Visualising the localisation of mRNA to granules in yeast

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Biology, Medicine and Health

2021

Emma Linney

School of Biological Sciences,
Division of Molecular and Cellular Biology

Contents

Word Count: 44,481

List of Figures

List of Tables

Abbreviations

4E-BPs	eIF4E binding proteins
ARE	AU-rich element
Cas13	CRISPR associated protein 13
Cas9	CRISPR associated protein 9
cDNA	Complementary DNA
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	crispr RNA
CTD	Carboxyl terminal domain
CXR	Carboxy-X-rhodamine
dCas13	Dead Cas13
dCas9	Dead Cas9
DDX6	DEAD-box helicase 6
DNA	Deoxyribonucleic acid
eGFP	Enhanced green fluorescent protein
EJC	Exon junction complex
ER	Endoplasmic reticulum
FISH	Fluorescent *in situ* hybridisation
GFP	Green fluorescent protein
HDR	Homology directed repair
IDRs	Intrinsically disordered regions
ISH	*In situ* hybridisation
LB	Luria-Bertani media
LiAc	Lithium acetate
LLPS	Liquid-liquid phase separation
MCP	MS2 coat protein
MRGs	Mitochondrial RNA granules
mRNA	Messenger ribonucleic acid
mRNP	mRNA protein complex/ messenger ribonucleoprotein complex
mtDNA	Mitochondrial DNA

NHEJ	Non-homologous end joining
NLS	Nuclear localisation signal
NMD	Nonsense mediated decay
NSBs	Nuclear stress bodies
ORF	Open reading frame
ORI	Origin of replication
PABP	Poly(A) binding protein
PAM	Protospacer adjacent motif
PAP	Poly(A) polymerase
PBs	P-bodies
PCR	Polymerase chain reaction
Pol II	RNA polymerase II
qRT-PCR	Quantitative real-time PCR
RBP(s)	RNA binding protein(s)
RFP	Red fluorescent protein
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulphate
sgRNA	Single-guide RNA
SGs	Stress granules
smFISH	Single-molecule fluorescent *in situ* hybridisation
SREs	Smaug recognition elements
SRP	Signal recognition particle
ssRNA	Single-stranded RNA
TFAM	Transcription factor A
tracrRNA	trans-activating crispr RNA
TREX	Transcription-export (complex)
TRICK	Translating RNA imaging by coat protein knock-off
UTR	Untranslated region
yeGFP	Yeast enhanced green fluorescent protein
ZBP1	Zip code binding protein 1

Abstract

mRNA localisation is a fundamental cellular process that provides a layer of protein regulation by directing protein synthesis in time and space. mRNAs may localise to cytosolic RNA granules, which contain distinct RNA and protein components. Cytosolic RNA granules are membrane-less regions that are often described as phase-separated from the surrounding cytosol, and appear to form via both specific and non-specific RNA-RNA, RNA-protein, and protein-protein interactions. RNA granules may regulate protein production by storing mRNAs in translationally repressed states, by encouraging either mRNA stabilisation or degradation, or by localising together specific mRNAs to produce proteins in the same multi-subunit complexes.

Previous work in the Ashe lab has identified a novel set of RNA granules in yeast. These granules are distinct from other RNA granules; they do not co-localise with classical markers for other types of granules, they are present in unstressed yeast cells, and they contain mRNAs that are actively translated. Further work has demonstrated that there are at least three types of granules harbouring actively translated mRNAs, which contain different, specific mRNAs and display differing cytosolic RNA granule patterns. Experiments were therefore performed in this project involving the mutagenesis of select mRNAs to determine the localisation elements that drive this differential mRNA localisation. These experiments showed that the ORFs of two mRNAs were capable of driving the localisation pattern of their respective full-length mRNAs, suggesting that either the ORFs of both mRNAs contain RNA sequence and/or structural localisation elements, or that the production of the nascent protein is capable of driving the localisation of its own mRNA.

To study mRNA localisation, microscopic visualisation techniques are used. Current systems all have potential drawbacks; the MS2 system involves modification of the mRNA structure, smFISH requires cell fixation/permeabilisation, and RNA aptamer systems require both mRNA modification and cell permeabilisation. A recently developed mRNA visualisation system, the dCas9-RNA targeting system, was therefore developed in this project for use in yeast. The dCas9 system is a live cell system that does not involve modification of the target mRNA, and is easily introduced into cells in a plasmid-based system. The dCas9 protein is specifically targeted to the mRNA of interest via a single-guide RNA, removing the need to modify the structure of the mRNA. The dCas9 system is shown here to be capable of accurately revealing patterns of mRNA localisation without affecting mRNA abundance, and is further used to target multiple mRNAs to identify novel localisation patterns. This dCas9 system has the potential to be used to rapidly screen endogenous, unmodified mRNA localisation patterns and study their dynamics in live cells.

Declaration

I declare that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

Acknowledgements

Firstly, I would like to thank my supervisor Mark Ashe, who has always been a positive and supportive influence in the lab, and whose advice and encouragement has allowed this PhD to happen.

I would also like to thank members of the Ashe, Grant and Pavitt labs for their friendship and support; particularly Jenny, Christian, Claire and Zorana, whose company made my time in the lab infinitely more enjoyable. I am also grateful for the other members of my PhD course who have always been willing to share advice, complaints, drinks and successes.

Thank you to my wonderful parents and brother, who have supported me all my life; I am particularly grateful for the copious number of Zoom calls to help get me through my last year of PhD.

Finally I would like to thank Mark, who has been with me throughout my PhD project and is a constant source of happiness and light.

Introduction

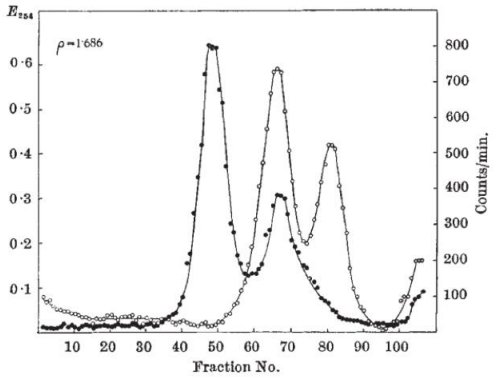
1. Introduction

1.1. General introduction

The study of molecular and cellular biology is always expanding; new knowledge does not narrow the field, but rather opens up new lines of enquiry and possibilities. DNA, discovered in 1869, has been a key component in the title of at least 100,000 published papers every year since 1987. One reason for the continual growth of scientific subjects is the development of new technology. One informs the other; new knowledge leads to the advent of new techniques (e.g. the use of DNA polymerases in PCR and cloning), and new techniques allow new knowledge to be gained (e.g. the use of PCR and cloning in DNA fingerprinting).

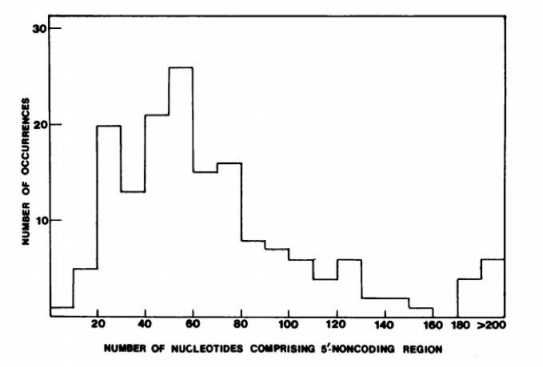
Of particular interest here are the techniques used to study mRNA dynamics and localisation. Early experiments to determine the existence and function of mRNA involved techniques such as isotope labelling and direct (non-computational) analysis of RNA sequences, and these techniques have been supplemented by the development of imaging techniques to directly visualise mRNAs and study their movement (Figure 1.1). An incredible amount of information about mRNA structure, function and dynamics was discovered using these early techniques before mRNAs were ever seen. For example, evidence of the existence of mRNAs was found through isotope labelling of host cells and invading phage cells. Following phage invasion the host cell transcription is halted and the phage DNA is then transcribed; differential isotope labelling of host and phage cell components demonstrated that host cell proteins may be produced after transcription is halted, indicating the existence of a semi-stable intermediary structure; mRNA (Brenner et al., 1961).

Nevertheless, the ability to detect and visualise mRNA in cells has led to many breakthroughs in mRNA function and dynamics, particularly with regards to mRNA localisation. For example, the development of techniques to visualise mRNAs movement in live cells has allowed mRNA localisation times to be measured (Bertrand et al., 1998), mRNA nuclear export events to be visualised (Grünwald and Singer, 2010) and rates of mRNA transcription and translation to be determined (Larson et al., 2011; Wu et al., 2016). Current experiments using mRNA visualisation techniques include investigation into the physical aspects of mRNA translation (e.g. whether or not the mRNA forms a closed-loop structure (Adivarahan et al., 2018)) and multi-colour imaging of mRNA to investigate translation dynamics and interactions with ribosomes (Boersma et al., 2019). Therefore, as new techniques are developed, both the amount of scientific knowledge and the number of new scientific questions to be answered increases.



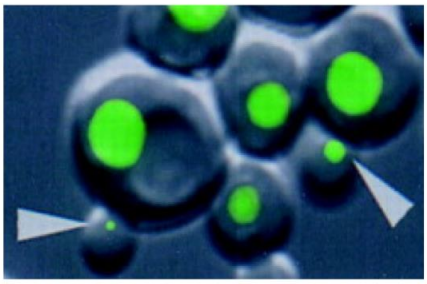
1961
Heavy isotope labelling –
Evidence for RNA

(Brenner et al., 1961)



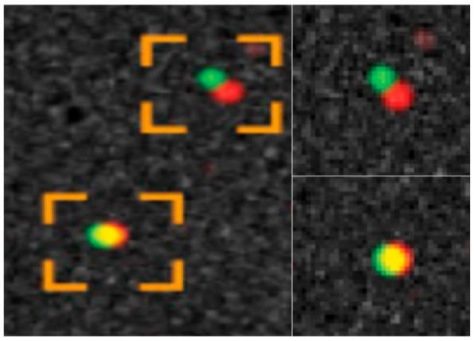
1984
Sequence analysis –
Kozak sequence

(Kozak, 1984)



1998
MS2 system –
Live cell imaging of bulk
mRNA localisation

(Bertrand et al., 1998)



2018
smFISH –
Single-molecule imaging of
mRNA

(Adivarahan et al., 2018)

Figure 1.1: Techniques for investigating mRNA function and dynamics. Early evidence of the existence of mRNA came from isotope labelling experiments performed in Brenner et al., 1961. The mammalian Kozak sequence was first discovered in Kozak, 1984, where mammalian mRNA sequences were manually aligned and analysed to discover the sequence. The MS2 system was the first live cell system able to visualise bulk mRNA localisation in yeast cells in Bertrand et al., 1998. smFISH techniques allow single molecule resolution of mRNA imaging to be achieved in mammalian cells in Adivarahan et al., 2018.

1.2. mRNA synthesis, structure and fate

Messenger RNA (mRNA) synthesis is the intermediary step necessary for the production of protein from DNA. In eukaryotes, this allows DNA to remain protected inside the nucleus of cells while allowing the information it contains to be translated outside the nucleus. In both eukaryotes and prokaryotes, the additional steps of mRNA synthesis, localisation and translation allow for multiple additional levels of the regulation of protein production. This regulation may include mRNA splicing, where different exons can be selected to produce different mRNAs which encode distinct protein isoforms from a single gene. As well as protein identity, the levels of protein can be regulated in many ways. The level of the intermediate mRNA can be by modified by increasing or decreasing the rates of mRNA synthesis (or transcription), or alternatively the rates of mRNA degradation can be altered to elicit the same effect. In addition, the level of a protein can be directly regulated by controlling its production via translation or degradation. Regulation of the protein directly has a more immediate effect and hence allows a cell to respond rapidly to changing environments. Finally, the site of protein production can be important depending on where a protein is required and whether there could be consequences if proteins are made in an inappropriate part of the cell. Therefore an mRNA can be localised to regulate where a protein is made.

1.2.1. mRNA transcription and processing

To enable appropriate protein production, an mRNA must be accurately transcribed and processed from the DNA sequence and it must be present in the correct quantity and location within a cell. Therefore, it is essential that the mRNA life cycle is strictly controlled. In eukaryotes, protein-coding genes are transcribed from DNA through the activity of RNA polymerase II (Pol II), where Pol II activity is regulated by transcription factor binding to DNA (reviewed in Fuda et al., 2009). Completion of Pol II-dependent steps of initiation, elongation, and termination produces RNA, including pre-mRNA. Before export out of the nucleus, pre-mRNA undergoes processing to produce mature mRNA, which involves 5’-end capping, splicing, and polyadenylation (reviewed in Hocine et al., 2010). mRNA processing can occur co-transcriptionally (reviewed in Bentley, 2014), and involves the binding of multiple proteins and protein complexes to the mRNA to produce a messenger ribonucleoprotein complex (mRNP). Many of these proteins remain bound to the mRNA after export to regulate mRNA localisation, rates of translation and mRNA degradation (Le Hir et al., 2000; Merz et al., 2007). mRNA transcription and processing are therefore both highly regulated steps of

mRNA production, and are linked by the activity of Pol II. In particular, the carboxyl terminal domain (CTD) of Pol II is capable of interacting with a wide variety of proteins, and acts as a scaffold for proteins associated with mRNA processing during mRNA synthesis (although not all steps of mRNA processing are associated with the CTD) (Verdecia et al., 2000; Kormarnitsky et al., 2000).

5'-end capping is the first step in mRNA processing, and occurs once 25-30 nucleotides have been transcribed (Banerjee, 1980) (Figure 1.2). The cap is produced by three enzymes; RNA triphosphatase, guanylyltransferase and 7-methyltransferase, and in eukaryotes, capping enzymes interact directly with Pol II via the CTD to couple capping with transcription initiation and elongation (Yue et al., 1997; Rodriguez et al., 2000). The 5'-end cap functions both to stabilise the mRNA by preventing 5' to 3' RNA degradation (Hsu and Stevens, 1993), and to allow the mRNA to interact with the translation factors during translation initiation (Tarun and Sachs, 1996; Wakiyama et al., 2000). The 5'-end cap of an mRNA is then followed by the 5' untranslated region (5'UTR), and then the protein coding sequence or open reading frame (ORF) followed by the 3'UTR.

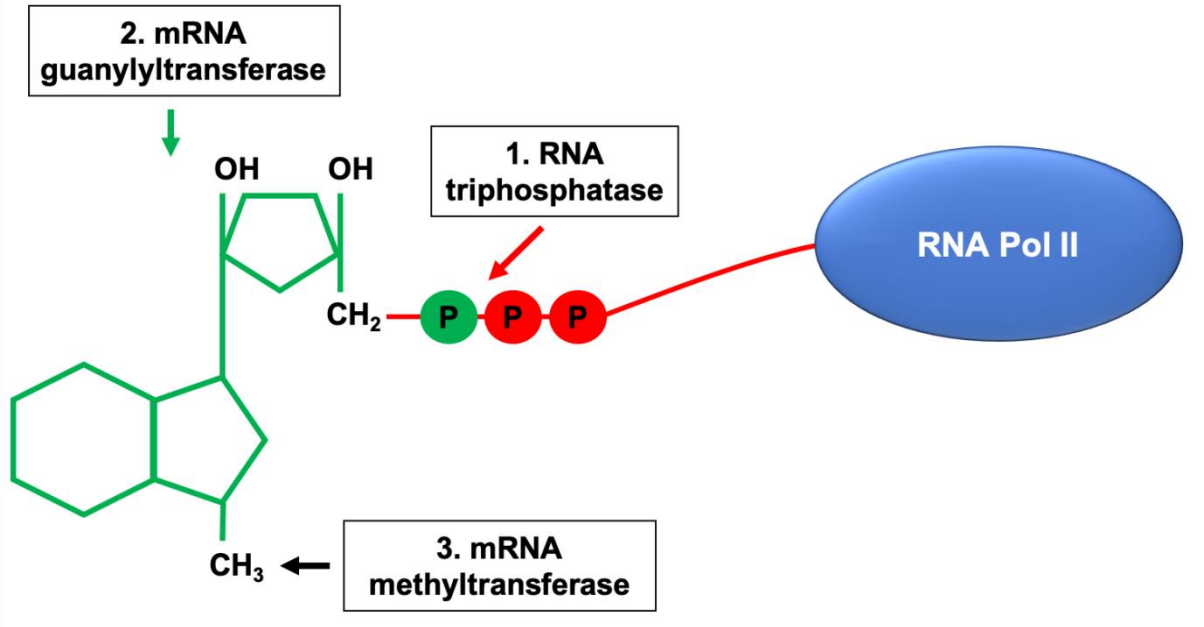


Figure 1.2: 5'end capping of mRNA. mRNA capping occurs in 3 main steps almost immediately after the mRNA begins to be transcribed. **(1)** In the first step, RNA triphosphates catalyses cleavage of the terminal phosphate of the mRNA, leaving a bisphosphate group. **(2)** mRNA guanylyltransferase catalyses the addition of a GTP molecule to the exposed phosphate (removing a phosphate from GTP in the process). **(3)** mRNA methyltranferase catalyses the addition of a methyl group to the 7-nitrogen of guanine. For certain the mRNAs, the 5' cap may undergo further modification once this is completed.

The final sequence of an mRNA is often different from the genomic DNA sequence that produced it. Regions from within the RNA sequence called introns are removed in a process termed splicing (reviewed in Lee and Rio, 2005). This means the RNA sequence to be translated into protein can vary greatly and is produced from the included exons of the gene sequence. This is because splicing profiles across an mRNA can vary greatly depending upon which exons are kept and which intronic regions are removed in a process termed alternative splicing. Splicing is catalysed by the spliceosome, a large protein-RNA complex consisting of several small ribonucleoproteins (snRNPs) and associated proteins (Moore et al., 1993). Splicing occurs via two transesterification steps, where the spliceosome first catalyses nucleophyllic attack by the 2'OH group of a branch point adenosine within the intron on the phosphodiester bond of the 5' splice site (Figure 1.3). This produces a free 5' exon and a lariat loop. The second step then involves nucleophyllic attack by the free 5' exon on the phosphodiester bond of the 3' splice site, ligating the exons together and excising the intron from the sequence (Konarska et al., 1984; Rodriguez et la., 1984) (Figure 1.3). As well as generating the correct mRNA sequence to enable the appropriate protein to be produced, splicing can facilitate export from the nucleus, mRNA stability, and translation initiation via the proteins that remain bound to the mRNA once the splicing reaction is complete (Le Hir et al., 2000; Merz et al., 2007). Although splicing is an important processing step for many mRNAs in higher eukaryotes, the majority of yeast mRNAs are not spliced, increasing the importance of regulated mRNA export in yeast as discussed below. For example, the exon junction complex (EJC) is a complex of proteins that remains bound to an mRNA after splicing (Le Hir et al., 2000), and can prove important for correct mRNA export (Cheng et al., 2006). Similar to 5'-end capping, splicing can occur co-transcriptionally and hence affect transcription efficiency (Zhang et al., 1994; McCracken et al., 1997; Howe et al., 2003), including via interactions with the CTD of Pol II (Hirose et al., 1999; Guo et al., 2019).

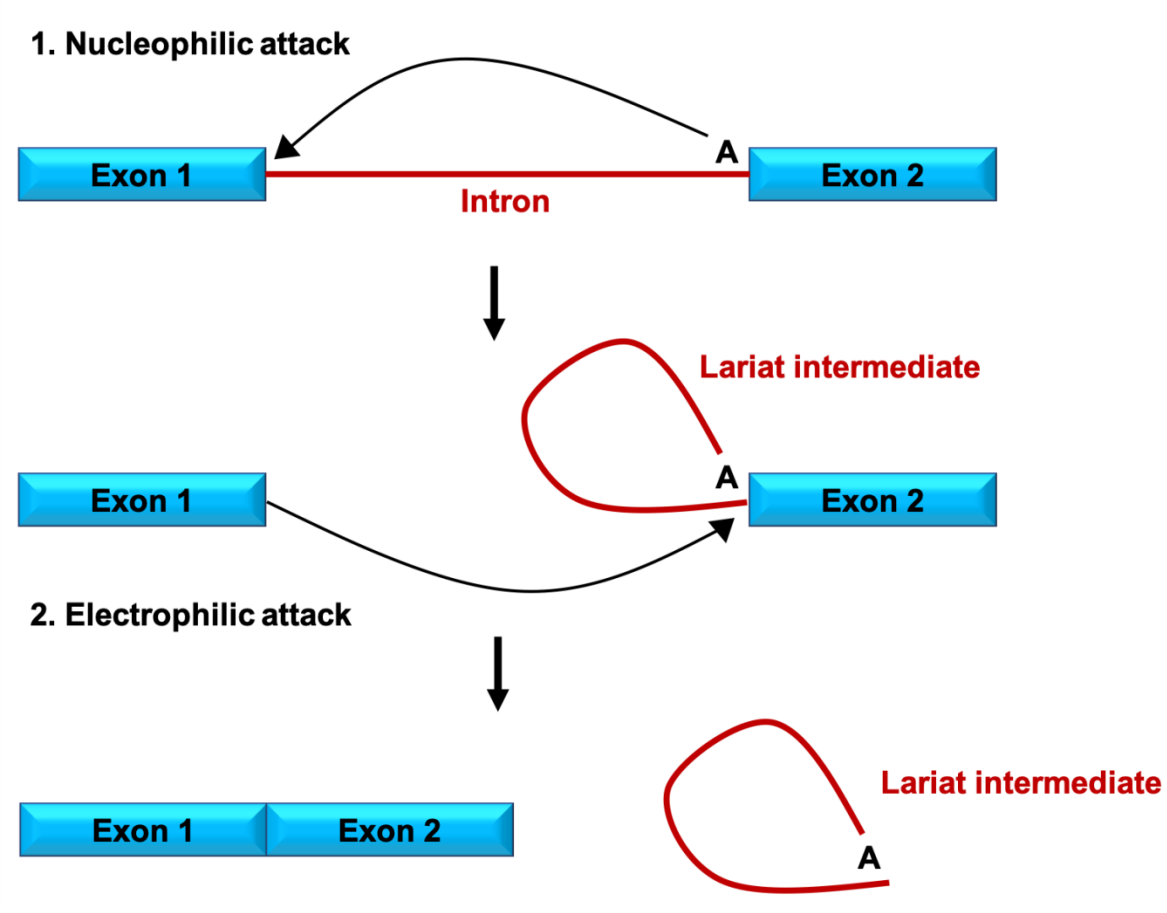


Figure 1.3: mRNA splicing. mRNA splicing functions to remove intronic regions from exon sequences, and involves in two transesterification steps. **(1)** The 2'OH group of an intronic nucleotide performs a nucleophilic attack on the first nucleotide of the intron, forming a lariat intermediate. **(2)** The exposed 3'OH group of the first exon then performs an electrophilic attack at the 3' splice site of the intron, joining the exons together.

The final step in the production of a mature mRNA is 3'-end processing or polyadenylation, which generates the 3'-end of the mRNA and therefore defines the extent of the 3'UTR. Polyadenylation is achieved by cleavage and polyadenylation of the nascent transcript at a 3'-end formation sequence. The cleavage reaction involves several associated proteins, including cleavage specificity proteins and cleavage factors (reviewed in Colgan and Manley, 1997; Lin et al., 2017; Neve et al., 2017) (Figure 1.4). Cleavage specificity factors are recruited to RNA sequences 10-30 nucleotides upstream of cleavage signal sequences; AAUAAA sequence in mammals and AU-rich sequences in yeast (Proudfoot and Brownlee, 1976; Dichtl and Keller, 2001; Sun et al., 2018). A poly(A) tail is then added to the exposed 3'-end of the RNA by the enzyme poly(A) polymerase (PAP) (Martin and Keller, 1996; Lemieux and Bachand, 2009) (Figure 1.4). This poly(A) tail has a range of functions including mRNA stabilisation and the facilitation of mRNA translation (Drummond et al., 1985). Furthermore, alternative polyadenylation can play a role in altering mRNA localisation and stability (Tian et al., 2005). As with 5'-end capping and splicing, polyadenylation is coupled with transcription, and 3'-end cleavage proteins interact with the CTD of Pol II (Hirose and Manley, 1998; Rigo and Martinson, 2009). Altogether, the steps of producing mature mRNA from pre-mRNA are tightly regulated, and each stage of processing (5'-end capping, splicing, and polyadenylation) is coupled to gene transcription.

As well as these global mRNA processing reactions that occur for most mRNAs produced, there are numerous modifications to an mRNA that tend to occur in an mRNA specific manner (reviewed in Nachtergaele and He, 2017). RNA editing is a process by which RNA nucleotide sequences can be altered post-transcriptionally to provide a variety of different functions (reviewed in Gott and Emeson, 2000; Eisenberg and Levanon, 2018; Kung et al., 2018). Hundreds of modifications to RNA have been discovered across biology and for many years it was largely assumed that these modifications were limited to non-translated mRNAs such as rRNA and tRNAs. However, recent evidence, especially for studies on the N^6-methyladenosine (m^6A) modification, suggest that mRNA can be modified and that this process can impact on the downstream fate of an mRNA, affecting its translation localisation and stability (reviewed in Zhao et al., 2017; Frye and Blanco, 2016; Andreassi et al., 2018). Although m^6A is the most prevalent post-transcriptional modification studied to date, over 150 chemical modifications of mRNA have been identified (Boccaletto et al., 2018), including other types of methylation (for example m^1A and m^5C (Zhang and Jia, 2018; Adams and Cory, 1975)), geranylation (Dumelin et al., 2012), and pseudouridylation (Carlile et al., 2014).

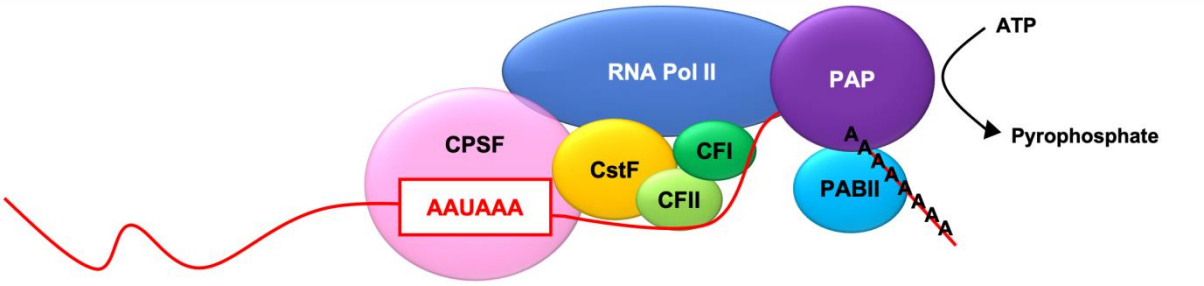


Figure 1.4: 3'-end processing or polyadenylation. Polyadenylation involves the addition of multiples adenine nucleotides (the poly(A) tail) to the 3'-end of the mRNA. First the mRNA is cleaved at the 3'-end by cleavage factors, including; CPSF (cleavage/polyadenylation specificity factor), CstF (cleavage stimulation factor), CFI (cleavage factor I), and CFII (cleavage factor II). CPSF recognises conserved RNA sequences near the 3'-end of the mRNA, and CstF and CFI increase the specificity of this interaction. CFII also appears to be involved in mRNA cleavage via an as-yet unknown mechanism. After cleavage, the poly(A) tail is added by PAP (polyadenylate polymerase). PABII (polyadenylate binding protein II) binds to short poly(A) sequences and increases the affinity of PAP to the mRNA. Both CPSF and CstF interact with Pol II and appear to have roles in signalling transcription termination.

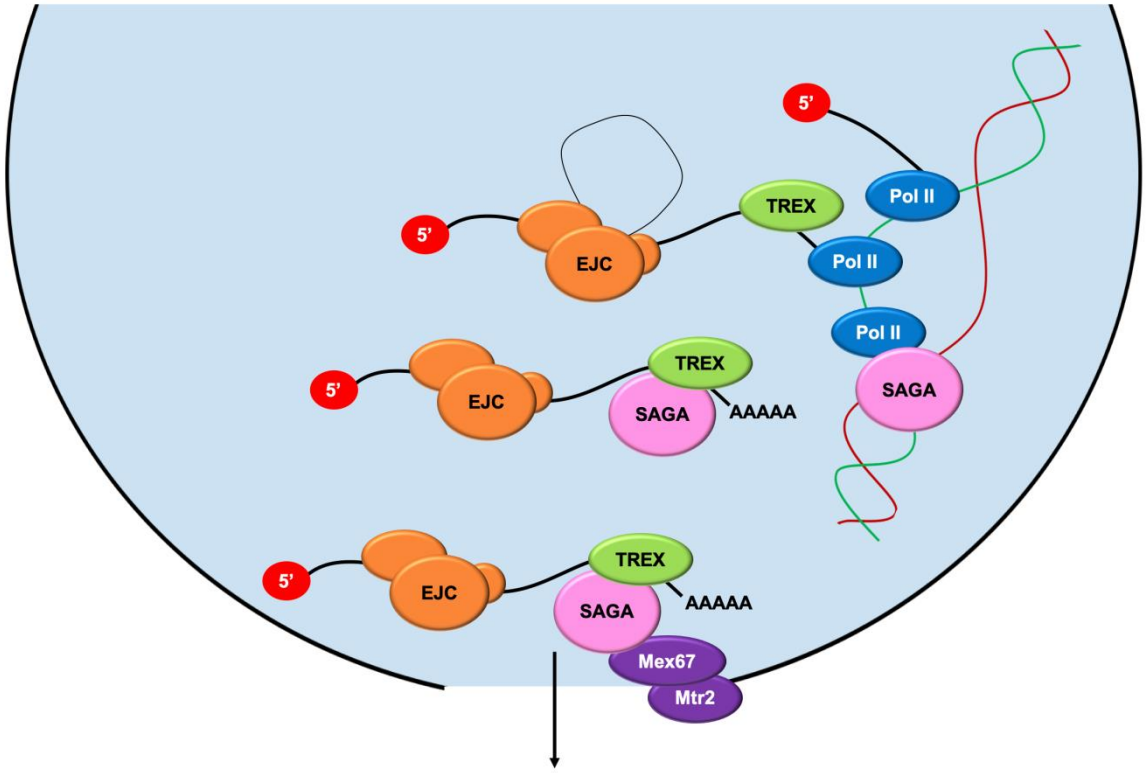


Figure 1.5: An overview of mRNA processing and export. mRNA processing involves three main steps; 5'-end capping, splicing and polyadenylation. These steps are mainly thought to occur co-transcriptionally, driven by protein interactions with Pol II. During processing, the pre-mRNA is bound by protein complexes including the TREX, SAGA and exon-junction complex (EJC). SAGA also acts as an activator of Pol II activity and, in yeast, both the TREX and SAGA complexes interact with nuclear export factors, including Mex67 and Mtr2, which regulate the interaction with the nuclear pore complex to allow mRNA nuclear export.

1.2.2. mRNA export

The proteins and RNA binding partners that mRNAs accumulate during their processing lead to the formation of an mRNP that is capable of exiting the nucleus. Many of these proteins are involved in mRNA export, and the formation of this mRNP is required for nuclear export (reviewed in Köhler and Hurt, 2007) (Figure 1.5). For example, the transcription-export (TREX) complex consists of multiple proteins that bind to mRNA during transcription and mRNA processing, and this association is dependent on 5'-end capping and splicing (Cheng et al., 2006; Masuda et al., 2005). In yeast, the TREX complex formation is particularly important as yeast possess many genes without introns and mRNAs may be processed and exported without splicing. Here, the TREX/THO complex has been proposed to couple gene transcription with the nuclear pore machinery in a process termed 'gene gating'. The TREX/THO complex has been shown to interact with SAGA, a protein complex and activator of Pol II activity (Daniel and Grant, 2007; Fischer et al., 2004), and the TREX/THO and SAGA complexes may interact with export factors such as Mex67-Mtr2 (Reed and Hurt, 2002) (Figure 1.6). These extensive interactions between multiple protein complexes and mRNA during transcription and processing allows both quality control of the mRNA, as if a processing step fails then the mRNA is likely to be unable to leave the nucleus and will be degraded, and also recruits proteins to form the mRNP that will determine the fate of the mRNA once exported.

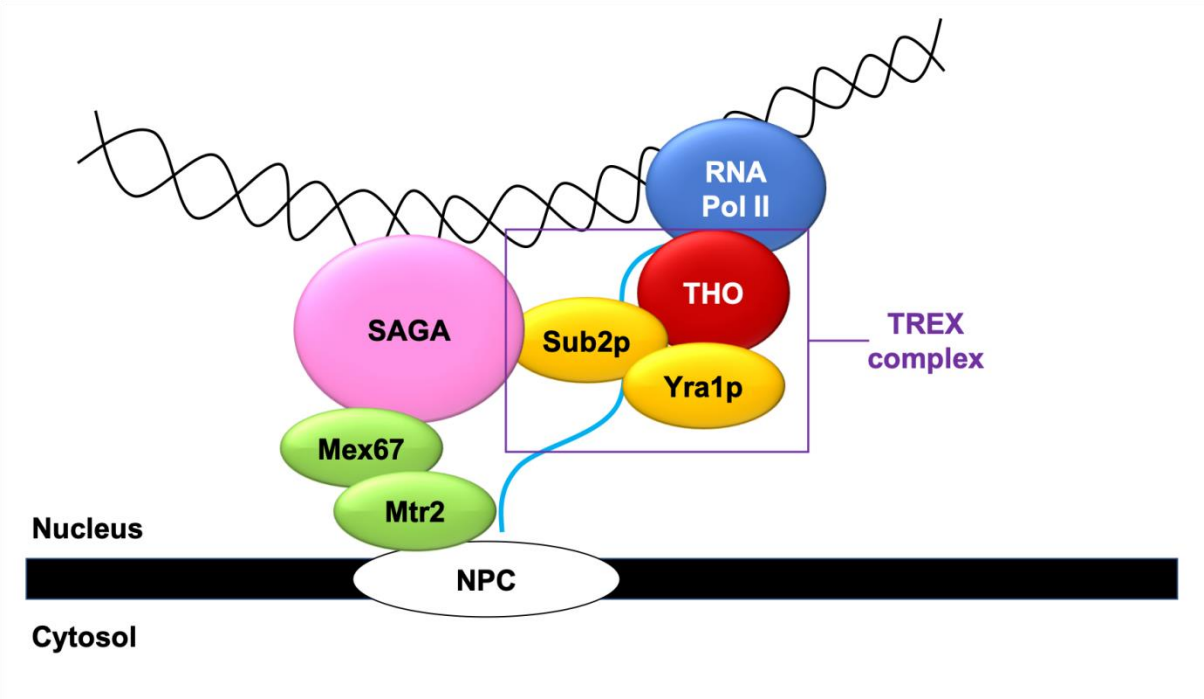


Figure 1.6: The TREX/THO complex in yeast. In yeast, the THO complex is made up of 5 protein subunits (6 in mammals), and interacts with both RNA Pol II and the mRNA being transcribed. Together with Sub2p and Yra1p (UAP56 and ALY in mammals, respectively), this forms the TREX/THO complex. The TREX/THO complex interacts with SAGA, which interacts with nuclear export factors Mex67 and Mtr2 (TAP and p15 in mammals, respectively) to help drive mRNA export.

1.2.3. mRNA fate and decay

Once exported from the nucleus as part of an mRNP, the mRNA may be directly translated, localised and/or stored, and then eventually degraded by the cell (Figure 1.7). Different mRNA fates are driven by RNA binding proteins (RBPs), which are determined by the sequence and structure of the mRNA, and the status of the cell as a whole. Many mRNAs are immediately translated upon interaction with ribosomes in the cytosol following a pioneer round of translation (reviewed in Maquat et al., 2010; Fortes et al., 2000) (Figure 1.8). During the pioneer round of translation, the progression of ribosomes along the mRNA ORF causes drop-off of the exon junction complexes that assemble during mRNA transcription (Ishigaki et al., 2001). If these complexes are not present (e.g. due to a premature stop codon) and not removed by ribosomes, the mRNA may be targeted for nonsense mediated mRNA decay (NMD) (Popp and Maquat, 2013). mRNAs translated immediately by cytosolic ribosomes include a large overall proportion of mRNAs (30-40% in human cells (Pyhtila et al., 2008)) that are destined for the endoplasmic reticulum (ER), to either function in the ER/Golgi pathway or to be secreted from the cell. The majority of these mRNAs are not directly regulated; their translation by the ribosome produces a signal peptide early on that is recognised by the signal recognition particle (SRP), which drives co-translational localisation to the ER membrane (reviewed in Akopian et al., 2013). mRNAs that are more highly regulated may be localised to RNA granules or to specific locations in the cell. mRNAs localised in this way are often translationally repressed until they reach their destination, since localised protein production is often the purpose of this localisation. mRNA localisation and RNA granules are discussed further in sections 1.3 and 1.4, respectively.

Finally, along with determining mRNA translation and localisation, the mRNA sequence and mRNP composition also contribute to determine the rate of degradation of the mRNA. mRNA decay is an important method for regulating protein production, and may also be triggered by aberrant or mutated mRNA sequences (Gatfield et al., 2003; Inada and Aiba, 2005), by widespread cellular stress, and by apparently random degradation events. mRNA decay is often linked to mRNA translation (reviewed in Garneau et al., 2007). The 5'-end cap of an mRNA may bind the translation initiation factor, eIF4E, and the 3' poly(A) tail may bind the poly(A) binding protein (PABP), where both binding events promote translation initiation and inhibit mRNA decay. mRNA degradation is most often initiated by deadenylation of the poly(A) tail. This can be followed either by 5' decapping then a 5' to 3' degradation via the exoribonuclease Xrn1, or by 3' to 5' degradation via the multi-protein exosome (Figure 1.9).

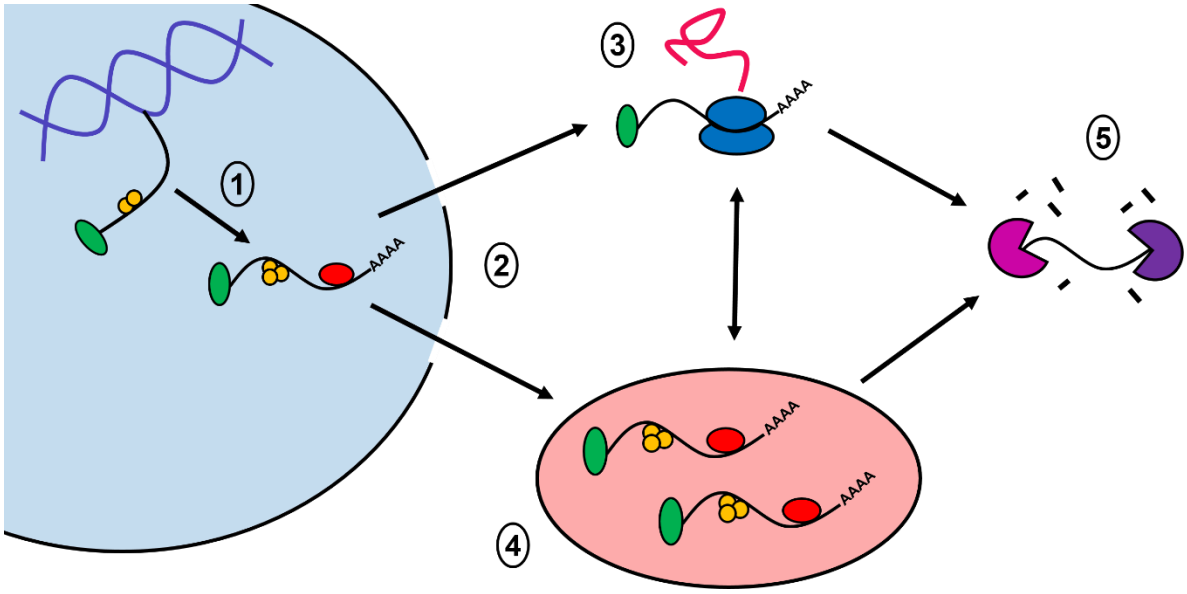


Figure 1.7: Overview of mRNA fate. mRNA fate is determined by the formation of mRNPs, which may target mRNA for translation, localisation, storage in cytosolic granules, and finally degradation. **(1)** mRNAs are bound by proteins cotranscriptionally to form mRNPs. Complex formation is necessary to allow mRNA processing and export from the nucleus. **(2)** mRNAs may then be translationally repressed until they are localised to specific sites in the cell. **(3)** mRNAs may be translated immediately upon entering the cytosol, or translated upon localisation and removal of translational repression. **(4)** mRNAs may be localised to granules to be stored until required for translation. These mRNAs may also be translationally repressed. (5) Finally, mRNAs are degraded by decapping enzymes and endonucleases.

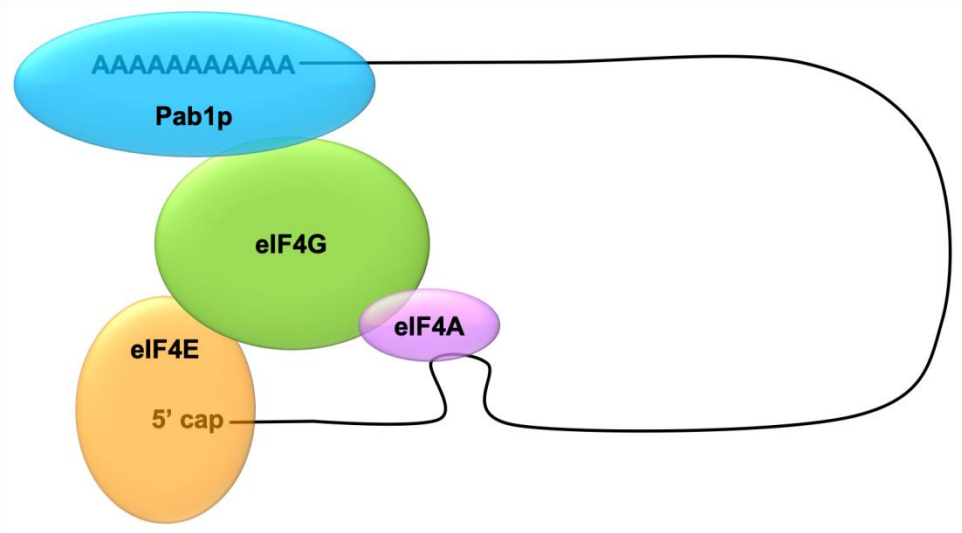


Figure 1.8: mRNA circularisation in yeast. mRNA circularisation is thought to enhance mRNA translation. Pab1p bound to the poly(A) tail and eIF4E bound to the 5' cap both interact with eIF4G to allow communication between the 5' and 3'-end of the mRNA. The interaction of eIF4A with both he mRNA and eIF4G are also suggested to increase mRNA translation efficiency.

There is some evidence that mRNA decay in eukaryotes can take place in P-bodies, which are cytoplasmic foci that contain high concentrations of specific, loosely associated proteins and RNA (Anderson and Kedersha, 2006). P-bodies have been shown to contain high levels of some 5' to 3' decay pathway proteins (Cougot et al., 2004; Sheth and Parker, 2003), and some cells contain P-bodies only under cellular stress, where the level of mRNA degradation is increased (Kedersha et al., 2005; Lui et al., 2014). However, the machinery for 5' to 3' decay is also present in the cytosol of cells and the exact role of P-bodies in mRNA decay is still unclear. mRNA decay may also occur by at least two other mechanisms; deadenylation-independent decapping (Muhlrad and Parker, 2005) and endoribonucleolytic decay (Tomecki and Dziembowski, 2010; Liu et al., 2004) (Figure 1.9).

It has been estimated that up to 20-50% of changes to mRNA levels in response to cellular signalling can be regulated by mRNA stability (Schwanhäusser et al., 2011). mRNA stability is predominantly controlled by the 3'UTR (the regulatory region of the mRNA) which subsequently effects the formation and activity of the mRNP. mRNPs may then act to increase or decrease the decay rate of the mRNA. For example, mRNA 3'UTRs may contain an AU-rich element (ARE). AREs may be different lengths and sequences, and can be affected by their flanking sequences to confer different levels of mRNA stability (Khabar, 2017). During mRNA decay, degradation machinery may bind directly to the ARE (such as some exosome components) (Otsuka et al., 2019), or be recruited to the ARE via other RBPs (such as AUF1 and TTP) (Chen et al., 2001; Lykke-Andersen and Wagner, 2005). An important advantage of this modulation of RNA stability by binding proteins is that the mRNA may be rapidly targeted for degradation to quickly reduce protein expression. Overall, the sequence and structure of a mature mRNA allows it to interact with multiple proteins and other RNAs in complex ways to allow mRNAs to fulfil their function of producing protein in a regulated manner.

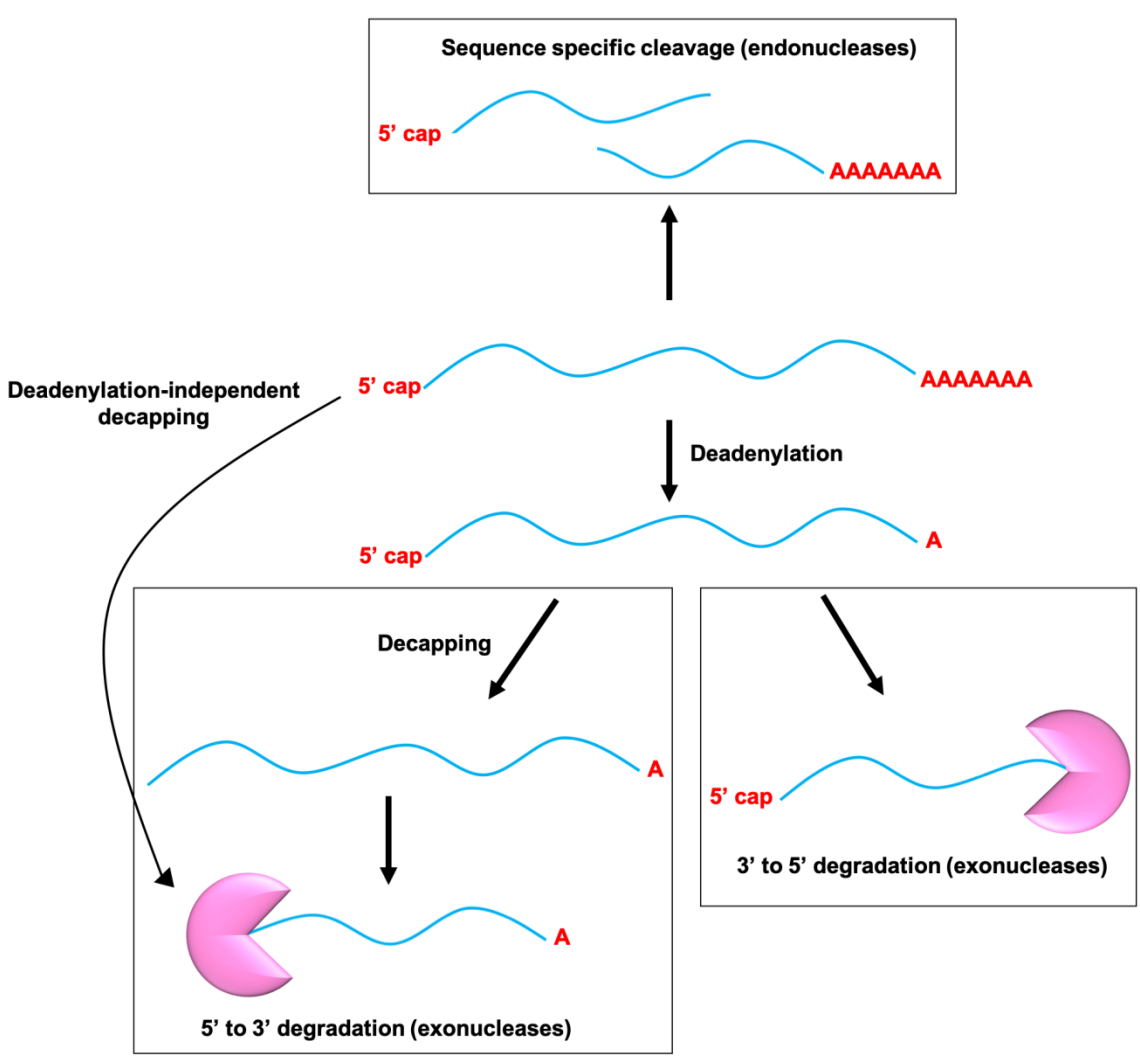


Figure 1.9: An overview of mRNA decay mechanisms. Global mRNA degradation most often occurs following deadenylation of the poly(A) tail. mRNA degradation can then continue in one of two ways; either by decapping and 5' to 3' degradation, or by 3' to 5' degradation. mRNA degradation can also occur via deadenylation-independent decapping. This is also termed nonsense mediated decay (NMD), and primarily occurs when the mRNA contains a premature stop codon. For individual mRNAs, degradation can also be catalysed by endonucleases that cleave the mRNA at specific sequences, allowing 5' to 3' and 3' to 5' degradation to continue without deadenylation/decapping.

1.3. mRNA localisation

mRNA localisation is the process by which the position of an mRNA is regulated in space and time inside the cell to regulate protein production. mRNA localisation is a conserved mechanism that occurs in all eukaryotes (reviewed in Parton et al., 2014; Martin and Ephrussi, 2009), and more recently examples have also been found in prokaryotes (reviewed in Fei and Sharma, 2018; Irastortza-Olaziregi and Amster-Choder, 2020). Although many cell types localise mRNA, the exact mechanisms involved in this localisation can vary widely depending on the encoded protein, cell type and cell status (e.g. developing cells or stressed cells). Generally, mRNA localisation first requires the presence of a localisation signal in the mRNA sequence that targets the mRNA for localisation. This localisation signal may be recognised by binding proteins which then drive the localisation of the mRNA, often involving motor proteins and the cytoskeleton. During transit, the mRNA is often translationally repressed to prevent protein production at unwanted times and places. Finally, tethering of the mRNA at its destination may occur, where translation becomes derepressed to allow protein production.

1.3.1. Examples and functions of mRNA localisation

A major function of mRNA localisation is to restrict protein production and activity to specific places in the cell. Along with localised protein production, mRNA localisation also has roles in allowing rapid responses to signalling, more energy efficient production of protein, and more efficient formation of protein complexes.

Some early examples of mRNA localisation can be found in *Drosophila*, where mRNA localisation and protein gradients are important during embryonic development. Perhaps the most extensively studied instances of mRNA localisation are that of the mRNAs *bicoid* and *oskar*. During oocyte development, *bicoid* mRNA is localised to the anterior pole, and *oskar* to the posterior pole (Berleth et al., 1988; Pokyrwka and Stephenson, 1991; Clark et al., 1994; Chang et al., 2011). This localisation results in the formation of two protein gradients; Bicoid protein at the anterior and Oskar protein at the posterior. This localised protein production eventually leads to the head of the fly forming at the anterior end and the abdomen forming at the posterior end of the egg (Driever and Nüsslein-Volhard, 1988; Ephrussi and Lehmann, 1992; Lasko, 2012). Defects in correct mRNA localisation during *Drosophila* development can lead to aberrant body formation and can be lethal (Kuchinke et al., 1998; Norvell et al., 2005).

mRNA localisation leading to localised protein production is also essential for the correct function of many adult cells. In particular cells that undergo migration, such as mammalian fibroblast cells, rely on the formation of protein gradients to ensure correct directionality of movement. In fibroblasts, cell migration is driven by the formation of protrusions at the leading edge of cells, followed by cell polarisation and detachment at the rear of the cell to allow translocation. Many proteins are therefore specifically concentrated at the poles of the cell. For example, β-actin and Arp2/3 are localised at the leading edge (Lawerence and Singer, 1986; Welch et al., 1997) and myosin IIB at the rear (Choi et al., 2008), and this localised protein activity is driven by localisation of the encoding mRNAs (Mingle et al., 2005; Kislauskis et al., 1994).

mRNA localisation also allows fast responses to cellular signalling, as protein production can be regulated more rapidly from mRNAs without the need for new transcription and localisation to occur. For example, neuronal cells rely on mRNA localisation to function correctly. This is due to the length of neurons; the nucleus, the site of transcription, is located in the cell body far from the active synapse of the cell. Therefore, many mRNAs are held in translationally repressed states along the axon and within the dendrites (Merianda et al., 2009; Yan et al., 2009; Biever et al., 2020). When the cell receives a signal, these mRNAs can be rapidly translated into proteins in response to the signal, allowing neuronal cells to respond far more quickly than they otherwise would be able to.

mRNA localisation to dendrites also highlights another advantage; that it is far more energy efficient to localise one mRNA than to localise every protein translated from it. It has been estimated that on average, thousands of proteins are translated from each mRNA; in yeast, this number is suggested to be up to 20,000 proteins per mRNA (Ghaemmaghami et al., 2003). Particularly in neuronal cells, where the proteins would have to travel (relatively) large distances to reach their destination (likely reaching the loading capacity of the cytoskeleton), localising a single mRNA is far more cost effective in terms of energy use.

Finally, recent studies have shown that mRNA localisation may be important in the co-translational formation of multi-subunit protein complexes (Schwarz and Beck, 2019). Co-translational assembly of protein complexes has been shown to promote assembly by limiting non-specific interactions between protein subunits and increasing the frequency of interaction between subunits in a crowded environment (Shieh et al., 2015; Wells et al., 2016). For example, it has been shown that *TAF8* mRNA specifically localises to TAF10 protein subunits to promote co-translational assembly of the TFIID complex in mammalian cells (Kamenova et al., 2019). Both mRNA localisation to protein complex subunits and co-

localisation of mRNAs that are actively translated may therefore drive protein complex formation.

Overall, these examples demonstrate that mRNA localisation has many varied functions in many different cell types. This is reflected in the multiple different mechanisms and proteins that may be involved in driving mRNA localisation, although many mRNAs are localised via generally similar methods.

1.3.2. Mechanisms of mRNA localisation

The eventual localisation of an mRNA is often pre-determined during mRNA transcription and processing by the protein contingent the mRNA interacts with, before the mRNA even reaches the cytosol (discussed more in section 1.2). During mRNA biogenesis, the mRNA is bound by multiple RNA binding proteins (RBPs) that form the mRNA-protein complex (mRNP). This mRNP often contains proteins that drive localisation of the mRNA once it reaches the cytosol, and/or proteins that translationally repress the mRNA until it reaches its destination (Figure 1.10). Conversely, for some mRNAs the proteins that carry out these functions reside in the cytosol and interact with the mRNA following nuclear export, or in a signal-dependent manner when certain conditions are met in the cell. RBPs then may play further roles in mRNA localisation, such as facilitating mRNA transport, tethering, and removal of translation repression.

Localised mRNAs contain *cis*-acting localisation elements (or zip codes), which are recognised by *trans*-acting recognition factors (usually RBPs). Zip codes are specific sequences found in the mRNA that promote binding to the correct recognition factor to drive localisation. Zip codes may be primary sequences that are directly recognised by binding proteins (for example both short and repeat nucleotide sequences in *vg1* mRNA in *Xenopus* (Mowry and Melton, 1992; Zhao et al., 2001)), or sequences that form secondary RNA structures (for example stem loop structures in the yeast *ASH1* mRNA (Gonzalez et al., 1999; Chartrand et al., 1999)). *ASH1* mRNA also displays two other possible features of zip codes; the mRNA contains more than one zip code and contains zip codes in both the ORF and 3'UTR, whereas often localisation signals are found only in the 3'UTR (Gonzalez et al., 1999; Chartrand et al., 1999). Multiple zip codes in one transcript can be redundant; any of the four zip codes present in *ASH1* is sufficient to drive mRNA localisation (Gonzalez et al., 1999). However mRNAs bearing these isolated sequences do not remain localised in the

same way as the full-length transcript, suggesting the multiple localisation signals may be required for tethering the mRNA at its destination (Chartrand et al., 1999; Jansen, 2001).

Zip code sequences and structures are then recognised by proteins that often form part of an mRNP. Unsurprisingly, given the multiple functions and mechanisms of mRNA localisation, many different proteins have been identified. It has been shown that some of these proteins can bind multiple mRNAs, and also some of these proteins are conserved in different organisms. For example, in *Drosophila*, the protein Staufen has been shown to bind and drive the localisation of *bicoid*, *oskar* and *prospero* mRNAs (Roegiers and Jan, 2000). Staufen contains five double-stranded RNA binding domains (dsRBDs) that are capable of binding non-specifically to mRNAs (St Johnston et al., 1992). Staufen specificity has therefore been suggested to be elicited by mRNA zip codes that alone are not sufficient to drive localisation (Ferrandon et al., 1994). Staufen also has homologues in mammalian cells, where the Staufen homologue in rats has been shown to associate with neuronal mRNA granules and be localised within dendrites (Köhrmann et al., 1999). Additionally, the well-characterised zip code binding protein 1 (ZBP1) has been shown to bind *β-actin* mRNA in fibroblasts (Hüttelmaier et al., 2005), while its *Xenopus* homologue Vg1RBP binds the zip code of *vg1* mRNA (Deshler et al., 1998). Interestingly, ZBP1 and Vg1RBP share a 78% structural similarity (Havin et al., 1998) while their respective zip codes are dissimilar (Jansen, 2001), although they both drive mRNA localisation by facilitating interaction between the mRNA and microtubules (Elisha et al., 1995; Song et al., 2015). Indeed, although zip code binding proteins may play roles in translational repression and anchoring, their main function is often to link the mRNAs to motor protein machinery to allow transport along the cytoskeleton.

Some of the most well studied mechanisms of mRNA transport involve mRNA association with motor proteins and subsequent directed transport along microtubules or actin filaments. Examples of alternative mechanisms for mRNA transport such as mRNA diffusion have also been identified, and are discussed more below. Transport of mRNAs facilitated by association with motor proteins occurs in multiple species, including *ASH1* transport via the She2/She3/Myo4 complex along actin cables in yeast (Bertrand et al., 1998) (Figure 1.10), and β-actin mRNA transport along microtubules via dynein and kinesin in neuronal cells (Leung et al., 2018; Turner-Bridger et al., 2020). There is also evidence for multiple mRNAs being transported as large multi-mRNP complexes, or RNA granules (Chekulaeva et al., 2006; Lange et al., 2008; Segal et al., 2020). During this transport process, localised mRNAs are often translationally repressed to restrict protein production to the destination site.

Specific mechanisms of mRNA translational repression again differ depending on the mRNA and system, although they generally involve the binding of proteins that form part of the mRNP (Figure 1.10). The protein may act to physically prevent the interaction of the mRNA with the translation machinery, as in the case of eIF4E binding proteins (4E-BPs). 4E-BPs interact with repressed mRNAs and compete with eIF4G for the binding of eIF4E, inhibiting formation of the eIF4F complex and translation initiation (Richter and Sonenburg, 2005). Translation repression may then be relieved on the mRNA by dissociation of repressor proteins, to either allow localised protein production or in response to specific signalling (Besse and Ephrussi, 2008; Martin and Ephrussi, 2009). Translational repression is particularly important during development in *Drosophila*, where around 96% of *nanos* mRNA is not localised and is translationally repressed (Bergsten and Gavis, 1999; Trcek et al., 2015), and subsequently degraded early in oocyte formation (Dahanukar and Wharton, 1996; Bashirullah et al., 1999). Repression of *nanos* requires the binding of the protein Smaug to Smaug recognition elements (SREs) in the *nanos* 3'UTR sequence (Dahanukar and Wharton, 1996; Smibert et al., 1999), and Smaug also drives degradation of *nanos* by recruiting the CCR4-NOT deadenylase complex (Jeske et al., 2006; Zaessinger et al., 2006). Translational repression of *nanos* mRNA is relieved when *nanos* is localised to the posterior pole of oocytes, where the presence of the Oskar protein is suggested to promote *nanos* derepression by directly interacting with the RNA binding domain of Smaug (Dahanukar et al., 1999). Production of the Nanos protein is therefore restricted to the posterior pole, even though the majority of its mRNA is not.

For mRNAs that are translated once they reach their destination, tethering may occur to maintain the position of the mRNA. Again, there are multiple methods that different mRNAs utilise; most involve mRNA anchoring to specific proteins on the cytoskeleton. For example, *ASH1* mRNA is anchored at the daughter bud tip after localisation, to ensure the daughter cell only receives the mRNA (Bertrand et al., 1998) (Figure 1.10). As mentioned above, *ASH1* anchoring relies on multiple zip codes in its mRNA sequence (Chartrand et al., 1999; Jansen, 2001), and the endogenous interactions of its localisation machinery She2/She3/Myo4 with the cytoskeleton to remain anchored (Bertrand et al., 1998; Palacios, 2007). In the *Drosophila* oocyte/embryo, it has been shown that multiple mRNAs appear to be anchored to either the cytoskeleton (*oskar* and *gurken* mRNAs (Delanoue et al., 2007; Vanzo and Ephrussi, 2002)) or to the anterior pole (*bicoid* mRNA (Trovisco et al., 2016)) following mRNA localisation, although many of the exact proteins and interactions involved in mRNA anchoring are not yet well understood.

Finally, there are other, more uncommon methods of mRNA localisation that do not involve one or more of the discussed key steps. In particular, mRNA localisation may occur without transport along the cytoskeleton, but via mRNA diffusion and trapping. For example in *Drosophila*, *nanos* mRNA localisation does not appear to involve the cytoskeleton, but is driven by mRNA diffusion, widespread mRNA degradation, and protection of *nanos* at the posterior pole (Forrest et al., 2003; Weil et al., 2006; Kugler and Lasko, 2009). A similar mechanism involving general mRNA degradation and localised protection has been shown to drive localisation of *hsp80* mRNA to the posterior pole in *Drosophila* oocytes (Ding et al., 1993). Therefore, although most mRNA localisation involves the steps of zip code recognition and mRNP formation, translational repression and cytoskeletal transport, and mRNA tethering and translation derepression, the large variety of different mRNAs, proteins and processes involved means there is no one conserved mechanism for mRNA localisation. Nevertheless, the necessity for mRNA localisation to provide both spatial and temporal regulation of protein production makes the continued study of novel mRNA localisation pathways worthwhile, especially since several proteins involved mRNA localisation have connections with human disease.

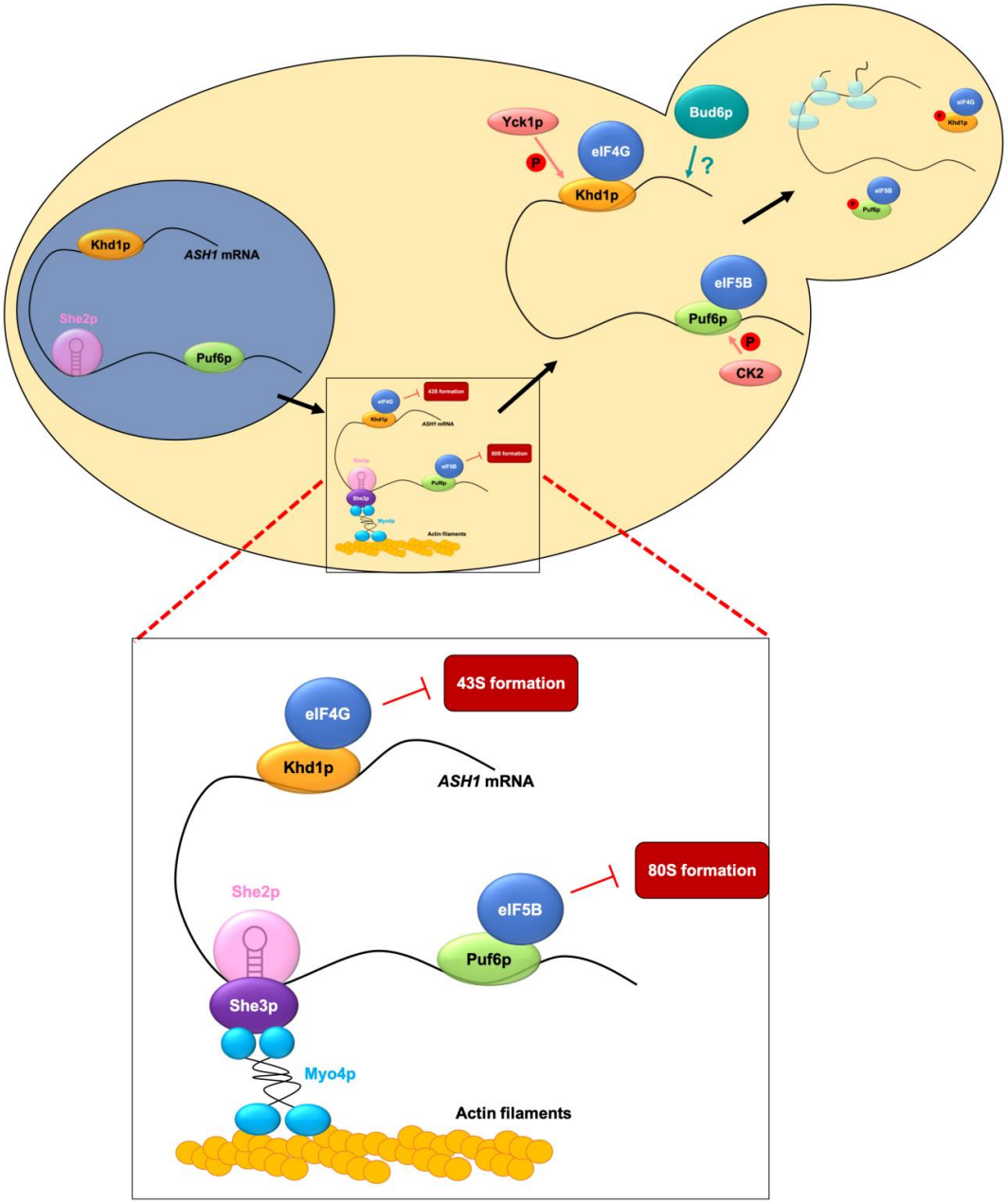


Figure 1.10: *ASH1* mRNA localisation in yeast. The localisation of *ASH1* mRNA in yeast highlights the four most common steps of mRNA localisation: zip code recognition and mRNP formation, mRNA association with motor proteins and mRNA transport, translational repression of mRNA during transport, and mRNA tethering and translational depression following mRNA localisation. Proteins important for mRNA transport and translational repression form part of the *ASH1* mRNP prior to mRNA export, including She2p, Khd1p and Puf6p. She2p associates with the motor protein Myo4p via interactions with She3p, allowing *ASH1* mRNA transport along actin filaments. *ASH1* mRNA is translationally repressed via at least two mechanisms. The first is the interaction between Khd1p and eIF4G, which sequesters the eIF4G and prevents the formation of the 43S pre-initiation complex. The second is the interaction between Puf6p and eIF5B, which inhibits the formation of the 80S ribosomal subunit. When *ASH1* mRNA reaches its destination at the bud, both Khd1p and Puf6p are phosphorylated, which is thought to reduce the affinity of both proteins for mRNA and cause their dissociation. Ribosomes may then bind and translate *ASH1* mRNA. Tethering of *ASH1* mRNA has been shown to involve at least one bud-associated protein, Bud6p, although the exact interaction between Bud6p and the *ASH1* mRNA is as yet unknown.

1.4. RNA granules

Eukaryotic cells contain multiple distinct compartments, which helps to regulate cellular functions and protect cell components. Yeast cells of the species *Saccharomyces cerevisiae* share many of these compartments with most other eukaryotic cell types, including both membrane-bound organelles and membrane-less RNA granules (Figure 1.11). These shared organelles include: the nucleus, where DNA is stored to aid the regulation of gene expression and to protect the DNA from damage; the endoplasmic reticulum and Golgi apparatus, which are involved in protein synthesis, folding, and secretion from the cell; and the mitochondria, which are sites of respiration and ATP production. Yeast cells also possess a vacuole, which functions to maintain PH and ion homeostasis in the cell, and also may be involved in protein and organelle degradation.

RNA granules are bodies found in the cytosol and organelles of cells, and contain high concentrations of specific RNAs and proteins. There are many types of RNA granules with different compositions and functions, found in different species and cell types. Generally, the function of these granules is to regulate RNA storage, translation, and degradation, where inhibition of granule formation can lead to aberrant RNA regulation. Although imaging evidence for RNA granules was found as early as 1865 (Metschnikoff, 1865), the exact purpose, components, and dynamics of many types of RNA granules is still unknown. Despite this, several driving factors of RNA granule formation are conserved and many new studies are emerging which provide insights into RNA granule regulation and dynamics.

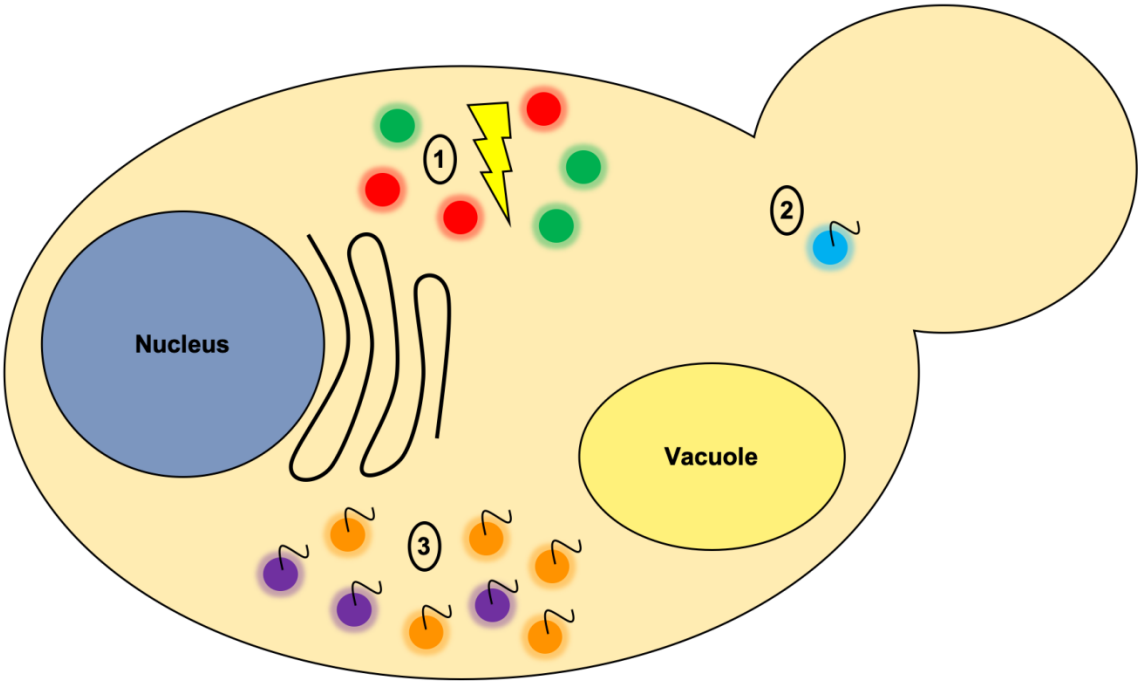


Figure 1.11: RNA granules in yeast. Yeast cells of the species *Saccharomyces cerevisiae* contain cell compartments including the nucleus, vacuole, endoplasmic reticulum, and multiple types of cytosolic RNA granules. **(1)** Stress granules and P bodies both form in response to cellular stress and have been shown to be present in between 2-20 granules per cell depending on cell conditions. The mRNAs they contain are translationally repressed. **(2)** Translation factor mRNA granules are present in unstressed yeast cells. They are present in 1-2 granules per cell and the mRNAs they contain are actively translated in unstressed conditions. **(3)** Glycolytic mRNA granules are also present in unstressed yeast cells. They are present in 10-20 granules per cell and the mRNAs they contain are actively translated in unstressed conditions.

1.4.1. Liquid-liquid phase separation and intrinsically disordered proteins

RNA granules have been suggested to form via low affinity interactions including; RNA-RNA, RNA-protein, and protein-protein interactions. The association of many of these components into large, multivalent structures is known as liquid-liquid phase separation (LLPS), where the granule is distinct from the surrounding cytosol but does not possess a lipid membrane. These RNA granules have been shown to be highly dynamic, where proteins and RNAs are shuttled between them and the surrounding cytosol throughout their lifespan. Within the granule, protein components may interact specifically with each other and with RNA to provide structure and functionality. Alternatively, intrinsically disordered regions (IDRs) within proteins may promote phase separation via relatively non-specific interactions with multiple proteins and RNAs.

For instance, well-studied examples of proteins that are capable of undergoing LLPS include the RNA binding protein FUS (Patel et al., 2015; Monahan et al., 2017), and the P granule protein LAF-1 (Elbaum-Garfinkle et al., 2015; Schuster et al., 2018). FUS is a prion-like protein that contains multiple IDRs and has been shown to condense into liquid-like compartments both *in vitro* and *in vivo* (Alberti et al., 2015). These liquid-like compartments may condense further into protein aggregations, which are associated with the neurodegenerative disease ALS (Alberti et al., 2015; Fawzi et al., 2017). LAF-1 is an RNA helicase that contains a disordered, arginine/glycine rich (RGG) domain. This RGG domain is necessary for LAF-1 to undergo LLPS (Brangwynne et al., 2015), and it has been shown that increasing numbers of RGG domains leads to phase separation more readily *in vitro* (both at lower temperatures and more quickly) (Hammer et al., 2018).

In general, it has been shown that proteins containing IDRs are more prevalent in RNA granules (Reijns et al., 2008; Kato et al.,2012; Jain et al., 2016), and that high concentrations of IDRs can drive LLPS of RNA granule proteins *in vitro* (Lin et al., 2015; Molliex et al., 2015; Patel et al., 2015; Zhang et al., 2015). To facilitate the formation of RNA granules via LLPS, while providing specificity of components and function, it is likely that both specific protein-protein and protein-RNA interactions are required along with mass interactions of protein IDRs (Protter et al., 2018). Examples of both these interactions can be demonstrated for different types of RNA granules.

1.4.2. Stress granules and P-bodies

Perhaps the most extensively studied examples of cytoplasmic RNA granules are stress granules (SGs) and processing bodies, or P-bodies (PBs). In eukaryotes, both granules form in response to stress, and PBs may also be constitutively present (Kedersha et al., 2005; Ohn et al., 2008). It was initially thought that SGs represent sites of mRNA storage and PBs were associated with mRNA decay, whereas more recent studies have shown that many proteins are shared between the two granules and both appear to play roles in storage. While the formation, composition, and dynamics of these granules have been well-studied, the exact functions of these RNA granules are therefore still unclear.

SG formation occurs following cellular stress. In response to stress, the translation initiation process is inhibited and any elongating ribosomes 'run-off' the mRNAs (Advani and Idanov, 2019). As a result mRNAs carrying stalled pre-initiation complexes accumulate and are bound by RBPs to target their localisation to stress granules (Kedersha et al., 2005; Sanders et al., 2020). Overexpression of some of these RBPs, for example the RasGAP-associated endoribonuclease G3BP, can drive the formation of SGs in unstressed cells (Tourrière et al., 2003). Proteins that contain IDRs (Molliex et al., 2015; Youn et al., 2018) and RNAs that can self-assemble (Van Treeck et al., 2018) also contribute to the formation of SGs.

SGs contain many proteins that are important for mRNA regulation, including regulators of mRNA processing, translation, and stability (Markmiller et al., 2018; Youn et al., 2018). SGs have also been shown to contain proteins found in the pre-initiation complex, for example ribosomal subunits, initiation factors, and PABP (Kedersha et al., 2002; Kedersha et al., 2005). SG composition is also different between cell types, and may change in response to stress (Fujimura et al., 2012; Aulas et al., 2018; Markmiller et al., 2018). Generally, the mRNAs found in SGs are translationally repressed (Kedersha et al., 1999; Souquere et al., 2009) and mRNA targeting appears widespread and non-specific (Khong et al., 2017). Exceptions are some heat shock proteins (HSP70 and HSP90), which have been shown to be excluded from SGs (Stöhr et al., 2006), and some evidence of low levels of mRNA translation in stress granules under certain conditions (Mateju et al., 2020). Many proteins are present in both SGs and PBs, discussed more below.

In *Saccharomyces cerevisiae*, SG formation has been suggested to be promoted by the presence of PBs, where inhibition of PB formation may result in inhibition of SG formation (Buchan et al., 2008) (Figure 1.11). Yeast SGs were first identified as bodies separate from PBs that contained high concentrations of eIF4E, eIF4G and Pab1p (unlike mammalian SGs) (Hoyle et al., 2007), where these so-called EGP bodies do not associate with mRNA

decay proteins indicative of PBs (Hoyle et al., 2007; Buchan et al., 2008). The composition of SGs appears to depend on the type of cellular stress, where the SGs that form in response to severe stress (e.g. heat shock rather than nutrient deprivation) possess proteins more similar to mammalian SGs, including the 40S ribosomal subunit and eIF3 (Groušl et al., 2009; Groušl et al., 2013). As in mammalian cells, translationally repressed mRNAs may leave yeast SGs and return to the translation pool upon alleviation of cellular stress (Buchan et al., 2008; Buchan et al., 2011).

PB formation is less well understood than SG formation. During stress, PB formation may be driven by the increased number of translationally stalled mRNAs as in SGs (Kedersha et al., 2005), though this does not account for constitutive PB formation. Proteins containing IDRs are enriched in PBs (Jonas and Izaurralde, 2013; Lin et al., 2015; Decker et al., 2007), although thus far only three proteins have been shown to be necessary for PB formation in mammals (Ayache et al., 2015; Ohn et al., 2008). Of note, both the helicase activity of DEAD-box helicase 6 (DDX6) in mammals (Ayache et al., 2015; Minshall et al., 2009), and the ATPase and scaffold activity of DEAD-box ATPase Dhh1 in yeast (Mugler et al., 2016) have been shown to be required for PB assembly. The helicase activity of other helicases, for example the DEAD-box protein eIF4A, have been shown to inhibit SG formation by interacting with RNAs and preventing other RNA-protein and RNA-RNA interactions that drive granule formation (Tauber et al., 2020). These differing functions of similar proteins highlight the importance of RNA interactions in the formation of RNA granules and the multiple levels of regulation required for granule assembly.

PBs contain multiple proteins required for mRNA degradation, including the CCR4/NOT deadenylase complex (Cougot et al., 2004; Andrei et al., 2005; Parker and Sheth, 2007) and decapping enzymes (van Dijk et al., 2002; Eulalio et al., 2007). Along with the lack of PABP found in PBs (Kedersha et al., 2005), which might indicate an absence of mRNA poly(A) tails, these results were initially thought to support the hypothesis that PBs are sites of mRNA decay. However, more recent studies have shown that mRNAs enriched in PBs do not possess shorter poly(A) tails than cytosolic mRNAs (Hubstenberger et al., 2017), and the absence of PBs does not inhibit mRNA degradation (Franks and Lykke-Andersen, 2008; Pitchiaya et al., 2019). Therefore, the exact role of PBs in mRNA decay is still being investigated. Like SGs, PBs contain translationally repressed mRNAs, though PBs show more selectivity for mRNAs that have a low GC content and codons associated with lower rates of translation (Hubstenberger et al., 2017; Courel et al., 2019). These poorly translated mRNAs are enriched in PBs when the cell is unstressed, suggesting a role in mRNA storage (Matheny et al., 2019; Riggs et al., 2020).

In yeast, PBs are absent in unstressed cells and appear to rapidly form in cells in response to stress (Ramachandran et al., 2011; Lui et al., 2014) (Figure 1.11). It appears that the purpose of PBs in yeast is similar to mammalian cells, where PBs have been associated with both mRNA decay and mRNA storage (Aizer et al., 2014). Many proteins associated with decapping and mRNA degradation have been identified in yeast PBs (Sheth and Parker, 2003; Cougot et al., 2004; Parker and Sheth, 2007), and it has also been shown that mRNA decay intermediates accumulate in yeast PBs (Nissan and Parker, 2008). Along with evidence that mRNAs can leave PBs and re-enter the translation pool following cell recovery from stress (Brengues et al., 2005; Bhattacharyya et al., 2006; Arribere et al., 2011), it appears that PBs may play dual roles both in mRNA decay and storage in yeast following stress.

It has been suggested that SGs and PBs are associated with at least 144 shared proteins (Youn et al., 2018), including the translation factor eIF4E (Kedersha et al., 2005) and many proteins that contain IDRs (Youn et al., 2018; Kershaw et al., 2020). Since the mRNAs in both granules are translationally repressed, it is unsurprising that they both also contain RNAs and proteins associated with repression, including miRNAs and Argonaute proteins (Liu et al., 2005; Sen and Blau, 2005). SG and PBs may also both play roles in mRNA storage. As mentioned above, PBs contain poorly translated mRNAs (Matheny et al., 2019), and are also able to release mRNAs back into the cytosol to re-bind polysomes and undergo translation (Brengues et al., 2005; Bhattacharyya et al., 2006; Hubstenberger et al., 2017). SGs may function similarly to temporarily sequester mRNAs away from cytosolic ribosome pools during stress to promote translation of mRNAs encoding stress response genes such as HSP70 and HSP90 (Kedersha and Anderson, 2002). If the cell successfully recovers from stress, SGs then may disassemble and release their mRNAs to undergo translation. Therefore, although the exact functions of SGs and PBs are still unclear, their conserved presence throughout eukaryotes and described roles in the regulation of mRNA translation and the cellular response to stress mean that further study of these RNA granules will be important in the future.

1.4.3. Other types of RNA granules

Some further examples of relatively well-characterised RNA granules include germ granules, mitochondrial RNA granules, nuclear RNA granules, and actively translating granules in yeast. The protein and mRNA compositions of these granules are often very different, although some protein homologues are conserved. The formation of these RNA granules requires specific protein-protein and protein-RNA interactions, along with the interactions of protein IDRs (Protter et al., 2018). Highlighted here are some features of these RNA granules that are similar or distinct from one another, particularly with regards to granule formation.

Germ granules are early examples of RNA granules found in germ line cells. They have been shown to play roles in mRNA regulation and have been identified in many species; P granules in *C. elegans* (Seydoux, 2018), polar granules in *Drosophila* (Jones and Macdonald, 2007; Thomson et al., 2008), and Balbiani bodies in *Xenopus* oocytes (Boke and Mitchison, 2017). P granules are localised to the perinucleus in *C.elegans* and have been suggested to interact with mRNAs as they are exported from the nucleus (Pitt et al., 2000; Sheth et al., 2010). Their formation involves the specific self-interactions of the protein homologues PGL-1 and PGL-3, which also contain RNA-binding domains and may therefore provide the scaffold for the P granule to form (Hanazawa et al., 2011; Aoki et al., 2016; Kawasaki et al., 1998). Balbiani bodies in *Xenopus* are relatively large RNA granules that contain mitochondria, ER, Golgi, and RNA, but do not possess a lipid bilayer. The main component of the Balbiani body is the protein Xvelo, which forms micron-scale networks *in vitro* and is a highly disordered protein (Boke et al., 2016). Other Balbiani body organisers have predicted prion-like domains and RNA-binding domains (Malinovska et al., 2013).

Mitochondrial RNA granules (MRGs) form in unstressed cells. MRGs contain maturation factors and are associated with mRNA processing, where they can interact with immature mRNAs transcribed from mitochondrial DNA (mtDNA) within mitochondrial nucleoids. As yet, no evidence has been found of low complexity domains in MRG associated proteins. However, it has been shown that just removing the mRNA can cause MRG disassembly, and so granule formation may be driven by an RNA-binding scaffold protein (Iborra et al., 2004; Jourdain et al., 2013). Alternatively, since mitochondrial nucleoids assemble via mass binding of the transcription factor A (TFAM) to mitochondrial DNA (mtDNA) (Kukat et al., 2015), a similar scaffolding process may occur in MRG formation. MRGs have been found in mouse and human cells, though not in yeast or green plants (which have larger mitochondrial genomes with multiple promoters and introns, and process mitochondrial mRNA differently (Jourdain et al., 2016)). In *Drosophila*, the mtDNA structure is more similar

to mammalian mtDNA, and *Drosophila* mitochondria contain proteins related to MRG components, and so may possibly possess MRGs that have not yet been identified.

Nuclear RNA granules include Cajal bodies (Morris, 2008) and nuclear stress bodies (Biamonti and Vourc'h, 2010). The Cajal body, or coiled body, is the initial processing site for multiple U snRNPs that eventually form the spliceosome (Staněk et al., 2008), and is identified by the presence of the protein coilin. Similar to the nucleolus, its mechanism of formation is unknown, although it has been suggested that protein interactions with histones are important (but not necessary) for its formation and for efficient snRNPs processing (Wang et al., 2016; Sawyer et al., 2016). Nuclear stress bodies (NSBs) form in response to heat shock and are thought to regulate gene transcription as part of the stress response (Chiodi et al., 2000; Aly et al., 2019). They are formed following the activation of the heat shock response protein HSF1, which promotes the transcription of the long non-coding RNA HSATIII (highly repetitive satellite III). Both accumulation of HSF1 and HSATIII act to nucleate the formation of the NSB (Jolly et al., 2004; Rizzi et al., 2004), providing another example of specific protein-protein and protein-RNA interactions driving RNA granule formation.

Overall, RNA granules in different cell types and organelles provide many different functions, particularly with regards to the regulation of mRNA. Although the formation of RNA granules share common features such as the interaction of protein IDRs, in order to achieve specificity of function and remain distinct, RNA granules rely on specific protein and RNA interactions to drive their formation. The importance of these granules in RNA regulation means that further study of their mechanisms and purposes may provide novel insights into mRNA dynamics moving forward.

Of particular interest here are RNA granules found in yeast. In *S. cerevisiae*, it has been shown that both SGs and PBs form in response to stress (Lui et al., 2014). More recently, yeast have also been shown to possess a novel type of RNA granule, called actively translating granules or 'translation factories' (Figure 1.11). These RNA granules are present in unstressed cells, contain mRNAs that are not translationally repressed, and do not share the same markers as SGs and PBs (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). These granules are discussed in more detail in section 1.7.

1.5. Visualisation of RNA

To study the localisation, dynamics and fate of mRNAs many techniques have been developed to visualise RNA in cells. Some of these techniques need to be implemented in fixed cells, such as *in situ* hybridisation (ISH), fluorescent *in situ* hybridisation (FISH), and single-molecule fluorescent *in situ* hybridisation (smFISH). While these fixed cell techniques often provide high resolution due to high signal-to-noise ratios, they can only provide information on mRNA localisation at a single point in time and rely upon fixatives that may alter subcellular structures. Live cell techniques negate concerns about the fixatives and can also provide more information on mRNA localisation dynamics. These techniques include so-called 'molecular beacons', RNA aptamers, M-tagging and RNA-targeting with CRISPR-associated proteins (Cas9 and Cas13). All of these techniques have distinct advantages and disadvantages and the most appropriate technique is often context-dependent, contingent on cell type and status, the specific RNA under investigation, and the experimental question to be answered (Table 1). In practise, most observations made on mRNA localisation would require confirmation using multiple different techniques.

RNA visualisation technique	Live or fixed cell imaging	Imaging resolution	Multi-colour imaging	mRNA targeting structure	Experimental methods
ISH	Fixed	Struggles with low copy number	Yes	cDNA, dsRNA, ssRNA	Fixation and permeabilisation
FISH	Fixed	Struggles with low copy number	Yes	cDNA, dsRNA, ssRNA	Fixation and permeabilisation
smFISH	Fixed	Single-molecule	Yes	cDNA, dsRNA, ssRNA	Fixation and permeabilisation
Molecular beacons	Fixed and Live	Single-molecule for reporter RNAs	Yes	ssRNA	Permeabilisation
RNA aptamers	Live	Single-molecule	Yes	ssRNA (incorporated)	Cloning and injection or washes
M-tagging	Live	Single-molecule	Yes	ssRNA (incorporated)	Cloning and plasmid transformation
Cas9 RNA-targeting	Live	Potential for single-molecule	Yes	dsRNA	Cloning and plasmid transformation

Table 1: A summary of RNA imaging techniques.

1.5.1. Fixed cell methods to visualise RNA

Early techniques used to visualise RNA were performed in fixed cells using *in situ* hybridisation (ISH). ISH was first developed for the detection of DNA (Gall and Pardue, 1969; Harper and Saunders, 1981), though the same technique is used for RNA detection (Jin and Llyod, 1996). ISH involves the binding of cDNA or RNA probes that have complimentary sequences to the DNA or RNA of interest (to form a hybrid), and these probes are labelled to allow detection (Jensen, 2014). This labelling may be achieved by radiolabelling (Simmons et al., 1989; Hoefler et al., 1986), histochemical labelling (Komminoth, 1992; Egger et al., 1994), or by fluorescent labelling. In particular, two types of fluorescent labelling methods are currently widely used: molecular beacons (which may be used in live cell imaging and are discussed more below) and fluorescent *in situ* hybridisation (FISH).

FISH involves fluorescent labelling and detection of the DNA or RNA probe using fluorescence microscopy. A major advantage of FISH is that different probes can be bound to different colour fluorophores to image multiple sites of DNA or RNA in the same cell. FISH has therefore been used extensively to map sites in the human genome (Langer-Safer et al., 1982; du Manoir et al., 1993), in prenatal screening to detect chromosome abnormalities (Ward et al., 1993; Li et al., 2011; Xu et al., 2014), and to visualise gene expression by targeting mRNAs in both eukaryotes and prokaryotes (Singer and Ward, 1982; Coleman et al., 2007). While the advantage of both ISH and FISH techniques is that almost any DNA or RNA sequence can be targeted and detected, they are often limited by detection sensitivity. In particular, low copy number DNA or RNA molecules can be difficult to detect. There have been several techniques developed to improve the sensitivity of ISH experiments, including PCR *in situ* (Long, 1998) and primed *in situ* (Coullin et al., 2002).

Perhaps the most effective method to improve the sensitivity of FISH experiments is the development of single-molecule FISH (smFISH) (Femino et al., 1998). smFISH involves the binding of multiple small nucleic acid probes to either the gene or RNA of interest. Probes are usually 20 nucleotides in length, and 48 probes has been suggested to be the optimal number for a strong signal with minimal off-target effects (Raj et al., 2008). Each of these probes is attached to a low intensity fluorophore, so that when multiple probes are bound to the gene or RNA a strong fluorescent signal is produced. The advantage of this method is that single probes alone are not bright enough to produce a strong signal and so low numbers of off-target binding or unbound probes should not decrease the signal-to-noise ratio. As the name suggests, this technique may be sensitive enough to visualise single mRNA molecules, though often this depends on the experimental conditions (Femino et al.,

1998; Noma et al., 2017; Chen et al., 2018). One drawback of this method is that smFISH cannot differentiate between genes or RNAs with very similar sequences, as there will not be sufficient unique nucleotide stretches for the probes to target. Nevertheless, smFISH is currently a widely used technique to visualise mRNA to study gene expression (Zenklusen et al., 2008; Raj et al., 2008; Vargas et al., 2011) and mRNA localisation (Park et al., 2012; Samacoits et al., 2018).

The main limitation of these hybridisation techniques is that they must be performed in fixed cells. This means that any visualisation is at a fixed point in time, and so dynamic processes, such as mRNA localisation and stress responses, cannot be easily studied. More technically, fixation often relies on cross-linking proteins, which may introduce artefacts into imaging experiments and have unknown effects on cell states and intracellular cell structures (Schnell et al., 2012; Whelan and Bell, 2015).

1.5.2. Live cell methods to visualise RNA

Live cell approaches may therefore be used to study mRNA dynamics, using techniques such as molecular beacons, RNA aptamers, M-tagging, and RNA-targeting via Cas proteins. All of these methods involve tethering a fluorescent marker to the RNA of interest, and so imaging quality has often improved as better fluorescence microscopy methods and brighter fluorophores have been developed. These techniques offer their own distinct advantages and disadvantages, most often associated with achieving a strong fluorescent signal and their technical ease of use (Table 1).

Molecular beacons are RNA probes that possess both a fluorophore and a quencher of that fluorophore (Tyagi and Kramer, 1996; Monroy-Contreras and Vaca, 2011). In the initial unbound state, the probe is structured so that the fluorophore and quencher interact and no signal is produced. Upon binding an RNA target, the probe undergoes a conformational change allowing the fluorophore and quencher to separate such that the fluorescent signal can be detected (Bonnet et al., 1999). As with fixed cell ISH techniques, molecular beacons also rely on probe hybridisation with the RNA of interest to achieve targeting specificity. Similar to smFISH, the main advantage of molecular beacons is that no signal is produced by the unbound probe, resulting in a high signal-to-noise ratio. Molecular beacons have therefore been used in live cell experiments to study mRNA localisation and dynamics (Bratu et al., 2003; Tyagi and Alsmadi, 2004). The main drawback of this technique is that each probe must be specific for the RNA sequence and be made with both a fluorophore and

quencher attached, which is expensive and difficult to produce. Also for live cell imaging, the molecular beacon must be introduced into the cell by microinjection or by cell permeabilisation/electroporation, which can be technically difficult and cause damage and/or stress to the cells (Monroy-Contreras and Vaca, 2011).

RNA aptamers are small, synthetic RNA structures that can be introduced into the mRNA of interest (Babendure et al., 2003). These RNA aptamers are then recognised by specific fluorescent dyes which, upon binding to their aptamer, massively increase their fluorescence. For example, RNA Mango is a short aptamer that can be genetically introduced into a gene sequence. Once the mRNA is transcribed, RNA Mango binds to its thiazole orange dye (TO1 or TO3) to increases the fluorescence of the fluorophore around 10^3-fold (Dolgosheina et al. 2014). Similar to smFISH and molecular beacons, the most commonly used fluorogenic dyes produce very limited background fluorescence when unbound and have a high signal-to-noise ratio (Jeng et al., 2016). RNA aptamers have been used to image mRNA in live cells in both eukaryotes and prokaryotes (Strack et al., 2013; Song et al., 2017; Zhang et al., 2015), and recent improvements to the fluorescent stability of some RNA aptamer techniques have allowed the visualisation of single mRNA molecules (Cawte et al., 2020). Drawbacks of using RNA aptamers are usually specific to the RNA aptamer and dye in question. Since RNA aptamers modify the structure of the mRNA, they are designed to be as small as possible (often 40-50 nucleotides) while binding their fluorescent dye with high affinity (Cawte et al., 2020; Bell et al., 2020). Ideal features of fluorescent dyes include a lack of cell toxicity, low background fluorescence, and organelle and cell membrane permeability (Dolgosheina et al. 2014; Bouhedda et al., 2018).

M-tag systems, or RNA stem-loop systems, are perhaps historically the most extensively used technique to visualise RNA in live cells. The most prevalent of these systems, the MS2 system, was first used to study the localisation of *ASH1* mRNA in budding yeast (Bertrand et al., 1998). The MS2 system involves the introduction of MS2 stem loops, most often at the start of the 3'UTR, into the targeted mRNA. These stem loops are recognised and bound by the bacteriophage protein MCP (MS2 coat protein), which can be fused to a fluorophore such as GFP, and this fluorescent signal tracked. M-tag systems can be used for multi-colour imaging via different stem loop structures and recognition partners. For example, one mRNA with MS2 stem loops can be visualised by an MCP-GFP fusion protein, while another mRNA can be viewed in the same cell via incorporation of PP7 stem loops bound to a PP7-mCherry fusion (Pizzinga et al., 2019). The brightness of these systems can vary based on the number of inserted stem loops and fused fluorophores, and the fluorescent intensity of the fluorophores. For example, it has been shown that 6x MS2 stem loops and a single

MCP-GFP fusion is capable of tracking bulk mRNA (Bertrand et al., 1998), while single-molecule resolution can be achieved using 24x MS2 stem loops and an MCP-RFP fusion (Wu et al., 2016) or MCP-JF646 fusion (Morisaki et al., 2016). Since its development, the MS2 system has been widely used to study mRNA localisation (Bertrand et al., 1998; Grünwald and Singer, 2010), along with transcription and translation (Darzacq et al., 2007; Martin et al., 2013; Wu et al., 2016). However the MS2 system has been criticised due to the requirement for modification of the mRNA structure. Experiments using the MS2 system most often include the addition of 12 or 24 stem loops at the start of the 3'UTR (Bertrand et al., 1988). This modification, along with the binding of multiple recognition proteins to the stem loops, may affect rates of mRNA translation and decay, and could lead to the accumulation of mRNA fragments in mRNA localisation experiments (Garcia and Parker, 2015; Haimovich et al., 2016; Garcia and Parker 2016; Heinrich et al., 2017). For certain MS2-tagged mRNAs, it has been shown that the MS2 mRNA binding complex may prevent the complete 5' to 3' degradation of mRNA by blocking the progression of the Xrn1 exonuclease (Sheth and Parker 2003; Heinrich et al., 2017). This leads to an accumulation of mRNA 3'UTR decay fragments that still contain the MS2 stem loops and so are still visualised by the MS2 system (Heinrich et al., 2017). These 3' fragments may then localise to P-bodies (Heinrich et al., 2017), which would lead to potential misinterpretation of patterns of mRNA localisation. Recently, MS2 stem loops have been used which bind the MCP recognition protein with a lower affinity. This strategy has been suggested to alleviate some adverse effects of tagging, although this technique still involves modification of the mRNA structure (Tutucci et al., 2018a; Tutucci et al., 2018b).

A relatively recently developed RNA live imaging system that does not involve the modification of mRNA is the Cas protein RNA-targeting system. The development of the CRISPR/Cas9 system, and the use of Cas9 in RNA-targeting are discussed in section 1.6. Briefly, it has been shown that catalytically dead Cas9 (dCas9) can be targeted to mRNAs by specific sgRNAs (single-guide RNAs) (Nelles et al., 2016; Sampson et al., 2013; Price et al., 2015). Similar to the MS2 system, the dCas9 protein can be fused to a fluorophore to allow visualisation of the mRNA in live cells. The advantage of this dCas9 system is that the mRNA is endogenously expressed and unmodified, therefore possibly providing an improved RNA-targeting system (Nelles et al., 2016). It has since been shown that Cas13 proteins can also be used to target RNA in live cells (Abudayyeh et al., 2017; Smargon et al., 2017; Yan et al., 2018), and RNA-targeting by Cas proteins is a rapidly developing area of research (Smargon et al., 2020).

1.6. Development of CRISPR/Cas9 and the dCas9 RNA-targeting system

A recently developed technique for studying endogenous mRNA localisation in live cells involved the targeting of dCas9 (dead CRISPR associated protein 9) to RNA, called the dCas9-RNA targeting system (Nelles et al., 2016). This technique is similar to the way dCas9 has been used to visualise DNA, which represents an extensive use of Cas9 since the development of the CRISPR (clustered regularly interspaced short palindromic repeats)/Cas9 system (Anton et al., 2014; Martens et al., 2019; reviewed in Xu and Qi, 2019). CRISPR/Cas9 is a genome-editing tool that can be used to specifically target regions of DNA and make alterations; typically gene/sequence deletions, insertions and point mutations (reviewed in Pickar-Oliver and Gersbach, 2019).

1.6.1. The CRISPR/Cas9 system in bacteria and genome-editing

The CRISPR/Cas9 system is a bacterial immune system used to destroy non-native DNA, which may enter the cell from pathogens (Mojica et al., 2005; Barrangou et al., 2007). This is an acquired immunity, where DNA sequences (protospacers) are acquired from invading pathogens and arranged in CRISPR repeats (as spacers) in the bacterial genome (Ishino et al., 1987; reviewed in Hsu et al., 2014). These 20 base pair (bp) spacer sequences are then transcribed into RNA typically referred to as crispr RNA (crRNA). This crRNA forms a complex with a tracrRNA (trans-activating crispr RNA) that is partially complimentary to the crRNA and contains a conserved scaffold domain to which Cas9 binds (Jansen et al., 2002; Brouns et al., 2008; Deltcheva et al., 2011). Cas9 is an endonuclease capable of unwinding and cleaving DNA (Garneau et al., 2010; Jinek et al., 2014). Therefore, when foreign DNA is present in a cell that has acquired resistance to it, Cas9 is directed to the pathogen via the crRNA-tracrRNA complex for cleavage. Cas9 cleavage also requires the presence of a protospacer adjacent motif (PAM) sequence (-NGG), which is recognised by the PAM-interacting domain of Cas9 (Bolotin et al., 2005) (Figure 1.12). This PAM sequence is located immediately downstream of the protospacer sequence in the pathogenic DNA and increases the specificity of Cas9 for the pathogen; if the PAM sequence is not present, cleavage does not occur (Deveau et al., 2008; Jinek et al., 2014). This helps to prevent accidental cleavage of the native bacterial genome by Cas9, particularly since Cas9 may bind (less efficiently) to sequences that are similar to the 20 bp guide sequence (Shah et al., 2013).

Specificity is also an important consideration when using the CRISPR/Cas9 system for genome-editing experiments (Veres et al., 2014). When a gene or genomic region is targeted for deletion/mutation, the targeting 20 base pairs must be upstream of a PAM sequence, and online tools can be used to select the sequence that has the optimal gene specificity and therefore the fewest off-target effects (Doench et al., 2016; Ran et al., 2013). In genome-editing the mechanism of Cas9 cleavage is generally the same as that found in the bacterial host system, although the crRNA and tracrRNA is often introduced as a single RNA molecule called the single-guide RNA (sgRNA) (Cong et al., 2013; Jinek et al., 2014) (Figure 1.12). Cleavage of the gene can then be repaired by either homology directed repair (HDR) or non-homologous end joining (NHEJ), depending on the desired outcome (reviewed in Jasin and Haber, 2016). HDR allows for insertion and mutation, where a DNA template is provided that is incorporated into and replaces the original gene sequence (Mali et al., 2013; Yang et al., 2013). NHEJ often results in random base insertion/deletion at the site of repair, and therefore is used to introduce frameshift mutations and essentially give gene knockouts (Wang et al., 2013; Yin et al., 2014). CRISPR/Cas9 is therefore a versatile and widely used tool for genome-editing that allows precise and efficient alterations to genomic DNA to be made.

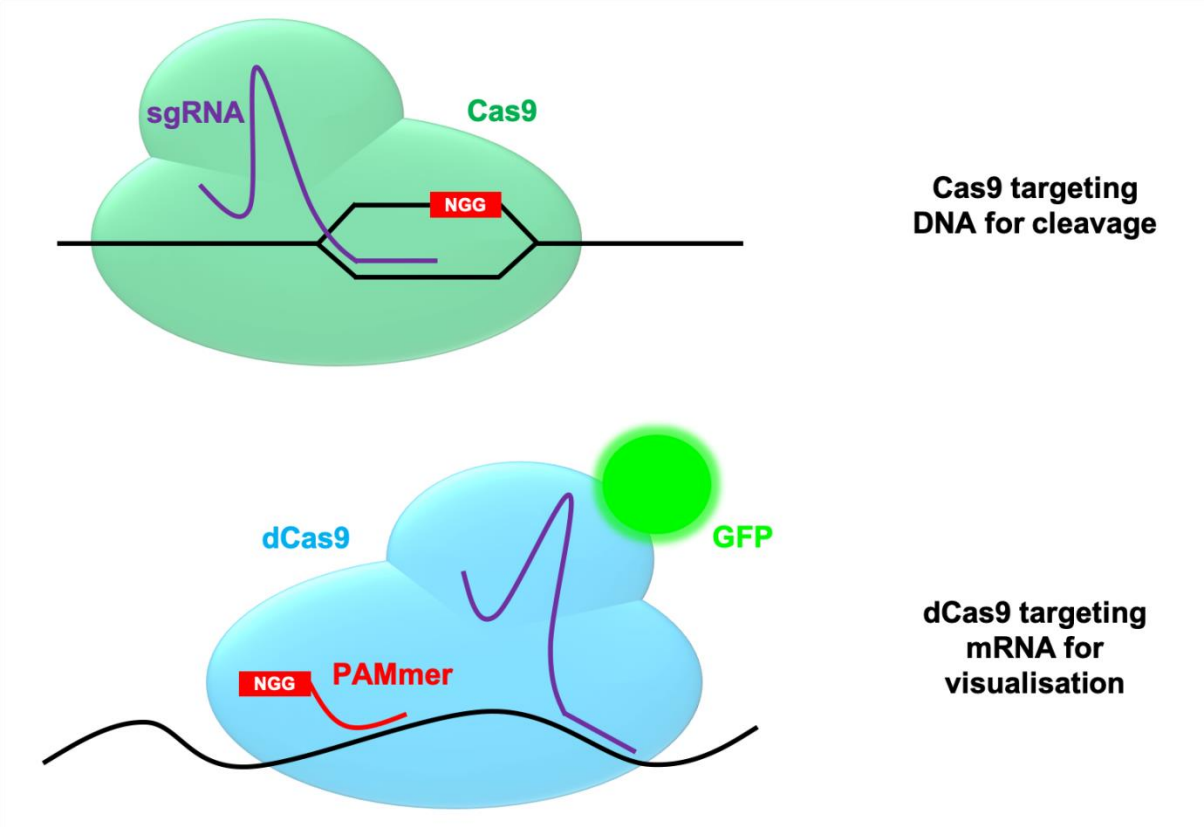


Figure 1.12: The CRISPR/Cas9 and dCas9 RNA-targeting systems. In the CRISPR/Cas9 system (top image), a single-guide RNA (sgRNA) is used to target the Cas9 protein to the gene sequence. The sgRNA contains both a 20 bp recognition region that hybridises to the gene sequence and a conserved scaffold domain region to which the Cas9 binds. Cas9 will then catalyse a double-stranded break in the DNA sequence directly adjacent to the PAM sequence (NGG). In the dCas9 RNA-targeting system (bottom image), the dCas9 cannot catalyse nucleotide cleavage and is fused to a fluorophore (or fluorophores). Again, the dCas9 is targeted to the RNA sequence via a sgRNA. A PAMmer may also be included, which contains a 20 bp recognition region for the mRNA and a PAM sequence (NGG), and provides additional targeting of the dCas9.

1.6.2. The use of dead Cas9 in gene regulation and live-cell imaging

More recently, CRISPR/Cas9 has been developed to target DNA without cleavage, allowing the system to be used for live-cell imaging, epigenetic modification and gene regulation (reviewed in Brocken et al., 2018). This is achieved through use of a catalytically inactive Cas9, called dead Cas9 (dCas9) (Jinek et al., 2012). dCas9 has two point mutations in its gene sequences that inhibit its catalytic activity, so that the protein is still capable of recognising sgRNA and being targeted to sequences, but not able to cleave (Jinek et al., 2012). This dCas9 construct has been used in a wide range of experiments; for example, dCas9 binding to genomic loci has been shown to occlude binding of transcription factors and RNA polymerase II to inhibit gene transcription (called CRISPRi) (Qi et al., 2013). Also, dCas9 fusion to enhancers/repressors of transcription has been used to provide epigenetic regulation of gene transcription (Pulecio et al., 2017).

Similarly, the use of dCas9 in live-cell imaging typically involves the fusion of the dCas9 protein to a fluorophore (Chen et al., 2013) (Figure 1.12, bottom image). This imaging system has been used extensively to study chromatin dynamics in live cells, including modelling telomere length and dynamics (Chen et al., 2013; Duan et al., 2018), and studying the dynamics of chromatin condensation (Xue and Acar, 2018). Initial experiments focused on targeting repetitive genome sequences, which allowed multiple sgRNAs and dCas9-fluorophores to bind to provide a strong signal (Chen et al., 2013). To target non-repetitive genome sequences, strategies to increase the fluorescent signal of the dCas9 system have been employed. Successful techniques include the use of multiple sgRNAs, addition of multiple peptide epitopes (e.g. SunTag) to the dCas9 protein, and/or MS2 tagging of the sgRNA (Wu et al., 2019):

- Targeting multiple sgRNAs to a genomic region allows multiple dCas9-fluorophore proteins to bind; the number of sgRNAs required for a signal depends on the cell type and fluorophore used, where around 30 sgRNAs has been shown to be sufficient to label DNA in mammalian cells using dCas9-eGFP (enhanced GFP) (Chen et al., 2013).
- SunTagging involves the attachment of a peptide repeat array to the protein of interest, to which genomically expressed multiple single-chain antibodies can bind (Tanenbaum et al., 2014). These antibodies are themselves each fused to a fluorophore, and so multiple fluorophores may be targeted to the genome via a single sgRNA/dCas9 (Shao et al., 2017).
- MS2 stem loops may be introduced into the structure of the sgRNA without affecting its binding to nucleotides or dCas9. This then allows multiple coat protein-fluorophore

fusion proteins to bind the MS2 sequences on the sgRNA and provide a strong signal; studies have shown that a 14x MS2 loop insertion into one sgRNA is sufficient to visualise specific genetic loci in the nucleus (Qin et al., 2017).

Variations of these dCas9 imaging systems can also be used to visualise multiple genomic foci in a single living cell. Often this is achieved using dCas9 orthologs derived from different species of bacteria; the first and most commonly used Cas9 is from *Streptococcus pyogenes* (Sp) (Jinek et al., 2012). For example, Cas9 orthologs from *Neisseria meningitidis* and *Streptococcus thermophilus* recognise and bind different sgRNA structures compared to SpCas9 (Ma et al., 2015). Therefore, unique sgRNAs can be targeted to different genomic loci and bound by their distinct dCas9 orthologs, each fused to discrete fluorophores, allowing the visualisation of multiple genomic regions in the same cell. Multi-colour imaging can also be achieved using modified sgRNAs. For example, one sgRNA can be tagged with MS2 stem loops and its corresponding coat protein and fluorophore, while another sgRNA may contain PP7 stem loops to be bound by its own, distinct coat protein fused to a different fluorophore (Fu et al., 2016). Overall, the ability of dead Cas9 to specifically bind nucleotides without cleavage means that dCas9 constructs have many uses for both gene regulation and live imaging of genomic loci.

Along with genome editing, CRISPR/Cas based systems have also been used to target RNA. It has been shown that Cas9 from the bacterial species *Francisella novicida* (FnCas9) is capable of driving degradation of endogenous mRNA (Sampson et al., 2013) and can be engineered to target other RNAs (Price et al., 2015). As in the endogenous system, this RNA targeting relies on a crRNA that forms a complex with the tracrRNA, which interacts with Cas9 and has a single-stranded (ss) RNA region complimentary to the RNA target (Price et al., 2015). It has also been shown that Cas9 proteins derived from *Staphylococcus aureus* and *Campylobacter jejuni* are capable of cleaving RNA when targeted by a sgRNA alone (Strutt et al., 2018).

Cas9 from *Streptococcus pyogenes* (SpCas9) has also been used to target RNA. SpCas9 can be targeted to ssRNA through the use of a sgRNA and a PAMmer, where the PAMmer is a single-stranded oligonucleotide that contains a complimentary region to the RNA target with a downstream PAM sequence (O'Connell et al., 2014). As in the endogenous bacterial system, the presence of this PAMmer stimulates Cas9 cleavage of the RNA (O'Connell et al., 2014). Both wild-type SpCas9 and catalytically inactive SpCas9 can also be targeted to RNA without the presence of a PAMmer (sgRNA-targeting alone) (Liu et al., 2016; Batra et al., 2017), although the presence of a PAMmer may increase the efficiency of this targeting (Nelles et al., 2016).

Since catalytically dead Cas9 (dCas9) can also be used to target RNAs in this way, dCas9 constructs can also be used to image RNA (Figure 1.12, bottom image). This technique, developed in 2016 by Nelles et al., has been used to study mRNA localisation in live mammalian cells, including mRNA accumulation in RNA granules. This method uses a dCas9-GFP fusion construct that includes a nuclear localisation signal (NLS), so that the fluorescent signal is confined to the nucleus in the absence of targeting and dCas9 may interact with mRNAs as they are synthesised. mRNA targeting is achieved through the use of a complimentary sgRNA, where the additional presence of a PAMmer improves dCas9 targeting (shown by a higher percentage of GFP leaving the nucleus when targeting a cytosolic mRNA) (Nelles et al., 2016). Further experiments from this study demonstrated that the targeting of mRNAs by dCas9 did not affect associated RNA and protein levels, and orthogonal assays using smFISH showed that the localisation patterns seen using the dCas9 system were accurate (Nelles et al., 2016). The ability of this dCas9-RNA targeting system to target endogenous, unmodified mRNAs in live cells may provide a useful tool for the visualisation of RNA in the future.

1.7. Translation factories in yeast (Previous work in the lab)

Previous work undertaken by the Ashe lab has identified a novel set of RNA containing granules in yeast (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). These granules appear distinct from previously discovered types of granules (e.g. stress granules and P-bodies), as they are present in unstressed yeast cells, they do not overlap with established P-body and stress granule markers, and the mRNAs they contain are not translationally repressed (Lui et al., 2014) (Figure 1.11).

Indeed, the formation of these granules appears to be dependent on mRNA translation; when stem loops were introduced into the mRNA structure that prevented translation initiation, the mRNA no longer localised to granules (Pizzinga et al., 2019; Morales-Polanco et al., 2021). Further, photobleaching experiments, studies using the TRICK technique and quantitative assessments of the proportion of translated or granule-enriched mRNA all suggest that mRNA translation occurs at these granules (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). Photobleaching experiments were done using fluorescently tagged Eno2p (Eno2p-mOrange), where fluorescence was seen to recover specifically in the same location as granules containing *PDC1* mRNA (which *ENO2* mRNA colocalises with), suggesting the Eno2p is being produced in the same place (Lui et al., 2014). TRICK (translating RNA imaging by coat protein knock-off; Halstead et al., 2015) experiments were done by introduction of stem loops before and after the stop codon of localised mRNAs (*NIP1* and *TIF4631*) (Pizzinga et al., 2019). PP7 stem loops were introduced before the stop codon to be recognised by PP7-GFP_3 and MS2 stem loops were introduced after the stop codon to be recognised by MCP-$mCherry_3$. In unstressed cells only the mCherry signal could be seen localised to RNA granules, indicating that ribosome elongation had caused dissociation of the PP7-GFP_3 during translation. When cells were stressed following glucose depletion and global translation was stalled, both mCherry and GFP could then be seen localised to the same RNA granules (Pizzinga et al., 2019). Given that these RNA granules house a functionally related set of mRNAs which are actively translated into protein, they are termed ‘translation factories’ or actively translating granules. Current work is underway to investigate the mechanism of mRNA localisation to these granules, the protein and RNA components of these granules, and finally, what the purpose of these granules may be.

Multiple mRNAs have been identified that localise to actively translating granules, with different mRNAs showing different localisation patterns. For example, the *PDC1* and *ENO2* glycolytic mRNAs are present in around 10-20 granules per cell, whereas the *NIP1* and *TIF1* translation initiation factor mRNAs are present in 1-2 granules per cell (Lui et al., 2014;

Pizzinga et al., 2019; Figure 3.1). Other mRNAs were not localised to granules under the same conditions; e.g. *GIP2*, *CIN5* and *NPC2* mRNA (Lui et al., 2014; Figure 3.1). Further work characterising these granules found that 19 different glycolytic mRNAs (such as *PDC1*, *ENO2,* and *FBA1*) show the same localisation pattern (several granules per cell) and also co-localise together in the same granules (Morales-Polanco et al., 2021). Similarly, mRNAs that encode translation factor proteins (such as *NIP1*, *TIF1* and *EFB1*) also demonstrate the same localisation patterns as each other (1-2 granules per cell) and show some level of co-localisation within cells (Pizzinga et el., 2019). Unpublished data from the Ashe lab also shows that mRNAs encoding aromatic amino acid biosynthetic enzymes (e.g. *ARO1*, *ARO3*, and *ARO4*) also localise to granules in unstressed yeast cells. Together, these data highlight the possibility that the purpose of these actively translating granules is to increase the efficiency of the pathways involved – the glycolytic pathway, protein translation, and aromatic amino acid biosynthesis. Further, for the translation factor mRNA granules, an additional role was suggested in providing a ready source of translation factor mRNAs to the developing daughter cell via a specific inheritance mechanism (Pizzinga et al., 2019).

To visualise mRNA localisation in cells, both the MS2 system and smFISH approaches have been used in the lab. As discussed in section 1.5, the MS2 system is a live cell system that involves the introduction of MS2 stem loops into the mRNA of interest, which are then bound by a coat protein fused to a fluorophore to allow visualisation (Bertrand et al., 1988). smFISH is a fixed cell method that involves the binding of multiple small oligomers, each fused to a low-intensity fluorophore, to the mRNA of interest (Femino et al., 1998). Both of these techniques have been used to show mRNA localisation to actively translating granules in yeast and used to study co-localisation of these mRNAs (Pizzinga et al., 2019; Morales-Polanco et al., 2021). The main reason for including the orthogonal approach of smFISH is due to the possible drawbacks of the MS2 system, which are discussed in more detail in section 1.5. Therefore in the Ashe lab, mRNA localisation and co-localisation results found using the MS2 system have been verified using smFISH on endogenous unmodified mRNAs (Pizzinga et al., 2019; Morales-Polanco et al., 2021). Although smFISH has been used successfully to validate results found using the MS2 system, there are drawbacks since smFISH is a fixed cell technique. Therefore, smFISH cannot be used to study mRNA dynamics; for example the effect on mRNA localisation following stress, or the lifetime of the actively translating granules. smFISH also requires fixation of the yeast cells, which may introduce imaging artefacts that affect mRNA localisation (Schnell et al., 2012; Whelan and Bell, 2015). Therefore, the further study of these actively translating granules in yeast would benefit from the development of novel imaging techniques to allow the study of mRNA localisation in live cells without the disadvantages of mRNA modification.

1.8. Project aims

As discussed in section 1.7, previous work in the Ashe lab has identified novel cytosolic RNA granules in yeast, called 'translation factories' or actively translating granules. Different mRNAs show strikingly different patterns of RNA granules; glycolytic mRNAs are shown to be present in 10-20 granules per cell and translation factor mRNAs in one or two granules per cell. One aim of this thesis was to understand the mechanisms by which these different patterns of mRNA localisation are achieved inside the cell. To investigate this, regions of mRNA (5'UTR, ORF, 3'UTR) were exchanged between mRNAs to attempt to identify the region(s) responsible for mRNA localisation. Swap constructs were made between the glycolytic mRNA *PDC1* and the translation factor mRNA *NIP1*, and imaged using the MS2 system to determine the pattern of mRNA granule localisation.

Since glycolytic mRNAs are more highly expressed than translation factor mRNAs, it is possible that this is the reason for the increased number of glycolytic mRNA granules. To investigate this, *NIP1*, *PDC1* and *NIP1/PDC1* swap mRNAs were expressed on a lower expression plasmid and qRT-PCR was performed on each mRNA to test whether mRNA abundance had an effect on the number of actively translating granules.

Investigation into mRNA localisation performed in the Ashe lab has thus far been achieved by imaging mRNA with the MS2 system and smFISH system. As discussed in section 1.7, each system has drawbacks; the MS2 system requires modification of the target mRNA structure which may affect mRNA dynamics, and smFISH is a fixed cell system that cannot be used to study mRNA in live cells. Therefore, another aim of this thesis was to develop a method that allows the visualisation of unmodified mRNAs in live cells and to make this system technically less onerous than the MS2 system, which requires three different yeast strains to be generated sequentially for every mRNA tagged. Hence, in this project a novel RNA targeting system was developed in yeast, called the dCas9 RNA-targeting system. The advantages of this system are that the targeted mRNA is endogenous and unmodified, the system can be used to visualise mRNA in live cells, and the system can be easily adapted to target many different mRNAs. The dCas9 RNA-targeting system was validated by assessing mRNAs with known localisation profiles established using the MS2 and smFISH systems. Then as a further demonstration of the ease and flexibility of the new system, its use was extended to study localisation profiles of mRNAs that have not been previously described.

Materials and Methods

2. Materials and Methods

2.1. Yeast strains and growth conditions

All yeast strains were from *Saccharomyces cerevisiae* and a complete list of strains used and made is listed in Table 2. Yeast strains here were almost all made by the transformation into parent strains of a bacterial plasmid(s), listed in Table 3. The exception was the yMK3433 strain where the nanobody-GFP fragment was genomically integrated into the yeast genome near the *TRP1* locus (Chromosome IV 463001). Yeast strains were grown in either complete media (YPD; 1% yeast extract, 2% Bacto peptone, 2% glucose) or in synthetic complete media with the appropriate amino acid(s) excluded to select for the plasmid(s) (SCD; 0.17% yeast nitrogen base, 0.5% ammonium sulphate, 2% glucose, 0.02 mg/ml adenine, arginine, uracil, methionine, histidine, tryptophan and threonine, 0.03 mg/ml lysine, tyrosine and isoleucine, 0.05 mg/ml phenylalanine, 0.06 mg/ml leucine, 0.1 mg/ml glutamic acid and aspartic acid, 0.15 mg/ml valine, 0.4 mg/ml serine). For selection using hygromycin, 200 µg/ml hygromycin B was also included (Thermo Fisher Scientific). For growth on solid media, 2% agar was also included. Yeast were grown at 30ºC either overnight for microscopy or for 2-3 days following plasmid transformation.

2.2. Molecular biology

2.2.1. sgRNA design and construction

sgRNAs were designed to target 20 base pairs (bps) of target mRNAs. In most cases the 3'UTR was targeted, except for *PDC1* and *PDC5* mRNAs which have very similar 3'UTR sequences; differing regions of their ORFs were targeted. sgRNAs were designed using the online ATUM CRISPR gRNA Design tool (https://www.atum.bio/eCommerce/cas9/input) which identifies the most unique sequence possible for targeting. The following design parameters were used: **Species:** *Saccharomyces cerevisiae*, **PAM:** NGG (not necessary here but may increase specificity (O'Connell et al., 2014)), and **Type:** Wild-type Cas9. Targeting region specificity was then verified using a BLAST search for each 20 bps against the *Saccharomyces cerevisiae* genome, where no sgRNAs appeared capable of binding off-target genes. sgRNAs were made by annealing together forward and reverse primers (listed in Table 5) containing the 20 bp target region flanked on either side by 25 bps corresponding to the sgRNA vector (pMEL14). Primers were annealed at a 1:1 ratio, boiled at 95ºC for 10 minutes then left to cool to room temperature. Annealed primers were stored at -20ºC until used in Gibson cloning.

2.2.2. PCR

PCRs were performed using primers listed in Table 4. PCRs were used to generate fragments for Gibson cloning (Gibson et al., 2009), to verify the length of MS2 stem loops regions and to verify construct insertion into the yeast genome. PCRs were done using either Phusion® High-Fidelity DNA Polymerase (NEB) or RANGER mix (MyTaq™ polymerase) (Bioline), and performed according to the manufacturer's instructions.

2.2.3. Restriction cloning and restriction enzyme mapping

All restriction enzymes were sourced from NEB and used with their corresponding NEB buffer. Restriction cloning was used to make CEN plasmid backbone constructs and $mNeonGreen_3$ constructs. In each case the appropriate restriction enzyme(s) were used to digest miniprep DNA samples; 5 µl of each enzyme in the appropriate buffer was incubated with 25 µl of DNA for at least 1 hour at 37ºC and then heat inactivated at the appropriate temperature. Digested samples were run on 1% agarose gels and the desired gel band excised with a scalpel. DNA was then isolated using Isolate II PCR and Gel kit (Bioline). DNA fragments were ligated using T4 ligase (NEB) according to the manufacturer's instructions. Samples were then either stored at -20ºC or directly transformed into bacteria.

Restriction enzyme mapping of constructs was done by incubating 1 µl of the appropriate restriction enzyme in its corresponding buffer with 5 µl of miniprep DNA. Samples were incubated for at least 2 hours at 37ºC and ran on 1% agarose gels.

2.2.4. Gibson cloning

Gibson cloning (Figure 2.1) was performed using Gibson Assembly® Master Mix (NEB) (Gibson et al., 2009). PCR fragments were produced to include 25 bp overhangs that correspond to the other annealing fragment(s) (see list of primers, Table 4). PCR fragments were either incubated together directly or following clean-up using the Isolate II PCR and Gel kit (Bioline). Gibson reactions were carried out using 10 µl Gibson master mix and 100-300 ng of DNA fragments, incubated together for 30 minutes at 50ºC. Samples were then either stored at -20ºC or directly transformed in bacteria.

2.2.5. Bacterial transformation and plasmid isolation

All cloning into bacteria was done using *Escherichia coli* cells (NEB 5-alpha cells). Bacterial transformations were carried out by incubating 2-20 µl of ligated DNA samples with 50 µl NEB 5-alpha cells for 30 minutes on ice. Cells were then heat-shocked for 30 seconds at 42ºC and rested on ice for 2 minutes. 950 µl SOC media (NEB) was then added to each sample and samples were incubated on rotation for at least 1 hour at 30ºC. Cells were then plated onto Luria-Bertani media plates (LB; 1% Bacto tryptone, 0.5% yeast extract, 10 mM sodium chloride, 2% agar) and incubated overnight at 37ºC. Single bacterial colonies were then grown in liquid LB (without agar) overnight at 30ºC on rotation and DNA plasmids were isolated using the Monarch® Plasmid Miniprep Kit (NEB) according to the manufacturer's instructions.

2.2.6. Yeast transformation

Yeast strains were grown overnight at 30ºC on rotation in the appropriate liquid media to an OD600 of 0.5 – 0.8 (exponential phase). Cells were transformed using the lithium acetate (LiAc) protocol (Gietz and Schiestl, 2007). Yeast cells were pelleted at 3000 xg for 5 minutes and resuspended in 100 mM LiAc. Cells were then incubated in 35% polyethylene glycerol (PEG), 100 mM LiAc, 0.46 µg/ml salmon sperm DNA (Sigma-Aldrich) and 10-100 ng of plasmid DNA or PCR DNA fragment. Salmon sperm DNA was boiled for 5 minutes at 95ºC and chilled on ice for 5 minutes before use. Cells were incubated for at least 1 hour at 30ºC on rotation, and then heat-shocked for 20 minutes at 42ºC. 1 ml of sterile water was added to each sample, and cells were pelleted at 3000 xg for 5 minutes and then plated onto appropriate SCD plates. Yeast colonies were grown for 2-3 days at 30ºC on plates and then streaked out onto fresh selection plates.

2.2.7. Yeast genomic DNA extraction

DNA was extracted from yeast to perform PCR verification of the nanobody-GFP construct insertion into the genome. DNA was extracted using the LiAc-SDS (lithium acetate - sodium dodecyl sulphate) method (Lõoke et al., 2011). Single yeast colonies were suspended in a 200 mM LiAc 1% SDS solution and vortexed for 1 minute. Samples were then incubated at 70ºC for 10 minutes and 300 µl of 96% ethanol was added. Samples were vortexed briefly and centrifuged at 15,000 xg for 5 minutes. The supernatant was removed and the DNA pellet washed with 70% ethanol. DNA was again washed with 100 µl of sterile water, and then resuspended in 50 µl sterile water. 1 µl of the genomic DNA solution was used for PCR.

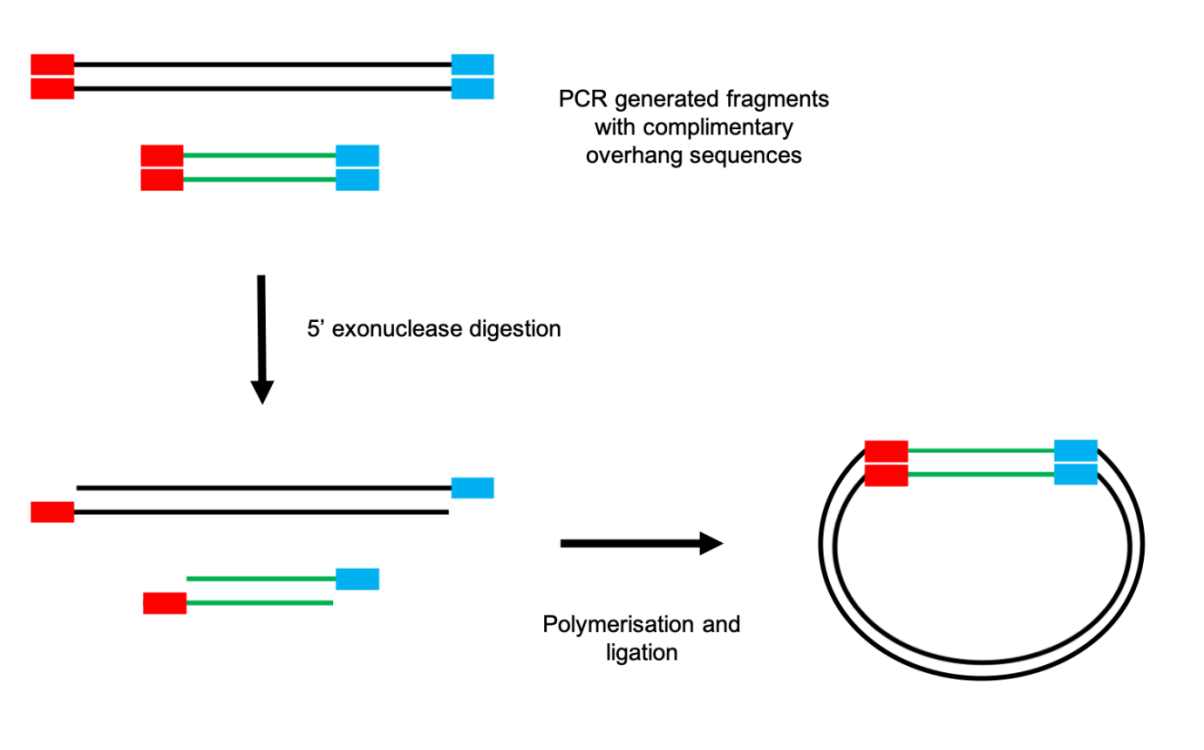


Figure 2.1: An overview of Gibson cloning. PCR fragments are generated using primers with complimentary overhang sequences to the appropriate fragment. 5' exonuclease digestion removes nucleotides at the 5' ends of each PCR fragment, and polymerisation and ligation allow fragments to ligate and missing nucleotides to be filled in.

2.3. Western blotting

Western blotting was performed on yeast strains transformed with dCas9-GFP_3, dCas9-$mNeonGreen_3$ and dCas9-MoonTag plasmids to verify the expression of these fusion proteins. Yeast strains were grown overnight at 30ºC on rotation in the appropriate liquid media to an OD600 of 0.7 (exponential phase) and harvested at 3000 xg for 5 minutes. Western blot samples were prepared using 100 µl of the cell sample, 30 µl of Novex® Bolt™ LDS Sample Buffer (Thermo Fisher Scientific) and 12 µl Novex® Bolt™ Sample Reducing Agent (Thermo Fisher Scientific). Samples were boiled at 95ºC for 10 minutes and stored at -20ºC before use.

All samples were run on Bolt™ 4-12% Bis-Tris Plus gels (Thermo Fisher Scientific). Samples were run in 1X Bolt™ MOPS SDS Running Buffer (Thermo Fisher Scientific) using a Spectra™ Multicolor Broad Range Protein Ladder as a marker (Thermo Fisher Scientific) (5 µl of marker and 10 µl of sample). The blot was ran for 45 minutes at 200 V and then moved to the transfer chamber. The transfer was ran for 90 minutes at 25 V in 1X transfer buffer (750 ml distilled water, 200 ml methanol, 50 ml 20X NuPAGE™ Transfer Buffer (Thermo Fisher Scientific)). Samples were then blocked with PBST with milk (1X PBS, 0.1% Tween-20, 5% milk) for 1 hour. Samples were incubated with primary antibodies overnight; PBST with milk and either 1:1000 anti-GFP antibody (AbCam), 1:2000 anti-dCas9 antibody (AbCam) or 1:000 anti-eIF4A antibody (AbCam). Samples were then washed with PBST 3x for 5 minutes. Secondary antibodies were incubated for an hour; PBST with milk and either 1:5000 anti-Rabbit antibody (AbCam) or 1:5000 anti-Mouse antibody (AbCam). Samples were then washed with PBST 3x for 5 minutes and imaged using an Odyssey® Fc imaging system (LI-COR).

2.4. RNA extraction and qRT-PCR

RNA was extracted from yeast cells grown overnight at 30ºC on rotation in the appropriate liquid media to an OD600 of 0.5 – 0.8 (exponential phase). Cells were harvested by centrifugation at 3000 xg for 5 minutes and resuspended in 1 ml TRIzol™ (Thermo Fisher Scientific). Cells were lysed by the addition of glass beads and vortexed 5x for 20 seconds with 40 seconds rest on ice. The supernatant was transferred to a fresh Eppendorf (StarLab) and 150 µl chloroform added. Samples were incubated for 3 minutes at room temperature and centrifuged at 12,000 xg for 15 minutes at 4ºC. The supernatant was transferred to a fresh Eppendorf and 1 µl Glycoblue (Thermo Fisher Scientific) and 350 µl 100% isopropanol was added to each sample. Samples were stored at -80ºC overnight and then centrifuged at

12,000 xg for 15 minutes at 4ºC. The supernatant was removed and the RNA pellet washed with 750 µl of 75% ethanol in nuclease-free water (Thermo Fisher Scientific). Samples were then centrifuged at 12,000 xg for 10 minutes at 4ºC, the supernatant was removed and RNA pellets were left to dry under a Bunsen burner for 20 minutes. RNA pellets were resuspended in 20 µl nuclease-free water (warmed to 55ºC) and stored at -80ºC.

RNA samples were then DNase treated using RQ1 RNase-Free DNase (Promega) according to the manufacturer's instructions. Around 1 µg of RNA was used per reaction. cDNA was then produced using ProtoScript® II First Strand cDNA Synthesis Kit (NEB) according to manufacturer's instructions with minor modifications. Namely, 0.5 µl RNaseIn Plus (Promega) was added to each sample, and the enzyme mix was not included for minus reverse transcriptase samples. Finally, qPCR was performed using GoTaq® qPCR Master Mix (Promega). Each sample contained 10 µl GoTaq qPCR master mix, 300 nM each of the appropriate forward and reverse primers (see primer list, Table 4), 0.2 µl carboxy-X-rhodamine (CXR) reference dye, around 300 ng of cDNA sample in a total volume of 20 µl (made up with nuclease-free water). qPCR reactions were made up in qPCR 96-well plates (Thermo Fisher Scientific) and performed in a Bio-Rad Cfx Connect plate reader (Bio-Rad). The qPCR protocol was run using CFX Manager™ Software (Bio-Rad), which consisted of polymerase activation at 95ºC for 2 minutes, followed by 40 cycles of denaturation (15 seconds at 95ºC), annealing/extension (1 minute at 60ºC) and a plate read. qPCR data was analysed using Microsoft excel and GraphPad Prism, and qPCR graphs were created using GraphPad Prism.

2.5. Microscopy and image analysis

All imaging was done on yeast cells grown overnight at 30ºC on rotation in the appropriate media to an OD600 of 0.5 – 0.8 (exponential phase). Images were taken using a Delta Vision (Applied Precision) or Nikon Eclipse E600 microscope with a 100x oil objective lens. Images were taken as either single plane images or z-stacks, where z-stacks were taken in 0.2 µM increments. Where appropriate, images were deconvolved using Huygens deconvolution software and displayed as maximum or average projected intensity images using ImageJ. An ImageJ plug-in, FindFoci (Herbert et al., 2014), was used to quantify mRNA granules where necessary. This software requires the user to manually define examples of foci (in this case cytosolic RNA granules) and then automatically applies this processing to further images. mRNA granule quantification was shown in graphs created using GraphPad Prism.

2.6. Computational analysis

Analysis to quantify RNA granules was done using FindFoci as described above. qPCR analysis was done in Microsoft Excel to derive ΔCt values and fold changes ($2^{-\Delta\Delta Ct}$) (Livak and Schmittgen, 2001). For both granule quantification and qPCR data, error bars and standard deviation were calculated using GraphPad Prism. GraphPad Prism was also used to carry out statistical analysis of granule quantification by performing one-way analysis of variance (ANOVA) with the parameters 'assume Gaussian distribution of residuals' and 'assume equals standard deviations'.

MEME Suite analysis was done to analyse the mRNA sequences of glycolytic mRNAs and translation factor mRNAs. MEME Suite is an online web-server that automates alignment and sequence analysis to search for conserved sequence motifs (Bailey et al., 2006; Bailey et al., 2009). For the MEME analysis done here mRNA sequences were input with additional 1kb sequences either side of the coding sequence to ensure promoter, UTR and terminator regions were included. The following MEME Suite parameters were used:

Motif discovery mode: Classical (one dataset is provided)

Sequence alphabet: DNA, RNA or protein

Site distribution: "zoops" (zero or one of each motif in each sequence)

Number of motifs: 3

Motif length: 2-600 nucleotides

Strain	Genotype	Source
yMK467	*MATα ADE2 his3-11,15 leu2-3,112 trp1-1 ura3-1*	Ashe strain collection
yMK2476	yMK467 *MCP-GFP$_3$*	Ashe strain collection
yMK2520	yMK467 *MCP-GFP$_3$ p[NIP1-MS2L URA3]*	Ashe strain collection
yMK2521	yMK467 *MCP-GFP$_3$ p[PDC1-MS2L URA3]*	Ashe strain collection
yMK1859	yMK467 *NPC2-MS2L p[MCP-GFP$_3$ HIS3]*	Ashe strain collection
yMK3557	yMK467 *MCP-GFP$_3$ p[NIP1 PDC1 ORF-MS2L URA3]*	This study
yMK3558	yMK467 *MCP-GFP$_3$ p[NIP1 PDC1 5'UTR ORF-MS2L URA3]*	This study
yMK3559	yMK467 *MCP-GFP$_3$ p[NIP1 PDC1 ORF 3'UTR-MS2L URA3]*	This study
yMK3560	yMK467 *MCP-GFP$_3$ p[PDC1 NIP1 ORF-MS2L URA3]*	This study
yMK3561	yMK467 *MCP-GFP$_3$ p[NIP1-MS2L CEN URA3]*	This study
yMK3562	yMK467 *MCP-GFP$_3$ p[PDC1-MS2L CEN URA3]*	This study
yMK3563	yMK467 *MCP-GFP$_3$ p[NIP1 PDC1 ORF-MS2L CEN URA3]*	This study
yMK3437	yMK467 *ENO1-MS2L p[MCP-GFP$_3$ HIS3]*	Ashe strain collection
yMK1577	yMK467 *ENO2-MS2L p[MCP-GFP$_3$ HIS3]*	Ashe strain collection
yMK1585	yMK467 *TIF1-MS2L p[MCP-GFP$_3$ HIS3]*	Ashe strain collection
yMK2462	yMK467 *PDC5-MS2L p[MCP-GFP$_3$ HIS3]*	Ashe strain collection
yMK2416	yMK467 *FBA1-MS2L p[MCP-GFP$_3$ HIS3]*	Ashe strain collection
yMK2124	yMK467 *SUP35-MS2L p[MCP-GFP$_3$ HIS3]*	Ashe strain collection
yMK3220	yMK467 *ARO3-MS2L p[MCP-GFP$_3$ HIS3]*	Ashe strain collection
yMK3084	yMK467 *p[dCas9-GFP$_3$ HIS3]*	This study
yMK3230	yMK467 *p[dCas9-GFP$_3$ HIS3] p[NIP1 sgRNA LEU2]*	This study
yMK3231	yMK467 *p[dCas9-GFP$_3$ HIS3] p[TIF1 sgRNA LEU2]*	This study
yMK3232	yMK467 *p[dCas9-GFP$_3$ HIS3] p[ENO1 sgRNA LEU2]*	This study

yMK3233	yMK467 *p[dCas9-GFP$_3$ HIS3] p[ENO2 sgRNA LEU2]*	This study
yMK3234	yMK467 *p[dCas9-GFP$_3$ HIS3] p[NPC2 sgRNA LEU2]*	This study
yMK3235	yMK467 *p[dCas9-GFP$_3$ HIS3] p[SCR1 sgRNA LEU2]* (non-targeting sgRNA)	This study
yMK3281	yMK467 *p[dCas9-mNeonGreen$_3$ HIS3]*	This study
yMK3566	yMK467 *p[dCas9-mNeonGreen$_3$ HIS3] p[NIP1 sgRNA LEU2]*	This study
yMK3564	yMK467 *p[dCas9-mNeonGreen$_3$ HIS3] p[ENO1 sgRNA LEU2]*	This study
yMK3565	yMK467 *p[dCas9-mNeonGreen$_3$ HIS3] p[ENO2 sgRNA LEU2]*	This study
yMK3567	yMK467 *p[MCP-mNeonGreen$_3$ HIS3]*	This study
yMK1514	yMK467 *ENO2-MS2L*	Ashe strain collection
yMK3568	yMK467 *ENO2-MS2L p[MCP-mNeonGreen$_3$ HIS3]*	This study
yMK3080	yMK467 *ARO1-MS2L*	Ashe strain collection
yMK3569	yMK467 *ARO1-MS2L p[MCP-mNeonGreen$_3$ HIS3]*	This study
yMK3432	yMK467 *Nanobody-GFP*	This study
yMK3433	yMK467 *NB-GFP p[dCas9-MoonTag HIS3]*	This study
yMK3573	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[SCR1 sgRNA LEU2]* (non-targeting sgRNA)	This study
yMK3572	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[NPC2 sgRNA LEU2]*	This study
yMK3434	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[NIP1 sgRNA LEU2]*	This study
yMK3570	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[TIF1 sgRNA LEU2]*	This study
yMK3435	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[ENO1 sgRNA LEU2]*	This study
yMK3571	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[ENO2 sgRNA LEU2]*	This study
yMK3574	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[PDC1 sgRNA LEU2]*	This study
yMK3575	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[PDC5 sgRNA LEU2]*	This study
yMK3576	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[FBA1 sgRNA LEU2]*	This study
yMK3577	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[SUP35 sgRNA LEU2]*	This study

yMK3578	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[ARO3 sgRNA LEU2]*	This study
yMK3579	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[ERG1 sgRNA LEU2]*	This study
yMK3580	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[ERG2 sgRNA LEU2]*	This study
yMK3581	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[ERG4 sgRNA LEU2]*	This study
yMK3582	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[ERG10 sgRNA LEU2]*	This study
yMK3583	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[ERG20 sgRNA LEU2]*	This study
yMK3584	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[RPL28 sgRNA LEU2]*	This study
yMK3585	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[RPL3 sgRNA LEU2]*	This study
yMK3586	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[RPS31 sgRNA LEU2]*	This study
yMK3587	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[CCW12 sgRNA LEU2]*	This study
yMK3588	yMK467 *NB-GFP p[dCas9-MoonTag HIS3] p[ASC1 sgRNA LEU2]*	This study

Table 2: A list of yeast strains used in this study.

Strain	Name/Description	Source
BMK565	*p[MCP-GFP$_3$ CEN HIS3]* **MCP-GFP$_3$**	Ashe strain collection
BMK768	*p[NIP1-MS2L 2 µm URA3]* ***NIP1* 12x MS2 stem loops**	Ashe strain collection
BMK769	*p[PDC1-MS2L 2 µm URA3]* ***PDC1* 12x MS2 stem loops**	Ashe strain collection
BMK925	*p[NIP1 PDC1 ORF-MS2L 2 µm URA3]* ***NIP1 PDC1* ORF MS2 swap**	This study
BMK943	*p[NIP1 PDC1 5'UTR ORF-MS2L 2 µm URA3]* ***NIP1 PDC1* 5'UTR ORF MS2 swap**	This study
BMK944	*p[NIP1 PDC1 ORF 3'UTR-MS2L 2 µm URA3]* ***NIP1 PDC1* ORF MS2 3'UTR swap**	This study
BMK945	*p[PDC1 NIP1 ORF-MS2L 2 µm URA3]* ***PDC1 NIP1* ORF MS2 swap**	This study
BMK466	*pRS316* **CEN origin of replication**	Addgene
BMK956	*p[NIP1-MS2L CEN URA3]* ***NIP1* 12x MS2 stem loops CEN**	This study
BMK957	*p[PDC1-MS2L CEN URA3]* ***PDC1* 12x MS2 stem loops CEN**	This study
BMK958	*p[NIP1 PDC1 ORF-MS2L CEN URA3]* ***NIP1 PDC1* ORF MS2 swap CEN**	This study
BMK876	*p[dCas9-NLS mKO2]* **dCas9 ORF NLS (mammalian vector)**	Kind gift from Andy Badrock
BMK864	*p[dCas9-GFP$_3$ CEN HIS3]* **dCas9-GFP$_3$**	This study
BMK825	*pMEL14* **sgRNA vector, contains scaffold and 20 bp variable regions**	Addgene
BMK878	*p[NIP1 sgRNA 2 µm LEU2]* ***NIP1* sgRNA**	This study
BMK881	*p[TIF1 sgRNA 2 µm LEU2]* ***TIF1* sgRNA**	This study
BMK884	*p[ENO1 sgRNA 2 µm LEU2]* ***ENO1* sgRNA**	This study
BMK887	*p[ENO2 sgRNA 2 µm LEU2]* ***ENO2* sgRNA**	This study
BMK891	*p[NPC2 sgRNA 2 µm LEU2]* ***NPC2* sgRNA**	This study
BMK893	*p[SCR1 sgRNA 2 µm LEU2]* **Non-targeting sgRNA**	This study
BMK902	*p[mNeonGreen-IFT88]* **mNeonGreen (mammalian vector)**	Kind gift from Mark Johnston
BMK921	*p[mNeonGreen$_3$ CEN HIS3]* **mNeonGreen$_3$**	This study
BMK922	*p[dCas9-mNeonGreen$_3$ CEN HIS3]* **dCas9-mNeonGreen$_3$**	This study

BMK942	*p[MCP-mNeonGreen$_3$ CEN HIS3]* **MCP-mNeonGreen$_3$**	This study
BMK930	*p[anti-gp41 nanobody-GFP]* **Nanobody-GFP (synthesised in pBluescript SK)**	Epoch Life Sciences
BMK928	*p[Kif18b-24xgp41 5aa linker]* **MoonTag sequence**	Kind gift from Tanenbaum lab
BMK931	*p[dCas9-MoonTag CEN HIS3]* **dCas9-MoonTag**	This study
BMK908	*p[PDC1 sgRNA 2 µm LEU2]* ***PDC1* sgRNA**	This study
BMK909	*p[PDC5 sgRNA 2 µm LEU2]* ***PDC5* sgRNA**	This study
BMK911	*p[FBA1 sgRNA 2 µm LEU2]* ***FBA1* sgRNA**	This study
BMK912	*p[SUP35 sgRNA 2 µm LEU2]* ***SUP35* sgRNA**	This study
BMK910	*p[ARO3 sgRNA 2 µm LEU2]* ***ARO3* sgRNA**	This study
BMK946	*p[ERG1 sgRNA 2 µm LEU2]* ***ERG1* sgRNA**	This study
BMK947	*p[ERG2 sgRNA 2 µm LEU2]* ***ERG2* sgRNA**	This study
BMK948	*p[ERG4 sgRNA 2 µm LEU2]* ***ERG4* sgRNA**	This study
BMK949	*p[ERG10 sgRNA 2 µm LEU2]* ***ERG10* sgRNA**	This study
BMK950	*p[ERG20 sgRNA 2 µm LEU2]* ***ERG20* sgRNA**	This study
BMK951	*p[RPL28 sgRNA 2 µm LEU2]* ***RPL28* sgRNA**	This study
BMK952	*p[RPL3 sgRNA 2 µm LEU2]* ***RPL3* sgRNA**	This study
BMK953	*p[RPS31 sgRNA 2 µm LEU2]* ***RPS31* sgRNA**	This study
BMK954	*p[CCW12 sgRNA 2 µm LEU2]* ***CCW12* sgRNA**	This study
BMK955	*p[ASC1 sgRNA 2 µm LEU2]* ***ASC1* sgRNA**	This study

Table 3: A list of bacterial plasmids used in this study.

Name	Purpose	Sequence 5’ to 3’
M13 **F**	Sequencing: *NIP1 PDC1* swaps dCas9 and MCP vectors sgRNA vectors	GTAAAACGACGGCCAGT
M13 **R**		CAGGAAACAGCTATGAC
PDC1 MS2 ORF **F**	*NIP1 PDC1* ORF swap PCR fragments for Gibson cloning	CCTGCCAAGCCCGAGATCTACGAAAA ATGTCTGAAATTACTTTGGGTAAA
PDC1 MS2 ORF **R**		TCGAAGTAGTTTTTGATTTTATTGAGC ATAGGCCACTAGTGGATCTGATA
NIP1 backbone (5’UTR and 3’UTR) **F**		TATCAGATCCACTAGTGGCCTATGCT CAATAAAATCAAAAACTACTTCGA
NIP1 backbone (5’UTR and 3’UTR) **R**		TTTACCCAAAGTAATTTCAGACATTTT TCGTAGATCTCGGGCTTGGCAGG
NIP1 MS2 ORF **F**	*PDC1 NIP1* ORF swap PCR fragments for Gibson cloning	AAAATAACACAGTCAAATCAATCAAAA TGTCCCGTTTCTTTTCGTCTAATT
NIP1 MS2 ORF **R**		AACTAATAATTAGAGATTAAATCGCGC ATAGGCCACTAGTGGATCTGATA
PDC1 backbone (5’UTR and 3’UTR) **F**		TATCAGATCCACTAGTGGCCTATGCG CGATTTAATCTCTAATTATTAGTT
PDC1 backbone (5’UTR and 3’UTR) **R**		AATTAGACGAAAAGAAACGGGACATT TTGATTGATTTGACTGTGTTATTT
NIP1 3'UTR **F**	*PDC1 NIP1* 5’UTR ORF swap PCR fragments for Gibson cloning	TATCAGATCCACTAGTGGCCTATGCT CAATAAAATCAAAAACTACTTCGA
NIP1 3'UTR **R**		AAAAGTCAGAAGAGCATACATAATCA TCGAATTCCCGCGGCCGCCATGGC
PDC1 5'UTR ORF MS2 **F**		GCCATGGCGGCCGCGGGAATTCGAT GATTATGTATGCTCTTCTGACTTTT
PDC1 5'UTR ORF MS2 **R**		TCGAAGTAGTTTTTGATTTTATTGAGC ATAGGCCACTAGTGGATCTGATA
NIP1 5'UTR **F**	*PDC1 NIP1* ORF 3’UTR swap PCR fragments for Gibson cloning	AAGTTGCCGACAGTCTGTTGAATTGA TCACTAGTGAATTCGCGGCCGCCT
NIP1 5'UTR **R**		TTTACCCAAAGTAATTTCAGACATTTT TCGTAGATCTCGGGCTTGGCAGG
PDC1 ORF 3’UTR MS2 **F**		CCTGCCAAGCCCGAGATCTACGAAAA ATGTCTGAAATTACTTTGGGTAAA
PDC1 ORF 3’UTR MS2 **R**		AGGCGGCCGCGAATTCACTAGTGATC AATTCAACAGACTGTCGGCAACTT
NIP1 stem loop seq **F**	To amplify stem loop regions for *NIP1 and NIP1 PDC1* ORF	CCACCATCAAATCGTCGTTGA
NIP1 stem loop seq **R**		TGATTTTATTGAGCATAGGCC
PDC1 stem loop seq **F**	To amplify stem loop regions for *PDC1 and NIP1 PDC1* ORF	TGGTTGAACAAGCTAAGTTGA
PDC1 stem loop seq **R**		AGATTAAATCGCGCATAGGCC
CEN sequencing **F**	To sequence across the CEN ORI	GGGTCCTTTTCATCACGTGC

NIP1 qRT-PCR **F**	Targets *NIP1* expressed on plasmid for 2 µm and CEN	CCACCATCAAATCGTCGTTGA
NIP1 qRT-PCR **R**		GCATACATTATACGAAGTTAT
PDC1 qRT-PCR **F**	Targets *PDC1* and *NIP1 PDC1* ORF expressed on plasmid for 2 µm and CEN	AAGGTTGGGACCACCTATCC
PDC1 qRT-PCR **R**		GCATACATTATACGAAGTTAT
ACT1 qRT-PCR **F**	Reference gene for qRT-PCR	CACCGCTTTGGCTCCATCTTCC
ACT1 qRT-PCR **R**		GTAGATAAAGTCAGTGCTTAAAC
dCas9 ORF Gibson **F**	To amplify dCas9 ORF NLS PCR fragments for dCas9-GFP_3 Gibson cloning	AGAGCCCTCAACCGGAGTTTGAAGCA TGGATAAGAAGTATAGCATCGGCC
dCas9 ORF Gibson **R**		CGAATTCCTGCAGCCCGGGGGATCC CACCTTTCTCTTCTTCTTAGGAGAT
MCP-GFP_3 Gibson **F**	To amplify GFP_3 without MCP PCR fragments for dCas9-GFP_3 Gibson cloning	ATCTCCTAAGAAGAAGAGAAAGGTGG GATCCCCCGGGCTGCAGGAATTCG
MCP-GFP_3 Gibson **R**		GGCCGATGCTATACTTCTTATCCATG CTTCAAACTCCGGTTGAGGGCTCT
mNeonGreen 1 **F**	PCR fragments designed for dCas9-$mNeonGreen_3$ Gibson cloning Made $mNeonGreen_3$ in the vector backbone only (usable)	CGGGCTGCAGGAATTCGATATCGTGA TGGTGAGCAAGGGCGAGGAGGATA
mNeonGreen 1 **R**		TATCCTCCTCGCCCTTGCTCACCATG GTACCCTTGTACAGCTCGTCCATG
mNeonGreen 2 **F**		CATGGACGAGCTGTACAAGGGTACCA TGGTGAGCAAGGGCGAGGAGGATA
mNeonGreen 2 **R**		TATCCTCCTCGCCCTTGCTCACCATG GTACCCTTGTACAGCTCGTCCATG
mNeonGreen 3 **F**		CATGGACGAGCTGTACAAGGGTACCA TGGTGAGCAAGGGCGAGGAGGATA
mNeonGreen 3 **R**		ATAACTAATTACATGACTCGACCAGCT ACTTGTACAGCTCGTCCATGCCCATC
dCas9 without GFP **F**	To amplify dCas9 backbone without GFP_3 PCR fragments for dCas9-MoonTag Gibson cloning	TCTGCACCGGTCAGCAGTGGAGGAC CTGGTCGAGTCATGTAATTAGTTAT
dCas9 without GFP **R**		GCATGGTGGCAAGCTTAAGTTTAAAA TCGAATTCCTGCAGCCCGGGGGAT
MoonTag array **F**	To amplify MoonTag array PCR fragments for dCas9-MoonTag Gibson cloning	ATCCCCCGGGCTGCAGGAATTCGATT TTAAACTTAAGCTTGCCACCATGCGT
MoonTag array **R**		ATAACTAATTACATGACTCGACCAGG TCCTCCACTGCTGACCGGTGCAGAGA
gp41 genome verification **F**	To verify anti-gp41 nanobody-GFP insertion into the yeast genome (Hygromycin resistance marker)	ACAAGATGTAAAGATAATGC
gp41 genome verification **R**		CTCCTGATTCCGCTAATAGG
Hygromycin verification **F**		CGTGGTTGGCTTGTATGGAG
Hygromycin verification **R**		CTCCATACAAGCCAACCACG

sgRNA qRT-PCR **F**	Targets the conserved scaffold region	GTTTTAGAGCTAGAAATAGCA
sgRNA qRT-PCR **R**		CACCGACTCGGTGCCACTTTT
NIP1 qRT-PCR **F**	Targets *NIP1* 3'UTR To test mRNA levels when *NIP1* is targeted by dCas9-MT	CTACTTCGAAAGCATTCTTAAC
NIP1 qRT-PCR **R**		GGTTTATACTCTCCTGCTTAG
TIF1 qRT-PCR **F**	Targets *TIF1* 3'UTR To test mRNA levels when *TIF1* is targeted by dCas9-MT	ACCTCTGTCAAGGCTCCTCA
TIF1 qRT-PCR **R**		CCTGGAGTACCAACGACGAT
ENO1 qRT-PCR **F**	Targets *ENO1* 3'UTR To test mRNA levels when *ENO1* is targeted by dCas9-MT	ACGTTTTGCCAGTTCCATTC
ENO1 qRT-PCR **R**		GTTTGGAGCAACACCACCTT
ENO2 qRT-PCR **F**	Targets *ENO2* 3'UTR To test mRNA levels when *ENO2* is targeted by dCas9-MT	ATCCGAAAGATTGGCTAAGTTG
ENO2 qRT-PCR **R**		CTTGTCACCGTGGTGGAAG

Table 4: A list of primers used in this study.

Name	Description	Sequence
NIP1 sgRNA **F**	Targeting region shown in red. Targets the 3'UTR.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCTTATACTCTCCTGCTTAGTTGTTTTAGAGCTAGAAA TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
NIP1 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACAACTAAGCAGGAGAGTATAAGATCATTTATCTTTCACT GCGGAGAAGTTTCGAACGCCGAAACATGCGCA
TIF1 sgRNA **F**	Targeting region shown in red. Targets the 3'UTR.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCTATGTTACGTTATCAAGATAGTTTTAGAGCTAGAAA TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
TIF1 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACTATCTTGATAACGTAACATAGATCATTTATCTTTCACT GCGGAGAAGTTTCGAACGCCGAAACATGCGCA
ENO1 sgRNA **F**	Targeting region shown in red. Targets the 3'UTR.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCATGAAAAATGCTTGTTTGCAGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
ENO1 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACTGCAAACAAGCATTTTTCATGATCATTTATCTTTCACT GCGGAGAAGTTTCGAACGCCGAAACATGCGCA
ENO2 sgRNA **F**	Targeting region shown in red. Targets the 3'UTR.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCAAAACTGTATCAATTTTTAAGTTTTAGAGCTAGAAA TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
ENO2 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACTTAAAAATTGATACAGTTTTGATCATTTATCTTTCACT GCGGAGAAGTTTCGAACGCCGAAACATGCGCA
NPC2 sgRNA **F**	Targeting region shown in red. Targets the 3'UTR.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCAGTCATCTTCCCACCAAGGTGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
NPC2 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACACCTTGGTGGGAAGATGACTGATCATTTATCTTTCAC TGCGGAGAAGTTTCGAACGCCGAAACATGCGCA

Non-targeting sgRNA **F**	Targeting region shown in red. Designed not to target any mRNA.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCAGGCATACGACTGAACGTTGGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
Non-targeting sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACCAACGTTCAGTCGTATGCCTGATCATTTATCTTTCACT GCGGAGAAGTTTCGAACGCCGAAACATGCGCA
PDC1 sgRNA **F**	Targeting region shown in red. Targets the ORF.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCCATTGACACCATCTTGGCTTGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
PDC1 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACAAGCCAAGATGGTGTCAATGGATCATTTATCTTTCAC TGCGGAGAAGTTTCGAACGCCGAAACATGCGCA
PDC5 sgRNA **F**	Targeting region shown in red. Targets the ORF.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCTGTTAGAACTGTTGTTGAATGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
PDC5 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACATTCAACAACAGTTCTAACAGATCATTTATCTTTCACT GCGGAGAAGTTTCGAACGCCGAAACATGCGCA
FBA1 sgRNA **F**	Targeting region shown in red. Targets the 3'UTR.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCGTTCTTCGACCCAAGAGTCTGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
FBA1 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACAGACTCTTGGGTCGAAGAACGATCATTTATCTTTCAC TGCGGAGAAGTTTCGAACGCCGAAACATGCGCA
SUP35 sgRNA **F**	Targeting region shown in red. Targets the 3'UTR.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCGTAGATTCACTTTGAGAGATGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
SUP35 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACATCTCTCAAAGTGAATCTACGATCATTTATCTTTCACT GCGGAGAAGTTTCGAACGCCGAAACATGCGCA
ARO3 sgRNA **F**	Targeting region shown in red. Targets the 3'UTR.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCGGAGTCCACCGAACAGGTATGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
ARO3 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACATACCTGTTCGGTGGACTCCGATCATTTATCTTTCAC TGCGGAGAAGTTTCGAACGCCGAAACATGCGCA

ERG1 sgRNA **F**	Targeting region shown in red. Targets the 3'UTR.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCGGAACCGTCAAACATTAAGCGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
ERG1 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACGCTTAATGTTTGACGGTTCCGATCATTTATCTTTCACT GCGGAGAAGTTTCGAACGCCGAAACATGCGCA
ERG2 sgRNA **F**	Targeting region shown in red. Targets the 3'UTR.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCCGTTCTGCGATTTTCTCATGGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
ERG2 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACCATGAGAAAATCGCAGAACGGATCATTTATCTTTCAC TGCGGAGAAGTTTCGAACGCCGAAACATGCGCA
ERG4 sgRNA **F**	Targeting region shown in red. Targets the 3'UTR.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCGCCCGACCTCACGAAATCAGGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
ERG4 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACCTGATTTCGTGAGGTCGGGCGATCATTTATCTTTCAC TGCGGAGAAGTTTCGAACGCCGAAACATGCGCA
ERG10 sgRNA **F**	Targeting region shown in red. Targets the 3'UTR.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCCTATCCACTTCATTAACTTAGTTTTAGAGCTAGAAA TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
ERG10 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACTAAGTTAATGAAGTGGATAGGATCATTTATCTTTCACT GCGGAGAAGTTTCGAACGCCGAAACATGCGCA
ERG20 sgRNA **F**	Targeting region shown in red. Targets the 3'UTR.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCCTAACGCTAATCGATAAAACGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
ERG20 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACGTTTTATCGATTAGCGTTAGGATCATTTATCTTTCACT GCGGAGAAGTTTCGAACGCCGAAACATGCGCA
RPL28 sgRNA **F**	Targeting region shown in red. Targets the 3'UTR.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCGTTTAATCTAATTCAACATCGTTTTAGAGCTAGAAA TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
RPL28 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACGATGTTGAATTAGATTAAACGATCATTTATCTTTCACT GCGGAGAAGTTTCGAACGCCGAAACATGCGCA

RPL3 sgRNA **F**	Targeting region shown in red. Targets the 3'UTR.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCGTTAGCTACATACAACAGTTGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
RPL3 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACAACTGTTGTATGTAGCTAACGATCATTTATCTTTCACT GCGGAGAAGTTTCGAACGCCGAAACATGCGCA
RPS31 sgRNA **F**	Targeting region shown in red. Targets the 3'UTR.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCATAGACAGTAGAGGAATATAGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
RPS31 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACTATATTCCTCTACTGTCTATGATCATTTATCTTTCACT GCGGAGAAGTTTCGAACGCCGAAACATGCGCA
CCW12 sgRNA **F**	Targeting region shown in red. Targets the 3'UTR.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCTGGCATTTAACAAATATATAGTTTTAGAGCTAGAAA TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
CCW12 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACTATATATTTGTTAAATGCCAGATCATTTATCTTTCACT GCGGAGAAGTTTCGAACGCCGAAACATGCGCA
ASC1 sgRNA **F**	Targeting region shown in red. Targets the 3'UTR.	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGAT AAATGATCATCGTCATAGATTTCGAAGTGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
ASC1 sgRNA **R**		GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTC TAAAACACTTCGAAATCTATGACGATGATCATTTATCTTTCACT GCGGAGAAGTTTCGAACGCCGAAACATGCGCA

Table 5: A list of sgRNA primers produced and used in this study.

Results

Chapter 3: Investigating the mechanisms underlying the differential localisation of *NIP1* and *PDC1* mRNAs.

3. Investigating the mechanisms underlying the differential localisation of *NIP1* and *PDC1* mRNAs.

Previous work in the Ashe lab has identified novel types of cytosolic RNA granules. These granules are distinct from other cytosolic granules (such as stress granules or P-bodies); they are present in unstressed yeast cells and the mRNAs they contain are actively translated. Hence, the granules have been termed 'translation factories' or actively translating granules (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). Further work has been done to characterise at least two different types of actively translating granule. The first, named Core Fermentation (CoFe) granules, contain mRNAs encoding glycolytic enzymes (such as *PDC1*, *ENO2*, and *FBA1*) and are present in 10-20 granules per cell (Morales-Polanco et al., 2021). The second, named translation factor granules, contain mRNAs encoding translation factors (such as *NIP1*, *TIF1*, and *SUP35*) and are present in 1-2 granules per cell (Pizzinga et al., 2019). Other mRNAs are not present in cytosolic granules under the same conditions (such as *NPC2*, *GIP2*, *CIN5*, *VPS24* and *ERP4*) (Lui et al., 2014). Further experiments in the lab are currently being performed to investigate the RNA and protein components of these granules and what the purpose of these granules may be, while experiments were performed here to begin to investigate the mechanism of mRNA localisation to these granules.

The initial goal here was to investigate whether *cis*-acting RNA sequences are involved in signalling the localisation status of a specific mRNA. A strategy to explore the different patterns of cytosolic RNA granules was settled upon where 'mRNA swap' constructs were made using *NIP1* and *PDC1* derived plasmids. The *PDC1* and *NIP1* mRNAs had been previously investigated using the mTAG system, and so MS2 stem loops had been inserted into the 3'UTR sequence at the genomic loci to allow mRNA visualisation in live yeast cells (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). Initially, the intention was to swap each part of these mRNAs; their 5'UTR, ORF, and 3'UTR and also their control regions (promoter and terminator) to identify region(s) that impact on mRNA localisation (Figure 3.1). However, due to a series of cloning issues, which most likely stemmed from the twelve repeated MS2 stem loop sequences, only a limited number of constructs were made here.

Nevertheless, a key observation was made that the exchange of the *NIP1* ORF for the *PDC1* ORF (while maintaining the *NIP1* promoter, 5'UTR and 3'UTR) was sufficient to drive the chimeric mRNA into the localisation pattern seen for *PDC1* mRNA rather than *NIP1* mRNA. The opposite mRNA swap, where the *NIP1* ORF replaced the *PDC1* ORF (maintaining the *PDC1* promoter, 5'UTR and 3'UTR) similarly changes the cytosolic RNA granule pattern to

appear more like endogenous *NIP1*. These results suggest that either RNA localisation elements are present in the ORFs of *NIP1* and *PDC1* mRNAs, or that the production of the nascent protein is capable of driving its own mRNA localisation. mRNA sequence analysis was also performed using MEME motif software (Bailey et al., 2006; Bailey et al., 2009). This analysis demonstrated conserved motifs in the ORFs of glycolytic mRNAs that were not present in translation factor mRNAs, suggesting that sequence and/or structural elements may be present in the ORF of these mRNAs that drive their localisation.

Experiments were also performed to investigate the effect of mRNA abundance on the pattern of cytosolic RNA granules. Glycolytic mRNAs are highly abundant mRNAs and localise to multiple granules per cell, while translation factor mRNAs are lower abundance mRNAs and localise to one or two granules per cell (Lui et al., 2004; Pizzinga et al., 2019; Morales-Polanco et al., 2021). Furthermore, *YEF3* and *TEF1* are two mRNAs that also encode translation factors, but are higher abundance than most other translation factor mRNAs (such as *NIP1* and *TIF1*) (Pizzinga et al., 2019). *YEF3* and *TEF1* have been shown to localise to multiple granules per cell, rather than one or two granules per cell, although they do not appear to co-localise with glycolytic mRNAs (Pizzinga et al., 2019). It is therefore possible that the abundance of an mRNA correlates with the number of cytosolic RNA granules they are able to form. To investigate this, *NIP1*, *PDC1* and the *NIP1 PDC1* ORF construct made here were expressed on a low-copy number (CEN) yeast expression plasmid to reduce their mRNA levels. The RNA granule pattern for each of these constructs remained the same when expressed from CEN plasmids, indicating that the multiple RNA granule pattern seen for the *NIP1 PDC1* ORF construct was not dependent on mRNA copy number but more likely be an intrinsic feature of the mRNA sequence and/or structure.

Overall experiments performed here provide an initial starting platform to investigate mechanisms of mRNA localisation to cytosolic RNA granules and suggest many interesting lines of further inquiry. Although the exact mechanisms of the differential mRNA localisation patterns of *NIP1* and *PDC1* have yet to be identified, results here suggest that either RNA localisation elements may be found in the ORFs of these mRNAs, or that the production or the nascent protein may act upon its own mRNA to drive localisation.

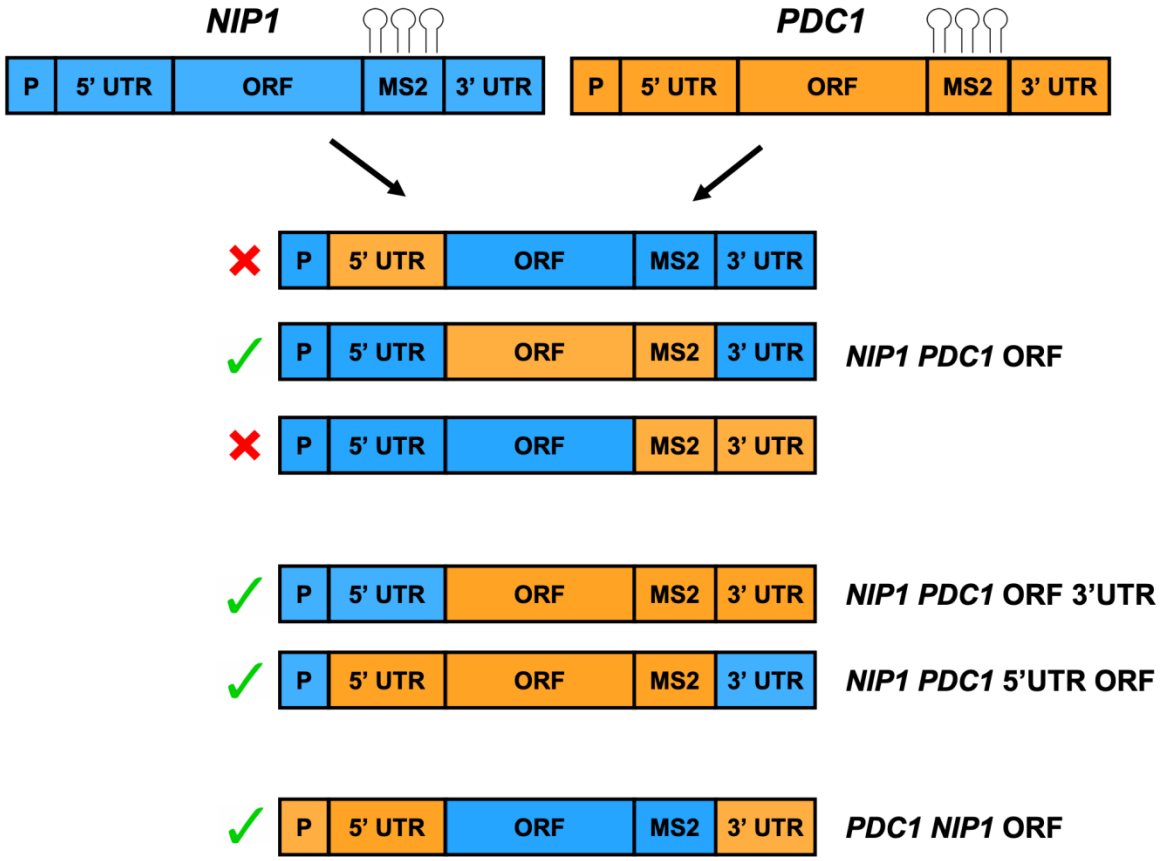


Figure 3.1: An overview of *NIP1*/*PDC1* swap constructs attempted and produced. Various swap constructs were made by swapping the 5'UTR, ORF and/or 3'UTR of *NIP1* and *PDC1* mRNA. Successfully made constructs are denoted with a green tick, and referred to by these names throughout the chapter. Constructs that were not successfully produced (*PDC1* 5'UTR and *PDC1* 3'UTR swaps) are shown by a red cross. The MS2 stem loop regions are identical in the parent *NIP1* and *PDC1* sequences, and the swap construct colours show which parent the MS2 region originated from (blue for *NIP1* and orange for *PDC1*). P denotes the promoter region for each construct.

3.1 Verification of the previously established mRNA localisation status for *NIP1* and *PDC1* mRNAs:

As stated above, previous studies in the Ashe lab had assessed the localisation of *NIP1* (encoding the translation factor eIF3b) and *PDC1* (encoding the glycolytic pyruvate decarboxylase enzyme) (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). To confirm these previous results, strains were retrieved from the Ashe lab strain collection that carried either MS2-tagged *NIP1* or *PDC1* and expressed an MS2 coat protein (MCP)-GFP_3 fusion protein. Fluorescent microscopy analysis revealed that these strains have mRNA granules (Figure 3.2) and that the number of mRNA granules per cell closely matches that described in published work from the Ashe lab (Figure 3.3; Pizzinga et al., 2019; Morales-Polanco et al., 2021). Importantly, these initial experiments also revealed that a strain expressing a previously characterised non-localised MS2-tagged mRNA *NPC2* was also non-localised here (Figure 3.2).

Therefore, experiments were performed here to determine the sequences present in the *NIP1* and *PDC1* mRNAs (and not present in *NPC2* mRNA) that are responsible for driving differential patterns of mRNA localisation to cytosolic RNA granules.

At the outset of the project, the genomic MS2-tagged strains were not considered optimal for an investigation of *cis*-acting sequences in the mRNAs. Instead, a plasmid-based system was deemed more suitable, as it should allow more straightforward alterations of the *NIP1* and *PDC1* mRNAs without affecting cell growth and metabolism.

Previous work in the lab had isolated the *PDC1* and *NIP1* MS2-tagged genes from the genomically tagged strains (Lui et al., 2014; Pizzinga et al., 2019) and inserted these into the yeast YEplac195 vector (Gietz and Sugino, 1988). These plasmids were individually transformed into the yeast strain, yMK2476, which had previously been constructed in the Ashe lab to genomically express the MCP-GFP_3 fusion required for MS2 imaging (Lui et al., 2014). A cross-comparison of the microscopic images and the quantitation revealed that both the *NIP1* and *PDC1* mRNAs generated similar profiles and numbers of granules per cell regardless of whether the mRNAs derive from the genomic copy (shown in previous work) (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021) of the gene or from a plasmid copy (Figures 3.2 and 3.3).

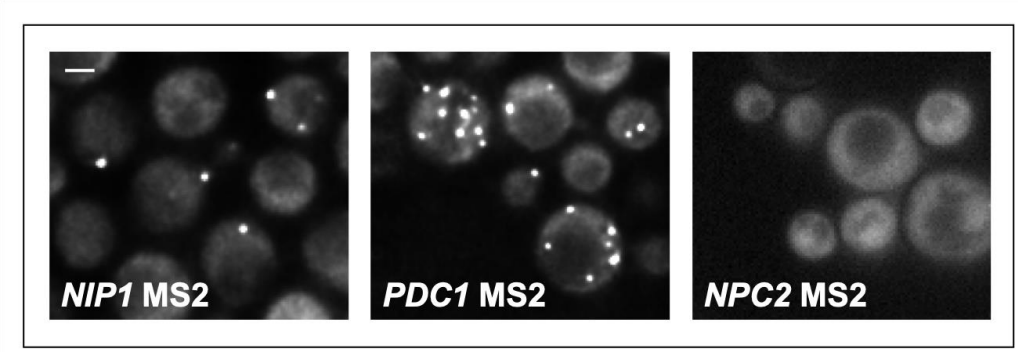


Figure 3.2: mRNA localisation of *NIP1*, *PDC1* and *NPC2* mRNAs. Images of yeast strains carrying MS2-tagged *NIP1*, *PDC1* and *NPC2* mRNAs. *NIP1* mRNA appears in one or two granules per cell, *PDC1* mRNA in several granules per cell and *NPC2* mRNA does not appear in cytosolic granules. Images were taken as Z-stacks and the maximum intensity projection is shown. Scale bar: 2µm.

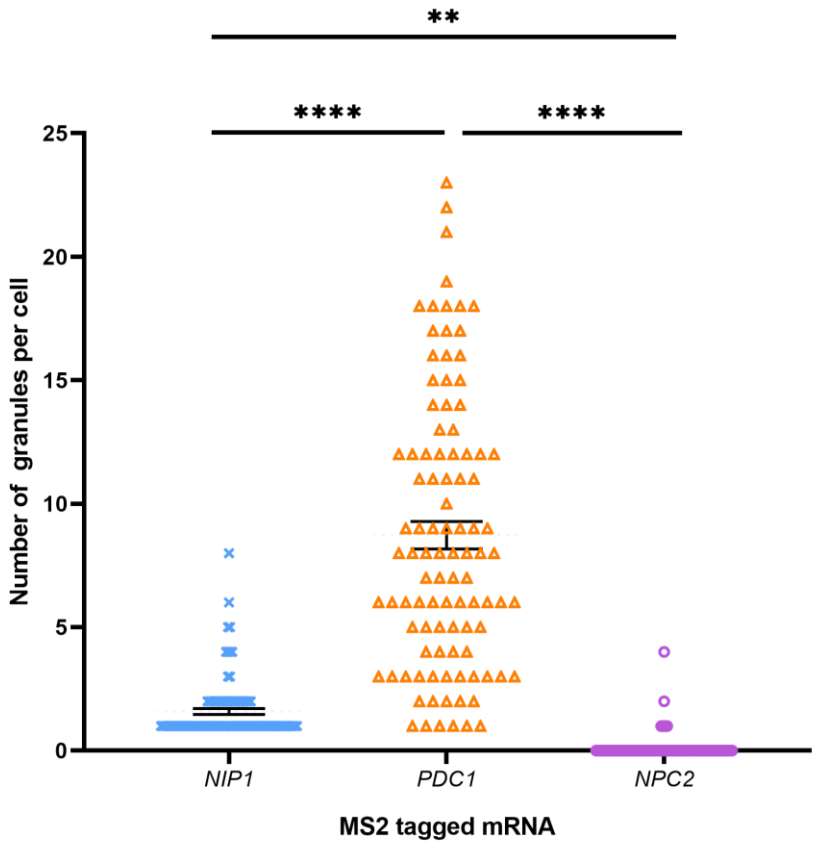


Figure 3.3: Abundance of mRNA granules for *NIP1*, *PDC1* and *NPC2* mRNAs. For each mRNA the number of granules per cell is shown (n=100). For *NIP1* (shown in blue) the average number is 1-2 granules per cell, for *PDC1* (orange) the average number is 8-9 granules per cell and for *NPC2* (purple) the average number is 0 granules per cell. Error bars are shown as the mean with standard deviation. One-way ANOVA statistical analysis showed significant differences between each of the mRNAs ($P<0.0001$ for *NIP1*/*PDC1* and *PDC1*/*NPC2* comparisons, and $P = 0.0049$ for the *NIP1*/*NPC2* comparison).

3.2. Construction and imaging of the *NIP1 PDC1* ORF and *PDC1 NIP1* ORF constructs:

In order to assess the whether *cis*-acting sequences might be important in the localisation of the *NIP1* and PDC1 mRNAs, a series of 'mRNA swap' plasmid constructs were designed. The first successful swap plasmid generated was the "*NIP1 PDC1* ORF" construct, which contained the backbone and untranslated regions of *NIP1* (promoter, 5'UTR and 3'UTR) and the ORF of *PDC1* (Figure 3.1). The MS2 stem loop region also originated from the *PDC1* parent vector, as PCR reactions were more often successful when the MS2 region was present on the shorter PCR fragment (in this case the *PDC1* ORF (~1600 bp) rather than the plasmid backbone and *NIP1* regions (~6600 bp)). This ORF swap construct was made by Gibson cloning, where the *PDC1* ORF MS2 region replaced the *NIP1* ORF MS2 region in the plasmid (Figure 3.4). The *NIP1 PDC1* ORF construct was then verified by restriction enzyme mapping and DNA sequencing across the untranslated region (Figure 3.5).

The *NIP1 PDC1* ORF construct was then transformed into the yeast strain yMK2476, which expresses the MCP-GFP_3 fusion protein from an integrated cassette on the genome (Lui et al., 2014), and imaged. Imaging demonstrated the *NIP1 PDC1* ORF construct displayed a pattern of cytosolic mRNA granules more similar to *PDC1* than *NIP1* (Figure 3.6). Several cytosolic granules could be seen in almost all cells imaged; a stark difference to *NIP1* mRNA, where cells most often showed a single granule per cell. Quantification of these cells (n=100) was done by computerised image analysis of 100 yeast cells, where the number of granules in each cell was counted (described fully in methods section 2.6). The quantification showed that the *NIP1 PDC1* ORF construct showed the same pattern of cytosolic RNA granules as *PDC1*, around 9 or 10 granules per cell (Figure 3.7). These results therefore suggested that the mRNA localisation signal for *PDC1* mRNA is either an RNA element present in the *PDC1* ORF sequence, or that the production of the Pdc1 nascent protein acts to drive localisation of its own mRNA.

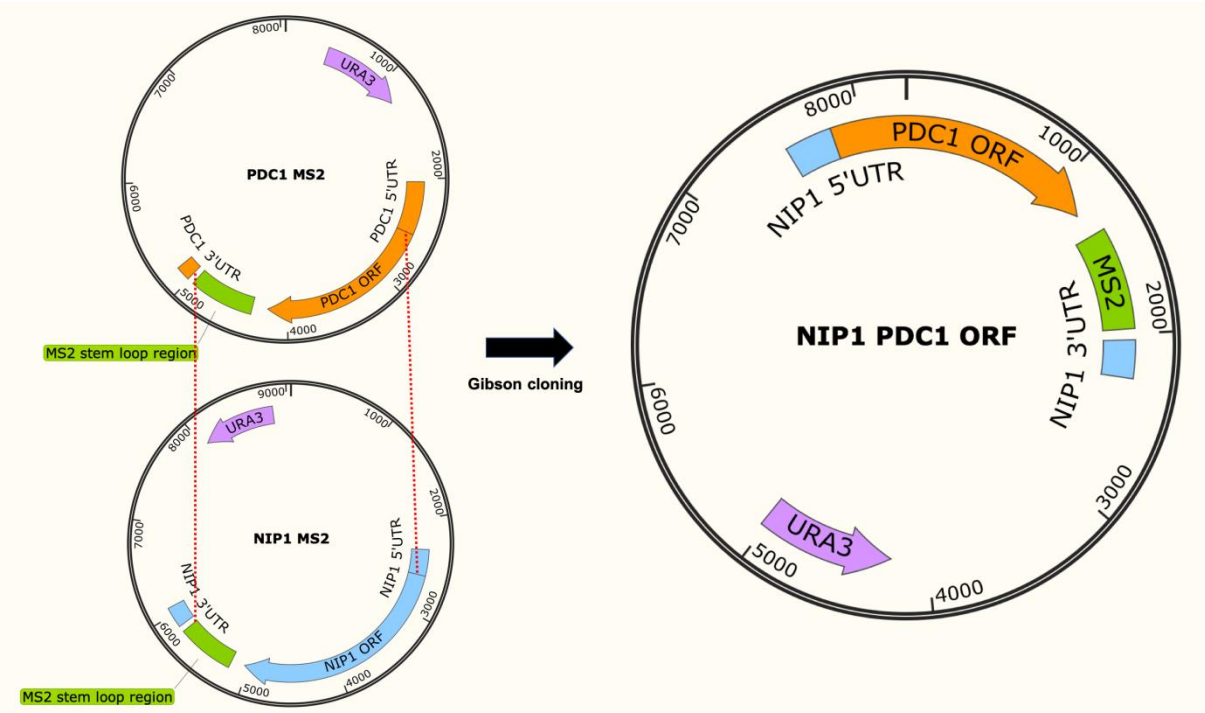


Figure 3.4: Gibson cloning of the *NIP1 PDC1* ORF swap construct. Gibson cloning was performed to replace the ORF and MS2 region of *NIP1* with the ORF and MS2 region the *PDC1* vector (the MS2 stem loops sequences were identical). The resulting vector contained the *NIP1* 5'UTR, *PDC1* ORF, MS2 region and *NIP1* 3'UTR, in a yeast expression vector.

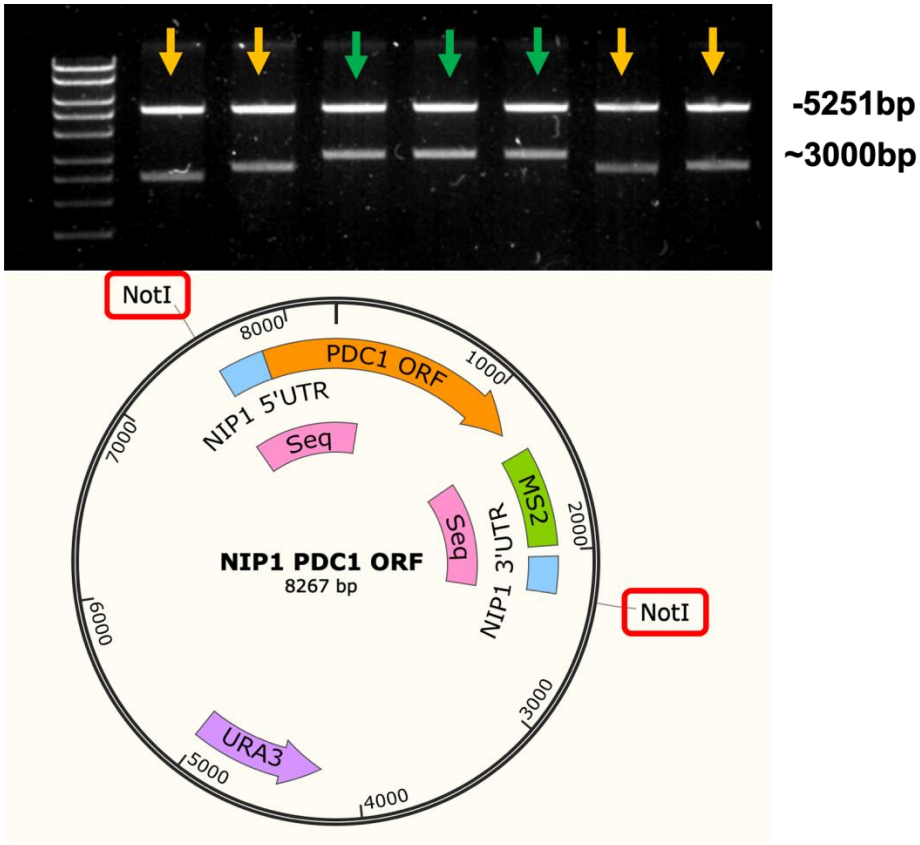


Figure 3.5: Restriction enzyme mapping of the *NIP1 PDC1* ORF swap construct. All gels are shown with HyperLadder 1kb markers. The gel shows 7 possible *NIP1 PDC1* ORF samples digested with *Not*I restriction enzyme. All samples show a band at 5251bp corresponding to the plasmid backbone. Potentially correct samples (green arrows) also produced a band at 3014bp which would correspond to the *NIP1* UTRs and *PDC1* ORF, with 12 MS2 stem loops (the *NIP1* ORF is longer than the *PDC1* ORF and would have produced a band at 3792bp). These constructs were verified by sequencing across the junction regions and MS2 region (shown highlighted in pink in the plasmid representation). Samples that showed lower bands on the gel (orange arrows) were shown by sequencing to contain the *NIP1* UTRs and *PDC1* ORF, but fewer numbers of MS2 stem loops (8-10 loops).

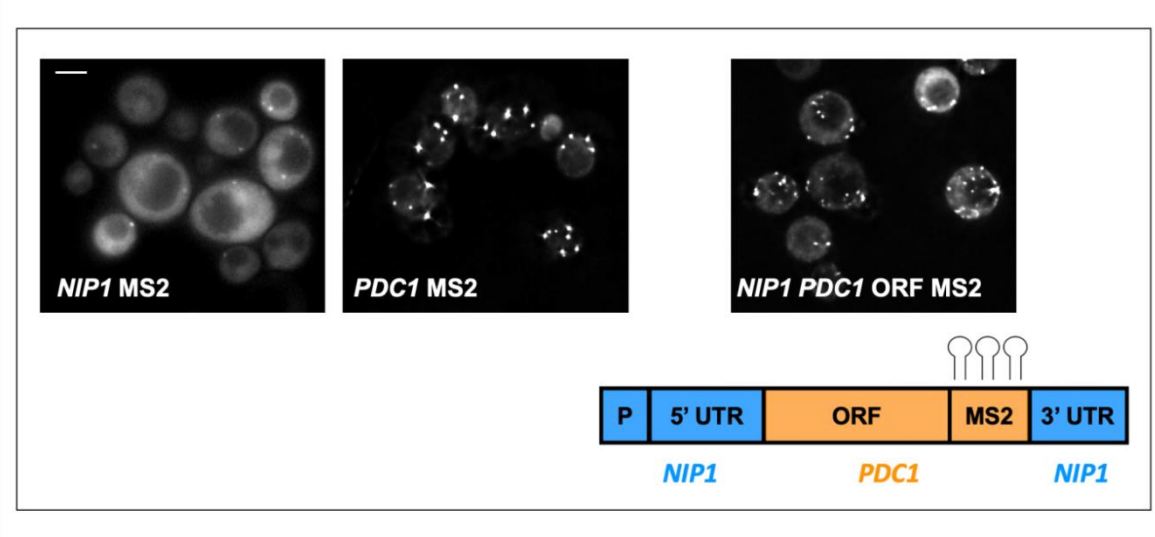


Figure 3.6: Imaging of the *NIP1 PDC1* ORF swap construct. The first two images show yeast cells expressing the MS2-tagged *NIP1* and *PDC1* parent vectors, along with genomically expressed MCP-GFP_3. The final image shows yeast cells also genomically expressing MCP-GFP_3 along with the *NIP1 PDC1* ORF construct plasmid, and with a visual representation of the construct. The MS2 region contained 12 stem loops. Images were taken as Z-stacks and the average intensity projection is shown. Scale bar: 3μm.

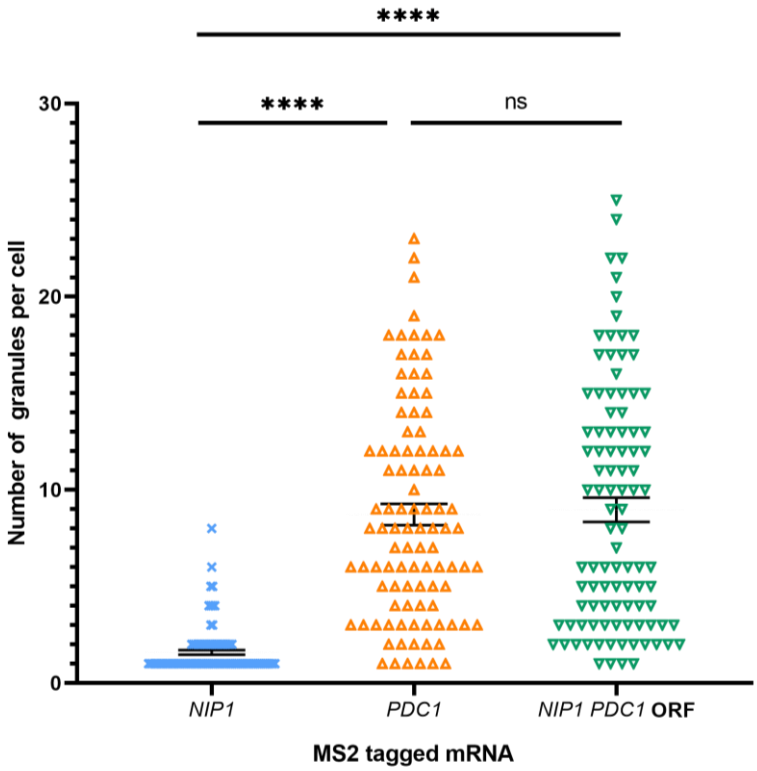


Figure 3.7: Quantification of the number of granules for the *NIP1 PDC1* ORF swap construct. For each mRNA the number of granules per cell is shown (n=100). For *NIP1* (shown in blue) the average number is 1-2 granules per cell, for *PDC1* (orange) the average number is 8-9 granules per cell and for the *NIP1 PDC1* ORF construct (green) the average number is 9-10 granules per cell. Error bars are shown as the mean with standard deviation. One-way ANOVA statistical analysis showed significant differences between *NIP1* and the *PDC1* and *NPORF* constructs (P<0.0001) and no significant differences between the *PDC1* and *NPORF* constructs.

The opposite mRNA swap was then made, containing the *NIP1* ORF in a *PDC1* background (*PDC1* promoter, 5'UTR and 3'UTR). This construct is referred to as the *PDC1 NIP1* ORF construct (Figure 3.1). This *PDC1 NIP1* ORF construct was also made by Gibson cloning and verified by restriction enzyme mapping and DNA sequencing (Figure 3.8). The *PDC1 NIP1* ORF was then transformed into yeast expressing MCP-GFP_3 (yMK2476) and imaged. Imaging showed that the *PDC1 NIP1* ORF construct produced cytosolic granules in a pattern that appeared more similar to *NIP1* than *PDC1* (Figure 3.9). Quantification of these granules (n=100) also showed that the *PDC1 NIP1* ORF construct produced 1-2 granules per cell on average (the same as *NIP1*) (Figure 3.10). Therefore, just as the *PDC1* ORF appears capable of driving a similar cytosolic granule pattern to full-length *PDC1*, the *NIP1* ORF appears to drive the same cytosolic granule pattern as full-length *NIP1*.

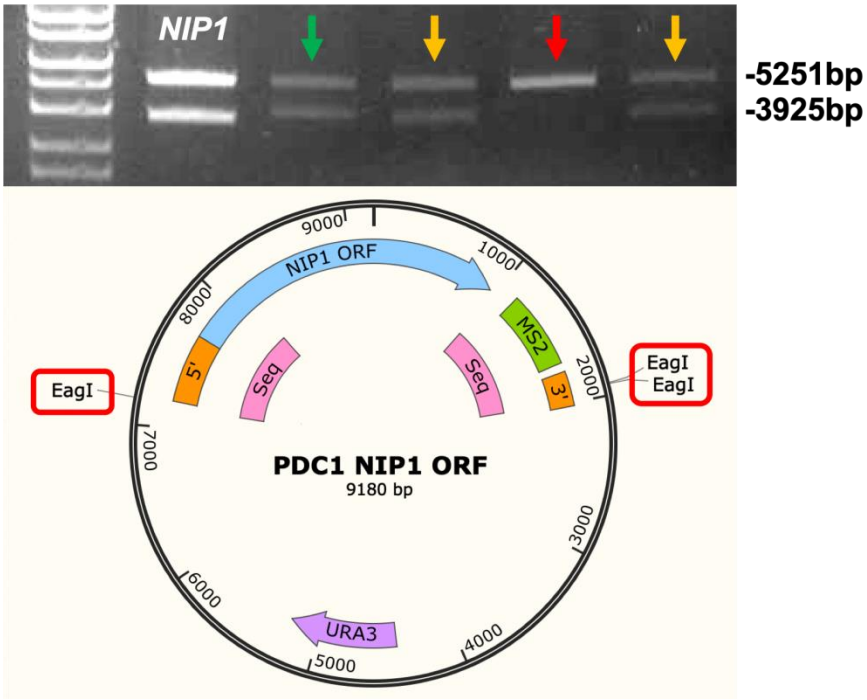


Figure 3.8: Restriction enzyme mapping of the *PDC1 NIP1* ORF swap construct. The gel shows 4 possible *PDC1 NIP1* ORF samples and a *NIP1* parent plasmid DNA control (first lane), all digested with *Eag*I restriction enzyme. All samples show a band at 5251bp corresponding to the plasmid backbone. Potentially correct samples (green and orange arrows) also produced a band around 3925bp that corresponded to the *PDC1* ORF with *NIP1* UTRs. As shown in the *NIP1* parent lane, the *NIP1* construct produced a slightly lower band at 3792bp (and the *PDC1* parent would have produced a band at 3153bp). The red arrow shows a band corresponding to re-ligated vector backbone. Possible ORF swap samples were verified by Sanger sequencing across the cloning junction regions (highlighted in pink on the plasmid representation). Sequencing showed that samples 2 and 4 (orange arrows) had the correct *PDC1 NIP1* swap regions but fewer MS2 stem loops (8-9 stem loops). Sample 1 (green arrow) had the correct *PDC1 NIP1* ORF swap and 12x MS2 stem loops.

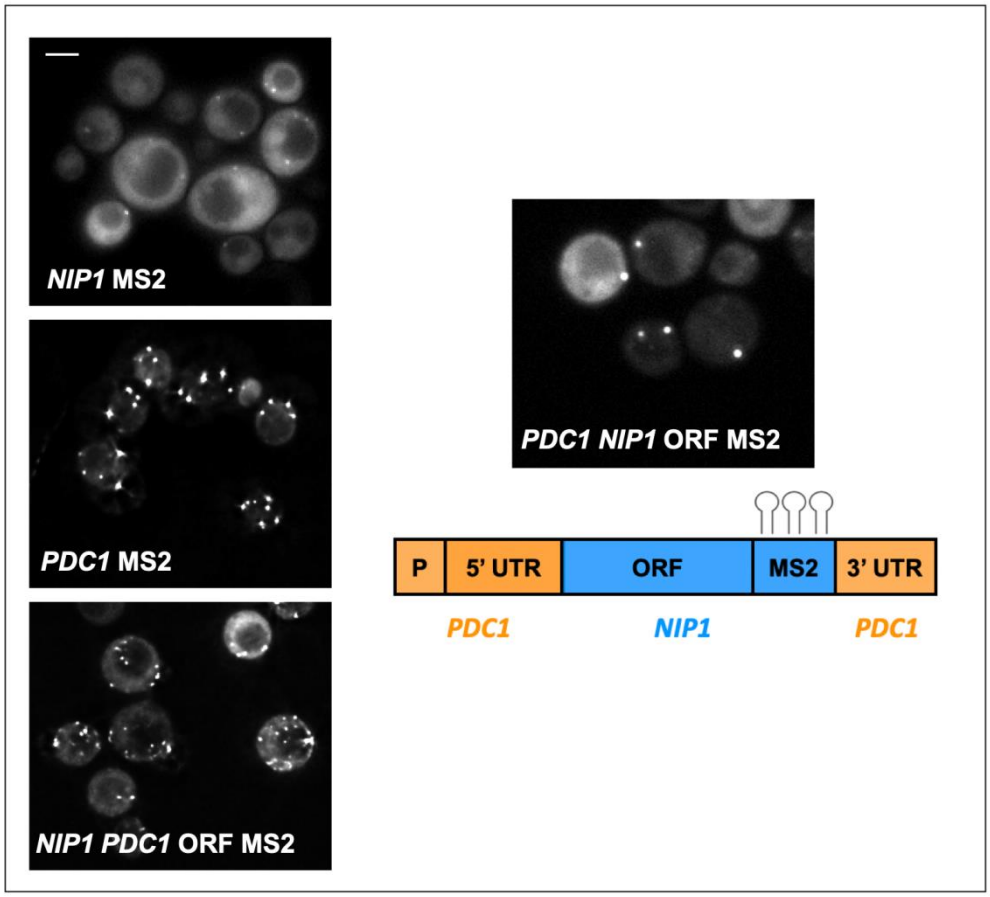


Figure 3.9: Imaging of the *PDC1 NIP1* ORF swap construct. Left-hand images show MS2-tagged *NIP1*, *PDC1*, and the *NIP1 PDC1* ORF swap constructs imaged in yeast cells genomically expressing MCP-GFP_3. The right-hand image shows the *PDC1 NIP1* ORF construct in yeast cells also expressing MCP-GFP_3, along with a visual representation of the construct. The MS2 region contained 12 stem loops. Images were taken as Z-stacks and the maximum intensity projection is shown. Scale bar: 3µm.

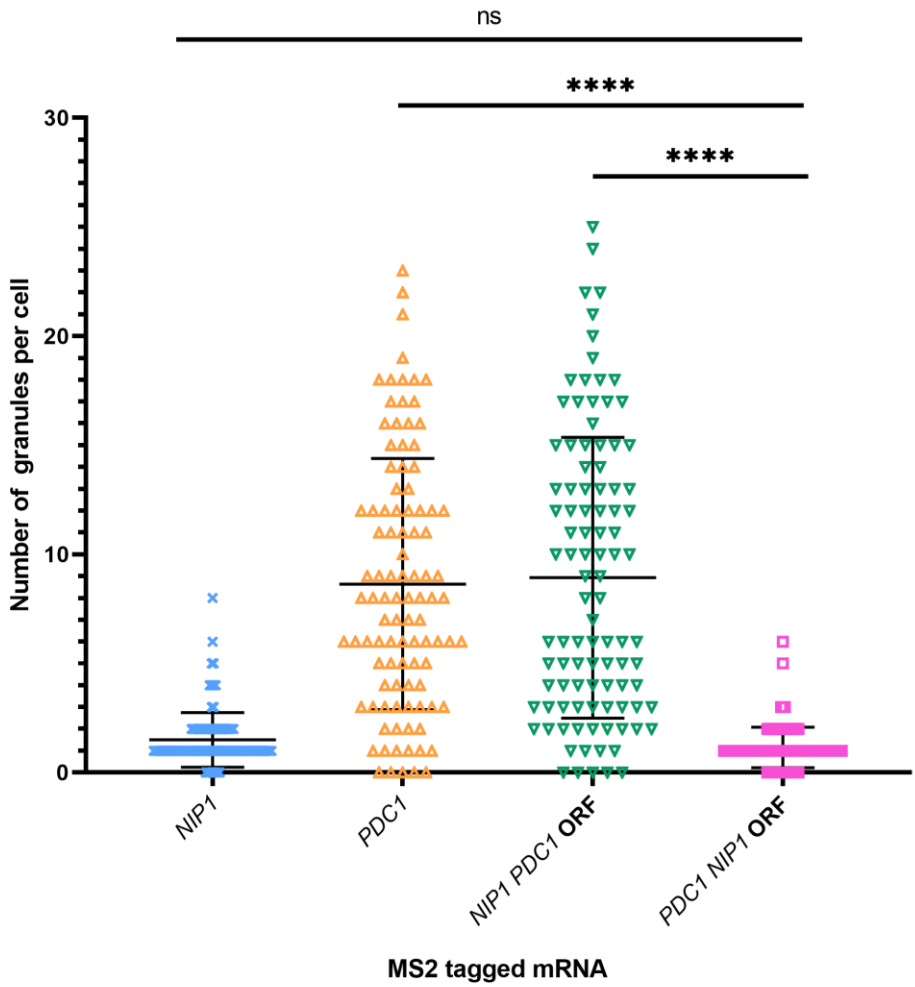


Figure 3.10: Quantification of the number of granules for the *PDC1 NIP1* ORF swap construct. For each mRNA the number of granules per cell is shown (n=100). For *NIP1* (shown in blue) the average number is 1-2 granules per cell, for *PDC1* (orange) the average number is 8-9 granules per cell, for the *NIP1 PDC1* ORF construct (green) the average number is 9-10 granules per cell and for the *PDC1 NIP1* ORF construct (pink) the average number is 1-2 granules per cell. Error bars are shown as the mean with standard deviation. One-way ANOVA statistical analysis showed significant differences between the *PNORF* and *PDC1* and *NPORF* constructs (P<0.0001) and no significant difference between *PNORF* and *NIP1*.

3.3. Construction and imaging of *NIP1 PDC1* 5'UTR ORF and *NIP1 PDC1* ORF 3'UTR constructs:

Due to difficulties in achieving long PCR fragments for vector backbone fragments, constructs with single UTR swaps (e.g. *NIP1 PDC1* 5'UTR and *NIP1 PDC1* 3'UTR) were not successfully generated (Figure 3.1). Instead, to reduce the size of the vector backbone fragments and ensure the shorter PCR fragment contained the MS2 stem loop region, two other constructs were made that swapped either the 5'UTR and ORF or the ORF and 3'UTR of *PDC1* with *NIP1* (Figure 3.1). These constructs were called *NIP1 PDC1* 5'UTR ORF and *NIP1 PDC1* ORF 3'UTR, respectively. Both constructs were made with Gibson cloning to swap the appropriate regions and verified with restriction enzyme mapping and DNA sequencing (Figure 3.11).

DNA sequencing showed that unlike the *NIP1 PDC1* ORF and *PDC1 NIP1* ORF swap constructs, which retained the 12x MS2 stem loops derived from the parent plasmid, both the *NIP1 PDC1* 5'UTR ORF and *NIP1 PDC1* ORF 3'UTR constructs contained fewer MS2 stem loops (6-8 loops) (Figure 3.12). This was possibly due to stem loop loss during PCRs, where the polymerase might skip one or more loops and the PCR then favours amplification of the smaller DNA fragment with fewer MS2 stem loops (Figure 3.12). Nevertheless, since it has been shown that as few as 6x MS2 stem loops are sufficient to visualise bulk mRNA in cells (Bertrand et al., 1998), both of these constructs were transformed into yeast cells genomically expressing MCP-GFP_3 (yMK2476) and imaged.

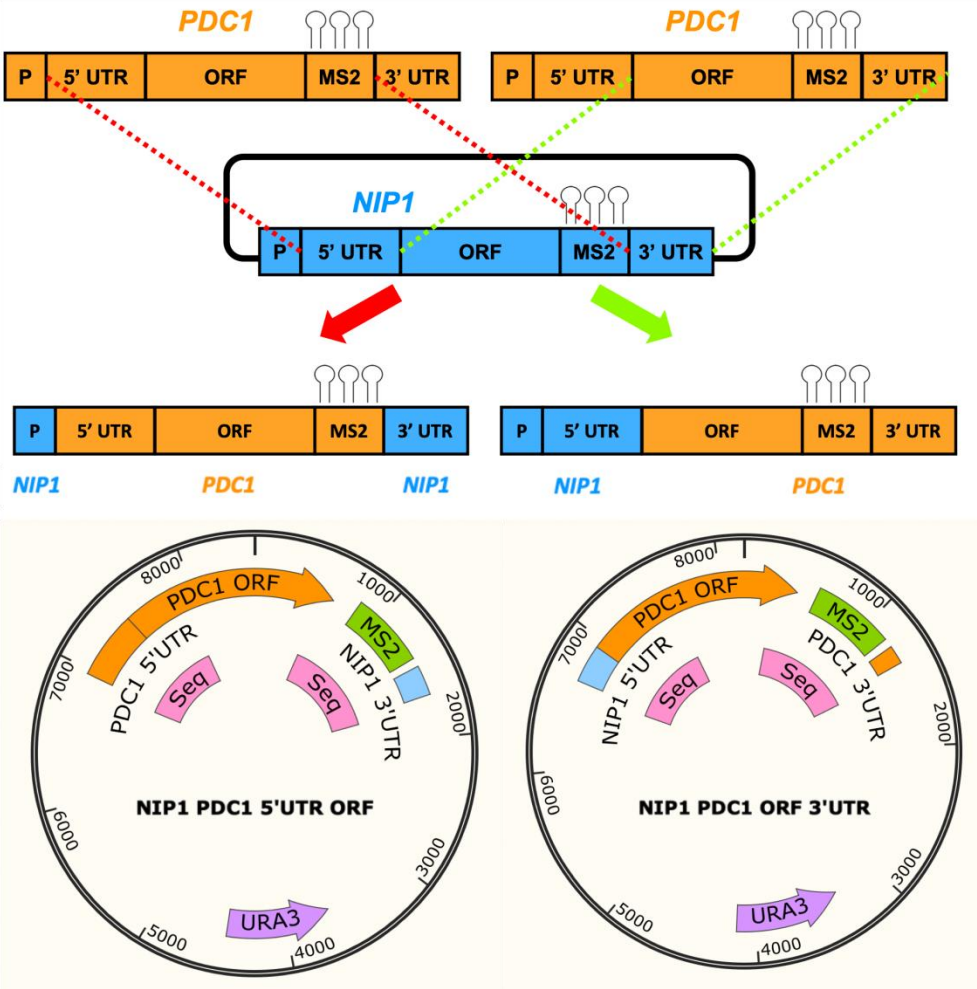


Figure 3.11: Gibson cloning of the *NIP1 PDC1* 5'UTR ORF and *NIP1 PDC1* ORF 3'UTR swap constructs. For each construct, the appropriate *PDC1* fragment was amplified and cloning was designed to replace those regions in the *NIP1* parent vector. Possible samples were verified by restriction enzyme mapping (not shown here) and then sequenced across the cloning junctions as shown highlighted in pink on the plasmid representations.

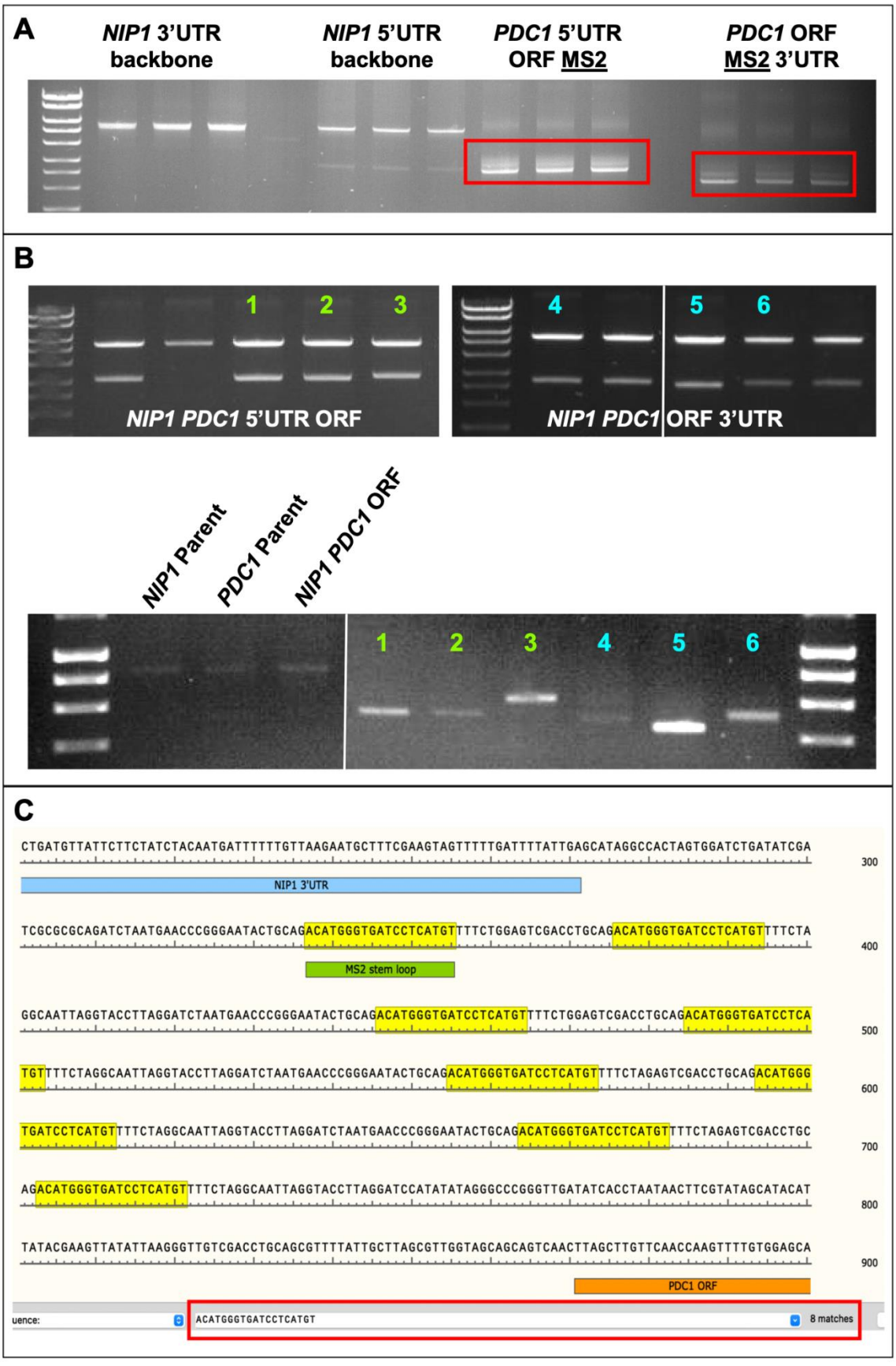

A
NIP1 3'UTR backbone
NIP1 5'UTR backbone
PDC1 5'UTR ORF MS2
PDC1 ORF MS2 3'UTR
B
1
2
3
NIP1 PDC1 5'UTR ORF
4
5
6
NIP1 PDC1 ORF 3'UTR
NIP1 Parent
PDC1 Parent
NIP1 PDC1 ORF
1
2
3
4
5
6
C
CTGATGTTATTCTTCTATCTACAATGATTTTTTGTTAAGAATGCTTTCGAAGTAGTTTTTGATTTTATTGAGCATAGGCCACTAGTGGATCTGATATCGA
300
NIP1 3'UTR
TCGCGCGCAGATCTAATGAACCCGGGAATACTGCAGACATGGGTGATCCTCATGTTTTCTGGAGTCGACCTGCAGACATGGGTGATCCTCATGTTTTCTA
400
MS2 stem loop
GGCAATTAGGTACCTTAGGATCTAATGAACCCGGGAATACTGCAGACATGGGTGATCCTCATGTTTTCTGGAGTCGACCTGCAGACATGGGTGATCCTCA
500
TGTTTTCTAGGCAATTAGGTACCTTAGGATCTAATGAACCCGGGAATACTGCAGACATGGGTGATCCTCATGTTTTCTAGAGTCGACCTGCAGACATGGG
600
TGATCCTCATGTTTTCTAGGCAATTAGGTACCTTAGGATCTAATGAACCCGGGAATACTGCAGACATGGGTGATCCTCATGTTTTCTAGAGTCGACCTGC
700
AGACATGGGTGATCCTCATGTTTTCTAGGCAATTAGGTACCTTAGGATCCATATATAGGGCCCGGGTTGATATCACCTAATAACTTCGTATAGCATACAT
800
TATACGAAGTTATATTAAGGGTTGTCGACCTGCAGCGTTTTATTGCTTAGCGTTGGTAGCAGCAGTCAACTTAGCTTGTTCAACCAAGTTTTGTGGAGCA
900
PDC1 ORF
uence:
ACATGGGTGATCCTCATGT
8 matches

Figure 3.12: The number of MS2 stem loops in swap constructs. Panel A shows the PCR fragments for each of the *NIP1* and *PDC1* fragments generated for Gibson cloning. For the PCR fragments where MS2 stem loops are present (*PDC1* fragments, red boxes) smears are visible on the gel, indicating differences in stem loop number. Panel B shows different DNA samples of possible swap constructs. The lower bands of *NIP1 PDC1* 5'UTR ORF and *NIP1 PDC1* ORF 3'UTR constructs are slightly different sizes, indicating differences in stem loop number. The lower gel shows PCR fragments specifically designed to amplify the stem loop region of all constructs, where *NIP1*, *PDC1*, and *NIP1 PDC1* ORF constructs show bands around 800bp, corresponding to 12x MS2 stem loops. *NIP1 PDC1* 5'UTR ORF and *NIP1 PDC1* ORF 3'UTR constructs show shorter fragments suggesting fewer numbers of MS2 stem loops are present. Panel C displays an example of DNA sequencing across the MS2 stem loop region (this example is *NIP1 PDC1* ORF 3'UTR DNA sample 6), where sequencing shows that 8x MS2 stem loops were present. Sequencing was performed in the same way for each construct DNA sample.

Images showed that both the *NIP1 PDC1* 5'UTR ORF and *NIP1 ORF* 3'UTR constructs displayed a cytosolic granule pattern that was different from both *NIP1* and *PDC1* (Figure 3.13), where some cells displayed a single granule per cell and some cells showed multiple granules per cell. Quantification of these cells (n=100) also showed that the granule pattern was dissimilar to either *NIP1* or *PDC1* (Figure 3.14). Although there were many cells that possessed only one or two granules per cell (like *NIP1*), both the *NIP1 PDC1* 5'UTR ORF and *NIP1 ORF* 3'UTR constructs also resulted in many cells expressing multiple granules per cell, at a higher frequency than is seen for *NIP1* only expressing cells (Figure 3.14). However these cells were also dissimilar from *PDC1* expressing cells, as the number of granules in these cells tended to be fewer; around 4-5 granules per cell whereas *PDC1* is usually present at an average of ~10 granules per cell (Figure 3.14).

Since both the *NIP1 PDC1* 5'UTR ORF and *NIP1 PDC1* ORF 3'UTR constructs possess the ORF of *PDC1*, which appears to be capable of driving the pattern of cytosolic granules to look more like the pattern seen for full-length *PDC1* (Figure 3.7), this result was unexpected. Since this version of the MS2 system does not achieve single molecule resolution, it requires multiple tagged mRNAs to localise together to allow detection (Bertrand et al., 1998; Pizzinga et al., 2019; Morales-Polanco et al., 2021). Therefore, the presence of fewer MS2 stem loops in the *NIP1 PDC1* 5'UTR ORF and *NIP1* ORF 3'UTR constructs may mean that granules containing fewer numbers of these mRNAs cannot be detected, as they can be for both *PDC1* and the *NIP1 PDC1* ORF construct containing 12x MS2 stem loops. It is also possible that the differences in granule patterns seen between cells expressing the same construct are due to differences in plasmid expression. In an attempt to address this, *NIP1*, *PDC1* and the *NIP1 PDC1* ORF swap construct were expressed on a lower expression plasmid (see below). Due to time constraints, further attempts to produce *NIP1 PDC1* 5'UTR ORF and *NIP1 ORF* 3'UTR constructs with 12x MS2 stem loops were not possible.

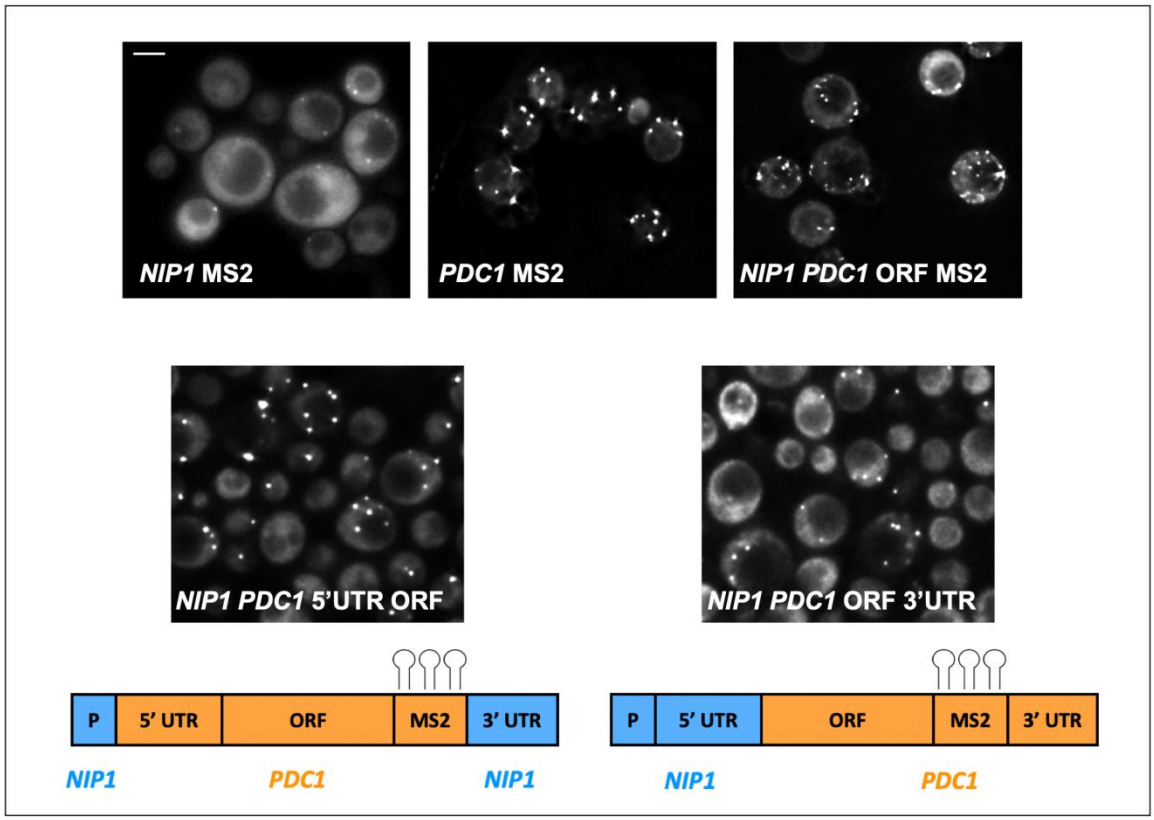


Figure 3.13: Imaging of the *NIP1 PDC1* 5'UTR ORF and *NIP1 PDC1* ORF 3'UTR swap constructs. Images show MS2 tagged parental and swap constructs in yeast cells also expressing MCP-GFP_3. *NIP1*, *PDC1* and *NIP1 PDC1* ORF images are the same as above. The *NIP1 PDC1* 5'UTR ORF and *NIP1 PDC1* ORF 3'UTR constructs are also shown along with their sequence representations. All images were taken as Z-stacks and the average intensity projection is shown. Scale bar: 3μm.

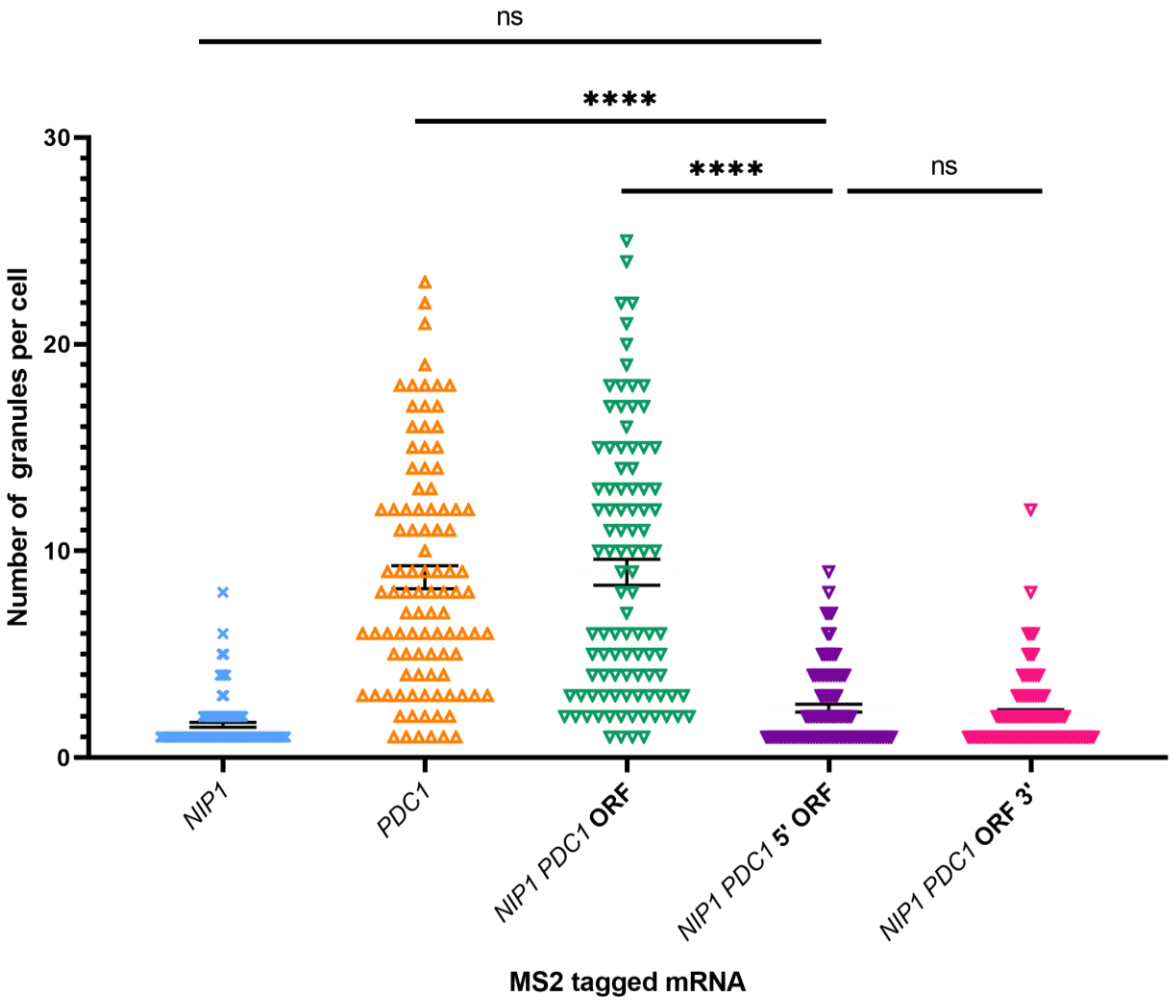


Figure 3.14: Quantification of the number of granules for the *NIP1 PDC1* 5'UTR ORF and *NIP1 PDC1* ORF 3'UTR swap constructs. For each mRNA the number of granules per cell is shown (n=100). For *NIP1* (shown in blue) the average number is 1-2 granules per cell, for *PDC1* (orange) the average number is 8-9 granules per cell, for the *NIP1 PDC1* ORF construct (green) the average number is 9-10 granules per cell, and for both the *NIP1 PDC1* 5'UTR ORF (purple) and *NIP1 PDC1* ORF 3'UTR (pink) constructs the average number is 2-3 granules per cell. Error bars are shown as the mean with standard deviation. One-way ANOVA statistical analysis showed no significant differences between *NP5'ORF* and *NPORF3'* constructs. When compared with the other constructs shown on the graph the significance was the same for both *NP5'ORF* and *NPORF3'*, and so is only shown here for *NP5'ORF* for ease of reading. *NP5'ORF* and *NPORF3'* were significantly different from *PDC1* and *NPORF* (P<0.0001) and not significantly different from *NIP1*.

3.4. Construction and imaging of CEN expression constructs:

Thus far in this chapter, *NIP1*, *PDC1* and swap constructs have been expressed on plasmids containing a yeast 2μ origin of replication (ORI). The 2μ ORI plasmids are relatively high copy number plasmids (copy number can be as high as 40-60 copies per cell (Chan et al., 2013)), and therefore may cause differential expression levels from plasmids and mixed expression profiles across cell populations. *NIP1*, *PDC1* and *NIP1 PDC1* ORF swap constructs were cloned into a yeast vector containing a low copy number ORI for two reasons; to investigate whether or not the pattern of cytosolic granules was altered for these constructs when they were expressed at lower levels, and to stabilise the expression of these constructs between cells. To do this, a bacterial plasmid (pRS316) containing an ORI derived from a centrosome segment (CEN), shown to produce low copy numbers of plasmids (Clarke and Carbon, 1985), was used to replace the 2μ ORI with the CEN ORI for *NIP1*, *PDC1* and *NIP1 PDC1* ORF swap constructs.

Restriction cloning was designed to replace the plasmid backbone (including the 2μ ORI) of *NIP1*, *PDC1* and *NIP1 PDC1* ORF swap constructs with the backbone (including the CEN ORI) of pRS316 (Figure 3.15). Constructs were verified with restriction enzyme mapping and DNA sequencing (Figure 3.15). Since cloning fragments were generated by restriction enzyme digest and not PCR, it was expected that the 12x MS2 stem loops from the parent vectors would be maintained, and this was also verified by DNA sequencing (Figure 3.15). CEN driven constructs were then imaged by transforming each construct into yeast cells genomically expressing MCP-GFP_3 (yMK2476).

Imaging showed that for *NIP1*, *PDC1* and the *NIP1 PDC1* ORF construct, the pattern of cytosolic RNA granules was the same for the CEN plasmids as the 2μ plasmids (Figure 3.16). Quantification of these cells (n=100) showed the same result; *NIP1* expressed from a 2μ or CEN plasmid showed on average one or two granules per cell, while both the *PDC1* and *NIP1 PDC1* ORF constructs expressed from the 2μ or CEN plasmids appeared in ~10 granules per cell (Figure 3.17). One difference between the 2μ and CEN plasmids was that the RNA granules for constructs expressed from the CEN plasmid were noticeably dimmer than from the 2μ plasmid, and longer exposure times were needed to visualise them (Figure 3.16). This was likely due to fewer numbers of mRNAs being present and therefore a lower fluorophore concentration in one place.

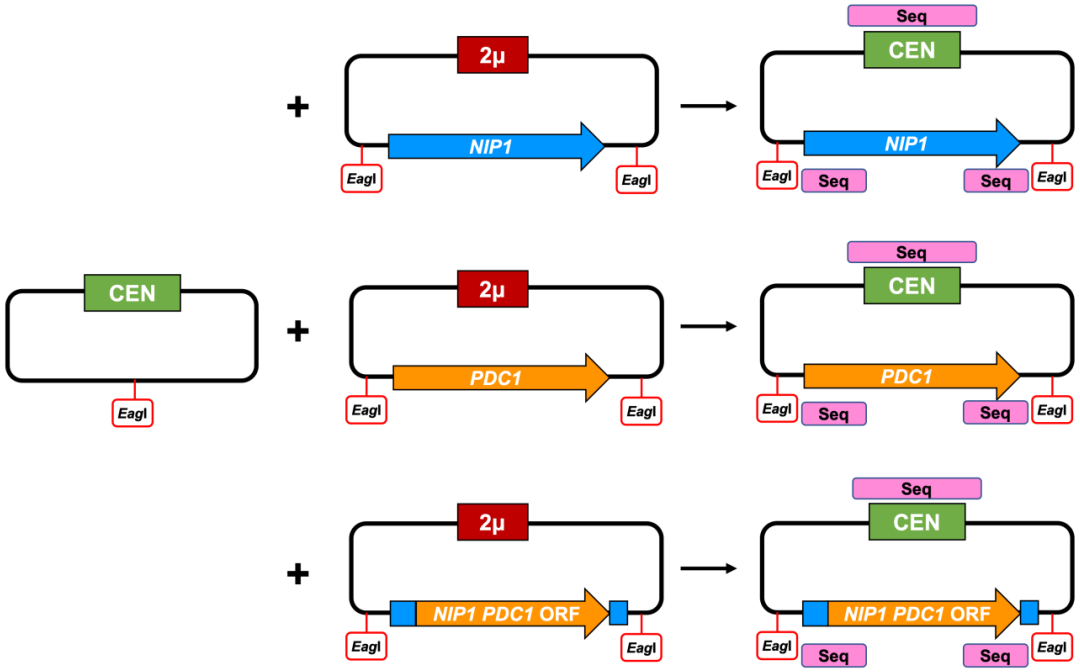


Figure 3.15: Construction of CEN plasmids for *NIP1*, *PDC1* and *NIP1 PDC1* ORF constructs. The pRS316 vector was linearised by digestion with *Eag*I and ligated with *NIP1*, *PDC1* and *NIP1 PDC1* ORF constructs (also digested with *Eag*I) to produce the CEN driven plasmids shown here. Each construct was verified with restriction enzyme mapping (not shown here) and DNA sequencing across the cloning junctions and the CEN sequence, shown highlighted in pink on the plasmid representations. Sequencing showed 12x MS2 stem loops for each construct.

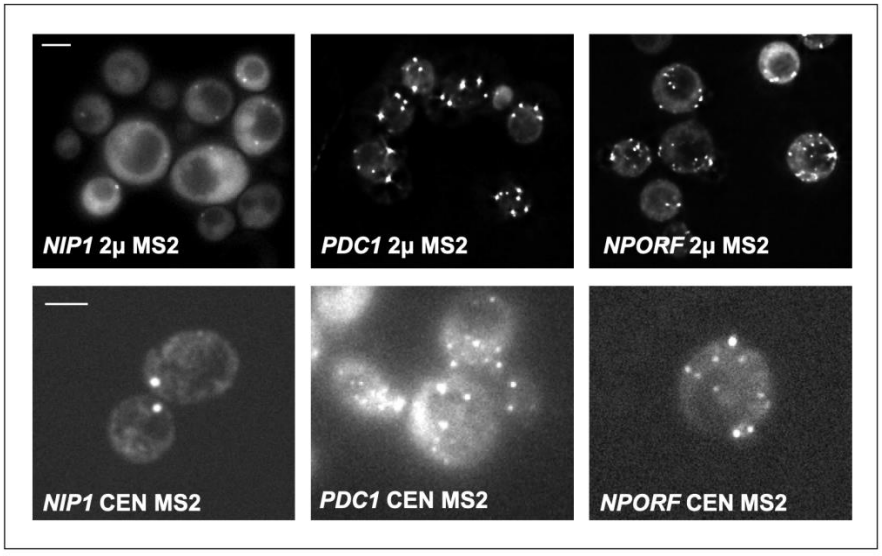


Figure 3.16: Imaging of *NIP1, PDC1* and *NIP1 PDC1* ORF CEN derived constructs. Images show yeast cells expressing MCP-GFP_3 and MS2 *NIP1*, *PDC1* and *NIP1 PDC1* ORF constructs. The top row of images show *NIP1*, *PDC1* and the *NIP1 PDC1* ORF constructs expressed from a 2μ plasmid. The bottom row of images show the same constructs expressed from a CEN plasmid. 2μ plasmid images were taken at exposure times of 200ms, while CEN plasmid images were taken at exposure times of 800ms. All images were taken as Z-stacks and the average intensity projection is shown. Scale bars: 3μm.

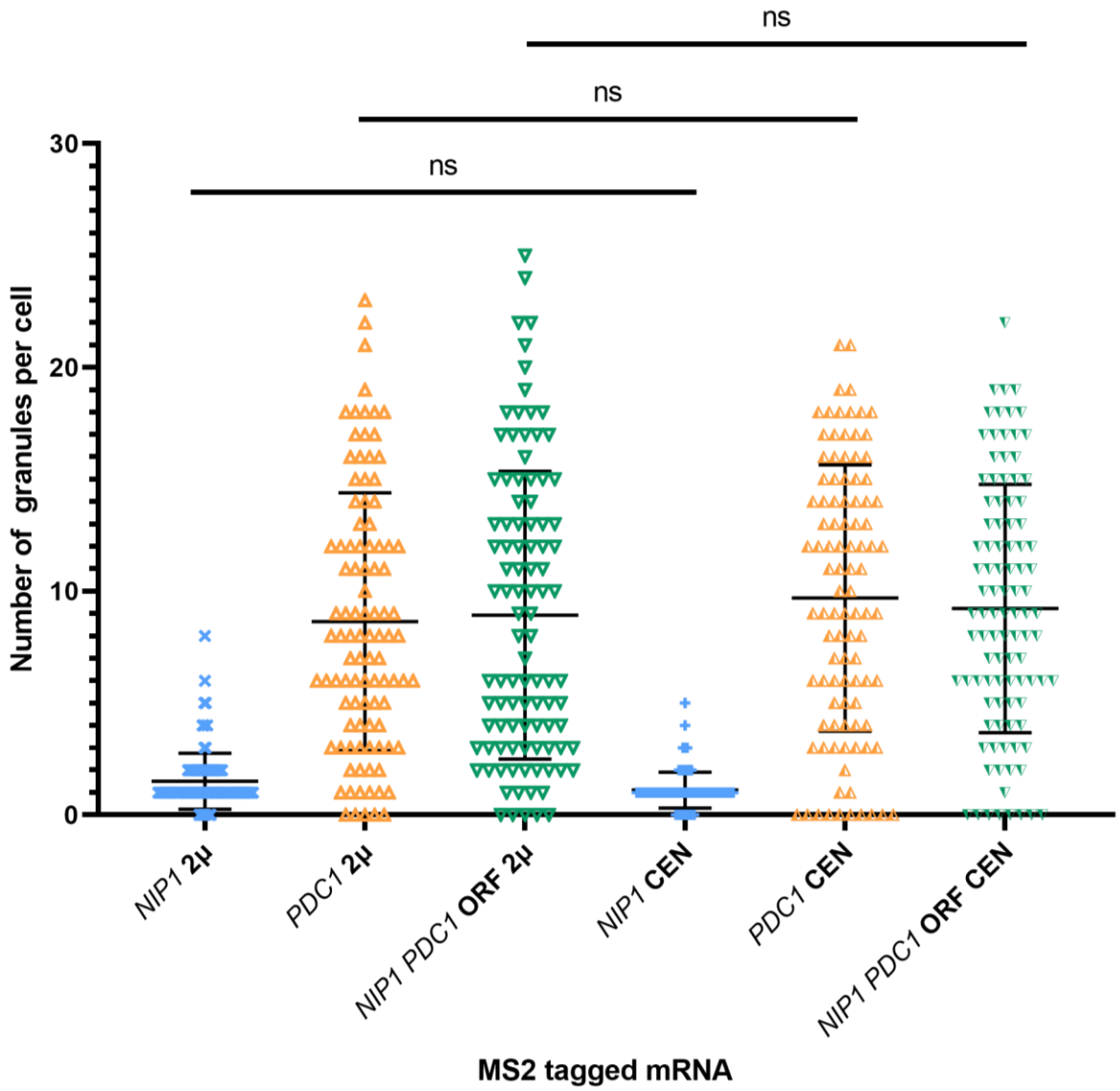


Figure 3.17: Quantification of the number of granules for *NIP1*, *PDC1* and *NIP1 PDC1* ORF CEN plasmids. For each mRNA the number of granules per cell is shown (n=100). For *NIP1* expressed from both the 2µ and CEN plasmids (shown in blue) the average number is 1-2 granules per cell, for *PDC1* expressed from the 2µ and CEN plasmids (orange) the average number is ~10 granules per cell, and for the *NIP1 PDC1* ORF construct expressed from the 2µ and CEN plasmids (green) the average number is also ~10 granules per cell. Error bars are shown as the mean with standard deviation. One-way ANOVA statistical analysis showed no significant differences between the 2µ and CEN derived constructs for *NIP1*, *PDC1*, or the *NPORF* construct.

qRT-PCR was then performed on the *NIP1*, *PDC1* and *NIP1 PDC1* ORF constructs expressed from both the 2µ and CEN plasmids to investigate differences in mRNA levels. First, it was tested whether or not mRNA abundance was altered between expression from 2µ or CEN plasmids. qRT-PCR showed that each construct was more abundant when expressed on the 2µ plasmid compared to the CEN plasmid (Figure 3.18). *NIP1* was ~5 times more abundant when expressed from the 2µ plasmid, *PDC1* was ~30 times more abundant from the 2µ plasmid, and the *NIP1 PDC1* ORF was ~40 times more abundant from the 2µ plasmid compared to the CEN plasmid (Figure 3.18). Since each CEN plasmid is most often present in a single copy per cell (Clarke and Carbon, 1985), this data suggested that both *PDC1* and *NIP1 PDC1* ORF constructs are also more highly expressed than *NIP1*, which was further studied using qRT-PCR.

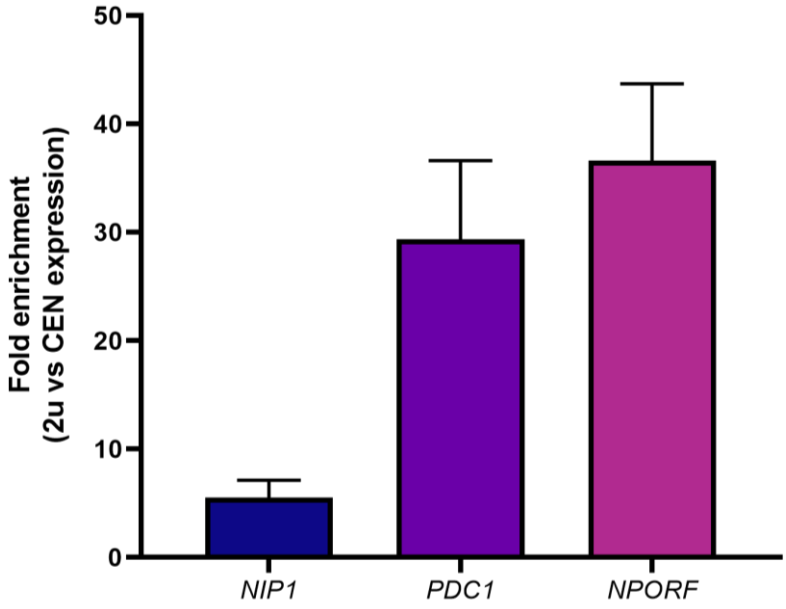


Figure 3.18: qRT-PCR of *NIP1*, *PDC1* and *NPORF* relative abundance in a 2µ plasmid compared to a CEN plasmid. *NPORF* denotes the *NIP1 PDC1* ORF construct, and the abundance of each mRNA expressed from a 2µ plasmid is compared to expression from a CEN plasmid. *NIP1* 2µ is ~5 times more abundant than *NIP1* CEN (blue), *PDC1* 2µ is ~30 times more abundant than *PDC1* CEN (purple), and *NPORF* 2µ is ~40 times more abundant than *NPORF* CEN (pink). For each sample *ACT1* was used as a reference gene, and n=3. Error bars show the mean with standard deviation.

qRT-PCR was therefore then performed to investigate the relative mRNA abundance of *NIP1*, *PDC1* and *NIP1 PDC1* ORF constructs compared to one another, in both the 2μ and CEN plasmids. When expressed from 2μ plasmids, *PDC1* is ~100 times more abundant *NIP1*, and the *NIP1 PDC1* ORF construct is around ~20 times more abundant than *PDC1*; accordingly, the *NIP1 PDC1* ORF construct is ~2000 times more abundant than *NIP1* (Figure 3.19). Other studies investigating mRNA abundance in yeast have estimated that endogenous *NIP1* is present at around 10 copies per cell and *PDC1* at around 160 copies per cell (Lawless et al., 2016; Lahtvee et al., 2017; Pizzinga et al., 2019), indicating that endogenous *PDC1* mRNA is ~16 times more abundant than endogenous *NIP1* mRNA. Therefore, expression of these constructs on a 2μ plasmid appears to drastically change mRNA levels (Figure 3.19), although the cytosolic RNA granule pattern does not appear to be affected for *NIP1* and *PDC1* mRNAs (Figures 3.2 and 3.3).

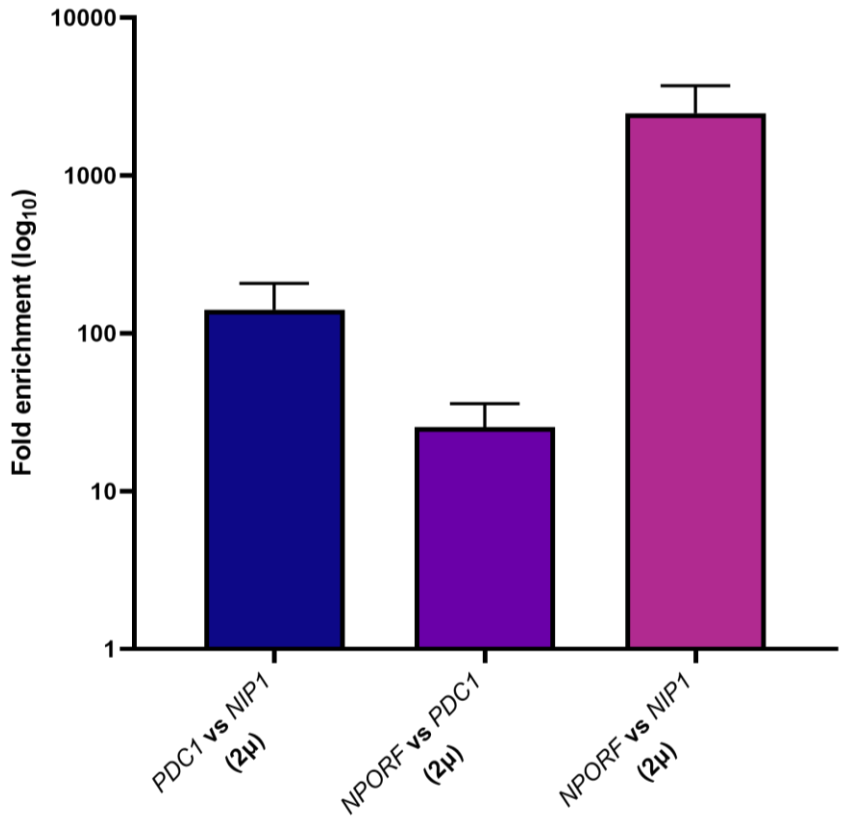


Figure 3.19: qRT-PCR of *NIP1*, *PDC1* and *NPORF* relative abundance expressed from a 2μ plasmid. *NPORF* denotes the *NIP1 PDC1* ORF construct, and the abundance of each mRNA is tested relative to another. *PDC1* is ~100 times more abundant than *NIP1* (blue), *NPORF* is ~20 times more abundant than *PDC1* (purple), and *NPORF* is ~2000 times more abundant than *NIP1* (pink). For each sample *ACT1* was used as a reference gene, and n=3. Error bars show the mean with standard deviation.

qRT-PCR performed on the *NIP1*, *PDC1* and *NIP1 PDC1* ORF constructs expressed from the CEN plasmid showed a similar result to the 2µ plasmid, where again *PDC1* was more abundant than *NIP1* and the *NIP1 PDC1* ORF construct was more abundant than *PDC1* (Figure 3.20). However, the differences in abundance levels between the mRNAs were lower; *PDC1* was ~20 times more abundant than *NIP1*, the *NIP1 PDC1* ORF construct was ~5 times more abundant than *PDC1* and ~100 times more abundant than *NIP1* (Figure 3.20). Therefore, when expressed on the single-copy CEN plasmid the enrichment of *PDC1* over *NIP1* is around the same level of enrichment as the untagged, endogenously expressed mRNAs (Lawless et al., 2016; Lahtvee et al., 2017; Pizzinga et al., 2019). In all cases it was unexpected that the *NIP1 PDC1* ORF construct showed higher levels of mRNA expression compared to *NIP1*, as the native promoter of *NIP1* is maintained in the swap construct. Therefore, this data indicates the possibility that the *NIP1 PDC1* ORF construct localisation pattern (similar to *PDC1*) is associated with the relatively high abundance of the *NIP1 PDC1* ORF construct (also similar to *PDC1*).

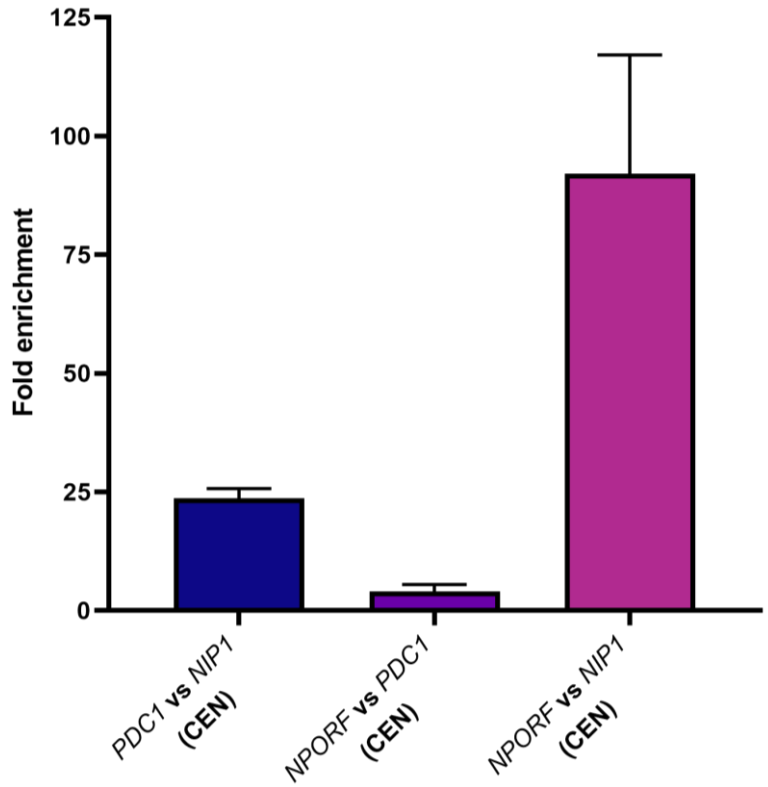


Figure 3.20: qRT-PCR of *NIP1*, *PDC1* and *NPORF* relative abundance expressed from a CEN plasmid. *NPORF* denotes the *NIP1 PDC1* ORF construct, and the abundance of each mRNA is tested relative to another. *PDC1* is ~20 times more abundant than *NIP1* (blue), *NPORF* is ~5 times more abundant than *PDC1* (purple), and *NPORF* is ~100 times more abundant than *NIP1* (pink). For each sample *ACT1* was used as a reference gene, and n=3. Error bars show the mean with standard deviation.

3.5. MEME analysis of localised mRNA sequences:

To investigate whether or not the sequences of both glycolytic and translation factor mRNAs contained conserved motifs that might drive mRNA localisation to granules, MEME motif analysis was carried out on their mRNA sequences. MEME Suite is a web-based server than can be used to align and analyse multiple sequences to search for conserved motifs (Bailey et al., 2006; Bailey et al., 2009). This analysis was performed on both sequences of glycolytic mRNAs and sequences of translation factor mRNAs to search for conserved elements that might be responsible for driving mRNA localisation.

Each analysis was first performed with control mRNA sequences present; sequences from mRNAs that do not share the same cytosolic RNA granules pattern. For the glycolytic mRNAs the control sequences were *NIP1* (a translation factor mRNA) and *PYK2* (a non-localised glycolytic mRNA), and for the translation factor mRNAs the control sequences were *PDC1* (a glycolytic mRNA) and *YEF3* (a translation factor mRNA found in many granules per cell) (Pizzinga et al., 2019). The MEME analysis is capable of searching for motif occurrences of zero or one (zoops), and so would be able to, for example, identify a motif present in glycolytic mRNAs that was not present in *NIP1*. However, since the motif analysis is more specific when fewer sequences are present (Bailey et al., 2016; Bailey et al., 2019), analysis was also then carried out with these control mRNAs removed.

Analysis on glycolytic mRNA sequences identified conserved motifs present in most glycolytic mRNAs seen to localise to cytosolic RNA granules (Figure 3.21 and 3.22). These sequences were most often found in the ORFs of the glycolytic mRNAs and were not present in the *NIP1* control mRNA (Figure 3.21). One of these motifs was found in the ORF of *PYK2*, while another was identified on the antisense strand of *PYK2* (and the sense strand of all other glycolytic mRNAs) (Figure 3.21). Although motifs present in the antisense strand of mRNAs could still be involved in the regulation of mRNA localisation, it is striking that this motif is only found on the antisense strand of the non-localised glycolytic mRNA *PYK2*. Analysis with control mRNAs included also identified A-rich motifs outside the coding region in all mRNA sequences; this was unsurprising as UTRs in yeast tend to be AU rich (Gilbert et al., 2007; Gingold and Pilpel, 2011; Sparks and Dieckerman, 1998) (Figure 3.21). When the control mRNAs were removed from the analysis this A-rich region was not identified in the top three conserved motifs, suggesting other motifs were more conserved between glycolytic mRNA sequences (Figure 3.22). Therefore, although it is as yet unclear whether or not these motifs affect mRNA localisation, glycolytic mRNAs appear to contain conserved motifs on their ORFs.

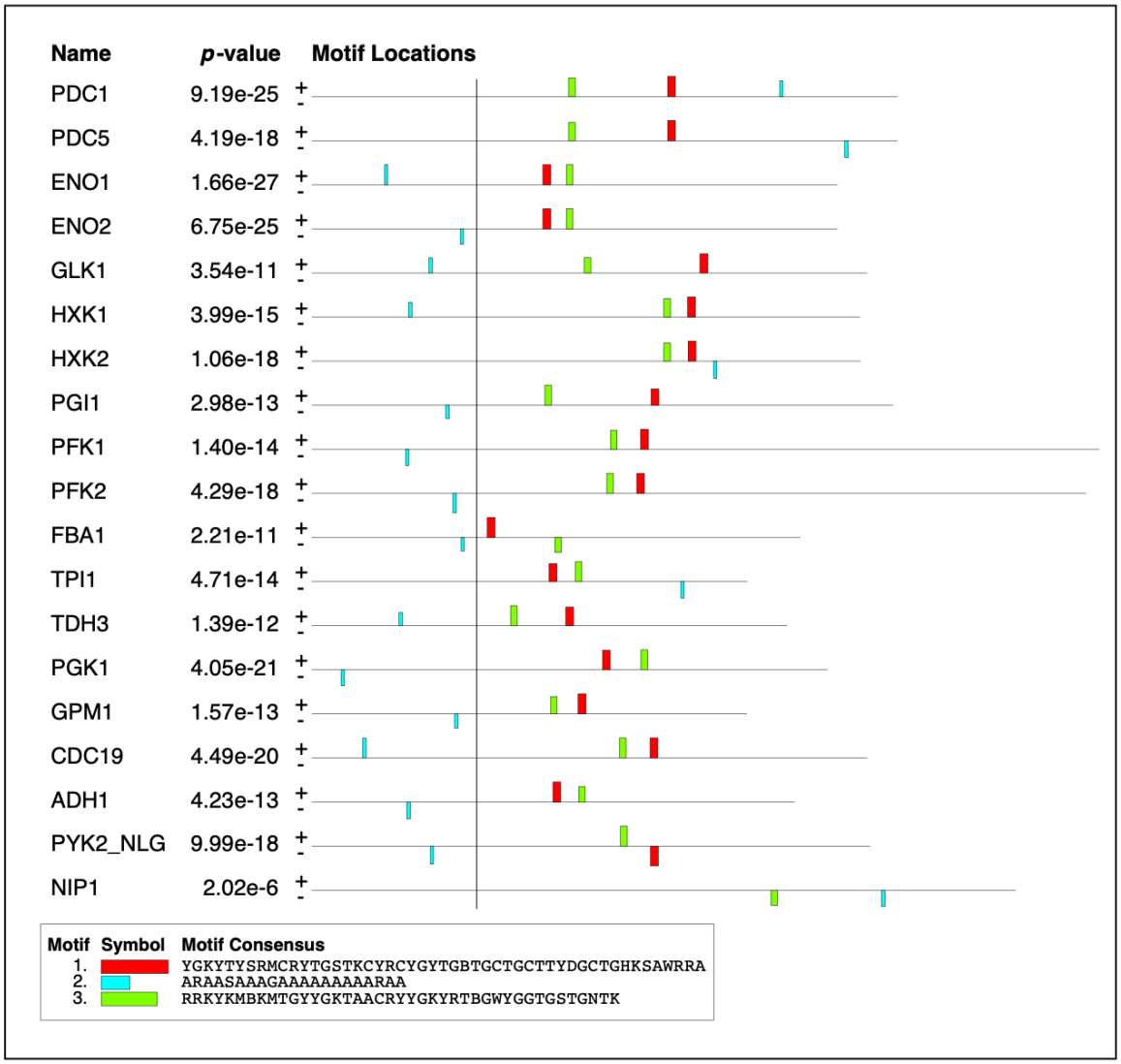


Figure 3.21: MEME analysis of glycolytic mRNA sequences with control sequences. MEME analysis was generated using MEME Suite (https://meme-suite.org/meme/tools/meme). For each mRNA, sequences were taken +/-1kb of the coding region, to ensure that UTRs and promoter/terminator regions were present. mRNAs are listed as their gene name and the *p*-value is an estimate of the probability that motifs are specific for each sequence (compared to a random sequence of the same length). The black vertical line represents the ATG start codon of each mRNA. In this analysis, the A-rich motif 2 (blue) is found in the UTRs of mRNAs, whereas motifs 1 and 3 (red and green) are found in the ORFs of mRNAs. Motif 1 is not found in *NIP1* mRNA and is present on the antisense strand of *PYK2*.

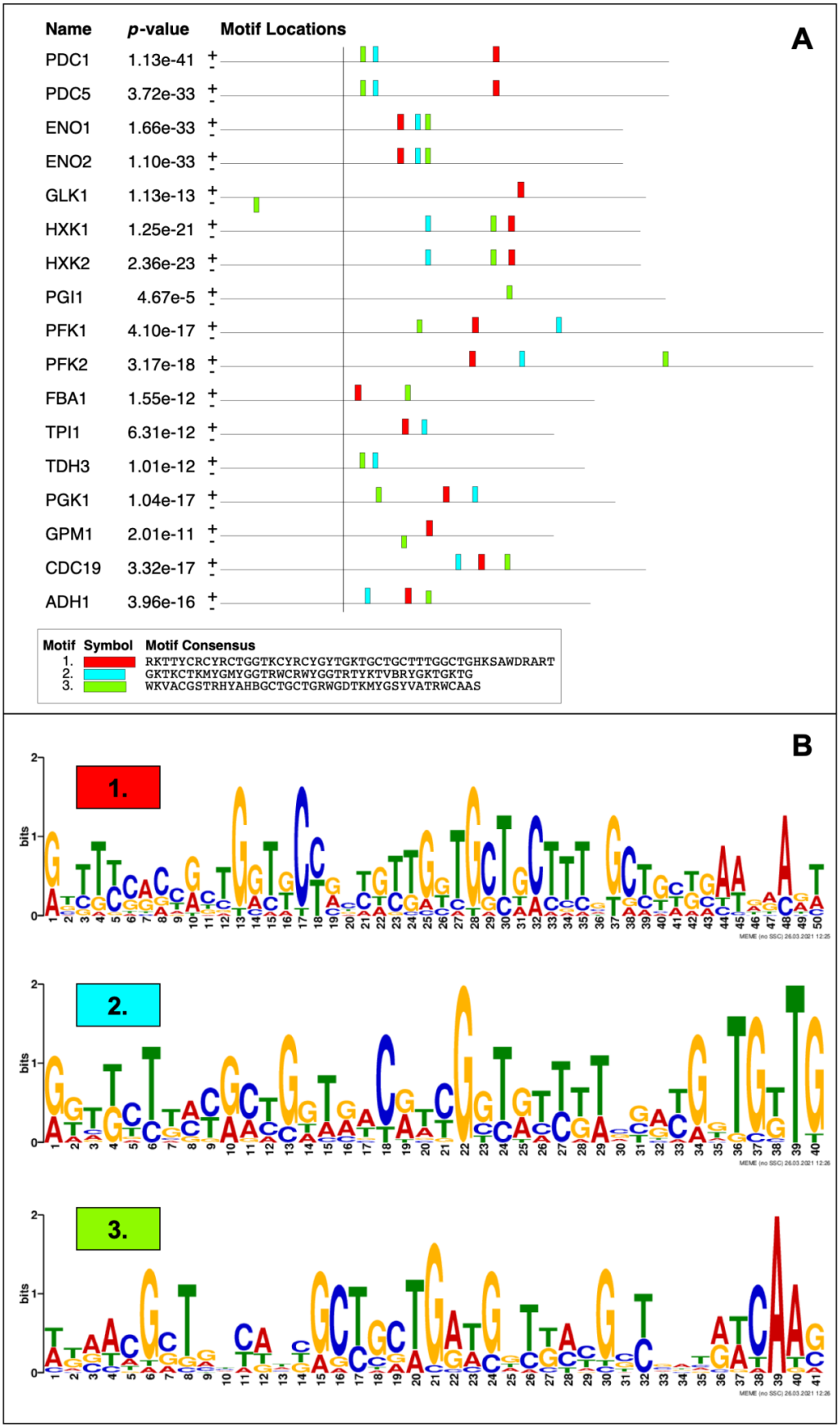

A
Name
p-value
Motif Locations
PDC1 1.13e-41
PDC5 3.72e-33
ENO1 1.66e-33
ENO2 1.10e-33
GLK1 1.13e-13
HXK1 1.25e-21
HXK2 2.36e-23
PGI1 4.67e-5
PFK1 4.10e-17
PFK2 3.17e-18
FBA1 1.55e-12
TPI1 6.31e-12
TDH3 1.01e-12
PGK1 1.04e-17
GPM1 2.01e-11
CDC19 3.32e-17
ADH1 3.96e-16
Motif Symbol Motif Consensus
1. RKTTYCRCYRCTGGTKCYRCYGYTGKTGCTGCTTTGGCTGHKSAWDRART
2. GKTKCTKMYGMYGGTRWCRWYGGTRTYKTVBRYGKTGKTG
3. WKVACGSTRHYAHBGCTGCTGRWGDTKMYGSYVATRWCAAS
B
1.
2.
3.
bits

Figure 3.22: MEME analysis of glycolytic mRNA sequences without control sequences. MEME analysis was generated using MEME Suite (https://meme-suite.org/meme/tools/meme). For each mRNA, sequences were taken +/-1kb of the coding region, to ensure that UTRs and promoter/terminator regions were present. mRNAs are listed as their gene name and the *p*-value is an estimate of the probability that motifs are specific for each sequence (compared to a random sequence of the same length). The black vertical line represents the ATG start codon of each mRNA. Panel A shows the motif locations for each mRNA, while Panel B shows the detailed sequences for each of the three motifs. With the exception of motif 3 in *GLK1*, the motifs appear on the ORFs of mRNAs.

Conversely, MEME analysis of translation factor mRNAs did not identify any particularly conserved motifs in either the controlled analysis or when control mRNAs were removed (Figures 3.23). A/T rich region motifs were identified for most mRNAs in both analyses, most often in the UTRs as expected. The analysis including control mRNAs did not identify motifs that were specific to translation factor mRNAs, or that were found in the same mRNA regions (Figure 3.23A). When the control mRNAs were removed other motifs were identified, but again these were not particularly conserved between mRNAs or found in the same mRNA regions (Figure 3.23B). Therefore, this analysis did not find consensus motifs present in translation factor mRNAs, raising the possibility that the mechanism of mRNA localisation does not rely on a conserved, RNA-binding process. Although it should be noted that many RNA binding proteins recognise structure and sequence rather than primary sequence and so a structural RNA motif is still a possibility.

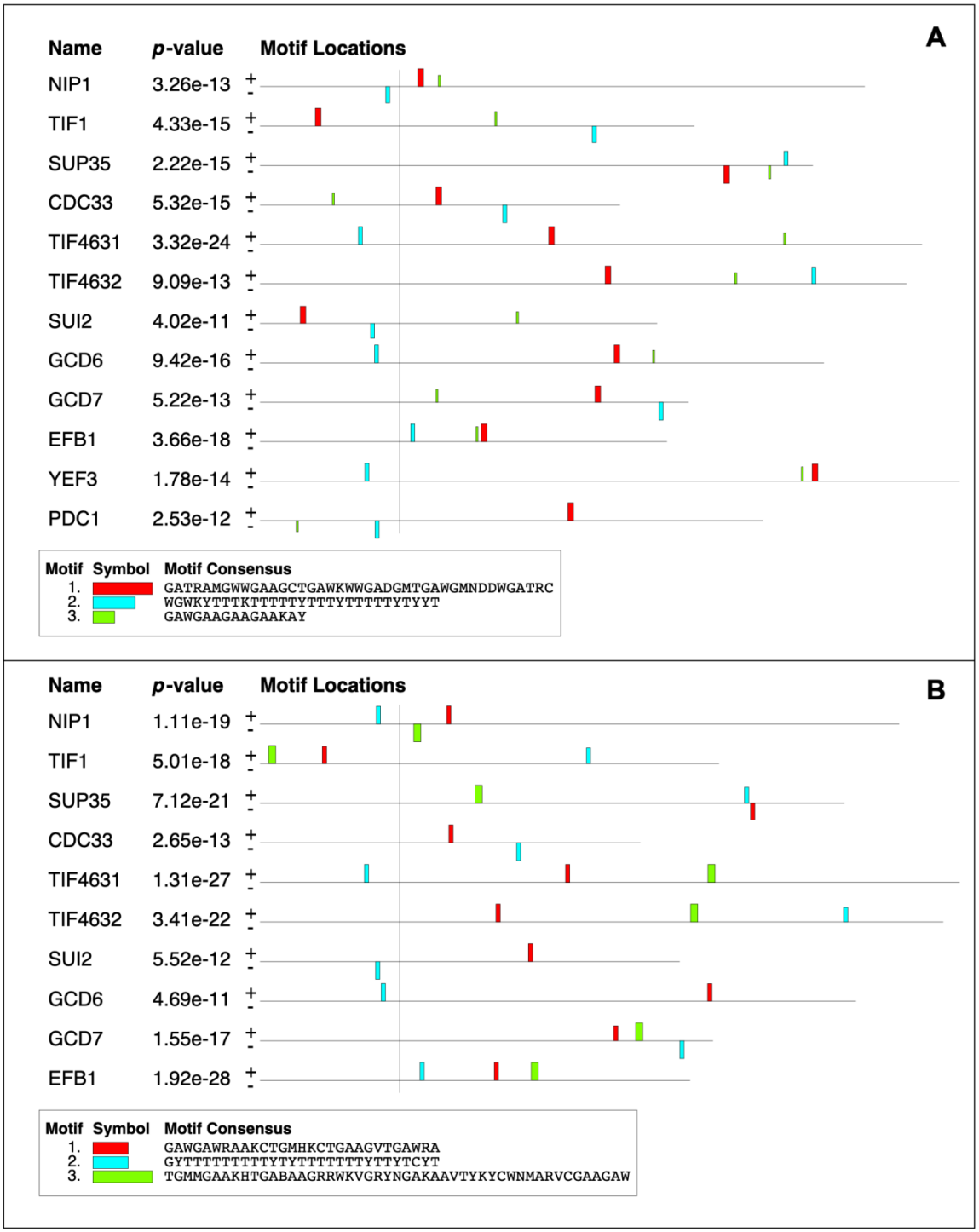


Figure 3.23: MEME analysis of translation factor mRNA sequences with and without control sequences. MEME analysis was generated using MEME Suite (https://meme-suite.org/meme/tools/meme). For each mRNA, sequences were taken +/-1kb of the coding region, to ensure that UTRs and promoter/terminator regions were present. mRNAs are listed as their gene name and the *p*-value is an estimate of the probability that motifs are specific for each sequence (compared to a random sequence of the same length). The black vertical line represents the ATG start codon of each mRNA. Panel A shows translation factor mRNA sequences with control mRNAs *YEF3* and *PDC1*, while Panel B shows the same translation factor sequences without controls. In each case, none of the three motifs are particularly conserved between these mRNA sequences.

3.6. Conclusions.

In this chapter, experiments were undertaken to investigate the differential localisation of *NIP1* and *PDC1* mRNAs. *NIP1* is a translation factor mRNA that is seen to localise to one or two cytosolic RNA granules per cell, while *PDC1* is a glycolytic mRNA seen to localise to ~10 cytosolic RNA granules per cell (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). Many glycolytic mRNAs co-localise together in these granules (such as *PDC1*, *ENO2* and *FBA1*) and translation factor mRNAs also show some level of co-localisation to their granules (such as *NIP1*, *TIF1* and *SUP35*) (Pizzinga et al., 2019; Morales-Polanco et al., 2021). Importantly, glycolytic granules and translation factor granules do not appear to co-localise together (Pizzinga et al., 2019; Morales-Polanco et al., 2021). Both glycolytic granules and translation factor granules contain mRNAs that are actively translated, and so the purpose of these granules may be to drive localisation of the respective proteins together to increase the efficiency or regulatability of these pathways.

To investigate mechanisms of mRNA localisation, 'swap mRNAs' were designed here, where regions of each mRNA (the ORF, UTR and promoter regions) were exchanged. The first successful mRNA swap made here was the *NIP1 PDC1* ORF swap, which contained the ORF of *PDC1* and maintained the promoter and UTRs of *NIP1*. This exchange was sufficient to drive mRNA localisation to ~10 cytosolic RNA granules, the same as full-length *PDC1*. The opposite mRNA swap, *PDC1 NIP1* ORF, showed a similar result; this construct was localised to one or two granules per cell, the same as full-length *NIP1*. Therefore, this data suggests that either RNA sequence/structural localisation elements are contained within the ORFs of *NIP1* and *PDC1*, or that the production of the nascent protein is capable of driving its own mRNA localisation (or a combination of the two).

It has been shown that the nascent chain of a protein may play roles in regulating both mRNA translation and mRNA localisation. In particular, features of the nascent peptide such as proline repeats and longer, more specific amino acids stretches have been shown to regulate rates of protein translation via ribosomal stalling (Ito and Chiba, 2013; Choi et al., 2018; Doerfel et al., 2013; Muto et al., 2006). With regards to mRNA localisation, perhaps the most well-known example is the localisation of mRNAs to the ER via the signal recognition particle (SRP). mRNAs destined for the ER encode a signal sequence near the start codon, so that the signal sequence is produced early during translation. This signal sequence is recognised and bound by SRP, which directs the mRNA to the ER membrane. mRNA translation is paused while this localisation is occurring, and then relieved once the mRNA reaches the ER membrane and the SRP dissociates (Siegel and Walter, 1986; Hermesh and Jansen, 2013). Other examples include the mammalian *Dia1* mRNA, which is

targeted to the perinuclear ER in a zip-code independent manner, and relies on the active translation of the mRNA (Liao and Liu, 2011), and the Drosophila *anillin* mRNA, the localisation of which requires both the protein coding region of a splice variant and the production of the associated nascent chain (Hirashima et al., 2018). It is therefore possible that *NIP1* and *PDC1* are also examples of mRNAs whose localisation relies on the production of the protein nascent chain.

Two other mRNA swaps were made here; a *NIP1 PDC1* 5'UTR ORF (*NP5'ORF*) and a *NIP1 PDC1* ORF 3'UTR (*NPORF3'*) construct. These constructs contained either the 5'UTR and ORF or the ORF and 3'UTR regions of *PDC1*, respectively, and all other mRNA regions were derived from *NIP1*. Since these constructs contained the ORF (and extra regions) of *PDC1*, it may have been expected that both constructs would also display an mRNA granule pattern that more closely resembled full-length *PDC1*; however both constructs displayed a granule pattern more similar to full-length *NIP1*. It may therefore be possible that *PDC1* UTRs have an inhibitory role in the localisation of the mRNA, possibly as a result of different mRNPs forming due to these additional sequences. However, it was also seen that both of these constructs (*NP5'ORF* and *NPORF3'*) possessed fewer MS2 stem loops than the full-length tagged mRNAs (and *NPORF* and *PNORF* constructs); around 6 stem loops instead of 12. Since the MS2 system does not reach single molecule resolution (and so only multiple mRNAs localised to granules can be visualised), it is therefore possible that fewer GFP molecules being targeted to the *NP5'ORF* and *NPORF3'* constructs meant that only granules containing higher numbers of these mRNAs could be seen, and some granules might have been present in these cells that could not be visualised. A valuable experiment would therefore be to make the *NP5'ORF* and *NPORF3'* constructs with 12 MS2 stem loops and investigate whether the mRNA localisation pattern remains the same or if a more *PDC1*-like granule pattern is seen.

To further investigate whether the ORFs of *NIP1* and *PDC1* are sufficient to drive mRNA localisation, other chimeric mRNA constructs could be made. For example, each ORF could be attached to a GFP reporter and imaged directly (as was done for *ASH1* RNA localisation elements (Bertrand et al., 1998; Gonzalez et al., 1999)), or to other non-localised sequences, to investigate the localisation of these ORFs without surrounding UTR sequences that may also affect mRNA localisation. To attempt to determine whether mRNA localisation is being driven by RNA elements or the nascent protein, a *PDC1* mutant ORF was designed here. This mutant ORF had its primary sequence altered as much as possible while encoding the same amino acids, and so would still produce a functional protein.

Unfortunately due to cloning issues this *PDC1* mutant ORF could not be introduced into yeast in time, although this remains a valuable future experiment.

The abundance of *NIP1*, *PDC1* and the *NIP1 PDC1* ORF mRNAs was then tested. In order to allow modification of these important mRNAs without the complication of growth phenotypes, MS2-tagged derivatives were expressed here using a plasmid-based system with the genomic copies of these genes remaining unaltered. The mRNA localisation profile of plasmid expressed MS2 tagged *NIP1* and *PDC1* appeared to be the same as the genomically MS2-tagged genes (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). Therefore, initially constructs were expressed from a 2µ plasmid; a high copy number plasmid present in around 40-60 copies per cell (Chan et al., 2013). Constructs were later expressed on CEN plasmids, which are expected to be present in a single copy per cell and therefore more closely represent endogenous mRNA expression (Clarke and Carbon, 1985). The pattern of cytosolic RNA granules was unchanged when mRNAs were expressed from CEN plasmids; *NIP1* localised to one or two granules per cell while both *PDC1* and the *NIP1 PDC1* ORF construct localised to ~10 granules per cell. This was expected for both *NIP1* and *PDC1*, as these are also the localisation patterns seen for the mRNAs from the genome (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). For the *NIP1 PDC1* ORF construct, this suggested that it was not merely high numbers of the construct that caused mRNA localisation to many granules per cell.

This was supported by qRT-PCR data that showed each construct was expressed in higher levels from 2µ plasmids compared to CEN plasmids. This was also interesting for *NIP1* mRNA localisation. Although most translation factor mRNAs are relatively low abundance and form one or two granules per cell, *YEF3* and *TEF1* are two translation factor mRNAs that are present in relatively high abundance and form multiple granules per cell (Pizzinga et al., 2019; Lawless et al., 2016). It might therefore have been possible that increasing the numbers of *NIP1* mRNA would result in its localisation to multiple granules, but that does not appear to be the case. This data supports the hypothesis that the purpose of the actively translating granules is to localise together mRNAs encoding proteins of a similar function; if all other translation factor mRNAs are present at their endogenous levels and mostly present in one or two granules, then *NIP1* mRNA would also only be required in the same granules.

Finally, MEME analysis on glycolytic and translation factor mRNA sequences was undertaken here to identify conserved motifs between these sequences. The translation factor mRNAs did not appear to contain any particular consensus sequences, suggesting that their mRNA localisation may not rely on motifs of similar primary RNA sequence. Instead, these mRNAs may rely on structural motifs, or possibly by nascent protein

interactions as discussed above. The glycolytic mRNAs did appear to contain reasonably conserved consensus sequences in their ORFs that were not present in the *NIP1* mRNA sequence. However, these motifs did not appear to carry a single specific sequence (no specific nucleotide was present in the same place for every mRNA), and so further experiments should be carried to determine if these motifs are involved in mRNA localisation. Specifically, removal of these motifs (if possible with regards to protein function) and introduction of these motifs into other sequences may allow their effect on mRNA localisation to be determined. These experiments, coupled with the investigation of the mutant *PDC1* ORF, may provide more evidence as to whether mRNA localisation to cytosolic RNA granules is driven by RNA elements or the production of the nascent protein chain.

Results

Chapter 4: Development of the dCas9-GFP_3 RNA-targeting system in yeast.

4. Development of the dCas9-GFP$_3$ RNA-targeting system in yeast.

A key shortcoming in the current technologies for mRNA visualisation in cells is that for live cell work the mRNA needs to be modified. So in order to study unmodified endogenous mRNA localisation, technologies utilise fixation approaches (Femino et al., 1998; Samacoits et al., 2018). As such, it is currently not possible to study endogenous mRNA at single molecule resolution in live cells (Kim et al., 2019). In an effort to overcome this shortcoming, in this chapter a dCas9 RNA-targeting system was developed and utilised to allow the study of endogenous mRNA localisation in live cells.

Although the MS2 bacteriophage system with its RNA stem loops and coat protein has been used extensively to study mRNA localisation in live cells, the validity of interpretations using this technique have been called into question (Garcia and Parker, 2015; Haimovich et al., 2016; Garcia and Parker 2016; Heinrich et al., 2017). For certain mRNAs, the modification of the target mRNA with inserted MS2 stem loop structures has been shown to affect rates of mRNA decay and translation, and as a result can impact upon localisation of the mRNA (Heinrich et al., 2017). Therefore, all mRNA localisation experiments previously described in the Ashe lab have been verified using orthogonal smFISH approaches to examine the localisation of unmodified mRNA (Pizzinga et al., 2019; Morales-Polanco et al., 2021). However, smFISH is a fixed cell method of visualising mRNA, and cellular fixation might also impact upon mRNA localisation. In addition, smFISH cannot be used to study mRNA dynamics over time and in response to changing conditions. Therefore, a dCas9 RNA-targeting system, previously described to study mRNA localisation in live mammalian cells (Nelles et al., 2016), has been developed in this chapter for use in yeast.

The dCas9 RNA-targeting system relies upon a protein fusion of GFP with an inactivated form of bacterial Cas9 protein (dCas9) (Figure 4.1). Targeting of this dCas9-GFP fusion to the mRNA is achieved by a single guide RNA (sgRNA), which contains 20 base pairs that are specific and complimentary to the target mRNA, followed by a conserved stem-loop scaffold region that is recognised and bound by the dCas9-GFP fusion. The main advantages of this system are that the mRNA remains unmodified and sgRNAs can be generated relatively easily for different mRNAs, whereas the MS2 system and smFISH can be more technically difficult and time-consuming. Therefore such a dCas9-based system might be well-suited to screen a greater number of mRNAs whose localisation patterns have not been previously investigated.

In this chapter, the development of a dCas9 system for use in yeast is described. A dCas9-GFP fusion protein plasmid is generated, as well as specific sgRNA plasmids targeted

against several mRNAs. Using this dCas9 system, similar mRNA localisation patterns are observed relative to those previously characterised using MS2 and smFISH techniques (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). However, the signal obtained for the new dCas9 system is weak and hence requires extended exposure times. As a result, further development to increase the signal to noise ratio of this system was undertaken to improve this system. An initial attempt to increase the system brightness was to replace the three GFPs fused to the dCas9 protein with three mNeonGreen fluorescent proteins, since mNeonGreen has been shown to fluoresce more brightly than GFP (Shaner et al., 2013; Hostettler et al., 2017). However, the mNeonGreen fusion proteins constructed here did not demonstrate the same behaviour as the dCas9-GFP system and appeared to prevent targeting of the dCas9 to mRNAs.

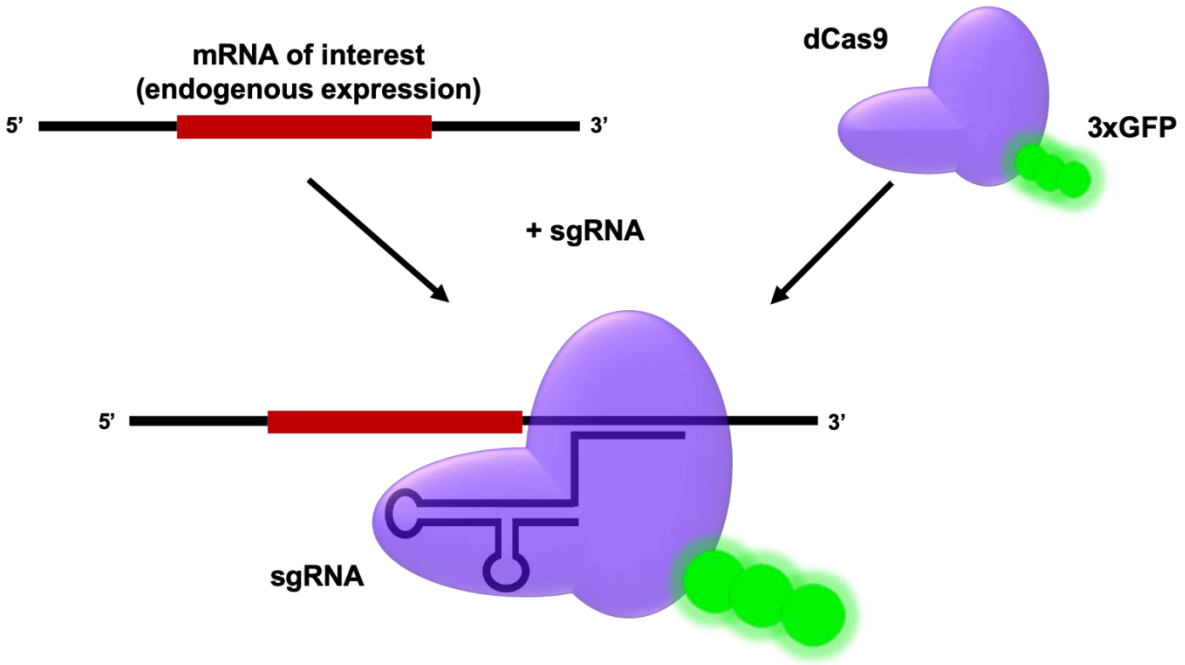


Figure 4.1: An overview of the dCas9 RNA-targeting system. The mRNA being targeted remains unaltered and the dCas9 system detects the endogenous mRNA. Constructs or cassettes driving production of the dCas9-GFP_3 fusion protein and a sgRNA are introduced into the cell. The sgRNA has a 20bp recognition region allowing binding to the mRNA of interest; most often this is designed to target the 3'UTR. The conserved stem-loop scaffold region of the sgRNA is recognised and bound by the dCas9-GFP_3 fusion protein, and therefore the position of the mRNA can be visualised in live cells.

4.1. Construction of the dCas9 RNA-targeting system in yeast:

In order to test whether or not a dCas9 RNA-targeting system will allow the study of mRNA localisation both the dCas9-GFP fusion protein and a sgRNA targeting the mRNA of interest must be co-expressed. To achieve this, two plasmids were designed; a dCas9-GFP plasmid and a specific sgRNA plasmid. The dCas9-GFP fusion plasmid was generated by Gibson cloning (Figure 4.2). Here the open reading frame of dCas9 with a C-terminal NLS was inserted in frame with three GFPs in a yeast expression plasmid. The construct was verified by mapping with appropriate restriction enzymes and sequencing of the plasmid DNA at the junction sites between DNA fragments (Figures 4.3 and 4.4).

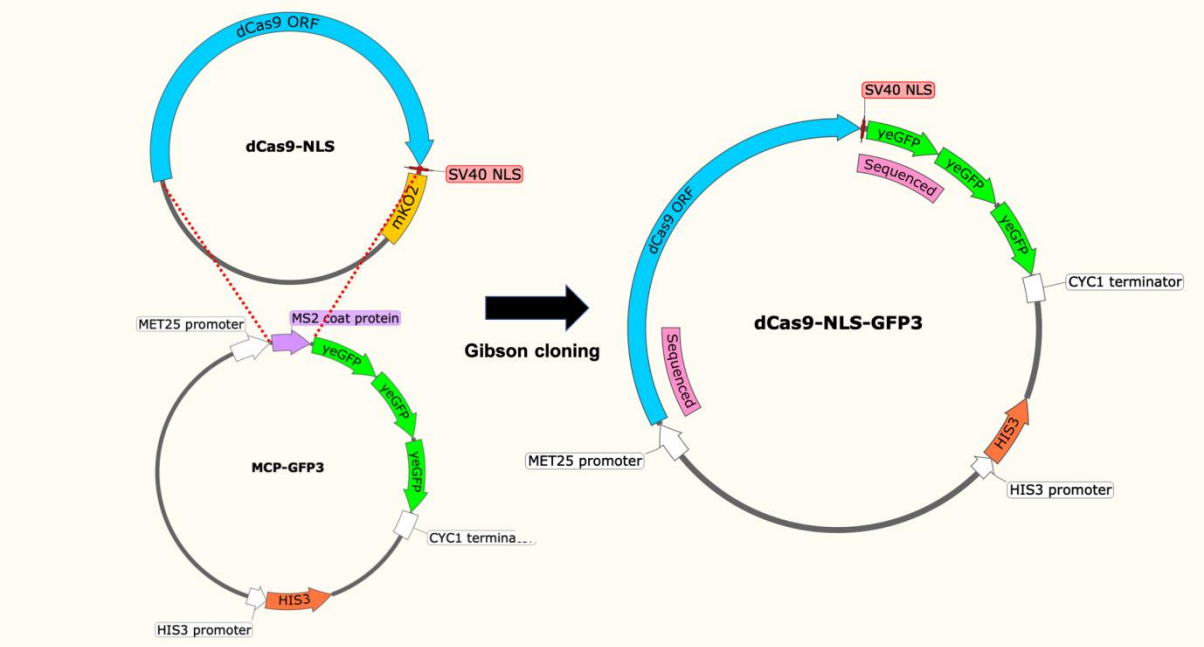


Figure 4.2: Gibson assembly of the dCas9-NLS-GFP$_3$ plasmid. The dCas9 ORF and NLS was isolated and amplified from a mammalian vector housing a dCas9-mKO2 orange fluorescent fusion protein cassette. This DNA fragment was then used to replace the MS2 coat protein sequences in a yeast vector originally used for the MS2 system. Ultimately, a yeast expression plasmid was obtained that contained yeast compatible promoters and terminators and could express the dCas9-GFP$_3$ fusion. A *HIS3* marker gene was also present allowing selection of transformants in appropriate yeast strains. The vector was verified using automated Sanger sequencing across the junction regions as shown in the figure.

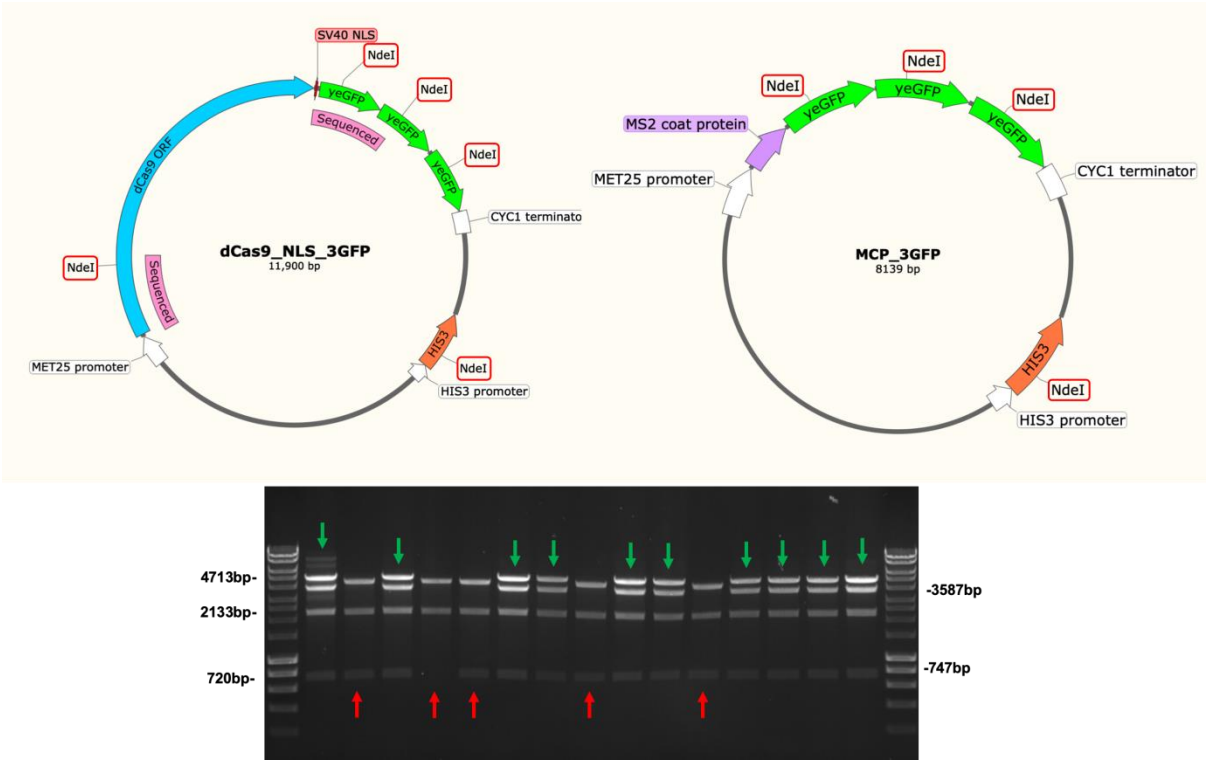
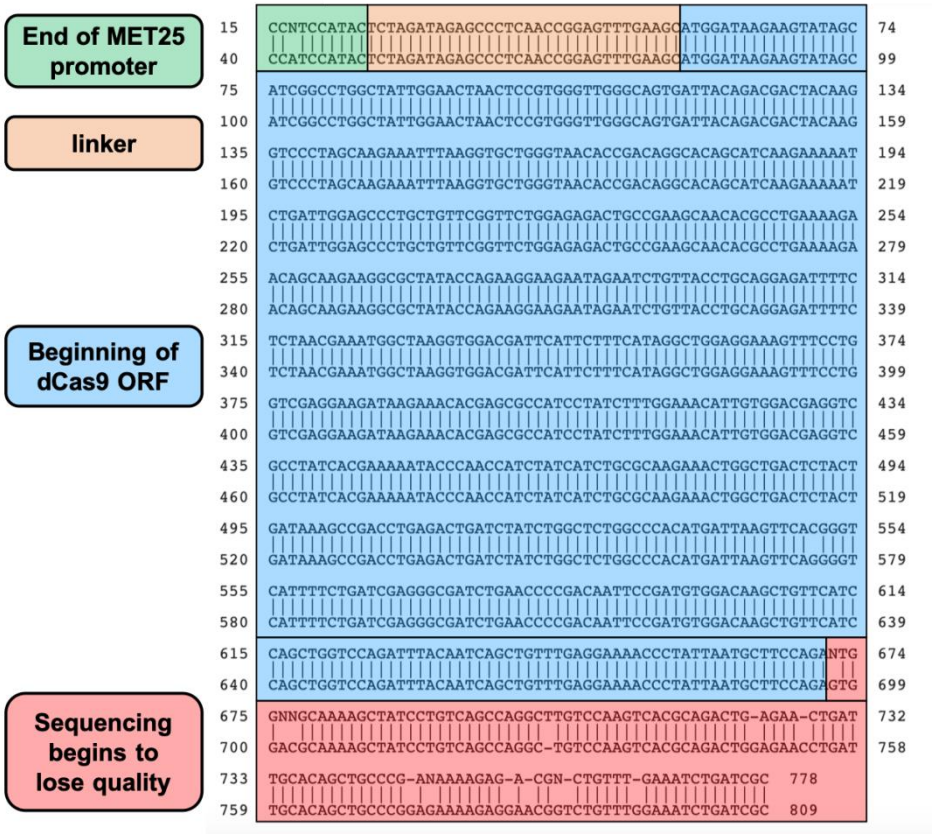


Figure 4.3: Restriction enzyme mapping of dCas9-GFP$_3$ constructs. All gels are shown with HyperLadder 1kb markers. Constructed plasmid DNA was digested using the restriction enzyme *Nde*I, which cuts at multiple sites in the dCas9-GFP$_3$ plasmid and the MCP-GFP$_3$ parent plasmid as shown in the figure. Both the dCas9-GFP$_3$ and MCP-GFP$_3$ digest produced bands at 720bp, 747bp and 2133bp, as shown on the gel. The green arrows on the gel indicate plasmid minipreps that showed the expected pattern for dCas9-GFP$_3$, including bands at 3587bp and 4713bp. The red arrows indicate minipreps that showed the expected pattern for the MCP-GFP$_3$ parent plasmid; one additional band at 4537bp. Several of the samples showing the correct pattern for the dCas9-GFP$_3$ construct were then sent for sequencing.


End of MET25 promoter
linker
Beginning of dCas9 ORF
Sequencing begins to lose quality

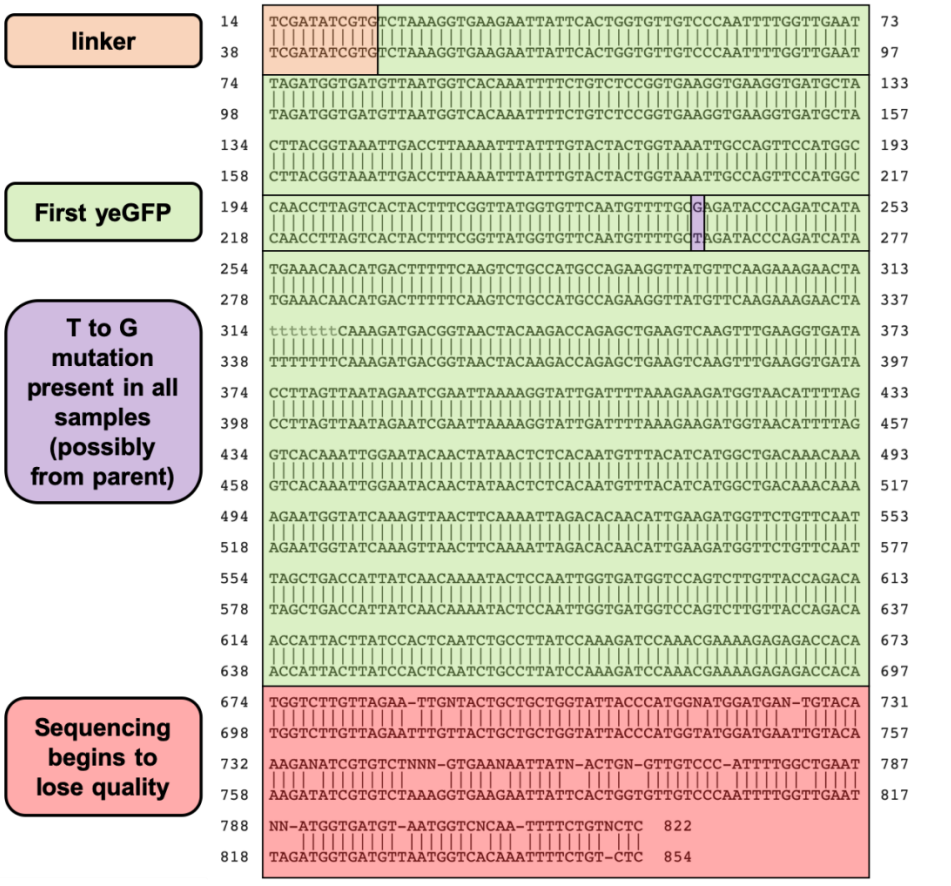

linker
First yeGFP
T to G mutation present in all samples (possibly from parent)
Sequencing begins to lose quality

Figure 4.4: Sanger sequencing of the dCas9-GFP$_3$ plasmid. A BLAST comparison of the sequenced dCas9-GFP$_3$ plasmid compared to the expected sequence. The plasmid was sequenced at both cloning junction regions (as shown in Figure 4.3). Sequencing at the first junction shows the end of the MET promoter, linker region and beginning of dCas9 ORF is correct, while the second junction sequencing shows the linker region and first yeGFP in-frame, with one T to G mutation in the yeGFP. This mutation was present in all sequenced minipreps, and so possibly originated from the parent plasmid. The sequencing is accurate for around 700 base pairs, where errors caused by polymerase drop-off near the sequence end are highlighted in red.

sgRNA expression vectors were also made using a Gibson assembly strategy. The sgRNA scaffold vector (Mans et al., 2015) contains a conserved stem-loop scaffold preceded by a 20bp variable region. The 20bp region is replaced with 20bps specific for the mRNA of interest. Oligomers were ordered that contained a 20bp mRNA targeting region flanked by 50bp regions complimentary to the sgRNA scaffold vector (Figure 4.5). Correct targeting sequence insertion was verified by sequencing across the variable and scaffold regions for each sgRNA vector made (Figure 4.6). Targeting sequences were designed to recognise 20bp in the 3'UTR for almost all target mRNAs except *PDC1* and *PDC5* mRNAs. *PDC1* and *PDC5* possess very similar 3'UTR sequences, and so sgRNAs were made to target differing regions of their ORFs. A complete list of sgRNAs designed and made is shown in Table 6.

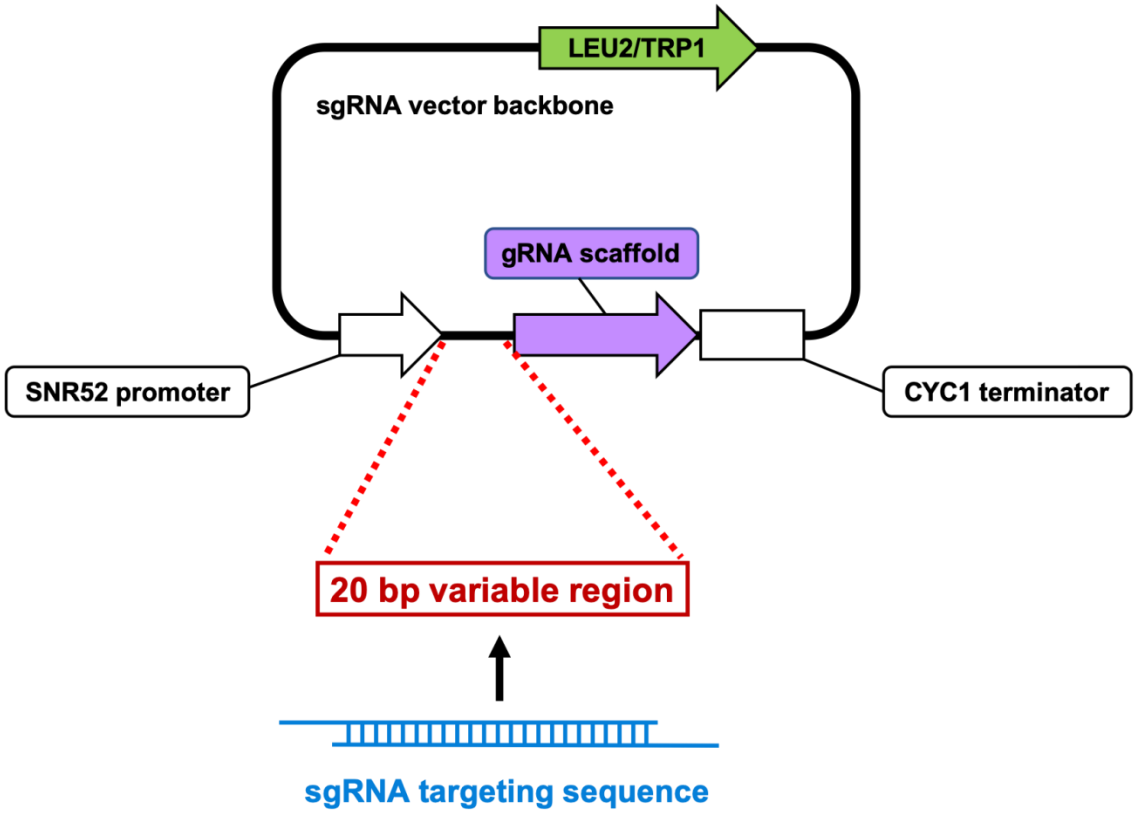


Figure 4.5: Gibson assembly of sgRNA vectors. sgRNA vectors were constructed using a Gibson assembly strategy to replace the 20bp variable region in the sgRNA backbone vector with 20bp specific to the mRNA of interest. The sgRNA expression plasmid contained a yeast marker (*LEU2* or *TRP1*) to allow transformation into yeast. The vector also contained the gRNA scaffold sequence necessary for recognition by dCas9. The sgRNA targeting sequence is a 120bp double-stranded oligomer that contains the 20bp targeting sequence flanked by 50bp complimentary to the sgRNA vector backbone.

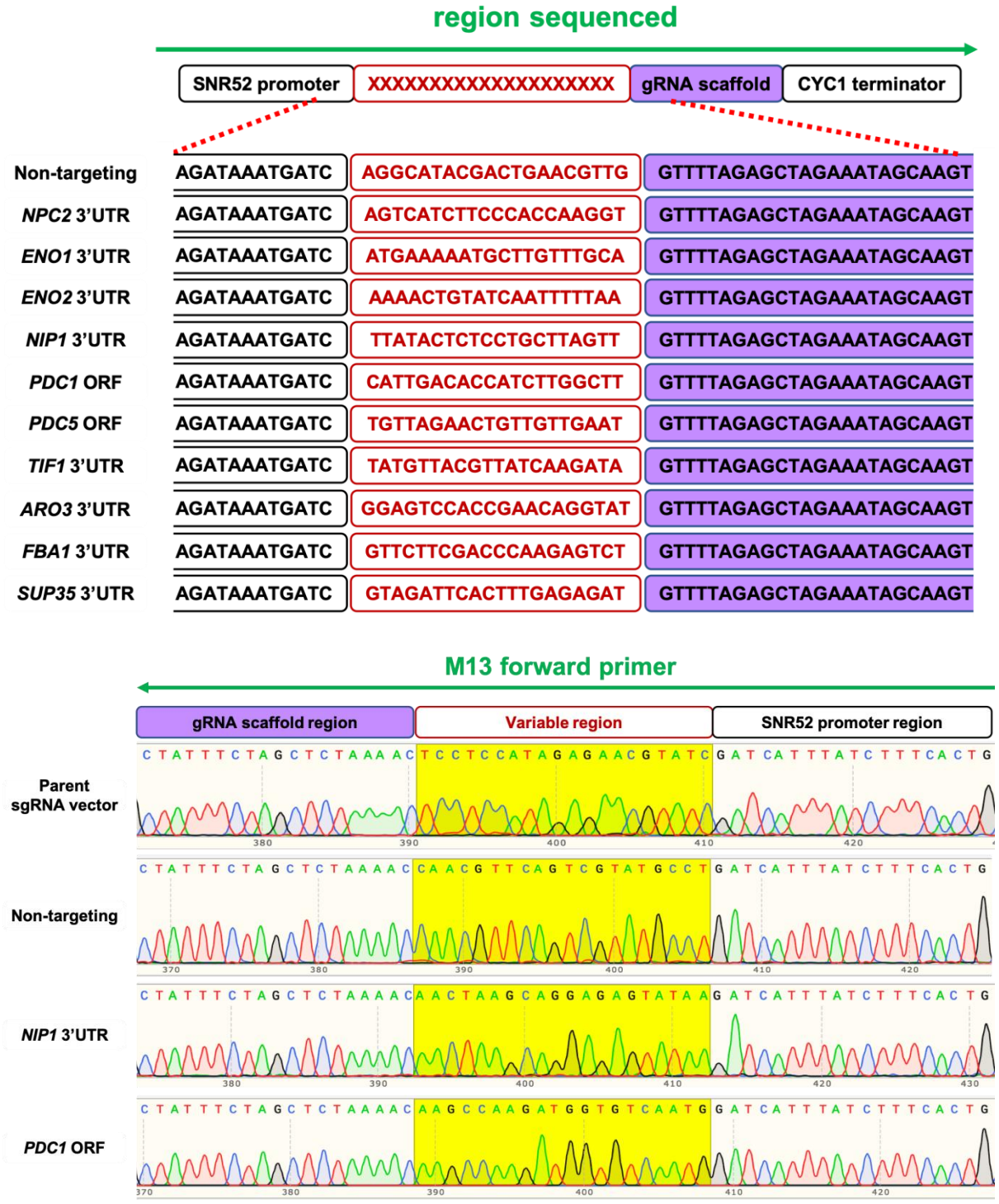


Figure 4.6: Sequencing verification of sgRNA vectors constructed. Each sgRNA vector constructed was verified by automated Sanger sequencing across the promoter, targeting region, gRNA scaffold, and terminator sequences. Sequencing was conducted using the M13 forward primer that binds the antisense DNA strand, so the example sequencing data shown in this figure displays the reverse complement of the sequence. A full list of sgRNA constructs made can be found in the table below.

mRNA Targeted	Target Region[1]	Imaged with dCas9-GFP$_3$[2]	Imaged with dCas9-MoonTag[3]
Non-targeting sequence	N/A	yes	yes
NPC2	3'UTR	yes	yes
ENO1	3'UTR	yes	yes
ENO2	3'UTR	yes	yes
NIP1	3'UTR	yes	yes
TIF1	3'UTR	yes	yes
PDC1	ORF		yes
PDC5	ORF		yes
FBA1	3'UTR		yes
SUP35	3'UTR		yes
ARO3	3'UTR		yes

Table 6: A list of sgRNA vectors constructed and their targets.

[1] See Figure 4.6.

[2] See Figures 4.8 and 4.10.

[3] See next results chapter (Section 5.).

4.2. Imaging of the dCas9 RNA-targeting system in yeast:

To verify that the plasmids constructed above are functioning as might be expected, the dCas9-GFP_3 plasmid was first transformed into yeast and imaged (Figure 4.7A). The expression of the dCas9-GFP_3 protein was verified by Western blotting (Figure 4.9), where antibodies against GFP and dCas9 both detected the expected protein band around 244 kDa. Intense fluorescent signal was observed using the GFP channel on the microscope in the nucleus of cells, where the fluorescent signal localised with 4′,6-diamidino-2-phenylindole (DAPI) (Figure 4.7A). DAPI also strongly stains mitochondrial DNA in live yeast microscopy (Williamson and Fennell, 1979), shown by localised DAPI signal outside the nucleus (Figure 4.7A). Since the dCas9 cassette is followed by an NLS, the observed localisation for the GFP signal represents the expected scenario. A nuclear localisation for the dCas9 protein is important since it would allow interaction with the sgRNAs as they are made, and allows the location of the dCas9 to be visualised when it is not targeted to mRNA.

Two control sgRNAs were then tested by transforming each sgRNA expression plasmid separately into yeast cells already expressing dCas9-GFP_3. The first plasmid carries a non-targeting sgRNA, where the 20bp targeting region was replaced with 20bps that are designed to not target any mRNA. This was achieved by generation of random 20 nucleotide sequences that were analysed by BLAST and ATUM webserver. For the non-targeting sequence selected, the ATUM webserver did not identify any possible sgRNAs for the sequence, indicating that no yeast sequences would be targeted by Cas9 when using this sgRNA. BLAST alignments of this sequence show the maximum alignment to be 14 nucleotides (Figure 4.8), where a minimum of 17 nucleotides is required for dCas9 targeting (Fu et al., 2014). The presence of this non-targeting sgRNA did not affect the localisation of dCas9-GFP_3 which remained in the nucleus (Figure 4.7B). This result is consistent with this sgRNA not being targeted. The second control plasmid carries a sgRNA designed to target the *NPC2* mRNA; *NPC2* is a previously characterised non-localised cytosolic mRNA (Lui et al., 2014; Figure 3.1). When this plasmid is present, the dCas9-GFP signal is altered and shifts from being predominantly nuclear to being present throughout the cytosol (Figure 4.7B). Not only does this result suggest that the *NPC2* sgRNA is expressed, but the result is also consistent with previous observations showing that the *NPC2* mRNA is non-localised within the cytosol (Lui et al., 2014; Figure 3.1). Significantly, neither of these control experiments provides any evidence for granular mRNA localisation.

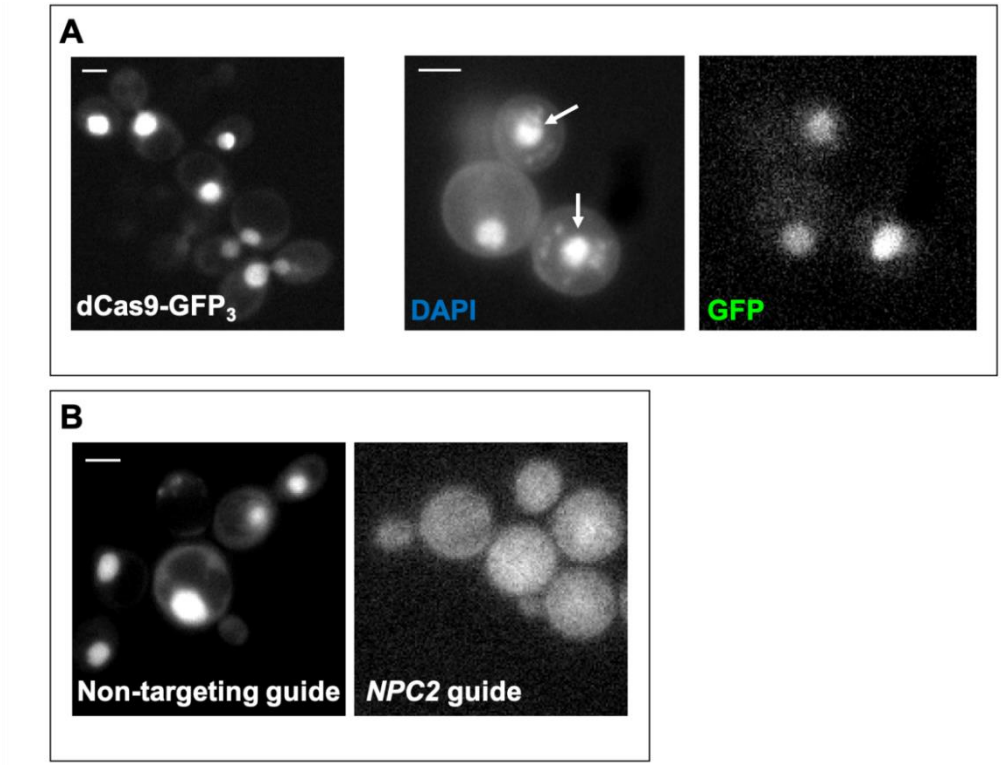


Figure 4.7: dCas9-GFP$_3$ in yeast and control sgRNAs. Panel A shows yeast cells expressing dCas9-GFP$_3$ in the GFP channel (left hand panel) and zoomed in images of these cells in either the DAPI or GFP channels (right hand images). White arrows indicate nuclear DAPI signal. Panel B shows yeast expressing both the dCas9-GFP$_3$ plus either a non-targeting sgRNA or a sgRNA targeting *NPC2* mRNA. The dCas9 signal is confined to the nucleus in both the dCas9-GFP$_3$ only yeast and yeast also expressing the non-targeting sgRNA. In yeast expressing the *NPC2* sgRNA, the fluorescent signal is dispersed throughout the cells. Scale bars: 3μm.

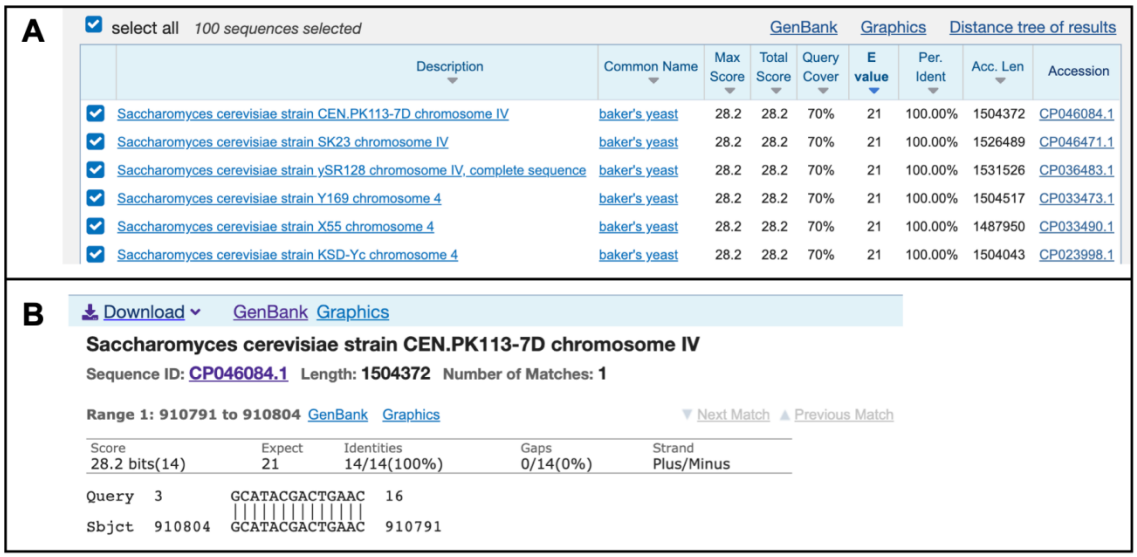


Figure 4.8: BLAST results for non-targeting sgRNA. The top panel (A) shows the top alignment results for the non-targeting sgRNA sequence for *Saccharomyces cerevisiae*, found on chromosome IV. Panel B shows the exact sequence match of the sgRNA to the yeast genome, which is 14 nucleotides in length. This 14 nucleotide sequence alignment was the same in all of the top results.

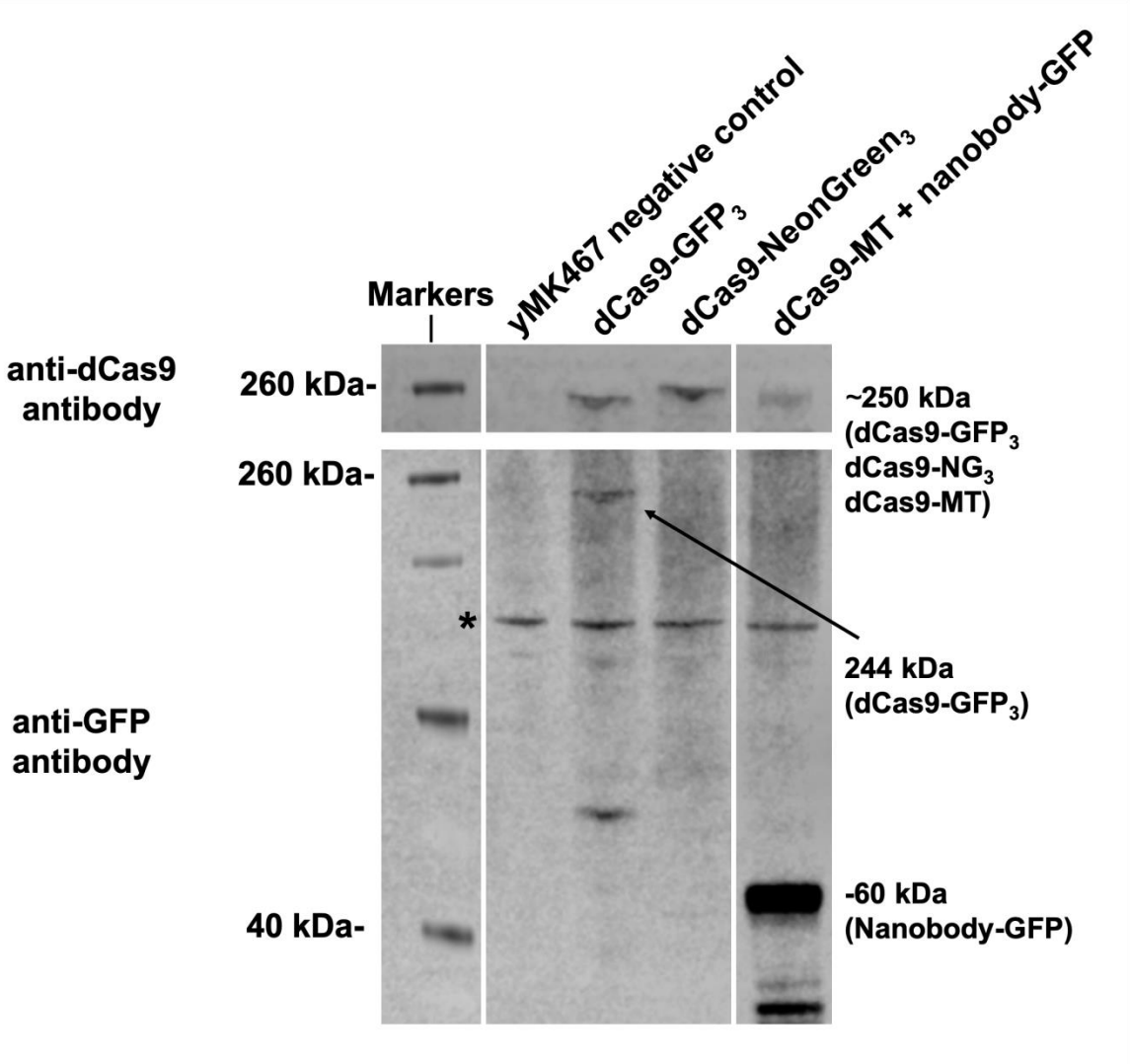


Figure 4.9: Western blot of dCas9-GFP_3 and dCas9-NG_3 proteins. Lane 1 contains protein extracts from the parent yeast strain yMK467 as a negative control. Lane 4 contains sample from the improved dCas9 system, discussed in the next chapter (Section 5). Lanes 2 and 3 contain dCas9-GFP_3 and dCas9-NG_3 protein extracts, respectively. Using an anti-dCas9 antibody (top blot), bands just below 260 kDa are shown, corresponding to the dCas9 fusion proteins (dCas9 is 160 kDa and both GFP_3 and NG_3 are around 84 kDa). Using an anti-GFP antibody (bottom blot), the dCas9-GFP_3 band is again present at just under 260 kDa. The anti-GFP antibody also appears to produce a non-specific band present in all samples (shown by the asterisk). Irrelevant lanes on the blot have been removed and replaced by white lines to separate distinct regions of the same gel.

The dCas9 system was used to visualise the localisation of mRNAs that had been shown previously to localise to cytoplasmic granules using the MS2 system and smFISH (see Figure 3.1 and Figure 4.10 bottom panel) (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). These mRNAs were *NIP1* and *TIF1*, both previously shown to localise to a few granules per cell (1 – 5), and *ENO1* and *ENO2*, previously shown to localise to multiple granules per cell (10 – 20) (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). The sgRNA plasmids targeted against these mRNAs were transformed into yeast strains that already carry the dCas9-GFP_3 plasmid and fluorescence microscopy was undertaken to compare the signal obtained using the dCas9 system relative to the MS2 system in live yeast cells,

For all four of the Cas9 strains tested, the presence of the sgRNA plasmid causes a different fluorescent pattern to be observed relative to the strain expressing just dCas9-GFP_3 (Figure 4.10 cf 4.8). Instead of a defined nuclear signal, the expression of sgRNAs targeted against *ENO1* and *ENO2* leads to defined aggregates of fluorescence in the cytosol. This profile is similar to what is observed using the MS2 system but less well resolved (Figure 4.10). The expression of sgRNAs against *NIP1* and *TIF1* leads to a small number of focussed granules in the cytosol within each cell - as well as a diffuse cytosolic background (Figure 4.10). Once again, this correlates well with what has been previously described for the MS2 system. So overall it seems likely that the dCas9-GFP_3 system has recapitulated previously described mRNA localisation patterns.

However, the dCas9-GFP system is less bright than the MS2 system due to the targeting of fewer GFPs to each mRNA. This is particularly evident for dCas9 system images targeting *TIF1* mRNA, where the fluorescent signal is weak and cytosolic granules are difficult to distinguish from the background in most cells (Figure 4.11 top panel). To be able to image cytosolic granules these images had to be taken using very high exposure times (around 1 second), which meant that multiple Z-stacked images could not be taken before bleaching occurred and a fluorescent signal could no longer be seen. When compared to the targeting of *TIF1* using the MS2 system at the same exposure time, the difference in brightness can be clearly seen between the two systems (Figure 4.11). Therefore, to make the dCas9 system a viable RNA-tracking system for use in yeast, experiments to increase the brightness of the system were undertaken.

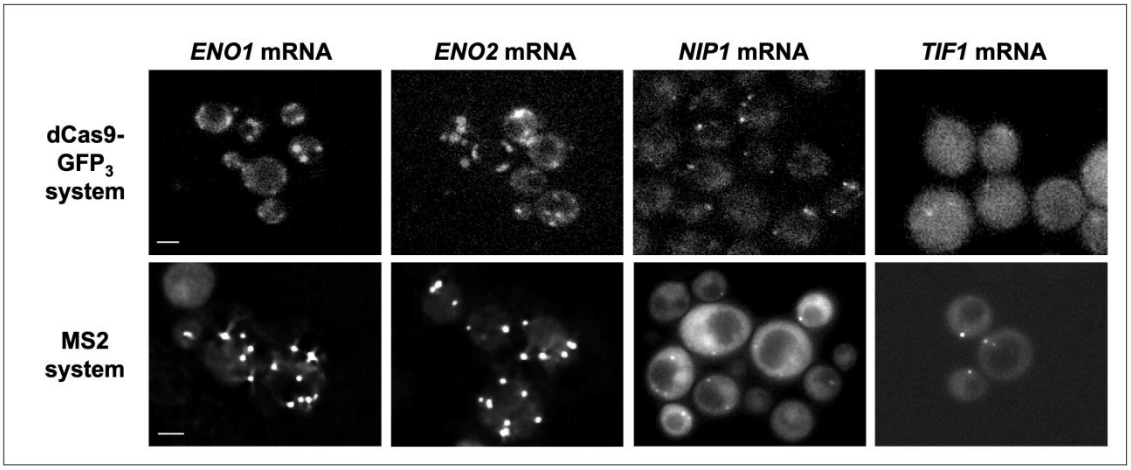


Figure 4.10: A comparison of mRNA localisation using the MS2 system and the dCas9 RNA-targeting system. The top panel shows the localisation patterns of *ENO1*, *ENO2*, *NIP1*, and *TIF1* mRNAs when visualised using the dCas9 system, while the bottom panel shows the same mRNAs viewed with the MS2 system. The mRNA localisation patterns appear similar, where *ENO1* and *ENO2* mRNAs are present in many granules per cell, whereas *NIP1* and *TIF1* are present in 1 – 5 granules per cell. However the dCas9-GFP system is less bright than the MS2 system and does not give a strong enough signal to be accurately quantified. Scale bars: 3μm.

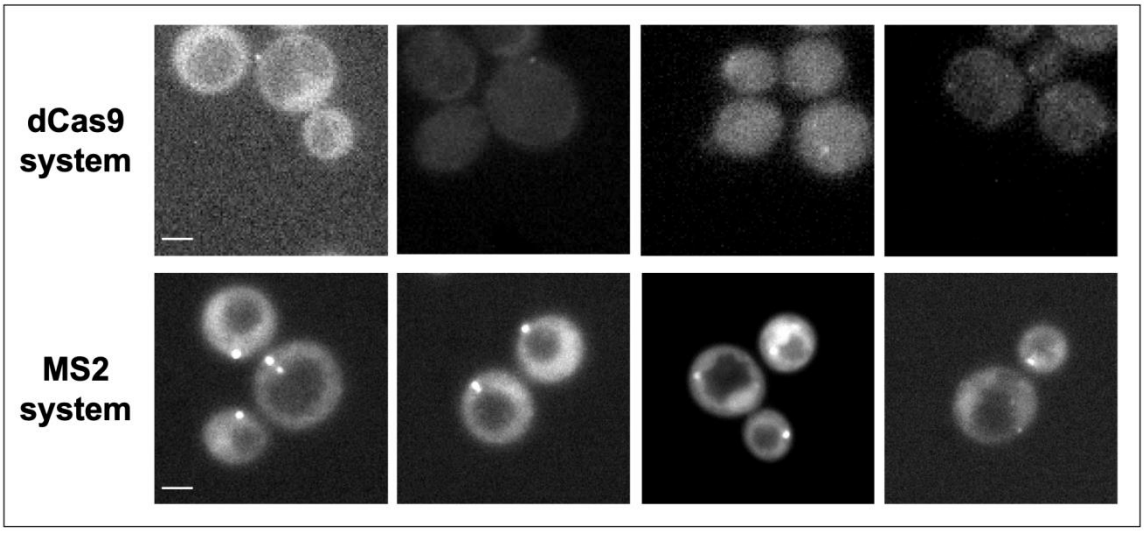


Figure 4.11: *TIF1* mRNA imaged with the dCas9 system and MS2 system at the same exposure time. Both panels show *TIF1* mRNA imaged at 1s exposure time. In the top panel *TIF1* mRNA is imaged using the dCas9 system, while the bottom panel shows *TIF1* imaged using the MS2 system. All images were taken as single plane images to allow comparison. Scale bars: 2μm.

4.3. Construction and imaging of dCas9-mNeonGreen$_3$ and MCP-mNeonGreen$_3$:

The first attempt to increase the brightness of the dCas9-GFP$_3$ system was to replace the three GFPs with three mNeonGreens, since mNeonGreen has been previously shown to fluoresce around three times more brightly than GFP (Shaner et al., 2013; Hostettler et al., 2017). An MCP-mNeonGreen$_3$ (MCP-NG$_3$) plasmid was also constructed to investigate the effect this had on the MS2 system. Constructs were made by Gibson assembly, where cloning was initially designed to amplify three mNeonGreen proteins from a vector housing a single mNeonGreen and replace the three GFPs in the dCas9-GFP$_3$ vector in one step. Gibson cloning resulted in the construction of a plasmid that contained three mNeonGreens in the backbone yeast vector, although this vector had lost the dCas9 ORF and NLS (Figure 4.12). This mNeonGreen$_3$ only vector was then used to construct both a dCas9-mNeonGreen$_3$ (dCas9-NG$_3$) and an MCP-mNeonGreen$_3$ plasmid using restriction cloning (Figure 4.12). The construction of all vectors was verified by restriction enzyme mapping and sequencing of the plasmid DNA at the junction sites between DNA fragments (Figure 4.13).

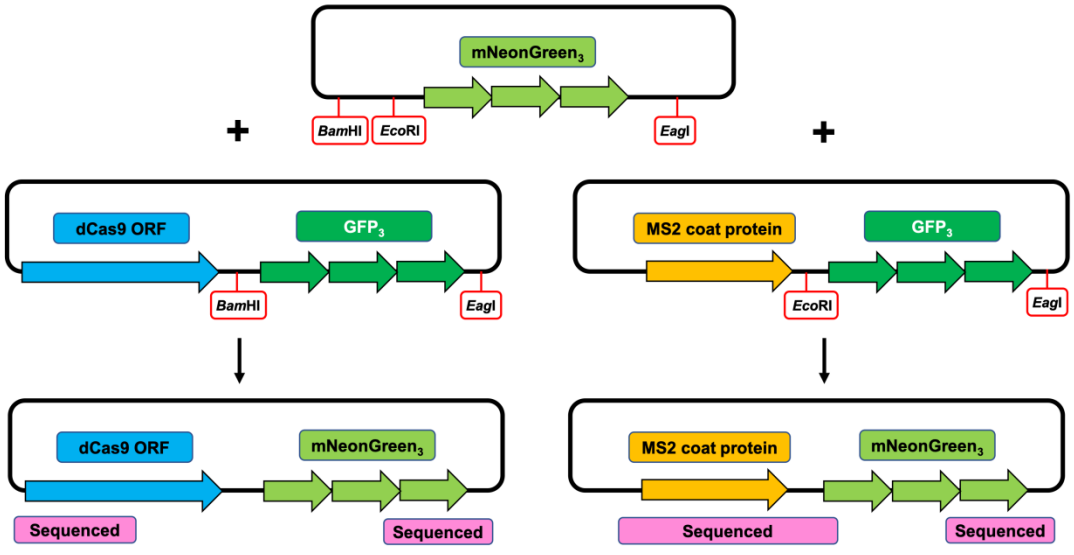


Figure 4.12: Construction of dCas9-mNeonGreen$_3$ and MCP-mNeonGreen$_3$. Constructs were made by restriction cloning to replace three GFPs for three mNeonGreens using *Bam*HI and *Eag*I sites for dCas9-NG$_3$, and *Eco*RI and *Eag*I sites for MCP-NG$_3$ as shown in the figure. NeonGreen constructs were then verified by restriction enzyme mapping and sequencing across junction regions, shown in purple boxes in the figure.

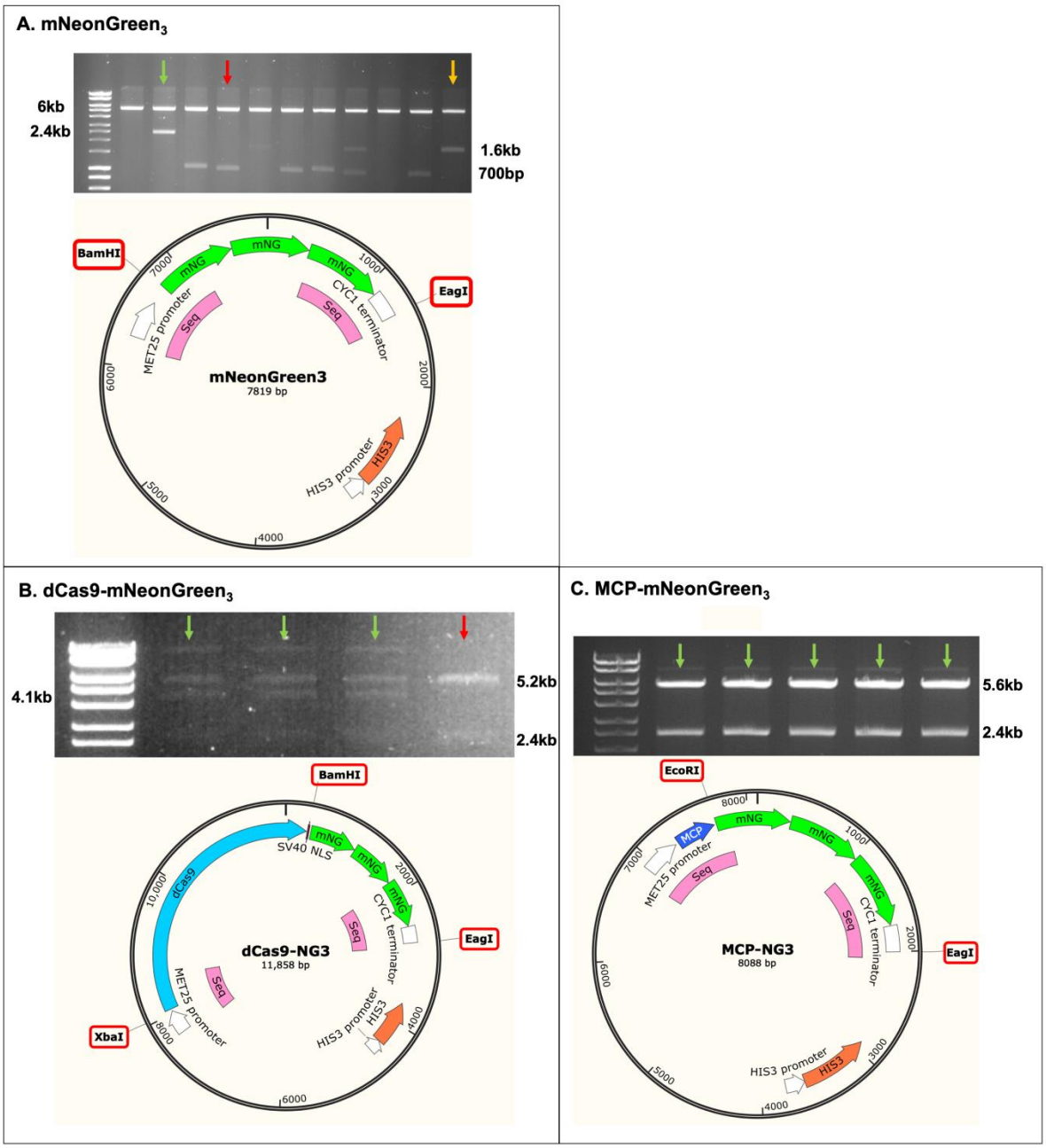


Figure 4.13: Restriction enzyme mapping and sequencing of mNeonGreen$_3$, dCas9-mNeonGreen$_3$ and MCP-mNeonGreen$_3$ constructs. Panel A shows the restriction digest for possible mNeonGreen3 samples. The *Eag*I/*Bam*HI digest was designed to cut on either side of the mNeonGreen region, giving a backbone band of 6kb and a band corresponding to the number of mNeonGreens present. The red arrow denotes a sample containing a single mNeonGreen at 747bp, the orange arrow denotes a sample containing two mNeonGreens at 1.6kb (including a linker region), and the green arrow denotes a sample containing three mNeonGreens at 2.4kb. Panel B shows the restriction digest for possible dCas9-NG$_3$ samples, where the restriction enzymes (*Eag*I, *Bam*HI and *Xba*I) were designed to cut out the backbone region (5.2kb), the dCas9 ORF (4.1kb), and the three mNeonGreen region (2.4kb). Panel C shows the digest for possible MCP-NG$_3$ samples, where enzymes (*Eag*I and *Eco*RI) were chosen to cut out the backbone region (5.6kb) and the three mNeonGreen region (2.4kb). For all panels, samples with the correct digest patterns were then sent for Sanger sequencing, where sequenced regions are highlighted in pink on the plasmid representations.

Both the dCas9-NG_3 and MCP-NG_3 plasmids were then transformed into yeast and imaged. The expression of the dCas9-NG_3 fusion protein was verified with Western blotting where antibodies against dCas9 detected a corresponding band around 244 kDa (Figure 4.9), however antibodies against mNeonGreen and the MS2 coat protein did not work in this project. Yeast cells expressing dCas9-NG_3 appeared as bright as or brighter than yeast expressing dCas9-GFP_3 (Figure 4.14B). As expected, the fluorescent signal was confined to the nucleus (Figure 4.14A top panel). MCP-NG_3 containing cells appeared relatively dim; this may have been expected since the MS2 coat protein is not attached to an NLS and so the NeonGreen signal would have been expected to be diffuse. However, in some cells a faint nuclear localisation could be seen (Figure 4.14A bottom panel).

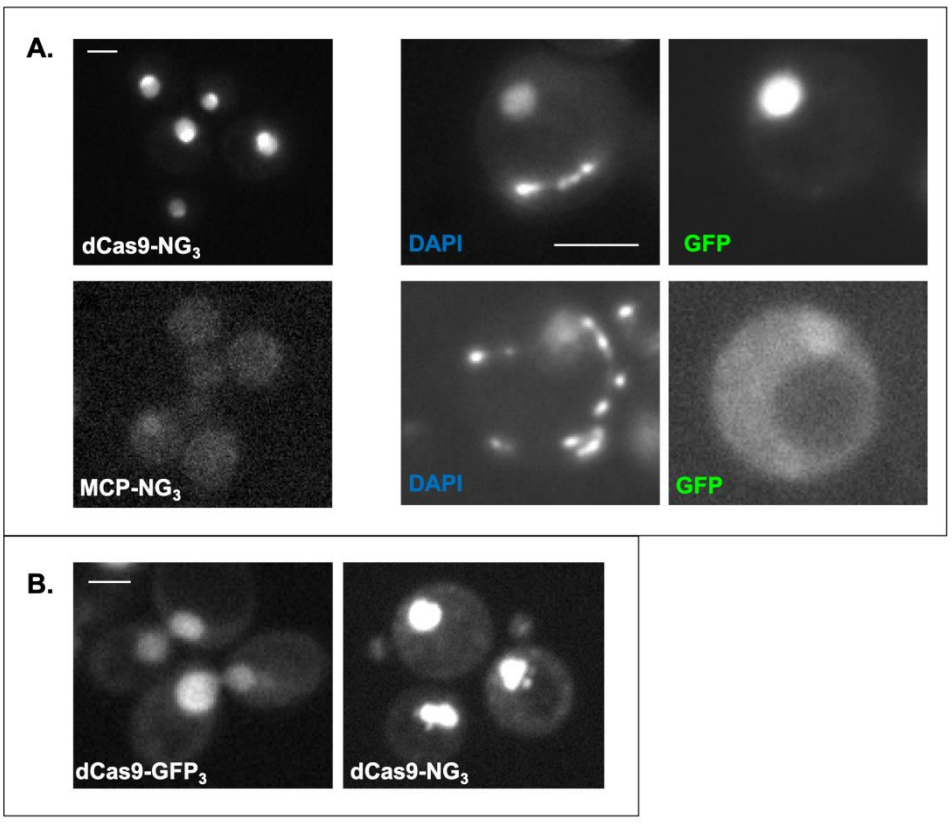


Figure 4.14: Images of dCas9-mNeonGreen3 and MCP-mNeonGreen3 in yeast. Panel A shows yeast cells expressing either dCas9-NG_3 or MCP-NG_3 plasmids only in the GFP channel (left hand side), along with zoomed in images taken in DAPI and GFP channels of the same cells (right hand images). The dCas9-NG_3 is confined to the nucleus, and some nuclear localisation can also be seen for the MCP-NG_3. Panel B shows images of cells expressing dCas9-GFP_3 or dCas9-NG_3 plasmids taken at the same exposure time (800 ms). All images were taken of a single image plane to avoid bleaching. Scale bars: 3μm.

Cells expressing dCas9-NG_3 were then transformed with a sgRNA targeting *ENO1*, *ENO2* or *NIP1* mRNA. None of these cells displayed the same fluorescence patterns previously seen using the dCas9-GFP_3 system (and MS2 system and smFISH); instead the dCas9-NG_3 appeared to remain in the nucleus (Figure 4.15 top panel). The MCP-NG_3 plasmid was transformed into cells containing either genomically MS2-tagged *ENO2* mRNA or genomically MS2-tagged *ARO1* mRNA. Both of these MS2-tagged mRNAs had been previously shown in the lab to localise to cytosolic granules when viewed using the MCP-GFP_3 plasmid. Contrastingly, these cells did not show any evidence of cytosolic mRNA localisation when viewed using MCP-NG_3 (Figure 4.15 bottom panel). Interestingly, in the *ENO2* tagged strain there was again some evidence for faint nuclear localisation (Figure 4.15 bottom panel). Both dCas9-NG_3 and MCP-NG_3 proteins therefore failed to recapitulate previously studied mRNA localisation, and were not used in subsequent experiments.

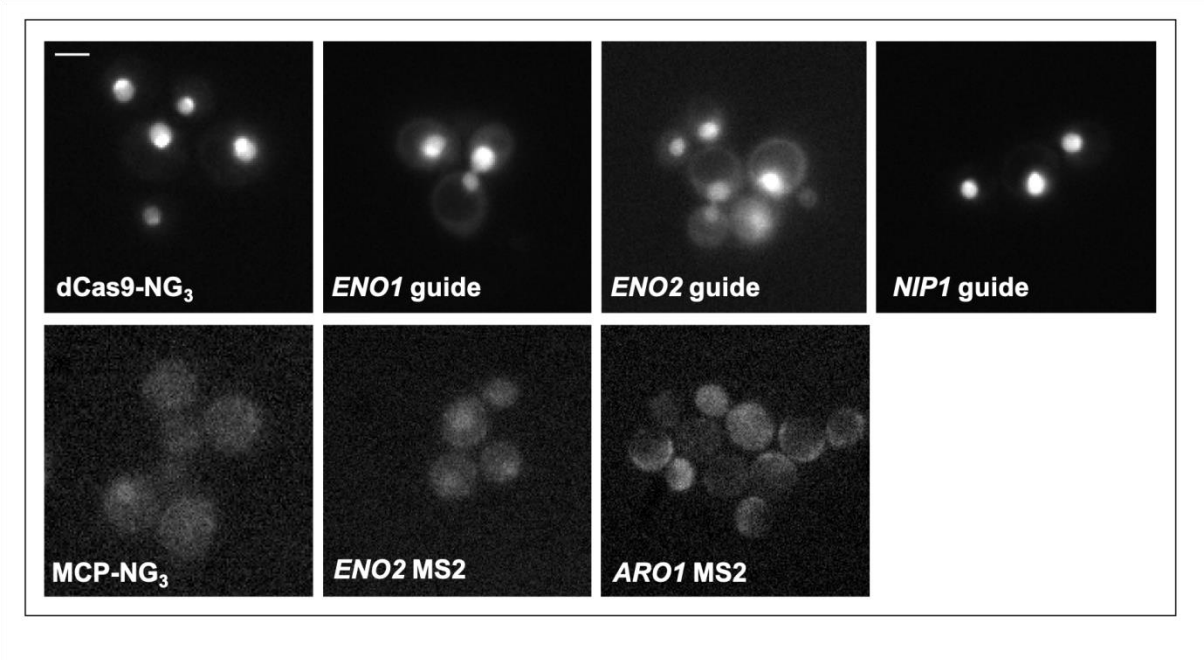


Figure 4.15: dCas9-NG_3 and MCP-NG_3 targeting to mRNAs. The top panel shows yeast cells expressing (A) dCas9-NG_3 only, and (B) cells expressing both dCas9-NG_3 and a sgRNA targeting *ENO1*, *ENO2*, or *NIP1* mRNAs. The bottom panel shows yeast cells expressing (A) MCP-NG_3 only, and (B) cells expressing MCP-NG_3 that contain genomically MS2-tagged *ENO2* mRNA or *ARO1* mRNA. No cytosolic mRNA localisation can be seen in these images, while cells containing dCas9-NG_3 and some cells containing MCP-GFP_3 appear to show different levels of nuclear localisation. Scale bar: 3μm.

4.4. Conclusions.

In this chapter, the dCas9 RNA-targeting system was adapted for use in yeast, where it seems to correctly target mRNAs to allow their localisation to be visualised. The mRNA localisation profiles obtained are similar to those observed using both smFISH and the MS2 system. This could be an important step forward, as it would represent the first time that unmodified mRNA has been visualised in live yeast cells. However, the signal from this dCas9-GFP_3 system is not intense enough to allow quantification of mRNA granules. This difficulty also precludes studies on mRNA granule dynamics, since extended exposures are required for mRNA granule visualisation, and from observations using the MS2 system, the granules likely move during this time.

The reduced sensitivity of the dCas9 system that is described above is not surprising if the number of GFP molecules associated with each mRNA is considered. For the dCas9 system, the fusion protein has three yeast enhanced GFP molecules and there is just a single fusion protein targeted to each mRNA. In contrast for the MS2 system, there are 12 stem loops inserted into the 3'UTR of the mRNA and each stem loop interacts with two coat proteins (Valegard et al., 1990; Peabody and Lim, 1996). Each coat protein has three enhanced GFP molecules fused to it. This gives a maximum occupancy of 72 GFP molecules. Hence, even though it has been established for multi-stem loop MS2 systems that not every stem loop is bound by a coat protein molecule (Rolfsson et al., 2016), there will still be substantially higher levels of GFP associated with the MS2 system relative to the dCas9 system.

An initial attempt to increase the brightness of the dCas9-GFP_3 system was to replace the GFP_3 with $mNeonGreen_3$. mNeonGreen has been suggested to fluoresce at least three times brighter than GFP using GFP filters (Shaner et al., 2013; Hostettler et al., 2017). However, fusion proteins made here using mNeonGreen did not recapitulate the mRNA localisation patterns previously seen for four different mRNAs tested. These mNeonGreen constructs showed no evidence of cytosolic mRNA granules previously described using dCas9-GFP_3 and MCP-GFP_3 targeting proteins, possibly suggesting the fusion of mNeonGreen to these proteins interfered with their ability to recognise and/or bind their targets. Indeed, it has been previously shown that although mNeonGreen itself is brighter than GFP, when fused to other proteins mNeonGreen may fluoresce less brightly and may affect protein function (Steiert et al., 2018; Ye et al., 2017). Given that MCP-NG_3 expressing cells also showed some evidence for nuclear localisation in the absence of an NLS, it is also possible that mNeonGreen contains a hidden or weak nuclear localisation signal that causes nuclear localisation and prevents the targeting of cytosolic mRNAs. Therefore the next aim

of this project, described in the next chapter, was to explore different approaches to improve the sensitivity of the dCas9 system to improve its ease of use, to make the results found more reliable, and to allow quantification.

Rather than directly modifying the structure of the mRNA, targeting of dCas9 to the RNA relies upon hybridisation of 20 bp of the sgRNA to the 3'UTR of the mRNA. Therefore, although the dCas9 initially targets endogenous mRNAs, the system could potentially affect mRNA dynamics due to the hybridisation of the targeting region of the sgRNA and/or by the attachment of the relatively large dCas9 GFP fusion protein. The hybridisation of the sgRNA to regions of the 3'UTR may prevent the binding of other, native RNA binding partners to the mRNA sequence, which may affect mRNA regulation and rates of degradation. The attachment of the dCas9 GFP fusion protein also has the potential to occlude other RNA binding partners, with the same drawbacks. The localisation of the mRNAs tested thus far appears to be unchanged, as these localisation patterns are the same as have been seen using both the MS2 system and smFISH (although these techniques also may affect mRNA dynamics). Nevertheless, control experiments including quantification of both protein and mRNA levels of the targeted mRNAs would be beneficial to investigate the effect of the dCas9 system on mRNA dynamics, and potential limitations should be kept in mind when using any of these visualisation systems.

As stated above, the main advantage of this dCas9 system is that it targets unmodified mRNAs, and so should not affect mRNA decay or translation rates as has been suggested for the MS2 system. Although this orthogonal approach to studying mRNA localisation provided similar results to the MS2 system and therefore increases their reliability, more controls could have been included here. In particular, the impact of dCas9 targeting on the levels of the targeted mRNAs could have been investigated, as a common criticism of the MS2 system is that the presence of stem loops prevents mRNA decay and causes an accumulation of mRNA decay fragments (which in turn may be visualised and affect localisation experiments). In the interest of time, qRT-PCR was not performed on the dCas9-sgRNA yeast strains targeting *NIP1*, *TIF1*, *ENO1* and *ENO2* mRNAs made here, as this was not the final version of the dCas9 system; instead qRT-PCR results investigating whether dCas9 targeting affects mRNA levels can be found in the next chapter for the improved version of the dCas9 system.

Results

Chapter 5: Development of a dCas9-MoonTag RNA-targeting system in yeast.

5. Development of a dCas9-MoonTag RNA-targeting system in yeast.

In the previous chapter, a dCas9 RNA-targeting system was developed for use in yeast. Although this system seemed to have successfully recapitulated mRNA localisation results found previously using the MS2 system, the signal from this dCas9 system was weaker than that obtained for either the MS2 system or smFISH. This is likely a result of each mRNA being targeted by a single dCas9 protein fused to three yeast enhanced GFPs in the dCas9 system, whereas in the MS2 system up to 72 GFPs might be targeted to each mRNA (Valegard et al., 1990; Peabody and Lim, 1996). To attempt to improve the brightness of the dCas9 system, the three GFPs were swapped with three mNeonGreen cassettes. Although this dCas9-mNeonGreen$_3$ appeared brighter when expressed alone, the fusion protein was unusable for visualising mRNA localisation in the cytosol as it remained in the nucleus under all conditions tested.

One method to improve the brightness in the system would be to include the targeting of multiple sgRNAs (and therefore multiple dCas9-GFP fusions) to an mRNA of interest. However, to limit the number of sgRNAs bound to the mRNA, and to make the system technically easier to use, an alternative strategy to increase the fluorescence of the system was pursued. This strategy involved the addition of a MoonTag peptide array to the dCas9 protein (Tanenbaum et al., 2014; Boersma et al., 2019), rather than a GFP fusion (Figure 5.1). The MoonTag array is a short repeating peptide derived from a viral protein gp41. In addition to the dCas9-MoonTag (dCas9-MT), a cassette is used which carries the gene for a single chain antibody (nanobody) against the MoonTag peptide fused to a single GFP (Figure 5.1). For the 24x MoonTag peptide array used here, it has been found that 10-12 nanobody-GFP molecules may bind (Boersma et al., 2019), and so this dCas9-MT system might be up to 4 times brighter than the previous dCas9-GFP$_3$ system.

Therefore, dCas9-MT system was used here to successfully visualise mRNA localisation for mRNAs previously studied using the MS2 system and smFISH. The dCas9-MT system is much brighter than the previous dCas9-GFP$_3$ and is hence more viable for the visualisation of mRNA. The patterns of cytosolic RNA granules found previously using the MS2 system and smFISH were recapitulated using this dCas9-MT system, and the abundance of these mRNAs was unchanged by dCas9 targeting. The dCas9-MT system was then used here to screen for novel mRNA localisation to cytosolic granules that had not previously been investigated. The dCas9-MT was used to study ergosterol biosynthesis mRNAs (*ERG1*, *ERG2*, *ERG4*, *ERG10* and *ERG20*), where evidence of cytosolic granule localisation was found for *ERG2*, *ERG4* and *ERG10* mRNAs. The dCas9-MT system was also used to investigate highly expressed mRNAs (*RPL28*, *CCW12*, *RPS31*, *RPL3* and *ASC1*) to explore

whether mRNA abundance has an effect on cytosolic RNA granules. In this case, it was found that these mRNAs did not appear to localise to cytosolic granules. Overall the dCas9-MT system appears to accurately report the localisation of mRNAs in yeast, and may therefore be a powerful tool for future use in studying endogenous, unmodified mRNA localisation in live cells and for screening novel mRNA localisation.

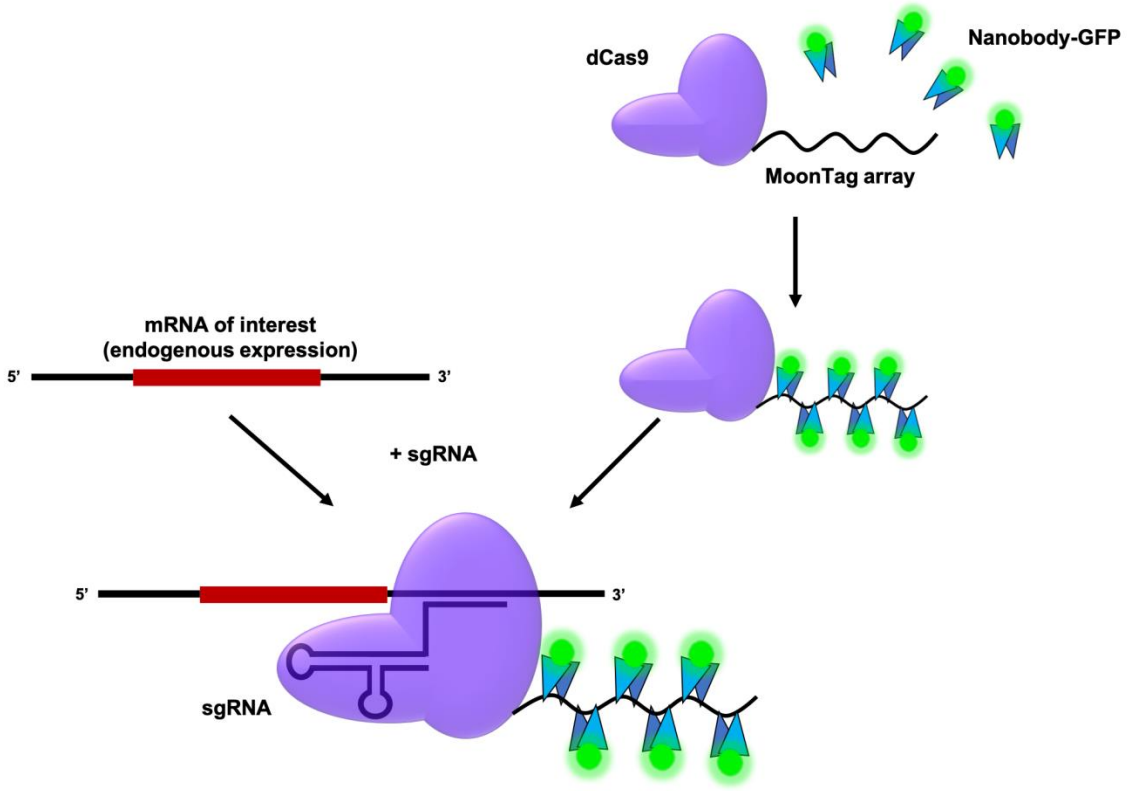


Figure 5.1: The dCas9 RNA-targeting system using a dCas9-MoonTag fusion protein. The dCas9 protein is fused to a 24x gp41 peptide array, called a MoonTag array. The peptides are recognised by anti-gp41 fragment antibodies (nanobodies), each of which is fused to GFP. The dCas9-MT construct is targeted to the mRNA of interest via a specific sgRNA, which contains a 20 bp recognition region for the mRNA and a stem-loop scaffold region that is recognised by dCas9. The targeting of multiple nanobody-GFP molecules to the mRNA of interest via dCas9 allows visualisation of mRNA localisation.

5.1. Construction of the dCas9-MoonTag RNA-targeting system in yeast:

To allow targeting and visualisation of mRNA in live cells, the dCas9-MT system requires three elements to be co-expressed; the sgRNA, the dCas9-MT protein, and the nanobody-GFP molecules. All sgRNAs were made by Gibson cloning as described in the previous chapter (Figures 4.5 and 4.6). The dCas9-MT plasmid was also produced by Gibson cloning, where three GFPs from the previously made dCas9-GFP_3 plasmid were replaced by a 24x MoonTag peptide array from a mammalian vector (Figure 5.2). The resulting dCas9-MT plasmid therefore consists of the dCas9 ORF followed by an NLS and the MoonTag array, with a *HIS3* selection marker to allow selection of transformants in yeast. This dCas9-MT construct was verified with a restriction digest and DNA sequencing (Figure 5.3).

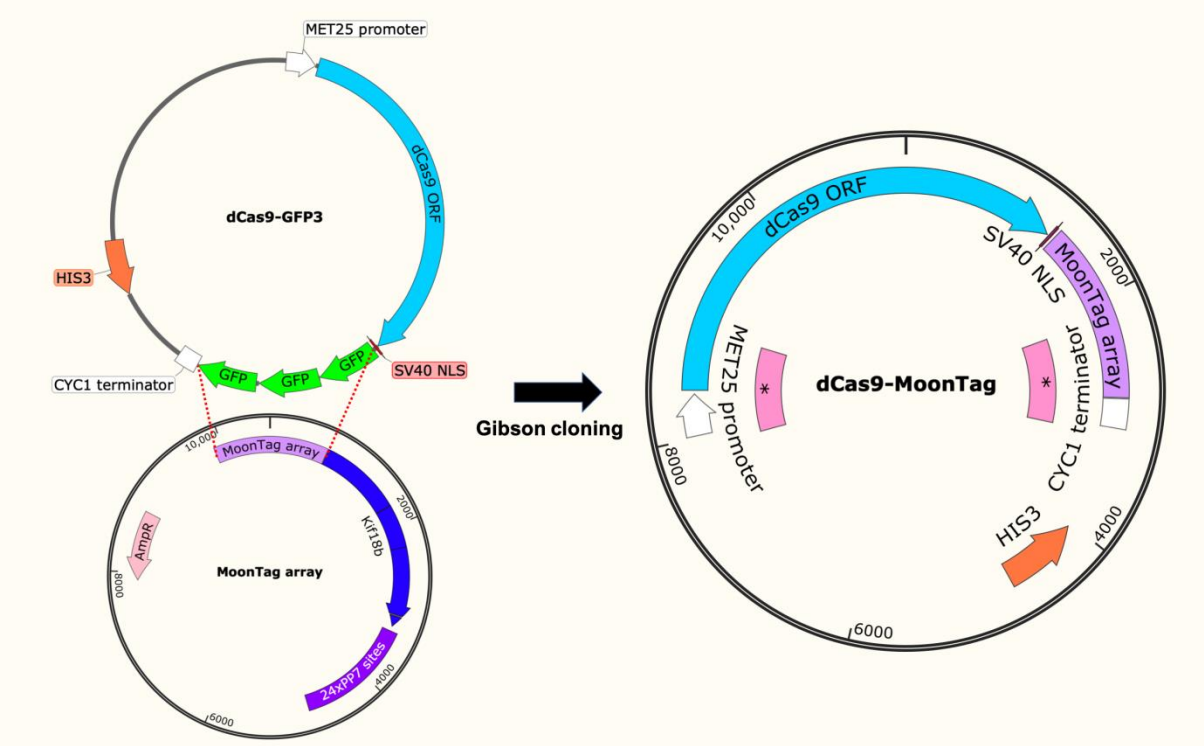


Figure 5.2: Gibson assembly of the dCas9-MoonTag plasmid. The three GFP molecules from a dCas9-GFP_3 vector were replaced with a 24x MoonTag array. The resulting dCas9-MT vector therefore contains the dCas9 ORF, NLS, and MoonTag array. The vector also includes a promoter, terminator, and *HIS3* marker to allow selection and expression in yeast cells. The regions denoted with an asterisk (highlighted in pink) are the junction regions verified by sequencing.

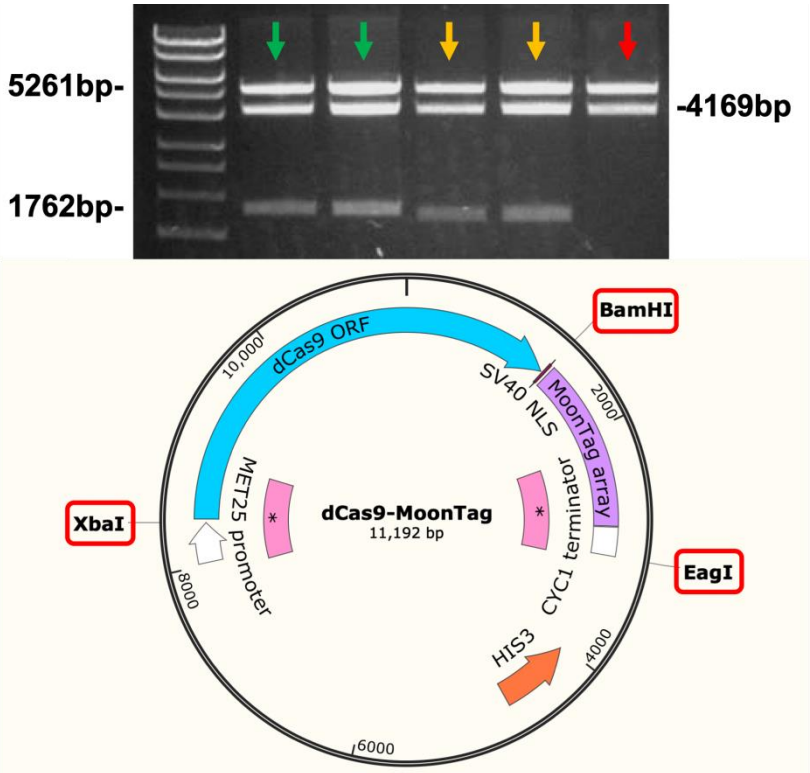


Figure 5.3: Restriction enzyme mapping of dCas9-MoonTag plasmid. All gels are shown with HyperLadder 1kb markers. The gel shows 5 possible dCas9-MT samples digested with *Bam*HI, *Eag*I and *Xba*I. All samples contain bands at 5261bp and 4168bp that derive from the plasmid backbone and the dCas9 ORF, respectively. Sample 5 (red arrow) only shows these two bands and is therefore incorrect. The other samples show a band around 1700bp, indicating that the MoonTag array is present (and not a band at 2470bp which would indicate the presence of 3x GFP). All samples were sent for sequencing, where the sequenced regions are shown in pink with an asterisk on the plasmid representation. Samples 1 and 2 (green arrows) were correct, while samples 3 and 4 (orange arrows) contained truncated MoonTag peptide arrays and not used.

To increase the ease of use of the dCas9-MT system, the nanobody-GFP construct was designed to be integrated into the yeast genome (Figure 5.4). The anti-gp41 nanobody sequence was kindly provided by the Tanenbaum lab (Boersma et al., 2019), which was then codon optimised here for efficient expression in yeast. The nanobody-GFP construct was synthesised by Epoch Life Science into a standard cloning vector (pBluescript SK), containing sequences for the nanobody-GFP fusion and a hygromycin resistance gene for selection, both with appropriate promoters and terminators for yeast (Figure 5.4). These sequences were flanked by yeast genomic regions to allow integration into the yeast genome near the *TRP1* locus (Chromosome IV 463001). The synthesised plasmid was amplified in bacteria to acquire large amounts of nanobody-GFP DNA. The plasmid samples were verified by restriction digest mapping and DNA sequencing (Figure 5.5). The verified plasmid was digested with *Bsp*QI to release the cassette (containing the nanobody-GFP and hygromycin resistance sequences and associated promoters and terminators) and then transformed into the yeast strain yMK467 (Figure 5.4).

Correct genomic integration of this cassette was verified using a PCR-based strategy on genomic DNA prepared from potential transformants (Figure 5.6). In the verification PCR, three fragments were expected to be produced from the nanobody-GFP integrated cells – a 4245 bp fragment spanning the entire region and two fragments at 849 bp and 3416 bp spanning shorter regions (Figure 5.6). Although the longest 4245 bp PCR failed (likely due to sub-optimal PCR conditions), the two shorter bands were present, indicating the presence of the nanobody-GFP cassette. Importantly, the same PCRs carried out on the control parental strain yMK467 only returned a band at 800 bp, corresponding to the genomic region if the nanobody-GFP insertion was not present. Neither PCR performed using hygromycin primers were successful (although they were successful for the nanobody-GFP integrated strain), as the parental strain did not contain hygroymcin (Figure 5.6).

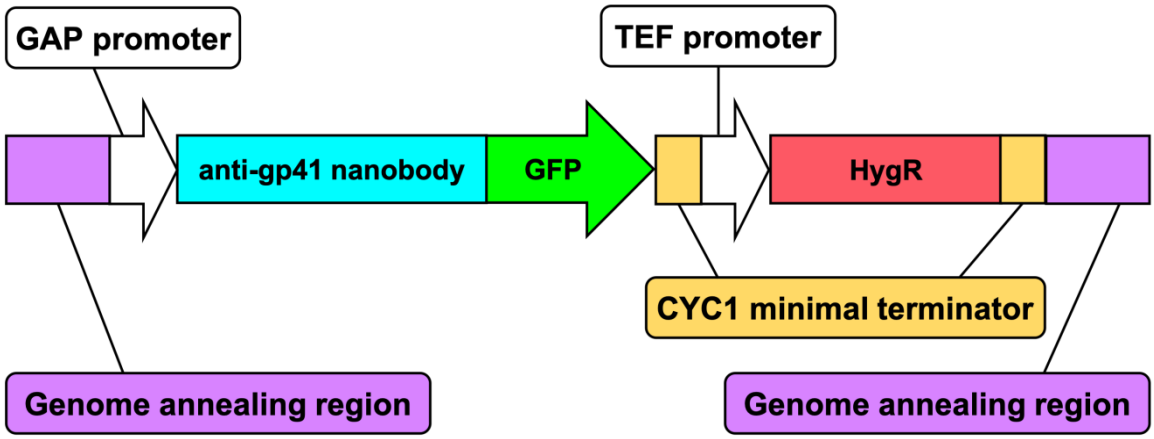


Figure 5.4: The nanobody-GFP construct for yeast genomic integration. The vector was designed to contain both the anti-gp41 nanobody sequence followed by a yeGFP sequence, and a hygromycin resistance (HygR) gene. Both sequences were flanked with promoters and terminators to allow expression in yeast. This entire cassette was flanked by genome annealing regions complementary to a region downstream of the *TRP1* locus to allow integration into the yeast genome. Sequences were also flanked by specific *Bsp*QI restriction sites to allow fragment release (not shown in figure).

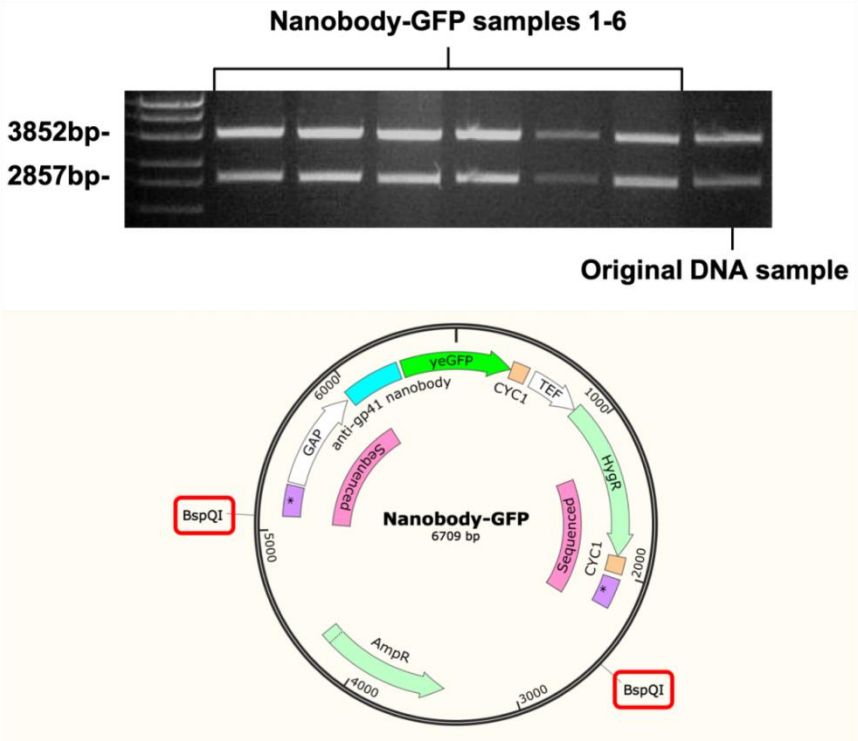


Figure 5.5: Restriction enzyme mapping of the nanobody-GFP plasmid. As shown in the gel, all samples produced bands at 3852bp and 2857bp, as expected from the *Bsp*QI digest. The final sample shows the same digest of the original nanobody-GFP DNA synthesised by Epoch Life Science as a positive control. Samples 1 and 2 were sent for Sanger sequencing and were verified as correct, where the sequenced regions are shown highlighted in pink on the plasmid representation.

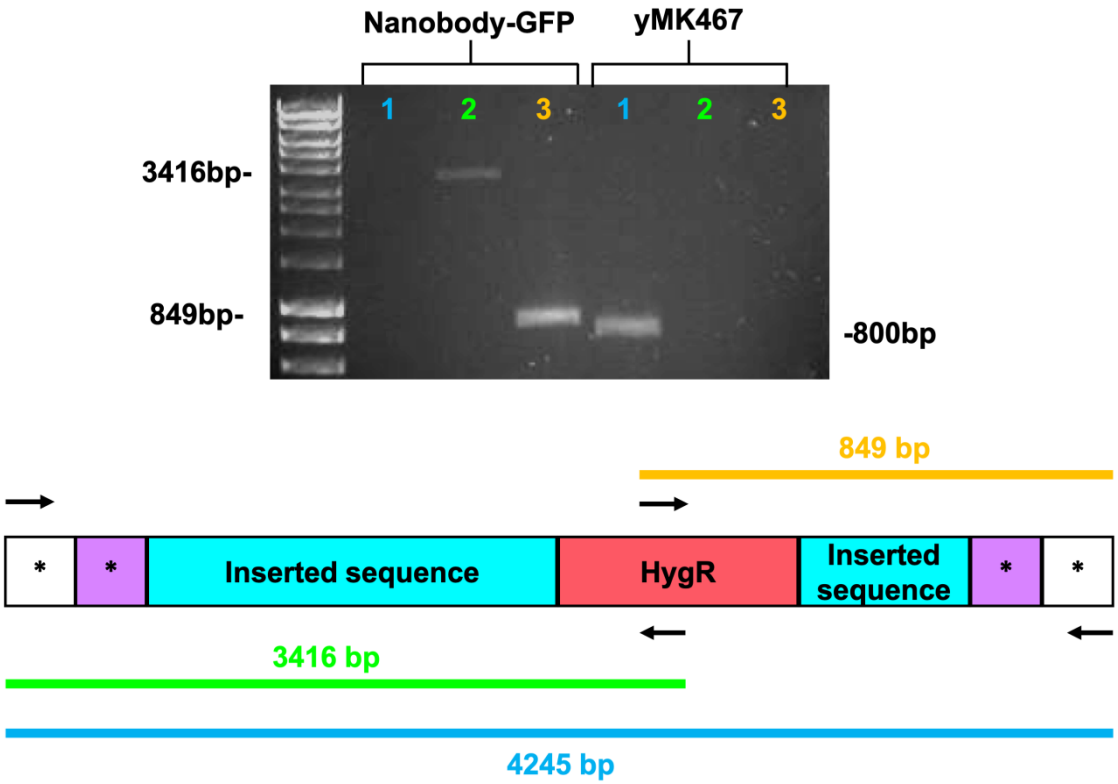


Figure 5.6: PCR verification of nanobody-GFP genomic integration. Primers used for PCR verification are shown as arrows on the sequence representation. Genomic regions are denoted by asterisks and are 200bp each, where regions highlighted in purple are the genomic annealing region included in the nanobody-GFP insertion fragment, and regions highlighted in white are flanking genomic sequences. On the gel, nanobody-GFP denotes the cells transformed with the nanobody-GFP fragment, while the untransformed cells are called yMK467. The gel shows that the PCR was successful for the two smaller regions (in green and orange) in the integrated cells, although the larger 4245bp band (in blue) was not produced. In yMK467, neither PCRs using the HygR primers produced bands, whereas the genomic binding primers produced a band at 800bp, indicating the insertion fragment was not present.

5.2. Development of the dCas9-MoonTag RNA-targeting system in yeast:

To allow targeting and visualisation of mRNA in yeast, the nanobody-GFP construct was integrated into the yeast genome and the dCas9-MT plasmid was then transformed into these cells. For each mRNA targeted, the specific sgRNA vector was then transformed into cells expressing both the nanobody-GFP and dCas9-MT constructs. The expression of the nanobody-GFP and dCas9-MT constructs was verified with Western blotting (Figure 4.9, replicated here as Figure 5.7 for ease of reading). Antibodies raised against GFP and dCas9 detected bands at approximately 58 and 222 kDa consistent with the successful expression of these proteins. For the nanobody-GFP construct sample, some evidence of degradation is evident, although the majority of the protein appeared to be full length (Figure 5.7).

Yeast expressing the nanobody-GFP construct exhibited a diffuse GFP signal throughout cells (Figure 5.8A). When dCas9-MT, which contains an NLS, was also expressed, the GFP signal was then mostly confined to the nucleus (Figures 5.8B and 5.8C). This is consistent with the nanobody interacting with the nuclear localised dCas9-MT protein. The dCas9-MT/ nanobody GFP signal is much brighter than either of the previous dCas9-GFP_3 and dCas9-NG_3 systems, as shown by images taken at the same exposure times (Figure 5.9). This means the system holds promise for the sensitive detection of mRNA.

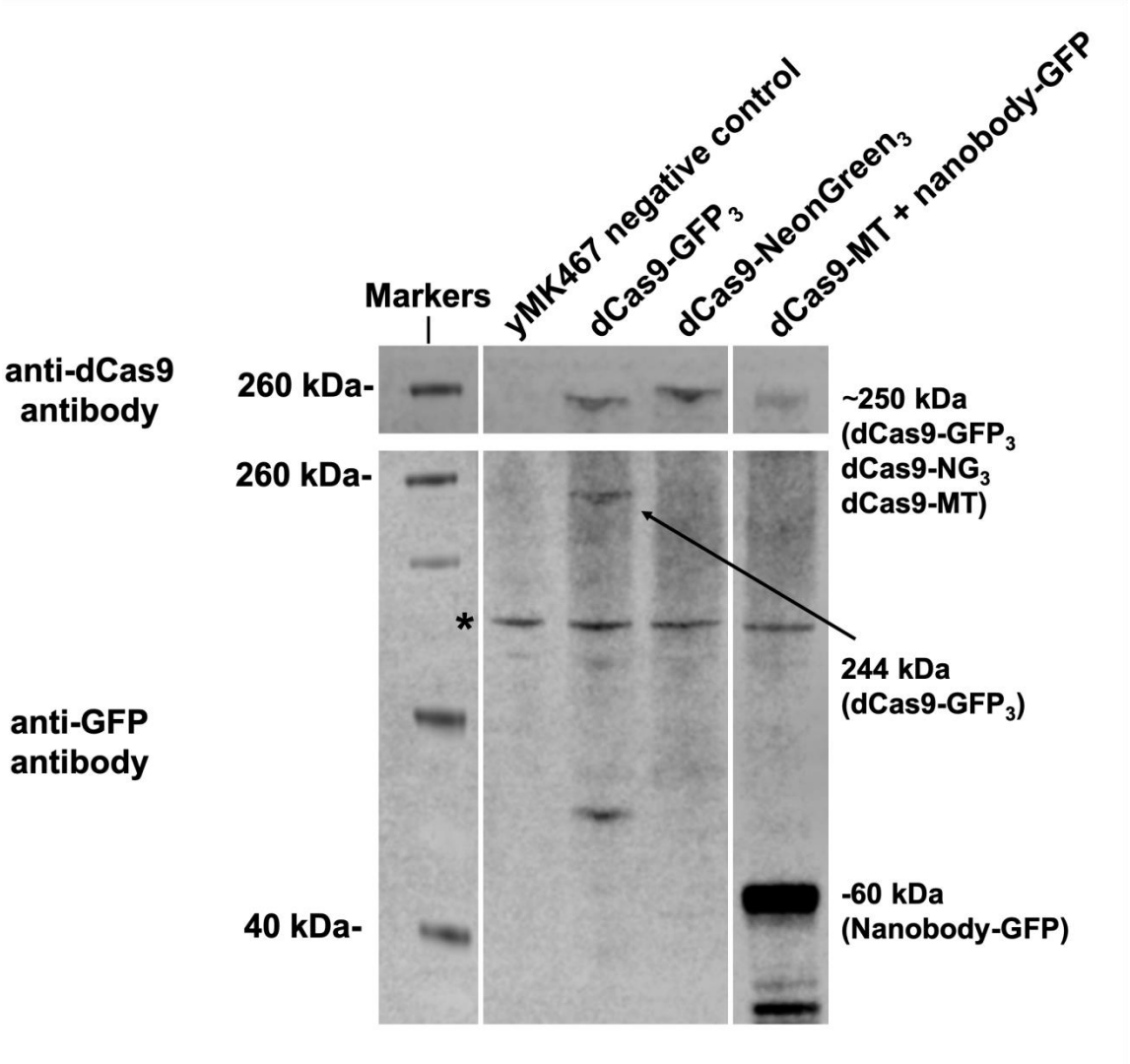


Figure 5.7: Western blots of nanobody-GFP and dCas9-MoonTag. This figure is replicated from the previous chapter (Figure 4.9). Lane 1 contains protein extracts from the parent yeast strain yMK467 as a negative control. Lanes 2 and 3 contain protein extracts from dCas9-GFP_3 and dCas9-NG_3 expressing strains, respectively, as discussed in the previous results chapter. Lane 4 contains proteins extracted from strains expressing the nanobody-GFP and dCas9-MT constructs. Using an anti-dCas9 antibody (top blot), a band just below 260 kDa is shown, corresponding to the dCas9-MT construct (dCas9 is 160 kDa and the MoonTag is estimated to be around 90kDa). Using an anti-GFP antibody (bottom blot), a large band is present above 40 kDa, corresponding to the nanobody-GFP construct (GFP is 28kDa and the nanobody is estimated to be around 30kDa), and a possible GFP only band is shown at 28 kDa. The anti-GFP antibody also appears to produce a non-specific band present in all samples (shown by the asterisk), and provides a loading control showing the samples were evenly loaded. Irrelevant lanes on the blot have been removed and replaced by white lines to separate distinct regions of the same images

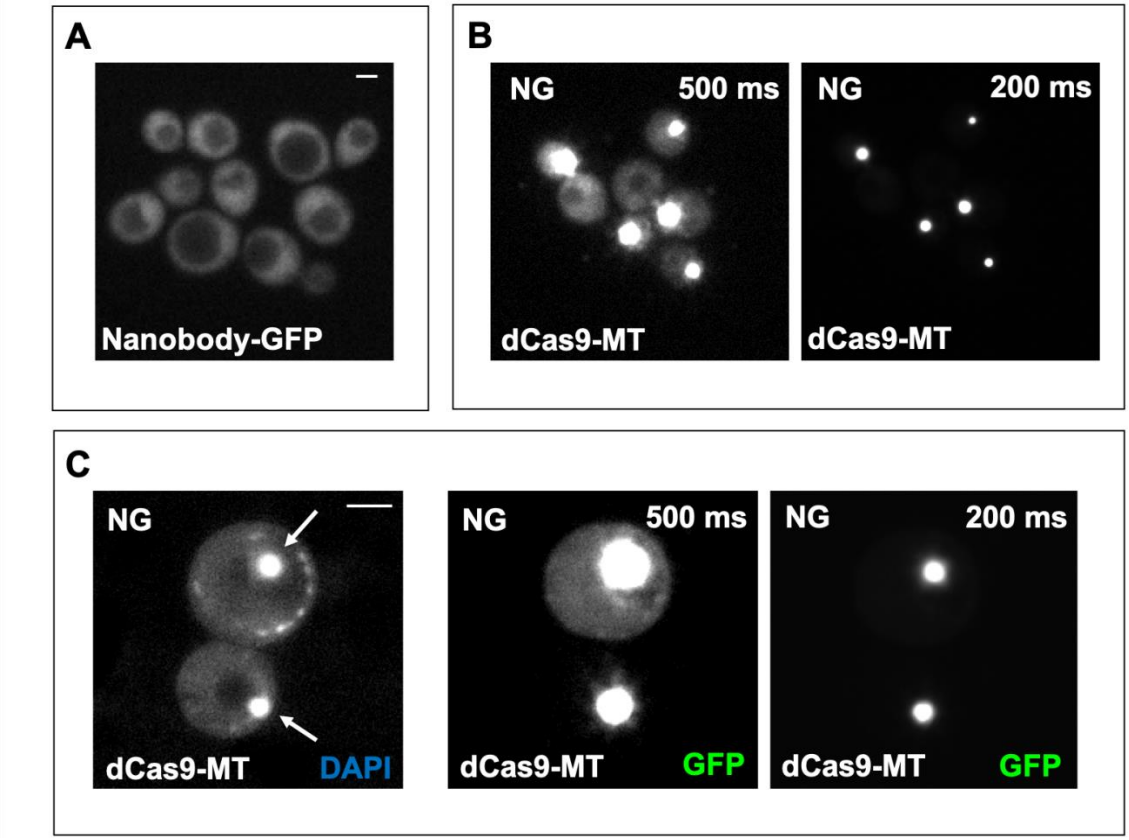


Figure 5.8: Nanobody-GFP and dCas9-MoonTag images in yeast. Yeast in panel A are expressing the nanobody-GFP construct alone, while in panel B cells are expressing both the nanobody-GFP (denoted by NG) and dCas9-MT constructs. Panel B shows the same single image at different exposure levels to allow cells to be shown, as the GFP is very bright. In panel C, cells are expressing both the nanobody-GFP and dCas9-MT constructs, DAPI and GFP images are shown with two different exposure levels for the GFP images. White arrows indicate nuclear DAPI signal. Images were taken as Z-stacks and the average intensity projection is shown. Scale bars: 2μm.

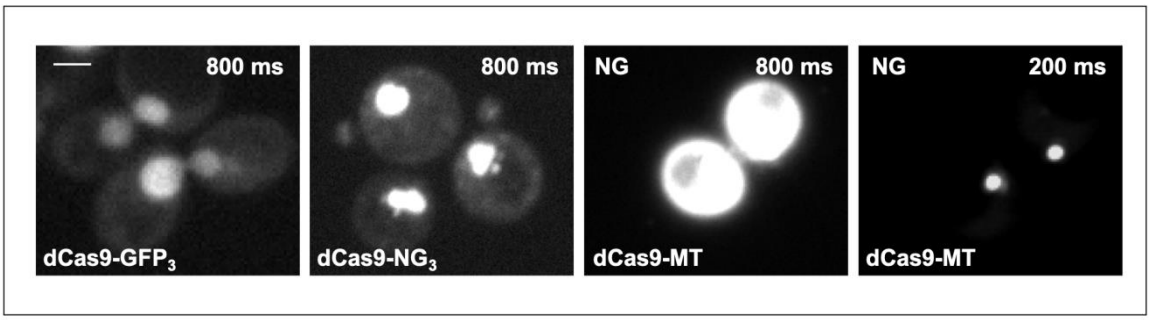


Figure 5.9: Comparison of dCas9-GFP$_3$, dCas9-NG$_3$ and dCas9-MT systems. Yeast cells expressing dCas9-GFP$_3$, dCas9-NG$_3$ or dCas9-MT (plus nanobody-GFP). Each image was taken at the same exposure time (800 ms) to compare relative brightness levels, and the dCas9-MT image is also shown at a shorter exposure time (200 ms). Images were taken of a single plane to prevent bleaching. Scale bar: 2μm.

5.3. Visualising mRNA localisation using the dCas9-MoonTag RNA-targeting system in yeast:

To evaluate the potential of this system for the study of mRNA localisation, yeast expressing both the nanobody-GFP and dCas9-MT constructs were transformed with a sgRNA specific to a targeted mRNA. The expression of these sgRNAs was verified using qRT-PCR, using primers that amplified the conserved scaffold region of the sgRNA sequence. qRT-PCR showed that the sgRNAs were expressed, where non-targeting, *NIP1*, *TIF1*, *ENO1* and *ENO2* sgRNAs were shown to be expressed in their respective cells compared to cells only expressing the nanobody-GFP and dCas9-MT constructs as a control (using *ACT1* as a reference gene) (Figure 5.10). As discussed in the previous chapter, two sgRNAs were used as controls – a non-targeting sgRNA designed to not recognise any yeast mRNAs, and a sgRNA targeting *NPC2* mRNA, which has been shown to be non-localised and found throughout the cytosol. The nucleus could still be seen in these cells (Figure 5.11); nevertheless, both of these controls showed no evidence of cytosolic RNA granules using either sgRNA (Figure 5.11).

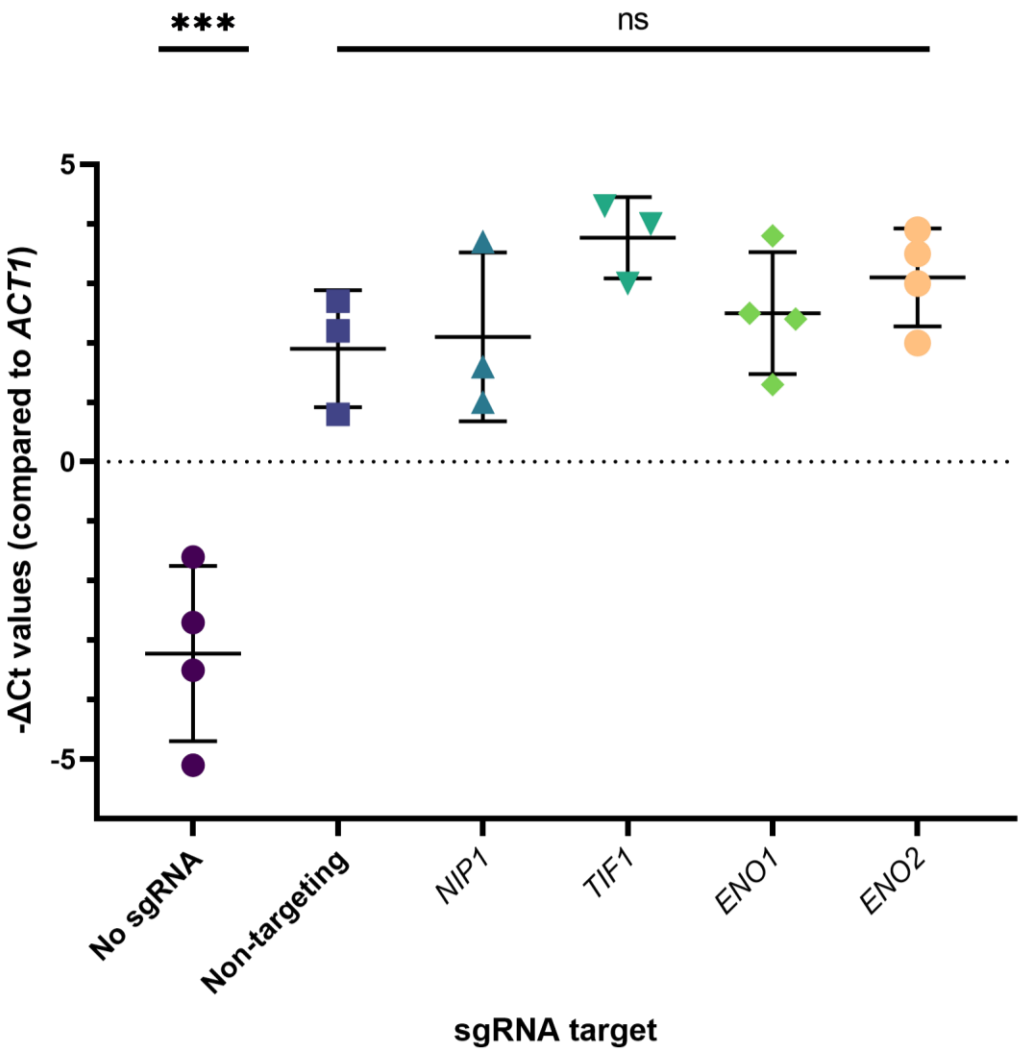


Figure 5.10: qRT-PCR of sgRNAs. Yeast cells were then tested for the presence of the appropriate sgRNA; a non-targeting sgRNA or a sgRNA targeting *NIP1*, *TIF1*, *ENO1* or *ENO2*. *ACT1* was used a reference gene for each sample, and the presence of the appropriate sgRNA was tested for in each case in comparison with yeast cells expressing the nanobody-GFP and dCas9-MT constructs (and no sgRNAs). Shown on the graph are –ΔCt values, calculated as (-) target gene – reference gene (*ACT1*) in each case. For each sample n = 3 or 4 and error bars are shown as the mean with standard deviation. One-way ANOVA statistical analysis showed no significant differences between any of the samples containing sgRNAs (all samples except 'No sgRNA'), and significant differences between the 'No sgRNA' and each of the other samples (P value range 0.0003 to <0.0001).

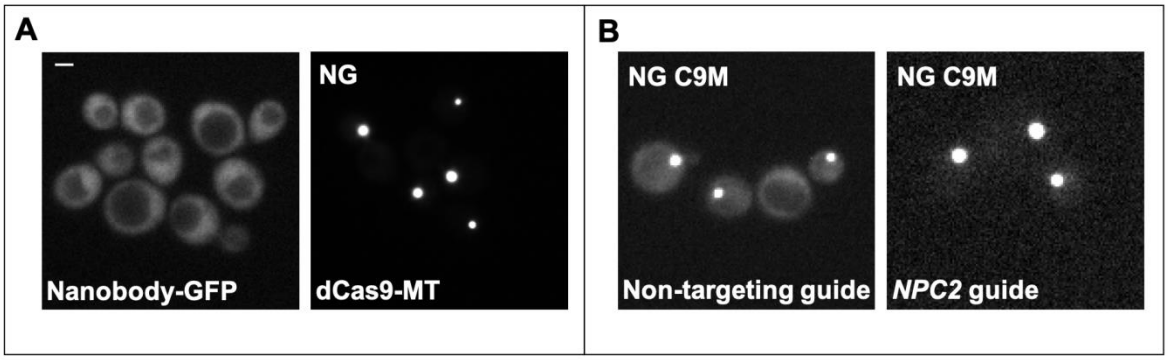


Figure 5.11: Controls sgRNAs for the dCas9-MT system. In all panels, NG denotes the presence of the nanobody-GFP construct and C9M denotes the presence of the dCas9-MoonTag plasmid. Panel A shows cells expressing the nanobody-GFP construct only or cells expressing both the dCas9-MoonTag and nanobody-GFP constructs. Panel B shows yeast cells expressing the nanobody-GFP and dCas9-MT constructs, along with either a non-targeting sgRNA or a sgRNA targeting *NPC2* mRNA. Images were taken as Z-stacks and the maximum intensity projection is shown. Scale bar: 2μm.

The dCas9-MT system was then used to visualise mRNA localisation for *ENO1*, *ENO2*, *NIP1* and *TIF1* mRNAs, which were previously visualised with the dCas9-GFP_3 system (Figure 4.10). *ENO1* and *ENO2* mRNAs show a localisation pattern of many granules per cell, while *NIP1* and *TIF1* mRNAs are generally present in one or two granules per cell (Figure 3.1 and Figure 5.12 middle panel) (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). While images taken with the dCas9-MT system often show the nucleus of the cells (as with the previous control sgRNAs), the system also recapitulates the pattern of mRNA localisation to cytosolic granules as shown by the dCas9-GFP_3 and MS2 systems (Figure 5.12). Accordingly, *ENO1* and *ENO2* are present in multiple cytosolic granules per cell while *NIP1* and *TIF1* are present in one or two granules per cell. The dCas9-MT system appears brighter than the dCas9-GFP_3 system and seems to show comparable brightness to the MS2 system, where distinct granules are easily visible in the cytosol (Figure 5.12). Critically, using the dCas9-MT system cytosolic granules are now clearly visible for the *TIF1* mRNA whereas using the dCas9-GFP_3 system this was not the case (Figure 4.11).

To test whether or not the targeting of dCas9-MT to these mRNAs had an effect on the abundance of the mRNA, qRT-PCR was carried on each of these strains. qRT-PCR primers were designed to bind to the 3'UTR of targeted mRNAs, and mRNA levels were compared to either cells expressing no sgRNAs or cells expressing the non-targeting sgRNA as controls. qRT-PCR showed that mRNA targeting by dCas9-MT had little effect on the abundance of the mRNA, where all mRNAs tested showed little change compared to controls cells (Figure 5.13).

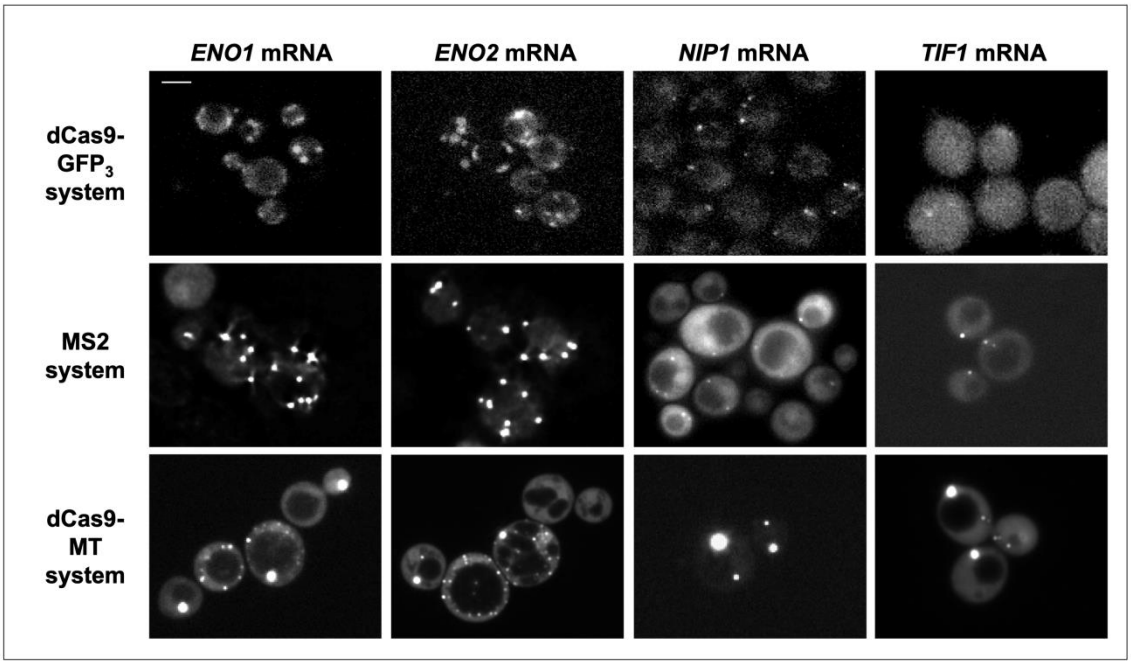


Figure 5.12: Comparison of dCas9-GFP$_3$, MS2 and dCas9-MT systems for visualising mRNA localisation. Each column shows the localisation pattern of *ENO1*, *ENO2*, *NIP1* or *TIF1* mRNAs using a different visualisation system. As described in previous chapter, the top row of images were taken with the dCas9-GFP$_3$ system and imaged as single plane images. Images taken with both the MS2 and dCas9-MT systems (shown in the middle row and bottom row, respectively) were taken as Z-stacks and the average projected intensity image is shown. Scale bar: 3μm.

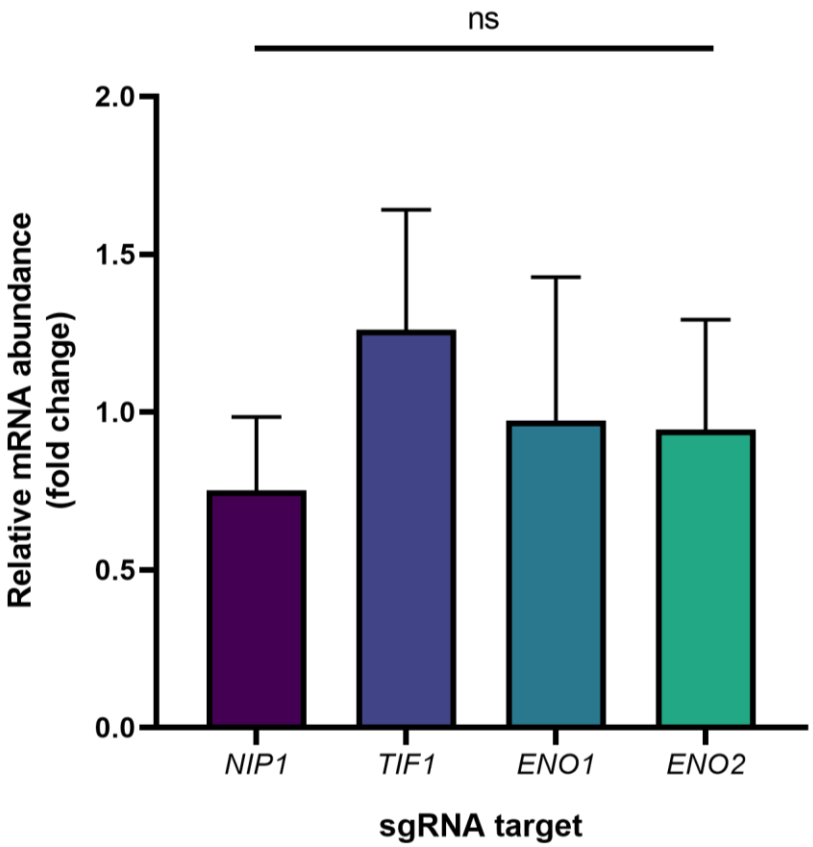


Figure 5.13: qRT-PCR of mRNA abundance. Yeast strains were tested for the mRNA abundance of the mRNA being targeted (e.g. *NIP1* mRNA for cells expressing the *NIP1* sgRNA). *ACT1* was used a reference gene for each sample and mRNA abundance was compared the mRNA abundance of the same gene in cells not expressing the sgRNA (e.g. *NIP1* mRNA levels were compared between cells expressing the nanobody-GFP and dCas9-MT constructs, and cells expressing both of these constructs plus the *NIP1* targeting sgRNA). *NIP1* mRNA appeared to be slightly lower in abundance in cells targeting the *NIP1* mRNA, *ENO1* and *ENO2* mRNA levels appeared to be around the same level of abundance when targeted (1-fold change), and *TIF1* mRNA appeared to be slightly more abundant in cells targeting *TIF1* mRNA. For *NIP1*, *ENO1* and *ENO2* mRNAs n=4 and for *TIF1* mRNA n=3, and error bars are shown as the mean with standard deviation. One-way ANOVA statistical analysis showed no significant differences between any of the mRNAs.

The dCas9-MT system was then used to image other mRNAs previously imaged using the MS2 system but not imaged using the dCas9-GFP_3 system (Figure 5.14 second column) (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). Again, the number of cytosolic granules found using the MS2 system for these mRNAs is recapitulated using the dCas9-MT system (Figure 5.14), demonstrating the visualisation system is accurately targeting mRNAs. As with the *ENO1* and *ENO2*, the glycolytic mRNAs *PDC1* and *FBA1* are present in multiple granules per cell (Figure 5.14) (Morales-Polanco et al., 2021). *PDC5*, another glycolytic mRNA, does not show cytosolic granules when imaged with the MS2 system, but does appear to form multiple granules per cell when imaged with this dCas9-MT system and smFISH (Figure 5.14) (Ashe lab, unpublished data). *SUP35* is a translation factor mRNA that is present in one or two granules per cell in both dCas9-MT and MS2 systems (Figure 5.14) (Pizzinga et al., 2019). Finally, *ARO3* is an aromatic amino acid biosynthesis mRNA that appears to localise to around 3-5 granules per cell in both the dCas9-MT and MS2 systems (Figure 5.14) (Ashe lab, unpublished data). Overall, the dCas9-MT system appears capable of correctly targeting mRNAs and recapitulating mRNA localisation data found previously using the MS2 and smFISH systems, and is much brighter than the previous dCas9-GFP_3 system.

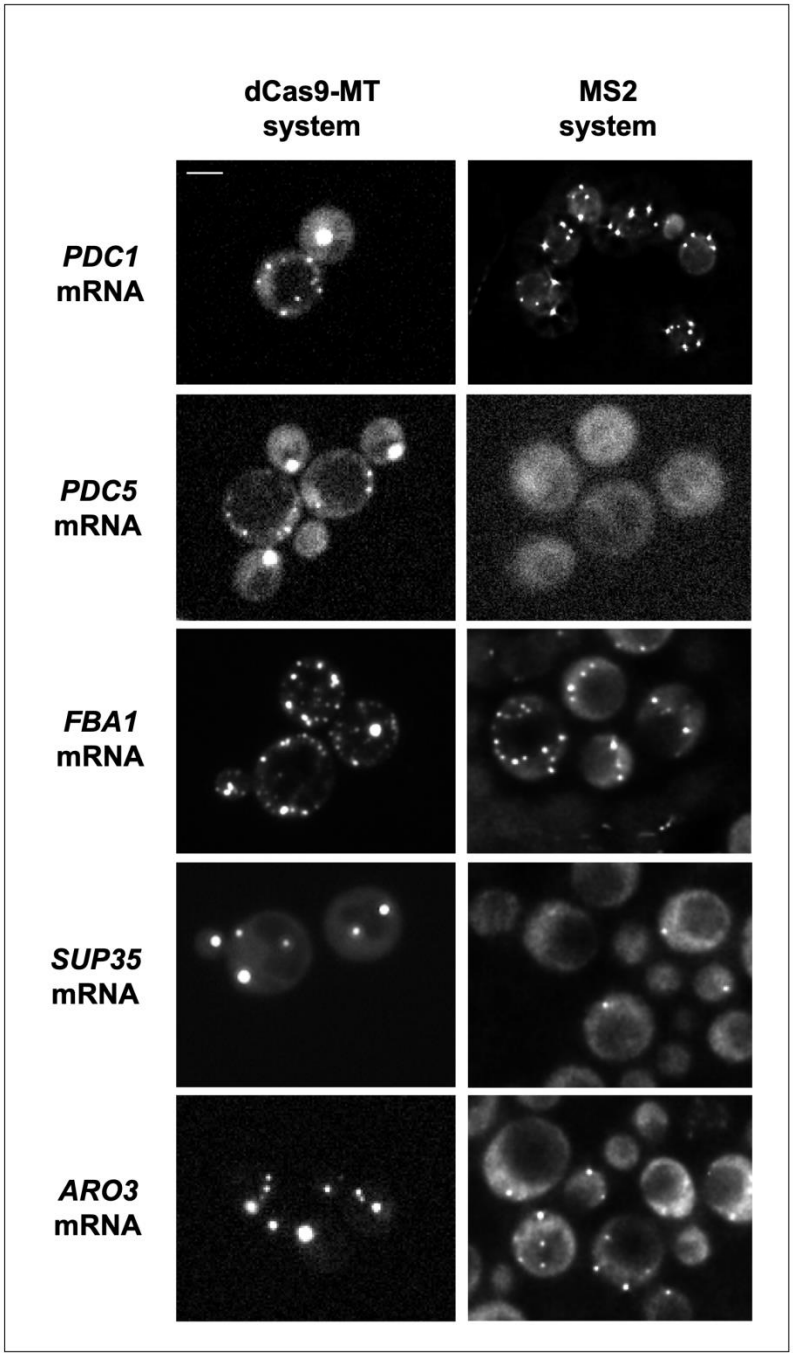


Figure 5.14: Comparison of the dCas9-MT and MS2 system targeting other mRNAs. Images were taken of the previously studied mRNAs listed down the side with either the dCas9-MT system (first column) or the MS2 system (second column). Images were taken as Z-stacks and the average projected intensity for each image is shown. The nucleus is visible in most cells using the dCas9-MT system, although mRNA cytosolic localisation patterns are generally similar in each system. Scale bar: 3μm.

5.4. Using the dCas9-MoonTag RNA-targeting system to screen mRNA localisation:

The three main advantages of the dCas9-MT system are that the system is capable of visualising unmodified mRNAs, the system can be used in live cells and that the method is quick and easy to use. Other methods of targeting mRNAs often involve complex strain construction to alter the mRNA structure (such as the MS2 system and aptamer systems), or relatively difficult protocols that involve cell fixation and/or permeabilisation (such as smFISH and aptamer systems). In contrast, once the cells expressing both the nanobody-GFP and dCas9-MT constructs are established, it is relatively straightforward to generate new sgRNAs and transform them into cells to study mRNA localisation.

In order to use the dCas9-MT system to visualise novel mRNA localisation, 10 new sgRNAs were produced to target mRNAs whose localisation has not been previously studied. These mRNAs were chosen to investigate two possible driving forces behind the patterns of mRNA localisation to cytosolic granules; highly expressed mRNAs and mRNAs encoding proteins of similar functions. Since glycolytic mRNAs (e.g. *ENO1*, *PDC1*) are expressed more highly than translation factor mRNAs (e.g. *NIP1*, *TIF1*), it is possible that this explains the abundance of the glycolytic mRNA granules. Therefore to investigate whether this 'abundant granule' pattern is repeated with other highly expressed mRNAs, the 5 most highly expressed genes in yeast, excluding glycolytic mRNAs, were targeted (*RPL28*, *CCW12*, *RPS31*, *RPL3* and *ASC1*) (Figure 5.15). The other 5 sgRNAs were made to target mRNAs encoding proteins involved in the ergosterol biosynthesis pathway (*ERG1*, *ERG2*, *ERG4*, *ERG10* and *ERG20*). Given that both glycolytic and translation factor granules contain actively translating mRNAs, it is possible that the purpose of the granules is to facilitate the cotranslation of proteins that have similar functions to increase the efficiency or allow coregulation of biochemical pathways. In more recent work, the Ashe lab have found evidence that mRNAs encoding proteins important in aromatic amino acid synthesis (*ARO1*, *ARO2*, *ARO3* and *ARO4*) also localise to cytosolic granules in yeast (unpublished data). To explore whether the coregulation of biochemical pathways might be a plausible rationale for presence of the mRNA granules, the ergosterol biosynthetic pathway was selected for initial studies.

Gene	Count	Gene	Count
FBA1	18588	*ERG10*	356
TDH3	16055		
PDC1	11952	*ERG2*	307
ENO2	9814		
PGK1	8884	*ARO2*	278
GPM1	7679		
RPL28	7546	*ERG4*	228
CDC19	5747		
ADH1	5525	*ERG20*	188
CCW12	4501		
RPS31	4240	*ARO1*	141
RPL3	3887		
ASC1	3612	*ARO3*	134
YEF3	1916	*SUP35*	77
ENO1	1824	*TIF1*	76
ARO4	373	*NIP1*	54
ERG1	373		

Figure 5.15: A list of relative mRNA abundance in yeast. Data is from Costello et al., 2015. Relative mRNA abundance in yeast calculated as the mean mRNA counts from immunoprecipitation data against Caf20, Eap1, eIF4E, eIF4G1, eIF4G2 and Eap1. Glycolytic mRNAs (highlighted in yellow) are the most abundant mRNAs in yeast. sgRNAs were designed to target the next most abundant non-glycolytic mRNAs (highlighted in blue), and ergosterol mRNAs (highlighted in pink). Aromatic amino acid biosynthesis mRNAs are highlighted in orange and translation factor mRNAs are highlighted in green. Spaces in the list show where other mRNAs have been removed from this dataset.

sgRNAs were made by Gibson cloning as described in the previous chapter (Figures 4.5 and 4.6). sgRNAs were then transformed into yeast cells expressing the nanobody-GFP and dCas9-MT constructs and imaged. For ergosterol mRNAs, images showed that *ERG2*, *ERG4*, and *ERG10* appeared to form cytosolic granules in unstressed yeast cells, while *ERG1* and *ERG20* did not appear to form granules under the same conditions (Figure 5.16). *ERG2*, *ERG4* and *ERG10* mRNAs appeared to form two distinct patterns of cytosolic mRNAs - either in many granules per cell (Figure 5.16 middle row) or in around 4-5 granules per cell (Figure 5.16 bottom row). Both imaging and quantification of these cells (n=100) showed that the most common pattern of cytosolic granules for these mRNAs was 4-5 granules per cell, and less than 10% of cells showed 10-20 granules per cell (Figure 5.17). Quantification of *ERG1* and *ERG20* mRNAs showed that generally cells did not appear to form any cytosolic granules, though rarely they showed one or two granules per cell (Figure 5.17). As with previous mRNAs, the nucleus could also be seen in most cells imaged.

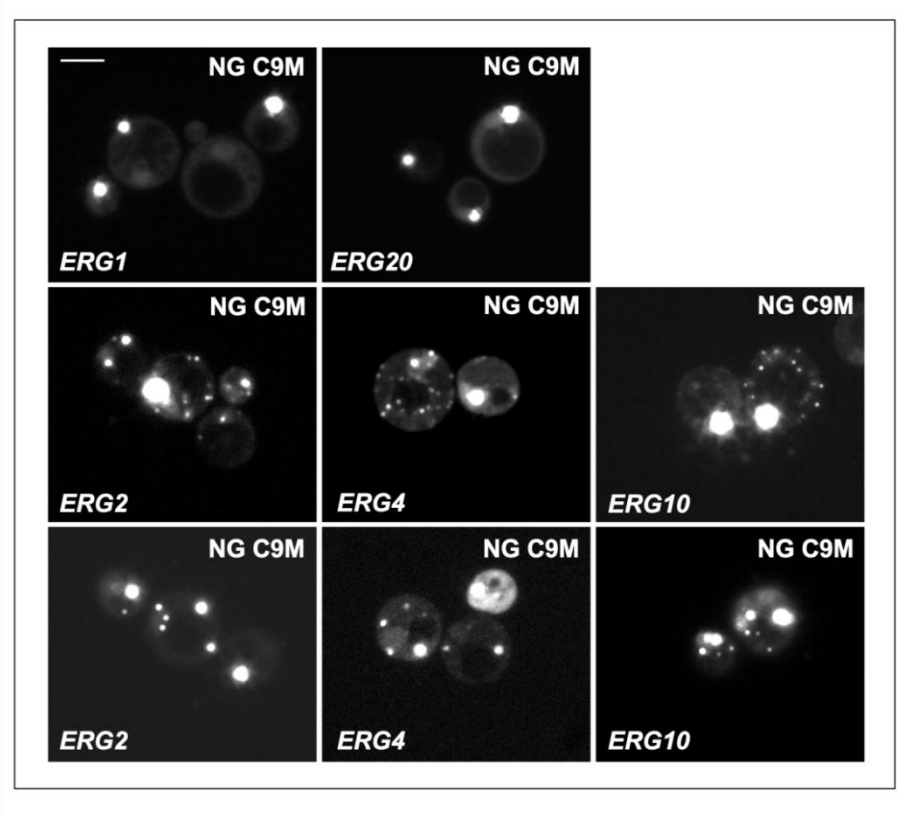


Figure 5.16: Ergosterol mRNA localisation imaged in yeast. All cells are expressing the nanobody-GFP and dCas9-MT constructs, along with a sgRNA for the mRNA being targeted (*ERG1*, *ERG2*, *ERG4*, *ERG10* or *ERG20*). *ERG1* and *ERG20* mRNAs (top row) do not appear to form cytosolic granules. Middle and bottom rows of images show the same three mRNAs (*ERG2*, *ERG4* and *ERG10*) appearing in either many or few granules per cell, respectively. Images were taken as Z-stacks and the maximum intensity image is shown. Scale bar: 3μm.

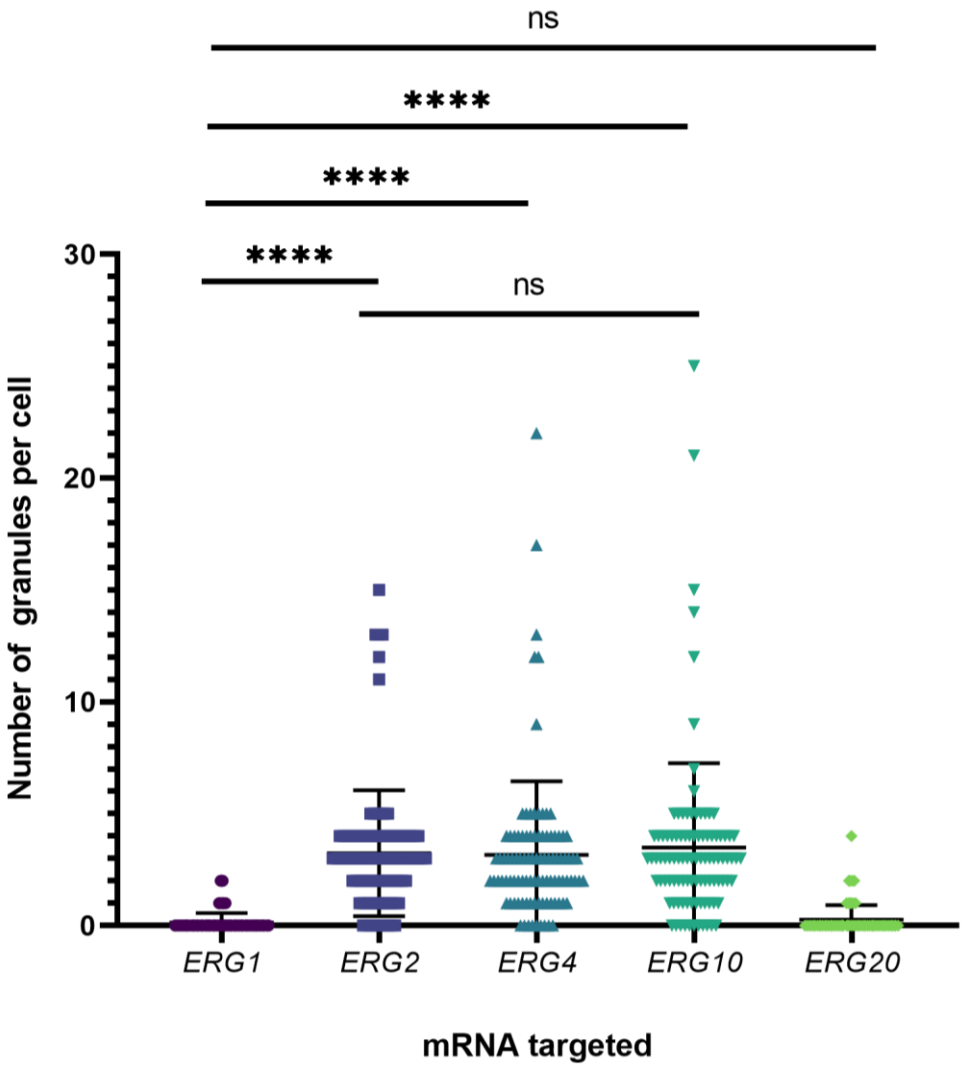


Figure 5.17: Quantification of ergosterol mRNA localisation in yeast. For each mRNA the number of granules per cell is shown (n=100). For *ERG1* and *ERG20* mRNAs, little evidence for RNA granules was found. For *ERG2*, *ERG4* and *ERG10* mRNAs, the average number of granules per cell was 4-5, and a small number of cells displayed 10-20 granules per cell. Error bars are shown as the mean with standard deviation. One-way ANOVA statistical analysis showed no significant differences between *ERG2*, *ERG4* and *ERG10*, and also between *ERG1* and *ERG20*. One-way ANOVA did show significant differences between *ERG1* and *ERG2*, *ERG4* and *ERG10*, and between *ERG20* and *ERG2*, *ERG4* and *ERG10* (not shown on graph). For all significant differences, P<0.0001.

In general, the highly expressed mRNAs targeted here did not appear to form cytosolic granules in unstressed yeast cells (Figure 5.18). Both imaging and quantification of *RPL3*, *RPS31*, *CCW12* and *ASC1* mRNAs showed that cytosolic granules were not present in almost all cells imaged (n=100) (Figure 5.18, bottom row and Figure 5.19). For *RPL28*, the majority of cells (~90%) also showed no evidence of cytosolic granules (Figure 5.18, top row and Figure 5.19). Interestingly, a small number of cells targeting *RPL28* showed many granules per cell (5-20 granules) (Figure 5.19). This expression pattern was seen almost exclusively in large mother cells and not smaller daughter cells (Figure 5.18, top row). This expression pattern may therefore be driven by a change in cell status (e.g. autophagy or stress), which would have been investigated further if there was time. Nevertheless this data indicated that the highly expressed mRNAs targeted here do not form cytosolic RNA granules in unstressed yeast cells.

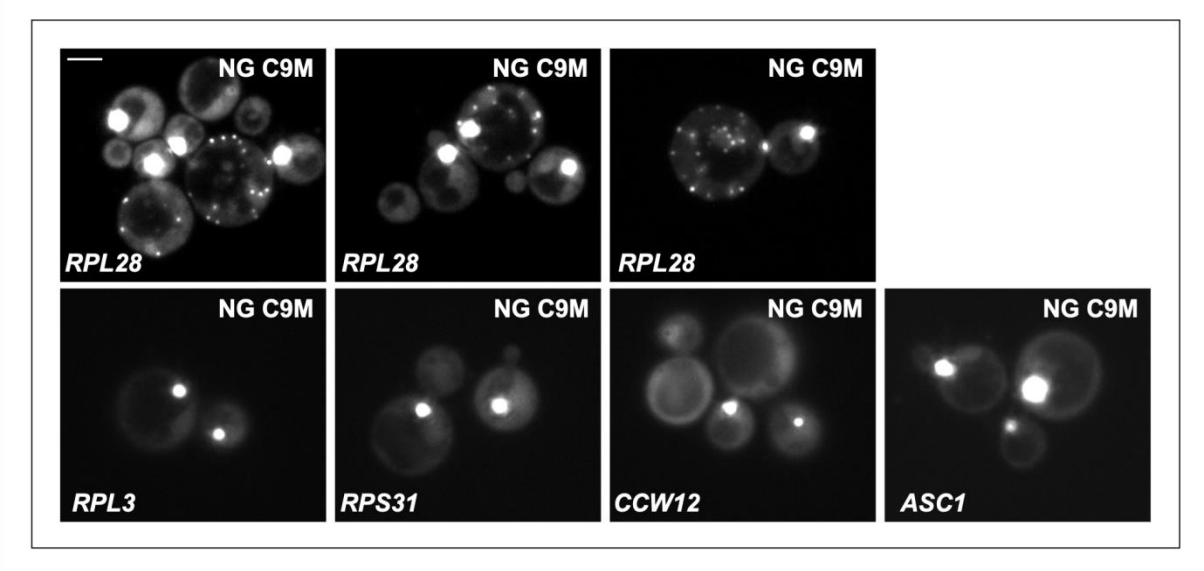


Figure 5.18: mRNA localisation of highly expressed mRNAs imaged in yeast. All cells are expressing the nanobody-GFP and dCas9-MT constructs, along with a sgRNA for the mRNA being targeted (*RPL28*, *RPL3*, *RPS31*, *CCW12* or *ASC1*). *RPL28* mRNA (top row) appears to either form many granules or no granules per cell. The bottom row shows *RPL3*, *RPS31*, *CCW12* and *ASC1* mRNAs that do not appear to form cytosolic granules. Images were taken as Z-stacks and the average intensity image is shown. Scale bar: 3μm.

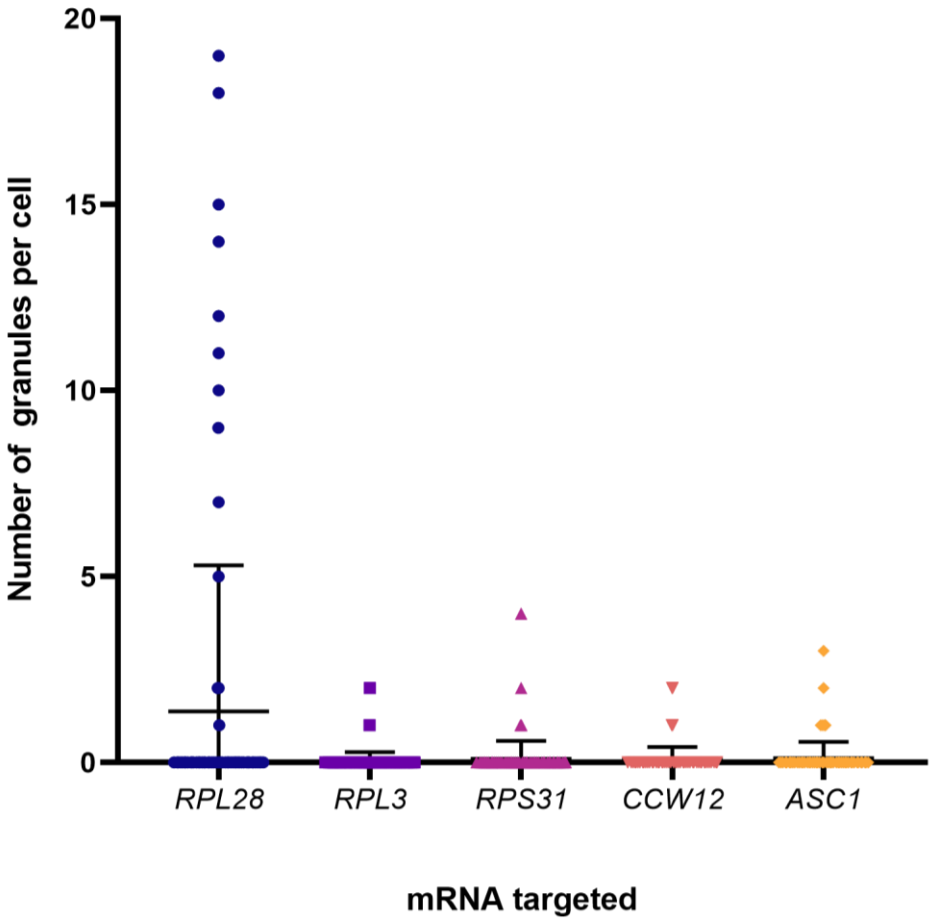


Figure 5.19: Quantification of mRNA localisation of highly expressed mRNAs in yeast. For each mRNA the number of granules per cell is shown (n=100). For *RPL28* mRNA, most cells show zero granules per cell, while a small number show 5-20 granules per cell. For *RPL3*, *RPS31, CCW12* and *ASC1* mRNAs, the average number of granules per cell was zero. Error bars are shown as the mean with standard deviation.

5.5. Conclusions.

In this chapter the dCas9 RNA-targeting system was improved by exchanging the dCas9-GFP_3 from the initial system (described in the previous chapter) for dCas9-MoonTag (dCas9-MT). The MoonTag is a repeat peptide array containing 24x gp41 sequence repeats taken from viral sequences (HIV-1) (Tanenbaum et al., 2014; Boersma et al., 2019). The MoonTag is recognised and bound by single chain anti-gp41 antibodies (nanobodies), which were here fused to a single GFP (nanobody-GFP). When the nanobody-GFP and dCas9-MT are expressed together in the same cell, multiple nanobody-GFP constructs bind to the MoonTag array attached to the dCas9. In the previous dCas9-GFP_3 system three GFP molecules were targeted to each mRNA, whereas in this dCas9-MT system the MoonTag array is bound by up to 12x nanobody-GFP molecules (Boersma et al., 2019). Therefore this dCas9-MT system may be up to 4 times brighter than the previous dCas9-GFP_3 system.

Although exact quantification was not done were, this dCas9-MT system is noticeably brighter than the dCas9-GFP_3 system when visualised both alone (nanobody-GFP and dCas9-MT) and when targeted to mRNAs (nanobody-GFP, dCas9-MT and a sgRNA). The dCas9-MT was first used to target mRNAs whose localisation had been previously studied in the lab using the MS2 system, smFISH and the dCas9-GFP_3 system here, where the same pattern of cytosolic RNA granules was shown for each mRNA using dCas9-MT. *ENO1* and *ENO2* mRNAs had been previously shown to localise to multiple (10-20) granules per cell (Lui et al., 2014; Morales-Polanco et al., 2021) and *NIP1* and *TIF1* mRNAs to one or two granules per cell (Lui et al., 2014; Morales-Polanco et al., 2021). These results were recapitulated here with the dCas9-MT system, where previously *TIF1* mRNA granules had not been bright to image convincingly using the dCas9-GFP_3 system. The brightness levels of the dCas9-MT system appear comparable with the version of the MS2 system used here, which introduces 12x MS2 stem loops into the target mRNA to allow binding of up to 72 GFP molecules per mRNA (Valegard et al., 1990; Peabody and Lim, 1996; Rolfsson et al., 2016).

The dCas9-MT system was then verified further by targeting other mRNAs previously imaged in the lab; *PDC1*, *PDC5*, *FBA1*, *SUP35* and *ARO3*. *PDC1* and *FBA1* mRNAs are glycolytic mRNAs previously shown to localise to multiple granules per cell (Lui et al., 2014; Morales-Polanco et al., 2021), *SUP35* is a translation factor mRNA previously shown to localise to one or two granules per cell (Lui et al., 2014; Pizzinga et al., 2019) and *ARO3* is an aromatic amino acid biosynthesis mRNA shown to localise to 3-5 granules per cell (Ashe lab, unpublished data). These results were all recapitulated here using the dCas9-MT system. *PDC5* is a glycolytic mRNA that does not appear to form cytosolic RNA granules in unstressed cells when visualised using the MS2 system. However, both unpublished data

from the Ashe lab using smFISH and data found here using the dCas9-MT system showed that *PDC5* does appear in 10-20 RNA granules per cell. Imaging was done carefully here to ensure the yeast cells were not stressed and BLAST analysis of the *PDC5* sgRNA was used to verify the sgRNA is not able to bind other mRNA sequences. Therefore the differences in *PDC5* localisation may be due to problems with the MS2 system targeting; either rearrangement of the genomic MS2 stem loops may have occurred or they may be issues with plasmid expression in these cells.

Since a major criticism of the MS2 system is the possibility that the introduction of the MS2 stem loops may affect rates of mRNA decay/transcription, qRT-PCR was performed here on mRNAs targeted with the dCas9-MT system. qRT-PCR showed that there were no significant differences from endogenous mRNA expression for *NIP1*, *TIF1*, *ENO1* or *ENO2* when targeted with the dCas9-MT, which was expected as the technique does not involve modification of the mRNA structure. Experiments have also been done previously in the lab to show that the abundance of these mRNAs is also unchanged when targeted using the MS2 system (Pizzinga et al., 2019; Morales-Polanco et al., 2021). It would therefore be interesting to target an mRNA whose expression levels are changed when targeted with the MS2 system (Garcia and Parker, 2015; Haimovich et al., 2016; Heinrich et al., 2017) and test whether targeting using the dCas9-MT system has the same effect.

Along with the ability of the dCas9-MT to target endogenous, unmodified mRNAs, another advantage of the system is that it is technically quick and easy to use. Once the initial constructs have been established in a yeast strain (the nanobody-GFP and dCas9-MT constructs), new sgRNAs can be rapidly generated and transformed into yeast to target new mRNAs. One application of this system could therefore be to screen mRNAs for novel mRNA localisation. This was tested here, where target mRNAs included 5 highly expressed, non-glycolytic mRNAs (*RPL28*, *RPL3*, *RPS31*, *CCW12* and *ASC1*) and 5 ergosterol biosynthesis mRNAs (*ERG1*, *ERG2*, *ERG4*, *ERG10*, *ERG20*).

Since glycolytic mRNAs are highly abundant and form many granules per cell (as opposed to translation factor mRNAs that are lower abundance and form one or two granules per cell), it was investigated here whether or not other highly expressed mRNAs would form cytosolic RNA granules in unstressed yeast cells. 4 of these mRNAs (*RPL28*, *RPL3*, *RPS31* and *ASC1*) encode ribosomal proteins, which in yeast are thought to be co-ordinately expressed at the level of transcription (Mager and Planta, 1991; Xiao and Grove, 2009). These mRNAs did not appear to localise to cytosolic granules, although in some cells (~10%), *RPL28* was present in multiple granules per cell. These cells were always large mother cells instead of daughter cells, and so this localisation pattern is possibly a result of

autophagy or cellular stress. If time had allowed, these cells would have been imaged with DAPI simultaneously to see if the nuclei remained intact, as the nuclei in these cells were not obvious. Given that these mRNAs do not share a similar function (as with the glycolytic, translation factor, and aromatic amino acid synthesis mRNAs), these results support the hypothesis that the purpose of the granules is to localise together functionally related proteins to improve the efficiency or regulatability of their pathways.

This is further supported by imaging of the ergosterol biosynthesis mRNAs, which were chosen because their encoded proteins function in the same pathway. Of the 5 ergosterol mRNAs tested here, three mRNAs (*ERG2*, *ERG4* and *ERG10*) displayed cytosolic RNA granules in unstressed cells, while the remaining two mRNAs (*ERG1* and *ERG20*) did not appear localised under the same conditions. *ERG2*, *ERG4* and *ERG10* appeared on average to be present in 4-5 granules per cell. This data supports the hypothesis that the higher the abundance of mRNA the higher the number of cytosolic mRNA granules, as glycolytic mRNAs (high abundance) show the highest numbers of granules per cell, and translation factor mRNAs (low abundance) show fewer numbers of granules per cell (Pizzinga et al., 2019; Morales-Polanco et al., 2021; Costello et al., 2015; Lawless et al., 2016; Lahtvee et al., 2017). Ergosterol mRNAs are expressed at levels somewhere between the two, along with aromatic amino acid mRNAs, and both of these types of mRNAs show around 3-5 cytosolic RNA granules per cell (Costello et al., 2015; Lawless et al., 2016; Lahtvee et al., 2017; Ashe lab, unpublished data). Furthermore, *YEF3* and *TEF1* are translation factor mRNAs that are present in higher copy numbers than most other translation factors mRNAs (such as *NIP1* and *TIF1*) and *YEF3* and *TEF1* are present in 10-20 granules per cells (Lawless et al., 2016; Lahtvee et al., 2017; Pizzinga et al., 2019). This data is also understandable from a numerical standpoint; to form lots of granules per cell (capable of being visualised with non-single molecule resolution techniques) lots of molecules of mRNAs would be required. *ERG2*, *ERG4* and *ERG10* did occasionally (~10%) appear in 10-20 granules per cell, possibly caused by small numbers of stressed yeast undergoing metabolic changes. Although this dCas9-MT has thus far demonstrated reliable targeting of mRNAs in yeast, the localisation patterns of these mRNAs would have been verified with the MS2 system or smFISH had time allowed.

Overall, the dCas9-MT RNA-targeting system developed here is capable of correctly visualising the localisation of endogenous, unmodified mRNAs in yeast, and the technique has the potential to provide further screening of mRNAs to identify novel localisation patterns in yeast.

Discussion

6. Discussion

6.1. General discussion

In this project, mRNA localisation to RNA granules in yeast was investigated using techniques that allow the visualisation of mRNA. This work was broadly split into two different aspects; one to use the MS2 system to visualise mRNA domain exchanges and investigate mechanisms of mRNA localisation, and the other to develop a novel mRNA visualisation system in yeast, the dCas9 system.

Previous work in the Ashe lab has identified cytosolic RNA granules that are present in unstressed yeast cells and do not localise with other RNA granule markers (such as stress granules or P-body markers) (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). These granules include glycolytic granules, which are present in 10-20 granules per cell and contain mRNAs encoding glycolytic enzymes (such as *PDC1*, *ENO2* and *FBA1*), and translation factor granules, which are present in one or two granules per cell and contain mRNAs encoding translation factors (such as *NIP1*, *TIF1* and *SUP35*) (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). To investigate the driving factors behind this differential mRNA localisation, mRNA swap constructs were made here between *NIP1* and *PDC1* mRNAs. The main results from these experiments showed that the exchange of the ORF between these mRNAs (e.g. the ORF of *PDC1* and the promoter, 5'UTR and 3'UTR of *NIP1*, and vice versa) resulted in a change of localisation pattern; both the *PDC1* ORF and the *NIP1* ORF are sufficient to drive the mRNA localisation pattern to appear the same as the respective native mRNA. Therefore, either the ORFs of both *NIP1* and *PDC1* contain RNA localisation elements, or the production of the protein in each case drives its own mRNA localisation.

A dCas9 RNA-targeting system was also developed here for use in yeast, adapted from a similar system used in mammalian cells (Nelles et al., 2016). The final version of the dCas9 system developed here showed comparable brightness to the MS2 system and successfully recapitulated mRNA localisation previously shown in the lab using the MS2 system and smFISH. It was shown both here and in mammalian cells that mRNAs targeted by dCas9 are not affected in terms of their overall levels (Nelles et al., 2016). Other advantages of the dCas9 system are that it is a live cell system and so can be used to study mRNA dynamics, and that the technique is quick and easy to use. The dCas9 system was therefore used here to screen mRNAs and study novel localisation patterns using a more rapid protocol than would have been possible for either the MS2 system or smFISH. To our knowledge, this is the first time a dCas9 RNA-targeting system has been utilised in yeast, and ultimately this

system has the potential to be used to screen every mRNA in yeast by generating a strain collection to study novel localisation patterns.

6.2. dCas9 RNA-targeting system in yeast

The development of the dCas9 system was driven by relatively recent criticism of the MS2 system. The MS2 system most often involves the introduction of multiple stem loops at the start of the 3'UTR of the target mRNA, which therefore alters the length and structure of the 3' regulatory region. It has been shown that for certain mRNAs, this modification may affect the rates of mRNA translation and decay (Garcia and Parker, 2015; Haimovich et al., 2016; Garcia and Parker 2016). Importantly, it has also been shown that the MS2 loops bound by the MCP recognition protein may prevent 5' to 3' mRNA decay, leading to an accumulation of mRNA decay fragments that are visualised by the MS2 system (Garcia and Parker, 2015; Heinrich et al., 2017). Therefore in the Ashe lab, mRNA localisation visualised using the MS2 system has been verified using smFISH. However smFISH is a fixed cell method, and so cannot be used to study mRNA dynamics and may also introduce imaging artefacts as a consequence of the fixation and permeabilisation of the yeast cells (Schnell et al., 2012; Whelan and Bell, 2015).

Therefore in this project, a dCas9 RNA-targeting system was developed for use in yeast. At the start of this project, it had already been shown that Cas9 proteins could be used to target mRNAs (Sampson et al., 2013; Price et al., 2015) and a dCas9 RNA-targeting system had been recently described to visualise mRNA localisation in mammalian cells (Nelles et al., 2016). Since then, it has been shown that Cas13 proteins specifically target RNAs (Yan et al., 2018; Smargon et al., 2017) and RNA imaging systems have been developed using dCas13a and dCas13b proteins (Yang et al., 2019; Huynh et al., 2020). If these dCas13 RNA-targeting systems had been available at the start of this project they might have been chosen instead to avoid genomic targeting by dCas9. Nevertheless, the dCas9 RNA-targeting system was successfully developed for use here in yeast, and has allowed both validation of previous mRNA localisation results found in the lab and the targeting of previously uncharacterised mRNAs to study their localisation in yeast.

In the mammalian cell system, dCas9 is targeted to mRNAs via both a sgRNA (with a 20bp recognition region for the mRNA of interest and a dCas9 binding region) and a PAMmer (Nelles et al., 2016) (Figure 1.12). The PAMmer is a short, single stranded RNA that also carries a recognition region for the mRNA of interest followed by the PAM sequence –NGG. Although it has been shown that dCas9 is capable of being targeted to mRNAs in the

absence of a PAMmer (by a sgRNA alone) (Strutt et al., 2018; Nelles et al., 2016), in the mammalian system it was shown that targeting was more efficient when a PAMmer was included (Nelles et al., 2016). The dCas9 system was to be used here to visualise mRNA localisation to RNA granules in unstressed yeast cells. Therefore, the dCas9 system was targeted to mRNAs using a sgRNA alone, as techniques to permeabilise/electroporate yeast to incorporate the PAMmer would have likely caused stress and structural effects on the cells that would potentially affect the localisation of the mRNAs being studied. Fortunately, it appears the introduction of a sgRNA alone is sufficient to target dCas9 to the mRNA of interest in yeast.

The initial version of this dCas9 system involved the targeting of a dCas9-GFP_3 fusion protein to the target mRNA. The dCas9 also possesses an NLS so that the signal is confined to the nucleus when not targeted and the dCas9 may interact with mRNAs as they are made. Although the dCas9-GFP_3 was shown to leave the nucleus and localise to cytosolic RNA granules when targeted to *NIP1*, *TIF1*, *ENO1* or *ENO2* mRNAs, the system was much less bright than the MS2 system and was difficult to use. The brightness of the dCas9 system was therefore improved by the addition of a MoonTag peptide array to the dCas9 protein (dCas9-MT) instead of the GFP_3 fusion. Along with dCas9-MT, a nanobody-GFP construct was also expressed in cells, where up to 12x nanobody-GFP molecules may bind the MoonTag peptide array (Tanenbaum et al., 2014; Boersma et al., 2019). MoonTag is a variation of the SunTag approach; SunTag uses a *GCN4* peptide array rather than the gp41 peptide array utilised by MoonTag. It is unclear whether SunTag can be used in yeast as *GCN4* is an endogenous yeast gene, while the gp41 peptide array is derived from viral sequences (HIV) (Tanenbaum et al., 2014; Boersma et al., 2019). At the time of writing this project this use of epitope labelling to increase the fluorescence of a fusion protein had not been demonstrated in yeast, though the development of the MoonTag array makes this a possibility.

The dCas9-MT was then used to recapitulate mRNA localisation data found previously in the lab using the MS2 system and smFISH. *NIP1*, *TIF1* and *SUP35* mRNAs were shown by all three systems to be present in one or two granules per yeast cell, while *ENO1*, *ENO2*, *PDC1* and *FBA1* mRNAs were shown to be present in 10-20 granules per cell (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). The dCas9-MT is therefore capable of targeting multiple different mRNAs, and sgRNAs appear specific for the target mRNA. Although Cas targeting systems may demonstrate off-target effects in some cases (Doench et al., 2016; Ran et al., 2013), it has been shown that at least a 17 nucleotide match is required for dCas9 targeting (Fu et al., 2014). Accordingly, each sgRNA designed was

verified by BLAST to not match more than 12 nucleotides of any other yeast sequence. Since the MS2 system has been criticised for possible effects on mRNA abundance levels, qRT-PCR was then performed here to investigate whether dCas9-MT targeting has an effect on target mRNA levels. Since the targeting of mRNA by dCas9 does not involve the modification of the mRNA structure in any way, it might have been anticipated that dCas9 targeting would not affect mRNA levels. However, the dCas9 targeting itself could affect diverse aspects of RNA fate such as mRNA processing and/or decay. Overall it appears that for the examples characterised here, the mRNA level was unaffected by the binding of the sgRNA and dCas9-MT.

Along with the lack of target mRNA alteration, another advantage of the dCas9-MT system is that is relatively quick to target different mRNAs, as a new sgRNA can be generated via a single cloning step and then directly transformed into yeast already expressing the nanobody-GFP and dCas9-MT constructs. The dCas9-MT system was therefore used here to screen mRNAs for localisation that had not been previously studied. To investigate whether or not mRNA expression levels affect mRNA localisation to cytosolic RNA granules, the 5 most highly abundant mRNAs in yeast (excluding previously assessed glycolytic and translation factor mRNAs) were targeted (Costello et al., 2015; Lawless et al., 2016; Lahtvee et al., 2017). These mRNAs encoded ribosomal proteins (*RPL3*, *RPL28*, *RPS31*, and *ASC1*) and a cell wall protein (*CCW12*). No evidence of cytosolic RNA granules could be seen for any of these mRNAs, suggesting that mRNA expression levels alone do not encourage the formation of cytosolic RNA granules and supporting the hypothesis that these actively translating granules form to increase the efficiency of their relevant pathways. Therefore the other 5 mRNAs targeted here were mRNAs that encode ergosterol biosynthesis enzymes (*ERG1*, *ERG2*, *ERG4*, *ERG10* and *ERG20*). Using the dCas9-MT system, cytosolic RNA granules could be seen in unstressed yeast cells for *ERG2*, *ERG4* and *ERG10* mRNAs, while *ERG1* and *ERG20* mRNAs did not appear to form cytosolic granules. *ERG2*, *ERG4* and *ERG10* mRNAs are more highly expressed than *NIP1* (~10 copies per cell), less highly expressed than *PDC1* (~100 copies per cell) and expressed at around the same levels as *ARO3* (Lawless et al., 2016; Lahtvee et al., 2017). The most common pattern of granule localisation for *ERG2*, *ERG4* and *ERG10* was 3-5 granules per cell; the same number as for *ARO3*, and more than *NIP1* and fewer than *PDC1*. Together these data seem to indicate that the higher the abundance of mRNA, the higher the numbers of their cytosolic granules that appear per cell (which is reasonable as more mRNA copies would be necessary to form multiple granules). Therefore, mRNA abundance may be a driving factor involved in determining the number of granules per cell for localised mRNAs, as discussed for the CEN derived plasmid constructs below.

Although it appears the dCas9-MT system does not affect the mRNA dynamics of the target mRNA, control experiments would still be valuable to validate the system and the novel mRNA localisation data found here. Further experiments that would have been completed if there was time involve co-localisation experiments of different mRNAs, and co-localisation experiments with the MS2 system/smFISH. Previous work in the lab has shown that several mRNAs co-localise together in cytosolic granules, for example *NIP1* and *TIF1*. Experiments have also shown that some of these mRNAs specifically do not co-localise, such as *ENO1* and *ENO2* (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). It therefore would have been ideal to repeat these experiments with the dCas9 system, where multi-colour experiments have been performed previously in mammalian cells using dCas9 proteins from different bacterial species. These dCas9 proteins recognise different sgRNA scaffolds, and so could be attached to different fluorophores and used to study two mRNAs simultaneously (Ma et al., 2015). Another good control would have been to target the same mRNA with the MS2 system and the dCas9 system in the same cell to confirm that the localisation patterns overlap. However this would have required difficult cloning to swap plasmid markers to express MS2 and dCas9 constructs in the same cell.

To improve the viability of the dCas9-MT as a system for screening mRNA localisation, it was attempted here to remove the CEN ORI from the dCas9-MT plasmid to allow integration into the yeast genome (as with the nanobody-GFP). This would have removed the requirement to maintain selection of the dCas9-MT plasmid, and may have improved expression stability between cells. Indeed, some differences in plasmid expression levels can be seen in yeast cells here. Unfortunately due to cloning issues the dCas9-MT integration could not be completed in time. The dCas9-MT system may also benefit from lower expression of the nanobody-GFP to reduce the nuclear signal seen in these cells, which would also have been attempted if there was time. Alternatively, it was attempted to knockout the NLS from dCas9 to see if there was an effect on mRNA targeting. Although there was a concern that the bright GFP signal from the nucleus could mask the presence of cytosolic RNA granules, this also would have investigated if time had allowed. Another interesting experiment would have been to increase the number of fluorophores attached to dCas9 to attempt to reach single molecule resolution with this system, although this could also result in an increase in nuclear signal.

Overall, the dCas9-MoonTag RNA-targeting system developed here for use in yeast is a quick and efficient live cell technique that can be used to screen localisation of endogenous, unmodified mRNAs without affecting mRNA dynamics. Ultimately, it would therefore be possible to use this dCas9-MT system to generate a strain collection facilitating the analysis

of every mRNA in yeast; this could be coupled to automated image collection and analysis to identify novel mRNA localisation patterns and to study mRNA localisation changes in response to a plethora of different conditions.

6.3. mRNA mutagenesis and localisation mechanisms

Previous work in the Ashe lab has identified at least three types of novel cytosolic RNA granules in yeast (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). These granules include glycolytic granules, which contain mRNAs encoding glycolytic enzymes (including *PDC1*, *ENO2* and *FBA1*) and translation factor granules containing translation factor mRNAs (including *NIP1*, *TIF1* and *SUP35*) (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). Since these granules contain mRNAs that are actively translated and that encode proteins with similar functions, the purpose of these granules may be to improve the efficiency or regulatability of the relevant pathway.

As the dCas9 system was under development throughout this project, experiments designed to investigate the mechanisms of mRNA localisation were performed using the MS2 system. This investigation focused on the mRNAs *NIP1* and *PDC1*, where *NIP1* localises to one or two cytosolic RNA granules and *PDC1* to many (~10-20) cytosolic RNA granules in unstressed yeast cells (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021). In an attempt to identify the localisation elements responsible for this differential mRNA localisation, mRNA swaps were made here, where regions of the mRNAs (5'UTR, ORF and 3'UTR) were exchanged between *NIP1* and *PDC1*. It was decided that the MS2 system was the appropriate visualisation system for these experiments for two main reasons:

1. The possible effects of MS2 modification on *NIP1* and *PDC1* mRNAs has already been investigated in the Ashe lab. qRT-PCR has been performed on *NIP1* and *PDC1* to show that the addition of MS2 stem loops does not alter the abundance of the mRNAs (Pizzinga et al., 2019; Morales-Polanco et al., 2021), and smFISH has been performed on these cells to verify the mRNA localisation pattern seen using the MS2 system (Pizzinga et al., 2019; Morales-Polanco et al., 2021). Therefore, MS2 tagging does not appear to affect *NIP1* or *PDC1* mRNA dynamics.
2. Since each mRNA swap construct was a mixture of *NIP1* and *PDC1* sequences, specifically targeting these constructs (and not the endogenous *NIP1* and *PDC1* mRNAs) would have been difficult using the dCas9 system and not possible using smFISH. dCas9 sgRNAs would have had to be designed to target the junction regions of swap constructs (e.g. the last 10 nucleotides of the *NIP1* 5'UTR followed

by the first 10 nucleotides of the *PDC1* ORF), which may not have been possible for all swap constructs. The smFISH system relies on multiple probes binding to the mRNA of interest, and so would not have been bright enough for visualisation with the one or two probes that would have been specific for each swap construct. Instead, swap constructs were produced from previously constructed MS2 tagged *NIP1* and *PDC1* on yeast expression vectors and immediately imaged using the MS2 system.

To investigate mechanisms of mRNA localisation to these granules various mRNA swap constructs between *NIP1* and *PDC1* were made in this project. Possibly the most interesting construct was the *NIP1 PDC1* ORF construct, containing the ORF of *PDC1* and the promoter and UTRs of *NIP1*. This construct demonstrated that the presence of the *PDC1* ORF was capable of driving mRNA localisation to look more like *PDC1*, where the construct appeared localised in ~10 cytosolic RNA granules per cell (unlike *NIP1*, which is present in one or two granules per cell). The opposite version of this construct, *PDC1 NIP1* ORF, demonstrated a similar outcome; this construct appeared to look more like *NIP1* than *PDC1*. These results therefore indicate that either RNA localisation elements are present in both the *NIP1* and *PDC1* ORFs, or that the production of the corresponding protein is responsible for driving its own mRNA localisation. The situation may be more complex, as constructs containing additional regions of *PDC1* (*NIP1 PDC1* 5'UTR ORF and *NIP1 PDC1* ORF 3'UTR) appeared to show a granule pattern more similar to full-length *NIP1*. However, due to the fewer numbers of MS2 stem loops present in these two constructs, it is unclear whether this is a biological or technical result, which could be examined in the future by producing the two constructs with the correct number (12) of MS2 stem loops. To further investigate the role of the ORF in mRNA localisation, it was attempted here to produce a mutated *PDC1* ORF sequence, where the primary sequence was changed as much as possible and codons were kept the same to produce a functional Pdc1 protein. Unfortunately due to cloning issues this construct was not introduced into yeast in time for this report, although this would be a valuable future experiment.

Along with encoding the associated protein, the ORF of mRNAs can also play roles in the regulation of mRNA translation. Codon usage has a strong effect on rates of translation, where the inclusion of non-optimal codons inhibits elongation and reduces the stability of an mRNA (Boël et al., 2016; Presnyak et al., 2015). This is likely important with regards to *NIP1* and *PDC1*. *PDC1* is a relatively high abundance mRNA and utilises optimal codons in almost every instance, whereas *NIP1* is a relatively low abundance mRNA and does not utilise many optimal codons. Therefore, the *PDC1* ORF mutant designed here will also have

an altered codon usage, which may affect mRNA dynamics and localisation patterns separately from RNA/nascent protein localisation elements. Chemical modifications may also occur to nucleotides in the mRNA coding region and affect translation; for example, m^6A methylation of nucleotides can inhibit translation elongation (Hoernes et al., 2016). Further, mRNAs may form secondary and tertiary structures depending on their sequence, which are unfolded by the ribosome during translation (Chen et al., 2013), and many of these structures also results in inhibition of translation (Kontos et al., 2001). Since it has been shown that the localisation of *NIP1* and *PDC1* relies on the mRNAs being translated (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021), these are important effects to keep in mind when performing mRNA mutagenesis experiments.

Although UTR swap constructs were also not successfully produced here due to cloning difficulties, they also would have provided valuable data. It is possible for an mRNA to contain multiple localisation elements sufficient to drive localisation, and these elements may be found in both the ORF and UTRs (such as *ASH1* (Gonzalez et al., 1999; Chartrand et al., 1999)), which may also be the case for *NIP1* and/or *PDC1*. This data would also be important in the context of protein or RNA elements driving mRNA localisation; if mRNA localisation is driven by the production of its protein, RNA localisation elements in the UTRs would be redundant. However, redundant mRNA localisation signals have been shown to play other roles important in mRNA localisation, for example tethering the mRNA at its destination (again *ASH1* (Gonzalez et al., 1999; Chartrand et al., 1999)), which may also be true for *NIP1* and/or *PDC1*.

If synthesis of these swap constructs could be attempted again, a more efficient cloning strategy could be used to focus on restriction cloning rather than Gibson cloning. Due to the lack of restriction enzyme sites between UTR and ORF regions for both *NIP1* and *PDC1*, swap constructs were made here by Gibson cloning, which does not rely on restriction sites (Gibson et al., 2009). However Gibson cloning does require PCR to generate fragments with overlapping ends. PCR fragments were difficult to produce for both long DNA fragments (e.g. the plasmid backbone with one or both UTRs) and for fragments that contained MS2 stem loops, as the repetitive loop sequences appeared to either prevent polymerase extension or resulted in MS2 stem loop loss. Therefore, a cloning strategy where restriction sites were introduced into the mRNA region junctions before cloning, followed by restriction cloning may have been more successful.

With regards to the mechanism of *NIP1* and *PDC1* mRNA localisation to RNA granules, relatively little is known. Previous work has shown that the localisation of *NIP1* is dependent on the She2p/She3p/Myo4p transport machinery, where the deletion of She2p disrupted

NIP1 mRNA localisation to granules (Pizzinga et al., 2019). This may suggest that the localisation of *NIP1* is similar to the well-characterised localisation of *ASH1*; *ASH1* mRNA contains RNA localisation elements in both its ORF and 3'UTR regions that are recognised by She2p and therefore linked (via She3p interactions) to the motor protein Myo4p to facilitate mRNA transport (Bertrand et al., 1998; Gonzalez et al., 1999; Chartrand et al., 1999). Since She2p binds to mRNA localisation elements, it is therefore possible that the *NIP1* ORF also contains mRNA localisation elements that are recognised by She2p. Conversely, it has also been shown that mRNAs whose localisation depends on the presence of their ORFs (and translation of the nascent protein) may be localised in a 'zip-code independent' manner (Liao and Liu, 2011; Hirashima et al., 2018). Since the localisation of *NIP1* and *PDC1* mRNAs also appear to be driven by the presence and translation of their ORFs (Lui et al., 2014; Pizzinga et al., 2019; Morales-Polanco et al., 2021) it is also possible that they may be localised by similar zip-code independent mechanisms. Overall there are many possibilities for the mRNA localisation mechanisms of *NIP1* and *PDC1*, which may involve a combination of zip-code dependent and independent methods.

In addition to mRNA sequence localisation elements, it was also investigated here whether or not the expression levels of *NIP1* and *PDC1* affected their localisation, particularly since two highly expressed translation factor mRNAs, *YEF3* and *TEF1*, are present in multiple granules per cell (Pizzinga et al., 2019). *PDC1* is a highly abundant mRNA (around 16 times more abundant than *NIP1* (Lawless et al., 2016; Lahtvee et al., 2017)), which is possibly the reason why *PDC1* is found in higher numbers of cytosolic RNA granules than *NIP1*. MS2 tagged *NIP1*, *PDC1* and the *NIP1 PDC1* ORF construct were therefore cloned into a lower expression yeast plasmid (CEN derived) in a bid to reduce the copy number of the mRNAs. These CEN constructs appeared dimmer than 2μ derived constructs and required longer exposure times to image. This is likely because there were fewer mRNAs to target and this version of the MS2 system does not achieve single molecule resolution (Pizzinga et al., 2019), so only bulk mRNA can be visualised. Nevertheless, each CEN construct displayed the same RNA granule localisation pattern as genomic or 2μ derived expression; *NIP1* CEN showed one or two granules per cell and *PDC1* and *NIP1 PDC1* ORF showed ~10 granules per cell.

qRT-PCR results showed that although the expression of each construct was reduced in the CEN derived plasmid compared to the 2μ plasmid, both *PDC1* and the *NIP1 PDC1* ORF constructs were still expressed at higher levels than *NIP1*, possibly explaining why these constructs still appeared in many granules per cell. Indeed, endogenous *NIP1* is present in

~10 copies per cell and *PDC1* in ~160 copies per cell (Lawless et al., 2016; Lahtvee et al., 2017). These data are concurrent with qRT-PCR data here that suggests CEN driven plasmid expression of constructs produces ~20 times more copies of *PDC1* than *NIP1*. Since the CEN plasmid is expected to be present in a single copy per cell (Clarke and Carbon, 1985), it is unsurprising that CEN expression of *PDC1* still results in multiple cytosolic RNA granules, as this would represent endogenous expression levels. Since *PDC1* appears to be always localised to multiple granules (whereas *NIP1* is localised to one or two), it is likely that reducing the levels of *PDC1* further would result in the mRNA being unable to be visualised by the MS2 system.

Nevertheless, the fact that reduction in the levels of mRNA abundance for the *NIP1 PDC1* ORF construct did not alter its expression pattern to RNA granules suggests that, like *PDC1*, the construct localises to multiple granules regardless of its expression levels. This may be explained by the fact that multiple mRNAs are co-localised together in these granules. Taking glycolytic granules as an example, if the function of these granules is to increase the efficiency of the glycolytic pathway by co-localising mRNAs encoding for glycolytic proteins, then each granule would need copies of each glycolytic mRNA. Therefore, since all other glycolytic mRNAs are expressed at their endogenous levels (and the endogenous, untagged mRNA is also being produced from the genome), the mRNA levels of *PDC1* and the *NIP1 PDC1* ORF do not alter the granule localisation pattern.

An interesting result from the qRT-PCR data showed that the *NIP1 PDC1* ORF construct was expressed slightly more than *PDC1* and many more times than *NIP1* in both the 2μ and CEN plasmids. Since the *NIP1 PDC1* ORF construct maintains the promoter, 5'UTR, 3'UTR and terminator region of *NIP1*, it would therefore possibly be expected that this construct would display similar expression levels to *NIP1*. For the 2μ driven plasmids, this may be explained by differences in 2μ plasmid abundance between *NIP1* and *NIP1 PDC1* ORF constructs (the 2μ plasmid may be present in 40-60 copies per cell and the *NIP1* promoter may result in ~10 mRNA copies (Chan et al., 2013; Lawless et al., 2016; Lahtvee et al., 2017)). However qRT-PCR data also showed that the *NIP1 PDC1* ORF was ~100 times more abundant than the *NIP1* mRNA when both were expressed on the CEN plasmid. One possible explanation is that the production of the Pdc1 protein affects the expression from the plasmid in some way, though this would need to be investigated further. It is also possible that mRNA decay rates for the *NIP1 PDC1* ORF construct were lower than *NIP1* as the construct was protected in RNA granules. The exchange of promoter regions of *NIP1* and *PDC1* mRNAs would have been interesting to investigate the effects of mRNA

expression levels on localisation patterns, which would have been completed here if time had allowed.

Further experiments that would have been performed include the construction of multiple other mRNA swap constructs and co-localisation experiments with other glycolytic mRNAs. As discussed above, UTR exchanges between *NIP1* and *PDC1* mRNA would have been studied if cloning had been successful. Finally, since the genomic MS2 tagged *PDC1* has been shown to co-localise with other glycolytic mRNAs, an interesting experiment would be to investigate whether or not this co-localisation also occurs for the *NIP1 PDC1* ORF construct. Unfortunately, the *NIP1 PDC1* ORF construct is expressed on a *URA3* yeast vector, and *URA3* selection is necessary for the function of the PP7 system (to allow dual imaging with the MS2 system) in these cells. A *URA3* marker exchange with either *TRP1* or *LEU2* followed by co-localisation experiments with the *NIP1 PDC1* ORF and other glycolytic mRNAs would have been completed here if time had allowed.

6.4. Overall Summary

Therefore, in this thesis significant inroads have been made toward the development of a robust system for the study of endogenous unmodified mRNA in live yeast cells. It can be envisaged that this approach could be used to generate a comprehensive strain collection, wherein a systems biology approach could be used to study the localisation of every mRNA in the yeast cell under a range of conditions. In addition, the question as to how different mRNAs localise to different granules in cells was addressed. While some evidence was acquired suggesting that the level of an mRNA may play a role, it was also shown that it is possible to switch the localisation pattern of an mRNA simply by changing the coding region. Future work will explore whether this effect relies upon RNA based *cis*-acting sequences, or the precise sequence of the nascent protein produced from the RNA.

Bibliography

Abudayyeh, O.O., Gootenberg, J.S., Essletzbichler, P., Han, S., Joung, J., Belanto, J.J., Verdine, V., Cox, D.B., Kellner, M.J., Regev, A. and Lander, E.S., 2017. RNA targeting with CRISPR–Cas13. *Nature*, *550*(7675), pp.280-284.

Adams, J.M. and Cory, S., 1975. Modified nucleosides and bizarre 5′-termini in mouse myeloma mRNA. *Nature*, *255*(5503), pp.28-33.

Adivarahan, S., Livingston, N., Nicholson, B., Rahman, S., Wu, B., Rissland, O.S. and Zenklusen, D., 2018. Spatial organization of single mRNPs at different stages of the gene expression pathway. *Molecular Cell*, *72*(4), pp.727-738.

Advani, V.M. and Ivanov, P., 2020. Stress granule subtypes: an emerging link to neurodegeneration. *Cellular and Molecular Life Sciences*, *77*, pp.4827-4845.

Aizer, A., Kalo, A., Kafri, P., Shraga, A., Ben-Yishay, R., Jacob, A., Kinor, N. and Shav-Tal, Y., 2014. Quantifying mRNA targeting to P-bodies in living human cells reveals their dual role in mRNA decay and storage. *Journal of Cell Science*, *127*(20), pp.4443-4456.

Akopian, D., Shen, K., Zhang, X. and Shan, S.O., 2013. Signal recognition particle: an essential protein-targeting machine. *Annual Review of Biochemistry*, *82*, pp.693-721.

Aly, M.K., Ninomiya, K., Adachi, S., Natsume, T. and Hirose, T., 2019. Two distinct nuclear stress bodies containing different sets of RNA-binding proteins are formed with HSATIII architectural noncoding RNAs upon thermal stress exposure. *Biochemical and Biophysical Research Communications*, *516*(2), pp.419-423.

Anderson, P. and Kedersha, N., 2006. RNA granules. *Journal of Cell Biology*, *172*(6), pp.803-808.

Andreassi, C., Crerar, H. and Riccio, A., 2018. Post-transcriptional processing of mRNA in neurons: the vestiges of the RNA world drive transcriptome diversity. *Frontiers in Molecular Neuroscience*, *11*, p.304.

Andrei, M.A., Ingelfinger, D., Heintzmann, R., Achsel, T., Rivera-Pomar, R. and Lührmann, R., 2005. A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. *RNA*, *11*(5), pp.717-727.

Anton, T., Bultmann, S., Leonhardt, H. and Markaki, Y., 2014. Visualization of specific DNA sequences in living mouse embryonic stem cells with a programmable fluorescent CRISPR/Cas system. *Nucleus*, *5*(2), pp.163-172.

Aoki, S.T., Kershner, A.M., Bingman, C.A., Wickens, M. and Kimble, J., 2016. PGL germ granule assembly protein is a base-specific, single-stranded RNase. *Proceedings of the National Academy of Sciences*, *113*(5), pp.1279-1284.

Arribere, J.A., Doudna, J.A. and Gilbert, W.V., 2011. Reconsidering movement of eukaryotic mRNAs between polysomes and P bodies. *Molecular Cell*, *44*(5), pp.745-758.

Aulas, A., Fay, M.M., Lyons, S.M., Achorn, C.A., Kedersha, N., Anderson, P. and Ivanov, P., 2017. Stress-specific differences in assembly and composition of stress granules and related foci. *Journal of Cell Science*, *130*(5), pp.927-937.

Ayache, J., Bénard, M., Ernoult-Lange, M., Minshall, N., Standart, N., Kress, M. and Weil, D., 2015. P-body assembly requires DDX6 repression complexes rather than decay or Ataxin2/2L complexes. *Molecular Biology of the Cell*, *26*(14), pp.2579-2595.

Babendure, J.R., Adams, S.R. and Tsien, R.Y., 2003. Aptamers switch on fluorescence of triphenylmethane dyes. *Journal of the American Chemical Society*, *125*(48), pp.14716-14717.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W. and Noble, W.S., 2009. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research*, *37*(suppl_2), pp.W202-W208.

Bailey, T.L., Williams, N., Misleh, C. and Li, W.W., 2006. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Research*, *34*(suppl_2), pp.W369-W373.

Banerjee, A.K., 1980. 5'-terminal cap structure in eucaryotic messenger ribonucleic acids. *Microbiological Reviews*, *44*(2), p.175.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A. and Horvath, P., 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, *315*(5819), pp.1709-1712.

Bashirullah, A., Halsell, S.R., Cooperstock, R.L., Kloc, M., Karaiskakis, A., Fisher, W.W., Fu, W., Hamilton, J.K., Etkin, L.D. and Lipshitz, H.D., 1999. Joint action of two RNA degradation pathways controls the timing of maternal transcript elimination at the midblastula transition in Drosophila melanogaster. *The EMBO Journal*, *18*(9), pp.2610-2620.

Batra, R., Nelles, D.A., Pirie, E., Blue, S.M., Marina, R.J., Wang, H., Chaim, I.A., Thomas, J.D., Zhang, N., Nguyen, V. and Aigner, S., 2017. Elimination of toxic microsatellite repeat expansion RNA by RNA-targeting Cas9. *Cell*, *170*(5), pp.899-912.

Bell, D.R., Weber, J.K., Yin, W., Huynh, T., Duan, W. and Zhou, R., 2020. In silico design and validation of high-affinity RNA aptamers targeting epithelial cellular adhesion molecule dimers. *Proceedings of the National Academy of Sciences*, *117*(15), pp.8486-8493.

Bentley, D.L., 2014. Coupling mRNA processing with transcription in time and space. *Nature Reviews Genetics*, *15*(3), pp.163-175.

Bergsten, S.E. and Gavis, E.R., 1999. Role for mRNA localization in translational activation but not spatial restriction of nanos RNA. *Development*, *126*(4), pp.659-669.

Berleth, T., Burri, M., Thoma, G., Bopp, D., Richstein, S., Frigerio, G., Noll, M. and Nüsslein-Volhard, C., 1988. The role of localization of bicoid RNA in organizing the anterior pattern of the Drosophila embryo. *The EMBO Journal*, *7*(6), pp.1749-1756.

Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S.M., Singer, R.H. and Long, R.M., 1998. Localization of ASH1 mRNA particles in living yeast. *Molecular Cell*, *2*(4), pp.437-445.

Besse, F. and Ephrussi, A., 2008. Translational control of localized mRNAs: restricting protein synthesis in space and time. *Nature Reviews Molecular Cell Biology*, *9*(12), pp.971-980.

Bhattacharyya, S.N., Habermacher, R., Martine, U., Closs, E.I. and Filipowicz, W., 2006. Relief of microRNA-mediated translational repression in human cells subjected to stress. *Cell*, *125*(6), pp.1111-1124.

Biamonti, G. and Vourc'h, C., 2010. Nuclear stress bodies. *Cold Spring Harbor Perspectives in Biology*, *2*(6), p.a000695.

Biever, A., Glock, C., Tushev, G., Ciirdaeva, E., Dalmay, T., Langer, J.D. and Schuman, E.M., 2020. Monosomes actively translate synaptic mRNAs in neuronal processes. *Science*, *367*(6477).

Boccaletto, P., Machnicka, M.A., Purta, E., Piątkowski, P., Bagiński, B., Wirecki, T.K., de Crécy-Lagard, V., Ross, R., Limbach, P.A., Kotter, A. and Helm, M., 2018. MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Research*, *46*(D1), pp.D303-D307.

Boël, G., Letso, R., Neely, H., Price, W.N., Wong, K.H., Su, M., Luff, J.D., Valecha, M., Everett, J.K., Acton, T.B. and Xiao, R., 2016. Codon influence on protein expression in E. coli correlates with mRNA levels. *Nature*, *529*(7586), pp.358-363.

Boersma, S., Khuperkar, D., Verhagen, B.M., Sonneveld, S., Grimm, J.B., Lavis, L.D. and Tanenbaum, M.E., 2019. Multi-color single-molecule imaging uncovers extensive heterogeneity in mRNA decoding. *Cell*, *178*(2), pp.458-472.

Boke, E. and Mitchison, T.J., 2017. The balbiani body and the concept of physiological amyloids. *Cell Cycle*, *16*(2), p.153.

Boke, E., Ruer, M., Wühr, M., Coughlin, M., Lemaitre, R., Gygi, S.P., Alberti, S., Drechsel, D., Hyman, A.A. and Mitchison, T.J., 2016. Amyloid-like self-assembly of a cellular compartment. *Cell*, *166*(3), pp.637-650.

Bolotin, A., Quinquis, B., Sorokin, A. and Ehrlich, S.D., 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, *151*(8), pp.2551-2561.

Bonnet, G., Tyagi, S., Libchaber, A. and Kramer, F.R., 1999. Thermodynamic basis of the enhanced specificity of structured DNA probes. *Proceedings of the National Academy of Sciences*, *96*(11), pp.6171-6176.

Bouhedda, F., Autour, A. and Ryckelynck, M., 2018. Light-up RNA aptamers and their cognate fluorogens: from their development to their applications. *International Journal of Molecular Sciences*, *19*(1), p.44.

Bratu, D.P., Cha, B.J., Mhlanga, M.M., Kramer, F.R. and Tyagi, S., 2003. Visualizing the distribution and transport of mRNAs in living cells. *Proceedings of the National Academy of Sciences*, *100*(23), pp.13308-13313.

Brengues, M., Teixeira, D. and Parker, R., 2005. Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. *Science*, *310*(5747), pp.486-489.

Brenner, S., Jacob, F. and Meselson, M., 1961. An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature*, *190*(4776), pp.576-581.

Brocken, D.J., Tark-Dame, M. and Dame, R.T., 2018. dCas9: a versatile tool for epigenome editing. *Current Issues Molecular Biology*, *26*, pp.15-32.

Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V. and Van Der Oost, J., 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*, *321*(5891), pp.960-964.

Buchan, J.R., Muhlrad, D. and Parker, R., 2008. P bodies promote stress granule assembly in Saccharomyces cerevisiae. *The Journal of Cell Biology*, *183*(3), pp.441-455.

Buchan, J.R., Yoon, J.H. and Parker, R., 2011. Stress-specific composition, assembly and kinetics of stress granules in Saccharomyces cerevisiae. *Journal of Cell Science*, *124*(2), pp.228-239.

Carlile, T.M., Rojas-Duran, M.F., Zinshteyn, B., Shin, H., Bartoli, K.M. and Gilbert, W.V., 2014. Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature*, *515*(7525), pp.143-146.

Cawte, A.D., Unrau, P.J. and Rueda, D.S., 2020. Live cell imaging of single RNA molecules with fluorogenic Mango II arrays. *Nature Communications*, *11*(1), pp.1-11.

Chan, A.P., Kloc, M. and Etkin, L.D., 1999. fatvg encodes a new localized RNA that uses a 25-nucleotide element (FVLE1) to localize to the vegetal cortex of Xenopus oocytes. *Development*, *126*(22), pp.4943-4953.

Chan, K.M., Liu, Y.T., Ma, C.H., Jayaram, M. and Sau, S., 2013. The 2 micron plasmid of Saccharomyces cerevisiae: a miniaturized selfish genome with optimized functional competence. *Plasmid*, *70*(1), pp.2-17.

Chang, C.W., Nashchekin, D., Wheatley, L., Irion, U., Dahlgaard, K., Montague, T.G., Hall, J. and St. Johnston, D., 2011. Anterior–Posterior Axis Specification in Drosophila Oocytes: Identification of Novel bicoid and oskar mRNA Localization Factors. *Genetics*, *188*(4), pp.883-896.

Chartrand, P., Meng, X.H., Singer, R.H. and Long, R.M., 1999. Structural elements required for the localization of ASH1 mRNA and of a green fluorescent protein reporter particle in vivo. *Current Biology*, *9*(6), pp.333-338.

Chekulaeva, M., Hentze, M.W. and Ephrussi, A., 2006. Bruno acts as a dual repressor of oskar translation, promoting mRNA oligomerization and formation of silencing particles. *Cell*, *124*(3), pp.521-533.

Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G.W., Park, J., Blackburn, E.H., Weissman, J.S., Qi, L.S. and Huang, B., 2013. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*, *155*(7), pp.1479-1491.

Chen, C., Zhang, H., Broitman, S.L., Reiche, M., Farrell, I., Cooperman, B.S. and Goldman, Y.E., 2013. Dynamics of translation by single ribosomes through mRNA secondary structures. *Nature Structural & Molecular Biology*, *20*(5), pp.582-588.

Chen, C.Y., Gherzi, R., Ong, S.E., Chan, E.L., Raijmakers, R., Pruijn, G.J., Stoecklin, G., Moroni, C., Mann, M. and Karin, M., 2001. AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. *Cell*, *107*(4), pp.451-464.

Chen, J., McSwiggen, D. and Ünal, E., 2018. Single molecule fluorescence in situ hybridization (smFISH) Analysis in budding yeast vegetative growth and meiosis. *Journal of visualized experiments: JoVE*, (135).

Cheng, H., Dufu, K., Lee, C.S., Hsu, J.L., Dias, A. and Reed, R., 2006. Human mRNA export machinery recruited to the 5′ end of mRNA. *Cell*, *127*(7), pp.1389-1400.

Chiodi, I., Biggiogera, M., Denegri, M., Corioni, M., Weighardt, F., Cobianchi, F., Riva, S. and Biamonti, G., 2000. Structure and dynamics of hnRNP-labelled nuclear bodies induced by stress treatments. *Journal of Cell Science*, *113*(22), pp.4043-4053.

Choi, C.K., Vicente-Manzanares, M., Zareno, J., Whitmore, L.A., Mogilner, A. and Horwitz, A.R., 2008. Actin and α-actinin orchestrate the assembly and maturation of nascent adhesions in a myosin II motor-independent manner. *Nature Cell Biology*, *10*(9), pp.1039-1050.

Choi, J., Grosely, R., Prabhakar, A., Lapointe, C.P., Wang, J. and Puglisi, J.D., 2018. How messenger RNA and nascent chain sequences regulate translation elongation. *Annual Review of Biochemistry*, *87*, pp.421-449.

Clark, I., Giniger, E., Ruohola-Baker, H., Jan, L.Y. and Jan, Y.N., 1994. Transient posterior localization of a kinesin fusion protein reflects anteroposterior polarity of the Drosophila oocyte. *Current Biology*, *4*(4), pp.289-300.

Clarke, L. and Carbon, J., 1985. The structure and function of yeast centromeres. *Annual Review of Genetics*, *19*(1), pp.29-55.

Coleman, J.R., Culley, D.E., Chrisler, W.B. and Brockman, F.J., 2007. mRNA-targeted fluorescent in situ hybridization (FISH) of Gram-negative bacteria without template amplification or tyramide signal amplification. *Journal of Microbiological Methods*, *71*(3), pp.246-255.

Colgan, D.F. and Manley, J.L., 1997. Mechanism and regulation of mRNA polyadenylation. *Genes & Development*, *11*(21), pp.2755-2766.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A. and Zhang, F., 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science*, *339*(6121), pp.819-823.

Costello, J., Castelli, L.M., Rowe, W., Kershaw, C.J., Talavera, D., Mohammad-Qureshi, S.S., Sims, P.F., Grant, C.M., Pavitt, G.D., Hubbard, S.J. and Ashe, M.P., 2015. Global mRNA selection mechanisms for translation initiation. *Genome Biology*, *16*(1), pp.1-21.

Cougot, N., Babajko, S. and Séraphin, B., 2004. Cytoplasmic foci are sites of mRNA decay in human cells. *The Journal of Cell Biology*, *165*(1), pp.31-40.

Coullin, P., Roy, L., Pellestor, F., Candelier, J.J., Bed-Hom, B., Guillier-Gencik, Z. and Bernheim, A., 2002. PRINS, the other in situ DNA labeling method useful in cellular biology. *American Journal of Medical Genetics*, *107*(2), pp.127-135.

Courel, M., Clément, Y., Bossevain, C., Foretek, D., Cruchez, O.V., Yi, Z., Bénard, M., Benassy, M.N., Kress, M., Vindry, C. and Ernoult-Lange, M., 2019. GC content shapes mRNA storage and decay in human cells. *ELife*, *8*, p.e49708.

Dahanukar, A. and Wharton, R.P., 1996. The Nanos gradient in Drosophila embryos is generated by translational regulation. *Genes & Development*, *10*(20), pp.2610-2621.

Dahanukar, A., Walker, J.A. and Wharton, R.P., 1999. Smaug, a novel RNA-binding protein that operates a translational switch in Drosophila. *Molecular Cell*, *4*(2), pp.209-218.

Daniel, J.A. and Grant, P.A., 2007. Multi-tasking on chromatin with the SAGA coactivator complexes. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *618*(1-2), pp.135-148.

Darzacq, X., Shav-Tal, Y., De Turris, V., Brody, Y., Shenoy, S.M., Phair, R.D. and Singer, R.H., 2007. In vivo dynamics of RNA polymerase II transcription. *Nature Structural & Molecular Biology*, *14*(9), pp.796-806.

Decker, C.J., Teixeira, D. and Parker, R., 2007. Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in Saccharomyces cerevisiae. *The Journal of Cell Biology*, *179*(3), pp.437-449.

Delanoue, R., Herpers, B., Soetaert, J., Davis, I. and Rabouille, C., 2007. Drosophila Squid/hnRNP helps Dynein switch from a gurken mRNA transport motor to an ultrastructural static anchor in sponge bodies. *Developmental Cell*, *13*(4), pp.523-538.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J. and Charpentier, E., 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, *471*(7340), pp.602-607.

Deshler, J.O., Highett, M.I., Abramson, T. and Schnapp, B.J., 1998. A highly conserved RNA-binding protein for cytoplasmic mRNA localization in vertebrates. *Current Biology*, *8*(9), pp.489-496.

Deveau, H., Barrangou, R., Garneau, J.E., Labonté, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P. and Moineau, S., 2008. Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. *Journal of Bacteriology*, *190*(4), pp.1390-1400.

Dichtl, B. and Keller, W., 2001. Recognition of polyadenylation sites in yeast pre-mRNAs by cleavage and polyadenylation factor. *The EMBO Journal*, *20*(12), pp.3197-3209.

Ding, D., Parkhurst, S.M., Halsell, S.R. and Lipshitz, H.D., 1993. Dynamic Hsp83 RNA localization during Drosophila oogenesis and embryogenesis. *Molecular and Cellular Biology*, *13*(6), pp.3773-3781.

Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R. and Virgin, H.W., 2016. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nature Biotechnology*, *34*(2), pp.184-191.

Doerfel, L.K., Wohlgemuth, I., Kothe, C., Peske, F., Urlaub, H. and Rodnina, M.V., 2013. EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. *Science*, *339*(6115), pp.85-88.

Dolgosheina, E.V., Jeng, S.C., Panchapakesan, S.S.S., Cojocaru, R., Chen, P.S., Wilson, P.D., Hawkins, N., Wiggins, P.A. and Unrau, P.J., 2014. RNA mango aptamer-fluorophore: a bright, high-affinity complex for RNA labeling and tracking. *ACS Chemical Biology*, *9*(10), pp.2412-2420.

Driever, W. and Nüsslein-Volhard, C., 1988. The bicoid protein determines position in the Drosophila embryo in a concentration-dependent manner. *Cell*, *54*(1), pp.95-104.

Drummond, D.R., Armstrong, J. and Colman, A., 1985. The effect of capping and polyadenylation on the stability, movement and translation of synthetic messenger RNAs in Xenopus oocytes. *Nucleic Acids Research*, *13*(20), pp.7375-7394.

du Manoir, S., Speicher, M.R., Joos, S., Schröck, E., Popp, S., Döhner, H., Kovacs, G., Robert-Nicoud, M., Lichter, P. and Cremer, T., 1993. Detection of complete and partial chromosome gains and losses by comparative genomic in situ hybridization. *Human Genetics*, *90*(6), pp.590-610.

Duan, J., Lu, G., Hong, Y., Hu, Q., Mai, X., Guo, J., Si, X., Wang, F. and Zhang, Y., 2018. Live imaging and tracking of genome regions in CRISPR/dCas9 knock-in mice. *Genome Biology*, *19*(1), pp.1-7.

Dumelin, C.E., Chen, Y., Leconte, A.M., Chen, Y.G. and Liu, D.R., 2012. Discovery and biological characterization of geranylated RNA in bacteria. *Nature Chemical Biology*, *8*(11), pp.913-919.

Egger, D., Troxler, M. and Bienz, K., 1994. Light and electron microscopic in situ hybridization: non-radioactive labeling and detection, double hybridization, and combined hybridization-immunocytochemistry. *Journal of Histochemistry & Cytochemistry*, *42*(6), pp.815-822.

Eisenberg, E. and Levanon, E.Y., 2018. A-to-I RNA editing—immune protector and transcriptome diversifier. *Nature Reviews Genetics*, *19*(8), pp.473-490.

Elbaum-Garfinkle, S., Kim, Y., Szczepaniak, K., Chen, C.C.H., Eckmann, C.R., Myong, S. and Brangwynne, C.P., 2015. The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proceedings of the National Academy of Sciences*, *112*(23), pp.7189-7194.

Elisha, Z., Havin, L., Ringel, I. and Yisraeli, J.K., 1995. Vg1 RNA binding protein mediates the association of Vg1 RNA with microtubules in Xenopus oocytes. *The EMBO Journal*, *14*(20), pp.5109-5114.

Ephrussi, A. and Lehmann, R., 1992. Induction of germ cell formation by oskar. *Nature*, *358*(6385), pp.387-392.

Eulalio, A., Rehwinkel, J., Stricker, M., Huntzinger, E., Yang, S.F., Doerks, T., Dorner, S., Bork, P., Boutros, M. and Izaurralde, E., 2007. Target-specific requirements for enhancers of decapping in miRNA-mediated gene silencing. *Genes & Development*, *21*(20), pp.2558-2570.

Fei, J. and Sharma, C.M., 2018. RNA localization in bacteria. *Regulating with RNA in Bacteria and Archaea*, pp.421-439.

Femino, A.M., Fay, F.S., Fogarty, K. and Singer, R.H., 1998. Visualization of single RNA transcripts in situ. *Science*, *280*(5363), pp.585-590.

Ferrandon, D., Elphick, L., Nüsslein-Volhard, C. and St Johnston, D., 1994. Staufen protein associates with the 3′ UTR of bicoid mRNA to form particles that move in a microtubule-dependent manner. *Cell*, *79*(7), pp.1221-1232.

Fischer, T., Rodríguez-Navarro, S., Pereira, G., Rácz, A., Schiebel, E. and Hurt, E., 2004. Yeast centrin Cdc31 is linked to the nuclear mRNA export machinery. *Nature Cell Biology*, *6*(9), pp.840-848.

Forrest, K.M. and Gavis, E.R., 2003. Live imaging of endogenous RNA reveals a diffusion and entrapment mechanism for nanos mRNA localization in Drosophila. *Current Biology*, *13*(14), pp.1159-1168.

Fortes, P., Inada, T., Preiss, T., Hentze, M.W., Mattaj, I.W. and Sachs, A.B., 2000. The yeast nuclear cap binding complex can interact with translation factor eIF4G and mediate translation initiation. *Molecular Cell*, *6*(1), pp.191-196.

Franks, T.M. and Lykke-Andersen, J., 2008. The control of mRNA decapping and P-body formation. *Molecular Cell*, *32*(5), pp.605-615.

Frye, M. and Blanco, S., 2016. Post-transcriptional modifications in development and stem cells. *Development*, *143*(21), pp.3871-3881.

Fu, Y., Rocha, P.P., Luo, V.M., Raviram, R., Deng, Y., Mazzoni, E.O. and Skok, J.A., 2016. CRISPR-dCas9 and sgRNA scaffolds enable dual-colour live imaging of satellite sequences and repeat-enriched individual loci. *Nature Communications*, *7*(1), pp.1-8.

Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M. and Joung, J.K., 2014. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nature Biotechnology*, *32*(3), pp.279-284.

Fuda, N.J., Ardehali, M.B. and Lis, J.T., 2009. Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature*, *461*(7261), pp.186-192.

Fujimura, K., Sasaki, A.T. and Anderson, P., 2012. Selenite targets eIF4E-binding protein-1 to inhibit translation initiation and induce the assembly of non-canonical stress granules. *Nucleic Acids Research*, *40*(16), pp.8099-8110.

Gall, J.G. and Pardue, M.L., 1969. Formation and detection of RNA-DNA hybrid molecules in cytological preparations. *Proceedings of the National Academy of Sciences*, *63*(2), pp.378-383.

Garcia, J.F. and Parker, R., 2015. MS2 coat proteins bound to yeast mRNAs block 5′ to 3′ degradation and trap mRNA decay products: implications for the localization of mRNAs by MS2-MCP system. *RNA*, *21*(8), pp.1393-1395.

Garcia, J.F. and Parker, R., 2016. Ubiquitous accumulation of 3′ mRNA decay fragments in Saccharomyces cerevisiae mRNAs with chromosomally integrated MS2 arrays. *RNA*, *22*(5), pp.657-659.

Garneau, J.E., Dupuis, M.È., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H. and Moineau, S., 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, *468*(7320), pp.67-71.

Garneau, N.L., Wilusz, J. and Wilusz, C.J., 2007. The highways and byways of mRNA decay. *Nature reviews Molecular Cell Biology*, *8*(2), pp.113-126.

Gatfield, D., Unterholzner, L., Ciccarelli, F.D., Bork, P. and Izaurralde, E., 2003. Nonsense-mediated mRNA decay in Drosophila: at the intersection of the yeast and mammalian pathways. *The EMBO Journal*, *22*(15), pp.3960-3970.

Ghaemmaghami, S., Huh, W.K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K. and Weissman, J.S., 2003. Global analysis of protein expression in yeast. *Nature*, *425*(6959), pp.737-741.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A. and Smith, H.O., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, *6*(5), pp.343-345.

Gietz, R.D. and Akio, S., 1988. New yeast-Escherichia coli shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites. *Gene*, *74*(2), pp.527-534.

Gietz, R.D. and Schiestl, R.H., 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols*, *2*(1), pp.31-34.

Gilbert, W.V., Zhou, K., Butler, T.K. and Doudna, J.A., 2007. Cap-independent translation is required for starvation-induced differentiation in yeast. *Science*, *317*(5842), pp.1224-1227.

Gingold, H. and Pilpel, Y., 2011. Determinants of translation efficiency and accuracy. *Molecular Systems Biology*, *7*(1), p.481.

Gonzalez, I., Buonomo, S.B., Nasmyth, K. and von Ahsen, U., 1999. ASH1 mRNA localization in yeast involves multiple secondary structural elementsand Ash1 protein translation. *Current Biology*, *9*(6), pp.337-340.

Gott, J.M. and Emeson, R.B., 2000. Functions and mechanisms of RNA editing. *Annual Review of Genetics*, *34*(1), pp.499-531.

Groušl, T., Ivanov, P., Frydlová, I., Vašicová, P., Janda, F., Vojtová, J., Malínská, K., Malcová, I., Nováková, L., Janošková, D. and Valášek, L., 2009. Robust heat shock induces eIF2α-phosphorylation-independent assembly of stress granules containing eIF3 and 40S ribosomal subunits in budding yeast, Saccharomyces cerevisiae. *Journal of Cell Science*, *122*(12), pp.2078-2088.

Grousl, T., Ivanov, P., Malcova, I., Pompach, P., Frydlova, I., Slaba, R., Senohrabkova, L., Novakova, L. and Hasek, J., 2013. Heat shock-induced accumulation of translation elongation and termination factors precedes assembly of stress granules in S. cerevisiae. *PloS One*, *8*(2), p.e57083.

Grünwald, D. and Singer, R.H., 2010. In vivo imaging of labelled endogenous β-actin mRNA during nucleocytoplasmic transport. *Nature*, *467*(7315), pp.604-607.

Guo, Y.E., Manteiga, J.C., Henninger, J.E., Sabari, B.R., Dall'Agnese, A., Hannett, N.M., Spille, J.H., Afeyan, L.K., Zamudio, A.V., Shrinivas, K. and Abraham, B.J., 2019. Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature*, *572*(7770), pp.543-548.

Haimovich, G., Zabezhinsky, D., Haas, B., Slobodin, B., Purushothaman, P., Fan, L., Levin, J.Z., Nusbaum, C. and Gerst, J.E., 2016. Use of the MS2 aptamer and coat protein for RNA localization in yeast: A response to "MS2 coat proteins bound to yeast mRNAs block 5′ to 3′ degradation and trap mRNA decay products: implications for the localization of mRNAs by MS2-MCP system". *RNA*, *22*(5), pp.660-666.

Halstead, J.M., Lionnet, T., Wilbertz, J.H., Wippich, F., Ephrussi, A., Singer, R.H. and Chao, J.A., 2015. An RNA biosensor for imaging the first round of translation from single cells to living animals. *Science*, *347*(6228), pp.1367-1671.

Hanazawa, M., Yonetani, M. and Sugimoto, A., 2011. PGL proteins self associate and bind RNPs to mediate germ granule assembly in C. elegans. *Journal of Cell Biology*, *192*(6), pp.929-937.

Harper, M.E. and Saunders, G.F., 1981. Localization of single copy DNA sequences on G-banded human chromosomes by in situ hybridization. *Chromosoma*, *83*(3), pp.431-439.

Havin, L., Git, A., Elisha, Z., Oberman, F., Yaniv, K., Schwartz, S.P., Standart, N. and Yisraeli, J.K., 1998. RNA-binding protein conserved in both microtubule-and microfilament-based RNA localization. *Genes & Development*, *12*(11), pp.1593-1598.

Heinrich, S., Sidler, C.L., Azzalin, C.M. and Weis, K., 2017. Stem–loop RNA labeling can affect nuclear and cytoplasmic mRNA processing. *RNA*, *23*(2), pp.134-141.

Herbert, A.D., Carr, A.M. and Hoffmann, E., 2014. FindFoci: a focus detection algorithm with automated parameter training that closely matches human assignments, reduces human inconsistencies and increases speed of analysis. *PloS One*, *9*(12), p.e114749.

Hermesh, O. and Jansen, R.P., 2013. Take the (RN) A-train: localization of mRNA to the endoplasmic reticulum. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1833*(11), pp.2519-2525.

Hirashima, T., Tanaka, R., Yamaguchi, M. and Yoshida, H., 2018. The ABD on the nascent polypeptide and PH domain are required for the precise Anillin localization in Drosophila syncytial blastoderm. *Scientific Reports*, *8*(1), pp.1-9.

Hirose, Y. and Manley, J.L., 1998. RNA polymerase II is an essential mRNA polyadenylation factor. *Nature*, *395*(6697), pp.93-96.

Hirose, Y., Tacke, R. and Manley, J.L., 1999. Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. *Genes & Development*, *13*(10), pp.1234-1239.

Hocine, S., Singer, R.H. and Grünwald, D., 2010. RNA processing and export. *Cold Spring Harbor Perspectives in Biology*, *2*(12), p.a000752.

Hoefler, H., Childers, H., Montminy, M.R., Lechan, R.M., Goodman, R.H. and Wolfe, H.J., 1986. In situ hybridization methods for the detection of somatostatin mRNA in tissue sections using antisense RNA probes. *The Histochemical Journal*, *18*(11-12), pp.597-604.

Hoernes, T.P., Clementi, N., Faserl, K., Glasner, H., Breuker, K., Lindner, H., Hüttenhofer, A. and Erlacher, M.D., 2016. Nucleotide modifications within bacterial messenger RNAs regulate their translation and are able to rewire the genetic code. *Nucleic Acids Research*, *44*(2), pp.852-862.

Hostettler, L., Grundy, L., Käser-Pébernard, S., Wicky, C., Schafer, W.R. and Glauser, D.A., 2017. The bright fluorescent protein mNeonGreen facilitates protein expression analysis in vivo. *G3: Genes, Genomes, Genetics*, *7*(2), pp.607-615.

Howe, K.J., Kane, C.M. and Ares, M., 2003. Perturbation of transcription elongation influences the fidelity of internal exon inclusion in Saccharomyces cerevisiae. *RNA*, *9*(8), pp.993-1006.

Hoyle, N.P., Castelli, L.M., Campbell, S.G., Holmes, L.E. and Ashe, M.P., 2007. Stress-dependent relocalization of translationally primed mRNPs to cytoplasmic granules that are kinetically and spatially distinct from P-bodies. *The Journal of Cell Biology*, *179*(1), pp.65-74.

Hsu, C.L. and Stevens, A., 1993. Yeast cells lacking 5'--> 3'exoribonuclease 1 contain mRNA species that are poly (A) deficient and partially lack the 5'cap structure. *Molecular and Cellular Biology*, *13*(8), pp.4826-4835.

Hsu, P.D., Lander, E.S. and Zhang, F., 2014. Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, *157*(6), pp.1262-1278.

Hubstenberger, A., Courel, M., Bénard, M., Souquere, S., Ernoult-Lange, M., Chouaib, R., Yi, Z., Morlot, J.B., Munier, A., Fradet, M. and Daunesse, M., 2017. P-body purification reveals the condensation of repressed mRNA regulons. *Molecular Cell*, *68*(1), pp.144-157.

Hüttelmaier, S., Zenklusen, D., Lederer, M., Dictenberg, J., Lorenz, M., Meng, X., Bassell, G.J., Condeelis, J. and Singer, R.H., 2005. Spatial regulation of β-actin translation by Src-dependent phosphorylation of ZBP1. *Nature*, *438*(7067), pp.512-515.

Huynh, N., Depner, N., Larson, R. and King-Jones, K., 2020. A versatile toolkit for CRISPR-Cas13-based RNA manipulation in Drosophila. *Genome Biology*, *21*(1), pp.1-29.

Iborra, F.J., Kimura, H. and Cook, P.R., 2004. The functional organization of mitochondrial genomes in human cells. *BMC Biology*, *2*(1), pp.1-14.

Inada, T. and Aiba, H., 2005. Translation of aberrant mRNAs lacking a termination codon or with a shortened 3′-UTR is repressed after initiation in yeast. *The EMBO Journal*, *24*(8), pp.1584-1595.

Irastortza-Olaziregi, M. and Amster-Choder, O., 2020. RNA localization in prokaryotes: Where, when, how, and why. *Wiley Interdisciplinary Reviews: RNA*, p.e1615.

Ishigaki, Y., Li, X., Serin, G. and Maquat, L.E., 2001. Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. *Cell*, *106*(5), pp.607-617.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. and Nakata, A., 1987. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *Journal of Bacteriology*, *169*(12), pp.5429-5433.

Ito, K. and Chiba, S., 2013. Arrest peptides: cis-acting modulators of translation. *Annual Review of Biochemistry*, *82*, pp.171-202.

Jain, S., Wheeler, J.R., Walters, R.W., Agrawal, A., Barsic, A. and Parker, R., 2016. ATPase-modulated stress granules contain a diverse proteome and substructure. *Cell*, *164*(3), pp.487-498.

Jansen, R., Embden, J.D.V., Gaastra, W. and Schouls, L.M., 2002. Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*, *43*(6), pp.1565-1575.

Jansen, R.P., 2001. mRNA localization: message on the move. *Nature reviews Molecular Cell Biology*, *2*(4), pp.247-256.

Jasin, M. and Haber, J.E., 2016. The democratization of gene editing: Insights from site-specific cleavage and double-strand break repair. *DNA Repair*, *44*, pp.6-16.

Jeng, S.C., Chan, H.H., Booy, E.P., McKenna, S.A. and Unrau, P.J., 2016. Fluorophore ligand binding and complex stabilization of the RNA Mango and RNA Spinach aptamers. *RNA*, *22*(12), pp.1884-1892.

Jensen, E., 2014. Technical review: In situ hybridization. *The Anatomical Record*, *297*(8), pp.1349-1353.

Jeske, M., Meyer, S., Temme, C., Freudenreich, D. and Wahle, E., 2006. Rapid ATP-dependent deadenylation of nanos mRNA in a cell-free system from Drosophila embryos. *Journal of Biological Chemistry*, *281*(35), pp.25124-25133.

Jin, L. and Lloyd, R.V., 1997. In situ hybridization: methods and applications. *Journal of Clinical Laboratory Analysis*, *11*(1), pp.2-9.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E., 2012. A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science*, *337*(6096), pp.816-821.

Jinek, M., Jiang, F., Taylor, D.W., Sternberg, S.H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S. and Kaplan, M., 2014. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science*, *343*(6176).

Jolly, C., Metz, A., Govin, J., Vigneron, M., Turner, B.M., Khochbin, S. and Vourc'h, C., 2004. Stress-induced transcription of satellite III repeats. *The Journal of Cell Biology*, *164*(1), pp.25-33.

Jonas, S. and Izaurralde, E., 2013. The role of disordered protein regions in the assembly of decapping complexes and RNP granules. *Genes & Development*, *27*(24), pp.2628-2641.

Jones, J.R. and Macdonald, P.M., 2007. Oskar controls morphology of polar granules and nuclear bodies in Drosophila. *Development*, *134*(2), pp.233-236.

Jourdain, A.A., Boehm, E., Maundrell, K. and Martinou, J.C., 2016. Mitochondrial RNA granules: compartmentalizing mitochondrial gene expression. *Journal of Cell Biology*, *212*(6), pp.611-614.

Jourdain, A.A., Koppen, M., Wydro, M., Rodley, C.D., Lightowlers, R.N., Chrzanowska-Lightowlers, Z.M. and Martinou, J.C., 2013. GRSF1 regulates RNA processing in mitochondrial RNA granules. *Cell Metabolism*, *17*(3), pp.399-410.

Kamenova, I., Mukherjee, P., Conic, S., Mueller, F., El-Saafin, F., Bardot, P., Garnier, J.M., Dembele, D., Capponi, S., Timmers, H.M. and Vincent, S.D., 2019. Co-translational assembly of mammalian nuclear multisubunit complexes. *Nature Communications*, *10*(1), pp.1-15.

Kato, M., Han, T.W., Xie, S., Shi, K., Du, X., Wu, L.C., Mirzaei, H., Goldsmith, E.J., Longgood, J., Pei, J. and Grishin, N.V., 2012. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell*, *149*(4), pp.753-767.

Kawasaki, I., Shim, Y.H., Kirchner, J., Kaminker, J., Wood, W.B. and Strome, S., 1998. PGL-1, a predicted RNA-binding component of germ granules, is essential for fertility in C. elegans. *Cell*, *94*(5), pp.635-645.

Kedersha, N. and Anderson, P., 2002. Stress granules: sites of mRNA triage that regulate mRNA stability and translatability. *Biochemical Society Transactions*, *30*(6), pp.963-969.

Kedersha, N., Chen, S., Gilks, N., Li, W., Miller, I.J., Stahl, J. and Anderson, P., 2002. Evidence that ternary complex (eIF2-GTP-tRNAi Met)–deficient preinitiation complexes are core constituents of mammalian stress granules. *Molecular Biology of the Cell*, *13*(1), pp.195-210.

Kedersha, N., Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fritzler, M.J., Scheuner, D., Kaufman, R.J., Golan, D.E. and Anderson, P., 2005. Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *The Journal of Cell Biology*, *169*(6), pp.871-884.

Kedersha, N.L., Gupta, M., Li, W., Miller, I. and Anderson, P., 1999. RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2α to the assembly of mammalian stress granules. *The Journal of Cell Biology*, *147*(7), pp.1431-1442.

Kershaw, C., Nelson, M.G., Lui, J., Bates, C.P., Jennings, M.D., Hubbard, S.J., Ashe, M.P. and Grant, C.M., 2020. Integrated multi-omics reveals common properties underlying stress granule and P-body formation. *BioRxiv*.

Khabar, K.S., 2017. Hallmarks of cancer and AU-rich elements. *Wiley Interdisciplinary Reviews: RNA*, *8*(1), p.e1368.

Khong, A., Matheny, T., Jain, S., Mitchell, S.F., Wheeler, J.R. and Parker, R., 2017. The stress granule transcriptome reveals principles of mRNA accumulation in stress granules. *Molecular Cell*, *68*(4), pp.808-820.

Kim, S.H., Vieira, M., Shim, J.Y., Choi, H. and Park, H.Y., 2019. Recent progress in single-molecule studies of mRNA localization in vivo. *RNA Biology*, *16*(9), pp.1108-1118.

Kislauskis, E.H., Zhu, X. and Singer, R.H., 1994. Sequences responsible for intracellular localization of beta-actin messenger RNA also affect cell phenotype. *The Journal of Cell Biology*, *127*(2), pp.441-451.

Köhler, A. and Hurt, E., 2007. Exporting RNA from the nucleus to the cytoplasm. *Nature Reviews Molecular Cell Biology*, *8*(10), pp.761-773.

Kohrmann, M., Luo, M., Kaether, C., DesGroseillers, L., Dotti, C.G. and Kiebler, M.A., 1999. Microtubule-dependent recruitment of Staufen-green fluorescent protein into large RNA-containing granules and subsequent dendritic transport in living hippocampal neurons. *Molecular Biology of the Cell*, *10*(9), pp.2945-2953.

Komarnitsky, P., Cho, E.J. and Buratowski, S., 2000. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes & Development*, *14*(19), pp.2452-2460.

Komminoth, P., 1992. Digoxigenin as an alternative probe labeling for in situ hybridization. *Diagnostic Molecular Pathology: the American Journal of Surgical Pathology, Part B*, *1*(2), pp.142-150.

Konarska, M.M., Padgett, R.A. and Sharp, P.A., 1984. Recognition of cap structure in splicing in vitro of mRNA precursors. *Cell*, *38*(3), pp.731-736.

Kontos, H., Napthine, S. and Brierley, I., 2001. Ribosomal pausing at a frameshifter RNA pseudoknot is sensitive to reading phase but shows little correlation with frameshift efficiency. *Molecular and ellular biology*, *21*(24), pp.8657-8670.

Kozak, M., 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Research*, *12*(2), pp.857-872.

Kuchinke, U., Grawe, F. and Knust, E., 1998. Control of spindle orientation in Drosophila by the Par-3-related PDZ-domain protein Bazooka. *Current Biology*, *8*(25), pp.1357-1365.

Kugler, J.M. and Lasko, P., 2009. Localization, anchoring and translational control of oskar, gurken, bicoid and nanos mRNA during Drosophila oogenesis. *Fly*, *3*(1), pp.15-28.

Kukat, C., Davies, K.M., Wurm, C.A., Spåhr, H., Bonekamp, N.A., Kühl, I., Joos, F., Polosa, P.L., Park, C.B., Posse, V. and Falkenberg, M., 2015. Cross-strand binding of TFAM to a single mtDNA molecule forms the mitochondrial nucleoid. *Proceedings of the National Academy of Sciences*, *112*(36), pp.11288-11293.

Kung, C.P., Maggi Jr, L.B. and Weber, J.D., 2018. The role of RNA editing in cancer development and metabolic disorders. *Frontiers in Endocrinology*, *9*, p.762.

Lahtvee, P.J., Sánchez, B.J., Smialowska, A., Kasvandik, S., Elsemman, I.E., Gatto, F. and Nielsen, J., 2017. Absolute quantification of protein and mRNA abundances demonstrate variability in gene-specific translation efficiency in yeast. *Cell Systems*, *4*(5), pp.495-504.

Lange, S., Katayama, Y., Schmid, M., Burkacky, O., Bräuchle, C., Lamb, D.C. and Jansen, R.P., 2008. Simultaneous transport of different localized mRNA species revealed by live-cell imaging. *Traffic*, *9*(8), pp.1256-1267.

Langer-Safer, P.R., Levine, M. and Ward, D.C., 1982. Immunological method for mapping genes on Drosophila polytene chromosomes. *Proceedings of the National Academy of Sciences*, *79*(14), pp.4381-4385.

Larson, D.R., Zenklusen, D., Wu, B., Chao, J.A. and Singer, R.H., 2011. Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science*, *332*(6028), pp.475-478.

Lasko, P., 2012. mRNA localization and translational control in Drosophila oogenesis. *Cold Spring Harbor Perspectives in Biology*, *4*(10), p.a012294.

Lawless, C., Holman, S.W., Brownridge, P., Lanthaler, K., Harman, V.M., Watkins, R., Hammond, D.E., Miller, R.L., Sims, P.F., Grant, C.M. and Eyers, C.E., 2016. Direct and absolute quantification of over 1800 yeast proteins via selected reaction monitoring. *Molecular & Cellular Proteomics*, *15*(4), pp.1309-1322.

Lawrence, J.B. and Singer, R.H., 1986. Intracellular localization of messenger RNAs for cytoskeletal proteins. *Cell*, *45*(3), pp.407-415.

Le Hir, H., Izaurralde, E., Maquat, L.E. and Moore, M.J., 2000. The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon–exon junctions. *The EMBO Journal*, *19*(24), pp.6860-6869.

Lee, Y. and Rio, D.C., 2015. Mechanisms and regulation of alternative pre-mRNA splicing. *Annual Review of Biochemistry*, *84*, pp.291-323.

Lemieux, C. and Bachand, F., 2009. Cotranscriptional recruitment of the nuclear poly (A)-binding protein Pab2 to nascent transcripts and association with translating mRNPs. *Nucleic Acids Research*, *37*(10), pp.3418-3430.

Leung, K.M., Lu, B., Wong, H.H.W., Lin, J.Q., Turner-Bridger, B. and Holt, C.E., 2018. Cue-polarized transport of β-actin mRNA depends on 3′ UTR and microtubules in live growth cones. *Frontiers in Cellular Neuroscience*, *12*, p.300.

Li, P., Pomianowski, P., DiMaio, M.S., Florio, J.R., Rossi, M.R., Xiang, B., Xu, F., Yang, H., Geng, Q., Xie, J. and Mahoney, M.J., 2011. Genomic characterization of prenatally detected chromosomal structural abnormalities using oligonucleotide array comparative genomic hybridization. *American Journal of Medical Genetics Part A*, *155*(7), pp.1605-1615.

Liao, G. and Liu, G., 2011. How and why does Dia1 mRNA localize?. *Communicative & Integrative Biology*, *4*(5), pp.560-562.

Lin, J., Xu, R., Wu, X., Shen, Y. and Li, Q.Q., 2017. Role of cleavage and polyadenylation specificity factor 100: anchoring poly (A) sites and modulating transcription termination. *The Plant Journal*, *91*(5), pp.829-839.

Lin, Y., Protter, D.S., Rosen, M.K. and Parker, R., 2015. Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. *Molecular Cell*, *60*(2), pp.208-219.

Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L. and Hannon, G.J., 2004. Argonaute2 is the catalytic engine of mammalian RNAi. *Science*, *305*(5689), pp.1437-1441.

Liu, J., Valencia-Sanchez, M.A., Hannon, G.J. and Parker, R., 2005. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nature Cell Biology*, *7*(7), pp.719-723.

Liu, X.S., Wu, H., Ji, X., Stelzer, Y., Wu, X., Czauderna, S., Shu, J., Dadon, D., Young, R.A. and Jaenisch, R., 2016. Editing DNA methylation in the mammalian genome. *Cell*, *167*(1), pp.233-247.

Livak, K.J. and Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *Methods*, *25*(4), pp.402-408.

Long, A.A., 1998. In-situ polymerase chain reaction: foundation of the technology and today's options. *European journal of histochemistry: EJH*, *42*(2), pp.101-109.

Lõoke, M., Kristjuhan, K. and Kristjuhan, A., 2011. Extraction of genomic DNA from yeasts for PCR-based applications. *Biotechniques*, *50*(5), pp.325-328.

Lui, J., Castelli, L.M., Pizzinga, M., Simpson, C.E., Hoyle, N.P., Bailey, K.L., Campbell, S.G. and Ashe, M.P., 2014. Granules harboring translationally active mRNAs provide a platform for P-body formation following stress. *Cell Reports*, *9*(3), pp.944-954.

Lykke-Andersen, J. and Wagner, E., 2005. Recruitment and activation of mRNA decay enzymes by two ARE-mediated decay activation domains in the proteins TTP and BRF-1. *Genes & Development*, *19*(3), pp.351-361.

Ma, N., Matsunaga, S., Takata, H., Ono-Maniwa, R., Uchiyama, S. and Fukui, K., 2007. Nucleolin functions in nucleolus formation and chromosome congression. *Journal of Cell Science*, *120*(12), pp.2091-2105.

Ma, X., Zhang, Q., Zhu, Q., Liu, W., Chen, Y., Qiu, R., Wang, B., Yang, Z., Li, H., Lin, Y. and Xie, Y., 2015. A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. *Molecular Plant*, *8*(8), pp.1274-1284.

Mager, W.H. and Planta, R.J., 1991. Coordinate expression of ribosomal protein genes in yeast as a function of cellular growth rate. *Molecular Mechanisms of Cellular Growth*, pp.181-187.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. and Church, G.M., 2013. RNA-guided human genome engineering via Cas9. *Science*, *339*(6121), pp.823-826.

Malinovska, L., Kroschwald, S. and Alberti, S., 2013. Protein disorder, prion propensities, and self-organizing macromolecular collectives. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1834*(5), pp.918-931.

Mans, R., van Rossum, H.M., Wijsman, M., Backx, A., Kuijpers, N.G., van den Broek, M., Daran-Lapujade, P., Pronk, J.T., van Maris, A.J. and Daran, J.M.G., 2015. CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in Saccharomyces cerevisiae. *FEMS Yeast Research*, *15*(2).

Maquat, L.E., Tarn, W.Y. and Isken, O., 2010. The pioneer round of translation: features and functions. *Cell*, *142*(3), pp.368-374.

Markmiller, S., Soltanieh, S., Server, K.L., Mak, R., Jin, W., Fang, M.Y., Luo, E.C., Krach, F., Yang, D., Sen, A. and Fulzele, A., 2018. Context-dependent and disease-specific diversity in protein interactions within stress granules. *Cell*, *172*(3), pp.590-604.

Martens, K.J., van Beljouw, S.P., van der Els, S., Vink, J.N., Baas, S., Vogelaar, G.A., Brouns, S.J., van Baarlen, P., Kleerebezem, M. and Hohlbein, J., 2019. Visualisation of dCas9 target search in vivo using an open-microscopy framework. *Nature Communications*, *10*(1), pp.1-11.

Martin, G. and Keller, W., 1998. Tailing and 3'-end labeling of RNA with yeast poly (A) polymerase and various nucleotides. *RNA*, *4*(2), pp.226-230.

Martin, K.C. and Ephrussi, A., 2009. mRNA localization: gene expression in the spatial dimension. *Cell*, *136*(4), pp.719-730.

Martin, R.M., Rino, J., Carvalho, C., Kirchhausen, T. and Carmo-Fonseca, M., 2013. Live-cell visualization of pre-mRNA splicing with single-molecule sensitivity. *Cell Reports*, *4*(6), pp.1144-1155.

Masuda, S., Das, R., Cheng, H., Hurt, E., Dorman, N. and Reed, R., 2005. Recruitment of the human TREX complex to mRNA during splicing. *Genes & Development*, *19*(13), pp.1512-1517.

Mateju, D., Eichenberger, B., Voigt, F., Eglinger, J., Roth, G. and Chao, J.A., 2020. Single-molecule imaging reveals translation of mRNAs localized to stress granules. *Cell*, *183*(7), pp.1801-1812.

Matheny, T., Rao, B.S. and Parker, R., 2019. Transcriptome-wide comparison of stress granules and P-bodies reveals that translation plays a major role in RNA partitioning. *Molecular and Cellular Biology*, *39*(24).

McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M. and Bentley, D.L., 1997. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature*, *385*(6614), pp.357-361.

Merianda, T.T., Lin, A.C., Lam, J.S., Vuppalanchi, D., Willis, D.E., Karin, N., Holt, C.E. and Twiss, J.L., 2009. A functional equivalent of endoplasmic reticulum and Golgi in axons for secretion of locally synthesized proteins. *Molecular and Cellular Neuroscience*, *40*(2), pp.128-142.

Merz, C., Urlaub, H., Will, C.L. and Lührmann, R., 2007. Protein composition of human mRNPs spliced in vitro and differential requirements for mRNP protein recruitment. *RNA*, *13*(1), pp.116-128.

Metschnikoff, E., 1865. Veber die Entwicklung der Cecidomyienlarven aus dem Pseudovum. *Arch. furr Naturg. Bd*, *1*.

Mingle, L.A., Okuhama, N.N., Shi, J., Singer, R.H., Condeelis, J. and Liu, G., 2005. Localization of all seven messenger RNAs for the actin-polymerization nucleator Arp2/3 complex in the protrusions of fibroblasts. *Journal of Cell Science*, *118*(11), pp.2425-2433.

Minshall, N., Kress, M., Weil, D. and Standart, N., 2009. Role of p54 RNA helicase activity and its C-terminal domain in translational repression, P-body localization and assembly. *Molecular Biology of the Cell*, *20*(9), pp.2464-2472.

Mojica, F.J., García-Martínez, J. and Soria, E., 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution*, *60*(2), pp.174-182.

Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A.P., Kim, H.J., Mittag, T. and Taylor, J.P., 2015. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell*, *163*(1), pp.123-133.

Monahan, Z., Ryan, V.H., Janke, A.M., Burke, K.A., Rhoads, S.N., Zerze, G.H., O'Meally, R., Dignon, G.L., Conicella, A.E., Zheng, W. and Best, R.B., 2017. Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *The EMBO Journal*, *36*(20), pp.2951-2967.

Monroy-Contreras, R. and Vaca, L., 2011. Molecular beacons: powerful tools for imaging RNA in living cells. *Journal of Nucleic Acids*, *2011*.

Moore, M.J. and Sharp, P.A., 1993. Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature*, *365*(6444), pp.364-368.

Morales-Polanco, F., Bates, C., Lui, J., Casson, J., Solari, C.A., Pizzinga, M., Forte, G., Griffin, C., Garner, K.E., Burt, H.E. and Dixon, H.L., 2021. Core Fermentation (CoFe) granules focus coordinated glycolytic mRNA localization and translation to fuel glucose fermentation. *IScience*, *24*(2), p.102069.

Morisaki, T., Lyon, K., DeLuca, K.F., DeLuca, J.G., English, B.P., Zhang, Z., Lavis, L.D., Grimm, J.B., Viswanathan, S., Looger, L.L. and Lionnet, T., 2016. Real-time quantification of single RNA translation dynamics in living cells. *Science*, *352*(6292), pp.1425-1429.

Morris, G.E., 2008. The cajal body. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1783*(11), pp.2108-2115.

Mowry, K.L. and Melton, D.A., 1992. Vegetal messenger RNA localization directed by a 340-nt RNA sequence element in Xenopus oocytes. *Science*, *255*(5047), pp.991-994.

Mugler, C.F., Hondele, M., Heinrich, S., Sachdev, R., Vallotton, P., Koek, A.Y., Chan, L.Y. and Weis, K., 2016. ATPase activity of the DEAD-box protein Dhh1 controls processing body formation. *ELife*, *5*, p.e18746.

Muhlrad, D. and Parker, R., 2005. The yeast EDC1 mRNA undergoes deadenylation-independent decapping stimulated by Not2p, Not4p, and Not5p. *The EMBO Journal*, *24*(5), pp.1033-1045.

Muto, H., Nakatogawa, H. and Ito, K., 2006. Genetically encoded but nonpolypeptide prolyl-tRNA functions in the A site for SecM-mediated ribosomal stall. *Molecular Cell*, *22*(4), pp.545-552.

Nachtergaele, S. and He, C., 2017. The emerging biology of RNA post-transcriptional modifications. *RNA Biology*, *14*(2), pp.156-163.

Nelles, D.A., Fang, M.Y., O'Connell, M.R., Xu, J.L., Markmiller, S.J., Doudna, J.A. and Yeo, G.W., 2016. Programmable RNA tracking in live cells with CRISPR/Cas9. *Cell*, *165*(2), pp.488-496.

Neve, J., Patel, R., Wang, Z., Louey, A. and Furger, A.M., 2017. Cleavage and polyadenylation: Ending the message expands gene regulation. *RNA Biology*, *14*(7), pp.865-890.

Nissan, T. and Parker, R., 2008. Analyzing P-bodies in Saccharomyces cerevisiae. *Methods in Enzymology*, *448*, pp.507-520.

Noma, A., Smith, C.S., Huisman, M., Martin, R.M., Moore, M.J. and Grunwald, D., 2017. Single-molecule FISH in Drosophila muscle reveals location dependent mRNA composition of megaRNPs. *BioRxiv*, p.156091.

Norvell, A., Debec, A., Finch, D., Gibson, L. and Thoma, B., 2005. Squid is required for efficient posterior localization of oskar mRNA during Drosophila oogenesis. *Development Genes and Evolution*, *215*(7), pp.340-349.

O'Connell, M.R., Oakes, B.L., Sternberg, S.H., East-Seletsky, A., Kaplan, M. and Doudna, J.A., 2014. Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature*, *516*(7530), pp.263-266.

Ohn, T., Kedersha, N., Hickman, T., Tisdale, S. and Anderson, P., 2008. A functional RNAi screen links O-GlcNAc modification of ribosomal proteins to stress granule and processing body assembly. *Nature Cell Biology*, *10*(10), pp.1224-1231.

Otsuka, H., Fukao, A., Funakami, Y., Duncan, K.E. and Fujiwara, T., 2019. Emerging evidence of translational control by AU-rich element-binding proteins. *Frontiers in Genetics*, *10*, p.332.

Palacios, I.M., 2007, April. How does an mRNA find its way? Intracellular localisation of transcripts. In *Seminars in cell & developmental biology* (Vol. 18, No. 2, pp. 163-170). Academic Press.

Park, H.Y., Trcek, T., Wells, A.L., Chao, J.A. and Singer, R.H., 2012. An unbiased analysis method to quantify mRNA localization reveals its correlation with cell motility. *Cell Reports*, *1*(2), pp.179-184.

Parker, R. and Sheth, U., 2007. P bodies and the control of mRNA translation and degradation. *Molecular Cell*, *25*(5), pp.635-646.

Parton, R.M., Davidson, A., Davis, I. and Weil, T.T., 2014. Subcellular mRNA localisation at a glance. *Journal of Cell Science*, *127*(10), pp.2127-2133.

Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T.M. and Pozniakovski, A., 2015. A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell*, *162*(5), pp.1066-1077.

Peabody, D.S. and Lim, F., 1996. Complementation of RNA binding site mutations in MS2 coat protein heterodimers. *Nucleic Acids Research*, *24*(12), pp.2352-2359.

Pickar-Oliver, A. and Gersbach, C.A., 2019. The next generation of CRISPR–Cas technologies and applications. *Nature reviews Molecular Cell Biology*, *20*(8), pp.490-507.

Pitchiaya, S., Mourao, M.D., Jalihal, A.P., Xiao, L., Jiang, X., Chinnaiyan, A.M., Schnell, S. and Walter, N.G., 2019. Dynamic recruitment of single RNAs to processing bodies depends on RNA functionality. *Molecular Cell*, *74*(3), pp.521-533.

Pitt, J.N., Schisa, J.A. and Priess, J.R., 2000. P granules in the germ cells of Caenorhabditis elegans adults are associated with clusters of nuclear pores and contain RNA. *Developmental Biology*, *219*(2), pp.315-333.

Pizzinga, M., Bates, C., Lui, J., Forte, G., Morales-Polanco, F., Linney, E., Knotkova, B., Wilson, B., Solari, C.A., Berchowitz, L.E. and Portela, P., 2019. Translation factor mRNA granules direct protein synthetic capacity to regions of polarized growth. *Journal of Cell Biology*, *218*(5), pp.1564-1581.

Pokrywka, N.J. and Stephenson, E.C., 1991. Microtubules mediate the localization of bicoid RNA during Drosophila oogenesis. *Development*, *113*(1), pp.55-66.

Popp, M.W.L. and Maquat, L.E., 2013. Organizing principles of mammalian nonsense-mediated mRNA decay. *Annual Review of Genetics*, *47*, pp.139-165.

Presnyak, V., Alhusaini, N., Chen, Y.H., Martin, S., Morris, N., Kline, N., Olson, S., Weinberg, D., Baker, K.E., Graveley, B.R. and Coller, J., 2015. Codon optimality is a major determinant of mRNA stability. *Cell*, *160*(6), pp.1111-1124.

Price, A.A., Sampson, T.R., Ratner, H.K., Grakoui, A. and Weiss, D.S., 2015. Cas9-mediated targeting of viral RNA in eukaryotic cells. *Proceedings of the National Academy of Sciences*, *112*(19), pp.6164-6169.

Protter, D.S., Rao, B.S., Van Treeck, B., Lin, Y., Mizoue, L., Rosen, M.K. and Parker, R., 2018. Intrinsically disordered regions can contribute promiscuous interactions to RNP granule assembly. *Cell Reports*, *22*(6), pp.1401-1412.

Proudfoot, N.J. and Brownlee, G.G., 1976. 3′ non-coding region sequences in eukaryotic messenger RNA. *Nature*, *263*(5574), pp.211-214.

Pulecio, J., Verma, N., Mejía-Ramírez, E., Huangfu, D. and Raya, A., 2017. CRISPR/Cas9-based engineering of the epigenome. *Cell Stem Cell*, *21*(4), pp.431-447.

Pyhtila, B., Zheng, T., Lager, P.J., Keene, J.D., Reedy, M.C. and Nicchitta, C.V., 2008. Signal sequence-and translation-independent mRNA localization to the endoplasmic reticulum. *RNA*, *14*(3), pp.445-453.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P. and Lim, W.A., 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, *152*(5), pp.1173-1183.

Qin, P., Parlak, M., Kuscu, C., Bandaria, J., Mir, M., Szlachta, K., Singh, R., Darzacq, X., Yildiz, A. and Adli, M., 2017. Live cell imaging of low-and non-repetitive chromosome loci using CRISPR-Cas9. *Nature Communications*, *8*(1), pp.1-10.

Raj, A., Van Den Bogaard, P., Rifkin, S.A., Van Oudenaarden, A. and Tyagi, S., 2008. Imaging individual mRNA molecules using multiple singly labeled probes. *Nature Methods*, *5*(10), pp.877-879.

Ramachandran, V., Shah, K.H. and Herman, P.K., 2011. The cAMP-dependent protein kinase signaling pathway is a key regulator of P body foci formation. *Molecular Cell*, *43*(6), pp.973-981.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A. and Zhang, F., 2013. Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, *8*(11), pp.2281-2308.

Reed, R. and Hurt, E., 2002. A conserved mRNA export machinery coupled to pre-mRNA splicing. *Cell*, *108*(4), pp.523-531.

Reijns, M.A., Alexander, R.D., Spiller, M.P. and Beggs, J.D., 2008. A role for Q/N-rich aggregation-prone regions in P-body localization. *Journal of Cell Science*, *121*(15), pp.2463-2472.

Richter, J.D. and Sonenberg, N., 2005. Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature*, *433*(7025), pp.477-480.

Riggs, C.L., Kedersha, N., Ivanov, P. and Anderson, P., 2020. Mammalian stress granules and P bodies at a glance. *Journal of Cell Science*, *133*(16).

Rigo, F. and Martinson, H.G., 2009. Polyadenylation releases mRNA from RNA polymerase II in a process that is licensed by splicing. *RNA*, *15*(5), pp.823-836.

Rizzi, N., Denegri, M., Chiodi, I., Corioni, M., Valgardsdottir, R., Cobianchi, F., Riva, S. and Biamonti, G., 2004. Transcriptional activation of a constitutive heterochromatic domain of the human genome in response to heat shock. *Molecular Biology of the Cell*, *15*(2), pp.543-551.

Rodriguez, C.R., Cho, E.J., Keogh, M.C., Moore, C.L., Greenleaf, A.L. and Buratowski, S., 2000. Kin28, the TFIIH-associated carboxy-terminal domain kinase, facilitates the recruitment of mRNA processing machinery to RNA polymerase II. *Molecular and Cellular Biology*, *20*(1), pp.104-112.

Rodriguez, J.R., Pikielny, C.W. and Rosbash, M., 1984. In vivo characterization of yeast mRNA processing intermediates. *Cell*, *39*(3), pp.603-610.

Roegiers, F. and Jan, Y.N., 2000. Staufen: a common component of mRNA transport in oocytes and neurons?. *Trends in Cell Biology*, *10*(6), pp.220-224.

Rolfsson, Ó., Middleton, S., Manfield, I.W., White, S.J., Fan, B., Vaughan, R., Ranson, N.A., Dykeman, E., Twarock, R., Ford, J. and Kao, C.C., 2016. Direct evidence for packaging signal-mediated assembly of bacteriophage MS2. *Journal of Molecular Biology*, *428*(2), pp.431-448.

Samacoits, A., Chouaib, R., Safieddine, A., Traboulsi, A.M., Ouyang, W., Zimmer, C., Peter, M., Bertrand, E., Walter, T. and Mueller, F., 2018. A computational framework to study sub-cellular RNA localization. *Nature Communications*, *9*(1), pp.1-10.

Sampson, T.R., Saroj, S.D., Llewellyn, A.C., Tzeng, Y.L. and Weiss, D.S., 2013. A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature*, *497*(7448), pp.254-257.

Sanders, D.W., Kedersha, N., Lee, D.S., Strom, A.R., Drake, V., Riback, J.A., Bracha, D., Eeftens, J.M., Iwanicki, A., Wang, A. and Wei, M.T., 2020. Competing protein-RNA interaction networks control multiphase intracellular organization. *Cell*, *181*(2), pp.306-324.

Sawyer, I.A., Sturgill, D., Sung, M.H., Hager, G.L. and Dundr, M., 2016. Cajal body function in genome organization and transcriptome diversity. *BioEssays*, *38*(12), pp.1197-1208.

Schnell, U., Dijk, F., Sjollema, K.A. and Giepmans, B.N., 2012. Immunolabeling artifacts and the need for live-cell imaging. *Nature Methods*, *9*(2), p.152.

Schuster, B.S., Reed, E.H., Parthasarathy, R., Jahnke, C.N., Caldwell, R.M., Bermudez, J.G., Ramage, H., Good, M.C. and Hammer, D.A., 2018. Controllable protein phase separation and modular recruitment to form responsive membraneless organelles. *Nature Communications*, *9*(1), pp.1-12.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W. and Selbach, M., 2011. Global quantification of mammalian gene expression control. *Nature*, *473*(7347), pp.337-342.

Schwarz, A. and Beck, M., 2019. The benefits of cotranslational assembly: a structural perspective. *Trends in Cell Biology*, *29*(10), pp.791-803.

Segal, R.A., Seo, H.S., Adelmant, G., Walensky, L., Marto, J.A., Dhe-Paganon, S., Fukuda, Y., Pazyra-Murphy, M.F., Tasdemir-Yilmaz, O.E., Li, Y. and Rose, L., 2020. Fast transport of RNA granules by direct interactions with KIF5A/KLC1 motors prevents axon degeneration. *BioRxiv*.

Sen, G.L. and Blau, H.M., 2005. Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies. *Nature Cell Biology*, *7*(6), pp.633-636.

Seydoux, G., 2018. The P granules of C. elegans: a genetic model for the study of RNA–protein condensates. *Journal of Molecular Biology*, *430*(23), pp.4702-4710.

Shah, S.A., Erdmann, S., Mojica, F.J. and Garrett, R.A., 2013. Protospacer recognition motifs: mixed identities and functional diversity. *RNA Biology*, *10*(5), pp.891-899.

Shaner, N.C., Lambert, G.G., Chammas, A., Ni, Y., Cranfill, P.J., Baird, M.A., Sell, B.R., Allen, J.R., Day, R.N., Israelsson, M. and Davidson, M.W., 2013. A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum. *Nature Methods*, *10*(5), pp.407-409.

Shao, S., Ren, C., Liu, Z., Bai, Y., Chen, Z., Wei, Z., Wang, X., Zhang, Z. and Xu, K., 2017. Enhancing CRISPR/Cas9-mediated homology-directed repair in mammalian cells by expressing Saccharomyces cerevisiae Rad52. *The International Journal of Biochemistry & Cell Biology*, *92*, pp.43-52.

Shaw, P. and Brown, J., 2012. Nucleoli: composition, function, and dynamics. *Plant Physiology*, *158*(1), pp.44-51.

Sheth, U. and Parker, R., 2003. Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science*, *300*(5620), pp.805-808.

Sheth, U., Pitt, J., Dennis, S. and Priess, J.R., 2010. Perinuclear P granules are the principal sites of mRNA export in adult C. elegans germ cells. *Development*, *137*(8), pp.1305-1314.

Shieh, Y.W., Minguez, P., Bork, P., Auburger, J.J., Guilbride, D.L., Kramer, G. and Bukau, B., 2015. Operon structure and cotranslational subunit association direct protein assembly in bacteria. *Science*, *350*(6261), pp.678-680.

Siegel, V. and Walter, P., 1986. Removal of the Alu structural domain from signal recognition particle leaves its protein translocation activity intact. *Nature*, *320*(6057), pp.81-84.

Simmons, D.M., Arriza, J.L. and Swanson, L.W., 1989. A complete protocol for in situ hybridization of messenger RNAs in brain and other tissues with radio-labeled single-stranded RNA probes. *Journal of Histotechnology*, *12*(3), pp.169-181.

Singer, R.H. and Ward, D.C., 1982. Actin gene expression visualized in chicken muscle tissue culture by using in situ hybridization with a biotinated nucleotide analog. *Proceedings of the National Academy of Sciences*, *79*(23), pp.7331-7335.

Smargon, A.A., Cox, D.B., Pyzocha, N.K., Zheng, K., Slaymaker, I.M., Gootenberg, J.S., Abudayyeh, O.A., Essletzbichler, P., Shmakov, S., Makarova, K.S. and Koonin, E.V., 2017. Cas13b is a type VI-B CRISPR-associated RNA-guided RNase differentially regulated by accessory proteins Csx27 and Csx28. *Molecular Cell*, *65*(4), pp.618-630.

Smargon, A.A., Shi, Y.J. and Yeo, G.W., 2020. RNA-targeting CRISPR systems from metagenomic discovery to transcriptomic engineering. *Nature Cell Biology*, *22*(2), pp.143-150.

Smibert, C.A., Lie, Y.S., Shillinglaw, W., Henzel, W.J. and Macdonald, P.M., 1999. Smaug, a novel and conserved protein, contributes to repression of nanos mRNA translation in vitro. *RNA*, *5*(12), pp.1535-1547.

Song, T., Zheng, Y., Wang, Y., Katz, Z., Liu, X., Chen, S., Singer, R.H. and Gu, W., 2015. Specific interaction of KIF11 with ZBP1 regulates the transport of β-actin mRNA and cell motility. *Journal of Cell Science*, *128*(5), pp.1001-1010.

Song, W., Filonov, G.S., Kim, H., Hirsch, M., Li, X., Moon, J.D. and Jaffrey, S.R., 2017. Imaging RNA polymerase III transcription using a photostable RNA–fluorophore complex. *Nature Chemical Biology*, *13*(11), p.1187.

Souquere, S., Mollet, S., Kress, M., Dautry, F., Pierron, G. and Weil, D., 2009. Unravelling the ultrastructure of stress granules and associated P-bodies in human cells. *Journal of Cell Science*, *122*(20), pp.3619-3626.

Sparks, K.A. and Dieckmann, C.L., 1998. Regulation of poly (A) site choice of several yeast mRNAs. *Nucleic Acids Research*, *26*(20), pp.4676-4687.

St Johnston, D., Brown, N.H., Gall, J.G. and Jantsch, M., 1992. A conserved double-stranded RNA-binding domain. *Proceedings of the National Academy of Sciences*, *89*(22), pp.10979-10983.

Staněk, D., Přidalová-Hnilicová, J., Novotný, I., Huranová, M., Blažíková, M., Wen, X., Sapra, A.K. and Neugebauer, K.M., 2008. Spliceosomal small nuclear ribonucleoprotein particles repeatedly cycle through Cajal bodies. *Molecular Biology of the Cell*, *19*(6), pp.2534-2543.

Steiert, F., Petrov, E.P., Schultz, P., Schwille, P. and Weidemann, T., 2018. Photophysical behavior of mneongreen, an evolutionarily distant green fluorescent protein. *Biophysical Journal*, *114*(10), pp.2419-2431.

Stephens, S.B., Dodd, R.D., Lerner, R.S., Pyhtila, B.M. and Nicchitta, C.V., 2008. Analysis of mRNA partitioning between the cytosol and endoplasmic reticulum compartments of mammalian cells. in post-transcriptional gene regulation. *Humana Press*, (pp. 197-214).

Stöhr, N., Lederer, M., Reinke, C., Meyer, S., Hatzfeld, M., Singer, R.H. and Hüttelmaier, S., 2006. ZBP1 regulates mRNA stability during cellular stress. *The Journal of Cell Biology*, *175*(4), pp.527-534.

Strack, R.L., Disney, M.D. and Jaffrey, S.R., 2013. A superfolding Spinach2 reveals the dynamic nature of trinucleotide repeat–containing RNA. *Nature Methods*, *10*(12), p.1219.

Strutt, S.C., Torrez, R.M., Kaya, E., Negrete, O.A. and Doudna, J.A., 2018. RNA-dependent RNA targeting by CRISPR-Cas9. *ELife*, *7*, p.e32724.

Sun, Y., Zhang, Y., Hamilton, K., Manley, J.L., Shi, Y., Walz, T. and Tong, L., 2018. Molecular basis for the recognition of the human AAUAAA polyadenylation signal. *Proceedings of the National Academy of Sciences*, *115*(7), pp.E1419-E1428.

Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S. and Vale, R.D., 2014. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell*, *159*(3), pp.635-646.

Tarun, S.Z. and Sachs, A.B., 1995. A common function for mRNA 5'and 3'ends in translation initiation in yeast. *Genes & Development*, *9*(23), pp.2997-3007.

Tauber, D., Tauber, G. and Parker, R., 2020. Mechanisms and regulation of RNA condensation in RNP granule formation. *Trends in Biochemical Sciences*.

Thomson, T., Liu, N., Arkov, A., Lehmann, R. and Lasko, P., 2008. Isolation of new polar granule components in Drosophila reveals P body and ER associated proteins. *Mechanisms of Development*, *125*(9-10), pp.865-873.

Tian, B., Hu, J., Zhang, H. and Lutz, C.S., 2005. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Research*, *33*(1), pp.201-212.

Tomecki, R. and Dziembowski, A., 2010. Novel endoribonucleases as central players in various pathways of eukaryotic RNA metabolism. *RNA*, *16*(9), pp.1692-1724.

Tourrière, H., Chebli, K., Zekri, L., Courselaud, B., Blanchard, J.M., Bertrand, E. and Tazi, J., 2003. The RasGAP-associated endoribonuclease G3BP assembles stress granules. *The Journal of Cell Biology*, *160*(6), pp.823-831.

Trcek, T., Grosch, M., York, A., Shroff, H., Lionnet, T. and Lehmann, R., 2015. Drosophila germ granules are structured and contain homotypic mRNA clusters. *Nature Communications*, *6*(1), pp.1-12.

Trovisco, V., Belaya, K., Nashchekin, D., Irion, U., Sirinakis, G., Butler, R., Lee, J.J., Gavis, E.R. and St Johnston, D., 2016. bicoid mRNA localises to the Drosophila oocyte anterior by random Dynein-mediated transport and anchoring. *ELife*, *5*, p.e17537.

Turner-Bridger, B., Caterino, C. and Cioni, J.M., 2020. Molecular mechanisms behind mRNA localization in axons. *Open biology*, *10*(9), p.200177.

Tutucci, E., Vera, M. and Singer, R.H., 2018b. Single-mRNA detection in living S. cerevisiae using a re-engineered MS2 system. *Nature Protocols*, *13*(10), pp.2268-2296.

Tutucci, E., Vera, M., Biswas, J., Garcia, J., Parker, R. and Singer, R.H., 2018a. An improved MS2 system for accurate reporting of the mRNA life cycle. *Nature Methods*, *15*(1), pp.81-89.

Tyagi, S. and Alsmadi, O., 2004. Imaging native β-actin mRNA in motile fibroblasts. *Biophysical Journal*, *87*(6), pp.4153-4162.

Tyagi, S. and Kramer, F.R., 1996. Molecular beacons: probes that fluoresce upon hybridization. *Nature Biotechnology*, *14*(3), pp.303-308.

Valegård, K., Liljas, L., Fridborg, K. and Unge, T., 1990. The three-dimensional structure of the bacterial virus MS2. *Nature*, *345*(6270), pp.36-41.

Van Dijk, E., Cougot, N., Meyer, S., Babajko, S., Wahle, E. and Séraphin, B., 2002. Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures. *The EMBO Journal*, *21*(24), pp.6915-6924.

Van Treeck, B., Protter, D.S., Matheny, T., Khong, A., Link, C.D. and Parker, R., 2018. RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome. *Proceedings of the National Academy of Sciences*, *115*(11), pp.2734-2739.

Vanzo, N.F. and Ephrussi, A., 2002. Oskar anchoring restricts pole plasm formation to the posterior of the Drosophila oocyte. *Development*, *129*(15), pp.3705-3714.

Vargas, D.Y., Shah, K., Batish, M., Levandoski, M., Sinha, S., Marras, S.A., Schedl, P. and Tyagi, S., 2011. Single-molecule imaging of transcriptionally coupled and uncoupled splicing. *Cell*, *147*(5), pp.1054-1065.

Verdecia, M.A., Bowman, M.E., Lu, K.P., Hunter, T. and Noel, J.P., 2000. Structural basis for phosphoserine-proline recognition by group IV WW domains. *Nature Structural Biology*, *7*(8), pp.639-643.

Veres, A., Gosis, B.S., Ding, Q., Collins, R., Ragavendran, A., Brand, H., Erdin, S., Cowan, C.A., Talkowski, M.E. and Musunuru, K., 2014. Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. *Cell Stem Cell*, *15*(1), pp.27-30.

Wakiyama, M., Imataka, H. and Sonenberg, N., 2000. Interaction of eIF4G with poly (A)-binding protein stimulates translation and is critical for Xenopus oocyte maturation. *Current Biology*, *10*(18), pp.1147-1150.

Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F. and Jaenisch, R., 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*, *153*(4), pp.910-918.

Wang, Q., Sawyer, I.A., Sung, M.H., Sturgill, D., Shevtsov, S.P., Pegoraro, G., Hakim, O., Baek, S., Hager, G.L. and Dundr, M., 2016. Cajal bodies are linked to genome conformation. *Nature Communications*, *7*(1), pp.1-17.

Ward, B.E., Gersen, S.L., Carelli, M.P., McGuire, N.M., Dackowski, W.R., Weinstein, M., Sandlin, C., Warren, R. and Klinger, K.W., 1993. Rapid prenatal diagnosis of chromosomal aneuploidies by fluorescence in situ hybridization: clinical experience with 4,500 specimens. *American Journal of Human Genetics*, *52*(5), p.854.

Weil, T.T., Forrest, K.M. and Gavis, E.R., 2006. Localization of bicoid mRNA in late oocytes is maintained by continual active transport. *Developmental Cell*, *11*(2), pp.251-262.

Welch, M.D., Iwamatsu, A. and Mitchison, T.J., 1997. Actin polymerization is induced by Arp 2/3 protein complex at the surface of Listeria monocytogenes. *Nature*, *385*(6613), pp.265-269.

Wells, J.N., Bergendahl, L.T. and Marsh, J.A., 2016. Operon gene order is optimized for ordered protein complex assembly. *Cell Reports*, *14*(4), pp.679-685.

Whelan, D.R. and Bell, T.D., 2015. Image artifacts in single molecule localization microscopy: why optimization of sample preparation protocols matters. *Scientific Reports*, *5*(1), pp.1-10.

Williamson, D.H. and Fennell, D.J., 1979. [62] Visualization of yeast mitochondrial dna with the fluorescent stain “DAPI”. *Methods in Enzymology* (Vol. 56, pp. 728-733). Academic Press.

Wu, B., Eliscovich, C., Yoon, Y.J. and Singer, R.H., 2016. Translation dynamics of single mRNAs in live cells and neurons. *Science*, *352*(6292), pp.1430-1435.

Wu, X., Mao, S., Ying, Y., Krueger, C.J. and Chen, A.K., 2019. Progress and challenges for live-cell imaging of genomic loci using CRISPR-based platforms. *Genomics, Proteomics & Bioinformatics*, *17*(2), pp.119-128.

Xiao, L. and Grove, A., 2009. Coordination of ribosomal protein and ribosomal RNA gene expression in response to TOR signaling. *Current Genomics*, *10*(3), pp.198-205.

Xu, F., Li, L., Schulz, V.P., Gallagher, P.G., Xiang, B., Zhao, H. and Li, P., 2014. Cytogenomic mapping and bioinformatic mining reveal interacting brain expressed genes for intellectual disability. *Molecular Cytogenetics*, *7*(1), pp.1-13.

Xu, X. and Qi, L.S., 2019. A CRISPR–dCas toolbox for genetic engineering and synthetic biology. *Journal of Molecular Biology*, *431*(1), pp.34-47.

Xue, Y. and Acar, M., 2018. Live-cell imaging of chromatin condensation dynamics by CRISPR. *IScience*, *4*, pp.216-235.

Yan, D., Wu, Z., Chisholm, A.D. and Jin, Y., 2009. The DLK-1 kinase promotes mRNA stability and local translation in C. elegans synapses and axon regeneration. *Cell*, *138*(5), pp.1005-1018.

Yan, W.X., Chong, S., Zhang, H., Makarova, K.S., Koonin, E.V., Cheng, D.R. and Scott, D.A., 2018. Cas13d is a compact RNA-targeting type VI CRISPR effector positively modulated by a WYL-domain-containing accessory protein. *Molecular Cell*, *70*(2), pp.327-339.

Yang, H., Wang, H., Shivalila, C.S., Cheng, A.W., Shi, L. and Jaenisch, R., 2013. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell*, *154*(6), pp.1370-1379.

Yang, L.Z., Wang, Y., Li, S.Q., Yao, R.W., Luan, P.F., Wu, H., Carmichael, G.G. and Chen, L.L., 2019. Dynamic imaging of RNA in living cells by CRISPR-Cas13 systems. *Molecular Cell*, *76*(6), pp.981-997.

Ye, H., Rong, Z. and Lin, Y., 2017. Live cell imaging of genomic loci using dCas9-SunTag system and a bright fluorescent protein. *Protein & Cell*, *8*(11), pp.853-855.

Yin, H., Xue, W., Chen, S., Bogorad, R.L., Benedetti, E., Grompe, M., Koteliansky, V., Sharp, P.A., Jacks, T. and Anderson, D.G., 2014. Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nature Biotechnology*, *32*(6), pp.551-553.

Youn, J.Y., Dunham, W.H., Hong, S.J., Knight, J.D., Bashkurov, M., Chen, G.I., Bagci, H., Rathod, B., MacLeod, G., Eng, S.W. and Angers, S., 2018. High-density proximity mapping reveals the subcellular organization of mRNA-associated granules and bodies. *Molecular Cell*, *69*(3), pp.517-532.

Yue, Z., Maldonado, E., Pillutla, R., Cho, H., Reinberg, D. and Shatkin, A.J., 1997. Mammalian capping enzyme complements mutant Saccharomyces cerevisiae lacking mRNA guanylyltransferase and selectively binds the elongating form of RNA polymerase II. *Proceedings of the National Academy of Sciences*, *94*(24), pp.12898-12903.

Zaessinger, S., Busseau, I. and Simonelig, M., 2006. Oskar allows nanos mRNA translation in Drosophila embryos by preventing its deadenylation by Smaug/CCR4. *Development*, *133*(22), pp.4573-4583.

Zenklusen, D., Larson, D.R. and Singer, R.H., 2008. Single-RNA counting reveals alternative modes of gene expression in yeast. *Nature Structural & Molecular Biology*, *15*(12), p.1263.

Zhang, C. and Jia, G., 2018. Reversible RNA Modification N1-methyladenosine (m1A) in mRNA and tRNA. *Genomics, Proteomics & Bioinformatics*, *16*(3), pp.155-161.

Zhang, G., Taneja, K.L., Singer, R.H. and Green, M.R., 1994. Localization of pre-mRNA splicing in mammalian nuclei. *Nature*, *372*(6508), pp.809-812.

Zhang, H., Elbaum-Garfinkle, S., Langdon, E.M., Taylor, N., Occhipinti, P., Bridges, A.A., Brangwynne, C.P. and Gladfelter, A.S., 2015. RNA controls PolyQ protein phase transitions. *Molecular Cell*, *60*(2), pp.220-230.

Zhang, J., Fei, J., Leslie, B.J., Han, K.Y., Kuhlman, T.E. and Ha, T., 2015. Tandem spinach array for mRNA imaging in living bacterial cells. *Scientific Reports*, *5*(1), pp.1-9.

Zhao, B.S., Roundtree, I.A. and He, C., 2017. Post-transcriptional gene regulation by mRNA modifications. *Nature Reviews Molecular Cell Biology*, *18*(1), pp.31-42.

Zhao, W.M., Jiang, C., Kroll, T.T. and Huber, P.W., 2001. A proline-rich protein binds to the localization element of Xenopus Vg1 mRNA and to ligands involved in actin polymerization. *The EMBO Journal*, *20*(9), pp.2315-2325.