ivt) and polyadenylation**

RNAs used to overexpress proteins in oocytes were *in vitro* generated using mMessage mMachine T3 Transcription Kit (Ambion) with 300 ng of linearized template and following manufacturer's instructions. An additional poly(A) tailing reaction was performed when necessary, using the Poly(A) Tailing Kit (Ambion), after Turbo DNase digestion and prior to RNA precipitation with lithium chloride. The final RNA precipitate was resuspended to 500 ng/ $\mu$ L and small working aliquots at 50 ng/ $\mu$ L were prepared.

## **SDS-PAGE and WB**

### **Lysates denaturation**

Lysates were denatured by adding 1/3 volumes of Laemmli sample buffer. The lysates were then boiled 5 minutes at 95°C and stored (-20°C) or immediately used.

### **Electrophoresis**

8-10% acrylamide/bis-acrylamide (1610148, BioRad) 0.1% SDS gels were casted. Electrophoresis were performed for 1.5-2.5 hours at 80-120 V in a Mini-Protean Tetra

Cell system (Bio-Rad). 10  $\mu\text{L}$  of molecular weight marker were routinely loaded on the left-hand-side of the gel (SeeBlue Plus2 Pre-stained Protein Standard, Thermo Fisher Scientific).

## Immunoblotting

Protein gels were transferred to PVDF membranes (Immobilon-P 0.45  $\mu\text{m}$ , IPVH00010, Merck) using a wet system (Mini Trans-Blot Cell, BioRad) for 1-2 hours at 125 V. Membranes were blocked with 5% milk in Tris-buffered Saline 0.05% Tween-20 (TBS-T) for at least 40 minutes. Primary antibodies were probed from 1 hour to overnight. Secondary antibodies were probed for 3 hours at RT. A summary of all antibodies used can be found on **Table S21**. Antibody-probed membranes were developed with GE Healthcare Amersham ECL WB Detection Reagents (Thermo Fisher Scientific) and the signal was captured in GE Healthcare Hyperfilm ECL (Thermo Fisher Scientific).

## B-isox precipitation

Enrichment of LC-containing proteins was performed by b-isox (900572, Merck) precipitation. Oocytes were lysed by pipetting in 9  $\mu\text{L}$ /oocyte cold IP lysis buffer. Homogenates were spun 10 minutes at 12000 x g 4°C, after which the clear phase was recovered. 1  $\mu\text{L}$  of 10x HIK phosphatase inhibitors were added per 9  $\mu\text{L}$  of recovered aqueous phase. The homogenates were subjected to a second clarification. At this point, the equivalent to 1-2 oocytes were put aside to be used as protein input. The remaining cleared lysates were exposed to 100 or 200  $\mu\text{M}$  of b-isox and rotated overnight at 4°C. The incubated reactions were precipitated by centrifugation for 10 minutes, 10000 x g at 4°C. The pellets were washed twice in cold lysis buffer (vortexing, incubating 10 minutes on ice and spinning for 10 minutes at 10000 x g 4°C) and resuspended in Laemmli sample buffer.

## Protein IP

### Lysis

Unless otherwise specified in the corresponding sections, oocytes expressing epitope-tagged proteins were lysed by pipetting in 9  $\mu\text{L}$ /oocyte cold IP lysis buffer. Ho-

mogenates were spun 10 minutes at 12000 x g 4°C, after which the clear phase was recovered. 1 µL of 10x H1K phosphatase inhibitors were added per 9 µL of recovered aqueous phase. At this point, the equivalent to 1-2 oocytes were put aside to be used as protein input.

### **Bead preparation**

IP of HA-tagged proteins was performed with HA-conjugated beads, either conjugated by us or purchased in the conjugated form (Anti-HA High Affinity rat monoclonal, 11867423001, Merck; Dynabeads Protein G, 10004D, Thermo Fisher Scientific; Pierce Anti-HA Magnetic Beads, 88837, Thermo Fisher Scientific).

In order to covalently conjugate the antibody to the beads, only in the case of non-preconjugated beads, the bead slurry was washed twice in PBS and incubated with the antibody with orbital shaking 2 hours at RT. 1 µg of antibody was used per 100 µL bead slurry. Beads were washed once in PBS, twice in crosslinking buffer (0.2 M triethanolamine pH 8) then incubated 1 hour 30 minutes in crosslinking buffer with crosslinking reagent (5.4 mg/mL dimethyl pimelimidate, 21666, Thermo Fisher Scientific). The crosslinking was stopped by two 5-minute washes in 50 mM Tris-HCl pH 8. After two washes in PBS, non-crosslinked antibody was removed by a 2-minute 0.1 M citric acid pH 3 wash. Last, two washes in lysis buffer were performed before using the beads. Preconjugated beads were washed twice PBS and twice in lysis buffer prior to use.

### **Immunoprecipitation**

In all IPs, the proportion of 150 µL conjugated-bead-slurry to 100 oocytes (or 1 mL clarified lysate) was maintained. The IPs and washes were performed at 4°C. Lysate and beads were incubated on the wheel 2-20 hours (depending on the protocol, incubation time specified on the corresponding sections). The beads were washed thrice with one volume of lysis buffer.

### **Elution**

The purified proteins were eluted from the beads by incubating with Laemmli sample buffer without DTT 20 minutes at 60°C. The supernatants were then separated from the beads with a magnetic rack (DynaMag-2, Thermo Fisher Scientific) and the beads discarded. DTT was added to a final concentration of 180 mM. The eluates were

further denaturalized by boiling 5 minutes 95°C and stored (-20°C) or immediately used for downstream applications (generally PAGE and immunoblotting).

## Lambda Protein Phosphatase ( $\lambda$ -PPase) assay

Two different setups were used for these experiments.

On the one hand, microinjected oocytes at different maturation stages were collected and lysed in 10  $\mu$ L/oocyte cold 1x  $\lambda$ -PPase buffer (from NEB; 50 mM Tris-HCl pH 7.5, 100 mM NaCl, 0.1 mM EGTA, 2 mM DTT and 0.01% Brij 35). Homogenates were spun 10 minutes at 12000 x g 4°C, after which the clear phase was recovered. 10  $\mu$ L of clarified oocyte lysate were supplemented with 2 mM  $MnCl_2$  and 400 U  $\lambda$ -PPase (P0753, NEB) and incubated for 1 hour 30 minutes at 30°C.

On the other hand, microinjected oocytes at different maturation stages were collected and lysed in 10  $\mu$ L/oocyte cold 1x H1K 0.4% NP40 lysis solution supplemented with 100 mM N-Ethylmaleimide (NEM, O4259 Sigma). Homogenates were spun 10 minutes at 12000 x g 4°C, after which the clear phase was recovered. One volume of lysate was mixed with one volume of 2x  $\lambda$ -PPase master mix [2x  $\lambda$ -PPase buffer, 4 mM  $MnCl_2$ , 800 U  $\lambda$ -PPase (P0753, NEB)] and incubated for 1 hour at 30°C.

In both cases, the reactions were stopped by addition of 1/3 sample volume of Laemmli sample buffer, incubated 5 minutes at 95°C and subsequently stored at -20°C or loaded on gel.

## CPEB3 p-site mapping by MS

Overexpressed HA-CPEB3 was immunoprecipitated as specified in the 'Protein IP' section (page 105) with the following specifications: 100 oocytes were used per IP; lysis was performed in 9  $\mu$ L/oocyte cold H1K 0.4% NP40 in order to preserve phosphorylation status, after clarification, 1  $\mu$ L of 10x IP lysis buffer [without NP40 nor Complete EDTA-free Protease Inhibitors (Roche)] was added per 9  $\mu$ L of recovered aqueous phase; the HA-antibody was conjugated to the beads as specified; the IPs were incubated for 1 hour 30 minutes. As with other protocols involving MS, recombinant protein was handled in Eppendorf LoBind microcentrifuge tubes (Thermo Fisher Scientific).

## SDS-PAGE and Silver staining

The IP elutions were run on precast 7.5% gels (Mini-Protean TGX, Bio-Rad). The gels were silver-stained using the Pierce Silver Stain for MS kit (24600, Thermo Fisher Scientific) following manufacturer's instructions. The bands of interest were cut, placed on microcentrifuge tubes and handed to the facility. Sample processing and analysis from this point on was performed at the IRB Barcelona Mass Spectrometry and Proteomics Core Facility.

## Sample preparation for MS

Gel bands were washed with 50 mM  $\text{NH}_4\text{HCO}_3$  (ammonium bicarbonate) and acetonitrile. Disulfide bond reduction was performed with 10 mM DTT and alkylation of thiol groups with 50 mM iodoacetamide. Some samples were digested with trypsin (V511, Sequencing Grade Modified Trypsin, Promega) and others with chymotrypsin (V1062, Chymotrypsin Sequencing Grade, Promega). The digestion was stopped by addition of 5% HCOOH (formic acid) and eluted. The samples were evaporated and reconstituted in 3% acetonitrile and 1% formic acid. For the nano-LC-MS/MS, 50% of the sample volume was used.

## Nano-LC-MS/MS

Samples were loaded to a C18 precolumn (Acclaim PepMap100, 5 mm, 100 Å, 300 mm x 5 mm, Thermo Fisher Scientific) at a flow rate of 15  $\mu\text{l}/\text{minute}$  using a Dionex UltiMate 3000 chromatographic system (Thermo Fisher Scientific). Peptides were separated using a C18 analytical column (NanoEase MZ HSS T3 1.8  $\mu\text{m}$ , 100 Å, 75  $\mu\text{m}$  x 250 cm, Waters) with a 90 minutes run, comprising three consecutive steps with linear gradients from 1 to 35% B in 90 minutes, from 35 to 50% B in 5 minutes, and from 50% to 85% B in 2 minutes, followed by isocratic elution at 85% B in 5 minutes and stabilization to initial conditions (where B was 0.1% formic acid in acetonitrile and the other eluate was 0.1% formic acid in water). The column outlet was directly connected to an Advion TriVersa NanoMate (Advion) fitted on an Orbitrap Fusion Lumos Tribrid (Thermo Fisher Scientific).

The mass spectrometer was operated in a data-dependent acquisition (DDA) mode. Survey MS scans were acquired in the Orbitrap with the resolution (defined at 200  $m/z$ ) set to 120,000. The lock mass was user-defined at 445.12  $m/z$  in each Orbitrap scan. The top speed (most intense) ions per scan were fragmented and detected in

the linear ion trap. The ion count target value was 400,000 for the survey scan and 50,000 for the MS/MS scan. Target ions already selected for MS/MS were dynamically excluded for 30 s. Spray voltage in the NanoMate source was set to 1.60 kV. RF Lenses were tuned to 30%. Minimal signal required to trigger MS to MS/MS switch was set to 25000. The spectrometer was working in positive polarity mode and singly charge state precursors were rejected for fragmentation.

### Peptide identification

MS/MS spectra were searched against UniprotKB *X. laevis* database, release 2018\_10, contaminants and user-defined proteins using both both MaxQuant software v1.6.2.6a with Andromeda search engine and Proteome Discoverer v2.1 (Thermo Fisher Scientific) with Sequest HT and Amanda search engines. Searches were run against targeted and decoy database to determine the false discovery rate (FDR). Search parameters included trypsin enzyme specificity, allowing for two missed cleavage sites, oxidation in methionine, phosphorylation in serine/threonine/tyrosine and acetylation in protein N-terminus as dynamic modifications and carbamidomethyl in cysteine as static modification. Peptide mass tolerance was 10 ppm, the MS/MS tolerance was 0.02 Da and minimal peptide length was 7 amino acids. Peptides with a q-value lower than 0.1 and FDR < 1% were considered as positive identifications with a high confidence level.

### Analysis

A phosphorylation ratio ( $r$ ) for each phosphorylation site (p-site) within CPEB3 protein was computed considering three search nodes: Andromeda, Amanda and Sequest. For each p-site, the number of position-specific phosphorylated peptide spectrum matches (PSMs) ( $N_{\text{Phos}}$ ) and the number of non-phosphorylated PSMs ( $N_{\text{NonPhos}}$ ) were counted; from which  $r$  was then computed as follows:  $r = N_{\text{Phos}} / (N_{\text{Phos}} + N_{\text{NonPhos}})$ . Only those p-sites with localization probability greater than 75% were considered.

## SVI *X. laevis* BioID

### In vivo biotinylation

Per condition, circa 150 oocytes were microinjected with the corresponding BirA-containing ivt RNAs. The injected oocytes were incubated in MBS with 20  $\mu\text{M}$

biotin (B4501, Merck) for 40 hours at 18°C. As with other protocols involving MS, recombinant protein was handled in Eppendorf LoBind microcentrifuge tubes (Thermo Fisher Scientific).

### **Biotinylated peptides affinity purification**

Oocytes were lysed in 6  $\mu\text{L}$ /oocyte cold BioID lysis buffer and centrifuged twice for clarification 15 minutes at 16000  $\times g$  4°C. 300  $\mu\text{L}$  of cold BioID lysis buffer were added to 200  $\mu\text{L}$  of cleared extract and the resulting 500  $\mu\text{L}$  were subjected to clearing with PD MiniTrap G-25 columns (28-9180-07, GE Healthcare). Triton-X100 and SDS were added to a final concentration of 1.6% and 0.04%, respectively, and the volume was increased to 1 mL, maintaining all concentrations.

Cleared extracts were incubated with 200  $\mu\text{L}$  MyOne Dynabeads Streptavidin C1 (65001, Invitrogen) 20 hours with orbital shaking at 4°C. After incubation, beads were washed thrice with wash buffer 1 (8 M Urea, 0.25% SDS in PBS), twice with wash buffer 2 (6 M Guanidine-HCl in PBS), once with wash buffer 3 (6.4 M Urea, 1 M NaCl, 0.2% SDS in PBS), thrice with wash buffer 4 (4 M Urea, 1 M NaCl, 10% isopropanol, 10% ethanol, 0.2% SDS in PBS), once with wash buffer 1, once with wash buffer 5 (8 M Urea and 1% SDS in PBS) and thrice with wash buffer 6 (2% SDS in PBS). The washed beads were further washed with 50 mM Tris-HCl pH 7.4 to remove detergents that could interfere with MS and 50 mM  $\text{NH}_4\text{HCO}_3$  pH 8 to eliminate the traces of Tris.

### **Sample preparation for MS**

For disulfide bond reduction, beads were resuspended in 500  $\mu\text{L}$  3 M urea, 50 mM  $\text{NH}_4\text{HCO}_3$  (ammonium bicarbonate) pH 8.0, 5 mM DTT for 1 hour with orbital shaking at RT. Alkylation of thiol groups was achieved by incubation in the dark with 10 mM iodoacetamide for 30 minutes at RT. Alkylation was stopped by addition of excess DTT (5  $\mu\text{L}$  of 500 mM). Sample volumes were brought to 1.5 mL with 50 mM  $\text{NH}_4\text{HCO}_3$  pH 8 to reduce urea concentration and proteins were digested on-bead with 2  $\mu\text{g}$  trypsin (V5111, Promega) 16 hours with orbital shaking at 37°C. Tryptic digestion was stopped by addition of 1% HCOOH (formic acid). The supernatant, containing released peptides, was recovered.

Sample processing from this point on was performed at the IRB Barcelona Mass Spectrometry and Proteomics Core Facility.

Digested peptide samples were desalted with PolyLC tips C18. Peptides were eluted with 80% acetonitrile and 1% formic acid. Samples were diluted to 20% acetonitrile and 0.1% formic acid and loaded into strong cation exchange columns. Peptides were eluted in 5%  $\text{NH}_4\text{OH}$  (ammonium hydroxide) and 30% methanol. Finally, samples were evaporated and reconstituted in 3% acetonitrile and 1% formic acid. For the nano-LC-MS/MS, 10% of the samples volume was used.

### **Nano-LC-MS/MS**

The Nano-LC-MS/MS was performed as specified in ‘CPEB3 p-site mapping by MS’ section ‘Nano-LC-MS/MS’ subsection (page 108) with a few modifications: peptides were separated using a C18 analytical column (Acclaim PepMap RSLC, 2  $\mu\text{m}$ , 100  $\text{\AA}$ , 75  $\mu\text{m}$  x 50cm, nanoViper, Thermo Fisher Scientific) with a 120 minutes run; the ion count target value was 10,000 for the MS/MS scan; target ions already selected for MS/MS were dynamically excluded for 15 s and minimal signal required to trigger MS to MS/MS switch was set to 5000 and activation Q was 0.250.

### **Peptide identification**

MS/MS spectra were searched against UniprotKB *Xenopodinae* release 2017\_02, contaminants and user-defined proteins using Proteome Discoverer v2.1.0.81 (Thermo Fisher Scientific) with Sequest HT search engine. Searches were run against targeted and decoy database to determine the false discovery rate (FDR). Search parameters included trypsin enzyme specificity, allowing for two missed cleavage sites, oxidation in methionine and acetylation in protein N-terminus as dynamic modifications. When specified, biotin in lysine was included as a dynamic modification as well. Peptide mass tolerance was 10 ppm and the MS/MS tolerance 0.6 Da. Peptides with a q-value lower than 0.1 and FDR < 1% were considered positive identifications with a high confidence level.

### **Statistical analysis**

The IRB Barcelona Bioinformatics and Biostatistics Core Facility performed results’ analysis from this point on.  $\text{Log}_{10}$ (iBAQ) values were used as protein intensity. Percentile normalization of the data was performed to minimize batch effects across biological replicates. K-nearest neighbours (kNN) was used for missing values imputation [function `impute.knn` from the `impute` R package (Hastie et al. 2020, R Core Team 2020)] with k set to 10. Only the cases with one or two missing values were imputed. Cases

with 3 or 4 missing observations were manually included only when the 3 or 4 missings occurred in the control condition and the test condition had 1 or none plus, the average  $\log_{10}(\text{iBAQ})$  of the test was not within the 25% percentile of values detected in the whole sample (so as to filter out weaker-intensity proteins). Candidate interactions were found following a differential expression analysis, using functions `lmFit` and `eBayes` from `limma` R package (Smyth 2005). Biological replicates were used as adjusting variable in the model when needed.

### **Validation of BioID candidates by co-IP**

Overexpressed HA-tagged proteins were immunoprecipitated as specified in the ‘Protein IP’ section (page 105) with the following specifications: 30-60 oocytes were used per IP, with pre-conjugated beads (Pierce Anti-HA Magnetic Beads, 88837, Thermo Fisher Scientific) and the IPs were incubated for 20 hours.

## **RIP-Seq**

The RIPs were performed as described in the ‘Protein IP’ section (page 105) with modifications in order to extract both protein and RNA. 120-140 oocytes were injected per condition; 120 oocytes were lysed in 9  $\mu\text{L}/\text{oocyte}$  cold IP lysis buffer, after clarification, 1  $\mu\text{L}$  of 10x H1K phosphatase inhibitors were added and a second clarification step was performed; 50  $\mu\text{L}$  of clarified lysate were put aside as protein input and 50  $\mu\text{L}$  to extract RNA and obtain the RNA input; 600-800  $\mu\text{L}$  of clarified lysate were incubated with the HA-beads for 2 hours, with Ribolock (Thermo Fisher Scientific) added at 12  $\mu\text{L}/\text{mL}$  (or 0.5 U/ $\mu\text{L}$ ); the wash buffer was also supplemented with 12  $\mu\text{L}/\text{mL}$  (or 0.5 U/ $\mu\text{L}$ ) of Ribolock (Thermo Fisher Scientific); before elution, 1/10 volume of lysate-beads slurry was put aside to elute the proteins, while the remaining fraction was used to elute the RNA. The protein fractions were treated as described in previous sections and subjected to SDS-PAGE and immunoblotting in order to control the IPs.

### **RNA elution and extraction**

After removal of all lysis buffer from the beads, 50  $\mu\text{L}$  of IP lysis buffer supplemented with 12  $\mu\text{L}/\text{mL}$  of Ribolock (Thermo Fisher Scientific) were added. Both RNA inputs and RNA IPs were treated at the same time in the same way. First, 400  $\mu\text{L}$  of Proteinase K buffer with 200  $\mu\text{g}/\text{mL}$  Proteinase K (3115887001, Proteinase K, recombinant, PCR grade, Roche) were added. The digestions were incubated 30 minutes at 37°C and

the supernatant (in the case of the IP) was recovered.

RNA was purified by organic phase extraction using Trizol Reagent (Thermo Fisher Scientific) and chloroform followed by ethanol precipitation. Briefly, 1 sample volume of Trizol was added, mixed by shaking and incubated for 5 minutes at RT. 0.2 volumes of chloroform were added and, after vigorous shaking and 10-minute centrifugation at 16000 g, 4°C, the aqueous phase was recovered in a clean tube. The chloroform extraction was repeated. The RNA contained in the aqueous phase was precipitated overnight at -20 °C by addition of 10 µL 3 M sodium acetate and 300 µL of ethanol. A visible pellet was formed after 20-minutes-centrifugation at 16000 g, 4°C. The supernatant was removed and the pellet was washed thrice in 75% ethanol prior to reconstitution in 40 µL (inputs) or 15 µL (eluates) of TURBO DNase premix [1x TURBO buffer (included in the TURBO DNA-free kit, AM1907, Invitrogen), 1 U/µL Ribolock (Thermo Fisher Scientific)]. Once reconstituted, 1 µL TURBO DNase was added and samples were incubated for 30 minutes at 37°C. The enzyme was inactivated following manufacturer's instructions. The recovered supernatants were brought to the IRB Barcelona Functional Genomics Core Facility for further processing.

### **Library preparation and sequencing**

RNA samples were quantified by fluorometry with Qubit RNA HS Assay kit (Thermo Fisher Scientific). RNA integrity was assessed with the Agilent RNA 6000 Pico chip (Agilent) and the Agilent 2100 Bioanalyzer instrument (Agilent). The mRNA of the Inputs was purified using the kit NEBNext Poly(A) mRNA Magnetic Isolation Module (New England BioLabs) following manufacturer's instructions. The mRNAs of the IPs samples were not purified. Library preparation was performed using the kit NEBNext Ultra II library prep kit for Illumina (New England Biolabs) following the manufacturer's instructions and 8-10 cycles of library amplification. Finally, an equimolar pool was generated with all the samples and sequenced in two 50 nt Single Read lanes of a HiSeq 2500 Sequencer (Illumina).

### **Mapping and differential enrichment analysis**

The IRB Barcelona Bioinformatics and Biostatistics Core Facility performed the analysis. Raw FastQ files were aligned against the *X. laevis* genome (UCSC version 9.2, excluding chrUn chromosomes), with Bowtie2 2.2.2 (Langmead et al. 2009), 1 mismatch and reporting best alignment site per read. FastQC v011 (Andrews 2010) was used to perform a quality control overview of FastQ files and aligned BAM files. PCA-like plots

to assess sample similarity using coverage correlations were performed using htSeqTools 1.20 (Planet et al. 2018). Binary tracks for all reads in TDF format were generated with IGVTools 2 (Robinson et al. 2011). Default IP versus Input peak calling for enrichment evaluation was performed using MACS 1.4.2 with default options.

For differential enrichment gene-based analysis, gene regions were extracted from the Xenbase *X. laevis* 9.2 annotation, using the GenomicFeatures package (Lawrence et al. 2013) from Bioconductor. Raw Bowtie2 aligned reads were used to generate gene level counts, using the options `allowMultiOverlap = TRUE`, `ignoreDup = FALSE`, `countMultimappingReads = FALSE`, `minMQS = 1`. Afterwards, DESeq2 was used to compare groups (Love et al. 2014). Differentially enriched genes in the IPs versus background were selected using a Benjamini-Hochberg adjusted p-value  $< 0.05$  and  $FC > 2$ . Adjustment for artifactual log2FC due to low count genes was performed with the `lfcShrink` function from DESeq2. Differential enrichment between CPEB1 and the CPEB2-4 group was performed with DESeq2. Differentially enriched genes were selected based on a `lfcShrink log2FC > 2` and Benjamini-Hochberg adjusted p-value  $< 0.05$ .

IP versus Input and CPEB1 versus CPEB2-4 GSEAs were performed with ROAST (D. Wu et al. 2010) within the limma package with GO and KEGG gene sets (Carbon et al. 2019, Kanehisa 2019).

### 3'UTR features analysis

Analysis of 3'UTR features in targets and non-targets was performed using the most abundant 3'UTR per transcript from Yang et al.'s dataset (F. Yang et al. 2020). PAS hex distances to the 3' end and 3'UTR length in targets of any CPEB versus non-targets were quantified using custom Perl scripts. 3'UTR architectures predictive of CPEB regulation were determined with the script published by Piqué et al. (Piqué et al. 2008) with some modifications: the maximum distance between the PAS hex and the 3' end was relaxed to 60 nucleotides and non-canonical PAS hex definitions were incorporated (Beaudoing et al. 2000).

A scan for *de novo* motifs in CPEB1-4 targets versus non-targets or CPEB1 targets versus CPEB2-4 and CPEB2-4 versus CPEB1 was performed with the `findMotifs` function from the Homer software (Benner et al. 2017) with motif lengths of 8, 10 and 12 and with the MEME Suite software (Bailey et al. 2009) specifying motif lengths between 6 and 15 nucleotides and default parameters otherwise. The motif's position probability matrices were drawn with the `ggseqlogo` R package.

## **Parsing, exploratory analyses and visualizations**

Exploratory data analysis, parsing and visualizations were done with R software using packages dplyr, ggplot2 and ComplexHeatmaps (Gu et al. 2016, Wickham 2016, Wickham et al. 2018).

## **Protein overexpression and purification from *E. coli***

### **Protein expression induction**

*E. coli* BL21 (DE3) competent cells were transformed with pET30a containing the proteins of interest. A transformed clone was grown overnight at 225 rpm 37°C in 200 mL Luria Broth (LB) with 50 µg/mL Kanamycin. The next morning the pre-cultures were diluted in 1 L fresh LB with Kanamycin to reach a starting OD<sub>600</sub> of 0.15 – 0.25 (usually requiring a 1/10-1/20 dilution). The cultures were grown at 225 rpm 37°C until OD<sub>600</sub> 0.4-0.6 was reached, at which point, protein expression was induced with 0.5 mM IPTG for 2 or 4 hours at 30°C. The cultures were spun 20 minutes at 6000 x g 4°C; the supernatants were discarded and the pellets were stored at -20°C.

Inductions were confirmed by SDS-PAGE followed by Coomassie and WB.

### **Bacterial cells lysis**

Upon confirmation of expression induction, the pellets were lysed in 40 mL of cold Protein Purification lysis buffer. The pellets were lysed by pipetting and sonication at 30% amplitude in pulses of 1 second on, 2 seconds off for as long as 3 minutes. The lysates were spun 10 minutes at 16000 x g 4°C and the resulting supernatants were used for protein purification.

### **Affinity purification**

All purification steps were carried out at 4°C. The His-tagged proteins were purified from the clarified lysates with 4 mL Ni-NTA Agarose beads (Qiagen) (4 mL of bead-slurry twice washed with lysis buffer). Upon incubation for 1 hour on the wheel the protein-bound beads were subjected to three 5-minute washes in batch in 10 mL Protein Purification wash buffer. The beads were packed in columns and the bound proteins

were eluted with three 1 mL incubations in Protein Purification elution buffer (two of 10 minutes and one overnight).

The resulting eluted fractions were dialyzed in Protein Purification dialysis buffer for 2 hours with gentle shaking (Spectra/Por Regenerated Cellulose Dialysis Membranes, 8,000 MWCO, Spectrum Europe). Last, 0.01% NaN<sub>3</sub> was added to the dialyzed purified proteins and these were stored in small aliquots at -80°C.

## ***In vitro* phosphorylation assays with oocyte lysates**

8 μL of clarified oocyte lysate were incubated with a 4 μL-premix containing 200 ng of recombinant protein [6xHis-CPEB2-Nter, 6xHis-CPEB2-20DE-Nter, 6xHis-CPEB3-Nter, 6xHis-CPEB3-18DE-Nter or Histone H1 (H5505, Sigma-Aldrich)], cold ATP (0.05 mM in the 12 μL reaction volume) and 2 μCi [γ-<sup>32</sup>P]ATP (Perkin-Elmer) in 1x H1K 0.4% NP40 lysis solution for 15 minutes at RT.

The reactions were stopped by addition of 8 μL of Laemmli sample buffer, incubated 5 minutes at 95°C and subsequently stored at -20°C or loaded on 8% polyacrylamide gels for PAGE.

Following PAGE, gels were stained with Coomassie Brilliant Blue [0.1% w/v Coomassie Brilliant Blue (Bio-Rad), 50% methanol, 20% acetic acid] 1 hour to overnight, destained (40% methanol, 10% acetic acid) 2 hours to overnight and rehydrated for 2 hours. The gels were analyzed by autoradiography (Amersham Hyperfilms MP, GE Healthcare; BAS-MS Imaging Plates, Fujifilm) and then dried at 80°C for 1 hour 30 minutes in a vacuum gel dryer system (SGD2000, Savant) for long-term storage.

## ***In vitro* phosphorylation assays with oocyte lysates and kinase inhibitors**

These assays were performed as described in the previous section with an added incubation step and a few modifications in the reaction volumes. 12 μL of clarified oocyte lysates were incubated for 30 minutes at 4°C with the corresponding inhibitors [Roscovitine (557360, Merck), U0126 (V112A, Promega) or FR180204 (328007, Merck)] at the indicated concentrations (relative to the final reaction volume) in a final volume of

16  $\mu\text{L}$ . 15  $\mu\text{L}$  of pretreated lysate were incubated with 6.4  $\mu\text{L}$  of premix containing 200 ng of recombinant protein, cold ATP (0.05 mM in the 12  $\mu\text{L}$  reaction volume) and 2  $\mu\text{Ci}$  [ $\text{Y}$ - $^{32}\text{P}$ ]ATP (Perkin-Elmer) in 1x H1K 0.4% NP40 lysis solution for 15 minutes at RT.

## ***In vitro* phosphorylation assays with recombinant ERK2 and Cdk1:CycB**

Phosphorylation by recombinant ERK2 (P6080, NEB) of 300 ng of recombinant protein substrate (6xHis-CPEB2-Nter, 6xHis-CPEB2-20DE-Nter, 6xHis-CPEB3-Nter and 6xHis-CPEB3-18DE-Nter) was assayed in a 30  $\mu\text{L}$  reaction, with 50 units of enzyme, 1x PK buffer (from NEB: 50 mM Tris-HCl pH 7.5, 10 mM  $\text{MgCl}_2$ , 1 mM EGTA, 2 mM DTT, 0.01% Brij 35), 0.1 mM cold ATP and 5  $\mu\text{Ci}$  [ $\text{Y}$ - $^{32}\text{P}$ ]ATP (Perkin-Elmer), for 30 minutes at 30°C.

Phosphorylation by recombinant Cdk1:CycB (PV3292, Thermo Fisher Scientific) of 300 ng of recombinant protein substrate [6xHis-CPEB2-Nter, 6xHis-CPEB2-20DE-Nter, 6xHis-CPEB3-Nter, 6xHis-CPEB3-18DE-Nter and Histone H1 (H5505, Sigma-Aldrich)] was assayed in a 30  $\mu\text{L}$  reaction, with 100 ng of enzyme, 20 mM Tris-HCl pH 7.5, 5 mM  $\text{MgCl}_2$ , 0.1 mM cold ATP and 5  $\mu\text{Ci}$  [ $\text{Y}$ - $^{32}\text{P}$ ]ATP (Perkin-Elmer), for 30 minutes at 30°C.

The reactions were stopped and endpoints visualized by PAGE followed by autoradiography and Commassie staining, exactly as described in the '*In vitro* phosphorylation assays with oocyte lysates' section (page 116).

## **Cell culture and transfection**

### **Cell culture**

U-2 OS cells were grown in Dulbecco Modified Eagle Medium (DMEM) (109380025, Thermo Fisher Scientific) supplemented with 10% Fetal Bovine Serum (FBS, 1027106, Thermo Fisher Scientific), 100 U/mL Penicillin-Streptomycin (15140122, Thermo Fisher Scientific) and 2 mM L-glutamine (25030024, Thermo Fisher Scientific). Mycoplasma tests were performed regularly every two weeks.

## Transfection

24 hours prior to transfection, cells were seeded to achieve 70% confluence at transfection. The seeding was done on 6-well plates with 2 or 3 12 mm Ø poly-Lysine-coated glass coverslips (O111550, Marienfeld Superior) added per well if the transfections were to be used for fixed-cell imaging. For live-cell imaging, cells were seeded on  $\mu$ -Slide 8 Well ibiTreat plates (80826, Ibidi).

Transient transfection with the plasmids specified on **Table S20** was performed with Lipofectamine LTX with Plus Reagent (15338030, Thermo Fisher Scientific) following the manufacturer's protocol. For a 6-well plate, per well, 3  $\mu$ g of DNA in combination with 3  $\mu$ L of Plus Reagent and 7  $\mu$ L Lipofectamine LTX were used. The transfections for live-cell imaging were scaled-down proportionally to the growth surface area. Protein extraction or fixation and mounting for imaging were performed 24 hours post-transfection.

## Protein extraction

Transfected cells were washed twice in PBS and gently scraped off the wells in 1 mL cold PBS. Scraped cells were pelleted by centrifugation (5 minutes at 0.5 g and 4°C), the PBS removed and lysed by pipetting in 50  $\mu$ L U-2 OS lysis buffer. The lysates were further disrupted by sonication (M2 setting, Bioruptor XL, Diagenode), centrifuged 5 minutes at 16000 g 4°C and the supernatants recovered. Clarified lysates were quantified against a BSA standard curve using DC Protein Assay kit (500-0116, BioRad) following manufacturer's instructions.

# CPEB2 and CPEB3 distribution in transfected U-2 OS cells

## Cell fixation and mounting

Transfected U-2 OS cells were washed twice with 3 mL of PBS and fixed with 4% paraformaldehyde in PBS for 10 minutes at RT. Two more washes with PBS were performed to remove the fixative agent and the coverslips were transferred to a clean, flat, covered surface where they were incubated for 10 minutes with 0.5  $\mu$ g/ $\mu$ L DAPI (D9542, Sigma-Aldrich). Excess DAPI was washed with PBS. The coverslips were dried, mounted onto glass slides (76 x 26 mm, Thermo Fisher Scientific) with Prolong Gold Antifade mounting medium (P36934, Invitrogen), left to dry overnight and sealed with

nail polish the following morning.

### **Image acquisition**

Confocal Z-series stacks were acquired at 1024 x 1024 pixels on a Leica SP5 confocal (Leica Microsystems) using the 63x 1.4NA oil immersion objective and 3.3 zoom factor. The Argon 488 nm (set to 20%) and Diode 405 nm (set to 9%) laser lines were used with spectral detection adjusted for the emission of GFP (PMT2 detector set at 500-550 nm, 20% gain) and DAPI (PMT2 detector set at 415-480 nm, 21% gain), respectively. Stacks were acquired with a fixed z-step size of 0.2  $\mu\text{m}$ . The multiposition setting (Mark and Find) was used for acquisition.

### **Analysis**

Max intensity projections (MIPs) of all images were obtained with ImageJ (version: 2.0.0-rc-69/1.52n). The MIPs were manually curated (edited when possible or discarded) to assure there was only one cell per image. The MIPs were classified as High, Medium and Low intensity, with mean intensity thresholds from 101 to 255, 71 to 100 and 24 to 70, respectively. The images were then blindly classified, manually, between either aggregated or diffuse. The condition associated to each image was added a posteriori.

## **CPEB2 and CPEB3 co-localization with CPEB1**

### **Image acquisition**

Images were acquired as described in the previous section with the additional channel DPSS 561 nm (set to 12%) laser line adjusted for the emission of mCherry (PMT4 set at 578-565, 10% gain). Only cells with an aggregated pattern were captured.

### **Analysis**

Channels were split and thresholded and the Colocalization test from ImageJ was applied. For each image, the observed Pearson correlation coefficient and the average correlation coefficient of 100 randomized images were obtained. Comparisons between sample and randomization pairs were performed with a Wilcoxon signed-rank test with Bonferroni correction for multiple testing.

# Live cell imaging and FRAP

## Sample preparation

The media of live transfected cells was changed to FluoroBrite DMEM (A189701, Thermo Fisher Scientific) supplemented with 10% FBS, L-glutamine and Penicillin-Streptomycin.

## Image acquisition

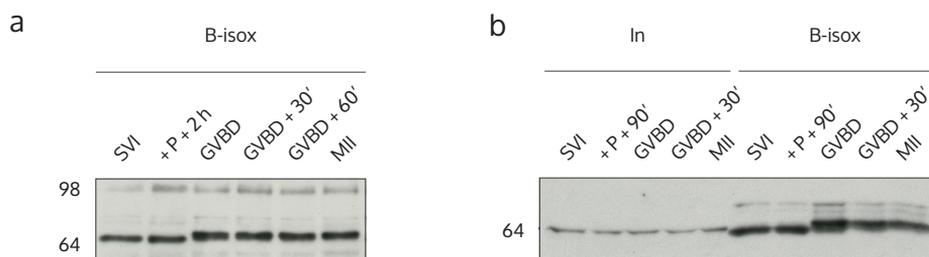
Live cell imaging and FRAP of cells expressing GFP-fusion-constructs were performed on Andor Revolution XD Spinning Disk Microscope (Andor Technology) equipped with a FRAPPA Photobleaching and Photoactivation module (Andor Technology). Per experimental unit, 350-512 x 512 pixel images were taken, 50 pre-bleaching and 300 post-bleaching, with a typical frame rate of 88 ms (11 images per second) and an exposure time of 50 ms in an iXon EMCCD camera (Andor Technology). The laser line used was AOTF 488 nm; it was set to 12% intensity for imaging and to 60% in two repeats with 40 ms dwell time for bleaching.

## Analysis

Per video, three regions of interest (ROIs) were defined: cell, background and bleaching area. Mean fluorescence intensity of these three ROIs throughout the 350 frames of the video were obtained and outputted in tabular format with ImageJ (version: 2.0.0-rc-69/1.52n). The tables were then entered to easyFRAP-web (Koulouras et al. 2018), where the “initial values to discard” variable was set to 20 (that is, 20 pre-bleach values were discarded) and double normalization was chosen. All the normalized curves were reviewed, and those that had abnormalities (e.g. if the bleached droplet disappeared halfway through the recovery time) were discarded, whereas some could be recovered by redefining one or more ROIs. The curves were fitted to single exponential, and half time of recovery ( $t_{\text{half}}$ ) and percentage of mobile were obtained. Comparison between group distributions was done using a Kluskal-Wallis test (0.05 significance) and post-hoc Dunn’s test with Bonferroni correction for multiple testing.

# Appendices





**Figure S1. CPEB3 expression time-courses.** Representative CPEB3 immunoblots of b-isox precipitated lysates. In: input; P: progesterone. MW, in kDa, indicated on the left side of the film scans.

**Table S1. CPEB3 p-site mapping by MS full results.** Number of PSM for each position in CPEB3 assigned with > 75% probability in which the modified/total ratio is greater in MII than SVI.

Res.	Modified		Unmodified		Total		Modified/Total		Ratio error	
	SVI	MII	SVI	MII	SVI	MII	SVI	MII	SVI	MII
S-021	6	35	168	44	174	79	0.03	0.44	0.01	0.01
S-023	1	3	173	76	174	79	0.01	0.04	0.01	0.01
S-078	4	9	33	4	37	13	0.11	0.69	0.02	0.06
S-145	63	106	174	34	237	140	0.27	0.76	0.00	0.01
S-148	54	104	183	36	237	140	0.23	0.74	0.00	0.01
S-173	3	7	6	9	9	16	0.33	0.44	0.08	0.04
S-178	0	3	9	13	9	16	0.00	0.19	0.11	0.05
S-181	0	3	10	13	10	16	0.00	0.19	0.10	0.05
S-195	0	2	2	2	2	4	0.00	0.50	0.50	0.18
S-231	46	52	153	14	199	66	0.23	0.79	0.00	0.01
T-236	8	6	239	167	247	173	0.03	0.03	0.00	0.01
S-251	0	20	38	18	38	38	0.00	0.53	0.03	0.02
T-254	0	3	12	10	12	13	0.00	0.23	0.08	0.06
T-266	2	6	139	111	141	117	0.01	0.05	0.01	0.01
S-353	9	10	61	17	70	27	0.13	0.37	0.01	0.03
S-625	0	1	0	0	0	1	0.00	1.00	0.00	1.00

**Table S2. BirA-CPEB1 BioID full results.** Mean: mean iBAQ, in logarithmic scale; NC: BirA-only negative control; miss.: number of missing observations, out of 4; FC: fold-change between condition and NC.

Name	Accessions	Mean	NC mean	Miss.	NC miss.	FC
ATXN2	AOA1L8HQR6	4.93	NA	1	4	NA
CNOT1	AOA1L8GFC1	5.05	3.81	0	2	17.19
CNOT1	AOA1L8GLF8	4.92	NA	0	4	NA
CON	CON_PO2768-1	5.68	4.47	1	1	16.07
CPEB1	AOA1L8GT79; Q52KN7	6.73	5.71	0	0	10.35
CPSF2	AOA1L8FOH9; Q9W799; AOA1L8F9H7	5.75	4.25	0	3	NA
CPSF3	AOA1L8G578	5.11	NA	1	4	NA
CSDE1	AOA1L8HE42; Q505M1; AOA1L8H6C1	4.87	NA	0	4	NA
CSTF64	Q68F10; Q7ZYT8; Q91581	5.05	NA	1	4	NA
DDX6	AOA1L8FLL9; P54824	6.70	5.60	0	0	12.46
EIF4E	AOA1L8GUZ9; AOA1L8GY11	7.12	6.17	0	0	8.84
EIF4ENIF1	AOA1L8HQ92; Q6DE09	7.26	6.35	0	0	7.97
EPAB	Q98SP8; AOA1L8EKZ2	5.75	4.43	0	2	21.16
EPAB	AOA1L8ES55; Q6GR16; AOA1L8ES40	5.77	4.64	0	1	13.31
FIPI	AOAUS9; Q63ZL7; AOA1L8HLC5	5.35	NA	0	4	NA
HUR	AOA1L8HY91; AOA1L8HNYO; Q919O3; AOA1L8GMC6; AOA1L8GME3; AOA1L8G- MD7; AOA1L8GMC8; AOA1L8GME1; AOA1L- 8GFU7; AOA1L8GMI8; Q7SZT7; Q91584	5.15	NA	0	4	NA
LSM14	AOA1L8ESZ1; Q68F11	6.11	NA	0	4	NA
LSM14	AOA1L8GL86; AOA1L8GL44; AOA8M2	5.22	4.76	0	3	NA
LSM14	Q8AVJ2	5.99	NA	0	4	NA
MIO5B	AOA1L8FVU1; AOA1L8FR11; Q5U5D4; A4QNS7	4.94	NA	1	4	NA
MOV10	AOA1L8HEX9	5.85	4.46	0	2	24.34
MOV10	AOA1L8GZO4	5.07	4.18	0	3	NA
PATL2	Q4V7K4; AOA1L8HCH4	6.84	5.38	0	0	28.94
SLBP2	B7ZQW3; AOA1L8GUY9; AOA1L8GUW5; Q9YGP6	5.92	NA	1	4	NA
STAU2	Q2VPF5; Q6DD72; Q7ZWR9; AOA1L8FT79	5.14	NA	0	4	NA
SYMPK	AOA1L8F3F9; Q7ZYV9	5.86	3.94	0	3	NA
TPR	AOA1L8GGX7; Q5EE04	6.17	5.14	0	0	10.63
TPR	Q6NU42; AOA1L8G560	5.19	NA	0	4	NA
WDR33	AOA1L8GAN7; AOA1L8G4F1	5.17	NA	0	4	NA
ZAR1	AOA1L8HBI7	5.85	NA	0	4	NA
ZAR1	AOA1L8HJK9	5.73	NA	0	4	NA
ZAR2	COSPG1	6.00	NA	0	4	NA

**Table S3. CPEB1-BirA BioID full results.**

Name	Accessions	Mean	NC mean	Miss.	NC miss.	FC
ATXN2	AOA1L8HQR6	5.07	NA	0	4	NA
ATXN2	AOA1L8I198; Q4V7W4	4.95	4.33	0	3	NA
CNOT1	AOA1L8GFC1	4.94	3.83	0	2	13.00
CNOT1	AOA1L8GLF8	5.00	NA	1	4	NA
CON	CON_P02768-1	5.88	4.51	2	1	23.81
CPEB1	AOA1L8GT79; Q52KN7	7.22	5.75	0	0	29.15
CPSF2	AOA1L8FOH9; Q9W799; AOA1L8F9H7	5.90	4.28	0	3	NA
CPSF3	AOA1L8G578	5.22	NA	0	4	NA
CPSF3	Q5XGZ1	4.90	NA	0	4	NA
CSDE1	AOA1L8HE42; Q505M1; AOA1L8H6C1	5.03	NA	1	4	NA
DAZL	B7ZS12; Q4V7Y4	5.55	NA	1	4	NA
DDX6	AOA1L8FLL9; P54824	6.85	5.64	0	0	16.28
EIF4E	AOA1L8GUZ9; AOA1L8GY11	7.32	6.21	0	0	12.96
EIF4ENIF1	AOA1L8HQ92; Q6DEO9	7.45	6.39	0	0	11.31
EPAB	Q98SP8; AOA1L8EKZ2	6.11	4.47	0	2	44.23
EPAB	AOA1L8ES55; Q6GR16; AOA1L8ES40	5.90	4.68	0	1	16.62
FIP1	AOAUS9; Q63ZL7; AOA1L8HLC5	5.56	NA	0	4	NA
HNRNPDLB	Q6NU14	5.26	NA	1	4	NA
IGF2BP3	AOA1L8FW10; O73932	5.44	NA	1	4	NA
IGF2BP3	O57526	5.24	NA	1	4	NA
LSM14	AOA1L8ESZ1; Q68F11	6.39	NA	0	4	NA
LSM14	AOA1L8GL86; AOA1L8GL44; AOA8M2	5.67	4.80	0	3	NA
LSM14	Q498K9	5.72	NA	1	4	NA
LSM14	Q8AVJ2	6.27	NA	0	4	NA
MOV10	AOA1L8HEX9	6.09	4.55	0	2	35.14
MOV10	AOA1L8GZO4	5.44	4.21	0	3	NA
PATL2	Q4V7K4; AOA1L8HCH4	7.00	5.43	0	0	37.88
PTBP1	Q7ZXB4; Q4QR55; Q9PTS5; AOA1L8HNA9; AOA1L8HX82	5.86	4.54	0	1	20.87
SLBP2	B7ZQW3; AOA1L8GUY9; AOA1L8GUW5; Q9YGP6	6.20	NA	1	4	NA
STAU2	Q2VPF5; Q6DD72; Q7ZWR9; AOA1L8FT79	5.55	NA	0	4	NA
SYMPK	AOA1L8F3F9; Q7ZYV9	6.02	3.95	0	3	NA
WDR33	AOA1L8GAN7; AOA1L8G4F1	5.14	NA	0	4	NA
YTHDF1	AOA1L8ELU2	5.70	NA	1	4	NA
YTHDF2	Q6DCK2; AOA1L8HE82; AOA1L8HEE7	5.23	NA	1	4	NA
ZAR1	AOA1L8HBI7	6.11	NA	0	4	NA
ZAR1	AOA1L8HJK9	5.96	NA	0	4	NA
ZAR2	COSPG1	6.23	NA	0	4	NA

**Table S4. BirA-CPEB2 BioID full results.**

Name	Accessions	Mean	NC mean	Miss.	NC miss.	FC
CNOT9	AOA1L8EVQ1; Q6IP65	5.46	NA	0	4	NA
EIF4ENIF1	AOA1L8HQ92; Q6DEO9	6.41	4.84	0	0	37.88
EPAB	AOA1L8ES55; Q6GR16; AOA1L8ES40	5.30	4.49	0	3	NA
EPAB	Q98SP8; AOA1L8EKZ2	5.37	NA	0	4	NA
G3BP1	Q6GQB6	4.84	NA	1	4	NA
LSM14	AOA1L8ESZ1; Q68F11	5.69	NA	1	4	NA
LSM14	Q8AVJ2	5.30	NA	0	4	NA
MOV10	AOA1L8HEX9	5.07	NA	0	4	NA

**Table S5. CPEB2-BirA BioID full results.**

Name	Accessions	Mean	NC mean	Miss.	NC miss.	FC
EML2	Q6NTK9; AOA1L8F893; AOA1L8F870	4.90	NA	1	4	NA
EPAB	AOA1L8ES55; Q6GR16; AOA1L8ES40	5.33	4.53	0	3	NA
EPAB	Q98SP8; AOA1L8EKZ2	5.33	NA	1	4	NA
HDAC1	AOA1L8HF18; Q91695; Q7ZYT5; O42227; Q7SYZ5; AOA1L8G2V6; Q66J55; AOA1L- 8G8Y0	5.07	NA	1	4	NA
HSPD1	Q6IP60; Q7ZTR6	5.45	NA	1	4	NA
LSM14	AOA1L8ESZ1; Q68F11	5.77	NA	1	4	NA
LSM14	Q8AVJ2	5.26	NA	0	4	NA
MOV10	AOA1L8HEX9	4.96	NA	0	4	NA
PFK	AOA1L8HIB5	5.07	NA	1	4	NA
POLD3	Q52KS2; Q6DDT8	5.02	4.81	0	3	NA
RCC2	B9VQ38; AOA1L8FPA6; Q52KW8	5.20	NA	1	4	NA
YWHAQ	Q7ZXH6	5.78	NA	1	4	NA

Table S6. BirA-CPEB3 BioID full results.

Name	Accessions	Mean	NC mean	Miss	NC miss.	FC
ABCF2	Q7ZWW5	5.04	NA	1	4	NA
CCTN	Q8UWC3; B7ZQN6; AOAI18GE11	4.98	NA	0	4	NA
CNOT1	AOAI18GLF8	5.10	NA	1	4	NA
CNOT2	AOAI18GYK2	5.45	NA	1	4	NA
CNOT9	AOAI18EVQ1; Q6IP65	5.69	NA	0	4	NA
DNJA1	Q8AVG6	5.04	NA	1	4	NA
EPAB	AOAI18ES55; Q6GR16; AOAI18ES40	5.34	4.58	0	3	NA
EPAB	Q98SP8; AOAI18EKZ2	5.43	NA	0	4	NA
G3BP1	Q6GQB6	5.14	NA	0	4	NA
HSPD1	Q6IP60; Q7ZTR6	5.18	NA	1	4	NA
LSM14	AOAI18ESZ1; Q68F11	6.11	NA	1	4	NA
LSM14	AOAI18GL86; AOAI18GL44; AOAI18M2	4.92	NA	1	4	NA
LSM14	Q8AVJ2	5.57	NA	0	4	NA
MOV10	AOAI18HEX9	5.05	NA	0	4	NA
RPA2A	AOAI18HE73; Q6IP18; AOAI18H7J6; AIL2H9	5.38	NA	1	4	NA
RPL3L	Q64OD7; AOAI18EYM2	5.35	NA	1	4	NA
SUCLG1	AOAI18HSF1; AOAI18HSG7	5.61	NA	1	4	NA
TPR	AOAI18GGX7; Q5EE04	6.06	4.40	0	1	45.81
UBQ-conjugat_E2	AOAI18EULO	4.96	NA	1	4	NA
WDR33	AOAI18GAN7; AOAI18G4F1	5.18	NA	0	4	NA
xCPEB3	BirA-xCPEB3	5.65	NA	0	4	NA

**Table S7. CPEB3-BirA BioID full results.**

Name	Accessions	Mean	NC mean	Miss	NC-miss.	FC
ABCF2	Q7ZWW5	4.94	NA	1	4	NA
AKAP1	AOA1L8H8J4	4.91	4.19	0	3	NA
ALG13	AOA1L8F7F0	5.10	4.12	0	1	9.58
CNOT2	AOA1L8GYK2	4.99	NA	0	4	NA
CNOT9	AOA1L8EVQ1; Q6IP65	5.68	NA	0	4	NA
DNJA1	Q8AVG6	4.96	NA	1	4	NA
DUSP12	Q0IH89; AOA1L8G7J0	5.24	NA	1	4	NA
EIF4ENIF1	AOA1L8HQ92; Q6DE09	7.09	4.90	0	0	152.64
EML2	Q6NTK9; AOA1L8F893; AOA1L8F870	4.97	NA	1	4	NA
EPAB	AOA1L8ES55; Q6GR16; AOA1L8ES40	5.97	4.55	0	3	NA
EPAB	Q98SP8; AOA1L8EKZ2	5.93	NA	0	4	NA
G3BP1	Q6GQB6	5.02	NA	0	4	NA
GRSF1	AOA1L8EVQ5; BIHIQ2; AOA1L8ENU2	5.16	NA	1	4	NA
HNRNPAB	Q6GM69; Q98UD3; Q7ZYE9	4.94	NA	1	4	NA
HSPD1	Q6IP60; Q7ZTR6	5.35	NA	1	4	NA
LSM14	AOA1L8ESZ1; Q68F11	6.35	NA	1	4	NA
LSM14	AOA1L8GL86; AOA1L8GL44; AOA8M2	5.82	NA	1	4	NA
LSM14	Q8AVJ2	6.09	NA	0	4	NA
MOV10	AOA1L8HEX9	5.52	NA	0	4	NA
PATL2	Q4V7K4; AOA1L8HCH4	6.47	4.92	0	1	35.47
PRKRA	AOA1L8EWC9; Q7ZYA5; Q91836	5.60	NA	0	4	NA
PTBP1	Q7ZXB4; Q4QR55; Q9PTS5; AOA1L8HNA9; AOA1L8HX82	5.03	4.46	0	3	NA
RAB2A	AOA1L8FZ34; Q2VPM6; AOA1L8FT25	5.07	NA	1	4	NA
SLBP2	B7ZQW3; AOA1L8GUY9; AOA1L8GUW5; Q9YGP6	5.99	NA	1	4	NA
SYMPK	AOA1L8F3F9; Q7ZYV9	5.16	NA	0	4	NA
TARDBP	AOA1L8FKV7; Q7ZXS6; AOA1L8FFBI; Q8JJ42	4.92	NA	0	4	NA
WDR33	AOA1L8GAN7; AOA1L8G4F1	5.11	NA	0	4	NA
xCPEB3	xCPEB3-BirA	5.18	NA	1	4	NA
ZAR1	AOA1L8HBI7	5.90	4.57	0	3	NA
ZAR2	COSPG1	6.23	4.51	0	2	52.81

**Table S8. BirA-CPEB4 BioID full results.**

Name	Accessions	Mean	NC-mean	Miss	NC miss.	FC
ATXN2	AOA1L8HQR6	5.12	NA	1	4	NA
ATXN2	AOA1L8I198; Q4V7W4	4.97	NA	1	4	NA
CNOT10	AOA1L8FWB8; Q6DE97	4.94	NA	1	4	NA
CNOT2	AOA1L8GYK2	5.41	NA	1	4	NA
CNOT9	AOA1L8EVQ1; Q6IP65	5.97	NA	1	4	NA
CPEB1	AOA1L8HO47; Q9I572	5.54	4.68	1	2	7.20
DDX6	AOA1L8FLL9; P54824	6.44	4.88	1	0	36.49
DYNC1	AOA1L8FO32	4.27	3.39	1	0	7.66
EPAB	Q98SP8; AOA1L8EKZ2	6.05	NA	1	4	NA
G3BP1	Q6GQB6	5.21	NA	1	4	NA
HNRNPAB	Q6GM69; Q98UD3; Q7ZYE9	4.93	NA	1	4	NA
LSM14	AOA1L8ESZ1; Q68F11	6.29	NA	1	4	NA
LSM14	Q8AVJ2	5.86	NA	1	4	NA
MACROD1	AOA1L8GJT9	5.71	NA	1	4	NA
MOV10	AOA1L8HEX9	5.29	NA	0	4	NA
PUM1	AOA1L8HE63; AOA1L8HE74; AOA1L8H7I9; Q66K16	4.99	NA	1	4	NA
RPA2A	AOA1L8HE73; Q6IP18; AOA1L8H7J6; A1L2H9	5.26	NA	1	4	NA
SYMPK	AOA1L8F3F9; Q7ZYV9	5.12	NA	1	4	NA
TAGLN2	AOA1L8FCQ5; Q7ZWS8	5.47	NA	1	4	NA
WDR33	AOA1L8GAN7; AOA1L8G4F1	5.03	NA	1	4	NA
ZAR1	AOA1L8HJK9	5.97	4.50	1	2	29.64
ZAR2	COSPG1	6.27	4.54	1	2	54.43

Table S9. CPEB4-BirA BioID full results.

Name	Accessions	Mean	NC mean	Miss.	NC miss.	FC
AKAP1	AOA1L8H8J4	5.36	4.15	0	3	NA
ALG13	AOA1L8F7F0	5.50	4.07	0	1	26.35
ATXN2	AOA1L8HQR6	4.93	NA	0	4	NA
ATXN2	AOA1L8I198; Q4V7W4	5.12	NA	0	4	NA
CAPRN2	AOA1L8GWA6; B1WB5	6.63	5.75	0	0	7.69
CEP85	AOA1L8H7B4	5.25	4.56	0	3	NA
CNOT1	AOA1L8GFC1	5.00	NA	0	4	NA
CNOT1	AOA1L8GLF8	5.14	NA	0	4	NA
CNOT10	AOA1L8FWB8; Q6DE97	5.16	NA	0	4	NA
CNOT2	AOA1L8GUI7; Q6GPN8	5.43	NA	0	4	NA
CNOT2	AOA1L8GYK2	5.57	NA	0	4	NA
CNOT9	AOA1L8EVQ1; Q6IP65	6.11	NA	0	4	NA
CPEB1	AOA1L8GT79; Q52KN7	7.49	5.74	0	0	57.26
CPEB1	AOA1L8HO47; Q91572	5.95	4.68	0	2	18.40
CPSF4	AOA1L8EM14	5.19	NA	1	4	NA
CSDE1	AOA1L8HE42; Q505M1; AOA1L8H6C1	5.07	NA	0	4	NA
DDX3X	Q52L23; Q7ZXJ0; P24346	5.33	4.65	0	3	NA
DDX6	AOA1L8FLL9; P54824	6.81	4.81	0	0	99.83
EIF4E	AOA1L8GUZ9; AOA1L8GY11	7.09	5.75	0	0	28.17
EIF4ENIF1	AOA1L8HQ92; Q6DE09	7.44	4.85	0	0	383.27
EIF4G1	AOA1L8G9T6	5.91	4.76	0	0	13.86
EML2	Q6NTK9; AOA1L8F893; AOA1L8F870	4.92	NA	0	4	NA
EML4	AOA1L8G192; AOA1L8G1A2; Q2TAF3; AOA1L8G6E1	4.91	3.81	1	0	12.48
EPAB	AOA1L8ES55; Q6GR16; AOA1L8ES40	6.44	4.50	0	3	NA
EPAB	Q98SP8; AOA1L8EKZ2	6.30	NA	0	4	NA
ESRP1	AOA1L8FTH7; AOA1L8FZJ1; Q7ZY29	5.93	NA	0	4	NA
FUBP3	Q6PAAO; AOA1L8F1K5	5.34	NA	1	4	NA
G3BP1	Q6GQB6	5.55	NA	0	4	NA
GRSF1	AOA1L8EVQ5; BIHIQ2; AOA1L8ENU2	5.34	NA	0	4	NA
HNRNPAB	Q6GM69; Q98UD3; Q7ZYE9	5.38	NA	0	4	NA
HSPD1	Q6IP60; Q7ZTR6	4.99	NA	1	4	NA
HUR	AOA1L8HX08; AOA1L8HWX9; Q1JQ73	5.10	NA	0	4	NA
HUR	Q5U259; AOA1L8HNJ9	5.65	4.52	0	3	NA
LSM14	AOA1L8ESZ1; Q68F11	6.47	NA	0	4	NA
LSM14	AOA1L8GL86; AOA1L8GL44; AOA8M2	6.10	NA	0	4	NA
LSM14	Q8AVJ2	6.42	NA	0	4	NA
MOV10	AOA1L8GZO4	5.43	4.36	0	2	11.89
MOV10	AOA1L8HEX9	6.01	NA	0	4	NA

Name	Accessions	Mean	NC mean	Miss.	NC miss.	FC
PABPN	B7ZRW2; B7ZRW4; Q6TY21	5.25	NA	0	4	NA
PABPN	Q804A5	5.98	NA	0	4	NA
PATL2	Q4V7K4; AOAIL8HCH4	6.86	4.85	0	1	101.52
PIWIL1	AOAIL8HK01	5.40	3.61	0	3	NA
PRKRA	AOAIL8EWC9; Q7ZYA5; Q91836	5.49	NA	0	4	NA
PRRC2	AOAIL8GN33	5.03	NA	0	4	NA
PTBP1	Q7ZXB4; Q4QR55; Q9PTS5; AOAIL8HNA9; AOAIL8HX82	5.51	4.41	0	3	NA
PUM1	AOAIL8HE63; AOAIL8HE74; AOAIL8H7I9; Q66KI6	5.13	NA	0	4	NA
RBPM2	AOAIL8GT66; Q9YGP5; Q66IX3	5.86	NA	1	4	NA
RPL3L	Q64OD7; AOAIL8EYM2	6.45	NA	1	4	NA
SLBP2	B7ZQW3; AOAIL8GUY9; AOAIL8GUW5; Q9YGP6	6.07	NA	0	4	NA
STAU2	Q2VPF5; Q6DD72; Q7ZWR9; AOAIL8FT79	5.49	NA	0	4	NA
SUCLG1	AOAIL8HSF1; AOAIL8HSG7	5.41	NA	1	4	NA
SYMPK	AOAIL8F3F9; Q7ZYV9	5.56	NA	0	4	NA
TARDBP	AOAIL8FKV7; Q7ZXS6; AOAIL8FFB1; Q8JJ42	5.19	NA	0	4	NA
TPR	AOAIL8GGX7; Q5EE04	5.52	4.33	0	1	15.52
WDR33	AOAIL8GAN7; AOAIL8G4F1	5.53	NA	0	4	NA
YTHDF1	AOAIL8ELU2	6.09	NA	1	4	NA
YTHDF1	Q6PA59; AOAIL8ET20	5.45	NA	0	4	NA
YTHDF2	Q6DCK2; AOAIL8HE82; AOAIL8HEE7	5.19	NA	0	4	NA
ZAR1	AOAIL8HJK9	6.46	4.45	0	2	102.51
ZAR1	AOAIL8HBI7	6.40	4.52	0	3	NA
ZAR1	B7ZPGO; AOAIL8HL44; Q7T3UO; M9VOQ8	5.28	NA	1	4	NA
ZAR2	COSPG1	6.73	4.50	0	2	171.33
ZC3H7B	Q6DCZ5; AOAIL8GNW3	5.24	NA	1	4	NA





**Table S15. CPEB2 mutagenesis primers.** Pos.: position; Res.: residue; Subs.: substitution.

ID	Pos.	Res.	Subs.	Sequence
BD164	87	S	D	GAGCGCCGAGTC <sub>gac</sub> CCGCCTGCCCTGC
BD165	87	S	A	GAGCGCCGAGTC <sub>gcc</sub> CCGCCTGCCCTGC
BD166	104	S	D	GAGGAAGGACTTC <sub>gac</sub> CCCCAGGAGG
BD167	104	S	A	GAGGAAGGACTTC <sub>gcc</sub> CCCCAGGAGG
BD168	137	S	D	CCAGCGACACGGA <sub>gac</sub> CCGGCGCAGAGG
BD169	137	S	A	CCAGCGACACGGA <sub>gcc</sub> CCGGCGCAGAGG
BD170	145	S	D	GCAGAGGAGCTGGAG <sub>gat</sub> CCGGACAAGAATC
BD171	145	S	A	GCAGAGGAGCTGGAG <sub>gct</sub> CCGGACAAGAATC
BD172	153	S	D	GAATCTGCCAGT <sub>gac</sub> CCTCCTCCTCG
BD173	153	S	A	GAATCTGCCAGT <sub>gcc</sub> CCTCCTCCTCG
BD174	179	S	D	GCTCATGCCGCAAC <sub>gac</sub> CCTCCTGCTGTGC
BD175	179	S	A	GCTCATGCCGCAAC <sub>gcc</sub> CCTCCTGCTGTGC
BD176	214	S	D	CCTTACCCCGA <sub>gac</sub> CCAGAATTAGCC
BD177	214	S	A	CCTTACCCCGA <sub>gcc</sub> CCAGAATTAGCC
BD178	224	S	D	CCCCGGGGGA <sub>gac</sub> CCCCCTGCCCTG
BD179	224	S	A	CCCCGGGGGA <sub>gcc</sub> CCCCCTGCCCTG
BD180	233	T	E	CCGGCTTCGGT <sub>gaa</sub> CCCTGGTCTGTAC
BD181	233	T	A	GGGCTTCGGT <sub>gcc</sub> CCCTGGTCTGTAC
BD182	241	S	D	TCTGTACAGACTTC <sub>gat</sub> CCTCCCCGCCG
BD183	241	S	A	TCTGTACAGACTTC <sub>gct</sub> CCTCCCCGCCG
BD184	276	S	D	CAACGCCCTGCAT <sub>gac</sub> CCCGACCCCGAC
BD185	276	S	A	CAACGCCCTGCAT <sub>gcc</sub> CCCGACCCCGAC
BD186	313	S	D	CCGGATTCAGC <sub>gac</sub> CCCTTCTGTCTCAG
BD187	313	S	A	CCGGATTCAGC <sub>gcc</sub> CCCTTCTGTCTCAG
BD188	331, 334	S, S	D, D	GCAGAGCCGGAG <sub>Agac</sub> CCTGT <sub>Agat</sub> CCCCAGCTCAACCC
BD189	331, 334	S, S	A, A	GCAGAGCCGGAG <sub>Agcc</sub> CCTGT <sub>Agct</sub> CCCCAGCTCAACCC
BD200	362	K	R	CCAGCCACTTGT <sub>Taga</sub> CAGTCTCCATGGGG
BD190	364	S	D	CTTGTTAAACAG <sub>gat</sub> CCATGGGGCGGC
BD191	364	S	A	CTTGTTAAACAG <sub>gct</sub> CCATGGGGCGGC
BD192	405	S	D	CCATGAACCAAAT <sub>Cgac</sub> CCACTAAAGAAGG
BD193	405	S	A	CCATGAACCAAAT <sub>Cgcc</sub> CCACTAAAGAAGG
BD194	425, 429	T, T	E, E	CCAAGTTCACTCGTTCC <sub>gaa</sub> CCATCGCT <sub>Ggaa</sub> CCCAAATCTTGG
BD195	425, 429	T, T	A, A	CCAAGTTCACTCGTTCC <sub>gct</sub> CCATCGCT <sub>Ggct</sub> CCCAAATCTTGG
BD196	525	S	D	GGTTCTGAAT <sub>gat</sub> CCCAATGCTATTCTGG
BD197	525	S	A	GGTTCTGAAT <sub>gct</sub> CCCAATGCTATTCTGG
BD198	619	S	D	CTGTGTGTTCA <sub>gac</sub> CCTACCATCAAGG
BD199	619	S	A	CTGTGTGTTCA <sub>gcc</sub> CCTACCATCAAGG
MC245	576	K	R	GTTGATTGCCCCA <sub>Caga</sub> GCAGAAAGCAAATCC
MC246	623	K	R	GTTTCAAGCCCTACCAT <sub>Cagg</sub> GACAAGCCAGTTCAG
MC247	729	K	R	GATAAAAGGGTGGAGGT <sub>Gagg</sub> CCATACGTGTTGGATG
MC248	786	K	R	CACAAGCCTCTGGT <sub>Gagg</sub> GAAGGAGCTGACCGG

**Table S16. CPEB3 mutagenesis primers.** Pos.: position; Res.: residue; Subs.: substitution.

ID	Pos.	Res.	Subs.	Sequence
BD153	21	S	A	GCGACATCAGCAG <sub>gct</sub> CCATCCTCC
BD155	21	S	D	GCGACATCAGCAG <sub>gat</sub> CCATCCTCC
BD095	47	S	D	GCAGCAGAGAGAGC <sub>gac</sub> CCCCCACTCACAGG
BD117	47	S	A	GCAGCAGAGAGAGC <sub>gca</sub> CCCCCACTCACAGG
BD096	59	S	D	GAAGATGCAGATGGAG <sub>gac</sub> CCCTTACTGCCTGG
BD118	59	S	A	GAAGATGCAGATGGAG <sub>gca</sub> CCCTTACTGCCTGG
BD159	74	T	E	CATGAGCCCCCACC <sub>gag</sub> CCATCTCTGGACCA
BD160	74	T	A	GAGCCCCCACC <sub>gcc</sub> CCATCTCTGG
BD097	78	S	D	CCACCCCATCTCTG <sub>gac</sub> CCATCTTTTGGCAGC
BD119	78	S	A	CCACCCCATCTCTG <sub>gca</sub> CCATCTTTTGGCAGC
BD098	103	T	E	GCTTCTTCCCAGGGATT <sub>gaa</sub> CCAGTGAATGGGACC
BD120	103	T	A	GCTTCTTCCCAGGGATT <sub>gca</sub> CCAGTGAATGGGACC
BD099	129	S	D	GGGGGCACCTTC <sub>gat</sub> CCCCAGCTGG
BD121	129	S	A	GGGGGCACCTTC <sub>gct</sub> CCCCAGCTGG
BD100	145, 148	S, S	D, D	GAGGCGA <sub>gac</sub> CCCGCC <sub>gac</sub> CCCAACAACC
BD122	145, 148	S, S	A	GAGGCGA <sub>gca</sub> CCCGCC <sub>gca</sub> CCCAACAACC
BD144	174	S	A	GCCCTCGTCC <sub>gcc</sub> CCCAACTCC
BD163	174	S	D	GCCCTCGTCC <sub>gat</sub> CCCAACTCC
BD145	181	S	A	CCTTGCCT <sub>gcc</sub> CCGTCCAAC
BD157	181	S	S	CCTTGCCT <sub>gac</sub> CCGTCCAAC
BD146	195	S	A	GCCTGGAAT <sub>gcc</sub> CCTTCTAACCC
BD158	195	S	D	GCCTGGAAT <sub>gat</sub> CCTTCTAACCC
BD101	225	S	D	GGTGTCGGGGTGCC <sub>gat</sub> CCGCTCAACC
BD123	225	S	A	GGTGTCGGGGTGCC <sub>gct</sub> CCGCTCAACC
BD102	231	S	D	CCATC <sub>gac</sub> CCAATGAAAAAACTTCTCCAGC
BD124	231	S	A	CCATC <sub>gcc</sub> CCAATGAAAAAACTTCTCCAGC
BD161	251, 254	S, T	D, E	GCACCCCTAAATTC <sub>CAAGAGCAgac</sub> CCATT <sub>Agag</sub> CCGAAATCGTGG-GTG
BD162	251, 254	S, T	A, A	CCCTAAATTC <sub>CAAGAGCAgca</sub> CCATT <sub>Agcc</sub> CCGAAATC
BD104	353	S	D	GGCCTGAGT <sub>gat</sub> CCAACGCATTGTCAGAACGG
BD126	353	S	A	GGCCTGAGT <sub>gct</sub> CCAACGCATTGTCAGAACGG
BD105	444	S	D	CCTGTGTGTGTC <sub>gac</sub> CCAACCATCAAGGATAAACC
BD127	444	S	A	CCTGTGTGTGTC <sub>gca</sub> CCAACCATCAAGGATAAACC

**Table S17. P-site mapping plasmid list. FL: full-length.**

<b>ID</b>	<b>Full name</b>	<b>Insert</b>
pBDO02	pBSK 9xHis-HA CPEB2	FL wt CDS
pBDO22	pBSK 9xHis-HA GFP-CPEB2	FL wt CDS
pBDO69	pBSK 9xHis-HA CPEB2 20DE	20 (S/T)P sites mutated to D/E: S87, S104, S137, S145, S153, S179, S214, S224, T233, S241, S276, S313, S331, S334, S364, S405, T425, T429, S525, S619
pBDO70	pBSK 9xHis-HA CPEB2 20A	20 (S/T)P sites mutated to A
pBDO71	pBSK 9xHis-HA CPEB2 K362R	K362R
pBD105	pBSK 9xHis-HA CPEB2 K576R	K576R
pBD106	pBSK 9xHis-HA CPEB2 K623R	K623R
pBD107	pBSK 9xHis-HA CPEB2 K729R	K729R
pBD108	pBSK 9xHis-HA CPEB2 K786R	K786R
pBDO03	pBSK 9xHis-HA CPEB3	FL wt CDS
pBDO61	pBSK 9xHis-HA CPEB3 13DE	11 (S/T)P sites and an SS motif mutated to D/E: S47, S59, S78, T103, S129, S145, S148, S225, S231, S328, S329, S353, S444
pBDO62	pBSK 9xHis-HA CPEB3 13A	Same 13 sites mutated to A
pBDO63	pBSK 9xHis-HA CPEB3 18DE	18 (S/T)P sites mutated to D/E: S21, S47, S59, T74, S78, T103, S129, S145, S148, S174, S181, S195, S225, S231, S251, S254, S353, S444
pBDO64	pBSK 9xHis-HA CPEB3 18A	Same 18 (S/T)P sites mutated to A
pBDO65	pBSK 9xHis-HA CPEB3 11DE	11 (S/T)P sites mutated to D/E, 7 sites less than the 18A/DE mutants, specifically: S21, S74, S174, S181, S195, S251, T254
pBDO66	pBSK 9xHis-HA CPEB3 11A	Same 11 (S/T)P sites mutated to A
pBDO67	pBSK 9xHis-HA CPEB3 13DE (18-5)	13 (S/T)P sites mutated to D/E, 5 sites less than the 18A/DE mutants, specifically: S47, S59, T103, S129, S195
pBDO68	pBSK 9xHis-HA CPEB3 13A (18-5)	Same 13 (S/T)P sites mutated to A

**Table S18. Recombinant protein production plasmid list.**

<b>ID</b>	<b>Full name</b>	<b>Insert</b>
pBDO04	pET30 6xHis CPEB2	FL wt CDS
pBDO09	pET30 6xHis CPEB2-Nter	N-ter: nucleotides 1-1587 (529 aminoacids)
pBD103	pET30 6xHis CPEB2-DE-Nter	N-ter with all (S/T)P sites mutated to D/E
pBDO05	pET30 6xHis CPEB3	FL wt CDS
pBDO10	pET30 6xHis CPEB3-Nter	N-ter: nucleotides 1-1056 (352 aminoacids)
pBD104	pET30 6xHis CPEB3-DE-Nter	N-ter with all (S/T)P sites mutated to D/E

**Table S19. BioID plasmid list.** MC: plasmid in Manuel Cañete's plasmid bank (Cañete 2020).

ID	Full name
Addgene 36047	pCDNA-3.1 MCS BirA(R118G)-HA
MC	pCDNA-3.1 CPEB1 BirA(R118G)-HA
pBD050	pCDNA-3.1 CPEB2 BirA(R118G)-HA
pBD051	pCDNA-3.1 CPEB3 BirA(R118G)-HA
pBD052	pCDNA-3.1 CPEB4 BirA(R118G)-HA
Addgene 35700	pcDNA-3.1 myc-BirA(R118G) MCS
MC	pcDNA-3.1 myc-BirA(R118G) CPEB1
MC	pcDNA-3.1 myc-BirA(R118G) CPEB2
MC	pcDNA-3.1 myc-BirA(R118G) CPEB3
MC	pcDNA-3.1 myc-BirA(R118G) CPEB4

**Table S20. U-2 OS transfections plasmid list.**

ID	Full name	Tags
pBD076	pPEU6 CPEB2	N-ter StrepII-eGFP
pBD077	pPEU6 CPEB2 20A	N-ter StrepII-eGFP
pBD078	pPEU6 CPEB2 20DE	N-ter StrepII-eGFP
pBD094	pPEU4 CPEB2	C-ter eGFP-6xHis
pBD095	pPEU4 CPEB2 20A	C-ter eGFP-6xHis
pBD096	pPEU4 CPEB2 20DE	C-ter eGFP-6xHis
pBD098	pPEU5 CPEB2	C-ter mCherry-6xHis
pBD073	pPEU6 CPEB3	N-ter StrepII-eGFP
pBD074	pPEU6 CPEB3 18A	N-ter StrepII-eGFP
pBD075	pPEU6 CPEB3 18DE	N-ter StrepII-eGFP
pBD091	pPEU4 CPEB3	C-ter eGFP-6xHis
pBD092	pPEU4 CPEB3 18A	C-ter eGFP-6xHis
pBD093	pPEU4 CPEB3 18DE	C-ter eGFP-6xHis
pBD097	pPEU5 CPEB3	C-ter mCherry-6xHis
MC	pPEU4 CPEB1	C-ter eGFP-6xHis
MC	pPEU5 CPEB1	C-ter mCherry-6xHis

**Table S21. Antibody list.**

<b>Epitope</b>	<b>Full name</b>	<b>Dilution</b>	<b>Reference</b>
HA	Rat anti-HA High Affinity [3F10]	1,000	11867423001, Merck
CPEB1	Rabbit anti-CPEB1 #1010	2,000	Custom antibody, CRG
CPEB2	Mouse anti-CPEB2	500	Custom antibody, Yi Shuiang Huang's lab
CPEB3	Rabbit anti-CPEB3	1,000	Ab10883, Abcam
DDX6	Rabbit anti-DDX6 (Xe5)	5,000	Custom antibody, Abbyntek
Flag	Mouse anti-FLAG [M2]	2,000	F1804, Sigma-Aldrich
GLD2	Rabbit anti-GLD2 (Xe7)	1,000	Custom antibody, Abbyntek
MOS	Rabbit anti-Mos	500	C237, Santa Cruz Biotechnology
MYC	Goat anti-Myc	3,000	Ab9132, Abcam
Sumo-1	Rabbit anti-Sumo-1 [Y299]	1,000	Ab32058, Abcam
Sumo-2/3	Mouse anti-Sumo-2/3 [1E7]	500	M114-3, MBL
Ubiquitin	Mouse anti-mono-and-poly-Ubiquitin [FK2]	1,000	PML-PW8810, Enzo
Biotin	Streptavidin, HRP conjugate	1,000	S911, Invitrogen
CNOT2	Rabbit anti-CNOT2	1,000	ABIN1106729, Antibodies-online
CPSF2	Rabbit anti-CPSF2 (Xe2)	1,000	Custom antibody, Abbyntek
eIF4E-1b	Rabbit anti-eIF4E-1b (Xe4)	1,000	Custom antibody, Abbyntek
eIF4E-T	Rabbit anti-eIF4E-T (Xe3)	1,000	Custom antibody, Abbyntek
HuR	Mouse anti-HuR [3A2]	2,000	SC-5261, Santa Cruz Biotechnology
Symplekin	Mouse anti-Symplekin [25/Symplekin]	1,000	6106644, BD Biosciences
His	Mouse anti-polyHis [HIS-1]	5,000	H1029, Sigma-Aldrich
Mouse	Goat anti-Mouse IgG (H+L) Secondary Antibody, HRP	3,000 - 5,000	31430, Invitrogen
Rabbit	Goat anti-Rabbit IgG (H+L) Secondary Antibody, HRP	3,000 - 5,000	G21234, Invitrogen
Rat	Goat anti-Rat IgG (H+L) Secondary Antibody, HRP	3,000 - 5,000	31741, Invitrogen
Actin	Mouse anti-beta-Actin	10,000	m8226, Abcam
GFP	Rabbit anti-GFP	2,000	A6455, Invitrogen
Tubulin	Mouse anti-tubulin [DM1A]	10,000	T9026, Sigma-Aldrich
Vinculin	Mouse anti-Vinculin [SPM227]	10,000	ab18058, Abcam
Vinculin	Mouse anti-Vinculin [VIN-54]	10,000	ab130007, Abcam

# References



- Alarcon, Hodgman, Theis, Huang, Kandel & Richter. (2004). Selective modulation of some forms of Schaffer collateral-CA1 synaptic plasticity in mice with a disruption of the CPEB-1 gene. *Learning and Memory*, 11(3), 318–327.
- Alberti, Gladfelter & Mittag. (2019). Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. In *Cell* (Vol. 176, Issue 3, pp. 419–434). Cell Press.
- Alberti, Saha, Woodruff, Franzmann, Wang & Hyman. (2018). A User's Guide for Phase Separation Assays with Purified Proteins. *Journal of Molecular Biology*, 430(23), 4806–4820.
- Alberti. (2017). Phase separation in biology. In *Current Biology* (Vol. 27, Issue 20, pp. R1097–R1102). Cell Press.
- Andrews. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Atkins, Nozaki, Shigeri & Soderling. (2004). Cytoplasmic polyadenylation element binding protein-dependent protein synthesis is regulated by calcium/calmodulin-dependent protein kinase II. *Journal of Neuroscience*, 24(22), 5193–5201.
- Bah & Forman-Kay. (2016). Modulation of Intrinsically Disordered Protein Function by Post-translational Modifications. *Journal of Biological Chemistry*, 291(13), 6696–6705.
- Bailey, Boden, Buske, Frith, Grant, Clementi, Ren, Li & Noble. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research*, 37(Web Server), W202–W208.
- Banani, Lee, Hyman & Rosen. (2017). Biomolecular condensates: Organizers of cellular biochemistry. In *Nature Reviews Molecular Cell Biology* (Vol. 18, Issue 5, pp. 285–298). Nature Publishing Group.
- Barnard, Ryan, Manley & Richter. (2004). Symplekin and xGLD-2 Are Required for CPEB-Mediated Cytoplasmic Polyadenylation. *Cell*, 119(5), 641–651.
- Beauclair, Bridier-Nahmias, Zagury, Saïb, Zamborlini, Säb & Zamborlini. (2015). JASSA: a comprehensive tool for prediction of SUMOylation sites and SIMs. *Bioinformatics*, 31(21), 3483–3491.
- Beaudoing, Freier, Wyatt, Claverie & Gautheret. (2000). Patterns of variant polyade-

- nylation signal usage in human genes. *Genome Research*, 10(7), 1001–1010.
- Belloc & Méndez. (2008). A deadenylation negative feedback mechanism governs meiotic metaphase arrest. *Nature*, 452(7190), 1017–1021.
- Benner, Heinz & Glass. (2017). HOMER - Software for motif discovery and next generation sequencing analysis. [Http://Homer.Ucsd.Edu/](http://Homer.Ucsd.Edu/).
- Berchowitz, Kabachinski, Walker, Carlile, Gilbert, Schwartz & Amon. (2015). Regulated Formation of an Amyloid-like Translational Repressor Governs Gametogenesis. *Cell*, 163(2), 406–418.
- Berger-Sweeney, Zearfoss & Richter. (2006). Reduced extinction of hippocampal-dependent memories in CPEB knockout mice. *Learning and Memory*, 13(1), 4–7.
- Boke, Ruer, Wühr, Coughlin, Lemaitre, Gygi, Alberti, Drechsel, Hyman & Mitchison. (2016). Amyloid-like Self-Assembly of a Cellular Compartment. *Cell*, 166(3), 637–650.
- Brangwynne, Eckmann, Courson, Rybarska, Hoegge, Gharakhani, Jülicher & Hyman. (2009). Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science*, 324(5935), 1729–1732.
- Brangwynne, Mitchison & Hyman. (2011). Active liquid-like behavior of nucleoli determines their size and shape in *Xenopus laevis* oocytes. *Proceedings of the National Academy of Sciences of the United States of America*, 108(11), 4334–4339.
- Brenner, Jacob & Meselson. (1961). An Unstable Intermediate Carrying Information from Genes to Ribosomes for Protein Synthesis. *Nature*, 190(4776), 576–581.
- Cañete. (2020). Composition, remodeling and dynamics of the CPEB RNP. Universitat de Barcelona.
- Carbon, Douglass, Dunn, Good, Harris, Lewis, Mungall, Basu, Chisholm, Dodson, Hartline, Fey, Thomas, Albou, Ebert, Kesling, Mi, Muruganujan, Huang, ... Westerfield. (2019). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research*, 47(D1), D330–D338.
- Carotenuto & Tussellino. (2018). *Xenopus laevis* oocyte as a model for the study of the cytoskeleton. In *Comptes Rendus - Biologies* (Vol. 341, Issue 4, pp. 219–227). Elsevier Masson SAS.

- Chaigne. (2014). Cortical stiffness : a gatekeeper for spindle positioning in mouse oocytes. Université Pierre et Marie Curie - Paris VI.
- Chang & Huang. (2014). Arsenite-Activated JNK Signaling Enhances CPEB4-Vinexin Interaction to Facilitate Stress Granule Assembly and Cell Survival. *PLoS ONE*, 9(9), e107961.
- Chang, Tung, Chen, Tu & Chu. (2018). SUMOgo: Prediction of sumoylation sites on lysines by motif screening models and the effects of various post-translational modifications. *Scientific Reports*, 8(1), 1–10.
- Chao, Tsai, Lu, Lin, Huang, Chou, Lu, Lin, Lee & Huang. (2013). Deletion of CPEB3 enhances hippocampus-dependent memory via increasing expressions of PSD95 and NMDA receptors. *Journal of Neuroscience*, 33(43), 17008–17022.
- Chen & Huang. (2012). CPEB2-eEF2 interaction impedes HIF-1-alpha RNA translation. *The EMBO Journal*, 31(4), 959–971.
- Chong, Vernon & Forman-Kay. (2018). RGG/RG Motif Regions in RNA Binding and Phase Separation. In *Journal of Molecular Biology* (Vol. 430, Issue 23, pp. 4650–4665). Academic Press.
- Christerson & McKearin. (1994). orb Is required for anteroposterior and dorsoventral patterning during Drosophila oogenesis. *Genes and Development*, 8(5), 614–628.
- Cobb. (2017). 60 years ago, Francis Crick changed the logic of biology. *PLOS Biology*, 15(9), e2003243.
- Colot & Rosbash. (1982). Behavior of individual maternal pA+ RNAs during embryogenesis of *Xenopus laevis*. *Developmental Biology*, 94(1), 79–86.
- Costa-Mattioli, Sossin, Klann & Sonenberg. (2009). Translational Control of Long-Lasting Synaptic Plasticity and Memory. In *Neuron* (Vol. 61, Issue 1, pp. 10–26). Neuron.
- Courel, Clément, Bossevain, Foretek, Vidal Cruchez, Yi, Bénard, Benassy, Kress, Vindry, Ernoult-Lange, Antoniewski, Morillon, Brest, Hubstenberger, Roest Crolius, Standart & Weil. (2019). GC content shapes mRNA storage and decay in human cells. *ELife*, 8.
- Crick. (1958). On protein synthesis. *Symposia of the Society for Experimental Biology*, 12, 138–163.

- Crick. (1970). Central dogma of molecular biology. *Nature*, 227(5258), 561–563.
- Dai, Jiang, Sha, Jiang, Ou & Fan. (2019). A combinatorial code for mRNA 3'-UTR-mediated translational control in the mouse oocyte. *Nucleic Acids Research*, 47(1), 328–340.
- Dickson, Bilger, Ballantyne & Wickens. (1999). The Cleavage and Polyadenylation Specificity Factor in *Xenopus laevis* Oocytes Is a Cytoplasmic Factor Involved in Regulated Polyadenylation. *Molecular and Cellular Biology*, 19(8), 5707–5717.
- Domingo-Sananes, Kapuy, Hunt & Novak. (2011). Switches and latches: A biochemical tug-of-war between the kinases and phosphatases that control mitosis. In *Philosophical Transactions of the Royal Society B: Biological Sciences* (Vol. 366, Issue 1584, pp. 3584–3594). Royal Society.
- Dosztanyi, Meszaros & Simon. (2009). ANCHOR: web server for predicting protein binding regions in disordered proteins. *Bioinformatics*, 25(20), 2745–2746.
- Drisaldi, Colnaghi, Fioriti, Rao, Myers, Snyder, Metzger, Tarasoff, Konstantinov, Fraser, Manley & Kandel. (2015). SUMOylation Is an Inhibitory Constraint that Regulates the Prion-like Aggregation and Activity of CPEB3. *Cell Reports*, 11(11), 1694–1702.
- Dumont. (1972). Oogenesis in *Xenopus laevis* (Daudin). I. Stages of oocyte development in laboratory maintained animals. *Journal of Morphology*, 136(2), 153–179.
- Dupre, Buffin, Roustan, Nairn, Jessus & Haccard. (2013). The phosphorylation of ARPP19 by Greatwall renders the auto-amplification of MPF independently of PKA in *Xenopus* oocytes. *Journal of Cell Science*, 126(17), 3916–3926.
- Dworkin & Dworkin-Rastl. (1985). Changes in RNA titers and polyadenylation during oogenesis and oocyte maturation in *Xenopus laevis*. *Developmental Biology*, 112(2), 451–457.
- Eichhorn, Subtelny, Kronja, Kwasnieski, Orr-Weaver & Bartel. (2016). mRNA poly(A)-tail changes specified by deadenylation broadly reshape translation in *Drosophila* oocytes and early embryos. *ELife*, 5.
- El-Brolosy, Kontarakis, Rossi, Kuenne, Günther, Fukuda, Kikhi, Boezio, Takacs, Lai, Fukuda, Gerri, Giraldez & Stainier. (2019). Genetic compensation triggered by mutant mRNA degradation. *Nature*, 568(7751), 193–197.
- Fabrizio, Dannenberg, Dube, Kastner, Stark, Urlaub & Lührmann. (2009). The Evolu-

tionarily Conserved Core Design of the Catalytic Activation Step of the Yeast Spliceosome. *Molecular Cell*, 36(4), 593–608.

Fakim & Fabian. (2019). Communication Is Key: 5'-3' Interactions that Regulate mRNA Translation and Turnover. In *Advances in Experimental Medicine and Biology* (pp. 149–164).

Fernández-Miranda & Méndez. (2012). The CPEB-family of proteins, translational control in senescence and cancer. *Ageing Research Reviews*, 11(4), 460–472.

Ferrell & Ha. (2014a). Ultrasensitivity part I: Michaelian responses and zero-order ultrasensitivity. *Trends in Biochemical Sciences*, 39(10), 496–503.

Ferrell & Ha. (2014b). Ultrasensitivity part II: multisite phosphorylation, stoichiometric inhibitors, and positive feedback. *Trends in Biochemical Sciences*, 39(11), 556–569.

Ferrell & Ha. (2014c). Ultrasensitivity part III: cascades, bistable switches, and oscillators. *Trends in Biochemical Sciences*, 39(12), 612–618.

Fioriti, Myers, Huang, Li, Stephan, Trifilieff, Colnaghi, Kosmidis, Drisaldi, Pavlopoulos & Kandel. (2015). The Persistence of Hippocampal-Based Memory Requires Protein Synthesis Mediated by the Prion-like Protein CPEB3. *Neuron*, 86(6), 1433–1448.

Flavell, Kim, Gray, Harmin, Hemberg, Hong, Markenscoff-Papadimitriou, Bear & Greenberg. (2008). Genome-Wide Analysis of MEF2 Transcriptional Program Reveals Synaptic Target Genes and Neuronal Activity-Dependent Polyadenylation Site Selection. *Neuron*, 60(6), 1022–1038.

Ford, Ling, Kandel & Fioriti. (2019). CPEB3 inhibits translation of mRNA targets by localizing them to P bodies. *Proceedings of the National Academy of Sciences*, 116(36), 18078–18087.

Fox, Sheets & Wickens. (1989). Poly(A) addition during maturation of frog oocytes: distinct nuclear and cytoplasmic activities and regulation by the sequence UUUUUAU. *Genes & Development*, 3(12b), 2151–2162.

Franzmann & Alberti. (2019). Prion-like low-complexity sequences: Key regulators of protein solubility and phase behavior. In *Journal of Biological Chemistry* (Vol. 294, Issue 18, pp. 7128–7136). American Society for Biochemistry and Molecular Biology Inc.

- Gagnon, Kreiling, Powrie, Wood & Mowry. (2013). Directional Transport Is Mediated by a Dynein-Dependent Step in an RNA Localization Pathway. *PLoS Biology*, 11(4), e1001551.
- Gallie. (1991). The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes & Development*, 5(11), 2108–2116.
- Gehring, Wahle & Fischer. (2017). Deciphering the mRNP Code: RNA-Bound Determinants of Post-Transcriptional Gene Regulation. *Trends in Biochemical Sciences*, 42(5), 369–382.
- Gerhart, Danilchik, Doniach, Roberts, Rowning & Stewart. (1989). Cortical rotation of the *Xenopus* egg: Consequences for the anteroposterior pattern of embryonic dorsal development. *Development*, 37–51.
- Gerovac & Tampé. (2019). Control of mRNA Translation by Versatile ATP-Driven Machines. *Trends in Biochemical Sciences*, 44(2), 167–180.
- Goascogne, Sananès, Gouézou & Baulieu. (1985). Testosterone-induced meiotic maturation of *Xenopus laevis* oocytes: Evidence for an early effect in the synergistic action of insulin. *Developmental Biology*, 109(1), 9–14.
- Gros, Hiatt, Gilbert, Kurland, Risebrough & Watson. (1961). Unstable Ribonucleic Acid Revealed by Pulse Labelling of *Escherichia Coli*. *Nature*, 190(4776), 581–585.
- Gu, Eils & Schlesner. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, 32(18), 2847–2849.
- Guillén-Boixet, Buzon, Salvatella & Méndez. (2016). CPEB4 is regulated during cell cycle by ERK2/Cdk1-mediated phosphorylation and its assembly into liquid-like droplets. *ELife*, 5.
- Hake & Richter. (1994). CPEB is a specificity factor that mediates cytoplasmic polyadenylation during *Xenopus* oocyte maturation. *Cell*, 79(4), 617–627.
- Han, Kato, Xie, Wu, Mirzaei, Pei, Chen, Xie, Allen, Xiao & McKnight. (2012). Cell-free Formation of RNA Granules: Bound RNAs Identify Features and Components of Cellular Assemblies. *Cell*, 149(4), 768–779.
- Hastie, Tibshirani, Narasimhan & Chu. (2020). impute: Imputation for microarray data.

- Heim, Tischer & Mayer. (2018). Calcineurin promotes APC/C activation at meiotic exit by acting on both XErp1 and Cdc20. *EMBO Reports*, 19(12).
- Hervas, Rau, Park, Zhang, Murzin, Fitzpatrick, Scheres & Si. (2020). Cryo-EM structure of a neuronal functional amyloid implicated in memory persistence in *Drosophila*. *Science*, 367(6483), 1230–1234.
- Herzel, Ottoz, Alpert & Neugebauer. (2017). Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. *Nature Reviews Molecular Cell Biology*, 18(10), 637–650.
- Hieronimus & Silver. (2004). A systems view of mRNP biology. In *Genes and Development* (Vol. 18, Issue 23, pp. 2845–2860). Cold Spring Harbor Laboratory Press.
- Horie & Kotani. (2016). Formation of mos RNA granules in the zebrafish oocyte that differ from cyclin B1 RNA granules in distribution, density and regulation. *European Journal of Cell Biology*, 95(12), 563–573.
- Hosoda, Funakoshi, Hirasawa, Yamagishi, Asano, Miyagawa, Ogami, Tsujimoto & Hoshino. (2011). Anti-proliferative protein Tob negatively regulates CPEB3 target by recruiting Caf1 deadenylase. *The EMBO Journal*, 30(7), 1311–1323.
- Huang, Carson, Barbarese & Richter. (2003). Facilitation of dendritic mRNA transport by CPEB. *Genes and Development*, 17(5), 638–653.
- Huang, Jung, Sarkissian & Richter. (2002). N-methyl-D-aspartate receptor signaling results in Aurora kinase-catalyzed CPEB phosphorylation and CaMKII mRNA polyadenylation at synapses. *EMBO Journal*, 21(9), 2139–2148.
- Huang, Kan, Lin, Richter, Alarcon, Hodgman, Theis, Kandel, Richter, Bailey, Kandel, Si, Barnard, Cao, Richter, Barnard, Ryan, Manley, Richter, ... Zuker. (2006). CPEB3 and CPEB4 in neurons: analysis of RNA-binding specificity and translational control of AMPA receptor GluR2 mRNA. *The EMBO Journal*, 25(20), 4865–4876.
- Igea & Méndez. (2010). Meiosis requires a translational positive loop where CPEB1 ensues its replacement by CPEB4. *The EMBO Journal*, 29(13), 2182–2193.
- Imataka. (1998). A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *The EMBO Journal*, 17(24), 7480–7489.

- Jacob & Monod. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, 3(3), 318–356.
- Jung, Lorenz & Richter. (2006). Translational Control by Neuroguidin, a Eukaryotic Initiation Factor 4E and CPEB Binding Protein. *Molecular and Cellular Biology*, 26(11), 4277–4287.
- Kamenska, Lu, Kubacka, Broomhead, Minshall, Bushell & Standart. (2014). Human 4E-T represses translation of bound mRNAs and enhances microRNA-mediated silencing. *Nucleic Acids Research*, 42(5), 3298–3313.
- Kamenska, Simpson, Vindry, Broomhead, Bénard, Ernoult-Lange, Lee, Harries, Weil & Standart. (2016). The DDX6–4E-T interaction mediates translational repression and P-body assembly. *Nucleic Acids Research*, 44(13), 6318–6334.
- Kanehisa. (2019). Toward understanding the origin and evolution of cellular organisms. In *Protein Science* (Vol. 28, Issue 11, pp. 1947–1951). Blackwell Publishing Ltd.
- Kato, Han, Xie, Shi, Du, Wu, Mirzaei, Goldsmith, Longgood, Pei, Grishin, Frantz, Schneider, Chen, Li, Sawaya, Eisenberg, Tycko & McKnight. (2012). Cell-free Formation of RNA Granules: Low Complexity Sequence Domains Form Dynamic Fibers within Hydrogels. *Cell*, 149(4), 753–767.
- Kaufmann, Martin, Friedlein, Langen & Keller. (2004). Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *EMBO Journal*, 23(3), 616–626.
- Khong, Matheny, Jain, Mitchell, Wheeler & Parker. (2017). The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules. *Molecular Cell*, 68(4), 808-820.e5.
- Kim & Richter. (2006). Opposing Polymerase-Deadenylase Activities Regulate Cytoplasmic Polyadenylation. *Molecular Cell*, 24(2), 173–183.
- Kim, Jensen, Noble, KC, Roux, Motamedchaboki & Roux. (2016). An improved smaller biotin ligase for BioID proximity labeling. *Molecular Biology of the Cell*, 27(8), 1188–1196.
- Kishimoto. (2018). MPF-based meiotic cell cycle control: Half a century of lessons from starfish oocytes. *Proceedings of the Japan Academy, Series B*, 94(4), 180–203.

- Kloc, Bilinski, Dougherty, Brey & Etkin. (2004). Formation, architecture and polarity of female germline cyst in *Xenopus*. *Developmental Biology*, 266(1), 43–61.
- Kotani, Yasuda, Ota & Yamashita. (2013). Cyclin b1 mRNA translation is temporally controlled through formation and disassembly of RNA granules. *Journal of Cell Biology*, 202(7), 1041–1055.
- Koulouras, Panagopoulos, Rapsomaniki, Giakoumakis, Taraviras & Lygerou. (2018). EasyFRAP-web: a web-based tool for the analysis of fluorescence recovery after photobleaching data. *Nucleic Acids Research*, 46(W1), W467–W472.
- Krey, Wilmarth, Shin, Klimek, Sherman, Jeffery, Choi, David & Barr-Gillespie. (2014). Accurate Label-Free Protein Quantitation with High- and Low-Resolution Mass Spectrometers. *Journal of Proteome Research*, 13(2), 1034–1044.
- Lackford, Yao, Charles, Weng, Zheng, Choi, Xie, Wan, Xing, Freudenberg, Yang, Jothi, Hu & Shi. (2014). Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. *The EMBO Journal*, 33(8), 878–889.
- Langmead, Trapnell, Pop & Salzberg. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, 10(3), R25.
- Lantz, Chang, Horabin, Bopp & Schedl. (1994). The *Drosophila orb* RNA-binding protein is required for the formation of the egg chamber and establishment of polarity. *Genes and Development*, 8(5), 598–613.
- Lawrence, Huber, Pagès, Aboyoun, Carlson, Gentleman, Morgan & Carey. (2013). Software for Computing and Annotating Genomic Ranges. *PLoS Computational Biology*, 9(8), e1003118.
- Li, P., Banjade, Cheng, Kim, Chen, Guo, Llaguno, Hollingsworth, King, Banani, Russo, Jiang, Nixon & Rosen. (2012). Phase transitions in the assembly of multivalent signaling proteins. *Nature*, 483(7389), 336–340.
- Li, Y. R., King, Shorter & Gitler. (2013). Stress granules as crucibles of ALS pathogenesis. In *Journal of Cell Biology* (Vol. 201, Issue 3, pp. 361–372). J Cell Biol.
- Love, Huber & Anders. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 550.
- Luong, Daldello, Rajkovic, Yang & Conti. (2020). Genome-wide analysis reveals a switch

in the translational program upon oocyte meiotic resumption. *Nucleic Acids Research*, 48(6), 3257–3276.

Maillo, Martín, Sebastián, Hernández-Alvarez, García-Rocha, Reina, Zorzano, Fernandez & Méndez. (2017). Circadian- and UPR-dependent control of CPEB4 mediates a translational response to counteract hepatic steatosis under ER stress. *Nature Cell Biology*, 19(2), 94–105.

Majumdar, Cesario, White-Grindley, Jiang, Ren, Khan, Li, Choi, Kannan, Guo, Unruh, Slaughter & Si. (2012). Critical role of amyloid-like oligomers of *Drosophila* Orb2 in the persistence of memory. *Cell*, 148(3), 515–529.

Martinez-Rucobo, Kohler, van de Waterbeemd, Heck, Hemann, Herzog, Stark & Cramer. (2015). Molecular Basis of Transcription-Coupled Pre-mRNA Capping. *Molecular Cell*, 58(6), 1079–1089.

Mayya & Duchaine. (2019). Ciphers and Executioners: How 3'-Untranslated Regions Determine the Fate of Messenger RNAs. *Frontiers in Genetics*, 10.

McGrew, Dworkin-Rastl, Dworkin & Richter. (1989). Poly(A) elongation during *Xenopus* oocyte maturation is required for translational recruitment and is mediated by a short sequence element. *Genes & Development*, 3(6), 803–815.

McSwiggen, Mir, Darzacq & Tjian. (2019). Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. In *Genes & development* (Vol. 33, Issues 23–24, pp. 1619–1634). NLM (Medline).

Meijer, Kong, Lu, Wilczynska, Spriggs, Robinson, Godfrey, Willis & Bushell. (2013). Translational repression and eIF4A2 activity are critical for microRNA-mediated gene regulation. *Science*, 340(6128), 82–85.

Mendez, Barnard & Richter. (2002). Differential mRNA translation and meiotic progression require Cdc2-mediated CPEB destruction. *EMBO Journal*, 21(7), 1833–1844.

Mendez, Hake, Andresson, Littlepage, Ruderman & Richter. (2000). Phosphorylation of CPE binding factor by Eg2 regulates translation of *c-mos* mRNA. *Nature*, 404(6775), 302–307.

Mendez, Murthy, Ryan, Manley & Richter. (2000). Phosphorylation of CPEB by Eg2 Mediates the Recruitment of CPSF into an Active Cytoplasmic Polyadenylation Complex. *Molecular Cell*, 6(5), 1253–1259.

- Meneau, Dupré, Jessus & Daldello. (2020). Translational Control of *Xenopus* Oocyte Meiosis: Toward the Genomic Era. In *Cells* (Vol. 9, Issue 6, p. 1502). NLM (Medline).
- Merkel, Wells, Hilburn, Elazzouzi, Pérez-Alvarado & Lee. (2013). The C-terminal region of cytoplasmic polyadenylation element binding protein is a ZZ domain with potential for protein-protein interactions. *Journal of Molecular Biology*, 425(11), 2015–2026.
- Mészáros, Erdős & Dosztányi. (2018). IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Research*, 46(W1), W329–W337.
- Minshall, Reiter, Weil & Standart. (2007). CPEB Interacts with an Ovary-specific eIF4E and 4E-T in Early *Xenopus* Oocytes. *Journal of Biological Chemistry*, 282(52), 37389–37401.
- Mochida & Hunt. (2007). Calcineurin is required to release *Xenopus* egg extracts from meiotic M phase. *Nature*, 449(7160), 336–340.
- Mochida, Rata, Hino, Nagai & Novák. (2016). Two Bistable Switches Govern M Phase Entry. *Current Biology*, 26(24), 3361–3367.
- Neil, Jeschonek, Cabral, O’Connell, Powrie, Wood & Mowry. (2020). L-bodies are novel RNA-protein condensates driving RNA transport in *Xenopus* oocytes. *BioRxiv*, 2020.05.08.084814.
- Nirenberg & Matthaei. (1961). The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proceedings of the National Academy of Sciences*, 47(10), 1588–1602.
- Nishiyama, Yoshizaki, Kishimoto & Ohsumi. (2007). Transient activation of calcineurin is essential to initiate embryonic development in *Xenopus laevis*. *Nature*, 449(7160), 341–345.
- Novoa, Gallego, Ferreira & Mendez. (2010). Mitotic cell-cycle progression is regulated by CPEB1 and CPEB4-dependent translational control. *Nature Cell Biology*, 12(5), 447–456.
- O’Brien, Gunawardena, Paulo, Chen, Ibrahim, Gygi & Qaqish. (2018). The effects of nonignorable missing data on label-free mass spectrometry proteomics experiments. *Annals of Applied Statistics*, 12(4), 2075–2095.

- Ogami, Hosoda, Funakoshi & Hoshino. (2014). Antiproliferative protein Tob directly regulates c-myc proto-oncogene expression through cytoplasmic polyadenylation element-binding protein CPEB. *Oncogene*, 33(1), 55–64.
- Ortiz-Zapater, Pineda, Martínez-Bosch, Fernández-Miranda, Iglesias, Alameda, Moreno, Eliscovich, Eyrales, Real, Méndez & Navarro. (2012). Key contribution of CPEB4-mediated translational control to cancer progression. *Nature Medicine*, 18(1), 83–90.
- Pan, Shai, Lee, Frey & Blencowe. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genetics*, 40(12), 1413–1415.
- Pascual, Martín, Salvador, Reina, Chanes, Millanes-Romero, Suñer, Fernández-Miranda, Bartomeu, Huang, Gomis & Méndez. (2020). The RNA binding protein CPEB2 regulates hormone sensing in mammary gland development and luminal breast cancer. *Science Advances*, 6(20).
- Pascual. (2018). CPEB2 in mammary gland homeostasis and breast cancer [Universitat de Barcelona].
- Patel, Malinowska, Saha, Wang, Alberti, Krishnan & Hyman. (2017). ATP as a biological hydrotrope. *Science*, 356(6339), 753–756.
- Pavlopoulos, Trifilieff, Chevalleyre, Fioriti, Zairis, Pagano, Malleret & Kandel. (2011). Neuralized1 activates CPEB3: A function for nonproteolytic ubiquitin in synaptic plasticity and memory storage. *Cell*, 147(6), 1369–1383.
- Perales & Bentley. (2009). “Cotranscriptionality”: The Transcription Elongation Complex as a Nexus for Nuclear Transactions. *Molecular Cell*, 36(2), 178–191.
- Peuchen, Cox, Sun, Hebert, Coon, Champion, Dovichi & Huber. (2017). Phosphorylation Dynamics Dominate the Regulated Proteome during Early Xenopus Development. *Scientific Reports*, 7(1), 1–14.
- Piqué, López, Foissac, Guigó & Méndez. (2008). A Combinatorial Code for CPE-Mediated Translational Control. *Cell*, 132(3), 434–448.
- Planet, Stephan-Otto, Reina, Flores & Rossell. (2018). htSeqTools: Quality Control, Visualization and Processing for High-Throughput Sequencing data (R package version 1.31.0).

- Powrie, Ciocanel, Kreiling, Gagnon, Sandstede & Mowry. (2016). Using in vivo imaging to measure RNA mobility in *Xenopus laevis* oocytes. *Methods*, 98, 60–65.
- R Core Team. (2020). R: A Language and Environment for Statistical Computing.
- Racki & Richter. (2006). CPEB controls oocyte growth and follicle development in the mouse. *Development*, 133(22), 4527–4537.
- Ramaswami, Taylor & Parker. (2013). Altered ribostasis: RNA-protein granules in degenerative disorders. In *Cell* (Vol. 154, Issue 4, pp. 727–736). Cell.
- Rasar & Hammes. (2006). The Physiology of the *Xenopus laevis* Ovary. In *Methods in molecular biology* (Clifton, N.J.) (Vol. 322, pp. 17–30).
- Rauh, Schmidt, Bormann, Nigg & Mayer. (2005). Calcium triggers exit from meiosis II by targeting the APC/C inhibitor XErp1 for degradation. *Nature*, 437(7061), 1048–1052.
- Robinson, Thorvaldsdóttir, Winckler, Guttman, Lander, Getz & Mesirov. (2011). Integrative genomics viewer. *Nature Biotechnology*, 29(1), 24–26.
- Roux, Kim, Raida & Burke. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *Journal of Cell Biology*, 196(6), 801–810.
- Ryu & Kim. (2017). Translation initiation mediated by nuclear cap-binding protein complex. *BMB Reports*, 50(4), 186–193.
- Salditt-Georgieff, Harpold, Chen-Kiang & Darnell. (1980). The addition of 5' cap structures occurs early in hnRNA synthesis and prematurely terminated molecules are capped. *Cell*, 19(1), 69–78.
- Sandberg, Neilson, Sarma, Sharp & Burge. (2008). Proliferating Cells Express mRNAs with Shortened 3' Untranslated Regions and Fewer MicroRNA Target Sites. *Science*, 320(5883), 1643–1647.
- Sato & Tokmakov. (2020). Toward the understanding of biology of oocyte life cycle in *Xenopus Laevis* : No oocytes left behind. *Reproductive Medicine and Biology*, 19(2), 114–119.
- Sato. (2014). Transmembrane Signal Transduction in Oocyte Maturation and Fertiliza-

tion: Focusing on *Xenopus laevis* as a Model Animal. *International Journal of Molecular Sciences*, 16(1), 114–134.

Sears, May & Roux. (2019). BioID as a Tool for Protein-Proximity Labeling in Living Cells. In *Methods in Molecular Biology* (pp. 299–313).

Session, Uno, Kwon, Chapman, Toyoda, Takahashi, Fukui, Hikosaka, Suzuki, Kondo, van Heeringen, Quigley, Heinz, Ogino, Ochi, Hellsten, Lyons, Simakov, Putnam, ... Rokhsar. (2016). Genome evolution in the allotetraploid frog *Xenopus laevis*. *Nature*, 538(7625), 336–343.

Setoyama, Yamashita & Sagata. (2007). Mechanism of degradation of CPEB during *Xenopus* oocyte maturation. *Proceedings of the National Academy of Sciences*, 104(46), 18001–18006.

Shi & Manley. (2015). The end of the message: multiple protein–RNA interactions define the mRNA polyadenylation site. *Genes & Development*, 29(9), 889–897.

Si & Kandel. (2016). The role of functional prion-like proteins in the persistence of memory. *Cold Spring Harbor Perspectives in Biology*, 8(4).

Si, Choi, White-Grindley, Majumdar & Kandel. (2010). Aplysia CPEB Can Form Prion-like Multimers in Sensory Neurons that Contribute to Long-Term Facilitation. *Cell*, 140(3), 421–435.

Si, Giustetto, Etkin, Hsu, Janisiewicz, Miniaci, Kim, Zhu & Kandel. (2003). A Neuronal Isoform of CPEB Regulates Local Protein Synthesis and Stabilizes Synapse-Specific Long-Term Facilitation in Aplysia. *Cell*, 115(7), 893–904.

Singh, Pratt, Yeo & Moore. (2015). The Clothes Make the mRNA: Past and Present Trends in mRNP Fashion. *Annual Review of Biochemistry*, 84(1), 325–354.

Smyth. (2005). Limma: linear models for microarray data. In S. D. R. I. W. H. R. Gentleman V. Carey (Ed.), *Bioinformatics and Computational Biology Solutions using R and Bioconductor* (pp. 397–420). Springer.

Söding, Zwicker, Sohrabi-Jahromi, Boehning & Kirschbaum. (2020). Mechanisms for Active Regulation of Biomolecular Condensates. *Trends in Cell Biology*, 30(1), 4–14.

Sousa Martins, Liu, Oke, Arora, Franciosi, Viville, Laird, Fung & Conti. (2016). DAZL and CPEB1 regulate mRNA translation synergistically during oocyte maturation.

*Journal of Cell Science*, 129(6), 1271–1282.

- Stebbins-Boaz, Cao, de Moor, Mendez & Richter. (1999). Maskin Is a CPEB-Associated Factor that Transiently Interacts with eIF-4E. *Molecular Cell*, 4(6), 1017–1027.
- Stebbins-Boaz, Hake & Richter. (1996). CPEB controls the cytoplasmic polyadenylation of cyclin, Cdk2 and c-mos mRNAs and is necessary for oocyte maturation in *Xenopus*. *The EMBO Journal*, 15(10), 2582–2592.
- Stephan, Fioriti, Lamba, Colnaghi, Karl, Derkatch & Kandel. (2015). The CPEB3 Protein Is a Functional Prion that Interacts with the Actin Cytoskeleton. *Cell Reports*, 11(11), 1772–1785.
- Stepien, Oppitz, Gerlach, Dag, Novatchkova, Krüttner, Stark & Keleman. (2016). RNA-binding profiles of *Drosophila* CPEB proteins Orb and Orb2. *Proceedings of the National Academy of Sciences of the United States of America*, 113(45), E7030–E7038.
- Szklarczyk, Gable, Lyon, Junge, Wyder, Huerta-Cepas, Simonovic, Doncheva, Morris, Bork, Jensen & Mering. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, 47(D1), D607–D613.
- Tang, Wu, Guo, Hansen, Perry, Freel, Nutt, Jackson & Kornbluth. (2008). Cdc2 and Mos Regulate Emi2 Stability to Promote the Meiosis I–Meiosis II Transition. *Molecular Biology of the Cell*, 19(8), 3536–3543.
- Tay & Richter. (2001). Germ Cell Differentiation and Synaptonemal Complex Formation Are Disrupted in CPEB Knockout Mice. *Developmental Cell*, 1(2), 201–213.
- Tay, Hodgman, Sarkissian & Richter. (2003). Regulated CPEB phosphorylation during meiotic progression suggests a mechanism for temporal control of maternal mRNA translation. *Genes and Development*, 17(12), 1457–1462.
- Theis, Si & Kandel. (2003). Two previously undescribed members of the mouse CPEB family of genes and their inducible expression in the principal cell layers of the hippocampus. *Proceedings of the National Academy of Sciences of the United States of America*, 100(16), 9602–9607.
- Tian & Graber. (2012). Signals for pre-mRNA cleavage and polyadenylation. *Wiley Interdisciplinary Reviews: RNA*, 3(3), 385–396.

- Tong. (2013). Structure and function of biotin-dependent carboxylases. *Cellular and Molecular Life Sciences*, 70(5), 863–891.
- Tushev, Glock, Heumüller, Bieber, Jovanovic & Schuman. (2018). Alternative 3' UTRs Modify the Localization, Regulatory Potential, Stability, and Plasticity of mRNAs in Neuronal Compartments. *Neuron*, 98(3), 495–511.e6.
- Udagawa, Swanger, Takeuchi, Kim, Nalavadi, Shin, Lorenz, Zukin, Bassell & Richter. (2012). Bidirectional Control of mRNA Translation and Synaptic Plasticity by the Cytoplasmic Polyadenylation Complex. *Molecular Cell*, 47(2), 253–266.
- Verlhac & Wingman Lee. (2010). Mechanisms of Asymmetric Division in Metazoan Meiosis. In *Oogenesis* (pp. 291–310). John Wiley & Sons, Ltd.
- Waghray, Williams, Coon & Wickens. (2015). *Xenopus* CAF1 requires NOT1-mediated interaction with 4E-T to repress translation in vivo. *RNA*, 21(7), 1335–1345.
- Walker, Minshall, Hake, Richter & Standart. (1999). The clam 3' UTR masking element-binding protein p82 is a member of the CPEB family. *RNA*, 5(1), 14–26.
- Wang, C. F. & Huang. (2012). Calpain 2 Activated through N-Methyl-D-Aspartic Acid Receptor Signaling Cleaves CPEB3 and Abrogates CPEB3-Repressed Translation in Neurons. *Molecular and Cellular Biology*, 32(16), 3321–3332.
- Wang, E. T., Sandberg, Luo, Khrebtkova, Zhang, Mayr, Kingsmore, Schroth & Burge. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature*, 456(7221), 470–476.
- Wang, X. P. & Cooper. (2010). Comparative in Silico analyses of Cpeb1-4 with functional predictions. *Bioinformatics and Biology Insights*, 4, 61–83.
- Weill, Belloc, Castellazzi & Méndez. (2017). Musashi 1 regulates the timing and extent of meiotic mRNA translational activation by promoting the use of specific CPEs. *Nature Structural & Molecular Biology*, 24(8), 672–681.
- Wells, Hillner, Vale & Sachs. (1998). Circularization of mRNA by Eukaryotic Translation Initiation Factors. *Molecular Cell*, 2(1), 135–140.
- White-Grindley, Li, Mohammad Khan, Ren, Saraf, Florens & Si. (2014). Contribution of Orb2A Stability in Regulated Amyloid-Like Oligomerization of *Drosophila* Orb2. *PLoS Biology*, 12(2), e1001786.

- Wickham, François, Henry & Mueller. (2018). *dplyr: A Grammar of Data Manipulation*.
- Wickham. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- Wilczynska, Aigueperse, Kress, Dautry & Weil. (2005). The translational regulator CPEB1 provides a link between dcp1 bodies and stress granules. *Journal of Cell Science*, 118(5), 981–992.
- Wootton & Federhen. (1993). Statistics of local complexity in amino acid sequences and sequence databases. *Computers & Chemistry*, 17(2), 149–163.
- Wu, D., Lim, Vaillant, Asselin-Labat, Visvader & Smyth. (2010). ROAST: rotation gene set tests for complex microarray experiments. *Bioinformatics*, 26(17), 2176–2182.
- Wu, Wells, Tay, Mendis, Abbott, Barnitt, Quinlan, Heynen, Fallon & Richter. (1998). CPEB-mediated cytoplasmic polyadenylation and the regulation of experience-dependent translation of CaMKII mRNA at synapses. *Neuron*, 21(5), 1129–1139.
- Wu, X. & Bartel. (2017). Widespread Influence of 3'-End Structures on Mammalian mRNA Processing and Stability. *Cell*, 169(5), 905-917.e11.
- Wühr, Freeman, Presler, Horb, Peshkin, Gygi & Kirschner. (2014). Deep proteomics of the *xenopus laevis* egg using an mRNA-derived reference database. *Current Biology*, 24(13), 1467–1475.
- Xie & Ren. (2019). Mechanisms of nuclear mRNA export: A structural perspective. *Traffic*, 20(11), 829–840.
- Xu, Hafer, Agunwamba & Schedl. (2012). The CPEB Protein Orb2 Has Multiple Functions during Spermatogenesis in *Drosophila melanogaster*. *PLoS Genetics*, 8(11), 1003079.
- Yang, C.-R., Rajkovic, Daldello, Luong, Chen & Conti. (2020). The RNA-binding protein DAZL functions as repressor and activator of mRNA translation during oocyte maturation. *Nature Communications*, 11(1), 1399.
- Yang, F., Wang, Cetinbas, Sadreyev & Blower. (2020). Genome-wide analysis identifies cis-acting elements regulating mRNA polyadenylation and translation during vertebrate oocyte maturation. *RNA*, 26(3), 324–344.

Zhao, Xie, Zheng, Jiang, Liu, Mu, Liu, Zhao, Xue & Ren. (2014). GPS-SUMO: a tool for the prediction of sumoylation sites and SUMO-interaction motifs. *Nucleic Acids Research*, 42(W1), W325–W330.







