

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

RNA regulons in Hox 5'UTRs confer ribosome specificity to gene regulation and embryonic development

Permalink

<https://escholarship.org/uc/item/42h50712>

Author

Xue, Shifeng

Publication Date

2014

Peer reviewed|Thesis/dissertation

RNA regulons in Hox 5'UTRs confer ribosome specificity to gene
regulation and embryonic development

by

Shifeng Xue

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Developmental Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Acknowledgements

I would like to thank my mentor, Dr Maria Barna, without whom this work will not have been possible. Her exciting research drew me to the lab and her continual support and confidence in me helped move much of my work forward. I would like to express my gratitude to my committee members, Drs Jeremy Reiter, Takashi Mikawa and Robert Belloch for many discussions and helpful advice. Many thanks to my collaborators Dr Rhiju Das, Siqi Tian and Ann Kladwang, for helping to determine the RNA structure of the *Hox* IRES elements. Their help was invaluable in this work. A big thank you also to the members of the Barna lab, past and present, including Nadya Kondrashov, Timothy Sanders, Esther Llagostera Martin, Khanh Ngo, Katrina Cabaltera, Kotaro Fujii, Elif Sarinay Cenik, Gerald Tiu, Jing Huang, Deniz Simsek, Zhen Shi, Gun Woo Byeon, Corinne Fairchild, Nicolas Denans and Olena Zhulyn. Thank you for showing me how to do great science and making the lab a fun place to work in. I would like to thank Dr Davide Ruggero and the Ruggero lab, for introducing me to translation, and for being great people to discuss translation with even after our move to Stanford. I would also like to thank the administrative staff: Danny Dam and Billy Luh in UCSF and Mimi Qian, Sergio Alcantera and Todd Galitz in Stanford. They made the non-scientific parts of my PhD much easier. I would like to use this opportunity to also thank my friends and family for their support over the years, and my funding agency A*STAR for generous funding.

Chapter 1 was adapted from Xue, S., and Barna, M. (2012). Specialized ribosomes: a new frontier in gene regulation and organismal biology. *Nature Reviews Molecular Cell Biology*, 13(6), 355–369.

Chapter 2 was originally published in Xue, S., Tian, S., Fujii, K., Kladwang, W., Das, R. and Barna, M. (2015). RNA regulons in *Hox* 5'UTRs confer ribosome specificity to gene regulation. *Nature*, 517, 33-38.

Abstract

RNA regulons in Hox 5'UTRs confer ribosome specificity to gene regulation and embryonic development

Shifeng Xue

Historically, the ribosome has been viewed as a ribozyme with constitutive rather than regulatory capacity in mRNA translation. However, emerging studies reveal that ribosome activity may be highly regulated to impart a new layer of specificity in control of gene expression and organismal biology. Heterogeneity in the composition of the ribosome may generate variations to the translation machinery that significantly impact on how the genomic template is translated into functional proteins.

Studies in our laboratory have shown that ribosomal protein RPL38 confers specificity to the ribosome and is responsible for specifically translating a subset of Homeobox (Hox) mRNAs to regulate the mammalian body plan. In *Rpl38* mutant embryos, global protein synthesis is unchanged while the translation of 8 of the 39 Hox mRNAs is specifically affected. Unexpectedly, *Rpl38* transcripts are highly enriched in a tissue specific manner, most notably in somites and within the neural tube. Together, these studies led us to propose that ribosome in distinct regions of the embryo may be highly specialized to exert transcript-specific translational control. However how transcripts are recognized and selected by RPL38 remains unknown.

Here I uncover a unique set of RNA regulons embedded in the 5'UTRs of Hox genes which are critical for ribosome-directed control of their spatiotemporal expression patterns. These structured RNA elements, resembling Viral Internal Ribosome Entry

Sites (IRES), are positioned close to the AUG of subsets of Hox mRNA to interact with ribosomal proteins and directly recruit the ribosome. Using *Hoxa9* as a model, I find that the IRES element folds into a stable structure *in vitro* and can recruit ribosomes directly to the mRNA. By generating the first targeted knock-out of an IRES element in the mouse, I show that the *Hoxa9* transcript, despite numerous layers of transcriptional control, requires an IRES-element for accurate spatiotemporal expression and axial skeletal patterning. The unexpected requirement for specialized translation control of Hox transcripts is further explained by the activity of a novel Translational Inhibitory Element (TIE), which inhibits general cap-dependent translation, thereby converting a generic mode of mRNA translation into a highly specialized one. Together, these studies uncover a new paradigm of specialized translational regulation that is achieved through a combination of unique RNA regulons to facilitate ribosome-mediate control of gene expression.

Table of contents

Chapter 1: Introduction.....	1
Chapter 2: RNA regulons in Hox 5' UTRs confer ribosome specificity to gene regulation.....	67
Chapter 3: Insights into IRES and TIE.....	136
Chapter 4: Discussion and future directions.....	161

List of Tables

Chapter 2: RNA regulons in Hox 5' UTRs confer ribosome specificity to gene regulation

Table 2.1: Primer sequences.....132

Chapter 3: Insights into IRES and TIE

Table 3.1: *Hox* genes classified based on their IRES activity.....158

Table 3.2: Primer sequences.....159

List of Figures

Chapter 1: Introduction

Figure 1.1: Ribosome biogenesis.....	48
Figure 1.2: Heterogeneity in the ribosome at the level of ribosomal protein composition and post-translational modifications across different kingdoms of life.....	51
Figure 1.3: Extraribosomal functions of ribosomal proteins in translation control.....	54
Figure 1.4: Ribosome associated factors affect ribosome function.....	58
Figure 1.5: rRNA heterogeneity.....	60
Figure 1.6: <i>Cis</i> -regulatory elements within mRNAs that interface with ribosomes or ribosomal proteins to confer transcript-specific regulation of gene expression.....	62
Figure 1.7: Ribosome specificity in cell and developmental biology.....	64

Chapter 2: RNA regulons in Hox 5' UTRs confer ribosome specificity to gene regulation

Figure 2.1: Select <i>HoxA</i> 5' UTRs contain IRES elements regulated by RPL38.	100
Figure 2.2: <i>HoxA</i> IRES controls confirming that Fluc activity from bicistronic vector is due to IRES activity.....	102
Figure 2.3: Disruption of <i>Rpl38</i> in C3H10T1/2 by TALEN nucleases.....	104
Figure 2.4: Alignment of <i>Hoxa9</i> IRES element between vertebrate species.....	106
Figure 2.5: The <i>Hoxa9</i> IRES is an RNA structure that facilitates recruitment of the ribosome and is required for normal translation.....	108
Figure 2.6: Chemical mapping and secondary structure modeling of full-length <i>Hoxa9</i> 5' UTR.	110
Figure 2.7: Chemical mapping and secondary structure modeling of full length <i>Hoxa5</i> 5' UTR.....	112
Figure 2.8: Secondary structure modeling and mutate-and-map (M ²) data set of <i>Hoxa9</i> IRES element.....	114

Figure 2.9: Mutation/rescue results of <i>Hoxa9</i> IRES structure (nt 944–1,266) probed by 1M7.....	116
Figure 2.10: Putative uORFs within the 5' UTRs of <i>Hoxa9</i> and <i>Hoxa5</i> do not inhibit cap-dependent translation and <i>Hoxa9</i> ^{ΔIRES} targeting strategy.....	118
Figure 2.11: Identification of minimal IRES domains in all IRES-containing <i>HoxA</i> 5' UTR.....	120
Figure 2.12: Identification of a common TIE that blocks cap-dependent translation of Hox IRES-mRNAs.....	122
Figure 2.13: The <i>Hoxa9</i> IRES is required for axial skeleton patterning but not for <i>Hoxa9</i> mRNA expression.....	124
Figure 2.14: The presence of Neo cassette in the <i>Hoxa9</i> locus is linked to the presence of an L1 → T13 homeotic transformation.....	126
Figure 2.15: The <i>Hoxa9</i> IRES is critically required for <i>Hoxa9</i> translation <i>in vivo</i>	128
Figure 2.16: Sucrose gradient fractionation shows no difference in β-actin association with polysomes in <i>Hoxa9</i> ^{+/+} and <i>Hoxa9</i> ^{ΔIRES/ΔIRES} embryos.....	130

Chapter 3: Insights into IRES and TIE

Figure 3.1: Factor requirements of IRES and TIE.....	150
Figure 3.2: <i>Hoxa9</i> mRNA but not protein is expressed in the limb.....	152
Figure 3.3: <i>Hoxa9</i> is translationally repressed in the limb.....	154
Figure 3.4: IRES activity in all <i>Hox</i> mRNAs.....	156

Chapter 1

Introduction

A major challenge in biology is to understand how protein expression is regulated with exquisite temporal and spatial precision to control unique cell behaviors and to give rise to the remarkable diversity of cell types that is characteristic of metazoan development. Decades of research have demonstrated numerous layers of regulation that orchestrate this process. Indeed, almost every step in control of gene expression has evolved many levels of regulatory specificity: from exquisitely modified histones tails, to the countless small RNAs that fine-tune the abundance of transcripts, as well as the remarkable repertoire of post-translational modifications that steer protein behavior. On the contrary, such regulatory control has not been fully realized at a crucial step in gene expression, that of protein production. In particular, although the ribosome is well known to be an immensely complex “molecular machine”, it is believed to function as somewhat of an automaton in translating the genetic code.

Recent findings are now, however, shifting the view of ribosome function, revealing unique and unexpected roles for the translation machinery itself in directing fundamental aspects of cell and organismal biology. Here I discuss the hypothesis that “specialized ribosomes”, which have a unique composition or specialized activity confer regulatory control in gene expression. Examples of heterogeneity in ribosomes can include diversity in the composition and post-translational modifications of subsets of ribosomal proteins, variations in rRNA sequences, or binding to distinct ribosome-associated factors, all of which may contribute to the occurrence of “specialized ribosomes” in different cell types. In addition, even core components of the ribosome that show little variation may exert more specialized activity by virtue of their interactions with specific *cis*-acting regulatory elements present within subsets of mRNAs. In a

multi-cellular organism, greater regulation in ribosome activity may provide an important new layer for control of gene expression in time and space.

The ribosome as a “molecular machine” and how it is constructed

The ribosome has been described as a “molecular machine par excellence” (Frank, 2000; Sonenberg and Hinnebusch, 2009). This metaphor can be traced back to an influential article written by Bruce Alberts in which the analogy of the “molecular machine” was first used to describe the highly coordinated behavior of large complexes such as the DNA replication machinery or the ribosome (Alberts, 1998). The machine analogy influenced the manner in which we conceptualized a molecular complex; that although it may be composed of a number of distinct elements, a high degree of coordination and precision of these parts serves a singular, unilateral function. The basic mechanisms of how the genetic code is translated into proteins fits well with this concept. For example, the ribosome has a universally conserved role in catalyzing protein synthesis. Moreover, protein synthesis is remarkably fast. Elongation of the polypeptide chain by one amino acid occurs in approximately 60ms and in a highly accurate fashion (Zaher and Green, 2009). This requires a sophisticated level of precision and automation. However, the powerful metaphor of a machine has also perhaps stereotyped our view of the ribosome as a passive participant in translation. A critical question is whether certain parts of the ribosome are changeable and/or activated by additional layers of regulation that depend on *context*. To understand this principle, requires examining the molecular machine at its very core, as ribosome biogenesis is a highly coordinated process (Figure 1.1). Indeed, ribosomes make up

much of the cell's mass, and ribosome biogenesis uses a significant proportion of the cell's energy (Warner, 1999). The eukaryotic ribosome contains 4 RNA species and 79 ribosomal proteins. Despite the constitutive nature of the ribosome, several components could also give rise to divergent regulation of the “molecular machine”. In understanding its construction (Figure 1.1), possible points of variation can also be envisioned. I will next discuss the evidence or lack thereof for individual components exerting unique and more regulatory control.

Heterogeneity in ribosome function and/or activity

Francis Crick first proposed, in relation to the “RNA world” hypothesis, that an ancestral proto-ribosome may have been comprised of only functional rRNA, which shares homology with present-day rRNAs (Crick, 1968). Biochemical and genetic experiments have revealed a direct involvement of conserved rRNA nucleotides, highlighting their central role in ribosome function (Nissen, 2000; Noller et al., 2005). Remarkably, early attempts in delineating the minimal components necessary for ribosome activity demonstrated that peptidyl transferase, the ribosomal activity responsible for catalysis of peptide bond formation, might only require rRNA in the absence of the majority of ribosomal proteins (Noller et al., 1992). This raises the question of the function of the ribosomal proteins that associate with the ribosome. Ribosomal proteins have likely co-evolved important roles in rRNA folding and function (Held et al., 1973; Röhl and Nierhaus, 1982). However, we speculate that the emergence of ribosomal proteins might, in addition, imbue greater specificity to the RNA-based translation machinery in control of protein synthesis. It is also important to

consider whether ribosomal proteins may be acting at multiple times and/or locations to exert unique activities. For example, they may play important roles in rRNA processing within the nucleolus and also exert more specialized functions in translation control as part of the mature ribosome within the cytoplasm. In this section, we examine the possibility of further regulation in ribosome composition and activity both with respect to the proteins and rRNA that comprise the ribosome.

Ribosomal Protein Paralogs

In mammals, although there are ~2,000 sequences encoding for ribosomal proteins, most of them are predicted to be pseudogenes (Balasubramanian et al., 2009) and most functional ribosomal proteins are encoded by a single gene. In stark contrast, in yeast, plants and flies many ribosomal proteins are encoded by more than one gene. *Saccharomyces cerevisiae*, for example, arose from an ancient genome duplication and 59 of the 78 ribosomal proteins retained two genomic copies (Kellis et al., 2004). Remarkably, despite having high sequence homology, the two RP copies do not always appear to be functionally redundant. Deletions of individual paralogs can give rise to completely unique phenotypes affecting cell size, bud size selection (Ni and Snyder, 2001), sporulation (Enyenihi and Saunders, 2003), susceptibility to virus (Carroll et al., 2008; Ohtake and Wickner, 1995) as well as drug sensitivity (Komili et al., 2007; Parenteau et al., 2011) and these phenotypes are as dissimilar as deletions in completely unrelated ribosomal proteins. For example, deletions of at least six ribosomal proteins, but not their paralogs, affect the localization of *ASH1* mRNA to the bud tip at the emerging daughter cell, which requires tight translational regulation

(Komili et al., 2007). It is however important to note that distinct phenotypes observed among some ribosomal protein paralogs may be confounded by additional suppressor mutations arising during the propagation of ribosomal protein deletion strains (Hughes et al., 2000; Steffen et al., 2012) and hence a more careful examination is needed to confirm these phenotypes. Despite these caveats, the many examples of unique functional roles for individual ribosomal protein paralogs suggest a differential requirement for individual paralogs in specific molecular and cellular events.

Recently, it has been shown that splicing of yeast ribosomal protein mRNAs can regulate expression of both ribosomal protein paralog genes at the transcript level, often in a non-reciprocal manner. For example, removal of the intron in *RPS29A* reduces its own expression as well as the expression of *RPS29B* (Parenteau et al., 2011) (Figure 1.2a). This is surprising, as a change in the expression of any single copy of the duplicated ribosomal protein gene pair would be expected to increase expression of the other if it was simply required to maintain a stoichiometric level of expression. On the contrary, in wild-type cells ~70% of all duplicated ribosomal protein genes are asymmetrically expressed (Parenteau et al., 2011). This suggests that the ratio of paralogs in the yeast cell is regulated and that paralogs are not merely equivalent, functional substitutes. It is tempting to speculate that variation in ribosome composition, for example by incorporating different paralogs, may be a mechanism to provide unique regulatory activity to the ribosome. Although loss-of-function phenotypes in distinct ribosomal protein gene copies would support this possibility, little is known to date on the functional roles of distinct ribosomal protein paralogs in translational control.

Plants have even more gene copies encoding for a single ribosomal protein. Each ribosomal protein in *Arabidopsis thaliana* is encoded by two to seven genes (Barakat et al., 2001). Rapeseed, *Brassica napus*, has 966 genes encoding for 79 ribosomal proteins (Whittle and Krochko, 2009). Some of the ribosomal protein gene copies are identical in sequence, but many of the paralogs have variations from each other and unique expression patterns during development (Falcone Ferreyra et al., 2010; Weijers et al., 2001). For example, in *Arabidopsis thaliana*, *RPS5A* is strongly expressed in dividing cells while *RPS5B* is expressed in cells undergoing differentiation (Weijers et al., 2001) (Figure 1.2b). *RPL11B* (previously known as *RPL16B*), is also highly expressed in all actively dividing cells while the expression pattern of *RPL11A* is much more tissue-specific (Williams and Sussex, 1995). Moreover, the two *Arabidopsis RPL23a* paralogs, which possess high sequence homology, are not of equivalent importance for normal plant development (Degenhardt and Bonham-Smith, 2008a, 2008b). Deletion of *Rpl23aA* but not of *Rpl23aB* impeded growth and led to morphological abnormalities, revealing specialized functions for specific ribosomal protein paralogs. A genome wide study of ribosomal proteins in *Brassica napus* shows that different paralogs of the same ribosomal protein are not regulated in the same pattern between tissues. Interestingly, more than two-thirds of RPs are encoded by paralogs that have opposite expression patterns between embryos and seedlings, suggesting that a different set of ribosome paralogs might be required for early development (Whittle and Krochko, 2009).

Heterogeneity in ribosomal protein paralog expression is also seen in *Drosophila melanogaster*. For example, some paralogs show differential expression levels in the

adult testes (Marygold et al., 2007). *rpL22* is transcribed ubiquitously but *rpL22-like* is enriched in the testes (Figure 1.2c). Both genes are essential in the fly suggesting that they are not functionally redundant. Moreover, both proteins are incorporated into ribosomes as they are detected in polysomes (translationally active ribosomes) and 80S fractions but not in pre-ribosomal fractions (Kearse et al., 2011). In addition, *RpS5b*, *RpS19a*, *RpL10Aa* and *RpL37b* also show enhanced expression in the testes as compared to their paralogs (Marygold et al., 2007). Such heterogeneity in ribosomal protein expression in the gonads suggests that the development of germ cells may require tissue-specific variations in the translational machinery.

In mammals, most ribosomal proteins are encoded by only a single gene copy. However, some notable exceptions exist. For example, *RPS4*, is encoded by three genes located on the X and Y chromosomes (*RPS4X*, *RPS4Y1*, and *RPSY2*) (Fisher et al., 1990; Lopes et al., 2010). In human males, *RPS4X* and *RPS4Y1* are expressed ubiquitously but *RPSY2* expression is restricted to the testis and prostate, suggesting a male-specific role of *RPSY2* and the possibility of testis-specific ribosomes (Lopes et al., 2010) (Figure 1.2d). Although the sequences of the three *RPS4* proteins are very similar, distinct amino acids in the C-terminus of *RPS4Y2* are speculated to possibly facilitate unique interactions with distinct ribosomal proteins or extraribosomal factors, some of which may be testis-specific (Lopes et al., 2010). In addition, mice carry paralogs for *RPL10*, *RPL22* and *RPL39* and they are also differentially expressed (Sugihara et al., 2010). Ribosomes isolated from rodent liver, mammary gland and testis were isolated and analyzed by 2D gel electrophoresis and mass spectrometry. Surprisingly, *RPL10*-like and *RPL39*-like were only found in ribosomes from the testis but not in the liver or

mammary. Interestingly, RPL39-like localized to the nucleoli and was found in 80S and polysome fractions, consistent with it being a RP (Sugihara et al., 2010). By contrast RPL22-like is absent from the testes but is specifically expressed in the liver and mammary glands. Once again, this suggests that ribosomes in different tissues contain unique ribosomal protein paralogs and similar to *Drosophila*, variation appears to be evolutionarily conserved and particularly evident in germ cells.

Ribosomal protein expression

Paralogs aside, the expression of core ribosomal proteins may not be ubiquitous, at least under certain circumstances. It has been previously suggested that the equimolar production of ribosomal proteins is important for proper ribosome biogenesis and that ribosomal proteins are transcribed at the same rate and have similar lifespans (Hariharan et al., 1989; Kim and Warner, 1983). However, most of these studies focus on only a handful of ribosomal proteins in a single cell type. Recent studies examining ribosomal proteins more globally, at an organismal level, have shown that different ribosomal proteins are expressed at distinct levels in unique cell types (Bortoluzzi et al., 2001; Kondrashov et al., 2011). In the social amoeba *Dictyostelium discoideum*, ribosomes show differential ribosomal protein compositions, at two different states of its life cycle, the vegetative stage and the spore stage (Figure 1.2f). Many ribosomal proteins are expressed at different levels in one state as compared to the other, suggesting that these proteins may be developmentally regulated during cell differentiation (Ramagopal, 1991a; Ramagopal and Ennis, 1981).

In the mouse, the levels of *Rpl38* transcripts exhibit a tissue-specific expression pattern (Figure 1.7c) (Kondrashov et al., 2011). For example, during embryonic development *Rpl38* expression is substantially enriched within developing somites and a specific subset of motor neurons within the spinal cord. Strikingly, this pattern of expression mirrors, to a large extent, the tissues that are affected by the loss of function of RPL38 in mutant embryos (Kondrashov et al., 2011). Moreover, *Rpl38* is expressed in some adult tissues, but not others (Sahin et al., 2005). RPL10 expression in the mouse embryo is also highly tissue-specific as it is enriched in the developing epidermis, limb buds (Mills et al., 1999), and shows a very specific expression pattern in fetal bone (Green et al., 2000). Notably, RPL10 is one of the few mammalian ribosomal proteins that has a paralogue, however, this cannot solely account for its restricted expression pattern as the RPL10-like paralogue is also differentially expressed (Sugihara et al., 2010). RPL10-like has only been found expressed in the testes, but not other tissues such as the liver.

More globally, a large-scale quantitative expression-profiling screen examined the expression of 72 ribosomal proteins in 14 tissue and cell types within the developing vertebrate embryo (Figure 1.2e). This revealed dramatic difference in their expression levels among different tissue and cell types and revealed co-regulated patterns of expression among specific groups of ribosomal proteins (Kondrashov et al., 2011). Additional studies also point to the possibility of greater heterogeneity in ribosomal protein expression in adult tissues (Bortoluzzi et al., 2001). Although most of these studies only look at the mRNA levels of ribosomal proteins, the huge range of expression levels between tissues strongly suggests that ribosomes in distinct cell and

tissue types may be heterogeneous. These findings might also suggest that some ribosomes may contain more than one copy of certain, highly expressed ribosomal proteins. This is analogous to the dimerization of the acidic ribosomal proteins in *E. coli*, where there are two copies of RpL7 and RpL12 per ribosome (Hardy, 1975; Oleinikov et al., 1998; Subramanian, 1975). Importantly, as ribosomal proteins may have extra-ribosomal functions (Figure 1.3), can be degraded at the protein level (Lam et al., 2007), or are translationally controlled (Hsieh et al., 2012), further proteomics studies will be key in confirming this level of heterogeneity.

Post-translational modifications

In addition to ribosomal protein expression differences, post-translational modifications on ribosomal proteins can add another layer of specificity. In many organisms, ribosomal proteins are post-translationally modified. In yeast, at least 16 ribosomal proteins are either methylated or acetylated (Carroll et al., 2008; Lee et al., 2002). In humans, at least 11 large subunit ribosomal proteins and most of the small subunit proteins are also post-translationally modified (Odintsova et al., 2003; Yu et al., 2005). In *Arabidopsis thaliana*, 23 of the 80 ribosomal protein families contain at least one covalent modification (Carroll et al., 2008). Modifications of RPs appear to be highly regulated. For example, ribosomes from unique phases of the *Dictyostelium* life cycle show specific changes in ribosomal protein phosphorylation and methylation patterns that may facilitate unique translational needs of the cell (Ramagopal, 1991b) (Figure 1.2f).

Historically, the post-translational modifications on RPS6 have been the best characterized. For example, in mammals, RPS6 is exquisitely phosphorylated on five residues in response to mitogen and growth factor signaling (Krieg et al., 1988). It has long been believed that RPS6 phosphorylation has an important function in translation control of a subclass of mRNAs harboring a 5' tract oligopyrimidine (5'TOP) sequence, and that this level of regulation may imbue the ribosome with greater specificity. However, unexpectedly, a knock-in mouse where all 5 phosphorylated residues of RPS6 were mutated to alanine did not show a decrease in general or 5' TOP mRNA translation, but paradoxically resulted in an increase in protein synthesis rates. Moreover, the *in vivo* physiological role of RPS6 phosphorylation appears to be rather limited as mice with defective RPS6 phosphorylation only display decreased cell size and impaired glucose homeostasis (Ruvinsky et al., 2005). Hence, the physiological significance of ribosomal protein post-translational modifications will need to be carefully evaluated in the future.

A recent study showed that in addition to acetylation, methylation and phosphorylation, many ribosomal proteins can be modified by the addition of O-linked B-N-acetylglucosamine(O-GlcNAc) is added to serine or threonine residues. Overexpression of O-GlcNAc transferase (OGT), the enzyme involved in O-GlcNAc modification, resulted in an increase in the 60S and 80S peaks in the polysome profile, suggesting that O-GlcNAc modification might be important in ribosome biogenesis.

The same ribosomal proteins can exhibit different dynamics of specific post-translational modifications, suggesting even more complex regulation to the ribosome (Zeidan et al., 2010). In addition to being phosphorylated and O-GlcNAcylated,

ribosomal proteins are also regulated by ubiquitylation. For example, Rpl28 in *S. cerevisiae* is heavily ubiquitylated during the S phase of the cell cycle, while ubiquitylation is reduced in G₁. Rpl28 is located at the peptidyl transferase center and its ubiquitylation has a stimulatory effect on translation. Ribosomes carrying poly-ubiquitylated Rpl28 are able to translate a reporter gene at a faster rate than those with mono-ubiquitylated Rpl28 (Spence et al., 2000). In addition, mass spectrometry studies have shown that ribosomal proteins can be methylated under different conditions (Yu et al., 2005). Together, these studies identify unique post-translational modifications on specific ribosomal proteins that may affect the function of the ribosome.

Ribosome-associated factors

Proteomics of *Saccharomyces cerevisiae* ribosomes identified 77 ribosome-associated proteins, suggesting that apart from ribosomal proteins, many additional factors may also modulate ribosome activity. Cells lacking some of these proteins show defects in the rate and fidelity of protein synthesis, as well as ribosome biogenesis (Fleischer et al., 2006). Interestingly, ribosome associated proteins may also regulate the translation of specific subsets of mRNAs. For example, in *Drosophila*, the apoptosis regulator Reaper binds directly and specifically to the 40S subunit and disrupts AUG recognition by the scanning 48S complex, thereby inhibiting cap-dependent translation (Figure 1.4a). Interestingly, Reaper may allow the selective translation of specific mRNAs harboring IRES-elements, which bypasses scanning, to augment the apoptotic program (Colón-Ramos et al., 2006). Another ribosome-associated protein is glycogen synthase 1 (GYS1). In HeLa cells, GYS1 associates with actively translating ribosomes.

Knock down of GYS1 causes a decrease in polysomes and a change in the subsets of mRNAs in polysomes (Fuchs et al., 2011). It is therefore clear that the translation of mRNAs may be enhanced or inhibited by distinct ribosome-associated factors, adding a further level of complexity to the translation machinery.

RACK1, the receptor for activated C kinase (PKC) is a scaffold protein that associates with the 40S subunit in a 1:1 ratio (Adams et al., 2011; Nilsson et al., 2004). Many signaling molecules interact with RACK1 on the ribosome and affect protein synthesis. For example, activated PKC β II binds to the ribosome-bound RACK1 and phosphorylates the initiation factor eIF6, releasing it from the 60S subunit and allowing subunit joining (Ceci et al., 2003). Recent work has also suggested that RACK1 can recruit the miRNA-induced silencing complex (miRISC) to the ribosome, facilitating miRNA-mediated gene repression (Jannot et al., 2011). Other binding partners of RACK1 include integrin receptors (Nilsson et al., 2004), which can potentially localize ribosomes to the cell membrane, and Scp160p in yeast, an RNA-binding protein (Baum et al., 2004; Coyle et al., 2009) (Figure 1.4b). Scp160p associates with specific mRNAs and loss of Scp160p reduces the association of those mRNAs with polysomes (Li et al., 2003). Since Scp160p interacts with the ribosome through RACK1 (Baum et al., 2004; Coyle et al., 2009), these results suggest that Scp160p brings specific mRNAs to the ribosome to enhance their translation. Yeast contains many ribosome-associated proteins (Fleischer et al., 2006) and it is intriguing to hypothesize that additional RNA-intermediaries could associate with the ribosome and may contribute towards transcript-specific translational regulation.

Ribosome-associated factors can also function in processes unrelated to translation. For example, the mammalian TOR complex 2 (mTORC2) is a highly conserved protein kinase that integrates cell growth and metabolism in response to nutrients and growth factors (Zoncu et al., 2011). Following insulin signaling, mTORC2 binds to ribosomes and this interaction activates mTORC2 independently of translation (Zinzalla et al., 2011). Moreover, it appears that ribosome-bound mTORC2 may phosphorylate its substrate AKT co- or postranslationally (Oh et al., 2010). Importantly, these findings suggest that ribosome-associated factors such as mTORC2 can also operate to control specific molecular events independently from a direct role in translational control.

Heterogeneity in rRNA

Heterogeneity in rRNA may also contribute to specialized ribosomal activity. The malaria parasite *Plasmodium sp.* carries 2 different classes of rDNA genes, which allow the use of different “types” of ribosomes during the *Plasmodium*’s life cycle. One form is predominantly found when the parasite is in the sporozoite form (type S) in the mosquito while the other is found when it is in the gametocyte form (type A) in the mouse bloodstream (Gunderson et al., 1987) (Figure 1.5). The two ribosome types seem to possess different functions as substitution of the *Saccharomyces cerevisiae* 25S rRNA with the rRNA of type A *Plasmodium* rRNA produces no phenotype whereas substitution with S *Plasmodium* rRNA is lethal (Velichutina et al., 1998).

Most of the differences between the 2 classes of rRNA in *Plasmodium falciparum* are in the variable regions of the RNA, with only a small number of differences in areas

that are conserved among all eukaryotes. During the course of evolution, from prokaryotes to eukaryotes, the ribosome increased in size and rRNA has changed substantially. For example, *Drosophila* has a 30 nucleotide 2S rRNA, in addition to the 18S, 5.8S, 28S and 5S (Pavlakakis et al., 1979). The most drastic changes in the rRNA are in the variable regions, also known as expansion segments. Expansion segments are generally located on the surface of the ribosome and are very flexible. These expansion segments may have more subtle roles in fine-tuning translation control. ES7^L (also known as D2) is the largest expansion segment. It has expanded from 53nt in *E.coli* to a dramatic 689nt in the mouse (Hassouna et al., 1984). In yeast, where it is 216nt, ES7^L contacts multiple eukaryotic specific ribosomal proteins and potentially interact with P0, to recruit G-proteins to the ribosome (Ben-Shem et al., 2010). It will therefore be very important to delineate the functional roles of expansion segments and whether changes on these sequences influence ribosome function and translation control specificity.

Ribosomal RNA is heavily modified with methylation of 2' hydroxyls, conversion of uridine to pseudouridine and methylation of bases. Most of the modifications are in functionally important areas of the ribosome (Decatur and Fournier, 2002). 2' hydroxyl methylation and conversion of uridine to pseudouridine is aided by C/D box and H/ACA box small nucleolar RNAs (snoRNAs) respectively. There is evidence for the heterogeneity in the expression levels of both C/D box and H/ACA box snoRNAs in different tissues (Castle et al., 2010). The expression of some C/D box snoRNAs has also been shown to oscillate in a circadian manner (Hughes et al., 2012). However, a careful examination on how tissue-specific differences in snoRNA expression directly

impact on rRNA modifications themselves has not been carried out. Morpholino knock down of different snoRNAs in zebrafish cause specific developmental phenotypes. For example, U26 snoRNA-deficient embryos have a decreased body size, and specific defects in brain and eye development (Higa-Nakamine et al., 2011). These results indicate that snoRNAs and snoRNA-mediated modifications on rRNA can have very specific effects in an organism.

Transcript-specific control of ribosome function

Several key questions need to be addressed in order to determine how changes in ribosome composition, structure, and/or activity affect translation control. A central unresolved question is the nature of specific *cis*-regulatory elements within mRNAs that may interface with specialized ribosomes to provide regulatory control as well as whether additional *trans*-acting factors, such as RNA-binding proteins, have a role in this process. In addition, even core components of the ribosome that remain constitutively associated with the translation machinery may exert more regulative effects by interfacing with specific regulatory elements present on subsets of mRNAs. Therefore, a greater understanding of the *cis*-acting elements within mRNAs that promote translational control is an important avenue of exploration. An intriguing hypothesis is that ribosomal proteins could themselves act as “specificity filters”, allowing the ribosome to interface with specific subsets of mRNAs to enhance translation (Mauro and Edelman, 2002). However, the precise functions of only a handful of ribosomal proteins have been identified to date. Here we discuss the nature

of regulatory elements in mRNAs that may imbue regulation specific to the ribosome in recognition and translation of specific subsets of mRNAs.

Internal ribosomal entry site (IRES) elements

In eukaryotic translational initiation, the 40S ribosomal subunit is recruited to the mRNA by translation initiation factors (Sonenberg and Hinnebusch, 2009). The 40S ribosome complex, with associated factors, scans along the 5' untranslated region (UTR) to the initiation codon. Translation initiation factors are released at the initiation codon and the 60S ribosomal subunit joins to form a translationally active 80S ribosome. Most eukaryotic mRNAs are translated via this cap-dependent mechanism, but it is approximated that upwards of 10% of cellular mRNAs rely on cap-independent mechanisms of translational initiation involving IRES elements (Spriggs et al., 2010). The IRES is an RNA structured element positioned at the 5' UTR of specific mRNAs that can recruit the ribosome directly to the initiation region of mRNA with a reduced requirement for canonical initiation factors (Sonenberg and Hinnebusch, 2009). IRES elements were first discovered in RNA viruses which produce uncapped RNA and require these elements for efficient translation of viral mRNAs (Jang et al., 1988; Pelletier and Sonenberg, 1988). Later many additional cellular mRNAs were discovered to contain IRES elements in their 5'UTR. This class of mRNAs is often thought to encompass stress response genes, such as *HIF1 α* , *p53* and *XIAP* and the presence of IRES elements within their 5'UTRs are thought to be active during cellular stress conditions when cap-dependent translation is inhibited.

Modifications in rRNA are important for IRES-mediated translation. In particular, alterations in rRNA pseudouridylation cause a specific defect in the translation of some IRES-containing mRNAs. In mice hypomorphic for dyskerin, the enzyme that converts rRNA uridine to pseudouridine, general cap-dependent translation is not impaired. However, translation of IRES-containing mRNAs including tumor suppressors p27 and p53 is perturbed (Bellodi et al., 2010a, 2010b; Montanaro et al., 2010; Yoon et al., 2006). As a result, these mice have a higher propensity to develop cancer. Importantly, ribosomes that lack pseudouridine modifications show a direct impairment in binding to IRES elements (Jack et al., 2011).

Emerging evidence suggests that ribosomal proteins themselves may function to recognize specific IRES elements. For example, RPS25 is specifically required for IRES-dependent translation of viral mRNAs (Landry et al., 2009). In yeast harboring a deletion in *Rps25*, cap-dependent translational initiation and ribosome fidelity are not affected. However, the translation of two distinct viral IRES elements from the hepatitis C virus (HCV) and cricket paralysis virus (CrPV), is defective (Figure 1.6a). Moreover, RPS25 deficient 40S subunits are unable to bind to the CrPV IRES. Structural studies have placed RPS25 in the head of the 40S subunit, where it can interact with 2 stem loops in the CrPV IRES (Muhs et al., 2011). The function of RPS25 in IRES-mediated translation of viral mRNAs is conserved in mammalian cells (Landry et al., 2009). However, it remains critically unknown whether RPS25 may regulate the translation of specific cellular IRES elements. Consistent with this idea, RPS19 and RPL11, two RPs that are mutated in the human disease Diamond-Blackfan anemia (DBA), regulate IRES-dependent translation of at least two mRNAs important for erythroid differentiation,

Bag1 and *Csde1* (Horos et al., 2011). In addition, mass spectrometry studies have shown that ribosomes bound to the HCV IRES have a different methylation pattern than ribosomes bound to host mRNAs, indicating a role of methylated ribosomal proteins in IRES-mediated translational control (Yu et al., 2005). Studies of these ribosomal proteins, may reflect an emerging theme: that ribosomal proteins may confer more specialized regulation of translation by interfacing with specific IRES-elements in subsets of mRNAs.

Upstream open reading frames (uORFs)

Almost half of all human and mouse transcripts contain at least one uORF (Calvo et al., 2009). Moreover, ribosomal profiling has recently uncovered many near-cognate uORFs that do not start with an AUG codon and, therefore have been difficult to predict computationally (Ingolia et al., 2009, 2011). uORFs can regulate the translation of the downstream main ORF. In cap-dependent translation, the ribosome scans along the 5'UTR to start translation at the first start codon. As it is difficult to reinitiate translation after translating a uORF, most uORFs are repressive to the translation of the downstream main ORF. However, ribosome occupancy on some uORFs with near-cognate, non-AUG start codons correlates positively with the translation of the main ORF. For example, many mRNAs needed for sporulation in yeast contain near-cognate uORFs. Ribosome occupancy at such uORFs is higher in meiotic cells versus vegetative cells, suggesting that there is a shift in the translation machinery during meiosis to translate uORFs (Brar et al., 2012). This transition can change the types of

mRNAs translated, promoting the translation of mRNAs with non-AUG uORFs over those with cognate ones.

In plants, RPL24 is implicated in translational reinitiation of polycistronic mRNAs and mRNAs containing uORFs (Figure 1.6b). *Arabidopsis* mutants in RPL24B have developmental defects in vascular patterning and gynoecium structure, due to reduction in translation of auxin response genes. Interestingly, many genes of the auxin response factor (ARF) class contain uORFs that inhibit the translation of the main ORF. The removal of the uORFs in these genes is able to partially suppress the RPL24B phenotype (Nishimura et al., 2005). RPL24 is located at the interface between 40S and 60S subunits, and may help in subunit joining during translational reinitiation (Ben-Shem et al., 2010). The cauliflower mosaic virus (CaMV) takes advantage of this RPL24 function to promote infection. RPL24 binds to the CaMV transactivator (TAV) and the binding is essential for the translation of the second ORF from the polycistronic CaMV RNA. TAV also binds eIF3, the initiation factor that stimulates the binding of the ternary complex to the 40S. TAV can remain associated with the 40S after termination at the first ORF and potentially help in recruiting eIF3 and RPL24-containing 60S subunits for reinitiation (Park et al., 2001). It will be interesting to see if RPL24 has a role in overcoming uORF-mediated repression also in mammalian cells as a mechanism to promote translational control of specific subsets of uORF containing mRNAs. Moreover, it will be important to determine whether additional ribosomal proteins modulate uORF-mediated regulation of translation control.

Ribosome specificity in cell and developmental biology

As translation rates are a much better predictor of protein levels than mRNA levels, translational control is an important mechanism to regulate the expression of many genes (Schwanhäusser et al., 2011). The possibility of greater specificity in translation through a more regulatory ribosomal function of the ribosome opens an important avenue of exploration into the cellular and developmental processes that are regulated by this new level of control. This is a very rapidly emerging field and although much remains to be addressed, clear examples of specialized ribosome functions are becoming evident across all kingdoms of life.

Ribosome specificity in bacteria

In bacteria, a new paradigm is emerging for heterogeneity in ribosomes as a mechanism for stress adaptation. Remarkably, formation of 'stress-ribosomes' that possess rRNAs with sequence-specific differences allows for the selective translation of specific mRNAs (Figure 1.7a). In *E.coli*, the sequence-specific endonuclease MazF that preferentially cleaves single-stranded mRNAs at ACA sequences is activated following stress. Although MazF induction blocks the synthesis of most proteins, the translation of ~10% of all transcripts is nevertheless maintained, and some of these mRNAs encode for proteins that promote cell death. Strikingly, MazF is able to selectively promote translation of these transcripts by cleaving ACA sites at or closely upstream of the AUG start codon to generate leaderless mRNAs. MazF also changes the type of ribosomes that are produced to translate these leaderless mRNA by cleaving the 16S rRNA near the 3' end, releasing the anti-Shine-Dalgarno (aSD) sequence. Translational initiation in

bacteria depends heavily on base pairing between the Shine Dalgarno (SD) sequence in the mRNA and the aSD in the 16S rRNA (Kozak, 1999). Removal of the aSD by MazF renders a subpopulation of “specialized ribosomes” less efficient in translating most mRNAs, but improves translation of leaderless mRNAs that no longer possess the SD sequence (Vesper et al., 2011).

The role of ribosome heterogeneity in plants

In plants, mutations in ribosomal proteins give a wide variety of phenotypes. Loss of function mutations in a number of ribosomal proteins are embryonic lethal while others have more varied patterning and developmental defects (Byrne, 2009; Szakonyi et al., 2011). For example, mutations in *RPL10a*, *RPL9* and *RPL5* give a piggyback phenotype, where mutants have pointed leaves and pronounced marginal serrations. These genes are crucial in dorsal-ventral patterning in leaves, possibly through translational regulation of dorsal genes of the HD-ZIPIII family and ventral genes of the KANADI family (Pinon et al., 2008). *Arabidopsis* mutants in *RPL24B* have developmental defects in vascular patterning and gynoecium structure, due to reduction in translation of auxin response genes (Nishimura et al., 2005). As discussed, the effect of *RPL24* may be specifically at the level of translation reinitiation of mRNAs that harbor uORFs (see above). Although a clear mechanism for translational regulation by these ribosomal proteins is not fully understood, it is clear that ribosomal proteins have an unexpected and important role in regulating specific developmental processes and pattern formation (Byrne, 2009).

Ribosome specificity in invertebrate embryonic development

In *Drosophila*, loss of a copy of several ribosomal proteins leads to a Minute phenotype. Minute flies are smaller than their wild type counterparts, have delayed larval development and short, thin bristles (Marygold et al., 2007). This phenotype has been attributed to a general decrease in protein synthesis. However, some ribosomal protein mutants show additional patterning defects. For example, some Minutes have a reduction in wing size, while at least two additional ribosomal protein mutants display an increase in wing size as well as wing vein patterning defects (Brehme, 1941; Marygold et al., 2005). This suggests that some ribosomal proteins may have more specific requirements for embryonic development and patterning.

Germline formation in *Drosophila* may require specific ribosomes that are only produced in the mitochondria (Figure 1.7b). Mitochondrial ribosomes, which are typically smaller, resemble prokaryotic ribosomes and are required to translate the mitochondria genome. Strikingly, however, these ribosomes can be localized outside of mitochondria in polar granules, which are large ribonucleoprotein complexes found at the posterior end of the *Drosophila* embryo required for germ cell formation (Amikura et al., 2001; Kobayashi et al., 1993). Polysomes on polar granules contain both mitochondrial and cytoplasmic ribosomes. Blocking mitochondrial ribosomes with prokaryotic translation inhibitors kasugamycin and chloramphenicol, both of which do not affect cytoplasmic ribosomes, causes a decrease in the amount of Gcl protein, a germ cell specific protein, in polar granules and a defect in germ cell formation. The translation of another germ cell mRNA, *Nos*, is not affected by the drugs, suggesting that mitochondrial ribosomes are only responsible for the proper translation of a subset of mRNAs (Amikura et al.,

2005). Importantly, further studies on the codon usage of the *Gcl* mRNA are critically required to determine whether mitochondrial ribosomes are indeed participating in this level of translational control. Although it is not entirely understood why mitochondrial ribosomes in the cytoplasm may possess unique properties that endow them with the ability to translate specific cytoplasmic mRNAs, these studies suggest specificity in translational regulation governed at the level of heterogeneous ribosomes.

Ribosome specificity in vertebrate embryonic development

Ribosomal proteins are important regulators of vertebrate embryonic development. For example, morpholino knockdown of ribosomal proteins in zebrafish produce phenotypes that can be both unique and specific, resulting in abnormalities in brain, body trunk, eye and ear development (Uechi et al., 2006). How can mutations in ribosomal proteins cause tissue-specific effects? One important paradigm that has informed us on the role of ribosomal proteins in control of specialized ribosome activity stems from analysis of RPL38, which links a component of the ribosome to formation of the mammalian body plan (Figure 1.7c). In a forward genetic screen, *Rpl38* was identified as the gene deleted in the Tail short (*Ts*) mouse. *Ts*/⁺ mice show highly specific defects in the formation of the axial skeleton (Kondrashov et al., 2011). These include homeotic transformations, a change in the identity of one skeletal element to another, suggesting an unexpected regulatory role for ribosomes during tissue patterning. Hox genes have been shown to be key regulators of vertebrae segment identity and the loss-of-function of specific Hox genes causes homeotic transformations (Alexander et al., 2009; Wellik, 2007). *Ts*/⁺ embryos also show a defect in the

specification of Pea3⁺ motor neurons, a process that also requires Hox-directed patterning. Strikingly, *Ts/+* embryos show no defects in general protein synthesis, however, the translation of 8 of the 39 Hox mRNAs is specifically reduced. At the molecular level, RPL38 facilitates the formation of 80S complex during translation initiation on these mRNAs, while being part of the ribosome. RPL38's effects are highly specific as other ribosomal protein deficiencies in the mouse do not give rise to the same phenotypes, even when general rates of protein synthesis are decreased. Interestingly, *Ts/+* phenotypes are inherited in a dominant form, as haploinsufficiency in RPL38 expression is sufficient to manifest in tissue patterning defects. This suggests that RPL38 is rate-limiting for the translation of selective transcripts, such as Hox mRNAs. Moreover, this is consistent with the fact that Hox mutants also display phenotypes in a heterozygous condition thereby revealing that precise levels of Hox proteins are necessary for establishing the mammalian body plan. It will be important to determine whether RPL38 recognizes unique sequences or structures within subsets of mRNAs either directly or through interactions with other RNA-binding proteins.

The tissue specific phenotypes associated with RPL38 loss of function are also, at least in part, explained by the observation that RPL38 expression is highly regulated during embryonic development. A striking enrichment of *Rpl38* transcripts is evident in specific regions of the embryo such as developing somites, the precursors of vertebrae, as well as within the neural tube that mirror where tissue specific phenotypes associated with loss of RPL38 expression are observed (Kondrashov et al., 2011). These findings suggest that the enriched expression of specific ribosomal proteins may

produce heterogeneous ribosomes in distinct cell and tissue types with unique specificities in translating specific classes of mRNAs.

In addition to the *Ts* mice, the few ribosomal protein mouse mutants that have been generated or characterized to date display specific phenotypes. For example, mice carrying a dominant mutation in *Rpl24* (*Bst/+*) display a white ventral midline spot, white hind feet and kinked tail. These mice also have a reduction in the number of retinal ganglion cells as well as extra digits in the limbs (Oliver et al., 2004). On the contrary, *Rpl29*^{-/-} mice are smaller in size and possesses specific defects in osteogenesis associated with increased bone fragility (Oristian et al., 2009). These phenotypes suggest that additional ribosomal proteins may have more specialized roles during embryonic development.

Linking ribosome specificity to disease

Loss of some ribosomal proteins has been associated with certain morphological abnormalities, and this phenotype may rely on the activation of a p53-dependent stress response. A subset of ribosomal proteins are able to interact with MDM2, an inhibitor of p53, and thereby affect p53 activity (Zhang and Lu, 2009). Morpholino knockdown of *Rpl11*, *Rps7* and *Rps19* in zebrafish causes an upregulation of p53 and certain phenotypes can be partially rescued by knockdown of *p53* (Chakraborty et al., 2009; Danilova et al., 2008; Duan et al., 2011). Importantly, p53 cannot fully explain the phenotypes in ribosomal protein mutants. For example the tissue-specific defects associated with *Rpl38* loss of function were not reversed following reduction of p53 levels (Kondrashov et al., 2011). Moreover, in *Rps6*^{+/-};*p53*^{-/-} mouse embryos, tissue

specific phenotypes are present and these include defects in placenta and liver development (Panić et al., 2006).

In humans, accumulating evidence links mutations in a number of ribosomal proteins including *RPS19*, *RPS24*, *RPS17*, *RPL5*, *RPL11*, and *RPL35a* to Diamond-Blackfan anemia (DBA) (Boria et al., 2010). Notably, DBA patients carrying ribosomal protein mutations also exhibit tissue-specific defects including limb defects, cleft palate, abnormalities in heart development, growth failure as well a predisposition to cancer. Mutations in different ribosomal proteins also give rise to different types of birth defects. For instance, most patients with *RPL5* mutations have a cleft palate, whereas most individuals with *RPL11* mutations do not have any craniofacial defects (Gazda et al., 2008). In addition, mutations in *RPL21* have been identified in hereditary hypotrichosis simplex (HHS) that is characterized by progressive hair loss early in childhood (Zhou et al., 2011). Together, genetic disorders of linked mutations in ribosomal proteins in humans show remarkably specific phenotypes, suggesting that ribosomal proteins have unique functions in different tissues.

Intracellular localization of specialized ribosomes

A clue into the possible roles for specialized ribosomal protein function in specific cells types comes from research in neurons. Neurons are extremely polarized cells where axons and dendrites perform very different functions from the cell body. Local protein synthesis in neuronal dendrites is important in remodeling synaptic connections and has a role in learning and memory (Bramham and Wells, 2007) (Figure 1.7d). Long term potentiation (LTP), a cellular measure of memory formation, can be induced by

translation in hippocampal dendrites, where cell bodies have been severed (Vickers et al., 2005). Polysomes are found in dendrites and the number of polysomes in dendritic spines increases after initiation of LTP (Ostroff et al., 2002). Thereby, mRNAs are translated in an extremely localized manner, although the mechanism for this level of regulation is poorly understood.

Strikingly, transcripts for specific subsets of ribosomal proteins are enriched up to 10 fold at these sites compared to the cell body (Moroz et al., 2006). Since most ribosomal proteins are assembled onto the ribosome in the nucleus, it is surprising to find stores of specific ribosomal protein transcripts in neuronal processes. Additional components of the translational machinery have also been found in neuronal processes and are transported as part of large ribonucleoprotein particles (RNP) that contain mRNA, ribosomes and other RNA binding proteins (Moccia et al., 2003; Sossin and DesGroseillers, 2006). It has been suggested that ribosomal subunits are stored as inactive forms in RNA granules and are activated by the addition of locally synthesized ribosomal proteins (Moccia et al., 2003). Thus, ribosomes at these processes may be distinct from ones in the cell body due to the presence of newly, locally synthesized ribosomal proteins that may influence the translation of mRNAs that are targeted for the dendrites. In support of this hypothesis, recent studies have shown that the transmembrane receptor DCC associates with specific ribosomal proteins and spatially regulates translation in neuronal axons and dendrites (Tcherkezian et al., 2010). Interestingly, some mRNAs present in dendrites are known to contain IRES-elements (Pinkstaff et al., 2001). By extension, similar localized mechanisms for translational control may provide unique regulation in multiple cellular contexts, including

translational regulation in a polarized cell, at the leading edge of a migrating cell, or during specific steps in cell differentiation.

Perspectives

Several parallels can be drawn between what has been termed a DNA code with that of a possible 'ribosome code'. Similar to DNA that is modified, rRNA is also extensively modified. Histones, once considered boring, housekeeping proteins are now clearly recognized as active participants in chromatin remodeling and transcriptional control through exquisite post-translational modifications identified within histone tails. Likewise, the view of ribosomal proteins as only carrying out rote-like functions is undergoing a paradigm shift. Post-translational modifications as well as differences in ribosomal protein composition, may add greater regulatory specificity to the RNA-based translation machinery, as do histones to DNA. However, several outstanding questions need to be addressed before a 'ribosome code' for translational control could be accepted.

First, as discussed in this chapter, many examples now exist across all kingdoms of life that greater heterogeneity characterizes the ribosome, both at the level of rRNA and ribosomal proteins. However, despite these examples, this has not been explored in a systematic manner. Most ribosomes are purified from tissues containing an amalgamation of multiple cell types, and any differences in ribosome composition in a particular cell may be lost in such an ensemble. Studies within cell lines have the advantage of being able to identify heterogeneity in a single cell type; however, this approach lacks the dynamics of an *in vivo* system where more variations may be

present. A systems approach is necessary to look more globally across different cell types, perhaps during early stages of organogenesis or during specific steps in cell differentiation to determine the extent of variation in ribosomes. It is also tempting to speculate that a heterogeneous pool of ribosomes may already exist in a single cell, for example in germ cells and neurons. If so, how is the stoichiometry and production of such heterogeneous pools of ribosomes regulated? What regulates the intracellular localization of heterologous ribosomes in a single cell? Second, although examples of heterogeneity are evident, little is known on how this impacts translational control. For example, to date classes of mRNAs that may respond to 'specialized' ribosomes remain elusive. The discovery of a specific requirement for RPL38 containing ribosomes in facilitating translational regulation of subsets of Homeobox-containing mRNAs hints at exciting possible future work into this problem. Third, virtually nothing is known about the upstream regulatory signals that might produce ribosome heterogeneity in the first place. The realization that core components of the ribosome may be more highly regulated opens a new avenue of exploration in identifying upstream signaling pathways and trans-acting factors that control their expression. Lastly, it will be crucial to identify *cis*-acting 'translational regulons' within mRNAs that interface with specialized ribosomes to confer translational specificity. As discussed, examples of these important elements may include IRES elements and uORFs. Moreover, it is equally important to determine whether even core components of the ribosome that show little variation may also confer more specialized *activity* by virtue of their ability to interact with such specific translation regulons. Emerging evidence suggests that there is a great disconnect between the expression of the genome at the level of the transcriptosome versus the

proteome and that this thereby leaves ample room for regulation at the level of translational control (Schwanhäusser et al., 2011). In this respect, variations in the levels at which specific protein products are made are likely to have tremendous importance to cell biology. Thereby, we hypothesize that it will be important to conceptualize translational control in the same light as transcriptional control where enhancers or attenuators may fine-tune protein abundance, culminating in unique readouts with important biological significance.

Ultimately, a key question is how ribosome-mediated gene regulation may shape our view of organismal biology. It has long been known that a small handful of signaling molecules (such as sonic hedgehog, fibroblast growth factors (FGF), Wnt ligands and bone morphogenetic proteins (Bmps)) act many times during embryonic development and in distinct locations to produce entirely different tissues or structures (for example limbs vs motor neurons). Transcriptional profiling has not provided a complete explanation for these different mechanisms of action. The hypothesis that developmentally regulated signals directly converge on ribosome specificity opens an entirely new avenue of exploration into whether rapid changes in the production of proteins from pre-existing pools of mRNAs may instead influence fundamental aspects of mammalian cell and embryonic development.

The following work is an investigation into the *cis*-acting translational regulons that allow for regulation by the ribosome. Previously our lab has shown that RPL38 confers specificity to the ribosome and is responsible for specifically translating a subset of Homeobox (Hox) mRNAs to regulate the mammalian body plan. In *Rpl38*^{Ts/+} embryos, global protein synthesis is unchanged while the translation of 8 of the 39 Hox mRNAs is

specifically affected. Unexpectedly, *Rpl38* transcripts are highly enriched in a tissue specific manner, most notably in somites and within the neural tube (Kondrashov et al., 2011). These studies led us to propose that ribosome in distinct regions of the embryo may be highly “specialized” to exert transcript-specific translational control. However, the mechanism by which mRNAs are recognized by specialized ribosomes remains unknown.

Using RPL38 and the Hox mRNAs as a model, I identified unique combinations of RNA regulons in 5'UTRs that confer ribosome specificity. In Chapter 2, I report the discovery of IRESs and a novel RNA regulon, the Translation Inhibitory Element (TIE), which together, facilitate ribosome-mediated control of gene expression in mammalian development. Chapter 3 delves further into the mechanism of translational regulation by the TIE and IRES. Together, this work uncovers a new paradigm of specialized translational regulation that is achieved through a combination of unique RNA regulons within the mRNA template to promote ribosome-mediated control of gene expression and organismal development.

Acknowledgements

We would like to thank Drs. Davide Ruggero, Cristian Bellodi, Mary McMahon, Craig Stumpf and members of the Barna lab for discussion and critical reading of the manuscript. S.X. is supported by the Agency of Science, Technology and Research of Singapore. This work was supported by the Program for Breakthrough Biomedical Research, UCSF (M.B.) and the NIH Director's New Innovator Award (M.B.).

References

- Adams, D.R., Ron, D., and Kiely, P.A. (2011). RACK1, A multifaceted scaffolding protein: Structure and function. *Cell Commun. Signal.* 9, 22.
- Alberts, B. (1998). The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell* 92, 291–294.
- Alexander, T., Nolte, C., and Krumlauf, R. (2009). Hox genes and segmentation of the hindbrain and axial skeleton. *Annu. Rev. Cell Dev. Biol.* 25, 431–456.
- Amikura, R., Kashikawa, M., Nakamura, A., and Kobayashi, S. (2001). Presence of mitochondria-type ribosomes outside mitochondria in germ plasm of *Drosophila* embryos. *Proc. Natl. Acad. Sci. U. S. A.* 98, 9133–9138.
- Amikura, R., Sato, K., and Kobayashi, S. (2005). Role of mitochondrial ribosome-dependent translation in germline formation in *Drosophila* embryos. *Mech. Dev.* 122, 1087–1093.
- Balasubramanian, S., Zheng, D., Liu, Y.-J., Fang, G., Frankish, A., Carriero, N., Robilotto, R., Cayting, P., and Gerstein, M. (2009). Comparative analysis of processed ribosomal protein pseudogenes in four mammalian genomes. *Genome Biol.* 10, R2.
- Barakat, A., Szick-Miranda, K., Chang, I.-F., Guyot, R., Blanc, G., Cooke, R., Delseny, M., and Bailey-Serres, J. (2001). The Organization of Cytoplasmic Ribosomal Protein Genes in the *Arabidopsis* Genome 1. *Plant Physiol.* 127, 398–415.
- Baum, S., Bittins, M., Frey, S., and Sedorf, M. (2004). Asc1p, a WD40-domain containing adaptor protein, is required for the interaction of the RNA-binding protein Scp160p with polysomes. *Biochem. J.* 380, 823–830.
- Baxter-Roshek, J.L., Petrov, A.N., and Dinman, J.D. (2007). Optimization of ribosome structure and function by rRNA base modification. *PLoS One* 2, e174.
- Bellodi, C., Krasnykh, O., Haynes, N., Theodoropoulou, M., Peng, G., Montanaro, L., and Ruggero, D. (2010a). Loss of function of the tumor suppressor DKC1 perturbs p27 translation control and contributes to pituitary tumorigenesis. *Cancer Res.* 70, 6026–6035.
- Bellodi, C., Kopmar, N., and Ruggero, D. (2010b). Deregulation of oncogene-induced senescence and p53 translational control in X-linked dyskeratosis congenita. *EMBO J.* 29, 1865–1876.
- Ben-Shem, A., Jenner, L., Yusupova, G., and Yusupov, M. (2010). Crystal structure of the eukaryotic ribosome. *Science* 330, 1203–1209.

Boria, I., Garelli, E., Gazda, H.T., Aspesi, A., Quarello, P., Pavesi, E., Ferrante, D., Meerpohl, J.J., Kartal, M., Da Costa, L., et al. (2010). The ribosomal basis of Diamond-Blackfan Anemia: mutation and database update. *Hum. Mutat.* **31**, 1269–1279.

Bortoluzzi, S., D'Alessi, F., Romualdi, C., and Danieli, G.A. (2001). Differential expression of genes coding for ribosomal proteins in different human tissues. *Bioinformatics* **17**, 1152–1157.

Bramham, C.R., and Wells, D.G. (2007). Dendritic mRNA: transport, translation and function. *Nat. Rev. Neurosci.* **8**, 776–789.

Brar, G. a, Yassour, M., Friedman, N., Regev, A., Ingolia, N.T., and Weissman, J.S. (2012). High-Resolution View of the Yeast Meiotic Program Revealed by Ribosome Profiling. *Science* **335**, 552–557.

Brehme, K.S. (1941). Development of the Minute phenotype in *Drosophila Melanogaster*. A comparative study of the growth of three Minute mutants. *J. Exp. Zool.* **88**, 135–160.

Byrne, M.E. (2009). A role for the ribosome in development. *Trends Plant Sci.* **14**, 512–519.

Calvo, S.E., Pagliarini, D.J., and Mootha, V.K. (2009). Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 7507–7512.

Carroll, A.J., Heazlewood, J.L., Ito, J., and Millar, a H. (2008). Analysis of the *Arabidopsis* cytosolic ribosome proteome provides detailed insights into its components and their post-translational modification. *Mol. Cell. Proteomics* **7**, 347–369.

Castle, J.C., Armour, C.D., Löwer, M., Haynor, D., Biery, M., Bouzek, H., Chen, R., Jackson, S., Johnson, J.M., Rohl, C. a, et al. (2010). Digital genome-wide ncRNA expression, including SnoRNAs, across 11 human tissues using polyA-neutral amplification. *PLoS One* **5**, e11779.

Ceci, M., Gaviraghi, C., Gorrini, C., Sala, L., Offenhauser, N., Marchisio, P.C., and Biffo, S. (2003). Release of eIF6 (p27BBP) from the 60S subunit allows 80S ribosome assembly. *Nature* **426**, 579–584.

Chakraborty, A., Uechi, T., Higa, S., Torihara, H., and Kenmochi, N. (2009). Loss of ribosomal protein L11 affects zebrafish embryonic development through a p53-dependent apoptotic response. *PLoS One* **4**, e4152.

Colón-Ramos, D.A., Shenvi, C.L., Weitzel, D.H., Gan, E.C., Matts, R., Cate, J., and Kornbluth, S. (2006). Direct ribosomal binding by a cellular inhibitor of translation. *Nat. Struct. Mol. Biol.* **13**, 103–111.

Coyle, S.M., Gilbert, W. V, and Doudna, J. a (2009). Direct link between RACK1 function and localization at the ribosome in vivo. *Mol. Cell. Biol.* 29, 1626–1634.

Crick, F.H.C. (1968). The origin of the genetic code. *J. Mol. Biol.* 38, 367–379.

Dabeva, M.D., and Warner, J.R. (1993). Ribosomal protein L32 of *Saccharomyces cerevisiae* regulates both splicing and translation of its own transcript. *J. Biol. Chem.* 268, 19669–19674.

Danilova, N., Sakamoto, K.M., and Lin, S. (2008). Ribosomal protein S19 deficiency in zebrafish leads to developmental abnormalities and defective erythropoiesis through activation of p53 protein family. *Blood* 112, 5228–5237.

Decatur, W.A., and Fournier, M.J. (2002). rRNA modifications and ribosome function. *Trends Biochem. Sci.* 27, 344–351.

Degenhardt, R.F., and Bonham-Smith, P.C. (2008a). Arabidopsis ribosomal proteins RPL23aA and RPL23aB are differentially targeted to the nucleolus and are disparately required for normal development. *Plant Physiol.* 147, 128–142.

Degenhardt, R.F., and Bonham-Smith, P.C. (2008b). Transcript profiling demonstrates absence of dosage compensation in Arabidopsis following loss of a single RPL23a paralog. *Planta* 228, 627–640.

Duan, J., Ba, Q., Wang, Z., Hao, M., Li, X., Hu, P., Zhang, D., Zhang, R., and Wang, H. (2011). Knockdown of ribosomal protein S7 causes developmental abnormalities via p53 dependent and independent pathways in zebrafish. *Int. J. Biochem. Cell Biol.* 43, 1218–1227.

Enyenihi, A.H., and Saunders, W.S. (2003). Large-scale functional genomic analysis of sporulation and meiosis in *Saccharomyces cerevisiae*. *Genetics* 54, 47–54.

Falcone Ferreyra, M.L., Pezza, A., Biarc, J., Burlingame, A.L., and Casati, P. (2010). Plant L10 ribosomal proteins have different roles during development and translation under ultraviolet-B stress. *Plant Physiol.* 153, 1878–1894.

Ferreira-Cerca, S., Pöll, G., Kühn, H., Neueder, A., Jakob, S., Tschochner, H., and Milkereit, P. (2007). Analysis of the in vivo assembly pathway of eukaryotic 40S ribosomal proteins. *Mol. Cell* 28, 446–457.

Fisher, E.M., Beer-Romero, P., Brown, L.G., Ridley, A., McNeil, J.A., Lawrence, J.B., Willard, H.F., Bieber, F.R., and Page, D.C. (1990). Homologous ribosomal protein genes on the human X and Y chromosomes: escape from X inactivation and possible implications for Turner syndrome. *Cell* 63, 1205–1218.

Fleischer, T.C., Weaver, C.M., McAfee, K.J., Jennings, J.L., and Link, A.J. (2006). Systematic identification and functional screens of uncharacterized proteins associated with eukaryotic ribosomal complexes. *Genes Dev.* **20**, 1294–1307.

Frank, J. (2000). The ribosome--a macromolecular machine par excellence. *Chem. Biol.* **7**, R133–R141.

Fuchs, G., Diges, C., Kohlstaedt, L. a, Wehner, K. a, and Sarnow, P. (2011). Proteomic Analysis of Ribosomes: Translational Control of mRNA Populations by Glycogen Synthase GYS1. *J. Mol. Biol.* **410**, 118–130.

Gazda, H.T., Sheen, M.R., Vlachos, A., Choesmel, V., O'Donohue, M.-F., Schneider, H., Darras, N., Hasman, C., Sieff, C. a, Newburger, P.E., et al. (2008). Ribosomal protein L5 and L11 mutations are associated with cleft palate and abnormal thumbs in Diamond-Blackfan anemia patients. *Am. J. Hum. Genet.* **83**, 769–780.

Green, H., Canfield, A.E., Hillarby, M.C., Grant, M.E., Boot-Handford, R.P., Freemont, A.J., and Wallis, G.A. (2000). The ribosomal protein QM is expressed differentially during vertebrate endochondral bone development. *J. Bone Miner. Res.* **15**, 1066–1075.

Gunderson, J., Sogin, M., Wollett, G., and Hollingdale, M. (1987). Structurally distinct, stage-specific ribosomes occur in *Plasmodium*. *Science* **238**, 933–937.

Hardy, S.J. (1975). The stoichiometry of the ribosomal proteins of *Escherichia coli*. *Mol. Gen. Genet.* **140**, 253–274.

Hariharan, N., Kelley, D.E., and Perry, R.P. (1989). Equipotent mouse ribosomal protein promoters have a similar architecture that includes internal sequence elements. *Genes Dev.* **3**, 1789–1800.

Hassouna, N., Michot, B., and Bachellerie, J. (1984). The complete nucleotide sequence of mouse 28S rRNA gene. Implications for the process of size increase of the large subunit rRNA in higher eukaryotes. *Nucleic Acids Res.* **12**, 3563–3583.

Held, W.A., Mizushima, S., and Nomura, M. (1973). Reconstitution from Purified of *Escherichia coli* 30 S Ribosomal Molecular Components. *Proteins* **248**, 5223–5739.

Henras, A.K., Soudet, J., G rus, M., Lebaron, S., Caizergues-Ferrer, M., Mougin, A., and Henry, Y. (2008). The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cell. Mol. Life Sci.* **65**, 2334–2359.

Higa-Nakamine, S., Suzuki, T., Uechi, T., Chakraborty, A., Nakajima, Y., Nakamura, M., Hirano, N., Suzuki, T., and Kenmochi, N. (2011). Loss of ribosomal RNA modification causes developmental defects in zebrafish. *Nucleic Acids Res.* **40**, 391–398.

Horos, R., IJspeert, H., Pospisilova, D., Sendtner, R., Andrieu-Soler, C., Taskesen, E., Nieradka, A., Cmejla, R., Sendtner, M., Touw, I.P., et al. (2011). Ribosomal deficiencies in Diamond-Blackfan-Anemia impair translation of transcripts essential for differentiation of murine and human erythroblasts. *Blood* 119, 262–272.

Hsieh, A.C., Liu, Y., Edlind, M.P., Ingolia, N.T., Janes, M.R., Sher, A., Shi, E.Y., Stumpf, C.R., Christensen, C., Bonham, M.J., et al. (2012). The translational landscape of mTOR signalling steers cancer initiation and metastasis. *Nature* 485, 55–61.

Hughes, M.E., Grant, G.R., Paquin, C., Qian, J., and Nitabach, M.N. (2012). Deep sequencing the circadian and diurnal transcriptome of *Drosophila* brain. *Genome Res.*

Hughes, T.R., Roberts, C.J., Dai, H., Jones, a R., Meyer, M.R., Slade, D., Burchard, J., Dow, S., Ward, T.R., Kidd, M.J., et al. (2000). Widespread aneuploidy revealed by DNA microarray expression profiling. *Nat. Genet.* 25, 333–337.

Ingolia, N.T., Ghaemmighami, S., Newman, J.R.S., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* 324, 218–223.

Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity and Dynamics of Mammalian Proteomes. *Cell* 147, 789–802.

Jack, K., Bellodi, C., Landry, D.M., Niederer, R.O., Meskauskas, A., Musalgaonkar, S., Kopmar, N., Krasnykh, O., Dean, A.M., Thompson, S.R., et al. (2011). rRNA Pseudouridylation Defects Affect Ribosomal Ligand Binding and Translational Fidelity from Yeast to Human Cells. *Mol. Cell* 44, 660–666.

Jang, S.K., Kräusslich, H.G., Nicklin, M.J., Duke, G.M., Palmenberg, A.C., and Wimmer, E. (1988). A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J. Virol.* 62, 2636–2643.

Jannot, G., Bajan, S., Giguère, N.J., Bouasker, S., Banville, I.H., Piquet, S., Hutvagner, G., and Simard, M.J. (2011). The ribosomal protein RACK1 is required for microRNA function in both *C. elegans* and humans. *EMBO Rep.* 12, 581–586.

Johnson, A.W., Lund, E., and Dahlberg, J. (2002). Nuclear export of ribosomal subunits. *Trends Biochem. Sci.* 27, 580–585.

Kapasi, P., Chaudhuri, S., Vyas, K., Baus, D., Komar, A. a, Fox, P.L., Merrick, W.C., and Mazumder, B. (2007). L13a blocks 48S assembly: role of a general initiation factor in mRNA-specific translational control. *Mol. Cell* 25, 113–126.

Kearse, M.G., Chen, A.S., and Ware, V.C. (2011). Expression of ribosomal protein L22e family members in *Drosophila melanogaster*: rpL22-like is differentially expressed and alternatively spliced. *Nucleic Acids Res.* 39, 2701–2716.

Kellis, M., Birren, B.W., and Lander, E.S. (2004). Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* 428, 617–624.

Kim, C.H., and Warner, J.R. (1983). Messenger RNA for Ribosomal Proteins in Yeast. *J. Mol. Biol.* 165, 79–89.

Kobayashi, S., Amikura, R., and Okada, M. (1993). Presence of Mitochondrial Large Ribosomal RNA Outside Mitochondria in Germ Plasm of *Drosophila melanogaster*. *Science* (80-.). 260, 1521–1524.

Komili, S., Farny, N.G., Roth, F.P., and Silver, P.A. (2007). Functional specificity among ribosomal proteins regulates gene expression. *Cell* 131, 557–571.

Kondrashov, N., Pusic, A., Stumpf, C.R., Shimizu, K., Hsieh, A.C., Xue, S., Ishijima, J., Shiroishi, T., and Barna, M. (2011). Ribosome-Mediated Specificity in Hox mRNA Translation and Vertebrate Tissue Patterning. *Cell* 145, 383–397.

Kozak, M. (1999). Initiation of translation in prokaryotes and eukaryotes. *Gene* 234, 187–208.

Kressler, D., Hurt, E., and Bassler, J. (2010). Driving ribosome assembly. *Biochim. Biophys. Acta* 1803, 673–683.

Krieg, J., Hofsteenge, J., and Thomas, G. (1988). Identification of the 40 S ribosomal protein S6 phosphorylation sites induced by cycloheximide. *J. Biol. Chem.* 263, 11473–11477.

Lam, Y.W., Lamond, A.I., Mann, M., and Andersen, J.S. (2007). Analysis of nucleolar protein dynamics reveals the nuclear degradation of ribosomal proteins. *Curr. Biol.* 17, 749–760.

Landry, D.M., Hertz, M.I., and Thompson, S.R. (2009). RPS25 is essential for translation initiation by the Dicistroviridae and hepatitis C viral IRESs. *Genes Dev.* 23, 2753–2764.

Lee, S.-W., Berger, S.J., Martinović, S., Pasa-Tolić, L., Anderson, G.A., Shen, Y., Zhao, R., and Smith, R.D. (2002). Direct mass spectrometric analysis of intact proteins of the yeast large ribosomal subunit using capillary LC/FTICR. *Proc. Natl. Acad. Sci. U. S. A.* 99, 5942–5947.

Li, A.-M., Watson, A., and Fridovich-Keil, J.L. (2003). Scp160p associates with specific mRNAs in yeast. *Nucleic Acids Res.* **31**, 1830–1837.

Lopes, A.M., Miguel, R.N., Sargent, C.A., Ellis, P.J., Amorim, A., and Affara, N. a (2010). The human RPS4 paralogue on Yq11.223 encodes a structurally conserved ribosomal protein and is preferentially expressed during spermatogenesis. *BMC Mol. Biol.* **11**, 33.

Malygin, A.A., Parakhnevitch, N.M., Ivanov, A. V, Eperon, I.C., and Karpova, G.G. (2007). Human ribosomal protein S13 regulates expression of its own gene at the splicing step by a feedback mechanism. *Nucleic Acids Res.* **35**, 6414–6423.

Martín-Marcos, P., Hinnebusch, A.G., and Tamame, M. (2007). Ribosomal protein L33 is required for ribosome biogenesis, subunit joining, and repression of GCN4 translation. *Mol. Cell. Biol.* **27**, 5968–5985.

Marygold, S.J., Coelho, C.M. a, and Leever, S.J. (2005). Genetic analysis of RpL38 and RpL5, two minute genes located in the centric heterochromatin of chromosome 2 of *Drosophila melanogaster*. *Genetics* **169**, 683–695.

Marygold, S.J., Roote, J., Reuter, G., Lambertsson, A., Ashburner, M., Millburn, G.H., Harrison, P.M., Yu, Z., Kenmochi, N., Kaufman, T.C., et al. (2007). The ribosomal protein genes and Minute loci of *Drosophila melanogaster*. *Genome Biol.* **8**, R216.

Mauro, V.P., and Edelman, G.M. (2002). The ribosome filter hypothesis. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 12031–12036.

Mazumder, B., and Fox, P.L. (1999). Delayed translational silencing of ceruloplasmin transcript in gamma interferon-activated U937 monocytic cells: role of the 3' untranslated region. *Mol. Cell. Biol.* **19**, 6898–6905.

Mazumder, B., Sampath, P., Seshadri, V., Maitra, R.K., Dicorleto, P.E., and Fox, P.L. (2003). Regulated Release of L13a from the 60S Ribosomal Subunit as A Mechanism of Transcript-Specific Translational Control. *Cell* **115**, 187–198.

Mills, A.A., Mills, M.J., Gardiner, D.M., Bryant, S. V, and Stanbridge, E.J. (1999). Analysis of the pattern of QM expression during mouse development. *Differentiation*. **64**, 161–171.

Moccia, R., Chen, D., Lyles, V., Kapuya, E., E, Y., Kalachikov, S., Spahn, C.M.T., Frank, J., Kandel, E.R., Barad, M., et al. (2003). An unbiased cDNA library prepared from isolated *Aplysia* sensory neuron processes is enriched for cytoskeletal and translational mRNAs. *J. Neurosci.* **23**, 9409–9417.

Montanaro, L., Calienni, M., Bertoni, S., Rocchi, L., Sansone, P., Storci, G., Santini, D., Ceccarelli, C., Taffurelli, M., Carnicelli, D., et al. (2010). Novel dyskerin-mediated

mechanism of p53 inactivation through defective mRNA translation. *Cancer Res.* 70, 4767–4777.

Moroz, L.L., Edwards, J.R., Puthanveetil, S. V, Kohn, A.B., Ha, T., Heyland, A., Knudsen, B., Sahni, A., Yu, F., Liu, L., et al. (2006). Neuronal transcriptome of Aplysia: neuronal compartments and circuitry. *Cell* 127, 1453–1467.

Muhs, M., Yamamoto, H., Ismer, J., Takaku, H., Nashimoto, M., Uchiumi, T., Nakashima, N., Mielke, T., Hildebrand, P.W., Nierhaus, K.H., et al. (2011). Structural basis for the binding of IRES RNAs to the head of the ribosomal 40S subunit. *Nucleic Acids Res.* 39, 5264–5275.

Ni, L., and Snyder, M. (2001). A genomic study of the bipolar bud site selection pattern in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 12, 2147–2170.

Nilsson, J., Sengupta, J., Frank, J., and Nissen, P. (2004). Regulation of eukaryotic translation by the RACK1 protein: a platform for signalling molecules on the ribosome. *EMBO Rep.* 5, 1137–1141.

Nishimura, T., Wada, T., Yamamoto, K.T., and Okada, K. (2005). The Arabidopsis STV1 Protein , Responsible for Translation Reinitiation , Is Required for Auxin-Mediated Gynoecium Patterning. *Plant Cell* 17, 2940–2953.

Nissen, P. (2000). The Structural Basis of Ribosome Activity in Peptide Bond Synthesis. *Science* 289, 920–930.

Noller, H.F., Hoffarth, V., and Zimniak, L. (1992). Unusual Resistance of Peptidyl Extraction Transferase to Protein Procedures. *Science* 256, 1416–1419.

Noller, H.F., Hoang, L., and Fredrick, K. (2005). The 30S ribosomal P site: a function of 16S rRNA. *FEBS Lett.* 579, 855–858.

Odintsova, T.I., Müller, E.-C., Ivanov, A. V, Egorov, T. a, Bienert, R., Vladimirov, S.N., Kostka, S., Otto, A., Wittmann-Liebold, B., and Karpova, G.G. (2003). Characterization and analysis of posttranslational modifications of the human large cytoplasmic ribosomal subunit proteins by mass spectrometry and Edman sequencing. *J. Protein Chem.* 22, 249–258.

Oh, W.J., Wu, C., Kim, S.J., Facchinetti, V., Julien, L.-A., Finlan, M., Roux, P.P., Su, B., and Jacinto, E. (2010). mTORC2 can associate with ribosomes to promote cotranslational phosphorylation and stability of nascent Akt polypeptide. *EMBO J.* 29, 3939–3951.

Ohtake, Y., and Wickner, R.B. (1995). Yeast virus propagation depends critically on free 60S ribosomal subunit concentration. *Mol. Cell. Biol.* 15, 2772–2781.

- Oleinikov, A. V, Jokhadze, G.G., and Traut, R.R. (1998). A single-headed dimer of *Escherichia coli* ribosomal protein L7/L12 supports protein synthesis. *Proc. Natl. Acad. Sci. U. S. A.* *95*, 4215–4218.
- Oliver, E.R., Saunders, T.L., Tarlé, S.A., and Glaser, T. (2004). Ribosomal protein L24 defect in belly spot and tail (Bst), a mouse Minute. *Development* *131*, 3907–3920.
- Oristian, D.S., Sloofman, L.G., Zhou, X., Wang, L., Farach-Carson, M.C., and Kirn-Safran, C.B. (2009). Ribosomal protein L29/HIP deficiency delays osteogenesis and increases fragility of adult bone in mice. *J. Orthop. Res.* *27*, 28–35.
- Ostroff, L.E., Fiala, J.C., Allwardt, B., and Harris, K.M. (2002). Polyribosomes redistribute from dendritic shafts into spines with enlarged synapses during LTP in developing rat hippocampal slices. *Neuron* *35*, 535–545.
- Panić, L., Tamarut, S., Sticker-Jantscheff, M., Barkić, M., Solter, D., Uzelac, M., Grabusić, K., and Volarević, S. (2006). Ribosomal protein S6 gene haploinsufficiency is associated with activation of a p53-dependent checkpoint during gastrulation. *Mol. Cell. Biol.* *26*, 8880–8891.
- Parenteau, J., Durand, M., Morin, G., Gagnon, J., Lucier, J.-F., Wellinger, R.J.J.J., Chabot, B., and Abou Elela, S. (2011). Introns within Ribosomal Protein Genes Regulate the Production and Function of Yeast Ribosomes. *Cell* *147*, 320–331.
- Park, H., Himmelbach, A., Browning, K.S., Hohn, T., and Ryabova, L.A. (2001). A Plant Viral “ Reinitiation ” Factor Interacts with the Host Translational Machinery. *Cell* *106*, 723–733.
- Pavlakakis, G.N., Jordan, B.R., Wurst, R.M., and Vournakis, J.N. (1979). Sequence and secondary structure of *Drosophila melanogaster* 5.8 S and 2S rRNAs and of the processing site between them. *Nucleic Acids Res.* *7*, 2213–2238.
- Pelletier, J., and Sonenberg, N. (1988). Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* *334*, 320–325.
- Peltz, S.W., Hammell, A.B., Cui, Y., Yasenchak, J., Puljanowski, L., and Dinman, J.D. (1999). Ribosomal Protein L3 Mutants Alter Translational Fidelity and Promote Rapid Loss of the Yeast Killer Virus. *Mol. Cell. Biol.* *19*, 384–391.
- Pinkstaff, J.K., Chappell, S.A., Mauro, V.P., Edelman, G.M., and Krushel, L.A. (2001). Internal initiation of translation of five dendritically localized neuronal mRNAs. *Proc. Natl. Acad. Sci. U. S. A.* *98*, 2770–2775.
- Pinon, V., Etchells, J.P., Rossignol, P., Collier, S.A., Arroyo, J.M., Martienssen, R.A., and Byrne, M.E. (2008). Three PIGGYBACK genes that specifically influence leaf patterning encode ribosomal proteins. *Development* *135*, 1315–1324.

Ramagopal, S. (1991a). Differences in accumulation and phosphorylation of proteins in vegetatively growing and aggregation-competent cells of *Dictyostelium discoideum*. *Biochem. Cell Biol.* 69, 751–753.

Ramagopal, S. (1991b). Covalent modifications of ribosomal proteins in growing and aggregation-competent *dictyostelium discoideum*: phosphorylation and methylation. *Biochem. Cell Biol.* 69, 263–268.

Ramagopal, S., and Ennis, H.L. (1981). Regulation of synthesis of cell-specific ribosomal proteins during differentiation of *Dictyostelium discoideum*. *Proc. Natl. Acad. Sci. U. S. A.* 78, 3083–3087.

Rhodin, M.H.J., Rakauskaitė, R., and Dinman, J.D. (2011). The central core region of yeast ribosomal protein L11 is important for subunit joining and translational fidelity. *Mol. Genet. Genomics* 285, 505–516.

Röhl, R., and Nierhaus, K.H. (1982). Assembly map of the large subunit (50S) of *Escherichia coli* ribosomes. *Proc. Natl. Acad. Sci. U. S. A.* 79, 729–733.

Ruvinsky, I., Sharon, N., Lerer, T., Cohen, H., Stolovich-Rain, M., Nir, T., Dor, Y., Zisman, P., and Meyuhas, O. (2005). Ribosomal protein S6 phosphorylation is a determinant of cell size and glucose homeostasis. *Genes Dev.* 19, 2199–2211.

Sahin, F., Qiu, W., Wilentz, R.E., Iacobuzio-Donahue, C.A., Grosmark, A., and Su, G.H. (2005). RPL38, FOSL1, and UPP2 are predominantly expressed in the pancreatic ductal epithelium. *Pancreas* 30, 158–167.

Schäfer, T., Maco, B., Petfalski, E., Tollervey, D., Böttcher, B., Aebi, U., and Hurt, E. (2006). Hrr25-dependent phosphorylation state regulates organization of the pre-40S subunit. *Nature* 441, 651–655.

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, 337–342.

Sonenberg, N., and Hinnebusch, A.G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* 136, 731–745.

Sossin, W.S., and DesGroseillers, L. (2006). Intracellular trafficking of RNA in neurons. *Traffic* 7, 1581–1589.

Spence, J., Gali, R.R., Dittmar, G., Sherman, F., Karin, M., and Finley, D. (2000). Cell Cycle – Regulated Modification of the Ribosome by a Variant Multiubiquitin Chain. *Cell* 102, 67–76.

Spriggs, K.A., Bushell, M., and Willis, A.E. (2010). Translational regulation of gene expression during conditions of cell stress. *Mol. Cell* 40, 228–237.

Steffen, K.K., McCormick, M.A., Pham, K.M., MacKay, V.L., Delaney, J.R., Murakami, C.J., Kaerberlein, M., and Kennedy, B.K. (2012). Ribosome Deficiency Protects Against ER Stress in *Saccharomyces cerevisiae*. *Genetics* 191, 107–118.

Subramanian, A.R. (1975). Copies of proteins L7 and L12 and heterogeneity of the large subunit of *Escherichia coli* ribosome. *J. Mol. Biol.* 95, 1–8.

Sugihara, Y., Honda, H., Iida, T., Morinaga, T., Hino, S., Okajima, T., Matsuda, T., and Nadano, D. (2010). Proteomic analysis of rodent ribosomes revealed heterogeneity including ribosomal proteins L10-like, L22-like 1, and L39-like. *J. Proteome Res.* 9, 1351–1366.

Szakonyi, D., Byrne, M.E., and Centre, J.I. (2011). Ribosomal protein L27a is required for growth and patterning in *Arabidopsis thaliana*. *Plant J.* 65, 269–281.

Tcherkezian, J., Brittis, P.A., Thomas, F., Roux, P.P., and Flanagan, J.G. (2010). Transmembrane receptor DCC associates with protein synthesis machinery and regulates translation. *Cell* 141, 632–644.

Uechi, T., Nakajima, Y., Nakao, A., Torihara, H., Chakraborty, A., Inoue, K., and Kenmochi, N. (2006). Ribosomal protein gene knockdown causes developmental defects in zebrafish. *PLoS One* 1, e37.

Velichutina, I. V, Rogers, M.J., McCutchan, T.F., and Liebman, S.W. (1998). Chimeric rRNAs containing the GTPase centers of the developmentally regulated ribosomal rRNAs of *Plasmodium falciparum* are functionally distinct. *RNA* 4, 594–602.

Vesper, O., Amitai, S., Belitsky, M., Byrgazov, K., Kaberdina, A.C.C., Engelberg-Kulka, H., and Moll, I. (2011). Selective Translation of Leaderless mRNAs by Specialized Ribosomes Generated by MazF in *Escherichia coli*. *Cell* 147, 147–157.

Vickers, C.A., Dickson, K.S., and Wyllie, D.J.A. (2005). Induction and maintenance of late-phase long-term potentiation in isolated dendrites of rat hippocampal CA1 pyramidal neurones. *J. Physiol.* 568, 803–813.

Vyas, K., Chaudhuri, S., Leaman, D.W., Komar, A. a, Musiyenko, A., Barik, S., and Mazumder, B. (2009). Genome-wide polysome profiling reveals an inflammation-responsive posttranscriptional operon in gamma interferon-activated monocytes. *Mol. Cell. Biol.* 29, 458–470.

Wan, F., Anderson, D.E., Barnitz, R.A., Snow, A., Bidere, N., Zheng, L., Hegde, V., Lam, L.T., Staudt, L.M., Levens, D., et al. (2007). Ribosomal protein S3: a KH domain subunit in NF-kappaB complexes that mediates selective gene regulation. *Cell* 131, 927–939.

Warner, J.R. (1999). The economics of ribosome biosynthesis in yeast. *Trends Biochem. Sci.* **24**, 437–440.

Warner, J.R., and McIntosh, K.B. (2009). How common are extraribosomal functions of ribosomal proteins? *Mol. Cell* **34**, 3–11.

Weijers, D., Franke-van Dijk, M., Vencken, R.J., Quint, a, Hooykaas, P., and Offringa, R. (2001). An Arabidopsis Minute-like phenotype caused by a semi-dominant mutation in a RIBOSOMAL PROTEIN S5 gene. *Development* **128**, 4289–4299.

Wellik, D.M. (2007). Hox patterning of the vertebrate axial skeleton. *Dev. Dyn.* **236**, 2454–2463.

Whittle, C.A., and Krochko, J.E. (2009). Transcript profiling provides evidence of functional divergence and expression networks among ribosomal protein gene paralogs in *Brassica napus*. *Plant Cell* **21**, 2203–2219.

Williams, M.E., and Sussex, I.M. (1995). Developmental regulation of ribosomal protein L16 genes in *Arabidopsis thaliana*. *Plant J.* **8**, 65–76.

Yoon, A., Peng, G., Brandenburger, Y., Brandenburg, Y., Zollo, O., Xu, W., Rego, E., and Ruggero, D. (2006). Impaired control of IRES-mediated translation in X-linked dyskeratosis congenita. *Science* **312**, 902–906.

Yu, Y., Ji, H., and Doudna, J.A. (2005). Mass spectrometric analysis of the human 40S ribosomal subunit: Native and HCV IRES-bound complexes. *Protein Sci.* **14**, 1438–1446.

Zaher, H.S., and Green, R. (2009). Fidelity at the molecular level: lessons from protein synthesis. *Cell* **136**, 746–762.

Zeidan, Q., Wang, Z., Maio, A. De, and Hart, G.W. (2010). O-GlcNAc Cycling Enzymes Associate with the Translational Machinery and Modify Core Ribosomal Proteins. *Mol. Biol. Cell* **21**, 1922–1936.

Zhang, Y., and Lu, H. (2009). Signaling to p53: ribosomal proteins find their way. *Cancer Cell* **16**, 369–377.

Zhou, C., Zang, D., Jin, Y., Wu, H., Liu, Z., Du, J., and Zhang, J. (2011). Mutation in ribosomal protein L21 underlies hereditary hypotrichosis simplex. *Hum. Mutat.* **32**, 710–714.

Zinzalla, V., Stracka, D., Oppliger, W., and Hall, M.N. (2011). Activation of mTORC2 by association with the ribosome. *Cell* **144**, 757–768.

Zoncu, R., Efeyan, A., and Sabatini, D.M. (2011). mTOR: from growth signal integration to cancer, diabetes and ageing. *Nat. Rev. Mol. Cell Biol.* *12*, 21–35.

Figures

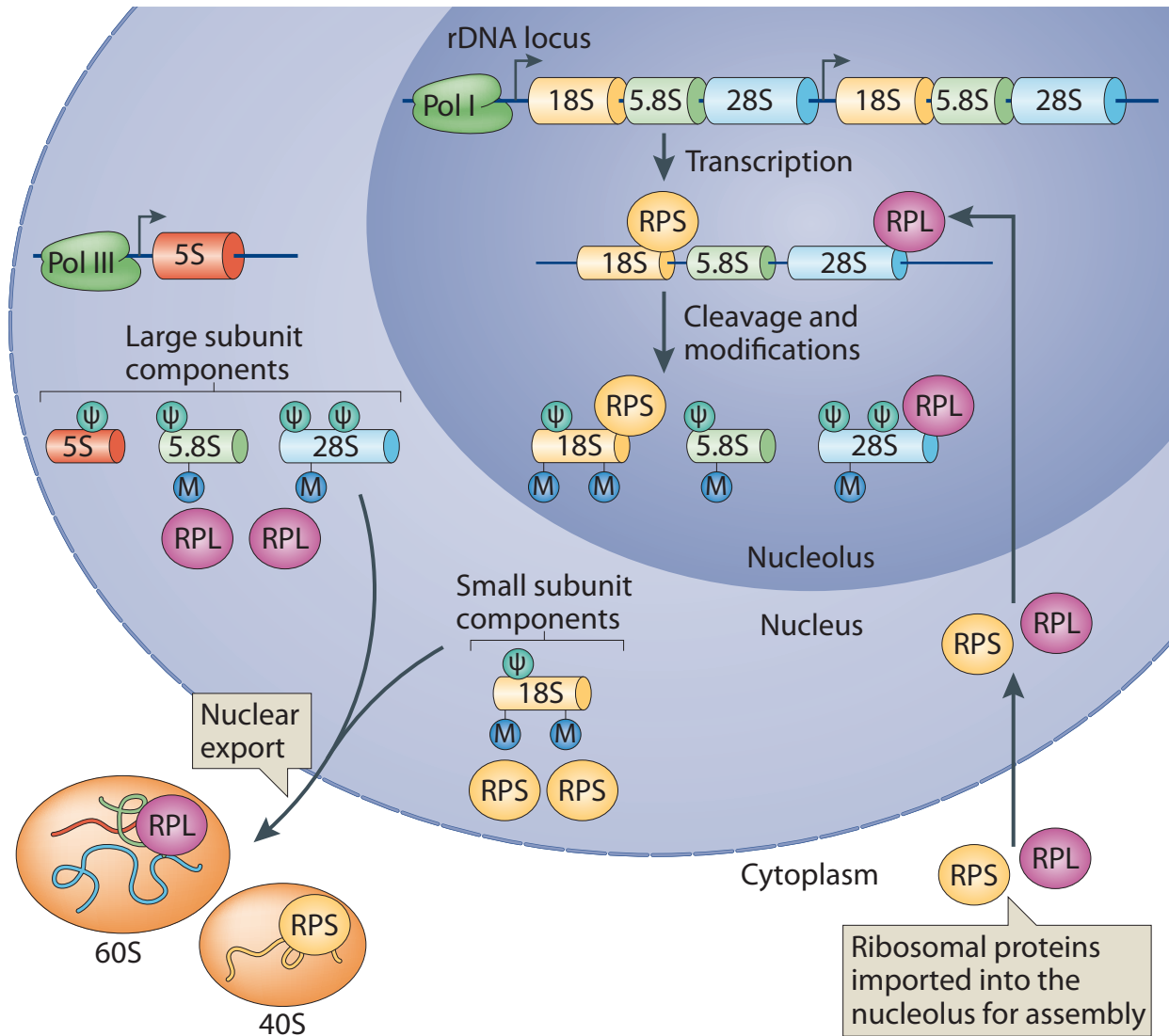


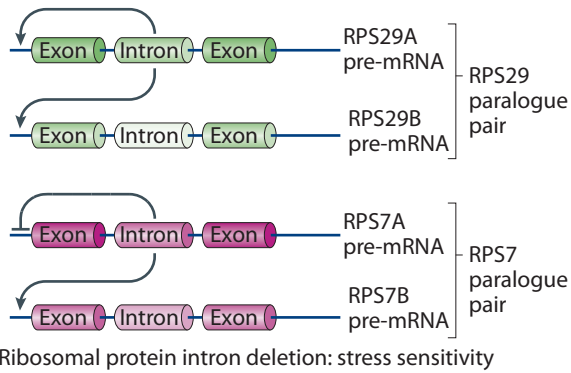
Figure 1.1: Ribosome biogenesis

Ribosome biogenesis is a highly orchestrated process involving all 3 RNA polymerases and over 150 non-ribosomal factors that take part in the maturation process (Kressler et al., 2010). Three of the four rRNA species, 18S, 5.8S and 28S, are transcribed as a single 47S transcript (pre-rRNA) by RNA Pol I in the nucleolus. The fourth rRNA, 5S, is transcribed by RNA Pol III in the nucleus. Strikingly, pre-rRNA is extensively modified at the level of pseudouridylation (ψ) as well as methylation (CH_3). This process is orchestrated by small nucleolar RNAs (snoRNAs) that guide site-specific modification of ~200 nucleotides in pre-rRNA. Although the role of rRNA modifications is not fully understood, they are important for ribosome function (Baxter-Roshek et al., 2007). For example, recent findings demonstrate that loss of rRNA pseudouridylation decreases translational fidelity, tRNA binding, and binding to structured RNA elements within mRNAs in an evolutionarily conserved manner from yeast to human cells (Jack et al., 2011).

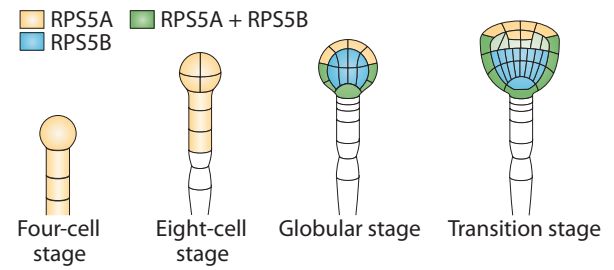
All ribosomal proteins are transcribed by RNA Pol II and translated in the cytoplasm. They are then imported into the nucleus and assembled on the pre-rRNA as it is being transcribed. Although less well studied, some ribosomal proteins have also been shown to be involved in rRNA processing, subunit assembly and export (Ferreira-Cerca et al., 2007; Martín-Marcos et al., 2007; Peltz et al., 1999; Rhodin et al., 2011; Schäfer et al., 2006). As the ribosomal proteins assemble, the pre-rRNA folds and undergoes a series of cleavages to generate the 18S, 5.8S and 28S rRNAs. The 18S rRNA along with 32 ribosomal proteins assemble into the small ribosomal subunit, 40S. The 28S, 5.8S, and 5S rRNA along with an additional 47 ribosomal proteins assemble

into the large 60S subunit. Both ribosomal subunits are exported into the cytoplasm and this process is also coordinated (Henras et al., 2008; Johnson et al., 2002).

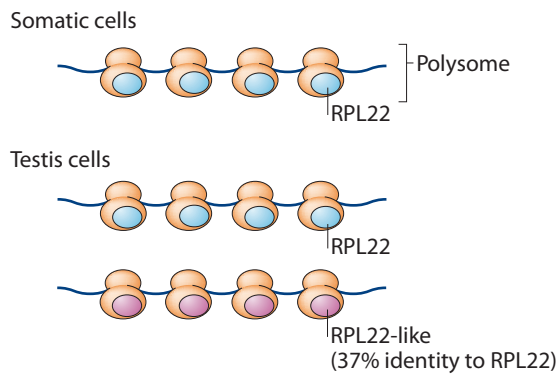
a *S. cerevisiae*



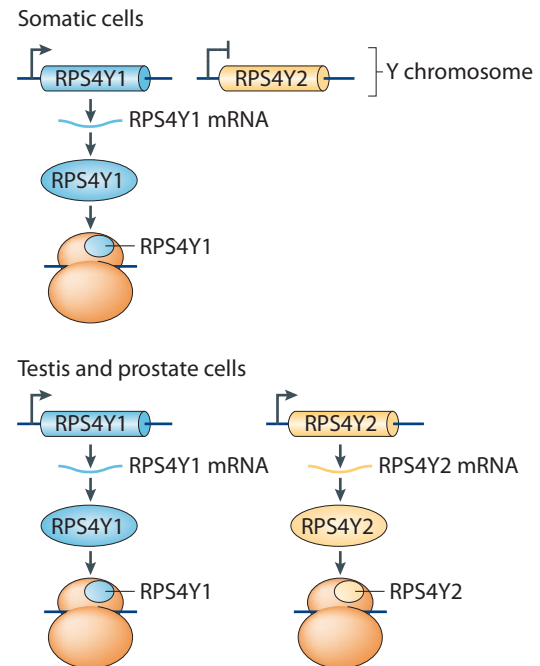
b *A. thaliana*



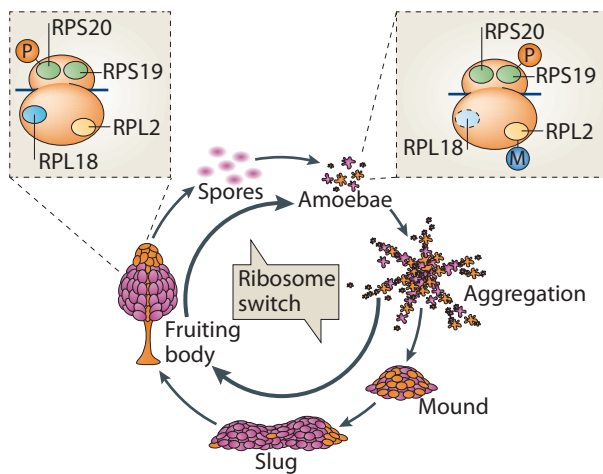
c *D. melanogaster*



d Human



e *D. discoideum*



f Mouse

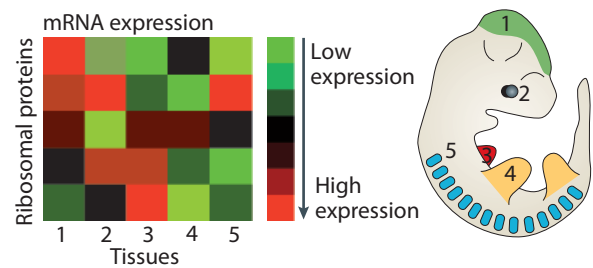


Figure 1.2 Heterogeneity in the ribosome at the level of ribosomal protein composition and post-translational modifications across different kingdoms of life.

(a) Introns of ribosomal proteins in *S. cerevisiae* regulate the expression of both themselves and their paralogues and this has important outcomes in a cell, e.g. increased fitness under stress. 70% of ribosomal protein paralogue pairs show differential expression. **(b)** In plants ribosomal protein paralogues have different functions and expression patterns. For example, in *Arabidopsis*, *RPS5A* is expressed in rapidly dividing cells early in embryonic development while *RPS5B* is expressed in cells undergoing differentiation. **(c)** In *Drosophila*, ribosomal protein paralogues show different expression patterns in the adult testes. For example, Rpl22 is expressed ubiquitously but Rpl22-like is specifically enhanced in the testes. Both proteins are incorporated into translationally active ribosomes (polysomes). **(d)** In humans there are fewer ribosomal protein paralogues, however notable examples exist RPS4Y1 is expressed ubiquitously but RPS4Y2 is restricted to the testis and prostate. **(e)** In the mouse, the mRNA expression pattern of ribosomal proteins varies dramatically between tissues. This can potentially translate into different ribosomal protein composition in ribosomes of distinct cell types. **(f)** In the lifecycle of the social amoeba *Dictyostelium discoideum*, ribosome switches are characterized by changes in ribosome composition at the level of ribosomal protein post-translational modifications. For example, Phosphorylation (P) on RPS19 and methylation (M) on RPL2 are lost as *Dictyostelium* aggregates from single cell amoeba to a multi-cellular fruiting body, while RPL20 gains

phosphorylation. RPL18 is exclusively found in ribosomes of developing cells and not in those of the amoeba stage.

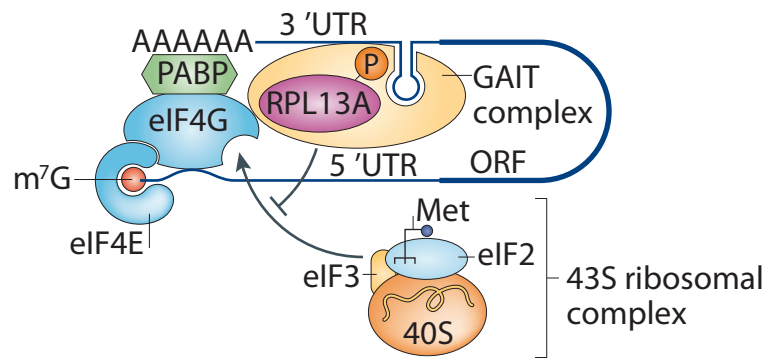


Figure 1.3 Extraribosomal functions of ribosomal proteins in translation control

Some ribosomal proteins have been shown to have additional functions off the ribosome (Warner and McIntosh, 2009). Most ribosomal proteins are RNA-binding proteins and some of these proteins can potentially bind to additional RNAs or protein complexes outside of the ribosome. For example, many ribosomal proteins regulate their own synthesis. In *Saccharomyces cerevisiae*, RPL30 inhibits splicing of its own transcript (Dabeva and Warner, 1993). In human cells, RPS13 also inhibits splicing of its own transcript by binding to its intron (Malygin et al., 2007). Ribosomal proteins can also function in other protein complexes (Wan et al., 2007).

The most interesting extraribosomal protein that functions in translation control is RPL13a. RPL13a is a ribosomal protein that regulates the translation of specific mRNAs as part of a non-ribosomal complex. Following interferon gamma (IFN γ) treatment, monocytes show a marked increase in transcription of ceruloplasmin (*Cp*) mRNA, an acute phase inflammatory protein. *Cp* protein synthesis is terminated after 16h even though *Cp* mRNA is still present (Mazumder and Fox, 1999). An interferon gamma activated inhibitor of translation (GAIT) complex was identified that binds to an RNA structure in the 3' UTR of *Cp* mRNA to inhibit translation. RPL13a is one of the four proteins that make up the GAIT complex. RPL13a is phosphorylated and released from the ribosome 24h after IFN γ stimulation (Mazumder et al., 2003). Phosphorylated RPL13a binds strongly to cap-bound eIF4G and blocks 43S ribosomal complex recruitment (Kapasi et al., 2007). It also has the same repressive effect on a number of other inflammatory response genes that have the same 3'UTR RNA structure, such as many chemokines and chemokine receptors (Vyas et al., 2009). Following this

paradigm, certain ribosomal proteins that are not essential for the catalytic activity of the ribosome may be co-opted for additional, specialized functions in translational control acting in complexes that function apart from the ribosome itself. PABP: poly-A binding protein. Met: Initiator tRNA

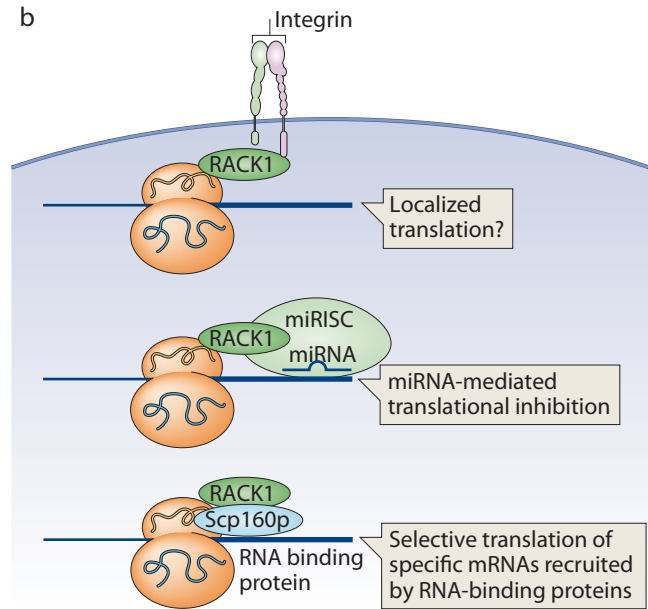
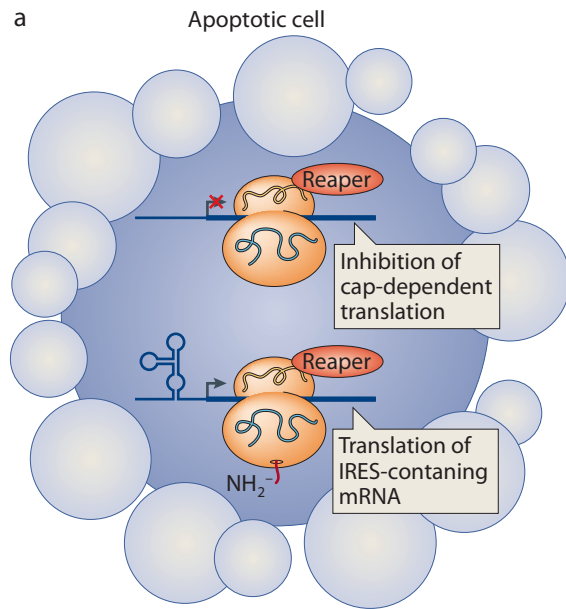


Figure 1.4 Ribosome associated factors affect ribosome function.

(a) Reaper protein in *Drosophila* associates with the 40S and disrupts the cap-scanning mechanism of translational initiation and promotes the translation of mRNAs containing IRES elements that may be required for cell death. **(b)** Ribosome-bound RACK1 can potentially localize ribosomes to the cell membrane by interacting with transmembrane receptors. RACK1 can recruit the miRNA-induced silencing complex (miRISC) to the ribosome to facilitate miRNA-mediated repression. RACK1 also binds to RNA-binding proteins such as Scp160p in yeast, which may recruit specific mRNAs to the ribosome.

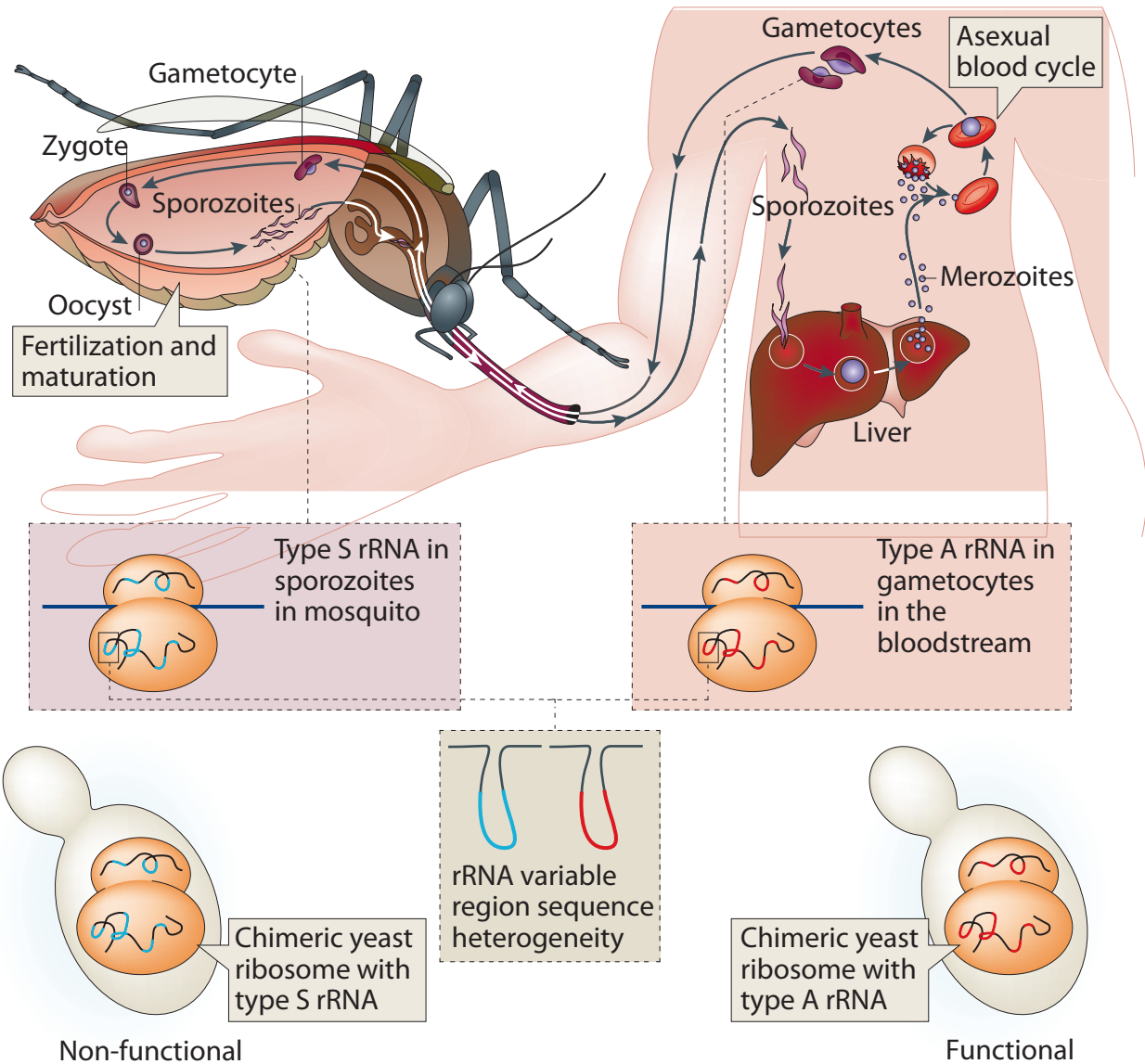


Figure 1.5 rRNA heterogeneity.

Plasmodium expresses different forms of rRNA when it is in sporozoite form (S) in the mosquito and when it is in the asexual form (A) in the bloodstream. The different forms of rRNA differ mainly in the variable regions with only a few nucleotides within the conserved areas. Chimeric yeast ribosomes where part of the 25S rRNA was replaced by the Type S *Plasmodium* rRNA are non-functional while replacement by the Type A *Plasmodium* rRNA does not produce any defects, suggesting the 2 types of rRNA are functionally distinct.

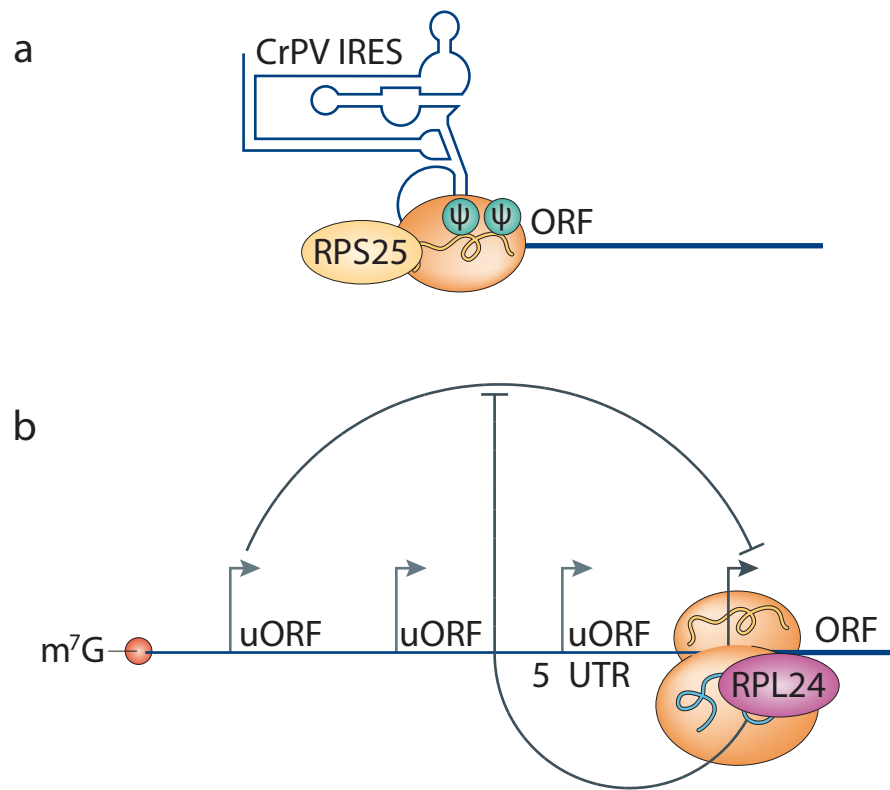


Figure 1.6 *Cis*-regulatory elements within mRNAs that interface with ribosomes or ribosomal proteins to confer transcript-specific regulation of gene expression.

(a) IRES elements: RPS25 is necessary for 40S recruitment to the Cricket Paralysis Virus (CrPV) intergenic region IRES for translational initiation by a cap-independent mechanism in yeast and mammals. rRNA modifications (such as pseudouridylation (ψ)) are important for ribosome binding to IRES elements. **(b) uORFs:** RPL24 promotes reinitiation of translation following a uORF, allowing the translation of the downstream main ORF.

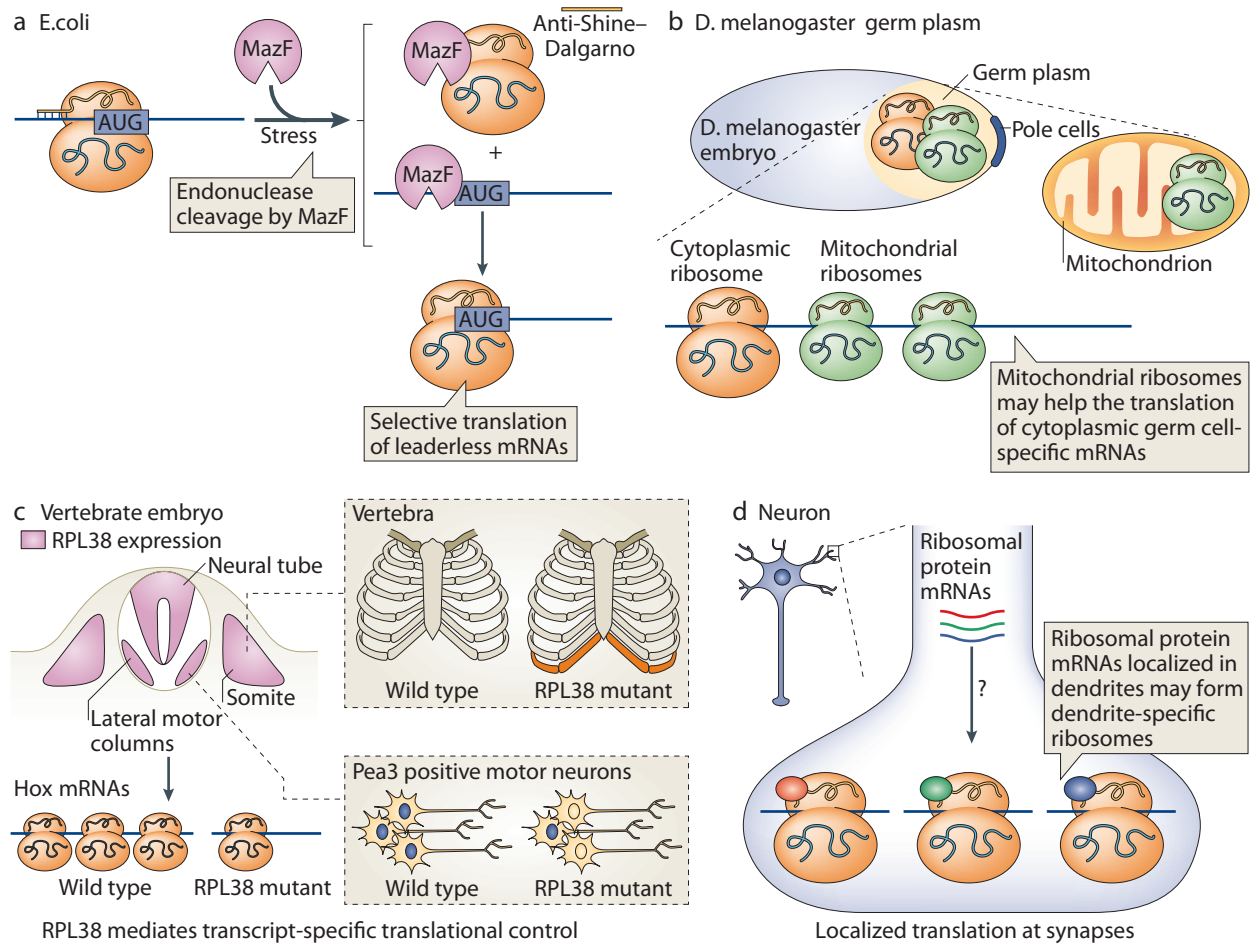


Figure 1.7 Ribosome specificity in cell and developmental biology.

(a) In *E. coli*, MazF cleaves the 3' end of the 16S rRNA and releases the anti Shine-Dalgarno (SD) sequence, generating specialized ribosomes that translate leaderless mRNAs involved in the stress response. **(b)** Mitochondrial ribosomes (green) in the germ plasm (yellow area) of *Drosophila* embryos are found outside of mitochondria and translate certain cytoplasmic germ cell specific mRNAs. **(c)** Expression of *Rpl38* (shaded) is highly enriched in the somites and neural tube of the developing mouse embryo and is responsible for translating specific *Hox* mRNAs necessary for axial skeletal patterning as well as the specification of *Pea3*⁺ motor neurons. **(d)** Localized translation takes place in dendrites, far away from the cell body. mRNAs for ribosomal proteins are found in dendrites, and can potentially make dendrite-specific ribosomes.

Chapter 2

RNA regulons in *Hox* 5' UTRs confer ribosome
specificity to gene regulation

Abstract

Emerging evidence suggests that the ribosome has a regulatory function in directing how the genome is translated in time and space. However, how this regulation is encoded in the messenger RNA sequence remains largely unknown. Here we uncover unique RNA regulons embedded in homeobox (*Hox*) 5' UTRs that confer ribosome-mediated control of gene expression. These structured RNA elements, resembling viral internal ribosome entry sites (IRESs), are found in subsets of *Hox* mRNAs. They facilitate ribosome recruitment and require the ribosomal protein RPL38 for their activity. Despite numerous layers of *Hox* gene regulation, these IRES elements are essential for converting *Hox* transcripts into proteins to pattern the mammalian body plan. This specialized mode of IRES-dependent translation is enabled by an additional regulatory element that we term the translational inhibitory element (TIE), which blocks cap-dependent translation of transcripts. Together, these data uncover a new paradigm for ribosome-mediated control of gene expression and organismal development.

Introduction

An important layer of post-transcriptional control of gene expression may be conferred through regulatory functions of the ribosome (Dinman, 2009; Komili et al., 2007; Kondrashov et al., 2011; Lee et al., 2013; Vesper et al., 2011; Xue and Barna, 2012). For example, RPL38, one of 80 ribosomal proteins (RPs) of the eukaryotic ribosome, helps establish the mammalian body plan by selectively

facilitating the translation of subsets of *Hox* mRNAs (Kondrashov et al., 2011), genes critically required for formation of the body plan (Alexander et al., 2009). However how ribosome-mediated regulation of gene expression is encoded within mRNA sequence remains an unanswered question.

All eukaryotic cellular mRNAs are capped and cap-dependent translation is considered a highly efficient and predominant means for translation of most transcripts in the mammalian genome (Livingstone et al., 2010; Sonenberg and Hinnebusch, 2009). In many viral mRNAs that are not capped, IRESs provide an alternative mechanism for ribosome recruitment to promote translation initiation (Plank and Kieft, 2012; Sonenberg and Hinnebusch, 2009). Interestingly, IRES elements have also been discovered in a handful of cellular mRNAs, including *c-myc*, *XIAP*, *Apaf-1* and *p53* (Bellodi et al., 2010a, 2010b; Ray et al., 2006; Riley et al., 2010; Stoneley et al., 2001; Ungureanu et al., 2006; Yoon et al., 2006). As these cellular mRNAs are capped, their IRES elements act as a 'fail-safe' mechanism to promote translation under stress conditions, such as apoptosis or hypoxia, when cap-dependent translation is downregulated (Holcik et al., 2000; Pyronnet et al., 2001). Cellular IRES elements have little sequence or structural homology (Xia and Holcik, 2009), and their relevance for translational control and gene expression in normal cell physiology, tissue patterning, or organismal development is unclear.

We report below the unexpected identification of IRES elements within *Hox* mRNAs, particularly those regulated by RPL38, which are critically required for accurate gene expression during normal development. We also identify the

TIE, an additional regulatory element in these mRNAs, which modulates the dependence of individual transcripts on canonical cap-dependent translation. This mechanism enables the IRES to become the predominant mode of translation initiation and confers greater gene regulatory potential by the ribosome.

Results

RPL38-regulated *Hox* mRNAs possess IRES elements

To address the nature of specific *cis*-regulatory elements within transcripts that confer greater regulation by the ribosome, we studied RPL38-mediated control of *Hox* mRNA translation. We showed previously that although RPL38 haploinsufficiency within the developing mouse embryo does not affect general cap-dependent translation, the translation of a subset of the eleven *HoxA* mRNAs (*Hoxa4*, *Hoxa5*, *Hoxa9*, *Hoxa11*) is perturbed (Figure 2.1a)(Kondrashov et al., 2011). We employed a bicistronic reporter assay in the murine mesenchymal stem cell line C3H10T1/2, which expresses *Hox* transcripts (Phinney et al., 2005), to delineate whether *Hox* 5' UTRs direct cap-independent translation (Figure 2.1a, b). In this reporter system, the first cistron (*Renilla* luciferase, RLuc) is translated by a cap-dependent mechanism, and the second cistron (Firefly luciferase, FLuc) is translated only if the preceding 5' untranslated region (UTR) element can recruit ribosomes by a cap-independent mechanism (Figure 2.1a); the latter is considered to reflect IRES-dependent translation initiation (Thompson, 2012).

These experiments revealed that many *Hox* 5' UTRs possess IRES activity that is as strong as, or even stronger than, the hepatitis C virus (HCV) IRES element, a bona fide viral IRES element (Figure 2.1b). To ensure that *Hox* 5' UTRs do not contain putative promoters or splice sites that may produce a monocistronic capped FLuc transcript, we performed additional controls. As expected for a single transcript containing both RLuc and Fluc: (i) the RNA levels of the two reporters are expressed in the same ratio among all bicistronic constructs; (ii) a short hairpin RNA (shRNA) against RLuc downregulated both RLuc and Fluc expression; (iii) these transcripts do not contain cryptic splice sites; and (iv) no ribosome read-through of the first reporter is evident (Figure 2.2). There is good correlation between RPL38 regulation and the presence of IRES elements; all RPL38-regulated mRNAs in the *HoxA* locus possess IRES elements. These include *Hoxa4*, which we show by 5' rapid amplification of cloned ends (RACE) possesses an IRES-containing 5' UTR of 1.1 kb, which is longer than previously annotated.

To see if cap-independent translation from the *Hox* 5' UTRs is controlled by RPL38 in C3H10T1/2 cells, a copy of *Rpl38* was disrupted by transcription activator-like effector nucleases (TALENs), yielding a 40% reduction in RPL38 protein expression (Figure 2.1c, 2.3a). In these cells, although there is no change in cap-dependent translation, there is a specific decrease in IRES-dependent translation of *Hox* target mRNAs that are regulated by RPL38 *in vivo* (Figure 2.1c, 2.3b). Consistent with the fact that RPL38 control of IRES-dependent translation is transcript-specific, *Hoxa3* IRES-dependent translation, which is not

regulated by RPL38 within the embryo (Kondrashov et al., 2011), and the HCV IRES, are similarly unaffected by RPL38 knockdown (Figure 2.1c). These results reveal the unexpected presence of IRES elements within *Hox* 5' UTRs and a critical function of RPL38 in regulating their IRES-dependent translation.

The *Hoxa9* minimal IRES is evolutionarily conserved

To further understand IRES-dependent regulation of *Hox* mRNAs, we made a series of deletions within the 1.2-kb *Hoxa9* 5' UTR. We localized the minimum fragment for RPL38-dependent IRES activity to nucleotides (nt) 944–1,266 (which we term the *Hoxa9* IRES element) (Figure 2.1d). Although this fragment does not fully recapitulate the IRES activity observed with full-length *Hoxa9* 5' UTR, it has strong IRES activity on its own and its removal abolishes all IRES activity (Figure 2.1d-e). This region of the *Hoxa9* 5' UTR shows remarkable evolutionary conservation in all vertebrates (from fish to mammals) that is greater than sequence conservation in the remaining 5' UTR or the 3' UTR sequence (Figure 2.1f, 2.4). IRES function is also evolutionarily conserved, as the zebrafish *hoxa9b* 5' UTR shows strong IRES activity in murine C3H10T1/2 cells (Figure 2.1g). These findings suggest that the *Hoxa9* IRES element may have arisen early during vertebrate evolution for post-transcriptional regulation of *Hox* expression.

***Hoxa9* is an RNA structure that recruits ribosomes**

Many viral IRESs possess structures such as conserved helices, asymmetric bulges, and pseudoknots that interact with initiation factors or the ribosome to promote translation initiation (Fraser and Doudna, 2007; Kieft et al., 2002; Spahn et al., 2004). We therefore examined whether the *Hoxa9* IRES element has structural properties *in vitro* that are functionally important.

Within the full-length *Hoxa9* 5' UTR, predictions of sub-regions of the 944–1,266 domain from selective 2'-hydroxyl acylation analysed by primer extension (SHAPE) (Wilkinson et al., 2006) were consistent with the minimal IRES domain forming a specific RNA structure *in vitro* (Figure 2.6). Automated modelling indicated that this domain formed a four-way junction with two long hairpin arms (P3 and P4; Figure 2.5a, 2.7) and a 'right-angle' asymmetric bulge (between P3b and P3c), but with little interaction with other domains of the 5' UTR (Figure 2.6). Analogous chemical mapping and functional studies on the 5' UTR from a different *Hox* mRNA, *Hoxa5*, also revealed a structured subdomain (nt 77–132) necessary for its full IRES activity (Figure 2.7a, d, e). Deletion of the subdomain substantially reduces IRES activity (Figure 2.7b). Structure conservation between different cellular IRES elements has so far been poorly defined (Xia and Holcik, 2009). Interestingly, the 'right angle' asymmetric bulges observed in *Hoxa5* and *Hoxa9* (Figure 2.7c) bear qualitative similarity to asymmetric bulges in, for example, the HCV IRES (Lukavsky et al., 2003), suggesting at least certain common structural features are present amongst cellular *Hox* IRESs and viral IRESs.

To gain further insight into structural and functional properties of a *Hox* IRES, we applied multidimensional chemical mapping to the *Hoxa9* minimal IRES subdomain (Figure 2.5a, 2.8). Comprehensive mutate-and-map analysis (Kladwang et al., 2011a) gave strong independent evidence for the same structural topology seen in the context of the full-length 5' UTR, revealing exposure of RNA segments upon mutation of their predicted pairing partners (Figure 2.8a-c). Importantly, compensatory mutation analysis (Tian et al., 2014) unambiguously confirmed specific base pairs in the P3 and P4 segments (green circles, Figure 2.5a, 2.9). After this *in vitro* structural characterization, we tested the functional significance of these hairpins for IRES-dependent translation through bicistronic assays. Deletion of P3, P4, or both hairpins in the context of the full-length *Hoxa9* 5' UTR resulted in a significant reduction in IRES activity. Most strikingly, the mutation of nucleotides in the base of P3a reduced IRES activity, whereas compensatory mutations gave quantitative rescue of IRES activity (Figure 2.5b). These results demonstrate the importance of the P3 and P4 hairpins in IRES function and connect the detailed *in vitro* *Hoxa9* IRES structural model to functional activity. Many viral IRES elements contain pseudoknots that may position the ribosome at the initiation codon (Berry et al., 2010). A potential pseudoknot was suggested by mutate-and-map analysis (Figure 2.5a), but this awaits further analysis. Together, these data confirm a structure defined at base-pair resolution for a *Hox* IRES element and indicate that a specific RNA structure is important for cellular *Hox* IRES activity.

To date, only a small class of viral IRES elements have been shown to interact with both the large and small ribosome subunits to form a translationally competent 80S ribosome (Jan and Sarnow, 2002; Stupina et al., 2011). However, a biotinylated full-length *Hoxa9* 5' UTR, as well as the minimal IRES element contained within nt 944–1,266, are able to pull down RPs from both the large and small subunits, including RPL38 (Figure 2.5c, d). The full-length 5' UTR also pulls down both 28S and 18S rRNAs (Figure 2.5e), suggesting that the 80S ribosome is able to form on the uncapped *Hoxa9* 5' UTR. Importantly, the full-length *Hoxa9* 5' UTR lacking the nt 944–1,266 IRES element (A9ΔIRES) or an antisense *Hoxa9* 5' UTR (A9 AS) is no longer able to efficiently recruit RPs or ribosomal RNA (Figure 2.5d, e), consistent with the fact that these fragments do not possess IRES activity (Figure 2.1d). These data support the idea that a structured element from a *Hox* IRES facilitates 80S ribosome recruitment to the mRNA, either directly or through intermediary RNA-binding proteins.

TIE inhibits cap-dependent translation

Because all eukaryotic mRNAs are capped and translated by a cap-dependent mechanism (Sonenberg and Hinnebusch, 2009), the function of an IRES-element in normal translational control remains unclear. Surprisingly, removing the IRES element from a *Hoxa9* 5' UTR monocistronic reporter construct dramatically decreases translational efficiency by 70-fold, despite the fact that the mRNA remains capped (Figure 2.5f). We therefore hypothesized that the *Hoxa9* 5' UTR may contain additional element(s) that intrinsically inhibit

cap-dependent translation. The *Hoxa9* 5' UTR contains 14 putative upstream open reading frames (uORFs), which are generally thought to inhibit the translation of the main open reading frame (Calvo et al., 2009). However, mutating all 14 ATGs to TTGs in the *Hoxa9* 5' UTR (A9ΔIRESΔuORF) (Figure 2.10a) does not affect cap-dependent translation of *Hoxa9*. We therefore placed various regions of the *Hoxa9* 5' UTR upstream of a monocistronic reporter to assay for effects on cap-dependent translation. Strikingly, nt 1–342 of the *Hoxa9* 5' UTR strongly inhibit cap-dependent translation (Figure 2.5g) despite this region not being particularly GC-rich (55%) and therefore unlikely to impede ribosome scanning. Other similarly sized pieces of the 5' UTR (nt 343–643), or the very long remaining portion of the 5' UTR (nt 343–1,266), do not inhibit cap-dependent translation and are probably translated by a combination of cap and IRES-mediated translation. To our knowledge, the 1–342 fragment of the *Hoxa9* 5' UTR is a unique example of a 5' UTR regulon that is sufficient to strongly inhibit cap-dependent translation under physiological conditions. We termed this 5' UTR mRNA element the translational inhibitory element (TIE). SHAPE analysis of the full-length *Hoxa9* 5' UTR lacks evidence for structural interactions between the TIE and the minimal IRES (Figure 2.6a), suggesting they function as independent modules. These findings reveal that the TIE enables a more specialized mode of translation initiation directed by the IRES.

5' UTR topology allows for specialized translation

We next asked whether the unique 5' UTR topology of both a TIE close to the 5' end of the mRNA and an IRES element is a more generalizable mechanism of translational regulation. Our findings reveal that all the minimal IRES domains of *HoxA* 5' UTRs lie within the 100–300 nt closest to the AUG, with the exception of *Hoxa5* (Figure 2.11). Deletions of these minimal IRES domains in monocistronic reporter constructs reveal that they are critically required for normal translation (Figure 2.12a-c). In addition, within these *HoxA* 5' UTRs, a TIE is also present near the mRNA cap (Figure 2.12a-c). Moreover, although the *Hoxa3*, *Hoxa4*, *Hoxa9*, and *Hoxa11* TIEs do not share obvious sequence similarity, they all strikingly block translation from the β -globin (*Hbb*) 5' UTR (Figure 2.12d), which initiates translation exclusively by a cap-dependent mechanism (Fletcher et al., 1990). *Hoxa5* is the only 5' UTR that does not appear to possess a TIE and a putative uORF also does not decrease cap-dependent translation (Figure 2.10b). However, *Hoxa5* has the shortest 5' UTR (237 nt) and almost the entire UTR is required for IRES activity, leaving little space for an additional TIE (Figure 2.11c). Interestingly, the *Hoxa5* transcript can be produced as a bicistronic mRNA with the neighbouring *Hoxa6* gene (Coulombe et al., 2010) and therefore may not need a TIE to inhibit cap-dependent translation in this context. These findings reveal a unique topology to *Hox* 5' UTRs that suppresses a more generic mode of cap-dependent translational control and facilitates specialized, ribosome-mediated control of protein expression.

Hoxa9* IRES is essential for translation *in vivo

To determine whether IRES-dependent translation is critically required for the conversion of the *Hoxa9* transcript into protein *in vivo*, we generated the first targeted mouse knockout of a cellular IRES by removing nt 944–1,193 of the *Hoxa9* 5' UTR, leaving the Kozak sequence and the rest of the 5' UTR intact (Figure 2.12a, 2.10c-d). A caveat to the previously generated HOXA9 targeted knockout is the retention of a Neomycin (Neo) cassette (Chen and Capecchi, 1997; Fromental-Ramain et al., 1996), which can have indirect effects on the expression of neighbouring *Hox* genes (Boulet and Capecchi, 1996). Indeed, our current genetic studies reveal that the presence of the Neo cassette within the *Hoxa9* targeting locus is sufficient to produce phenotypes previously attributed to *Hoxa9* loss-of-function (Figure 2.13). Therefore, in our *Hoxa9*^{ΔIRES/ΔIRES} mouse, we were careful to remove the Neo cassette from the targeted locus by Cre-mediated excision.

Hoxa9^{ΔIRES/ΔIRES} embryos do not show any major change in *Hoxa9* transcript boundaries or expression levels (Figure 2.12b, c). Strikingly, however, these mice show a fully penetrant homeotic transformation of the thirteenth thoracic vertebra (T13) to the first lumbar vertebra (L1), in which a full set of ribs normally present on T13 is missing (Figure 2.12d). Although we cannot at present ascertain whether the *Hoxa9* IRES deletion completely phenocopies a *Hoxa9* knockout mouse, this dramatic homeotic transformation is precisely at the anterior boundary of *Hoxa9* expression at the thoracic to lumbar transition (Fromental-Ramain et al., 1996). This transformation is also evident in ~50% of

Hoxa9^{ΔIRES/+} mice, although small rudimentary ribs can still form on T13.

Consistently, there is a striking decrease in HOXA9 protein levels in the neural tube and somites of *Hoxa9*^{ΔIRES/ΔIRES} embryos from the thoracic to sacral level (Figure 2.14a) revealing a block in HOXA9 protein production. Furthermore, in *Hoxa9*^{ΔIRES/ΔIRES} embryos there is a dramatic decrease in ribosome association of the *Hoxa9* transcript. The majority of the *Hoxa9* transcript in *Hoxa9*^{ΔIRES/ΔIRES} embryos, but not control transcripts such as β-actin, accumulates in pre-polysomal fractions, reflecting an accumulation of mRNA not bound by translationally active ribosomal subunits (Figure 2.14b, 2.15). These findings reveal that the *Hoxa9* IRES element is crucial in recruiting the ribosome to this transcript to pattern the mammalian body plan *in vivo*.

Discussion

Our studies reveal how expression information encoded within the mRNA template is sufficient to confer greater gene regulatory potential by the ribosome. In particular, our findings suggest that IRES elements are key regulatory elements within the mammalian genome that facilitate important, ribosome-mediated control of gene expression in development (Figure 2.14c). Cellular IRES elements have been controversial because of their weak IRES activity as compared to viral IRES elements under non-stress conditions. Their relevance to normal gene regulation is also poorly understood, as all eukaryotic mRNAs are capped and undergo cap-dependent translation (Komar and Hatzoglou, 2005; Thompson, 2012). Importantly, the requirement for *Hox* IRES elements is

explained by our identification of the TIE within these transcripts, which strongly inhibits general, cap-dependent translation (Figure 2.14c). We speculate that such IRES elements may be more widespread in the genome and suggest a two-state mechanism for gene regulation, controlled by transcription and an additional layer of *cis*-acting RNA regulons within 5' UTRs that direct protein production. Both IRES and TIE elements may be differentially regulated, providing a versatile toolbox for controlling the ultimate expression of transcripts. Although *Hox* IRES elements are highly conserved, and may already be present in *Drosophila* orthologues (Oh et al., 1992), the TIE lacks conservation and is absent even in vertebrate species such as fish and amphibians (Figure 2.1f). These species often lay their eggs in harsh environments (Johnston, 2006) and may have originally used IRES elements as a fail-safe mechanism to ensure proper expression of key transcripts under stress conditions. Therefore, the acquisition of a TIE later in evolution may have enabled an additional, independent level of regulatory control to IRES-dependent translation and gene expression during normal development. A mechanism that regulates whether a pre-existing transcript is converted into a protein offers a fast and dynamic means to provide greater variation and control to gene expression.

It will be interesting to determine if additional RPs may promote specialized translation through control of unique subsets of IRES-containing mRNAs, either directly or through RNA-binding proteins. For example, RPS25 is required for IRES-dependent translation of certain viral IRES mRNAs (Landry et al., 2009). Moreover, rRNA modifications both at the level of pseudouridylation,

and RPL13a-dependent methylation also appear to regulate the translation of certain cellular IRES containing mRNAs (Bellodi et al., 2010a, 2010b; Chaudhuri et al., 2007; Yoon et al., 2006). We therefore speculate that similar to the complex and highly regulated system of transcriptional control, in which specific DNA sequences and histone marks regulate gene expression, *cis*-acting RNA regulons, in conjunction with more specialized ribosome activity, provide new found regulatory control to gene expression critical for mammalian development.

Acknowledgements

We would like to thank members of the Barna laboratory and D. Ruggero for discussion and critical reading of the manuscript. We thank E. Sarinay Cenik for helpful advice with RNA pull down experiments; Y. Rim, A. Sapiro and A. Mateo for technical assistance. This work was supported by the Agency of Science, Technology and Research of Singapore (S.X.), Stanford Graduate Fellowship (S.T.), Human Frontier Science Program Fellowship (K.F.), NIH R01 GM102519 (R.D.), NIH Director's New Innovator Award, 7DP2OD00850902 (M.B.), Alfred P. Sloan Research Fellowship (M.B.) and Pew Scholars Award (M.B.).

Author Contributions

M.B. conceived and supervised the project; S.X. and M.B. designed experiments; S.X. performed all experiments with help from K.F. for 5' RACE and sucrose gradient fractionation; S.T. and R.D. designed the RNA structure experiments; S.T. and W.K.. performed the RNA structure experiments and analysed results; S.X. and M.B. wrote the manuscript with contributions from all authors.

Author Information

Hoxa4 5' UTR sequence has been deposited in GenBank under accession number [KM596709](#). All chemical mapping datasets have been deposited at the RNA Mapping Database (<http://rmdb.stanford.edu>) under the following accession codes: (1) Full-length: HOXA5_STD_0000, HOXA9_STD_0000; (2) *Hoxa9* TIE: HOXA9A_STD_0001; (3) *Hoxa9* IRES: HOXA9D_STD_0001, HOXA9D_STD_0002, HOXA9D_1M7_0001, and HOXA9D_RSQ_0001.

METHODS

Plasmids

Hox 5' UTRs were amplified from mouse E10.5–12.5 cDNA and cloned into the bicistronic pRF vector (Yoon et al., 2006) using EcoRI and NcoI sites. For monocistronic assays, *Hox* 5' UTRs are cloned into pGL3 control vector (Promega) using HindIII and NcoI sites. 5' UTRs of *Hoxa3* and *Hoxa10* were made by gene synthesis (IDT). pRF-HCV was a kind gift from D. Ruggero (UCSF). Zebrafish *hoxa9b* 5' UTR was amplified from zebrafish 24 hpf cDNA. Sequences were based off the Ensembl database (Flicek et al., 2014). See Supplementary Table 1 for a list of PCR primers used.

To ensure no ribosome read-through of the first cistron in bicistronic assays, a hairpin with a linker (GAATTCAGATCTGGTACCGAGCTCCCCGGGCTGCAGGATATCCTGCAGCCCGGGGACCTCGGTACCAGATCTAGGATCCGGATCCGGCGACACTCCACCATGAATCACTCCCCTGTGAGGAACTACTGTCTTCACGCAGAAAGCGTCTAGCATGG) was synthesized and ligated into the pRF vector at the EcoRI and NcoI sites. *Hox* 5' UTRs and HCV IRES were cloned into this new vector pRF-hpC in the NcoI site. Clones were sequenced to ensure directionality.

5' RACE

RNA was purified from dissected E11.5 neural tubes and somites using TRIzol (Invitrogen). The reverse transcription template was produced using ExactSTART Eukaryotic mRNA 5'-& 3'-RACE Kit (Epicentre) and reverse transcribed using Superscript III (Invitrogen) with a primer in the coding region of

Hoxa4. The 5' UTR fragment was amplified using Phusion High-Fidelity DNA polymerase (Thermo Scientific) with a gene specific reverse primer and a forward primer at the 5' linker. PCR products were adenylated, gel-extracted and sub-cloned into TOPO TA cloning vector (Invitrogen) for sequencing.

TALEN

2 sets of TALEN pairs were designed against the second and third exons of *Rpl38*. TALENs were generated towards the following sequences: F1 5'-GCGGACGCCAGAGCCGA, R1 5'-CCCTCGTCCCGCACTCA, F2 5'-CATCTTCTGTCCCTC, R2 5'-TGATCTCCTCAATTTTC. Intermediate arrays were produced for each TALEN-pair and assembled into pC-GoldyTALEN (Carlson et al., 2012). The 2 sets of TALEN pairs were transfected together into C3H10T1/2 cells. Cells were trypsinized into single cell suspension and plated for clonal expansion 2 days after transfection. Positive clones were identified by PCR followed by restriction digest. *Rpl38* disruption was confirmed by Sanger sequencing and western blotting.

Luciferase assay

C3H10T1/2 cells were transfected in 24-well plates with various pRF constructs using Lipofectamine 2000 (Invitrogen) as per manufacturer instructions. Cells were collected 24 h post-transfection and assayed using Dual luciferase kit (Promega). IRES activity is expressed as a ratio between Fluc and Rluc. To verify the integrity of the bicistronic construct, an shRNA plasmid against Rluc or scrambled plasmid (Bellodi et al., 2010a) was cotransfected with the pRF constructs. Cells were harvested 72 h post-transfection for luciferase

assays. For monocistronic luciferase assays, various pGL3 constructs were transfected as above into 12 well plates. Cells were collected 24 h later. 50% of the cells were used for qPCR analysis with Fluc primers to quantify total transcript abundance, the rest for luciferase assays. Fluc activity was normalized to total luciferase mRNA expression levels.

RNA pull down

Template DNA was made by PCR from pRF-HoxA9 using a 5' primer containing the T7 promoter and a 3' primer containing a linker sequence T2 (Supplementary Table 1). Biotinylated RNA was *in vitro* transcribed with Megascript T7 kit (Ambion) using a final concentration of 6.8 mM each of ATP, CTP and GTP, 6.5 mM UTP and 0.3 mM Biotin-16-UTP (Roche). Synthesized RNAs were treated with TURBO DNase (Ambion) and purified by phenol/chloroform extraction. Unincorporated nucleotides were removed by running the RNA through a G50 column (GE Healthcare). The RNA was analysed by electrophoresis for quality and Nanodrop for quantity.

A total of four 15-cm dishes of subconfluent C3H10T1/2 cells were collected for analysis using trypsinization. Cell pellets were washed and lysed on ice with 1.5 ml lysis buffer (10 mM Tris pH 7.5, 10 mM MgCl₂, 150 mM NaCl, 0.1% TritonX, 0.25% NP40, 10 mM DTT, RNaseOUT, Halt Protease inhibitor) for 30 min. Lysate was centrifuged at 14,000 rpm for 10 min, 4 °C. Supernatant was transferred to new tube and cleared with 400 µl avidin-agarose beads (Thermo) for 50 min, 4 °C. Cleared lysate was transferred to new tube and quantified by spectrometry.

Folding buffer was added to 100 pmol of biotinylated RNA of each species to a final concentration of 100 mM Tris pH8, 100 mM NaCl, 10 mM MgCl₂. RNA was heated to 95 °C for 1.5 min, ice for 2 min and 37 °C for 20 min. Then 100 pmol of folded RNA and 1 mg cleared lysate were incubated together with 100 µg ml⁻¹ heparin and 100 µg ml⁻¹ yeast tRNA in 2 ml lysis buffer for 1 h at room temperature.

Streptavidin M280 Dynabeads (Invitrogen) were washed according to manufacturer's instructions. 100 µl beads were incubated with each RNA per lysate sample for 1.5 h at 4 °C. Beads were then washed with lysis buffer with increasing concentrations of salt up to 600 mM NaCl. One-tenth of the bead sample was kept for RNA and the rest was used for protein analysis. RNA was eluted with 50 µl RNA denaturing buffer (95% formamide, 10 mM EDTA) at 65 °C for 5 min. Eluted RNA was purified by phenol/chloroform extraction. RNA was converted to cDNA and analysed by qPCR using primers *HoxA9* 5' UTR F and *HoxA9* 342 R (Supplementary Table 1). Proteins were eluted with 40 µl protein elution buffer (5 mM biotin, 3× SDS sample buffer) at 95 °C for 5 min. Proteins were analysed by western blotting.

Western blotting

Western blots were performed with antibodies against RPS6 (Cell Signaling #2217) and RPL10a (Santa Cruz sc-100827). Polyclonal RPL38 antibodies were generated against C-EKAEKCLKQSLPPGLAVKDLK.

RNA structural probing and data analysis

Double-stranded DNA templates were prepared by PCR assembly of DNA oligomers with maximum length of 60 nt ordered from IDT (Integrated DNA Technologies). DNA templates for mutate-and-map (M^2) and mutation/rescue were designed as previously described using NA_Thermo (https://github.com/DasLab/NA_thermo) and HiTRACE (Kladwang et al., 2011a; Tian et al., 2014). DNA template and RNA transcript preparation and quality checks were carried out as previously described (Kladwang et al., 2011a; Tian et al., 2014).

All chemical mapping were carried out in 96-well format as described previously (Kladwang et al., 2011a, 2011b; Tian et al., 2014). Briefly, before chemical modification, RNA was heated up to 90 °C and cooled on bench top to room temperature to remove secondary structure heterogeneity, and folded for 20 min at 37 °C in 10 mM $MgCl_2$, 50 mM Na-HEPES, pH 8.0. RNA was modified by adding 1/4 volume freshly made 5 mg ml^{-1} 1-methyl-7-nitroisatoic anhydride (1M7) (Mortimer and Weeks, 2007) in anhydrous DMSO or neat anhydrous DMSO as control. Modification reactions were incubated at room temperature and then quenched with 0.5 M Na-MES, pH 6.0. Quenches also included poly(dT) magnetic beads (Ambion) and FAM-labelled Tail2-A20 primer for reverse transcription. Samples were separated and purified using magnetic stands, washed with 70% ethanol twice, and air-dried. Beads were resuspended in ddH₂O and reverse transcription mix, then incubated at 55 °C for 30 min. RNA was degraded by addition of 1 volume of 0.4M NaOH and incubation at 90 °C for

3 minutes, then cooled and neutralized with an additional volume of acid quench (prepared as 2 volumes of 5 M NaCl, 2 volumes of 2 M HCl, and 3 M NaOAc pH 5.2). Fluorescent labelled cDNA was recovered by magnetic bead separation, rinsed twice with 70% ethanol and air-dried. The beads were resuspended in Hi-Di formamide containing ROX-350 ladder (Applied Biosystems), then loaded on capillary electrophoresis sequencer (ABI3100).

For full length *Hox* transcripts, an Illumina MiSeq-based ligSHAPE (ligation-based SHAPE) on long RNAs was carried out, using a protocol analogous to Mod-seq (Talkish et al., 2014). RNA was folded similarly and probed by SHAPE as above, then recovered by ethanol precipitation. To obtain 3' ends for ligating primer-binding sites at unstructured regions, modified RNA was folded again and fragmented by 5 mM TbCl_3 for 5 min (or no Tb^{3+} as control) (Harris and Walter, 2001), quenched by 0.5 M Na-EDTA pH 8.0, then purified by RNA cleanup and concentration columns (Zymo Research). Fragmented RNA was prepared through 3' phosphate removal by treatment with T4 polynucleotide kinase (New England Biolabs), T4 RNA-ligase-based ligation of a 5'-adenylated universal miRNA cloning linker (New England Biolabs) to the RNA 3' end, reverse transcription to cDNA using sequencing primers, and circLigase-based Illumina adapter ligation to the cDNA 3' end as described (Cheng et al., 2014).

The HiTRACE 2.0 package was used to analyse CE data, available as MATLAB toolbox at (<https://github.com/hitrace>) (Yoon et al., 2011) and web server at (<http://hitrace.org>) (Kim et al., 2013)). Electrophoretic traces were aligned and baseline subtracted using linear and nonlinear alignment routines as

previously described (Kim et al., 2009). Band intensities were fitted to Gaussian peaks, then background subtracted, signal attenuation corrected and normalized. MiSeq results were analysed using the MAPseeker software package (Cheng et al., 2014; Seetin et al., 2014). The SHAPE reactivity profile was generated by projection of pair-ended reads. Briefly, a correction factor of 0.25 was used for full-length read ligation bias correction, accounting for the empirically observed low circLigase ligation efficiency for full-length cDNAs ending in CC (Kladwang et al., 2014). Reads shorter than 8 nt were discarded due to ambiguity in their positional alignment in the full-length RNA. Modification fractions were determined by the number of cDNAs in which reverse transcription stopped at a given nucleotide divided by the number of cDNAs in which reverse transcription stopped or proceeded beyond that nucleotide (Kladwang et al., 2014). Traces were background subtracted (SHAPE minus no-SHAPE data, and SHAPE-Tb minus Tb-alone data) and normalized through a boxplot-based heuristic (Deigan et al., 2009) for *Hoxa5*, or using a scanning window of 350 nt for *Hoxa9*. Errors were estimated as the standard deviation between SHAPE/no-SHAPE and SHAPE-Tb/Tb-alone measurements.

Data-guided secondary structure models were obtained using the Fold executable of the RNAstructure package (Reuter and Mathews, 2010). Pseudo-energy parameters of RNAstructure version 5.4 were used by default. To obtain 2D-data-guided secondary structure models, Z score matrices for mutate-and-map data sets were calculated and used as base-pair-wise pseudo-energies with a slope and intercept of $1.0 \text{ kcal mol}^{-1}$ and 0 kcal mol^{-1} as previously described

(Kladwang et al., 2011a). The ShapeKnots executable of RNAstructure version 5.6 (Hajdin et al., 2013) enables modelling of pseudoknots but with different parameters for using SHAPE data in the modelling; this newer method was used for pseudoknot prediction of *Hoxa9* 957–1,132 when using 1D and 2D SHAPE data combined. Extension of ShapeKnots to accept 2D data are being incorporated into the next version of RNAstructure. Helix-wise confidence values were calculated via bootstrapping as described previously (Kladwang et al., 2011a; Tian et al., 2014). All predictions were modelled in context with full-length sequences. Secondary structures were visualized by the VARNA software (Darty et al., 2009).

Mice

A 2.8-kb 5' arm and a 4.9-kb 3' arm of *Hoxa9* were cloned into PGKneolox2DTA.2 (P. Soriano, Addgene plasmid 13449). This results in a replacement of nt 945–1,194 of *Hoxa9* 5' UTR with a *LoxP*-flanked Neo cassette in the same transcriptional direction as *Hoxa9* (Figure 2.10c). The plasmid was linearized with *Xho*I and electroporated into JM8NA3.N1 ESCs for homologous recombination by the UCSF ES cell core. The correctly targeted ESC clones were identified by long-range PCR and confirmed by Southern blotting. The clones were injected into C57BL/6 blastocysts by the Gladstone Institute Transgenic Core Facility. Resulting chimaeras were crossed with C57BL/6 mice to generate F1 (*Hoxa9*^{Neo/+}). *Hoxa9*^{Neo/+} were crossed with CMV-Cre mice (Jackson Laboratory) to precisely remove the *LoxP*-flanked Neo cassette within the 5' UTR to generate *Hoxa9*^{ΔIRES/+} mice. All mice used for experiments were

between 2 and 6 months old. All animal studies were performed in accordance to with Stanford University Animal Care and Use guidelines.

Sucrose gradient fractionation

E11.5 neural tubes and somites were dissected and trypsinized as previously described¹. The neural tube and somites of each embryo were lysed on ice in 200 μ l of lysis buffer (20 mM Tris pH 7.5, 200 mM NaCl, 15 mM $MgCl_2$, 1 mM DTT, 8% glycerol, 1% Triton X-100, 100 μ g ml^{-1} cycloheximide, 100 μ g ml^{-1} heparin, 200 U ml^{-1} SUPERase In (Ambion)). After lysis, nuclei and membrane debris was removed by centrifugation (1,300g, 5 min at 4 °C, and then 14,000 rpm, 5 min, at 4 °C). The supernatant was layered onto a linear sucrose gradient (10–40% sucrose(w/v), 20 mM Tris, pH 7.5, 100 mM NaCl, 15 mM $MgCl_2$, 100 μ g ml^{-1} cycloheximide) and centrifuge in a SW41Ti rotor (Beckman) for 2.5 h at 40,000 rpm at 4 °C. Fractions were collected by Density Gradient Fraction System (Brandel).

qPCR analysis

For cells, RNA was isolated using NucleoSpin RNA kit (Clontech) and DNase-treated on column. For embryos, E11.5 neural tubes and somites were dissected and trypsinized as previously described¹. RNA was isolated using TRIzol (Invitrogen) and treated with TURBO DNA-free kit (Ambion). For sucrose gradient fractions, RNA was purified by acid phenol/chloroform followed by isopropanol precipitation. 0.5 μ g of RNA was converted to cDNA using iScript supermix (BioRad). cDNA was diluted twofold and 1 μ l used to run a SYBR green detection qPCR assay (SsoAdvanced SYBR Green supermix and CFX384,

BioRad). Data was analysed and converted to relative RNA quantity using CFX manager (BioRad). For sucrose gradient fractions, amount of RNA from individual fractions was expressed as a fraction of the total RNA collected from all fractions. Primers were used at 250 nM per reaction. β -actin and *Hoxa9* qPCR primers were previously described (Kondrashov et al., 2011). Fluc F 5'-CATCACGTACGCGGAATACTT, Fluc R 5'-AAGAGATACGCCCTGGTTC. 18S F 5'-ACATCCAAGGAAGGCAGCAG, 18S R 5'-CATTCCAATTACAGGGCCTC. 28S F 5'-GGGGAGAGGGTGTAATCTC, 28S R 5'-TCCTTATCCCGAAGTTACGG.

Whole mount *in situ* hybridization

Whole mount *in situ* hybridization was carried out using digoxigenin-labelled antisense RNA probes as previously described (Kondrashov et al., 2011). *Hoxa9* probe was amplified from mouse embryo cDNA. Primers sequences are described in Table 2.1.

Skeletal staining

Alcian blue and alizarin red staining of cartilage and bone were performed on neonates as previously described (Kondrashov et al., 2011).

Immunofluorescence

E11.5 embryos were dissected, fixed and cryosectioned as previously described (Kondrashov et al., 2011). WT and *Hoxa9* ^{Δ IRES/ Δ IRES} littermates sections were placed onto the same slide for direct comparison and identical processing. Sections were equilibrated to room temperature and rehydrated in blocking buffer (1% heat-inactivated goat serum, 0.1% Triton X-100 in PBS) for

1 h at room temperature. Sections were incubated with anti-HOXA9 antibody (Abcam ab140631) at 1:1,000 in blocking buffer overnight at 4 °C, washed three times, 30 min each and incubated with Alexa 488 anti-rabbit secondary antibodies (Invitrogen) for 1 h at room temperature. Slides were washed with blocking buffer three times, 30 min each and incubated with 1 $\mu\text{g ml}^{-1}$ DAPI (Cell Signaling) for 15 min at room temperature. They were washed again in blocking buffer for 20 min and mounted with VectaShield Mounting Media (Vector Laboratories).

Statistical analysis

No normalization, blinding or exclusion was used in any experiments. No statistical method was used to predetermine sample size. In all cases, multiple independent experiments were performed on different days to verify the reproducibility of experimental findings. For mice experiments, embryos from multiple litters were used to avoid litter-specific bias.

Each variable was analysed using the two-tailed unpaired Student's *t* test or χ^2 test. For all analyses, a *P* value of less than 0.05 was considered significant. Results are given as means $\pm$ s.d.

References

- Alexander, T., Nolte, C., and Krumlauf, R. (2009). Hox genes and segmentation of the hindbrain and axial skeleton. *Annu. Rev. Cell Dev. Biol.* 25, 431–456.
- Bellodi, C., Kopmar, N., and Ruggero, D. (2010a). Deregulation of oncogene-induced senescence and p53 translational control in X-linked dyskeratosis congenita. *EMBO J.* 29, 1865–1876.
- Bellodi, C., Krasnykh, O., Haynes, N., Theodoropoulou, M., Peng, G., Montanaro, L., and Ruggero, D. (2010b). Loss of function of the tumor suppressor DKC1 perturbs p27 translation control and contributes to pituitary tumorigenesis. *Cancer Res.* 70, 6026–6035.
- Berry, K.E., Waghray, S., and Doudna, J.A. (2010). The HCV IRES pseudoknot positions the initiation codon on the 40S ribosomal subunit. *RNA* 16, 1559–1569.
- Boulet, A.M., and Capecchi, M.R. (1996). Targeted disruption of *hoxc-4* causes esophageal defects and vertebral transformations. *Dev. Biol.* 177, 232–249.
- Calvo, S.E., Pagliarini, D.J., and Mootha, V.K. (2009). Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc. Natl. Acad. Sci. U. S. A.* 106, 7507–7512.
- Carlson, D., Tan, W., Lillico, S.G., Stverakova, D., Proudfoot, C., Christian, M., Voytas, D.F., Long, C.R., Whitelaw, B.A., and Fahrenkrug, S.C. (2012). Efficient TALEN-mediated gene knockout in livestock. *Proc. Natl. Acad. Sci. U. S. A.* 109, 17382–17387.
- Chaudhuri, S., Vyas, K., Kapasi, P., Komar, A.A., Dinman, J.D., Barik, S., and Mazumder, B. (2007). Human ribosomal protein L13a is dispensable for canonical ribosome function but indispensable for efficient rRNA methylation. *RNA* 13, 2224–2237.
- Chen, F., and Capecchi, M.R. (1997). Targeted mutations in *Hoxa-9* and *Hoxb-9* reveal synergistic interactions. *Dev. Biol.* 181, 186–196.
- Cheng, C., Chou, F., Kladwang, W., Tian, S., Cordero, P., and Das, R. (2014). MOHCA-seq : RNA 3D models from single multiplexed proximity-mapping experiments. *bioRxiv*.
- Coulombe, Y., Lemieux, M., Moreau, J., Aubin, J., Joksimovic, M., Bérubé-Simard, F.-A., Tabariès, S., Boucherat, O., Guillou, F., Laroche, C., et al. (2010). Multiple promoters and alternative splicing: *Hoxa5* transcriptional complexity in the mouse embryo. *PLoS One* 5, e10600.

- Darty, K., Denise, A., and Ponty, Y. (2009). VARNAs: interactive drawing and editing of the RNA secondary structure. *Bioinformatics* 25, 1974–1975.
- Deigan, K.E., Li, T.W., Mathews, D.H., and Weeks, K.M. (2009). Accurate SHAPE-directed RNA structure determination. *Proc. Natl. Acad. Sci.* 106, 97–102.
- Dinman, J.D. (2009). The eukaryotic ribosome: current status and challenges. *J. Biol. Chem.* 284, 11761–11765.
- Fletcher, L., Corbin, S.D., Browning, K.S., and Ravel, J.M. (1990). The absence of a m7G cap on beta-globin mRNA and alfalfa mosaic virus RNA 4 increases the amounts of initiation factor 4F required for translation. *J. Biol. Chem.* 265, 19582–19587.
- Flicek, P., Amode, M.R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., et al. (2014). Ensembl 2014. *Nucleic Acids Res.* 42, D749–D755.
- Fraser, C.S., and Doudna, J.A. (2007). Structural and mechanistic insights into hepatitis C viral translation initiation. *Nat. Rev. Microbiol.* 5, 29–38.
- Fromental-Ramain, C., Warot, X., Lakkaraju, S., Favier, B., Haack, H., Birling, C., Dierich, A., Dollé, P., and Chambon, P. (1996). Specific and redundant functions of the paralogous Hoxa-9 and Hoxd-9 genes in forelimb and axial skeleton patterning. *Development* 122, 461–472.
- Hajdin, C.E., Bellaousov, S., Huggins, W., Leonard, C.W., Mathews, D.H., and Weeks, K.M. (2013). Accurate SHAPE-directed RNA secondary structure modeling, including pseudoknots. *Proc. Natl. Acad. Sci. U. S. A.* 110, 5498–5503.
- Harris, D.A., and Walter, N.G. (2001). Probing RNA Structure and Metal-Binding Sites Using Terbium(III) Footprinting. In *Current Protocols in Nucleic Acid Chemistry*, (John Wiley & Sons, Inc.),.
- Holcik, M., Sonenberg, N., and Korneluk, R.G. (2000). Internal ribosome initiation of translation and the control of cell death. *Trends Genet.* 16, 469–473.
- Jan, E., and Sarnow, P. (2002). Factorless Ribosome Assembly on the Internal Ribosome Entry Site of Cricket Paralysis Virus. *J. Mol. Biol.* 324, 889–902.
- Johnston, I.A. (2006). Environment and plasticity of myogenesis in teleost fish. *J. Exp. Biol.* 209, 2249–2264.

Kieft, J.S., Zhou, K., Grech, A., Jubin, R., and Doudna, J.A. (2002). Crystal structure of an RNA tertiary domain essential to HCV IRES-mediated translation initiation. *Nat. Struct. Biol.* 9, 370–374.

Kim, H., Cordero, P., Das, R., and Yoon, S. (2013). HiTRACE-Web: an online tool for robust analysis of high-throughput capillary electrophoresis. *Nucleic Acids Res.* 41, W492–W498.

Kim, J., Yu, S., Shim, B., Kim, H., Min, H., Chung, E.-Y., Das, R., and Yoon, S. (2009). A robust peak detection method for RNA structure inference by high-throughput contact mapping. *Bioinformatics* 25, 1137–1144.

Kladwang, W., VanLang, C.C., Cordero, P., and Das, R. (2011a). A two-dimensional mutate-and-map strategy for non-coding RNA structure. *Nat. Chem.* 3, 954–962.

Kladwang, W., Cordero, P., and Das, R. (2011b). A mutate-and-map strategy accurately infers the base pairs of a 35-nucleotide model RNA. *RNA* 17, 522–534.

Kladwang, W., Mann, T.H., Becka, A., Tian, S., Kim, H., Yoon, S., and Das, R. (2014). Standardization of RNA chemical mapping experiments. *Biochemistry* 53, 3063–3065.

Komar, A.A., and Hatzoglou, M. (2005). Internal ribosome entry sites in cellular mRNAs: mystery of their existence. *J. Biol. Chem.* 280, 23425–23428.

Komili, S., Farny, N.G., Roth, F.P., and Silver, P.A. (2007). Functional specificity among ribosomal proteins regulates gene expression. *Cell* 131, 557–571.

Kondrashov, N., Pusic, A., Stumpf, C.R., Shimizu, K., Hsieh, A.C., Xue, S., Ishijima, J., Shiroishi, T., and Barna, M. (2011). Ribosome-Mediated Specificity in Hox mRNA Translation and Vertebrate Tissue Patterning. *Cell* 145, 383–397.

Landry, D.M., Hertz, M.I., and Thompson, S.R. (2009). RPS25 is essential for translation initiation by the Dicistroviridae and hepatitis C viral IRESs. *Genes Dev.* 23, 2753–2764.

Lee, A.S., Burdeinick-Kerr, R., and Whelan, S.P.J. (2013). A ribosome-specialized translation initiation pathway is required for cap-dependent translation of vesicular stomatitis virus mRNAs. *Proc. Natl. Acad. Sci. U. S. A.* 110, 324–329.

Livingstone, M., Atas, E., Meller, A., and Sonenberg, N. (2010). Mechanisms governing the control of mRNA translation. *Phys. Biol.* 7, 021001.

- Lukavsky, P.J., Kim, I., Otto, G.A., and Puglisi, J.D. (2003). Structure of HCV IRES domain II determined by NMR. *Nat. Struct. Biol.* **10**, 1033–1038.
- Mortimer, S.A., and Weeks, K.M. (2007). A fast-acting reagent for accurate analysis of RNA secondary and tertiary structure by SHAPE chemistry. *J. Am. Chem. Soc.* **129**, 4144–4145.
- Oh, S., Scott, M.P., and Sarnow, P. (1992). Homeotic gene Antennapedia mRNA confer translational initiation by internal ribosome binding. *Genes Dev.* **6**, 1643–1653.
- Phinney, D.G., Gray, A.J., Hill, K., and Pandey, A. (2005). Murine mesenchymal and embryonic stem cells express a similar Hox gene profile. *Biochem. Biophys. Res. Commun.* **338**, 1759–1765.
- Plank, T.-D.M., and Kieft, J.S. (2012). The structures of nonprotein-coding RNAs that drive internal ribosome entry site function. *Wiley Interdiscip. Rev. RNA* **3**, 195–212.
- Pyronnet, S., Dostie, J., and Sonenberg, N. (2001). Suppression of cap-dependent translation in mitosis. *Genes Dev.* **15**, 2083–2093.
- Ray, P.S., Grover, R., and Das, S. (2006). Two internal ribosome entry sites mediate the translation of p53 isoforms. *EMBO Rep.* **7**, 404–410.
- Reuter, J.S., and Mathews, D.H. (2010). RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics* **11**, 129.
- Riley, A., Jordan, L.E., and Holcik, M. (2010). Distinct 5' UTRs regulate XIAP expression under normal growth conditions and during cellular stress. *Nucleic Acids Res.* **38**, 4665–4674.
- Seetin, M.G., Kladwang, W., Bida, J.P., and Das, R. (2014). Massively Parallel RNA Chemical Mapping with a Reduced Bias MAP-Seq Protocol. In *RNA Folding: Methods and Protocols*, C. Waldsich, ed. (Humana Press), pp. 95–117.
- Sonenberg, N., and Hinnebusch, A.G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* **136**, 731–745.
- Spahn, C.M.T., Jan, E., Mulder, A., Grassucci, R.A., Sarnow, P., and Frank, J. (2004). Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. *Cell* **118**, 465–475.
- Stoneley, M., Spencer, J.P., and Wright, S.C. (2001). An internal ribosome entry segment in the 5' untranslated region of the *mnt* gene. *Oncogene* **20**, 893–897.

Stupina, V.A., Yuan, X., Meskauskas, A., Dinman, J.D., and Simon, A.E. (2011). Ribosome binding to a 5' translational enhancer is altered in the presence of the 3' untranslated region in cap-independent translation of turnip crinkle virus. *J. Virol.* **85**, 4638–4653.

Talkish, J., May, G., Lin, Y., Woolford Jr, J.L., and Mcmanus, C.J. (2014). Mod-seq : high-throughput sequencing for chemical probing of RNA structure. *RNA* **20**, 1–8.

Thompson, S.R. (2012). So you want to know if your message has an IRES? *Wiley Interdiscip. Rev. RNA* **3**, 697–705.

Tian, S., Cordero, P., Kladwang, W., and Das, R. (2014). High-throughput mutate-map-rescue evaluates SHAPE-directed RNA structure and uncovers excited states. *RNA* **20**, 1815–1826.

Ungureanu, N.H., Cloutier, M., Lewis, S.M., de Silva, N., Blais, J.D., Bell, J.C., and Holcik, M. (2006). Internal ribosome entry site-mediated translation of Apaf-1, but not XIAP, is regulated during UV-induced cell death. *J. Biol. Chem.* **281**, 15155–15163.

Vesper, O., Amitai, S., Belitsky, M., Byrgazov, K., Kaberdina, A.C., Engelberg-Kulka, H., and Moll, I. (2011). Selective Translation of Leaderless mRNAs by Specialized Ribosomes Generated by MazF in *Escherichia coli*. *Cell* **147**, 147–157.

Wilkinson, K.A., Merino, E.J., and Weeks, K.M. (2006). Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nat. Protoc.* **1**, 1610–1616.

Xia, X., and Holcik, M. (2009). Strong eukaryotic IRESs have weak secondary structure. *PLoS One* **4**, e4136.

Xue, S., and Barna, M. (2012). Specialized ribosomes: a new frontier in gene regulation and organismal biology. *Nat. Rev. Mol. Cell Biol.* **13**, 355–369.

Yoon, A., Peng, G., Brandenburger, Y., Brandenburg, Y., Zollo, O., Xu, W., Rego, E., and Ruggero, D. (2006). Impaired control of IRES-mediated translation in X-linked dyskeratosis congenita. *Science* **312**, 902–906.

Yoon, S., Kim, J., Hum, J., Kim, H., Park, S., Kladwang, W., and Das, R. (2011). HiTRACE: high-throughput robust analysis for capillary electrophoresis. *Bioinformatics* **27**, 1798–1805.

Figures

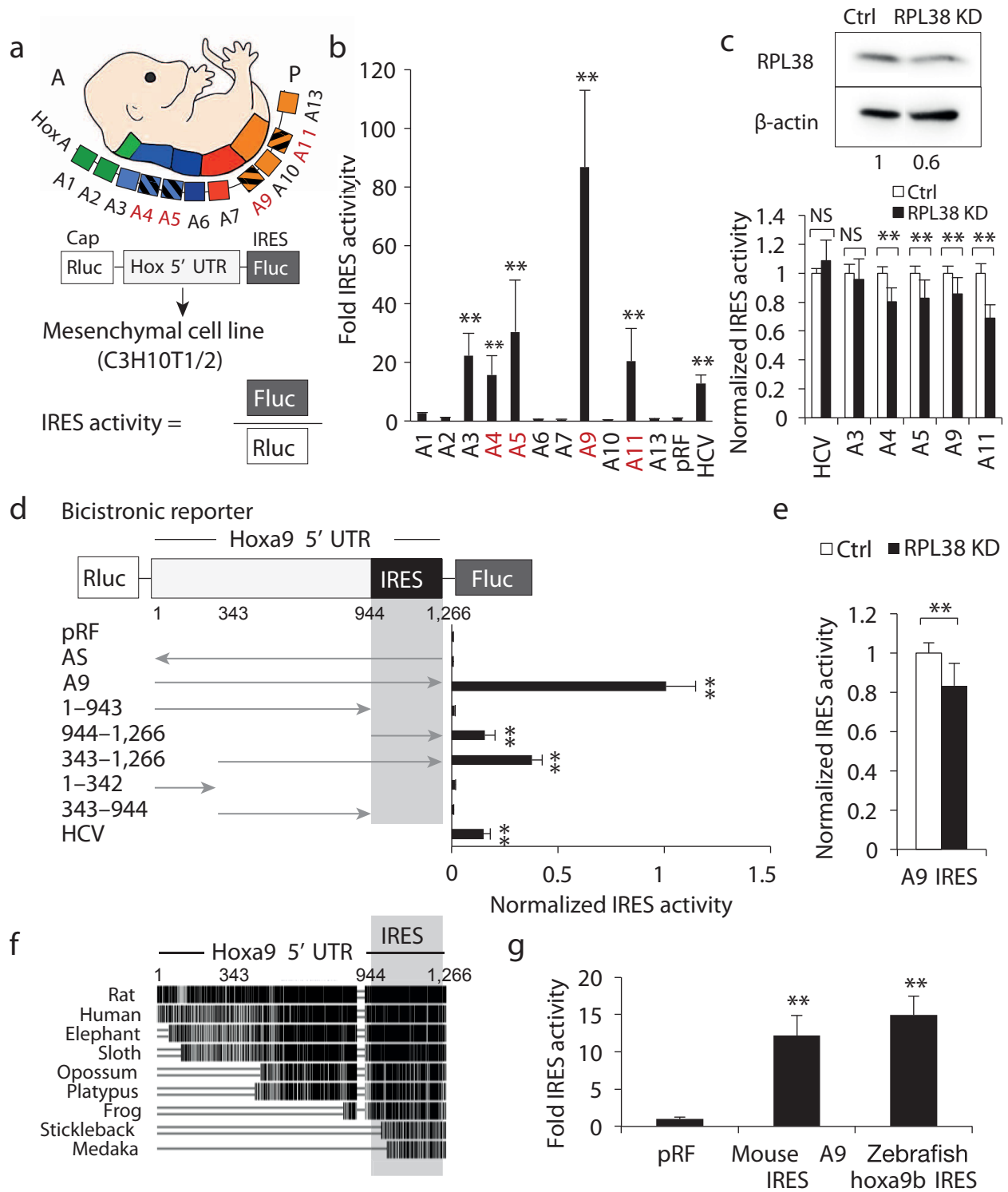


Figure 2.1 Select *HoxA* 5' UTRs contain IRES elements regulated by RPL38.

a, *HoxA* genes are located in tandem along a chromosomal locus, and their location directs their anterior expression boundaries along the embryo. Striped boxes, *Hox* genes translationally regulated by RPL38. IRES activity is the ratio between Fluc and Rluc. **b**, IRES activity of *HoxA* 5' UTRs. IRES activity of UTRs was normalized to that of empty vector (pRF) and RPL38-regulated *Hox* genes are denoted in red. $**P < 0.01$ (χ^2 test against pRF). $n = 5$. **c**, RPL38 regulates translation of specific *HoxA* IRES elements. Bicistronic reporters were transfected into control cells (Ctrl) or cells where one copy of *Rpl38* was deleted (RPL38 KD). Top: western blot of RPL38 and β -actin with quantification of RPL38 levels below blot. Bottom: IRES activity of specific *Hox* 5' UTRs in RPL38 knockdown cells normalized to that of control cells. $**P < 0.01$ (t -test), $n = 3$. **d**, Deletion analysis of the *Hoxa9* 5' UTR. IRES activity was normalized to that of the full-length *Hoxa9* 5' UTR (A9). AS: anti-sense. $**P < 0.01$ (χ^2 test against pRF), $n = 3$. **e**, IRES activity of the *Hoxa9* IRES (nt 944–1,266) in RPL38 knockdown cells. $**P < 0.01$ (t -test), $n = 3$. **f**, Alignment of the *Hoxa9* 5' UTR in vertebrates. **g**, The 5' UTR of one of two zebrafish *Hoxa9* paralogs (*hoxa9b*) has similar IRES activity as the mouse *Hoxa9* IRES. IRES activity was normalized to pRF. $**P < 0.01$ (t -test as compared to pRF). $n = 3$. All experiments were performed in triplicates. Data represent mean $\pm$ s.d.

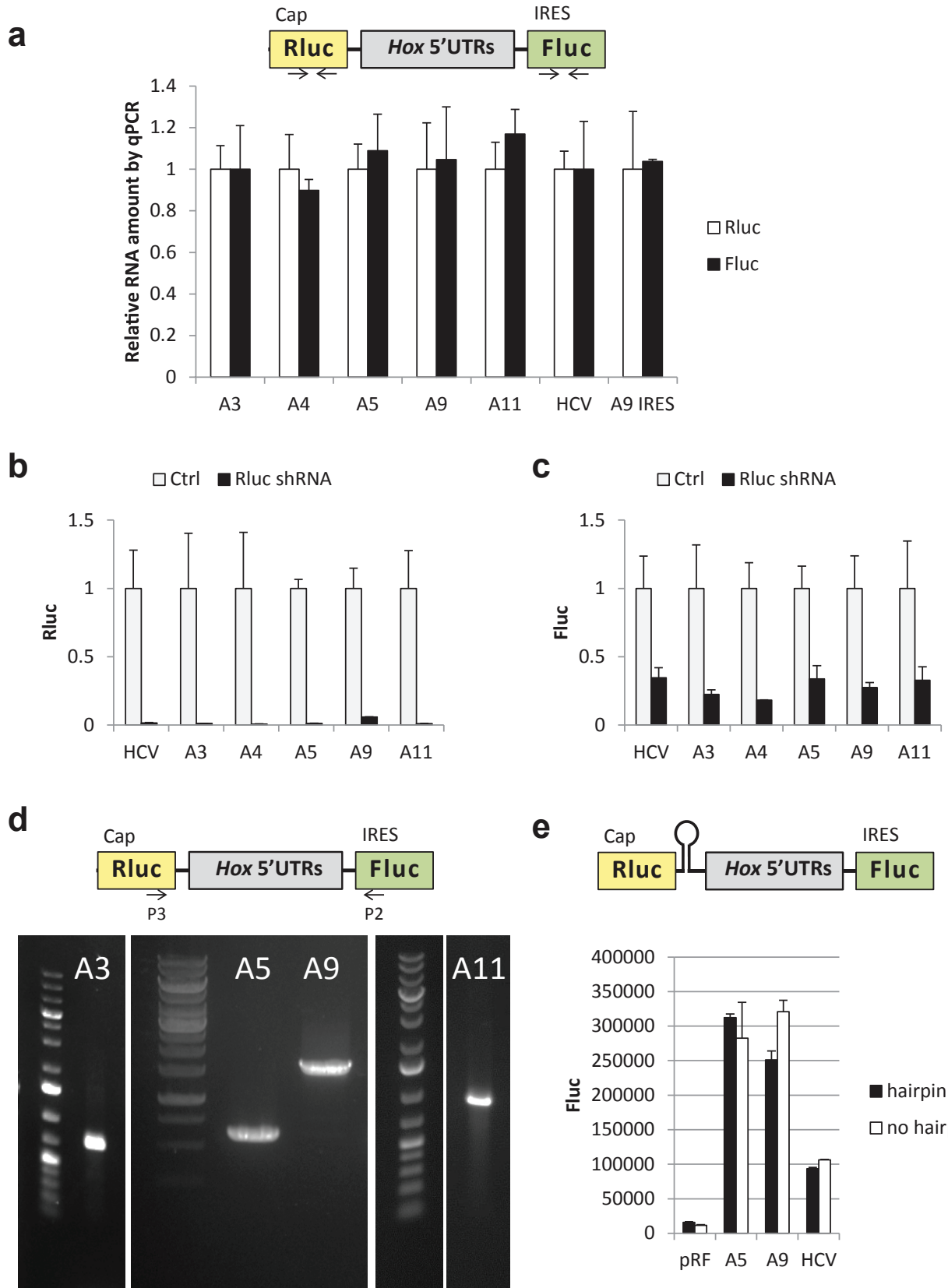


Figure 2.2 *HoxA* IRES controls confirming that Fluc activity from bicistronic vector is due to IRES activity.

a, qPCR of both Rluc and Fluc from transfected cells shows that Rluc and Fluc are produced at the same ratio. All Rluc and Fluc values are normalized to that of HCV (set to 1). $n = 3$, individual experiments performed in duplicate. **b**, **c**, shRNA against Rluc decreased reporter activity of both Rluc (**a**) and Fluc (**b**), confirming that Rluc and Fluc were transcribed on the same mRNA. $n = 3$, individual experiments performed in triplicate. **d**, RT-PCR using primers in Rluc and Fluc show that there is no cryptic splice site in the cloned *Hox* 5' UTR. Primer locations are shown as arrows in the diagram. **e**, Inserting a strong hairpin ($-67 \text{ kcal mol}^{-1}$) after the Rluc reporter did not affect Fluc activity, suggesting Fluc activity was not due to ribosome read-through.

a



b

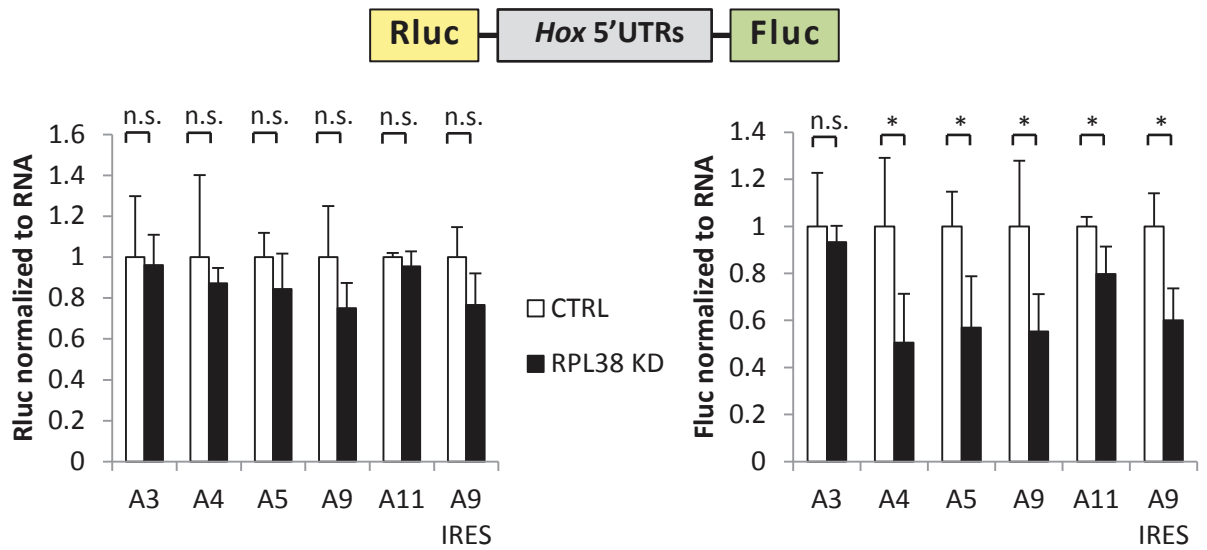


Figure 2.3 Disruption of *Rpl38* in C3H10T1/2 by TALEN nucleases.

a, Location of TALEN pairs. Two pairs of TALENs were designed to bind at the end of exon 2 and the beginning of exon 3 to make a genomic break close to the ATG. Sequencing of a positive clone shows a deletion of the ATG and most of the intron after it. Coding sequence is highlighted in green. **b** *Rpl38* knockdown does not change cap-dependent translation (Rluc) but decreases IRES-dependent translation (Fluc) from specific *Hox* 5' UTRs. Luciferase activity was normalized to amount of Fluc RNA in the cells as quantified by qPCR. * $P < 0.05$ (t -test compared to control). $n = 2$, individual experiments performed in duplicate.

Figure 2.4 Alignment of *Hoxa9* IRES element between vertebrate species.

Nucleotides 945 to 1,266 of the mouse *Hoxa9* 5' UTR were aligned with sequences from other vertebrates and show high sequence homology.

Nucleotides are coloured based on their homology where darker colours represent higher conservation.

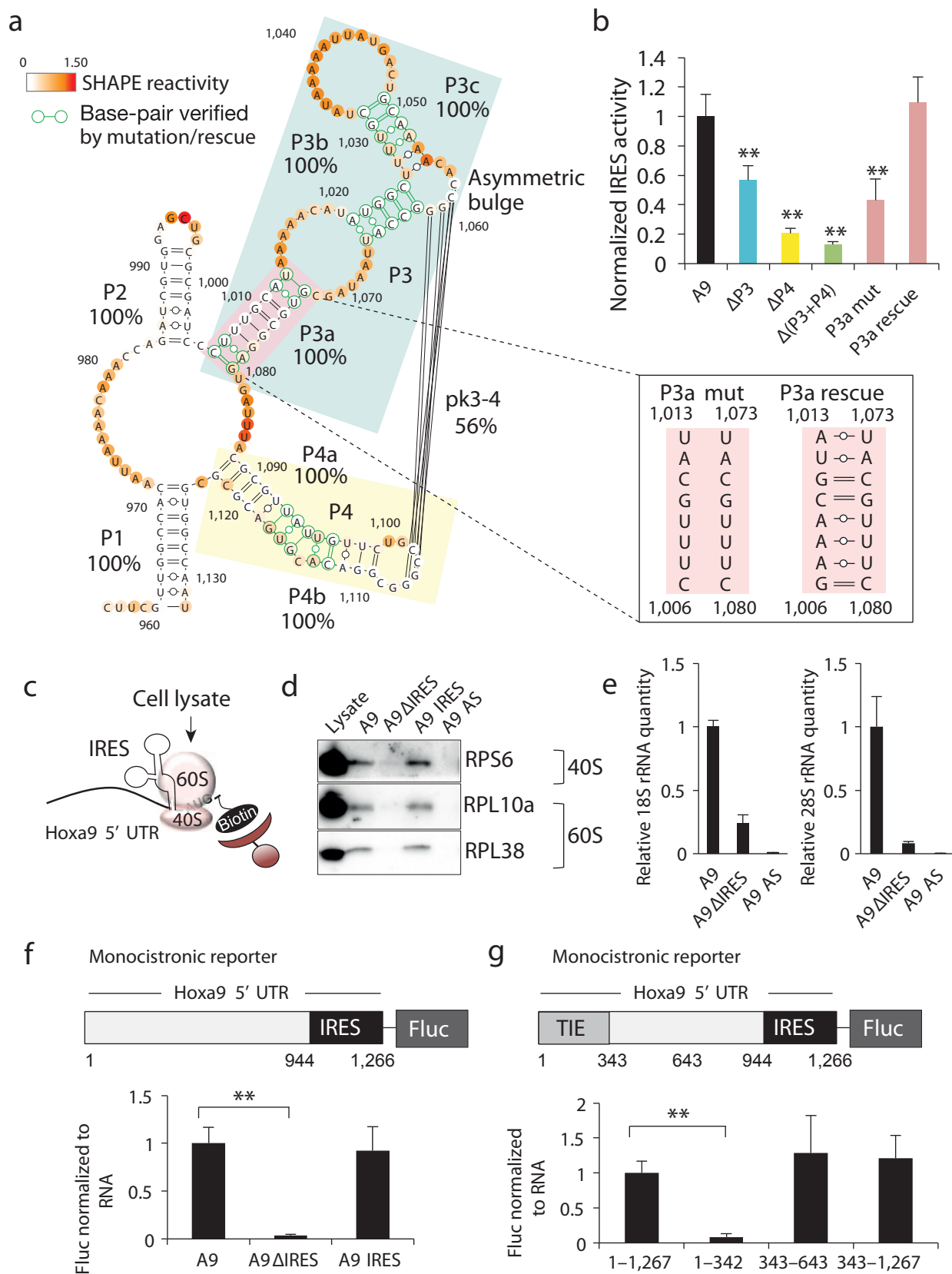


Figure 2.5 The *Hoxa9* IRES is an RNA structure that facilitates recruitment of the ribosome and is required for normal translation.

a, Secondary structure of *Hoxa9* 5' UTR (nt 957–1,132) by SHAPE and mutate-and-map analysis. Nucleotides are coloured with SHAPE reactivities.

Percentages give bootstrap support values for each helix, including a tentative pseudoknot (pk). Base-pairs verified by mutation and rescue are circled green.

Inset: mutations made in helix P3a for functional testing in **b**. **b**, Deletions as well

as mutation-and-rescue of specific hairpins identified in *Hoxa9* IRES structure

shown in **a** affects IRES activity in bicistronic reporter assays. IRES activity was

normalized to full-length *Hoxa9* 5' UTR (A9, set to 1). ** $P < 0.01$ (t -test as

compared to A9). $n = 3$ and experiments performed in triplicates. **c**, Schematic of

RNA pull-down experiments. **d**, **e**, Representative western blot for RPs (**d**) and

qPCR for rRNA (**e**) from eluted material of two independent RNA pull-down

experiments. The 18S and 28S rRNA amounts of all *Hoxa9* deletion mutants

were normalized to amount of full-length *Hoxa9* 5' UTR (set to 1) eluted from

beads. $n = 2$. **f**, Deletions within the *Hoxa9* 5' UTR in a monocistronic reporter

assay. Fluc activity was normalized to amount of luciferase mRNA determined by

quantitative PCR (qPCR). All values are normalized to *Hoxa9* full-length 5' UTR

(A9), set to 1; $n = 4$. **g**, A TIE that inhibits cap-dependent translation was

identified in the *Hoxa9* 5' UTR (nt 1–342); $n = 3$ and experiments were performed

in duplicates. Data represent mean $\pm$ s.d.

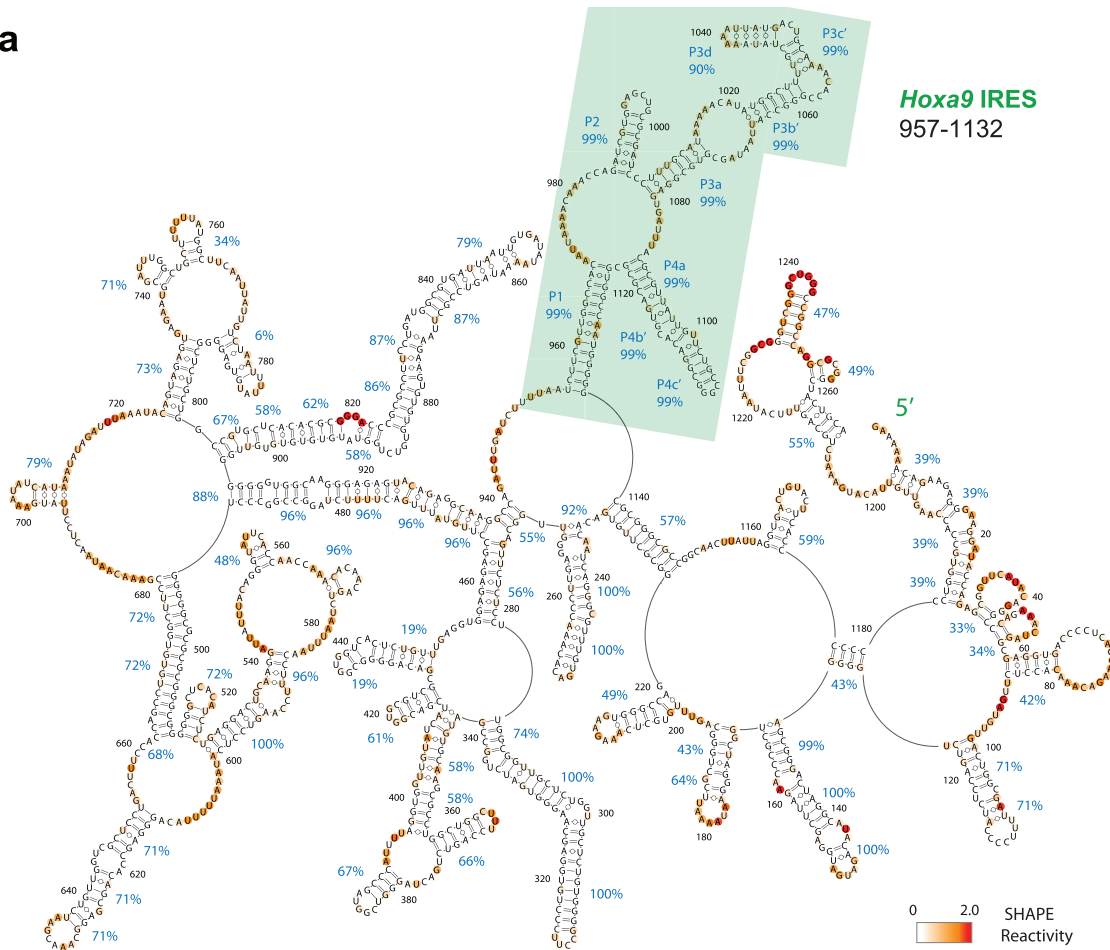
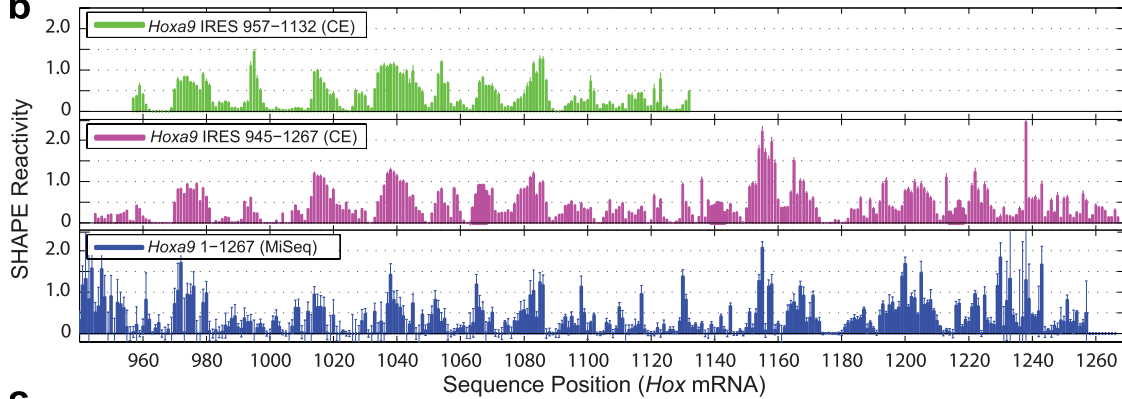
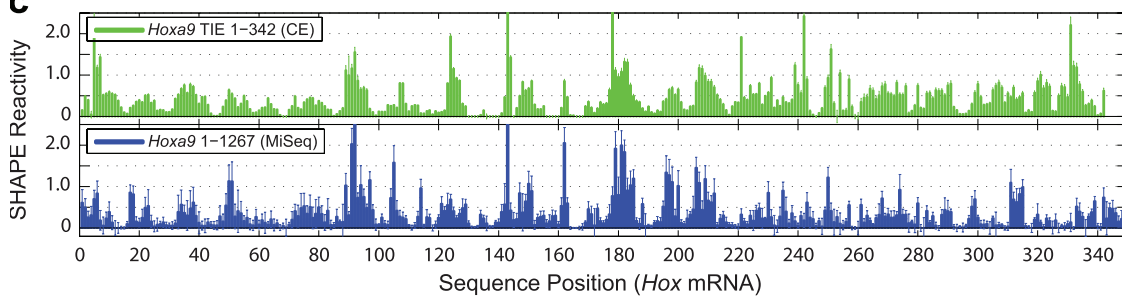
a**b****c**

Figure 2.6 Chemical mapping and secondary structure modeling of full-length *Hoxa9* 5' UTR.

a, Secondary structure modeling of full-length *Hoxa9* using ligSHAPE data. The *Hoxa9* IRES element (nt 957–1,132, shaded in green) is modeled as the same secondary structure shown in Figure **2.5a**. Confidence values from bootstrapping (bulge percentages) exceed 90% for this element, suggesting a well-determined subdomain, but are generally low outside this region, indicating poor certainty in other regions. **b**, Normalized SHAPE reactivity of *Hoxa9* IRES (nt 957–1,132 and 944–1,266 from one-dimensional SHAPE read out through capillary electrophoresis (CE), full-length 1–1,266 from MiSeq-based ligSHAPE). **c**, Normalized SHAPE reactivity of *Hoxa9* TIE (nt 1–342 from CE-based one-dimensional SHAPE, full-length 1–1,266 from MiSeq-based ligSHAPE).

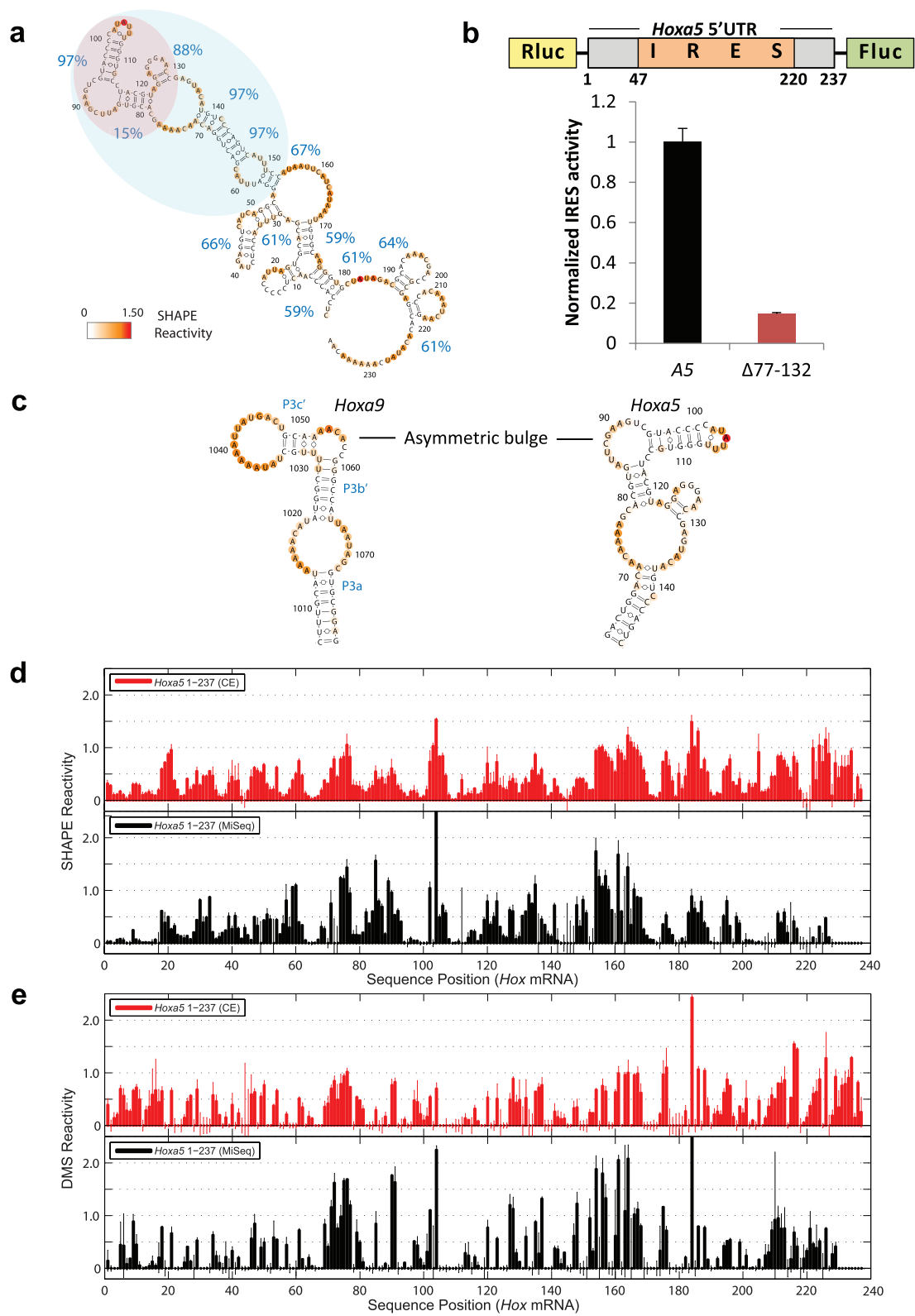


Figure 2.7 Chemical mapping and secondary structure modeling of full length *Hoxa5* 5' UTR.

a, Secondary structure modeling of *Hoxa5* using one-dimensional SHAPE data. Nucleotides are coloured with SHAPE reactivities. Percentage labels give bootstrap support values for each helix. The feature highlighted in blue resembles P3 in *Hoxa9* and the tip highlighted in pink is deleted in **b**. **b**, The deletion of the tip identified in *Hoxa5* IRES structure shown in **a** decreases IRES activity in bicistronic reporter assays. IRES activity was normalized to full length *Hoxa5* 5' UTR (A5, set to 1). $**P < 0.01$ (*t*-test as compared to A5). $n = 2$ experiments, performed in triplicate. **c**, Both *Hoxa9* and *Hoxa5* contain an asymmetric bulge in a region important for IRES activity. **d**, **e**, Normalized SHAPE (**d**) and DMS (**e**) reactivity of *Hoxa5* (CE-based and MiSeq-based).

Figure 2.8 Secondary structure modeling and mutate-and-map (M^2) data set of *Hoxa9* IRES element.

a-b, Entire M^2 data set and Z-score contact-map of *Hoxa9* nt 957–1,132 across 177 single mutants probed by 1M7. **c**, Secondary structure model of *Hoxa9* nucleotides 957–1,132 using M^2 data alone. **d**, Secondary structure model of *Hoxa9* nt 957–1,132 using one-dimensional SHAPE data alone. Nucleotides are coloured with SHAPE reactivity. **e**, Secondary structure model of *Hoxa9* nt 944–1,266 using one-dimensional SHAPE data. The models in c-e contain the same helices as the model from combined SHAPE/ M^2 analysis in Figure **2.5a**, up to register shifts and edge base pairs; the small rearrangements are labelled P3b', P3c', P3d, P4b', and P4c'.

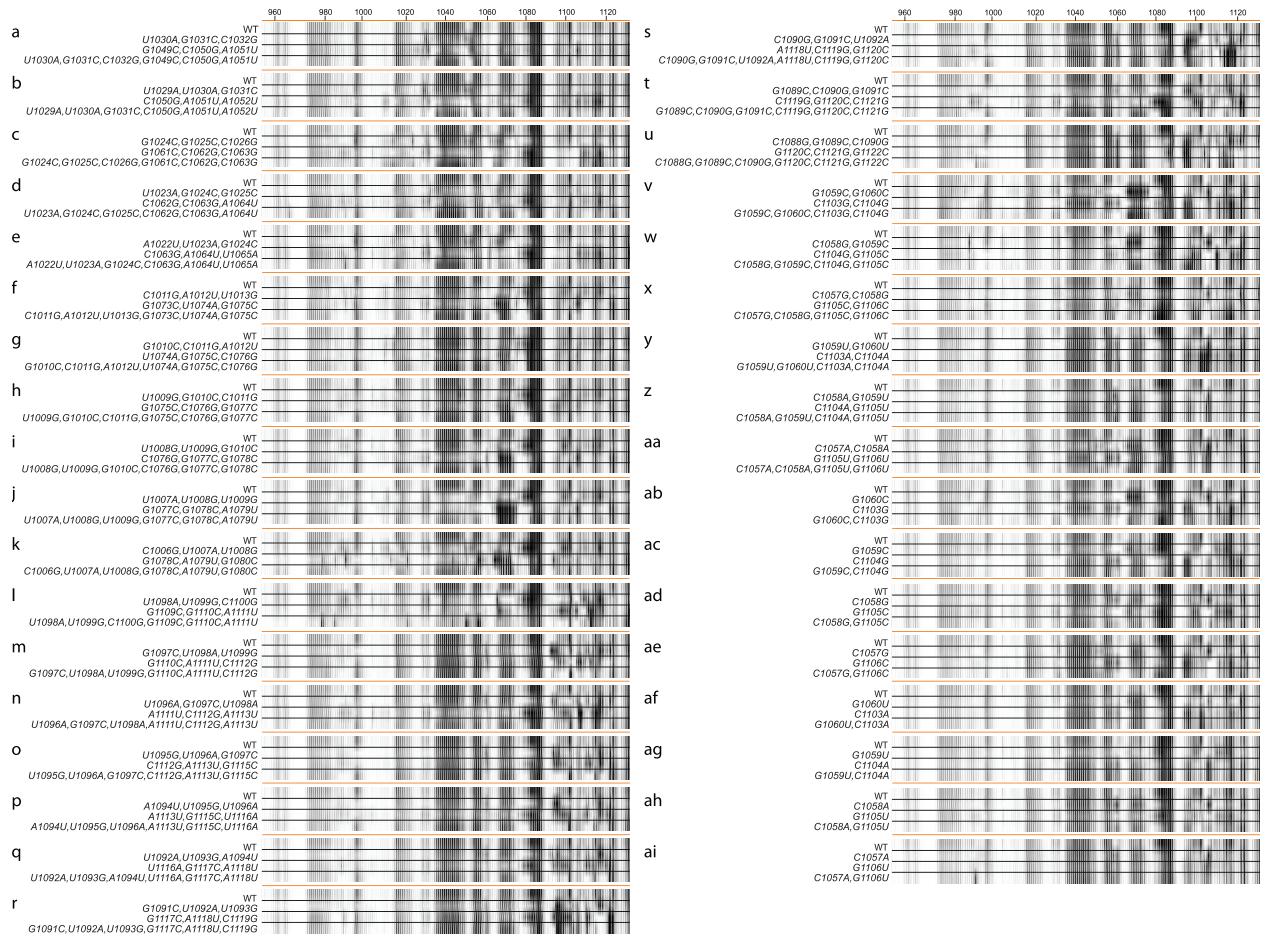


Figure 2.9 Mutation/rescue results of *Hoxa9* IRES structure (nt 944–1,266) probed by 1M7.

Electropherograms of mutation/rescue to test base-pairings in P3c (**a, b**), P3b (**c–e**), P3a (**f–k**), P4b (**l–o**), P4a (**p–u**) and pk3-4 (**v–ai**). Perturbation of the chemical mapping reactivities by mutations of one strand and restoration by mutations in the other strand provide strong evidence for the tested pairings in P3c (**a, b**), P3b (**c–e**), P3a (**f–k**), P4b (**m–o**) and P4a (**p–q, t**). Near perfect restoration by compensatory mutations in (**x**) and (**ad**) support pseudoknot pk3-4. Lack of rescue in other tested pairings is consistent with either absence of those pairings or higher-order structure (for example, base triples) interacting with those pairings.

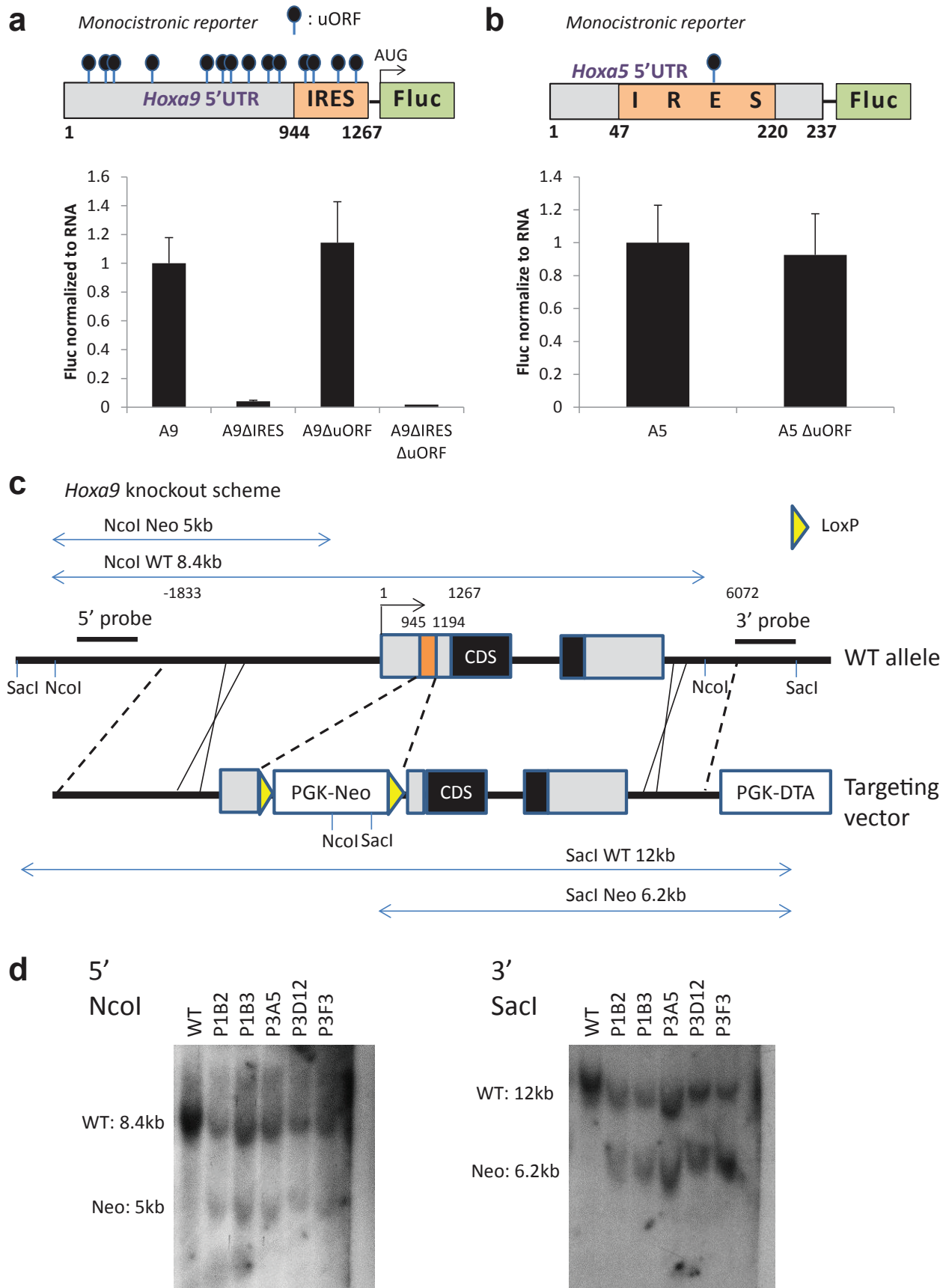
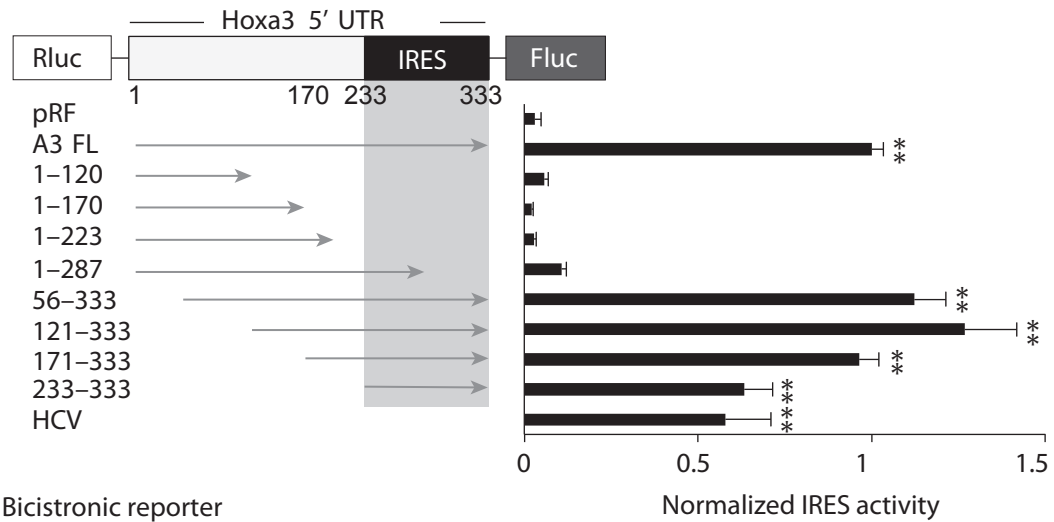


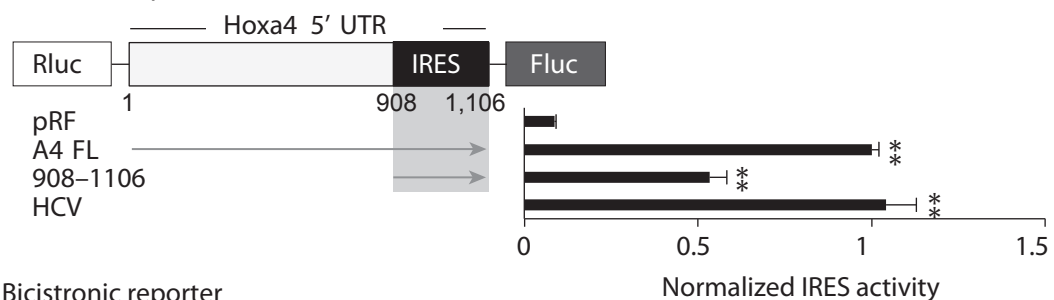
Figure 2.10 Putative uORFs within the 5' UTRs of *Hoxa9* and *Hoxa5* do not inhibit cap-dependent translation and *Hoxa9*^{ΔIRES} targeting strategy.

uORFs are marked by black circles on the diagram of monocistronic reporter for the *Hoxa9* (a) and *Hoxa5* (b) 5' UTR. All the ATGs in the 5' UTR were mutated to TTG in A9ΔuORF construct and GTG in A5ΔuORF. The IRES element (944–1,266) was removed in A9ΔIRES construct. The IRES element was removed from the A9ΔuORF construct in A9ΔIRESΔuORF. *n* = 3 individual experiments in duplicates. Data represent mean ± s.d. c, Diagrams of the *Hoxa9* locus and the targeting vector. Boxes represent exons, grey boxes represent UTRs and black boxes represent the coding sequence. Nucleotides 944–1,145 were replaced by a floxed Neo cassette in the targeting vector. Locations of Southern blot probes, restriction enzymes used for Southern analysis and expected sizes are marked on the diagrams. d, Southern blot analysis of targeted cells using both the 5' and 3' probes showing that both arms integrated correctly into the *Hoxa9* locus. Mice were generated from clone P3A5.

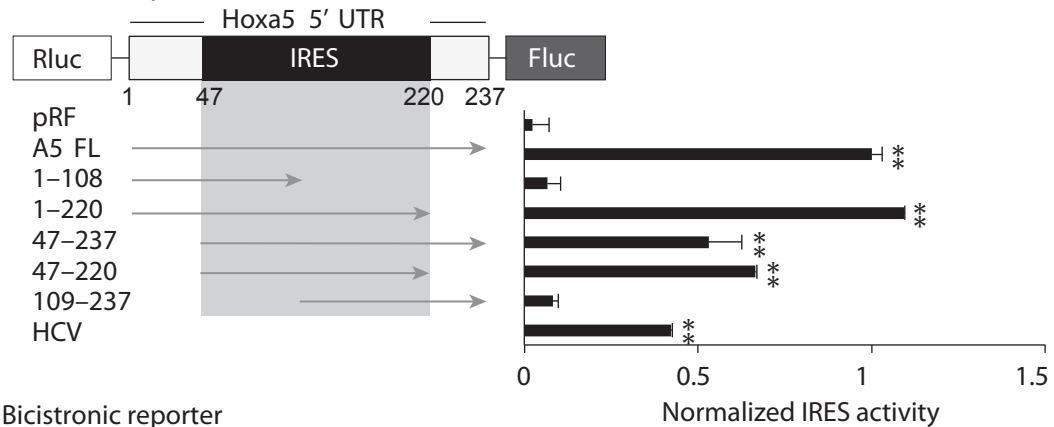
a Bicistronic reporter



b Bicistronic reporter



c Bicistronic reporter



d Bicistronic reporter

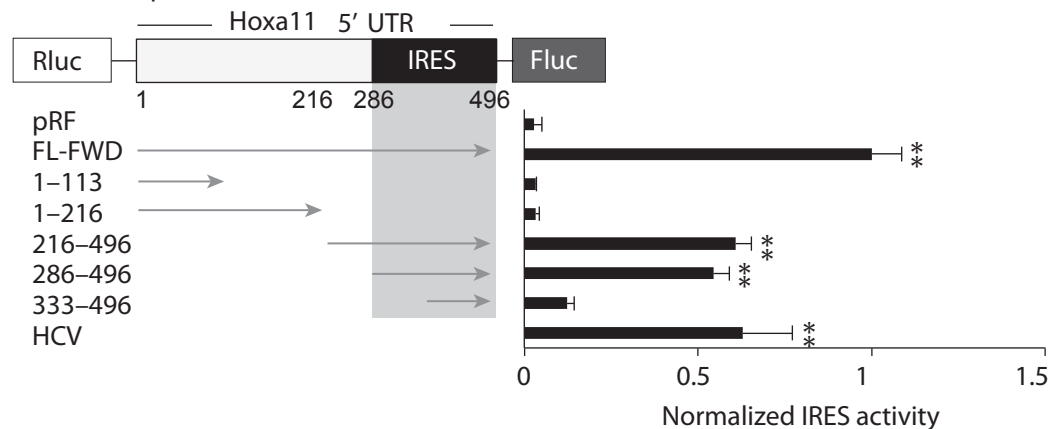
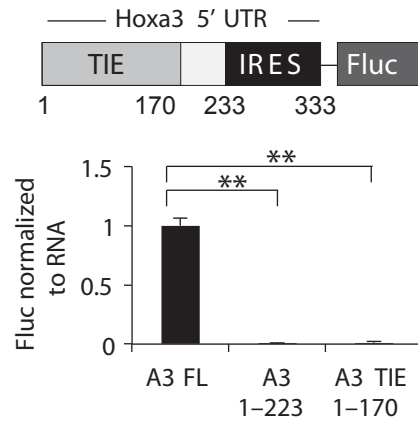


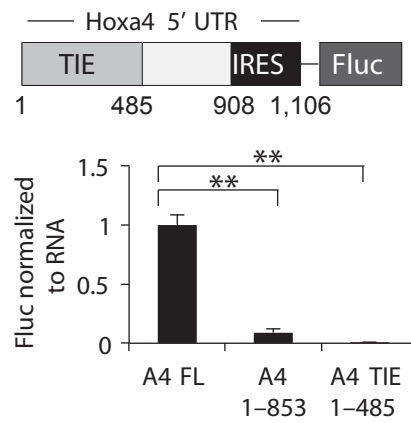
Figure 2.11 Identification of minimal IRES domains in all IRES-containing *HoxA* 5' UTR.

a–d, Deletion analysis of *HoxA* 5' UTRs performed in bicistronic constructs maps the IRES element in *Hoxa3* (**a**), *Hoxa4* (**b**), *Hoxa5* (**c**) and *Hoxa11* (**d**). IRES activity was normalized to full-length 5' UTRs (FL). ** $P < 0.01$ (χ^2 test against pRF). $n = 3$ and experiments were performed in triplicate. Data represent mean $\pm$ s.d.

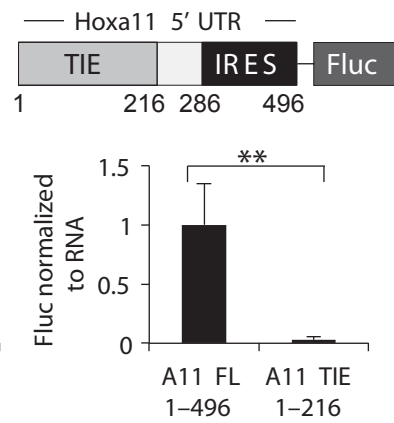
a Monocistronic reporter



b Monocistronic reporter



c Monocistronic reporter



d

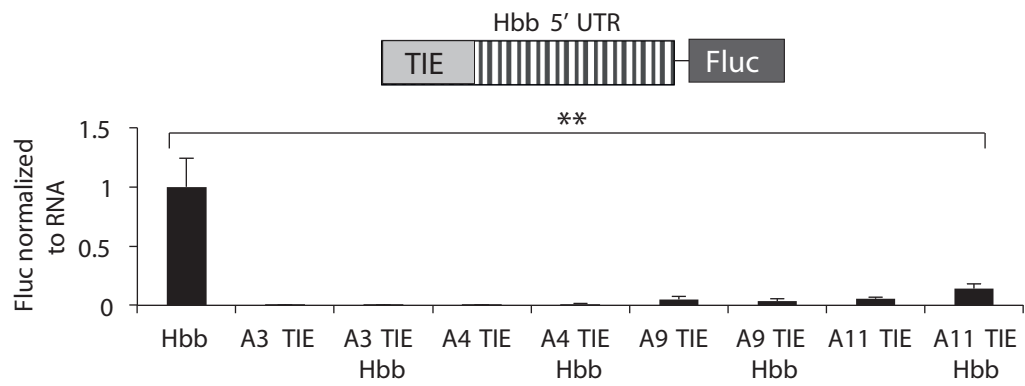


Figure 2.12 Identification of a common TIE that blocks cap-dependent translation of Hox IRES-mRNAs.

a–c, Deletion of the minimal IRES elements within monocistronic reporters of *HoxA* 5' UTR and identification of TIEs close to the mRNA cap in *Hoxa3* (**a**), *Hoxa4* (**b**) and *Hoxa5* (**c**). All values are normalized to the full-length constructs (FL); $n = 3$. **d**, *HoxA* TIEs were placed upstream of the β -globin 5' UTR (*Hbb*) in a monocistronic reporter to assay inhibition of cap-dependent translation; $n = 3$. All experiments were performed in duplicate. Data represent mean $\pm$ s.d. $**P < 0.01$ (t-test).

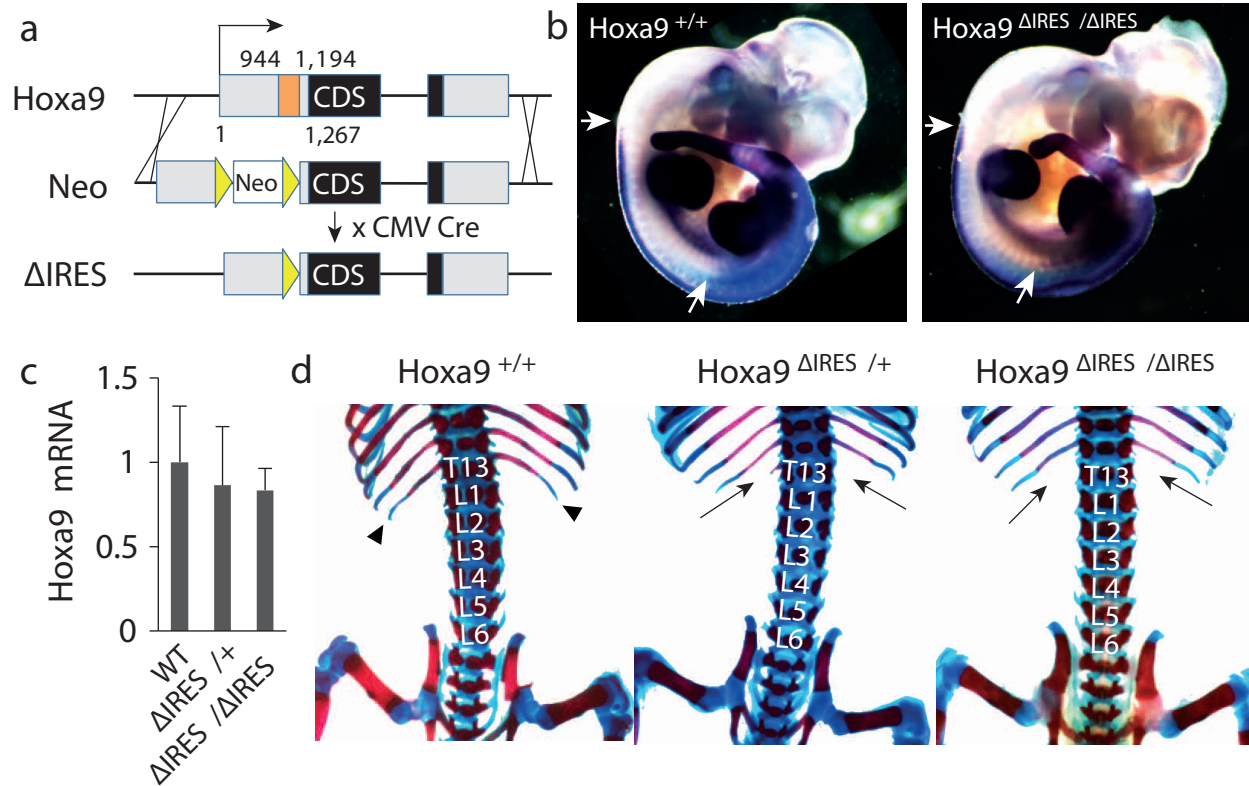


Figure 2.13 The *Hoxa9* IRES is required for axial skeleton patterning but not for *Hoxa9* mRNA expression.

a, Scheme of *Hoxa9* IRES targeted deletion in the mouse. *LoxP* sites are depicted by yellow triangles. The region deleted is highlighted in orange. **b**, Representative *in situ* hybridization of *Hoxa9* mRNA expression at E11.5; $n = 3$ embryos of each genotype. Arrows indicate the anterior boundaries of expression in the neural tube and somites. **c**, qPCR shows little difference in *Hoxa9* transcript levels in neural tube and somites between *Hoxa9*^{+/+}, *Hoxa9*^{ΔIRES/+} and *Hoxa9*^{ΔIRES/ΔIRES} embryos. *Hoxa9*^{ΔIRES/+}, $P = 0.06$; *Hoxa9*^{ΔIRES/ΔIRES} $P = 0.03$, *t*-test as compared to *Hoxa9*^{+/+}. $n = 5$ embryos of each genotype. **d**, Representative skeletons of *Hoxa9*^{+/+}, *Hoxa9*^{ΔIRES/+} and *Hoxa9*^{ΔIRES/ΔIRES} mice. Arrows point to the missing rib(s) on T13, normally present in WT mice (arrowheads) revealing a posterior homeotic transformation to L1. $n = 5$ mice of each genotype.

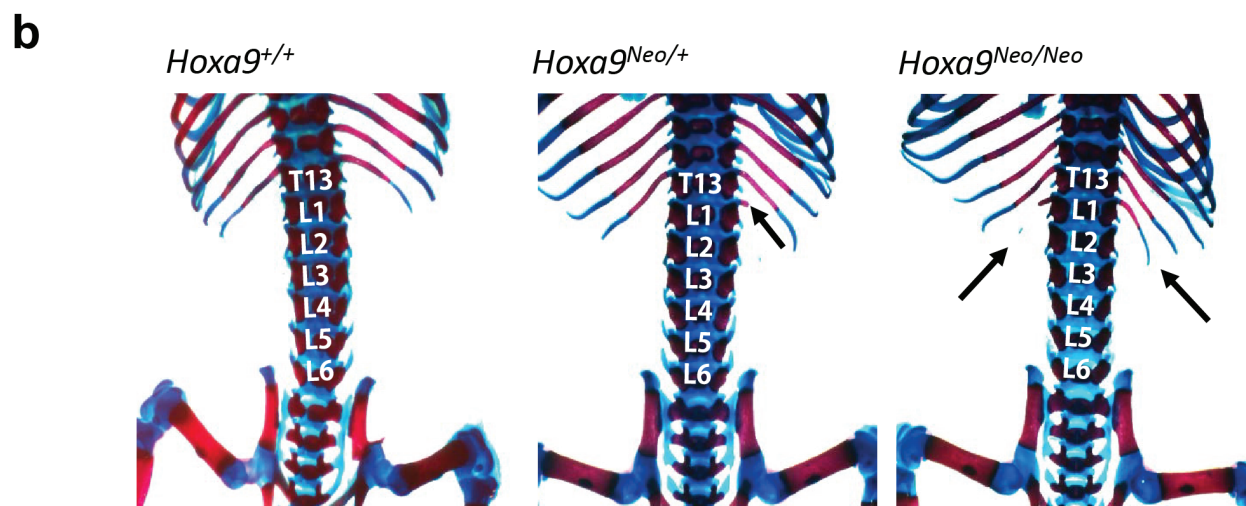
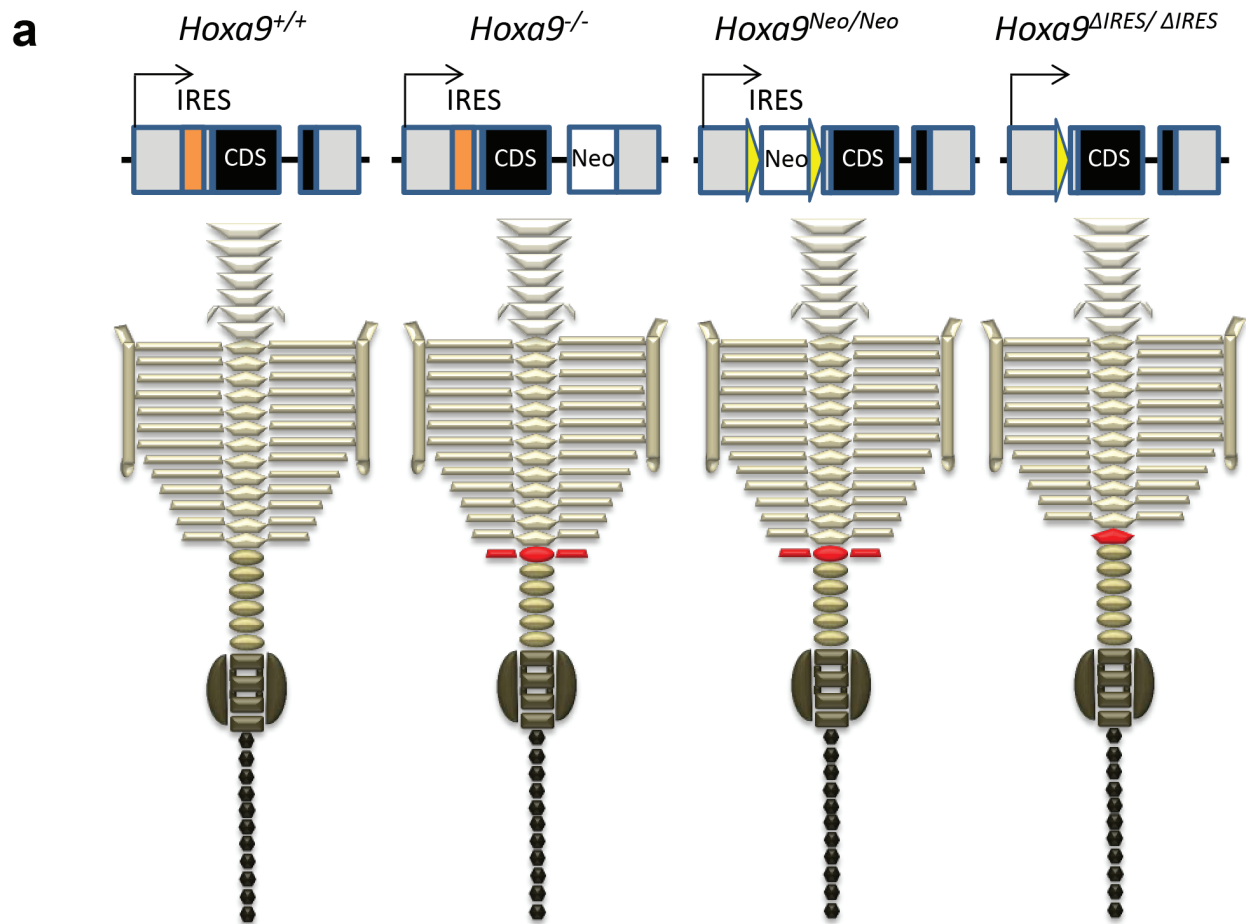


Figure 2.14 The presence of Neo cassette in the *Hoxa9* locus is linked to the presence of an L1 → T13 homeotic transformation.

a, Diagram of the *Hoxa9* locus (top) and axial skeleton phenotype (bottom) in different *Hoxa9* mouse mutants. The original *Hoxa9*^{-/-} was made by replacing the homeodomain with a Neo cassette. Vertebra with homeotic transformation is coloured red. **b**, Representative skeletons of *Hoxa9*^{+/+}, *Hoxa9*^{Neo/+} and *Hoxa9*^{Neo/Neo}. Arrows point to the additional rib(s) on L1, revealing a homeotic transformation to T13. These results show that it is the presence of Neo in the targeting locus, which may affect the expression of neighbouring Hox gene, that is sufficient to cause the L1 → T13 phenotype. When the Neo cassette is removed from the targeting locus by crossing the *Hoxa9*^{Neo/+} mouse with a CMV Cre line, the L1 → T13 phenotype is no longer present. *n* = 3 skeletons of each genotype.

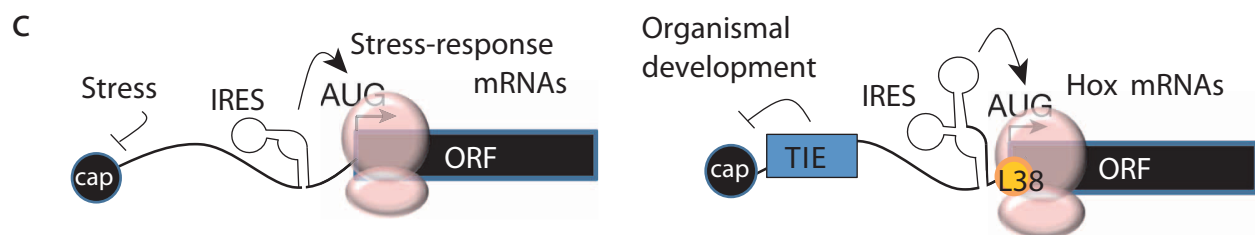
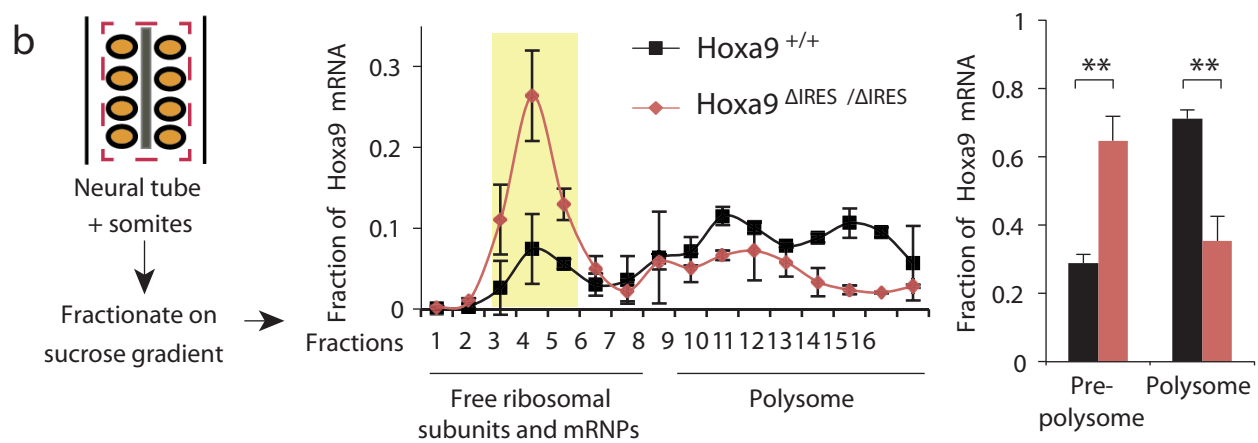
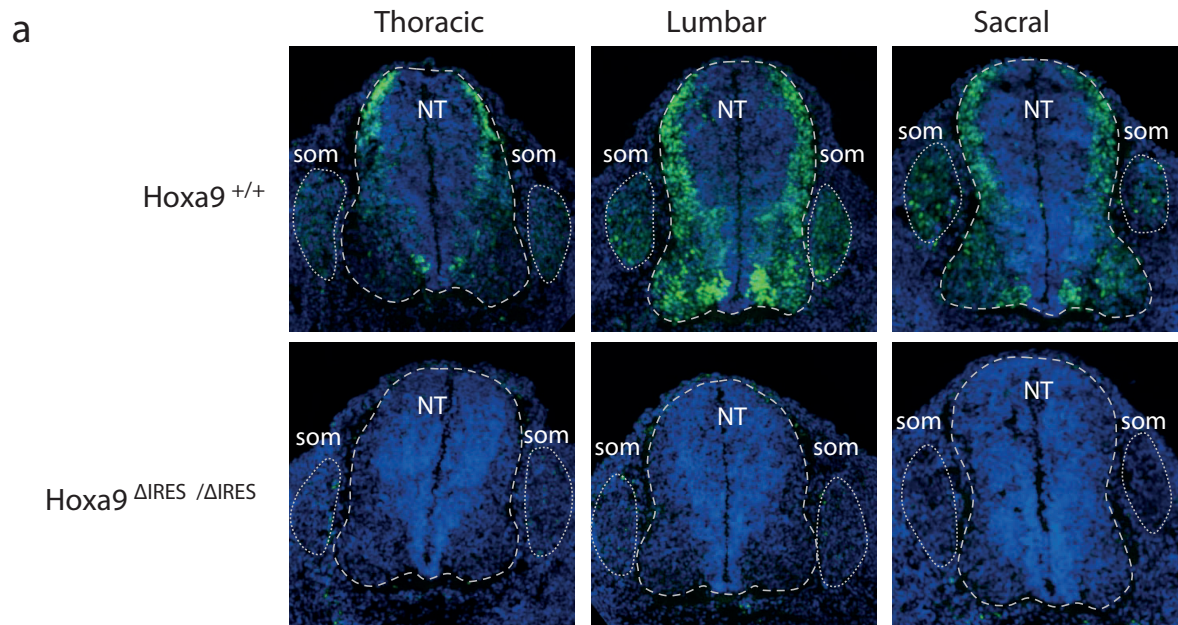


Figure 2.15 The *Hoxa9* IRES is critically required for *Hoxa9* translation *in vivo*.

a, Representative immunostaining of HOXA9 (green) in cross-section of E11.5 *Hoxa9*^{+/+} (top) and *Hoxa9*^{ΔIRES/ΔIRES} (bottom) embryos from the thoracic to sacral levels. DAPI (4',6-diamidino-2-phenylindole) staining is in blue. NT: neural tube, som: somite. *n* = 3 embryos of each genotype. **b**, Somites and neural tubes of E11.5 *Hoxa9*^{+/+} and *Hoxa9*^{ΔIRES/ΔIRES} embryos were microdissected and fractionated on a sucrose gradient. Middle: qPCR of *Hoxa9* mRNA from each fraction. Early pre-polysome fractions are highlighted. Right: quantification of *Hoxa9* mRNA in fractions. Fractions 1–8 are pre-polysome and 9–16 polysome fractions. ***P* < 0.01 (*t*-test compared to *Hoxa9*^{+/+}). *n* = 3 embryos of each genotype. **c**, Model for how ribosome-mediated regulation of gene expression is encoded within the mRNA sequence. Stress-responsive cellular IRES elements are only active when cap-dependent translation is downregulated through a decrease in eIF4F activity. *HoxA* mRNAs possess a TIE, which normally inhibits cap-dependent translation. The TIE enables ribosome-mediated control of *HoxA* IRES expression during embryonic development.

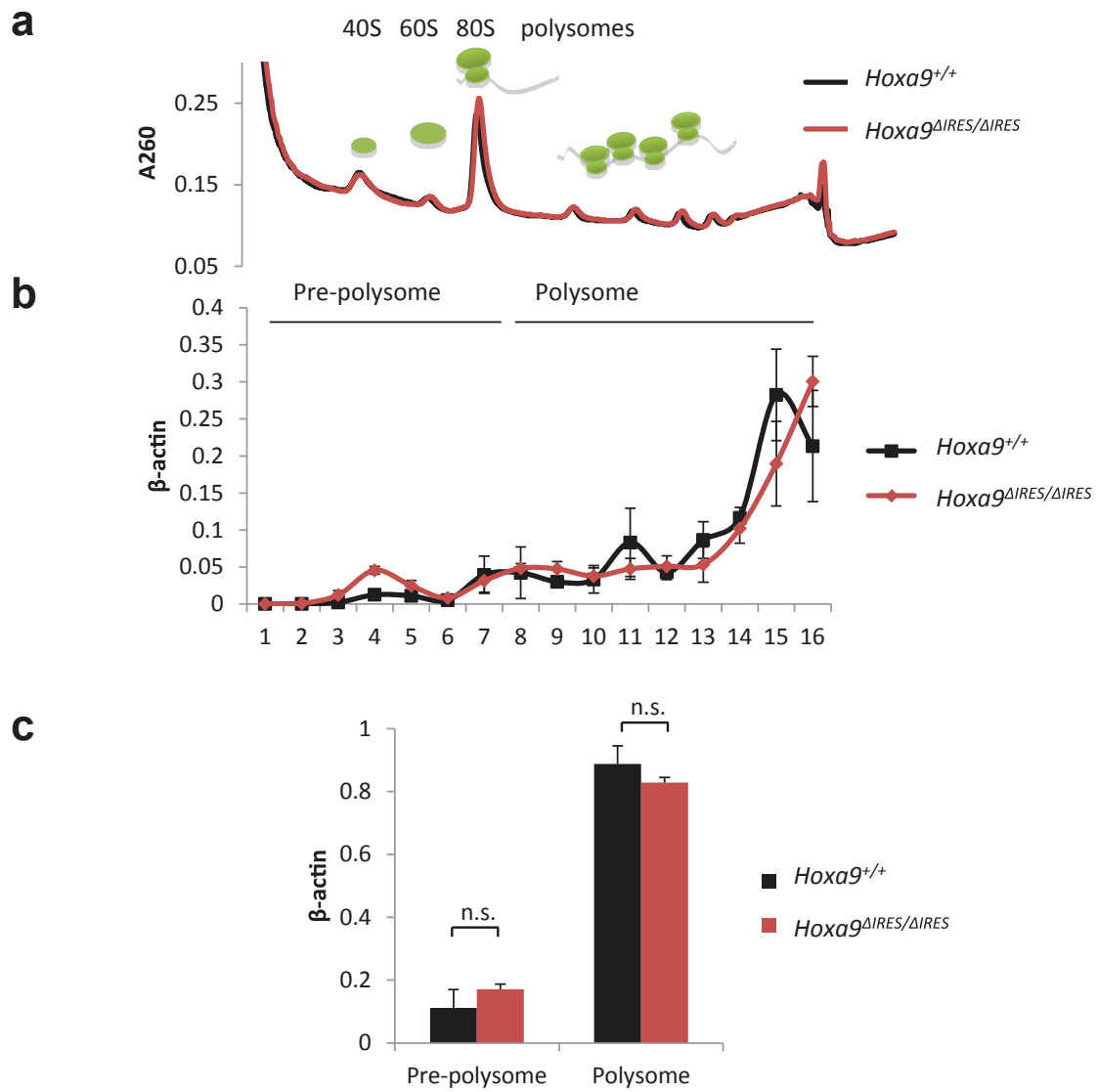


Figure 2.16 Sucrose gradient fractionation shows no difference in β -actin association with polysomes in *Hoxa9*^{+/+} and *Hoxa9* ^{Δ IRES/ Δ IRES} embryos.

a, Overlay of A₂₆₀ trace during fractionation showing no difference in polysome profiles between E11.5 *Hoxa9*^{+/+} and *Hoxa9* ^{Δ IRES/ Δ IRES} embryos. **b**, qPCR from each fraction reveals no difference in β -actin mRNA accumulation between *Hoxa9*^{+/+} and *Hoxa9* ^{Δ IRES/ Δ IRES} embryos. **c**, Quantification of β -actin mRNA in fractions. Fractions 1–8 are pre-polysomes and 9–16 are polysome fractions. *n* = 3 embryos of each genotype.

Table 2.1 Primer sequences

Name	Sequence (5' --> 3')	Description
Primers for amplifying HoxA 5'UTRs		
HoxA1 5UTR F	GATCAAGCTTGAATTCCTGCAGacattcatatcttttctctctg	used with HoxA1 5UTR R to amplify <i>Hoxa1</i>
HoxA1 5UTR R	GATCCCATGGCTGCAGctgtcactgagtgttgatc	5'UTR
HoxA2 5UTR F	GATCAAGCTTGAATTCCTGCAGaatagtttaatagtagcgcg	used with HoxA2 5UTR R to amplify <i>Hoxa2</i>
HoxA2 5UTR R	GATCCCATGGCTGCAGggccttctctcttgagc	5'UTR
HoxA3 5UTR F	CACCAAGCTTGAATTCctggagccggtctg	used with HoxA3 5UTR R to amplify <i>Hoxa3</i>
HoxA3 5UTR R	GATCCCATGGcgcatgtttcacgatcgtg	5'UTR
A4 5UTR F5	CACCAAGCTTGAATTCagttgaaaagaccgattgacaattag	used with HoxA4 5UTR R to amplify <i>Hoxa4</i>
HoxA4 5UTR R	GATCCCATGGCTGCAGtaattgtgaagtcaaaaatact	5'UTR
HoxA5 5UTR F	GATCAAGCTTGAATTCCTGCActccaccaactccccattag	used with HoxA5 5UTR R to amplify <i>Hoxa5</i>
HoxA5 5UTR R	GATCCCATGGCTGCAGgtgtttttgatatgtgtgttgattg	5'UTR
HoxA6 5UTR F	GATCAAGCTTGAATTCCTGCAttcggccatccagaaaacaccag	used with HoxA6 5UTR R to amplify <i>Hoxa6</i>
HoxA6 5UTR R	GATCCCATGGCTGCAGttgcgccccctctgca	5'UTR
HoxA7 5UTR F	GATCAAGCTTGAATTCCTGCAGtgcttagaaagggaagc	used with HoxA7 5UTR R to amplify <i>Hoxa7</i>
HoxA7 5UTR R	GATCCCATGGCTGCAGaatgttgacctgtgtgtgtg	5'UTR
HoxA9 5UTR F	CACCAAGCTTGAATTCgaaaaacagaaagggaaggataccag	used with HoxA9 5UTR R to amplify <i>Hoxa9</i>
HoxA9 5UTR R	GATCCCATGGCTGCAGtgtagtgcgcgcgcct	5'UTR
A11 5'UTR F	GATCAAGCTTGAATTCGCTAGCtatcctcaccaccggggag	used with HoxA11 5UTR R to amplify
HoxA11 5UTR R	GATCCCATGGGCTAGCtgggctaccttgggctc	<i>Hoxa11</i> 5'UTR
HoxA13 5UTR F2	CACCAAGCTTGAATTCCTGCAGgagccccccccctcg	used with HoxA13 5UTR R to amplify
HoxA13 5UTR R2	GATCCCATGGCTGCAGagccccgagccgatggag	<i>Hoxa13</i> 5'UTR
HoxA10 5UTR	GAT CAA GCT TGA ATT CCT GCA AAG AAA TGT CAG CCA GAA AGG GCT ATC TGC TCC CTT CGC CAA ATT ATC CCA CAA CAC TGC AGC CAT GGG ATC	synthesized double stranded oligos
Primers for deletion analysis of Hoxa9 5'UTR		
HoxA9 342 R	GATCCCATGGcgccagatcaccccttgct	used with HoxA9 5UTR F for 1-343
HoxA9 343 F	GATCAAGCTTGAATTCatgtcaagcgccctggctg	used with HoxA9 5UTR R for 343-1267
HoxA9 643 R	GATCCCATGGaacagattgttgcctcgtg	used with HoxA9 343 F for 343-643

HoxA9 943 R	GATCCCATGGatctggccttgctctgtac	used with HoxA9 5'UTR F for 1-944
HoxA9 944 F	GATCAAGCTTGAATTCtgatctttaatctctgtggcca	used with HoxA9 5'UTR R for 944-1267
Primers for zebrafish Hoxa9b 5'UTR		
zHoxa9b 5' EcoRI F1	CACCAAGCTTGAAATTCccttcataaaagctatgactctac	used with zHoxa9b 5'EcoRI F1 for zebrafish
zHoxa9b 5' NcoI R	GATCCCATGGaacaacaagaataagagagagcag	<i>Hoxa9b</i> 5'UTR
Primers for RNA pull down constructs		
HoxA9_F_T7	TTCTAATACGACTCACTATAGGGAATACATTAAAGAAAAACAGAAAGAGG GAAGGATACCA	used with HoxA9_R_T2 to make template for A9
HoxA9_R_T2	GTTGTTGTTGTTGTTCTTTGTTTTGTTTGCCATACTCTGGCTGCAGTAG CCCGCG	used with HoxA9_F_T7 to make template for A9ΔIRES
HoxA9C_R_T2	GTTGTTGTTGTTGTTCTTTGTTTTGTTTGCCATACTCTGGCatctggccttg cctctgtac	used with HoxA9_F_T7 to make template for A9ΔIRES
HoxA9D_F_T7	TTCTAATACGACTCACTATAGGGaatacataatctaatacatgtagttgTGATCTT TTAATCTTCGTTGGCCA	used with HoxA9_R_T2 to make template for A9 IRES
HoxA9_R_T7	TTCTAATACGACTCACTATAGGGAATACATTAAatgcagtagcccgcgct GTTGTTGTTGTTGTTCTTTGTTTTGTTTGCCATACTCTGGCgaaaaaacag	used with HoxA9_R_T7 to make template for A9 AS
A9 F T2	aagagggaaggataccag	used with A9 F T2 to make template for A9 AS
Primers for deletion analysis of Hoxa3 5'UTR		
A3 287 R	GATCCCATGGaccgtaaggagcacatTgg	used with HoxA3 5'UTR F for A3 1-287
A3 223 R	GATCCCATGGcgcgcaaacctccaagatg	used with HoxA3 5'UTR F for A3 1-223
A3 170 R	GATCCCATGGcaggtgctcagggaatcttttttcc	used with HoxA3 5'UTR F for A3 1-170
A3 120 R	GATCCCATGGaggagccaTtttcagagttttgag	used with HoxA3 5'UTR F for A3 1-120
A3 56 F	CACCAAGCTTGAAATTCgtgtcaatttaggtccagaagtgtcc	used with HoxA3 5'UTR R for A3 56-333
A3 121 F	CACCAAGCTTGAAATTCcAtgttaagagtcgctggacac	used with HoxA3 5'UTR R for A3 121-333
A3 171 F	CACCAAGCTTGAAATTCgaactggagaccattcctgtgttc	used with HoxA3 5'UTR R for A3 171-333
A3 233 F	CACCAAGCTTGAAATTCtgcgcgaggagtgtcac	used with HoxA3 5'UTR R for A3 233-333

Primers for deletion analysis of Hoxa4 5'UTR

A4 485 R	GATCCCATGGgcctacggaccattttctcg	used with A4 5UTR F5 for A4 1-485
A4 853 R	GATCCCATGGcattggcggatcgacgac	used with A4 5UTR F5 for A4 1-853
A4 5UTR F4	CACCAAGCTTGAAATTCaggagcggcgcggacttc	used with HoxA4 5UTR R for A4 908-1106

Primers for deletion analysis of Hoxa5 5'UTR

A5 mut F	gaggaaccgagtagcGgtcccgagtcatttc	used with A5 mut F to mutate uORF in <i>Hoxa5</i> 5'UTR
A5 mut R	gaaatgactgggacaCgtactcggttccttc	
A5 47 F	GATCAAGCTTGAAATTCatcaggcaggatttacgactgga	used with HoxA5 5UTR R for A5 47-237
A5 220 R	GATCCCATGGtgcttgatttggtgcgcg	used with HoxA5 5UTR F for A5 1-220
A5 108 R	GATCCCATGGcaaatatgggtacgacttcg	used with HoxA5 5UTR for A5 1-108
A5 109 F	GATCAAGCTTGAAATTCggtgcctacgtaggagggaac	used with HoxA5 5UTR R for A5 109-237

Primers for deletion analysis of Hoxa11 5'UTR

A11 113 R	GATCCCATGGcccggcagctctctcttg	used with HoxA11 5UTR F for 1-113
A11 216 R	GATCCCATGGcaacagatcgctcactcggtgttc	used with HoxA11 5UTR F for 1-216
A11 216 F	CACCAAGCTTGAAATTCctccctgaggtggctac	used with HoxA11 5UTR R for 216-496
A11 286 F	CACCAAGCTTGAAATTCggttaaaaaaaacccgggactagc	used with HoxA11 5UTR R for 286-496
A11 333 F	CACCAAGCTTGAAATTCcaacatcggtcacatgaccag	used with HoxA11 5UTR R for 333-496

Primers for chimeric UTRs

A3 170 Hbb R	TGCACCATgggtgctgtttgaggttgctagtgaacacagttgtgtcagaagcaaatgtcaggt	used with HoxA3 5UTR F to make A3 TIE
	gctcagggaatcttttttcc	Hbb
A4 485 Hbb R	TGCACCATgggtgctgtttgaggttgctagtgaacacagttgtgtcagaagcaaatgtggcct	
	acggaccatttttctcg	
A9 A Hbb R	TGCACCATgggtgctgtttgaggttgctagtgaacacagttgtgtcagaagcaaatgtcgcc	used with A4 5UTR F5 to make A4 TIE Hbb
	agatcaccttgct	Hbb
A11 A Hbb R	TGCACCATgggtgctgtttgaggttgctagtgaacacagttgtgtcagaagcaaatgtgcaa	used with HoxA9 5UTR F to make A9 TIE
	cagatcgtcactcggtgttc	Hbb

Primers to make Hoxa9ΔIRES targeting construct

A9 5'arm F	gcacccgcggcgctcttgtg	used with A9 5'arm R 3 for 5'arm of targeting construct
A9 5'arm R 3	GATCGTTTAAACatctggccttgctctgtactctcccttgc	
A9 3'arm F 3	GATCACTAGTAgttgttacatgaatactgcagtttcataatttcg	used with A9 3'arm R 2 for 3'arm of targeting construct
A9 3'arm R 2	GATCGATATCggtggctagtagccctggtac	

Extended Data Figure 1

P2 Fluc	GATGTTCACCTCGATATGTGC	used with P3 Rluc to check integrity of bicistronic RNA
P3 Rluc	GTTTATTGAATCGGACCCAGG	

Southern blot probes

A9 5' probe F2	cagcatcagaactggtcggtagttag	used with A9 5' probe R2 to make template for 5' Southern blot probe
A9 5' probe R2	aagaaacagcggaactgagagtag	
A9 3' probe F2	ttagcggtgaaaaattggtataaagtc	used with A9 3' probe R2 to make template for 3' Southern blot probe
A9 3' probe R2	agatacgtcagccaggatcgagtc	

HoxA4 RACE

Kota1562	TGGGCTCGATGTAGTTGGAG	used for reverse transcription
A4 5UTR R4	aactgggtcggctgagaag	used with Primer 1 to amplify 5' UTR
Primer 1	TCATACACATACGATTTAGGTGACACTATAGAGCGGCCCTGCAGGAA	
	A	used with A4 5UTR R4 to amplify 5' UTR

In situ hybridization probes

HoxA9 in situ F3	GGGACCCCTGGGTCAACC	used with HoxA9 in situ F3 to make template for in situ hybridization probe
HoxA9 in situ R3	CTTCGGTGAGGTGAGCAGC	

Chapter 3

Insights into IRES and TIE

Abstract

Previous work has shown that multiple Hox mRNAs contain internal ribosome entry sites (IRES) in their 5'UTR. These IRES elements are essential for their translation in normal development due to the presence of a translational inhibitory element (TIE), also in the 5'UTR, which inhibits cap-dependent translation. However the mechanism and regulation of the TIE and IRES are unknown. Here I examine the initiation factor requirements of both the TIE and IRES. In addition, I find a discrepancy between the Hoxa9 mRNA and protein expression, suggesting that its IRES is regulated differently in different tissues in the developing mouse embryo. Lastly I performed a screen and identified more IRES-containing Hox 5'UTRs. These findings demonstrate that IRES-dependent translation is exquisitely controlled in development and that many more developmentally regulated genes may be regulated similarly through IRES elements.

Introduction

Most cellular mRNAs depend on the cap-dependent mechanism for translation initiation. The 5' cap structure is recognized by eIF4E, the cap binding protein. eIF4E exists in a complex with the scaffold eIF4G, and RNA helicase eIF4A. This complex recruits the 43S complex to the mRNA. The 43S is comprised of the 40S subunit, the initiator tRNA with eIF2 and other initiation factors that assists in scanning and start codon recognition. The 43S scans to the start codon where the 60S subunit joins, initiation factors are released and peptide synthesis can begin (Sonenberg and Hinnebusch, 2009).

Much of our understanding of cap-independent translation comes from viruses. Viral internal ribosome entry sites (IRESs) can be classified into different types, based on their requirements for canonical translational initiation factors. They range from the Encephalomyocarditis virus (EMCV) IRES, which requires all initiation factors except for eIF4E to the Cricket Paralysis Virus (CrPV) IRES, which does not require any initiation factors (Jackson et al., 2010). Cellular IRES elements are generally thought to require more initiation factors than viral ones. For example, the c-myc IRES requires eIF4G, eIF4A, eIF3, poly A binding protein (PABP), the 43S, and other unidentified IRES trans-acting factors (ITAF) (Spriggs et al., 2009).

It has been observed that the *Drosophila* Hox IRES elements have different IRES activity in different tissues (Ye et al., 1997). Some mammalian IRES elements have also been shown to vary in activity in different tissues (Audigier et al., 2008; Créancier et al., 2000). We hypothesized that the mammalian Hox IRES elements could also be regulated differently in different tissues.

A striking finding in the previous chapter is the number of IRES elements that were found in the *HoxA* locus. Cellular IRES elements were first identified in stress conditions as their mRNAs remain actively translated under stress. Approximately 3-5% of expressed transcripts remain associated with polysomes in different types of stress (Johannes et al., 1999; Spriggs et al., 2008). It has been estimated that 10-15% of all mRNAs contain IRES elements (Spriggs et al., 2008). We have identified four IRES-containing mRNAs within eleven *HoxA* genes (36%), a much higher percentage than expected. This led us to hypothesize that IRES elements may be much more prevalent in the genome than previously expected.

In the previous chapter, we have discovered IRES elements in Hox mRNAs. We have also uncovered a novel RNA regulon, the Translation Inhibitory Element (TIE). Here we examined the initiation factor requirements of both elements. To determine if the IRES is regulated in a tissue-dependent manner, we characterized the RNA and protein levels of *Hoxa9* in different tissues in the mouse embryo. We also look beyond the *HoxA* locus to identify IRES elements in other Hox mRNAs.

Results

Factor requirements of IRES and TIE

Since there is no consensus set of initiation factors required for cellular IRES elements, we set out to dissect the initiation factor requirement for the Hox IRES. We transfected cells with the bicistronic constructs containing either the Hepatitis C Virus (HCV) IRES, or the *Hoxa9* IRES. The cells were then treated with an inhibitor of eIF4A pateamine A (PatA), or a DMSO control (Bordeleau et al., 2005). We observe a strong reduction in cap-dependent translation as measured by Rluc. HCV IRES activity is insensitive to PatA treatment as measured by Fluc activity (Fig. 3.1a). This is expected as HCV IRES does not require eIF4A (Bordeleau et al., 2005). However the *Hoxa9* IRES decreases significantly upon PatA treatment, indicating that the *Hoxa9* IRES requires eIF4A for its activity (Fig. 3.1a).

The TIE is a novel RNA regulon that we have uncovered in the previous chapter. Since it inhibits cap-dependent translation, we reasoned that overexpressing translation initiation factors required for cap-dependent translation may overcome the inhibition. Interestingly, overexpression of eIF4EI is able to partially overcome the cap inhibition

(Fig. 3.2b, c). eIF4E has been thought to be the rate-limiting factor in translational initiation (Gingras et al., 1999). It has also been shown that translation of mRNAs with structured 5'UTRs, such as ornithine decarboxylase (ODC), is enhanced by overexpression of eIF4E (Koromilas et al., 1992). It is possible that the overexpression of eIF4E can recruit other initiation factors to the 5'UTR, for example helicases eIF4A and eIF4B, that can unwind the TIE for more efficient 40S scanning. Overexpression of eIF4A was not efficient in our experiments (Fig. 3.1b) and hence more experiments will be necessary to determine if overexpression of other initiation factors can overcome the cap-dependent inhibition.

RNA and protein expression of *Hoxa9*

Previous work has suggested that there is a discrepancy between the appearance of *Hox* mRNAs and *Hox* proteins in the neural tube. Some *Hox* proteins are only detected days after the first appearance of the *Hox* mRNAs (Dasen et al., 2003). We therefore examined the RNA and protein expression of *Hoxa9* in mouse embryos.

At E11.5, *Hoxa9* mRNA is expressed strongly in the neural tube, somites and limbs, as have been previously reported (Burke et al., 1995) (Fig. 3.2a). Interestingly, the HOXA9 protein is found in the neural tube and somites, but missing in both the fore and hindlimbs (Fig. 3.3b). The lack of protein in the limb is also seen at E13.5 (Fig. 3.3c). *Hoxa9*^{-/-} mice do not show any phenotype in the limbs (Fromental-Ramain et al., 1996) and hence does not contradict the absence of HOXA9 protein in wildtype limbs. To see if the lack of HOXA9 in the limb is caused by a decrease in its translation, we isolated limb buds from E11.5 mouse embryos and fractionated the cell lysate on a sucrose

gradient. Compared to E11.5 neural tube and somites, *Hoxa9* mRNA in the limb is accumulated in the early pre-polysomal fractions, indicating that it is not translated as efficiently. Therefore the lack of HOXA9 protein in the limb is at least partially due to a decrease in its translation. Together, these experiments suggest that the *Hoxa9* IRES is less active in the limb than in the neural tube and somites.

Prevalence of IRES in *Hox* loci

To test if IRES elements are more prevalent in the genome, we have extended the IRES analysis to all *Hox* genes. We identified 4 additional IRES-containing mRNAs. This includes *Hoxd10*, a target of RPL38 (Fig. 3.1a). The *Hox* genes show a correlation between the length of the 5'UTR and IRES activity (Fig. 3.1b) and between delta G of the 5'UTR and IRES activity (Fig. 3.1c).

There is an enrichment of IRES-containing mRNAs in the ones which are regulated by RPL38. Five out of eight RPL38-regulated *Hox* genes contain IRES elements (63%) (Table 3.1). This is a much higher percentage than in the non-RPL38 targets (15%), suggesting that identifying RPL38 targets may be an effective way of identifying IRES-containing mRNAs. In the *HoxA* cluster, RPL38 regulates translation by regulating IRES-dependent translation. It is therefore puzzling that some RPL38 targets do not contain IRES elements. How does RPL38 regulate the translation of these mRNAs? One explanation is that 5'UTRs are poorly annotated. The *HoxA* IRES-containing 5'UTRs range from 237 to 1266 nucleotides. *Hoxa4* has an annotated 5'UTR of 60 nucleotides and we performed 5'RACE to identify its 1105-nucleotide, IRES-containing 5'UTR. The IRES-negative RPL38 targets have annotated 5'UTRs between

128-480 nucleotides. Hence it is possible that 5'RACE may yield longer 5'UTRs which contain IRES elements. Alternatively, RPL38 may regulate the translation of these genes by a yet unknown mechanism.

Hox genes are highly conserved genes in the genome. However, it is interesting to note that most paralogous groups only have one IRES-containing member. For example, *Hoxa11* has an IRES element but its paralogs, *Hoxc11* and *Hoxd11* do not (Fig. 3.1). While their coding sequences are largely similar, their 5'UTRs are rather different. This shows that the UTRs may have evolved under different selection pressures than from the coding sequence. A tempting hypothesis is that the UTRs may have been evolving to be different from other paralogs to have a different protein expression pattern.

Together we have identified more cellular IRES elements in the *Hox* cluster. 26% of all tested *Hox* genes contain IRES elements, much higher than the previous estimates. This suggests that many IRES elements may be more prevalent in cellular mRNAs than previously thought.

Discussion

We have shown that the *Hoxa9* IRES requires eIF4A for its activity. It will be interesting to determine if other initiation factors are necessary. In particular, it is useful to identify the minimal set of factors necessary to reconstitute ribosome binding and IRES activity *in vitro*. This can open many more possibilities in determining the biochemical properties of the IRES, such as how the IRES interacts with the ribosome.

The *Hoxa9* mRNA is found in limbs of embryos but not its protein, suggesting that the IRES may be differentially regulated in different tissues. A caveat to the experiment is that we are measuring RNA and protein amounts but not IRES activity directly. Nevertheless we observe a decrease in polysome association of *Hoxa9* mRNA in the limb, suggesting that the absence of HOXA9 protein is at least due in part to a decrease in translation. Future experiments will be necessary, for example by using a *Hoxa9* IRES reporter mouse, to show directly that IRES activity is lower in the limb. IRES elements found in other cellular mRNAs, such as FGF2 (Créancier et al., 2000), utrophin A (Miura et al., 2010) and c-myc (Créancier et al., 2001), have shown to be active only in some tissues, giving strong evidence that IRES activity can be regulated between tissues.

We have shown in the previous chapter that the *Hoxa9* IRES functions as an RNA structure (Figure 2.2). The IRES structure may be different in limbs versus neural tube, causing differences in IRES activity. Recent RNA structural studies show that RNA structures are much more dynamic *in vivo* than *in vitro* (Rouskin et al., 2014), and therefore IRES structures may only be stable in some tissues. Furthermore, cellular IRES elements often require ITAFs for stabilizing their structure or to recruit the ribosome. Hence differential expression of ITAFs may cause tissue-specific activation of the IRES.

So far, there has been no sequence or structural homology found between cellular IRES elements and it has been difficult to predict for cellular IRESs. We have discovered IRES elements in nine out of 34 tested *Hox* mRNAs (26%). This is much higher than previous estimates of 10-15% across the genome. IRES elements may be

enriched in the *Hox* loci, for more regulation in this set of master regulators. It is also likely that IRES elements are much more prevalent than previously thought, and a bigger search for IRES elements outside the *Hox* clusters will be highly informative. With the discovery of more cellular IRES elements, it may be possible to compare them for similar functional sequences and structures to assist in IRES prediction.

Materials and Methods

Plasmids

Hox 5' UTRs were amplified from mouse E10.5–12.5 cDNA and cloned into the bicistronic pRF vector using EcoRI and NcoI sites. Sequences were based off the Ensembl database. See Table 1 for a list of PCR primers used. eIF4AI and eIF4EI were amplified from mouse E10.5–12.5 cDNA and cloned into pENTR Gateway vector. The entry vectors were then recombined into PBX destination vectors (Sanders et al., 2013) for expression driven by the CAG promoter. All plasmids were sequenced verified by Sanger sequencing.

Luciferase assay

C3H10T1/2 cells were transfected in 24-well plates with various pRF constructs using Lipofectamine 2000 (Invitrogen) as per manufacturer instructions. Cells were collected 24 h post-transfection and assayed using Dual luciferase kit (Promega). IRES activity is expressed as a ratio between Fluc and Rluc. For assays involving overexpression of initiation factors, plasmids containing eIF4AI or eIF4EI were co-transfected with the full length *Hoxa9* or *A9* TIE monocistronic reporters in a 1:1 ratio into 12 well plates. Cells were collected 24 h later. 50% of the cells were used for qPCR analysis with Fluc primers to quantify total transcript abundance, the rest for luciferase assays. Fluc activity was normalized to total luciferase mRNA expression levels. For eIF4A inhibition studies, cells were first transfected with bicistronic pRF constructs. 24h post-transfection, the cells were treated with 50nM Pateamine A (a kind gift from J. Pelletier) for 4h. Cells were harvested for luciferase assays.

Antibodies

Western blots were performed with antibodies against eIF4E (BD 610269), eIF4A (Cell Signaling #2013) and B-actin (Cell Signaling #3700). Immunofluorescence was performed as before (Chapter 2) using antibodies against Hoxa9 (Abcam ab140631) and E-cadherin (Invitrogen 13-1900). Alexa 488 anti-rabbit antibodies (Invitrogen A-11008) and Alexa 568 anti-rat antibodies (Invitrogen A-11077) were used as secondary antibodies.

Sucrose gradient fractionation

E11.5 neural tubes and somites were dissected as before. Lysate from limbs of 3 E11.5 embryos or 1 neural tube and somites were loaded onto each sucrose gradient as previously described (Chapter 2). Fractions were collected, RNA extracted and qPCR performed as Chapter 2.

References

- Audigier, S., Guiramand, J., Prado-Lourenco, L., Conte, C., Gonzalez-Herrera, I.G., Cohen-Solal, C., Récasens, M., and Prats, A.-C. (2008). Potent activation of FGF-2 IRES-dependent mechanism of translation during brain development. *RNA* **14**, 1852–1864.
- Bordeleau, M.-E., Matthews, J., Wojnar, J.M., Lindqvist, L., Novac, O., Jankowsky, E., Sonenberg, N., Northcote, P., Teesdale-Spittle, P., and Pelletier, J. (2005). Stimulation of mammalian translation initiation factor eIF4A activity by a small molecule inhibitor of eukaryotic translation. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 10460–10465.
- Burke, A.C., Nelson, C.E., Morgan, B.A., and Tabin, C. (1995). Hox genes and the evolution of vertebrate axial morphology. *Development* **121**, 333–346.
- Créancier, L., Morello, D., Mercier, P., and Prats, A.C. (2000). Fibroblast growth factor 2 internal ribosome entry site (IRES) activity ex vivo and in transgenic mice reveals a stringent tissue-specific regulation. *J. Cell Biol.* **150**, 275–281.
- Créancier, L., Mercier, P., Prats, A., and Morello, D. (2001). c- myc Internal Ribosome Entry Site Activity Is Developmentally Controlled and Subjected to a Strong Translational Repression in Adult Transgenic Mice. *Mol. Cell. Biol.* **21**, 1833–1840.
- Dasen, J.S., Liu, J.-P., and Jessell, T.M. (2003). Motor neuron columnar fate imposed by sequential phases of Hox-c activity. *Nature* **425**, 926–933.
- Fromental-Ramain, C., Warot, X., Lakkaraju, S., Favier, B., Haack, H., Birling, C., Dierich, A., Dollé, P., and Chambon, P. (1996). Specific and redundant functions of the paralogous Hoxa-9 and Hoxd-9 genes in forelimb and axial skeleton patterning. *Development* **122**, 461–472.
- Gingras, A., Raught, B., and Sonenberg, N. (1999). eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu. Rev. Biochem.* **68**, 913–963.
- Jackson, R.J., Hellen, C.U.T., and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.* **11**, 113–127.
- Johannes, G., Carter, M.S., Eisen, M.B., Brown, P.O., and Sarnow, P. (1999). Identification of eukaryotic mRNAs that are translated at reduced cap binding complex eIF4F concentrations using a cDNA microarray. *Proc. Natl. Acad. Sci.* **96**, 13118–13123.

Koromilas, A.E., Lazaris-Karatzas, A., and Sonenberg, N. (1992). mRNAs containing extensive secondary structure in their 5' non-coding region translate efficiently in cells overexpressing initiation factor eIF-4E. *EMBO J.* **11**, 4153–4158.

Miura, P., Coriati, A., Bélanger, G., De Repentigny, Y., Lee, J., Kothary, R., Holcik, M., and Jasmin, B.J. (2010). The utrophin A 5'-UTR drives cap-independent translation exclusively in skeletal muscles of transgenic mice and interacts with eEF1A2. *Hum. Mol. Genet.* **19**, 1211–1220.

Rouskin, S., Zubradt, M., Washietl, S., Kellis, M., and Weissman, J.S. (2014). Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature* **505**, 701–705.

Sanders, T.A., Llagostera, E., and Barna, M. (2013). Specialized filopodia direct long-range transport of SHH during vertebrate tissue patterning. *Nature* **497**, 628–632.

Sonenberg, N., and Hinnebusch, A.G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* **136**, 731–745.

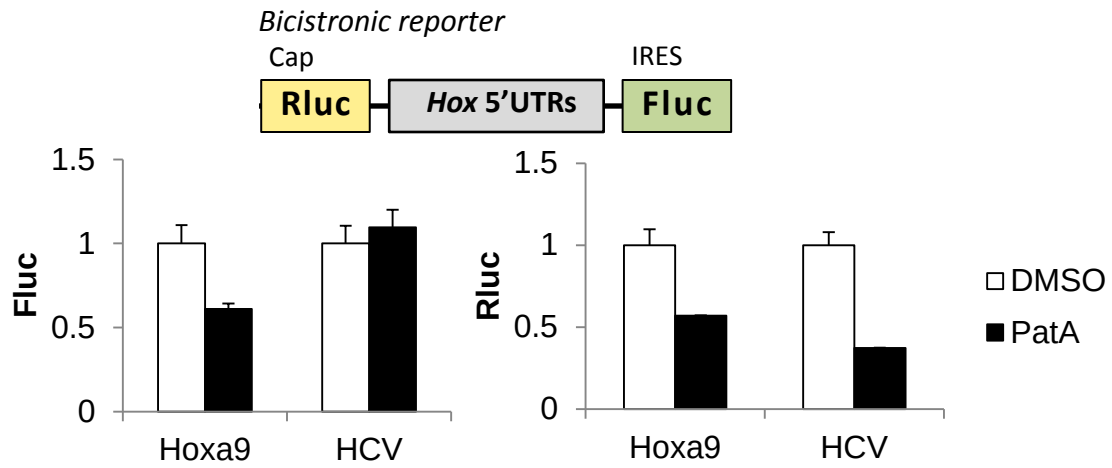
Spriggs, K.A., Stoneley, M., Bushell, M., and Willis, A.E. (2008). Re-programming of translation following cell stress allows IRES-mediated translation to predominate. *Biol. Cell* **100**, 27–38.

Spriggs, K.A., Cobbold, L.C., Jopling, C.L., Cooper, R.E., Wilson, L.A., Stoneley, M., Coldwell, M.J., Poncet, D., Shen, Y.-C., Morley, S.J., et al. (2009). Canonical initiation factor requirements of the Myc family of internal ribosome entry segments. *Mol. Cell. Biol.* **29**, 1565–1574.

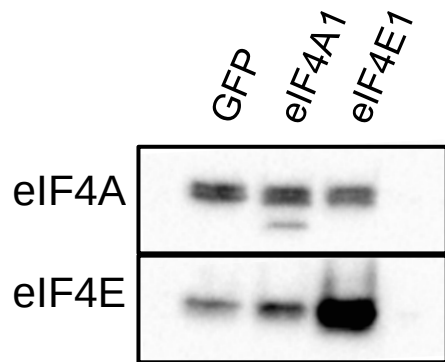
Ye, X., Fong, P., Iizuka, N., Choate, D., and Cavener, D.R. (1997). Ultrabithorax and Antennapedia 5' untranslated regions promote developmentally regulated internal translation initiation. *Mol. Cell. Biol.* **17**, 1714–1721.

Figures

a



b



c

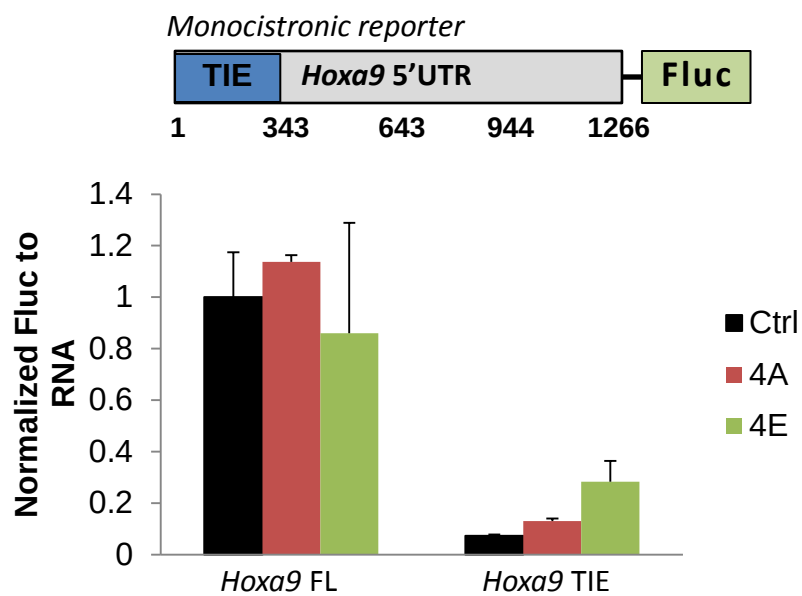


Figure 3.1 Factor requirements of IRES and TIE

a, *Hoxa9* IRES requires eIF4A for activity. Cells expressing IRES bicistronic reporters were treated with 50nM Pat A to inhibit eIF4A activity. Cap-dependent translation measured by Rluc and IRES-dependent translation by Fluc. **b**, Western blot showing overexpression of eIF4E1. eIF4A1 was not overexpressed to a significant degree. **c**, Cap inhibition by the *Hoxa9* TIE can be partially overcome by overexpression of eIF4E. Monocistronic reporters carrying the full length *Hoxa9* 5'UTR or just the TIE were used. n = 2 experiments performed in duplicates.

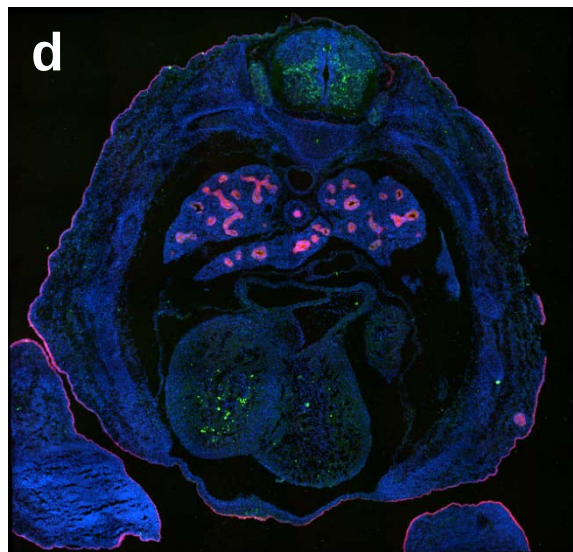
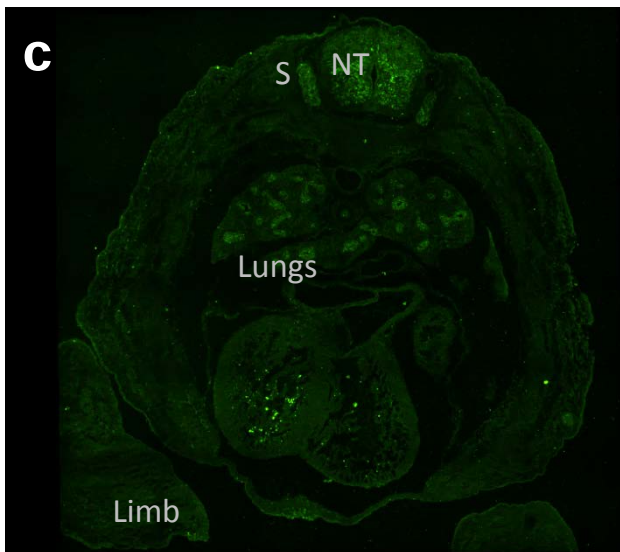
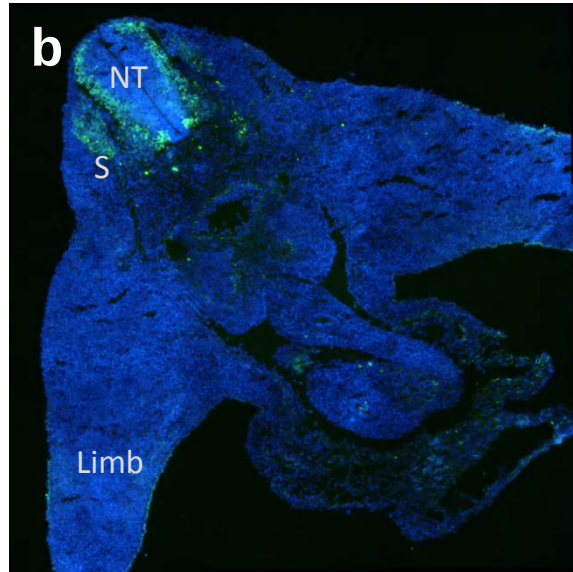
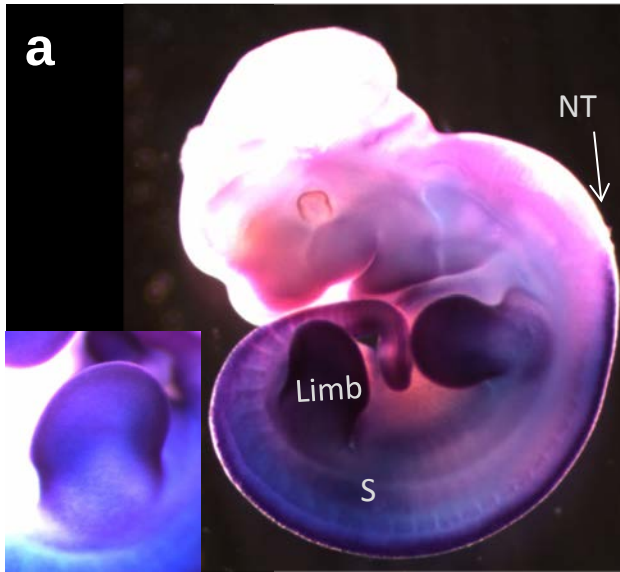
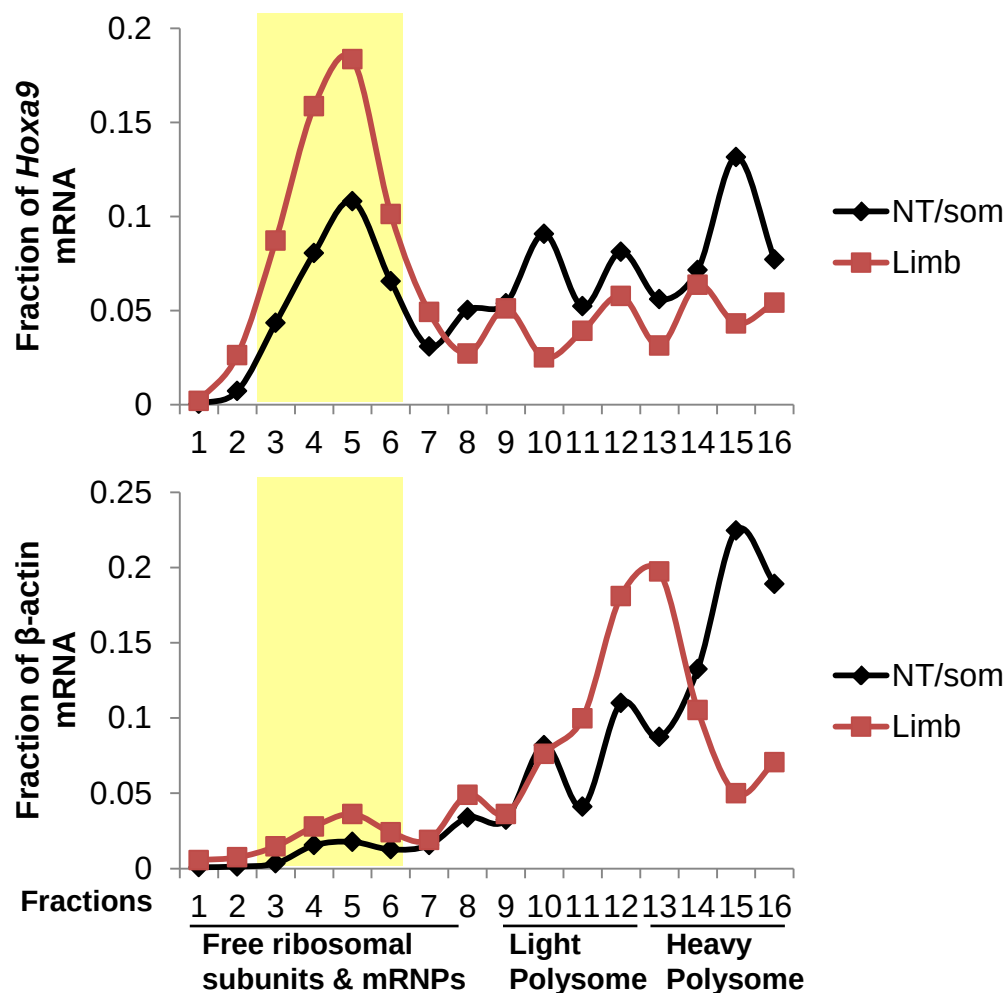


Figure 3.2 *Hoxa9* mRNA but not protein is expressed in the limb

a, Representative whole mount in situ hybridization of *Hoxa9* in E11.5 embryo shows strong *Hoxa9* mRNA expression in the limbs. Inset: Higher magnification of expression in the hindlimb. **b**, Representative immunofluorescence of HOXA9 (green) in the hindlimb region cross-section of E11.5 (**b**), and forelimb region of E13.5 embryo (**c**, **d**). HOXA9 is expressed in neural tube, somites and epithelium of the lung but not in limbs. DAPI staining is in blue. **d**, Overlay of HOXA9 (green), E-cadherin (red) and DAPI (blue). E-cadherin marks epithelia. NT: neural tube; S: somite

a



b

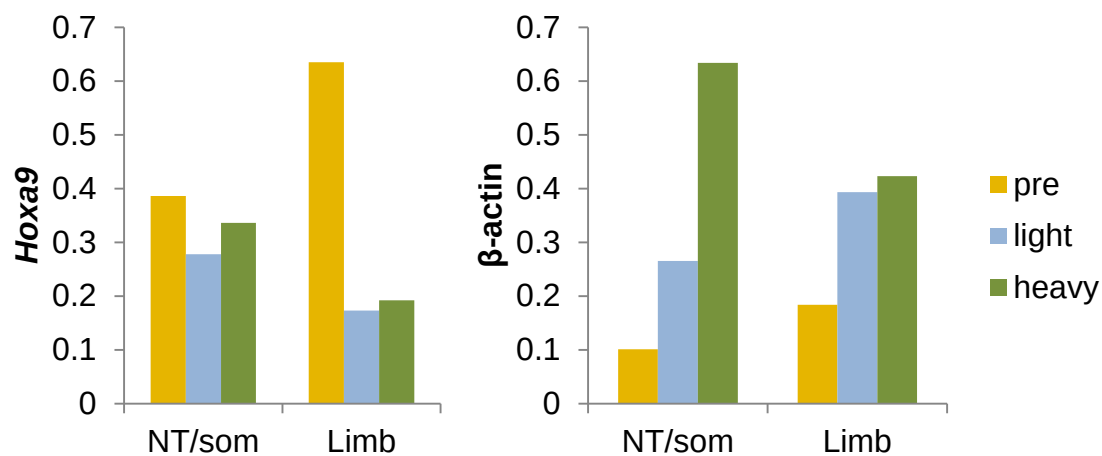


Figure 3.3 *Hoxa9* is translationally repressed in the limb

a, Limbs or somites and neural tubes of E11.5 embryos were microdissected and fractionated on a sucrose gradient. qPCR of *Hoxa9* (Top) and β -actin (Bottom) mRNA from each fraction is shown. Early pre-polysome fractions are highlighted. **b**, Quantification of *Hoxa9* and β -actin mRNA in fractions. Fractions 1-8 are pre-ribosome, 9-12 are light polysome and 13-16 are heavy polysome fractions. *Hoxa9* is accumulated in the pre-ribosomal fractions in the limb.

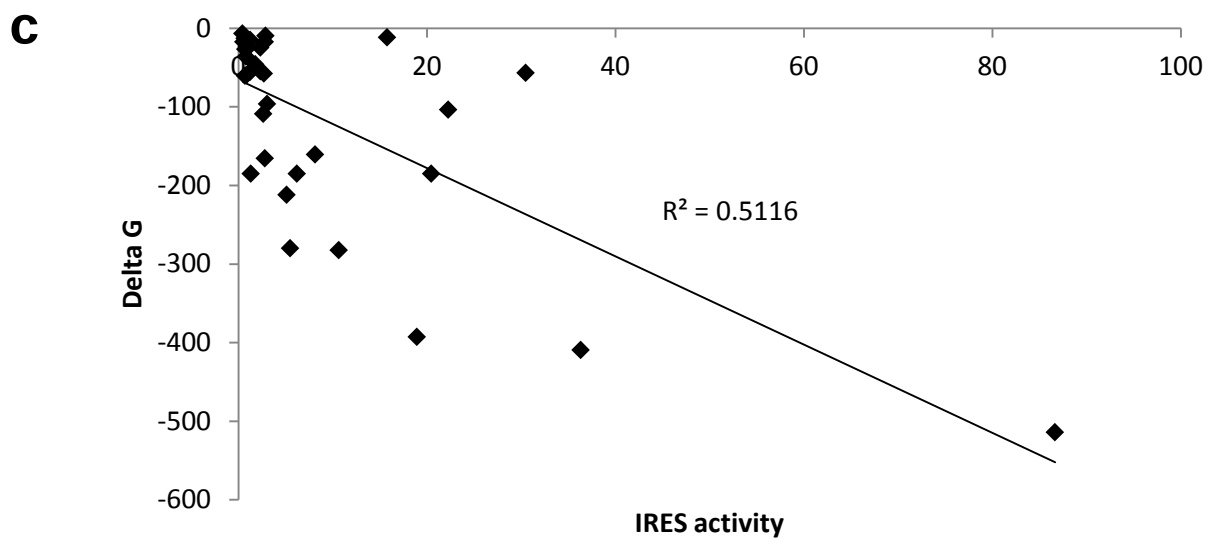
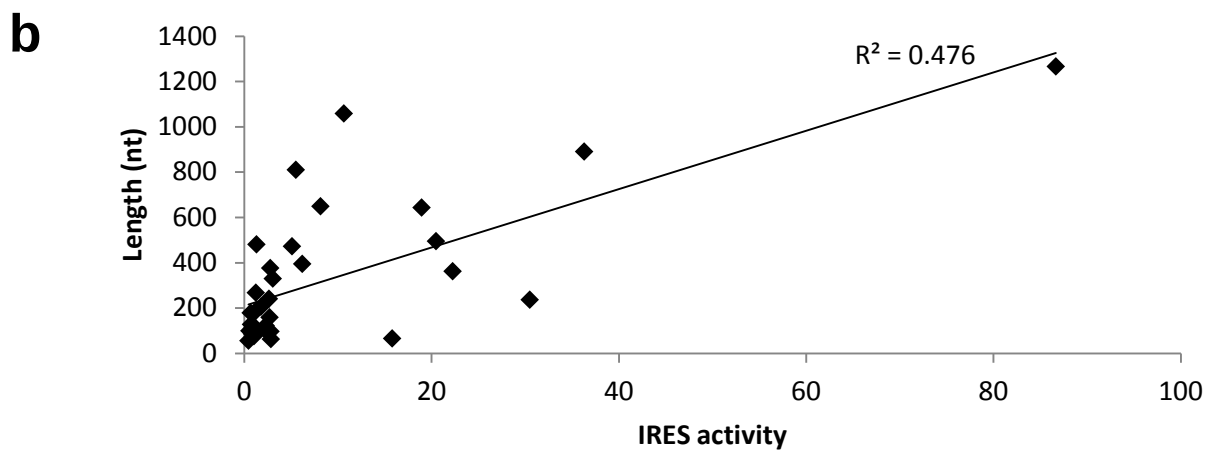
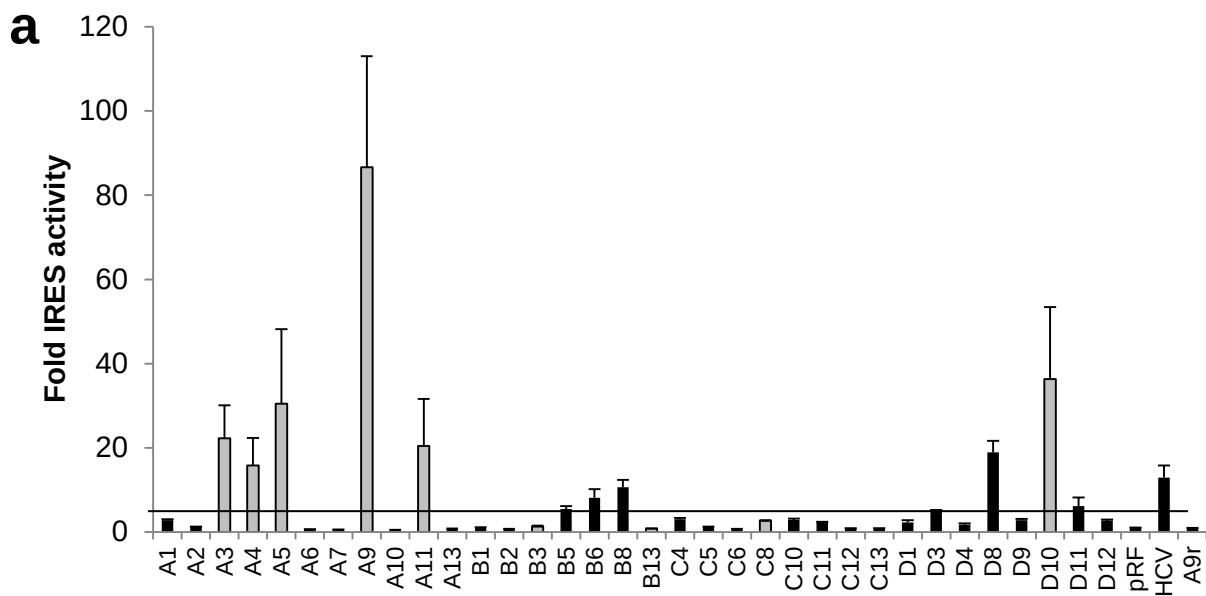


Figure 3.4 IRES activity in all *Hox* mRNAs

a, IRES activity of *HoxA* 5' UTRs. IRES activity of UTRs was normalized to that of empty vector (pRF) and RPL38-regulated *Hox* genes are denoted in grey. All UTRs with IRES activity higher than the horizontal black line (five-fold of pRF) are considered IRES positive. **b-c**, Slight correlation between length (**c**), ΔG (**d**) of 5'UTRs and IRES activity.

Table 3.1: *Hox* genes classified based on their IRES activity

	IRES positive	IRES negative
Regulated by RPL38	A5, A4, A9, A11, D10 (5)	B3, B13, C8 (3)
Not regulated by RPL38	A3, B8, D3, D8 (4)	A1, A2, A6, A7, A10, A13, B1, B2, B5, B6, C4, C5, C6, C10, C11, C12, C13, D1, D4, D9, D11, D12 (22)

Unknown: B4, B7, B9, C9, D13 (5)

Number of genes in each category is given in parenthesis.

Table 3.2 Primer sequences

Name	Sequence (5' --> 3')
Primers for amplifying Hox 5'UTRs	
HoxB1 5UTR F	GATCAAGCTTGAATTCCTGCAGgtgacatactgccgaaag
HoxB1 5UTR R	GATCCCATGGCTGCAGgtagaggccgaaggagggtc
HoxB2 5UTR F	GATCAAGCTTGAATTCCTGCAGatcagcaatatacaattataaag
HoxB2 5UTR R	GATCCCATGGCTGCAGggctttcaatgggtg
HoxB3 5UTR F2	CACCAAGCTTGAATTCCTGCAGgggtatttcgggggtgccagag
HoxB3 5UTR R	GATCCCATGGCTGCAGATcgctgggcgaggc
HoxB5 5UTR F	CACCGAATTCgtttaaaccccaaactcccgca
HoxB5 5UTR R	GATCGAATTCttgggtgattttgagggtc
HoxB6 5UTR F	CACCAAGCTTGAATTCATATGtaatgcagcattcttttgac
HoxB6 5UTR R	GATCCCATGGCATATGATcgggaggggaggcg
HoxB8 5UTR F	GATCAAGCTTGAATTCCTGCAGtaaatatgactacctcgtttg
HoxB8 5UTR R	GATCCCATGGCTGCAGtttattgaattttgaggcg
HoxB13 5UTR F	GATCAAGCTTGAATTCCTGCAGtcagcgccgttcgag
HoxB13 5UTR R	GATCCCATGGCTGCAGggagccaatgatcggtcatg
HoxC4 5UTR F	CACCAAGCTTGAATTCCTGCAGacattggagtgcagctggga
HoxC4 5UTR R	GATCCCATGGCTGCAGtaatttctggagtggaaaataat
HoxC5 5UTR F	CACCAAGCTTGAATTCgggctttgccaacatatcaagatcggttc
HoxC5 5UTR R	GATCCCATGGggcggggaggggcccaaaaaaatttc
HoxC6 5UTR F	CACCAAGCTTGAATTCATATGcagctgactttgtcattttg
HoxC6 5UTR R	GATCCCATGGCATATGccctttgcctttacctggtc
HoxC8 5UTR F	GATCAAGCTTGAATTCCTGCAGgttagaggaaagagctaag
HoxC8 5UTR R	GATCCCATGGCTGCAGgctgggtacatgaaaacc
HoxC10 5UTR F	CACCAAGCTTGAATTCCTGCAGgggaaaaaagatgtcagctc
HoxC10 5UTR R	GATCCCATGGCTGCAGtttcagagagagggttttaag
HoxC11 5UTR F	CACCAAGCTTGAATTCATATGgactaagacggataacgcgtc
HoxC11 5UTR R	GATCCCATGGCATATGcgttctcttctcctgcctc
HoxC12 5UTR F	CACCAAGCTTGAATTCCTGCAGcgtgagtagagcgccaatg
HoxC12 5UTR R	GATCCCATGGCTGCAGttccgcggggccga
HoxC13 5UTR F	CACCAAGCTTGAATTCCTGCAGggggagggaaaaagagagcga
HoxC13 5UTR R	GATCCCATGGCTGCAGgacatgatctgctccagct
HoxD1 5UTR F	CACCAAGCTTGAATTCCTGCAGatatgaagccactatttacctcg
HoxD1 5UTR R	GATCGAATTCCTGCAGggctcgctccgcacgac
HoxD3 5UTR F	CACCAAGCTTGAATTCCTGCAGgcattctgctttccagagc
HoxD3 5UTR R	GATCCCATGGCTGCAGtgactaccacggctccttctc
HoxD4 5UTR F	CACCAAGCTTGAATTCAGTAGTacttgattacacgtatgttatttag
HoxD4 5UTR R	GATCCCATGGACTAGTtaatttcagaaggtagaaaatac
HoxD8 5UTR F	CACCAAGCTTGAATTCATATGatctcgtccctctcctcgag
HoxD8 5UTR R	GATCCCATGGCATATGcggggcgggcgccgc
HoxD9 5UTR F	CACCAAGCTTGAATTCGCTAGCggagatcgaggacccagag
HoxD9 5UTR R	GATCCCATGGGCTAGCgggtgaggcgccgctcgctc

D10 5'UTR F	GATCAAGCTTGAATTCCTGCAGcgaataggggaacggctgg
HoxD10 5UTR R2	GATCCCATGGtttggggaggaatttgaagaag
HoxD11 5UTR F	CACCAAGCTTGAATTCCTGCAGggcgggcccagtcgggtg
HoxD11 5UTR R	GATCGAATTCggctgcgcggcgtgcacg
HoxD12 5UTR F	CACCAAGCTTGAATTCGCTAGCggtgggggtggggacgggtc
HoxD12 5UTR R	GATCCCATGGGCTAGCctccaacaaggcctcggctg
elF4A F	caccATGAATCTCTCAGAGTCCCTCCTC
elF4A R	gatcTCAAATGAGGTCAGCAACGTTGAG
elF4E F	caccATGGCGACTGTGGAACC
elF4E R	gatcTTAAACAACAAACCTATTTTATAGTGGTGG

Chapter 4:

Discussion and Future Directions

This chapter provides additional discussion not covered in previous chapters and follow-up questions and possible experiments.

Identifying IRESs

Known cellular IRES elements to date do not have a consensus sequence or structure and hence it has been challenging to predict for and identify cellular IRES elements. We have functionally identified a number of IRES elements in Hox genes. With these new found IRESs, one can now attempt to use common features between them, such as their RNA structures or RNA binding proteins to identify other IRES elements in the genome.

We have shown that *Hoxa9* requires its RNA structure for IRES function (Fig 2.2a-b). We have performed preliminary experiments on determining the structures of other IRES elements in the *HoxA* locus. In particular, the asymmetric bulge found in the *Hoxa9* IRES is also present in *Hoxa5* (Fig 2.?). A similar asymmetric bulge is found in domain II of the HCV IRES (Lukavsky et al., 2003), allowing the IRES to bend and adopt an L-shape conformation in solution and also when bound to the ribosome (Lukavsky et al., 2003; Spahn et al., 2001). It will be interesting to chemically probe secondary structures of other *Hox* IRES elements to see if the asymmetric bulge is conserved and could be potentially a common feature between cellular IRES elements. The chemical probing methods we have used gives nucleotide-resolution details on the secondary structures of the IRES and is great for comparing secondary structures (Deigan et al., 2009; Kladwang et al., 2011). However it is difficult to model these structures in 3D, or to predict which parts of the structures are responsible for

interactions with the ribosome. A different approach is to solve the 3D structure of the *Hoxa9* or *Hoxa5* IRES when bound to the ribosome by cryo-electron microscopy (cryoEM). This allows better mapping of regions of the IRES which are important for function and can serve as a guide to identify similar sequences or structures in other 5'UTRs.

Most cellular IRESs require ITAFs for function. We know very little about the ITAF requirements for the *Hox* IRES elements. One method of identifying ITAFs for *Hox* IRES elements is to pull down associated proteins with a *Hox* IRES and analyze them through mass spectrometry. We have performed an *in vitro* pull down using biotinylated *Hoxa9* 5'UTR and cell lysate. We were able to detect ribosomal proteins in the pull down (Fig. 2.?), but the protein yield was too low to detect other less abundant RNA binding proteins (RBP). A major limitation of our pull down was the inefficiency in elution of proteins from the RNA. We had difficulties releasing the RNA from the strong biotin-streptavidin bond and were therefore unable to collect enough proteins in solution.

There are multiple new ways of pulling down RBP without using biotinylated RNA. One approach is to tag the *Hoxa9* RNA with a small 16 nucleotide tag that can be recognized by an inactive version of a CRISPR nuclease, Csy4 (Lee et al., 2013). The RNA could be transfected into the cell or incubated with cell lysate for association with ITAFs. The RNA can then be pulled down using an inactive biotinylated Csy4 bound to streptavidin beads. Csy4 can be activated by the addition of imidazole, which changes the conformation of the inactive Csy4 and allows it to cut the RNA. This releases the RNA and the associated proteins from the beads. The eluted proteins can then be identified by mass spectrometry.

A more physiological approach will be to pull out the *Hox* mRNA and associated ITAFs from the embryo. Wildtype or *Hoxa9*^{ΔIRES/ΔIRES} embryos can be dissociated into individual cells then crosslinked with formaldehyde to preserve RNA-protein interactions (Selth et al., 2009). The *Hoxa9* mRNA and associated proteins can then be pulled out using tiled biotinylated DNA probes against the *Hoxa9* mRNA. Using multiple probes will ensure that some probes will be able to bind to an unstructured region of the mRNA. Non-specific binding can also be eliminated by splitting the probes into two groups and comparing the proteins pulled down in both groups. Only proteins common to both groups are considered specific (Chu et al., 2011). The proteins can then be eluted either by reversing the crosslink or by treating with RNase. This approach could give more physiological results since it does not involve overexpression of the IRES RNA or forming protein-RNA interactions *in vitro*. After identifying ITAFs using either the *in vitro* or *in vivo* method, the opposite cross-linking and immunoprecipitation (CLIP) experiment can be performed. The ITAF could be immunoprecipitated to determine what other RNAs it binds to, and this could be an efficient way of identifying more IRES elements.

TIE and mechanism

An interesting discovery we have made is the presence of a TIE in many IRES-containing mRNAs. The TIE inhibits cap-dependent translation, allowing IRES-dependent translation to be the predominant way of translating these genes. Previous work has found elements in 5'UTRs that inhibit translation, such as uORFs and strong hairpins. These elements are thought to inhibit with 40S scanning along the 5'UTR.

Some of these elements have been shown to be regulated. The most famous example of uORFs is in GCN4. GCN4 is a transcriptional activator of amino acid biosynthesis in yeast and contains 4 uORFs in its 5'UTR. The uORFs act as an amino acid sensor and their translation affects whether GCN4 gets translated (Miller and Hinnebusch, 1990). Strong hairpins have also been found to inhibit translation. The 5'UTRs of ODC (Manzella and Blackshear, 1990) and TGFB1 (Jenkins et al., 2010) for example, contain high GC contents and form stable hairpins that inhibit cap-dependent translation.

The TIEs that we have discovered do not have high GC contents and their mechanism of cap inhibition is unknown. These mechanisms could involve blocking the nuclear export of mRNAs, 40S recruitment or 40S scanning. To determine if nuclear export is affected, RNA fluorescence in situ hybridization (FISH) can be used to look at the localization of the *Hoxa9* mRNA in wildtype versus *Hoxa9* ^{Δ IRES/ Δ IRES} embryos. This could also give information on whether the Δ IRES mRNA is sequestered into a particular cytoplasmic compartment, for example in P bodies. An alternative method is to pull down the *Hox* mRNA and determine which cap-binding complex is on the mRNA. The nuclear cap-binding complex is loaded onto the mRNA cap co-transcriptionally and helps export the mRNA into the cytoplasm. Once there, the eIF4F complex takes over as the cap binding complex in the cytoplasm (Topisirovic et al., 2011). Determining which cap-binding complex is on the Δ IRES mRNA can help in understanding whether the mRNA has a defect in nuclear export.

To distinguish between a defect in 40S recruitment and 40S scanning, one could test in an *in vitro* translation system. Using the beta-globin (*Hbb*) or a TIE-*Hbb* chimeric mRNA, one can determine if the eIF4F complex can bind to the TIE-*Hbb* mRNA. If the

TIE blocks 40S recruitment, for example by physically occluding the cap from eIF4F, we will see a dramatic decrease in the amount of eIF4E, eIF4G and eIF4A bound to the TIE-*Hbb* mRNA. We can also treat the *in vitro* translation with a GTP analog such as GMP-PNP to isolate 48S complexes. 48S consists of the 40S subunit, the initiator tRNA and the mRNA. 48S complexes can be isolated on a sucrose gradient in the presence of a GTP inhibitor. Association of the TIE-*Hbb* and the *Hbb* mRNA with the 48S complex can be examined to determine if there is a defect in 40S recruitment. Together these experiments will give much more mechanistic insights on the TIE.

Function of RPL38

It was a surprising discovery that RPL38, a single ribosomal protein, can regulate the translation of specific mRNAs. I show in Chapter 2 that RPL38 regulates the IRES-dependent translation of specific *Hox* mRNAs. However how RPL38 is able to do that is still unknown. Is RPL38 able to bind to the IRES directly and recruit the ribosome to the mRNA? Does it work through intermediary proteins or RNA? What are other targets of RPL38 besides *Hox* genes? These are questions that await further investigation.

RPL38 is a ribosomal protein on the large subunit of the ribosome and it functions exclusively on the ribosome. One of the biggest mysteries is how a large subunit ribosomal protein is involved in IRES dependent translation. Most viral IRESs recruit the small subunit, and in cap-dependent translation, the large subunit does not encounter the mRNA until the small subunit reaches the start codon. Hence the large subunit has been mainly thought to have roles in elongation, as opposed to initiation. As mentioned earlier, the *Hox* IRESs contain an asymmetric bulge, similar to the one found in Domain

II of HCV. Domain II in HCV has been implicated in 80S formation, in particular in the hydrolysis and release of eIF2-GTP, the last step before subunit joining (Locker et al., 2007). Previous work has shown that in *RPL38^{Ts/+}* embryos, the translation defect takes place at or before subunit joining (Kondrashov et al., 2011). It will be interesting to determine if RPL38 is required for interacting with the *Hox* IRES in helping with eIF2 release and subunit joining.

RPL38 is located near the base of a large rRNA expansion segment, ES27. Expansion segments are regions of the rRNA that have increased in size tremendously from bacteria to eukaryotes. Although the expansion segments in metazoans are so extensive that they cover the ribosome as a layer of RNA, their functions are largely unknown (Anger et al., 2013). ES27 is one of the largest expansion segments and it is ~600nt in the mouse. It is also highly dynamic and has been shown to rotate ~110° about its base from the L1 stalk to the tunnel exit. RPL38 has been thought to stabilize ES27 in one of the positions in the rotation (Armache et al., 2010). RPL38 is a small protein of 8kD. It is possible that its actions are manifested through the much larger ES27. Deletion of ES27 in *Tetrahymena* is lethal, suggesting its essential function in ribosomes (Sweeney et al., 1994). One experiment would be to stop the movement of ES27 or to shorten it to see if it has an effect on translation of *Hox* IRESs.

We initially focused our attention on *Hox* genes in the *RPL38^{Ts/+}* embryos because of the strong homeotic transformation phenotype in the axial skeleton (Kondrashov et al., 2011). *RPL38^{Ts/+}* embryos have other phenotypes, for example cleft palates and various eye abnormalities. *RPL38^{Ts/+}* adult mice have also shown to suffer hearing loss (Noben-Trauth and Latoche, 2011). It will be interesting to determine how

much of these phenotypes are due to defects in translation of specific mRNAs versus non-translation effects of RPL38. One way is to do ribosome profiling in the *RPL38*^{Ts/+} embryos to discover other genes RPL38 regulates. *RPL38*^{Ts/Ts} embryos die at E5.5, making studies of complete loss of RPL38 difficult in older embryos (Kondrashov et al., 2011; Paterson, 1980). Rpl38 is dispensable in yeast (Steffen et al., 2012), and we speculate that it is also not required for viability in the mammalian cell. Hence studying an *Rpl38*-null tissue could be possible with a conditional deletion of *Rpl38* in specific tissues. *Rpl38* shows tissue-specific expression and it is possible that some tissues may require RPL38 more than others. It is therefore worthwhile to make a conditional *Rpl38* mouse to get at its function in different tissues.

Post-translational regulation of Hox mRNAs

Hox genes are expressed in a highly regulated manner. *Hox* expression boundaries are established early in embryonic development by collinear transcriptional activation. Post-transcriptional regulation mechanisms have been discovered in *Hox* genes to regulate their protein levels independent from transcription (Yekta et al., 2008). We have found that *Hox* mRNAs also undergo translational regulation. Because of their tight collinear transcriptional regulation, *Hox* genes may have been undergoing selective pressure to evolve alternative ways of regulating expression.

It has been observed in some systems that the mRNA and protein expression of *Hox* genes do not agree with each other. For example, *Hoxb4* mRNA is found at high levels in the neural tube at the forelimb level of E10.5 embryos, but the protein is absent (Brend, 2003). Similarly *Hoxc6* mRNA is found in the thoracic neural tube of chick

embryos, but not the protein (Dasen et al., 2003). Some of these differences have been attributed to miRNAs, but they could also be the result of translational regulation. Most studies on *Hox* genes use *in situ* hybridization as an assay for *Hox* expression. This is because antibodies against HOX proteins are notoriously difficult to generate due to the high homology between different HOX proteins. A careful study of expression of *Hox* mRNA and proteins may reveal many discrepancies. I have shown that in Chapter 3 that *Hoxa9* mRNA is expressed highly in the limbs but the protein is missing. One could do similar expression studies with other IRES-containing *Hox* mRNAs to see if their mRNA expression correlates with protein expression. A follow up experiment will be to generate a reporter mouse with the reporter hooked up to the IRES in question to see if the difference in protein expression is indeed caused by IRES-dependent translation. As expression of *Hox* genes are regulated by signals such as retinoic acid and fibroblast growth factors (FGFs), one could also determine if such differences between mRNA and protein expression is a response to signaling.

Conclusion

Together, I have shown here a novel way of regulating *Hox* gene expression using RNA regulons and specialized ribosomes. Historically developmental biologists have relied on *in situ* hybridization as the main way of determining gene expression. Our studies show that transcription is only the beginning of gene expression, and that protein expression may be a much more accurate depiction of where genes are functional. We also demonstrated that the ribosome is not a ubiquitous constitutive

machine, but has highly regulatory functions. This study opens avenues for future studies, including many topics in translation, RNA biology and developmental biology.

References

- Anger, A.M., Armache, J.-P., Berninghausen, O., Habeck, M., Subklewe, M., Wilson, D.N., and Beckmann, R. (2013). Structures of the human and *Drosophila* 80S ribosome. *Nature* **497**, 80–85.
- Armache, J.-P., Jarasch, A., Anger, A.M., Villa, E., Becker, T., Bhushan, S., Jossinet, F., Habeck, M., Dindar, G., Franckenberg, S., et al. (2010). Cryo-EM structure and rRNA model of a translating eukaryotic 80S ribosome at 5.5-Å resolution. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 19748–19753.
- Brend, T. (2003). Multiple levels of transcriptional and post-transcriptional regulation are required to define the domain of *Hoxb4* expression. *Development* **130**, 2717–2728.
- Chu, C., Qu, K., Zhong, F.L., Artandi, S.E., and Chang, H.Y. (2011). Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol. Cell* **44**, 667–678.
- Dasen, J.S., Liu, J.-P., and Jessell, T.M. (2003). Motor neuron columnar fate imposed by sequential phases of *Hox-c* activity. *Nature* **425**, 926–933.
- Deigan, K.E., Li, T.W., Mathews, D.H., and Weeks, K.M. (2009). Accurate SHAPE-directed RNA structure determination. *Proc. Natl. Acad. Sci.* **106**, 97–102.
- Jenkins, R.H., Bennagi, R., Martin, J., Phillips, A.O., Redman, J.E., and Fraser, D.J. (2010). A conserved stem loop motif in the 5′ untranslated region regulates transforming growth factor- β (1) translation. *PLoS One* **5**, e12283.
- Kladwang, W., Cordero, P., and Das, R. (2011). A mutate-and-map strategy accurately infers the base pairs of a 35-nucleotide model RNA. *RNA* **17**, 522–534.
- Kondrashov, N., Pusic, A., Stumpf, C.R., Shimizu, K., Hsieh, A.C., Xue, S., Ishijima, J., Shiroishi, T., and Barna, M. (2011). Ribosome-Mediated Specificity in *Hox* mRNA Translation and Vertebrate Tissue Patterning. *Cell* **145**, 383–397.
- Lee, H.Y., Haurwitz, R.E., Apffel, A., Zhou, K., Smart, B., Wenger, C.D., Laderman, S., Bruhn, L., and Doudna, J. a (2013). RNA-protein analysis using a conditional CRISPR nuclease. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 5416–5421.
- Locker, N., Easton, L.E., and Lukavsky, P.J. (2007). HCV and CSFV IRES domain II mediate eIF2 release during 80S ribosome assembly. *EMBO J.* **26**, 795–805.
- Lukavsky, P.J., Kim, I., Otto, G.A., and Puglisi, J.D. (2003). Structure of HCV IRES domain II determined by NMR. *Nat. Struct. Biol.* **10**, 1033–1038.

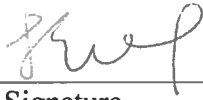
- Manzella, J.M., and Blackshear, P.J. (1990). Regulation of rat ornithine decarboxylase mRNA translation by its 5'-untranslated region. *J. Biol. Chem.* **265**, 11817–11822.
- Miller, P.F., and Hinnebusch, A.G. (1990). cis-Acting sequences involved in the translational control of GCN4 expression. *Biochim. Biophys. Acta* **1050**, 151–154.
- Noben-Trauth, K., and Latoche, J.R. (2011). Ectopic mineralization in the middle ear and chronic otitis media with effusion caused by RPL38 deficiency in the Tail-short (Ts) mouse. *J. Biol. Chem.* **286**, 3079–3093.
- Paterson, H.F. (1980). In vivo and in vitro studies on the early embryonic lethal tail-short (Ts) in the mouse. *J. Exp. Zool.* **211**, 247–256.
- Selth, L.A., Gilbert, C., and Svejstrup, J.Q. (2009). RNA Immunoprecipitation to Determine RNA-Protein Associations In Vivo Protocol. *Cold Spring Harb. Protoc.* **4**.
- Spahn, C.M., Kieft, J.S., Grassucci, R. a, Penczek, P. a, Zhou, K., Doudna, J. a, and Frank, J. (2001). Hepatitis C virus IRES RNA-induced changes in the conformation of the 40s ribosomal subunit. *Science* **291**, 1959–1962.
- Steffen, K.K., McCormick, M.A., Pham, K.M., MacKay, V.L., Delaney, J.R., Murakami, C.J., Kaeberlein, M., and Kennedy, B.K. (2012). Ribosome Deficiency Protects Against ER Stress in *Saccharomyces cerevisiae*. *Genetics* **191**, 107–118.
- Sweeney, R., Chen, L., and Yao, M.C. (1994). An rRNA variable region has an evolutionarily conserved essential role despite sequence divergence. *Mol. Cell. Biol.* **14**, 4203–4215.
- Topisirovic, I., Svitkin, Y. V, Sonenberg, N., and Shatkin, A.J. (2011). Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip. Rev. RNA* **2**, 277–298.
- Yekta, S., Tabin, C.J., and Bartel, D.P. (2008). MicroRNAs in the Hox network: an apparent link to posterior prevalence. *Nat. Rev. Genet.* **9**, 789–796.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.



Author Signature

12/31/14

Date