Functions of long non-coding RNAs in human disease and their conservation in *Drosophila* development.

Oliver M. Rogoyski[1], Jose Ignacio Pueyo[1], Juan Pablo Couso[1,2] and Sarah F. Newbury[1]*

[1] Brighton and Sussex Medical School, Medical Research building, University of Sussex, Falmer, Brighton, BN1 9PS, U.K.

[2] Centro Andaluz de Biología de Desarrollo, CSIC-Universidad Pablo de Olavide, Ctra. de Utrera, Km.1, Sevilla, 41013, Spain

*Corresponding author: s.newbury@bsms.ac.uk +44 (0)1273 877874

Abstract

Genomic analysis has found that the transcriptome in both humans and *Drosophila melanogaster* features large numbers of long non-coding RNA transcripts (lncRNAs). This recently discovered class of RNAs regulates gene expression in diverse ways, and has been involved in a large variety of important biological functions. Importantly, an increasing number of lncRNAs have also been associated with a range of human diseases, including cancer. Comparative analyses of their functions among these organisms suggest that some of their modes of action appear to be conserved. This highlights the importance of model organisms such as *Drosophila*, which shares many gene regulatory networks with humans, in understanding lncRNA function and its possible impact in human health. This review discusses some known functions and mechanisms of action of lncRNAs and their implication in human diseases, together with their functional conservation and relevance in *Drosophila* development.

Introduction

The central dogma of molecular biology as proposed by Crick in 1958, often paraphrased as "DNA encodes RNA, RNA encodes protein", implicates RNA as a molecular intermediate in the process of protein synthesis from the relevant encoding gene. As early as the 1950s however, other roles for non-coding RNAs, such as transfer RNAs and ribosomal RNAs, have been known to be vital to biological function. This showed the central dogma to be an over-simplified, if eloquent, summary of the flow of genetic information. Since then, many other types of non-coding RNA have been shown to exist, and furthermore, to be biologically relevant. In the 1990s, several studies began investigating the biological purpose of longer non protein-coding RNAs, such as *Xist* [1], which did not fit well into the RNA classifications existing at the time. With further advances in molecular techniques suggesting that only 2% of the human genome is comprised of protein-coding genes [2], and rapidly revealing lncRNAs with biological functions (including in human diseases), the topic has become an extremely promising and popular avenue of investigation.

In this review, we have used the definition of lncRNAs as being RNA transcripts longer than 200 nucleotides, which lack a significant open reading frame (greater than 100 amino acids in length) [3]. This definition is routinely used in the annotation of the *Drosophila* and other genomes. LncRNAs are highly abundant, and are found in many organisms across different taxa, including humans, mice, *Xenopus tropicalis*, *Drosophila melanogaster*, *Schizosaccharomyces pombe*, *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Arabidopsis thaliana*, *Medicago truncatula*, and *Zea mays* [4]. LncRNAs have been shown to regulate gene expression transcriptionally [5-8] and post-transcriptionally [9-13], and have a wide

range of cellular and molecular functions. Despite these proven non-coding functions, there exist a handful of lncRNAs that have been shown to encode small open reading frame (smORF) peptides with proven cellular functions [14-19]. Recent work has shown that lncRNAs can simultaneously display biological function as both a coding, and a non-coding RNA, for example where primary transcripts of microRNAs encode regulatory peptides [20, 21]. Additionally, ribosome profiling and bioinformatics analyses have identified the existence of thousands of lncRNAs containing putatively functional translated smORFs [19, 22-25], the extent of which may depend on developmental or tissue specific context. We have therefore used the accepted definition above, which coincides with genome annotations.

Drosophila melanogaster, the common fruit fly, is a well-established model organism for geneticists, and one in which lncRNAs are known to be abundant. With an estimated 75% of human disease-linked genes having a functional orthologue in *Drosophila*, and many basic molecular and biological functions conserved between species [26, 27], *Drosophila* are an appealing whole animal model for understanding human disease. In addition to their genetic similarities, the fly genome has been incredibly well studied and fully sequenced, with a wide range of genetic tools and gene-specific knockdown and mutant lines readily available. Combined with their low maintenance cost, short generation time, high fecundity, and compound factors lending themselves to ease of establishing genetic crosses, it is easy to see why *Drosophila* have emerged as one of the foremost systems for studying the genetic components of human disease, and have already been successfully used to dissect the roles and mechanisms of certain lncRNAs [28].

As well as the general excellence of *Drosophila* as a model organism, they stand out as particularly apt for the study of lncRNA. LncRNAs evolve rapidly, and can act as flexible scaffolds tethering together one or more functional elements [29]. *Drosophila* lncRNAs also appear to accumulate relatively few deleterious changes, due to genetic drift, compared to mammalian lncRNAs [30], and therefore can be useful in developing strategies to identify lncRNA orthologues, as shown for *roX* lncRNA orthologues in Drosophilid species [31]. Additionally, *Drosophila* is an excellent model system to functionally characterise lncRNA-protein complexes, for example by using the GAL4-UAS system to express lncRNAs in specific tissues or by characterising the localisation of RNA-proteins within cells (e.g. 7SK snRNA [32]).

Molecular functions and mechanisms of lncRNAs, such as their binding to protein complexes, definitively need to be tested *in vivo* in order to be well characterized. For example, *in vivo* experiments have shown that only the lncRNA transcribed in the reverse direction from the Polycomb/Trithorax response elements can bind the the Polycomb Repressive Complex 2 component Enhancer of Zeste, which provides the critical Histone Methyl Transferase activity required for transcriptional silencing. This level of understanding of such complex mechanisms and interactions would be extremely difficult to achieve without the use of a tractable *in vivo* system such as that provided by *Drosophila*.

In this review, we will be examining the emerging roles and relevance of lncRNAs using recent work illustrating their biological and molecular functions in *Drosophila*. We aim to examine these recent advances in our understanding of lncRNAs through the lens of their potential relevance to humans, and particularly human disease. By doing so, we hope to provide a concise synopsis of the topic, and demonstrate the value of using *Drosophila* as a model organism for understanding the roles of lncRNAs at molecular and cellular levels, and their implications in human disease.

Abundance and localisation of lncRNAs in the human and *Drosophila* genomes

According to the Ensembl database, lncRNAs comprise 7841 of the 63898 annotated genes in the human genome, and 2366 of the 17559 in the *Drosophila* genome. In both species, they account for a similar and substantial proportion of the entire genome (12.4% and 13.5% respectively). Although only a fraction of these have been investigated experimentally, information on their sequences and loci are readily available through various genomic databases, both non-specific (such as Ensembl), and dedicated non-coding RNA databases (such as LNCipedia, lncRNome, and lncRNAdb). Additionally, significant bioinformatic work has been carried out on them in terms of their expression and conservation within and across species [33]. With so much information on lncRNA now available, exploring this class of genes with a thorough experimental approach has become more feasible in recent years.

LncRNAs vary significantly in their distribution throughout cellular compartments, with the majority of transcripts residing predominantly in the nucleus, others in the cytoplasm, and some distributed more evenly between the two [34, 35]. For example, the *roX* transcripts in *Drosophila* are found in the nucleus, while *yar* is cytoplasmic [35]. The localisation of lncRNAs can give clues about their function; in the case of a chromatin restructuring lncRNA such as *roX1* or *roX2* it must be nuclear in order to access the chromatin. Localisation of particular lncRNAs can also affect their susceptibility to suppression by RNA interference and antisense oligonucleotides. An example of this is the suppression of nuclear lncRNAs *MALAT1* and *NEAT1* which in humans is more efficient using antisense methods, whereas cytoplasmic lncRNAs *DANCR* and *OIP5-AS1* are better suppressed with RNAi methods [35].

However, the sub-cellular localisation of the majority of lncRNAs has not been well characterised, with the localisation of relatively few being experimentally visualised. Single molecule RNA fluorescence *in situ* hybridisation has now been used to give high resolution data for the distribution of lncRNAs in human cells [34], and a systematic investigation of lncRNA localisation has been suggested as an important next step in expanding our understanding of their function; as well as a useful way to shed light on the potential relevance of lncRNAs to a particular mechanism.

LncRNA in human disease

LncRNAs have now been implicated as important factors linked to a range of human diseases. The broad range of biological functions of lncRNAs is reflected in the variety of different pathologies in which their aberrant expression is thought to be a contributing factor. Many lncRNAs have been shown to either be expressed at aberrant levels in cancerous cells [36-67], or their levels shown to affect the growth and behaviour of cancerous cells [46, 47, 49, 50, 52-56] (Table 1). This has prompted speculation that if better characterised, this class of genes may present many promising biomarkers, and even novel potential therapeutic targets. We cannot comprehensively cover this topic within the scope of this review, and point the reader to a comprehensive review of the topic for more information [57], but instead demonstrate this point with two well documented examples, below.

MALAT1, a highly conserved mammalian lncRNA, has been found to be overexpressed in human osteosarcoma cells and cell lines [46, 47]. It is hypothesised to function as a molecular scaffold for ribonucleoprotein complexes, acting as a transcriptional regulator for certain genes. Higher levels of *MALAT1* have been shown to be associated with “aggressive” cancer traits such as increased migration, metastasis, and clonogenic growth in non-small cell lung cancer [36-38] pancreatic [58], and prostate cancer cells [39]. Indeed, inducing a knockdown of *MALAT1* in osteosarcoma cell lines inhibited cell proliferation and invasion [46, 47].

The *HOTAIR* lncRNA, transcribed from an antisense Hox gene, plays an important role in the epigenetic regulation of genes thought to be due to its interactions with the Polycomb Repressive Complex 2 (PRC2) [43, 59], although recent work has indicated that PRC2 recruitment may be a downstream consequence of gene silencing, rather than initiating it [68]. *HOTAIR* is thought to act as a molecular scaffold, and is required for histone modification of particular genes across different chromosomes. Higher levels of *HOTAIR* have been found in colorectal cancer tissues, and are associated with increased tumour invasion, metastasis, vascular invasion, advanced tumour stage, and a worse prognosis in patients [43, 44]. *HOTAIR* has since been suggested for use as a biomarker for the progression and prognosis of certain cancers [44]. A *Drosophila* homologue for *HOTAIR* has not been identified, but given the similarities in polycomb regulation between species, it is likely that a targeted search might reveal such an equivalent.

Aside from cancer, strong evidence now exists linking certain lncRNAs to certain neurological pathologies [60]. LncRNAs have been shown to be relevant factors in amyotrophic lateral sclerosis, multiple sclerosis [61, 62], Alzheimer's disease [10, 63], Huntington's disease [64, 65], and Parkinson's disease, among others. For example, the *BACE1* antisense transcript (*BACE1-AS*) regulates mRNA stability of *BACE1*, a key enzyme in Alzheimer's disease pathology [10]. This subsequently affects amyloid-β 1-42 abundance, the increased expression of which is a hallmark of Alzheimer's disease. One mechanism by which lncRNAs have been hypothesized to impact neurodegenerative disease is through their induction of R-loop formation (which may be triggered by trinucleotide repeat expansion). R-loops have been shown to be capable of controlling the fate of neuroprotective genes [69], and are thought to contribute to the pathogenesis of fragile X syndrome and Friedrich's Ataxia [70, 71] by their silencing of certain genes. Additionally, work in *S. pombe* and *Arabidopsis* has suggested that R-loops may regulate lncRNA expression [72, 73], although whether this is true of lncRNAs linked to neurodegenerative diseases remains unclear. Trinucleotide repeats in lncRNAs are also known to be important in the pathogenesis of SCA8, by production of toxic noncoding CUG expansion RNAs from the ataxin 8 opposite strand (*ATXN8OS*), thought to cause a toxic gain of function at both the RNA and protein level [74, 75].

Another area of disease in which lncRNAs have been proven relevant is cardiovascular disease [66, 67]. Evidence now shows that lncRNAs are an important factor in susceptibility to coronary artery disease and myocardial infarction, prognosis in recovery from myocardial infarction, cardiovascular disease mortality, and heart failure [67]. Once again their correlations with prognosis and susceptibility have placed lncRNAs in the spotlight as a promising avenue of investigation in finding novel biomarkers.

Interestingly, *Drosophila* lncRNAs have been shown hold functional roles very relevant to these pathologies. *Hsromega* [76-80] and *bft* [81] are required for proper apoptosis process and cell differentiation, *yar* [82] and *CRG* [83] serve regulatory roles in the nervous system, and *sclA* and *sclB* are required for normal calcium transients and cardiac muscle contractility [19]. This is particularly promising given that these links can be made from the limited pool of *Drosophila* lncRNAs that have been experimentally characterised.

Molecular functions of lncRNA conserved in *Drosophila*

LncRNAs have been shown to function via a wide range of molecular mechanisms, falling under the broad categories of signals, molecular decoys, guide RNAs, or scaffolds [84]. Some lncRNAs have convincingly been shown to be translated, with the small peptide products (smORFs) having important biological functions [14-19, 22-25]. Through these various mechanisms (Figure 1), they have been implicated in

regulation of a diverse array of processes, such as differentiation, development, cell proliferation, nervous system function, and cardiovascular function in both *Drosophila* and humans, despite the lack of sequence conservation in lncRNAs across species. Importantly, similarities in the modes of action of lncRNAs have been found at the molecular level between organisms, discussed below.

LncRNA in the regulation of chromatin structure and gene expression

One of the most extensively studied molecular mechanisms of lncRNA modes of action is their role in sex chromosome dosage compensation pathways. Due to the difference in the number of X chromosome copies between males and females, there exists a compensation pathway required to maintain a similar level of expression for genes located on the X chromosome. In *Drosophila*, this is achieved by transcriptional hyperactivation of the single copy of the genes in males, allowing their expression at comparable levels to that given by the two copies of the gene found in females [85]. In humans, by contrast, the genes located on the X-chromosome in human females are partially transcriptionally repressed, giving a similar level of expression to that seen in males [86].

In *Drosophila*, the *RNA on the X* genes, *roX1* and *roX2*, are expressed in males, and regulate the assembly of the Male Specific Lethal (MSL) complex in *Drosophila*; a chromatin modifier that functions in histone modification [87-90]. The recruitment and binding of MSL proteins by high affinity sequences on the nascent *roX* transcripts covering the X chromosome allows the assembly of the active MSL complex, which can then spread in cis, allowing chromatin restructuring and hyperactivation of specific regions of the chromosome.

An immediate comparison can be made between the *roX* genes in *Drosophila*, and lncRNAs involved in the sex chromosome dosage compensation pathway in humans and other mammals; *X-inactive specific transcript* (*Xist*) and its antisense transcript, *Tsix*. Like the *roX* genes, *Xist* coats the X chromosome, where it regulates chromatin modifications, with consequent effects on the expression of particular target genes [91, 92]. Unlike *roX*, *Xist* is expressed in females, and regulates the inactivation of the X chromosome by facilitating the initiation and stabilising of the X chromosome inactivation process [86].

Although these lncRNA genes differ in their sequence, there are striking similarities between their role in specific regulation of the X-chromosome and the molecular mechanisms by which they are thought to achieve this. Interestingly, a subset of lncRNAs involved in chromatin looping, called topological anchor point RNAs (tapRNAs), have been identified in the human and mouse genomes, with conserved zinc-finger motifs capable of binding DNA and RNA [93]. Whether these are conserved in *Drosophila* has not yet been studied, but given the involvement of lncRNAs in *Drosophila* chromatin regulation so far, this may be a promising avenue to explore, and may reveal a wider conservation of this class of lncRNA chromatin regulators.

LncRNAs in the production of small peptides

The *Drosophila sarcolamban* (*scl*) gene, originally classified as a lncRNA *pncr003* [94], is transcribed into a 992 base-pair mRNA, which is translated to produce two related peptides of less than 30 amino acids [19]. The *scl* gene is expressed in muscle cells, and *scl* null mutants show arrhythmic cardiac contractions, a phenotype produced by abnormal intracellular calcium levels in contracting muscle cells [19].

Interestingly, the *scl* genes were found to have homologues in humans, namely *sarcolipin* (*sln*) and its longer paralogue, *phospholamban* (*pln*), encoding peptides of 31 and 52 amino acids respectively [19]. Phylogenetic analysis suggests that these genes belong to the same gene family, derived from a single ancestral gene, conserved for more than 550 million years. Furthermore, their function also seems to be conserved, with Sln and Pln regulating calcium transport in mammalian muscle cells, via dampening of Sarco-endoplasmic Reticulum Ca^{2+} adenosine triphosphate (SERCA) pump function. Scl peptides were able to colocalise and interact with *Drosophila* SERCA. Exogenous expression of the human Pln and Sln peptides in *Drosophila scl* mutant muscle cells were sufficient to rescue muscle function. Importantly, aberrant levels of Sln in humans have been linked to heart arrhythmias [95]. Regulation of SERCA by micropeptides (encoded by lncRNAs) has been extensively exploited in mammals; with tissue specific positive and negative regulators being found [22, 96, 97]. In addition, the number of characterized lncRNA genes encoding micropeptides is rapidly increasing, with roles found in a myriad of essential, conserved cellular functions, from phagocytosis [17] and cellular motility [98] to RNA degradation [18]. Thus, these examples show that lncRNAs that produce biologically relevant peptides may be conserved in structure, function, and relevance to pathologies between humans and *Drosophila* [19, 22].

Future directions

As previously shown in *sarcolamban*, proving the protein-coding potential of lncRNAs is a painstaking process, and an extremely difficult topic to broach; with genes having previously been catalogued as "non-coding" by arbitrary rules. Definitively showing the translation, or lack thereof, of an RNA using experimental techniques can be an arduous process, making this approach impractical to apply to the entire catalogue of identified lncRNAs. Ribosome profiling (in which a protease digestion is used to degrade RNA not protected by a bound ribosome,) and polysome profiling (where RNAs are separated by the number of ribosomes that are attached to different transcripts) have been used to provide a translational snapshot for several lncRNAs so far. This data has given a profile for lncRNA translation, but the threshold for significant translation is difficult to define in a non-arbitrary fashion. Therefore, use of model organisms to determine the biological function of any particular lncRNA remains crucial to gaining a meaningful understanding of the function of these molecules. A thorough and processive approach to clarifying this aspect of the gene class, as well as standardising measures and cut-offs for translational activity is an important priority for those in the field.

Bioinformatic approaches to elucidating the possible biological functions of lncRNAs are also being developed, although this method is not without its difficulties. Due to the poor sequence conservation characteristic of lncRNAs, standard approaches used to identify biologically relevant transcripts by their conservation within and across species are significantly less effective within this gene class. However, recent work has noted distinctive selection patterns in lncRNAs based on secondary structure [99], which may be of help in future analyses.

To conclude, we suggest that the studies currently being carried out on lncRNA in *Drosophila* should be of interest to a far wider audience than just fly geneticists, having shown that as a model organism, *Drosophila* is a logical choice both for better characterising this gene class, and for precursor studies to highlight genes and mechanisms that can be carried forward into more expensive and laborious large animal and human work. The superb annotation of the *Drosophila* genome and transcriptome, coupled with further increases in RNA-sequencing data available, will provide a candidate pool of lncRNAs for a rapid functional characterization (using the sophisticated genetic tools available in *Drosophila*). Therefore, further lncRNA

studies in *Drosophila*, of a suitably high calibre, are likely to provide us not only with a better understanding of the basic science behind this gene class, but promise to highlight potential biomarkers, elucidate genetic mechanisms behind a range of diseases, and perhaps provide novel targets for next generation therapeutics.

Abbreviations

lncRNA, long non-coding RNA; MSL, Male Specific Lethal; pln, phospholamban; PRC2, Polycomb Repressive Complex 2; roX, RNA on the X; scl, sarcolamban; sln, sarcolipin; small open reading frame, smORF; tapRNA, topological anchor point RNA; Xist, X-inactive specific transcript.

Funding

This work was funded by a Brighton and Sussex Medical School studentship [grant number WB002-61] and the Biotechnology and Biological Sciences Research Council [grant number BB/N001753/1].

Competing Interests

The Authors declare that there are no competing interests associated with the manuscript**.**

Acknowledgements

The authors thank Benjamin Towler and Pedro Patraquim for their input and advice.

References

1 Brown, C. J., Ballabio, A., Rupert, J. L., Lafreniere, R. G., Grompe, M., Tonlorenzi, R. and Willard, H. F. (1991) A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature. **349**, 38-44

2 Taft, R. J., Pheasant, M. and Mattick, J. S. (2007) The relationship between non-protein-coding DNA and eukaryotic complexity. Bioessays. **29**, 288-299

3 Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D. G., Lagarde, J., Veeravalli, L., Ruan, X., Ruan, Y., Lassmann, T., Carninci, P., Brown, J. B., Lipovich, L., Gonzalez, J. M., Thomas, M., Davis, C. A., Shiekhattar, R., Gingeras, T. R., Hubbard, T. J., Notredame, C., Harrow, J. and Guigó, R. (2012) The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res. **22**, 1775-1789

4 Au, P. C., Zhu, Q. H., Dennis, E. S. and Wang, M. B. (2011) Long non-coding RNA-mediated mechanisms independent of the RNAi pathway in animals and plants. RNA Biol. **8**, 404-414

5 Hirota, K., Miyoshi, T., Kugou, K., Hoffman, C. S., Shibata, T. and Ohta, K. (2008) Stepwise chromatin remodelling by a cascade of transcription initiation of non-coding RNAs. Nature. **456**, 130-134

6 Tian, D., Sun, S. and Lee, J. T. (2010) The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. Cell. **143**, 390-403

7 Yoo, E. J., Cooke, N. E. and Liebhaber, S. A. (2012) An RNA-independent linkage of noncoding transcription to long-range enhancer function. Mol Cell Biol. **32**, 2020-2029

8 Lai, F., Orom, U. A., Cesaroni, M., Beringer, M., Taatjes, D. J., Blobel, G. A. and Shiekhattar, R. (2013) Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. Nature. **494**, 497-501

9 Yoon, J. H., Abdelmohsen, K., Srikantan, S., Yang, X., Martindale, J. L., De, S., Huarte, M., Zhan, M., Becker, K. G. and Gorospe, M. (2012) LincRNA-p21 suppresses target mRNA translation. Mol Cell. **47**, 648-655

10 Faghihi, M. A., Modarresi, F., Khalil, A. M., Wood, D. E., Sahagan, B. G., Morgan, T. E., Finch, C. E., St Laurent, G., Kenny, P. J. and Wahlestedt, C. (2008) Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. Nat Med. **14**, 723-730

11 Wang, H., Iacoangeli, A., Lin, D., Williams, K., Denman, R. B., Hellen, C. U. and Tiedge, H. (2005) Dendritic BC1 RNA in translational control mechanisms. J Cell Biol. **171**, 811-821

12 Gong, C. and Maquat, L. E. (2011) lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. Nature. **470**, 284-288

13 Tripathi, V., Ellis, J. D., Shen, Z., Song, D. Y., Pan, Q., Watt, A. T., Freier, S. M., Bennett, C. F., Sharma, A., Bubulya, P. A., Blencowe, B. J., Prasanth, S. G. and Prasanth, K. V. (2010) The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol Cell. **39**, 925-938

14 Galindo, M. I., Pueyo, J. I., Fouix, S., Bishop, S. A. and Couso, J. P. (2007) Peptides encoded by short ORFs control development and define a new eukaryotic gene family. PLoS Biol. **5**, e106

15 Kondo, T., Hashimoto, Y., Kato, K., Inagaki, S., Hayashi, S. and Kageyama, Y. (2007) Small peptide regulators of actin-based cell morphogenesis encoded by a polycistronic mRNA. Nat Cell Biol. **9**, 660-665

16 Pueyo, J. I. and Couso, J. P. (2008) The 11-aminoacid long Tarsal-less peptides trigger a cell signal in Drosophila leg development. Dev Biol. **324**, 192-201

17 Pueyo, J. I., Magny, E. G., Sampson, C. J., Amin, U., Evans, I. R., Bishop, S. A. and Couso, J. P. (2016) Hemotin, a Regulator of Phagocytosis Encoded by a Small ORF and Conserved across Metazoans. PLoS Biol. **14**, e1002395

18 D'Lima, N. G., Ma, J., Winkler, L., Chu, Q., Loh, K. H., Corpuz, E. O., Budnik, B. A., Lykke-Andersen, J., Saghatelian, A. and Slavoff, S. A. (2017) A human microprotein that interacts with the mRNA decapping complex. Nat Chem Biol. **13**, 174-180

19 Magny, E. G., Pueyo, J. I., Pearl, F. M., Cespedes, M. A., Niven, J. E., Bishop, S. A. and Couso, J. P. (2013) Conserved regulation of cardiac calcium uptake by peptides encoded in small open reading frames. Science. **341**, 1116-1120

20 Lauressergues, D., Couzigou, J. M., Clemente, H. S., Martinez, Y., Dunand, C., Bécard, G. and Combier, J. P. (2015) Primary transcripts of microRNAs encode regulatory peptides. Nature. **520**, 90-93
21 Williamson, L., Saponaro, M., Boeing, S., East, P., Mitter, R., Kantidakis, T., Kelly, G. P., Lobley, A., Walker, J., Spencer-Dene, B., Howell, M., Stewart, A. and Svejstrup, J. Q. (2017) UV Irradiation Induces a Non-coding RNA that Functionally Opposes the Protein Encoded by the Same Gene. Cell. **168**, 843-855.e813
22 Anderson, D. M., Anderson, K. M., Chang, C. L., Makarewich, C. A., Nelson, B. R., McAnally, J. R., Kasaragod, P., Shelton, J. M., Liou, J., Bassel-Duby, R. and Olson, E. N. (2015) A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. Cell. **160**, 595-606
23 Aspden, J. L., Eyre-Walker, Y. C., Phillips, R. J., Amin, U., Mumtaz, M. A., Brocard, M. and Couso, J. P. (2014) Extensive translation of small Open Reading Frames revealed by Poly-Ribo-Seq. Elife. **3**, e03528
24 Mackowiak, S. D., Zauber, H., Bielow, C., Thiel, D., Kutz, K., Calviello, L., Mastrobuoni, G., Rajewsky, N., Kempa, S., Selbach, M. and Obermayer, B. (2015) Extensive identification and analysis of conserved small ORFs in animals. Genome Biol. **16**, 179
25 Ruiz-Orera, J., Messeguer, X., Subirana, J. A. and Alba, M. M. (2014) Long non-coding RNAs as a source of new peptides. Elife. **3**, e03523
26 Reiter, L. T., Potocki, L., Chien, S., Gribskov, M. and Bier, E. (2001) A systematic analysis of human disease-associated gene sequences in Drosophila melanogaster. Genome Res. **11**, 1114-1125
27 Bier, E. (2005) Drosophila, the golden bug, emerges as a tool for human genetics. Nat Rev Genet. **6**, 9-23
28 Schoenfelder, S., Smits, G., Fraser, P., Reik, W. and Paro, R. (2007) Non-coding transcripts in the H19 imprinting control region mediate gene silencing in transgenic Drosophila. EMBO Rep. **8**, 1068-1073
29 Mercer, T. R. and Mattick, J. S. (2013) Structure and function of long noncoding RNAs in epigenetic regulation. Nat Struct Mol Biol. **20**, 300-307
30 Haerty, W. and Ponting, C. P. (2013) Mutations within lncRNAs are effectively selected against in fruitfly but not in human. Genome Biol. **14**, R49
31 Quinn, J. J., Zhang, Q. C., Georgiev, P., Ilik, I. A., Akhtar, A. and Chang, H. Y. (2016) Rapid evolutionary turnover underlies conserved lncRNA-genome interactions. Genes Dev. **30**, 191-207
32 Nguyen, D., Krueger, B. J., Sedore, S. C., Brogie, J. E., Rogers, J. T., Rajendra, T. K., Saunders, A., Matera, A. G., Lis, J. T., Uguen, P. and Price, D. H. (2012) The Drosophila 7SK snRNP and the essential role of dHEXIM in development. Nucleic Acids Res. **40**, 5283-5297
33 Ulitsky, I. (2016) Evolution to the rescue: using comparative genomics to understand long non-coding RNAs. Nat Rev Genet. **17**, 601-614
34 Cabili, M. N., Dunagin, M. C., McClanahan, P. D., Biaesch, A., Padovan-Merhar, O., Regev, A., Rinn, J. L. and Raj, A. (2015) Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. Genome Biol. **16**, 20
35 Lennox, K. A. and Behlke, M. A. (2016) Cellular localization of long non-coding RNAs affects silencing by RNAi more than by antisense oligonucleotides. Nucleic Acids Res. **44**, 863-877
36 Ji, P., Diederichs, S., Wang, W., Böing, S., Metzger, R., Schneider, P. M., Tidow, N., Brandt, B., Buerger, H., Bulk, E., Thomas, M., Berdel, W. E., Serve, H. and Müller-Tidow, C. (2003) MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. Oncogene. **22**, 8031-8041
37 Schmidt, L. H., Spieker, T., Koschmieder, S., Schäffers, S., Humberg, J., Jungen, D., Bulk, E., Hascher, A., Wittmer, D., Marra, A., Hillejan, L., Wiebe, K., Berdel, W. E., Wiewrodt, R. and Muller-Tidow, C. (2011) The long noncoding MALAT-1 RNA indicates a poor prognosis in non-small cell lung cancer and induces migration and tumor growth. J Thorac Oncol. **6**, 1984-1992
38 Tano, K., Mizuno, R., Okada, T., Rakwal, R., Shibato, J., Masuo, Y., Ijiri, K. and Akimitsu, N. (2010) MALAT-1 enhances cell motility of lung adenocarcinoma cells by influencing the expression of motility-related genes. FEBS Lett. **584**, 4575-4580
39 Ren, S., Liu, Y., Xu, W., Sun, Y., Lu, J., Wang, F., Wei, M., Shen, J., Hou, J., Gao, X., Xu, C., Huang, J. and Zhao, Y. (2013) Long noncoding RNA MALAT-1 is a new potential therapeutic target for castration resistant prostate cancer. J Urol. **190**, 2278-2287
40 Hao, Y., Crenshaw, T., Moulton, T., Newcomb, E. and Tycko, B. (1993) Tumour-suppressor activity of H19 RNA. Nature. **365**, 764-767

41 Li, H., Yu, B., Li, J., Su, L., Yan, M., Zhu, Z. and Liu, B. (2014) Overexpression of lncRNA H19 enhances carcinogenesis and metastasis of gastric cancer. Oncotarget. **5**, 2318-2329
42 Yang, F., Bi, J., Xue, X., Zheng, L., Zhi, K., Hua, J. and Fang, G. (2012) Up-regulated long non-coding RNA H19 contributes to proliferation of gastric cancer cells. FEBS J. **279**, 3159-3165
43 Kogo, R., Shimamura, T., Mimori, K., Kawahara, K., Imoto, S., Sudo, T., Tanaka, F., Shibata, K., Suzuki, A., Komune, S., Miyano, S. and Mori, M. (2011) Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. Cancer Res. **71**, 6320-6326
44 Wu, Z. H., Wang, X. L., Tang, H. M., Jiang, T., Chen, J., Lu, S., Qiu, G. Q., Peng, Z. H. and Yan, D. W. (2014) Long non-coding RNA HOTAIR is a powerful predictor of metastasis and poor prognosis and is associated with epithelial-mesenchymal transition in colon cancer. Oncol Rep. **32**, 395-402
45 Pang, Q., Ge, J., Shao, Y., Sun, W., Song, H., Xia, T., Xiao, B. and Guo, J. (2014) Increased expression of long intergenic non-coding RNA LINC00152 in gastric cancer and its clinical significance. Tumour Biol. **35**, 5441-5447
46 Dong, Y., Liang, G., Yuan, B., Yang, C., Gao, R. and Zhou, X. (2015) MALAT1 promotes the proliferation and metastasis of osteosarcoma cells by activating the PI3K/Akt pathway. Tumour Biol. **36**, 1477-1486
47 Cai, X., Liu, Y., Yang, W., Xia, Y., Yang, C., Yang, S. and Liu, X. (2016) Long noncoding RNA MALAT1 as a potential therapeutic target in osteosarcoma. J Orthop Res. **34**, 932-941
48 Wang, J. Z., Xu, C. L., Wu, H. and Shen, S. J. (2017) LncRNA SNHG12 promotes cell growth and inhibits cell apoptosis in colorectal cancer cells. Braz J Med Biol Res. **50**, e6079
49 Sun, J., Wang, X., Fu, C., Zou, J., Hua, H. and Bi, Z. (2016) Long noncoding RNA FGFR3-AS1 promotes osteosarcoma growth through regulating its natural antisense transcript FGFR3. Mol Biol Rep. **43**, 427-436
50 Cong, M., Li, J., Jing, R. and Li, Z. (2016) Long non-coding RNA tumor suppressor candidate 7 functions as a tumor suppressor and inhibits proliferation in osteosarcoma. Tumour Biol. **37**, 9441-9450
51 Ma, B., Li, M., Zhang, L., Huang, M., Lei, J. B., Fu, G. H., Liu, C. X., Lai, Q. W., Chen, Q. Q. and Wang, Y. L. (2016) Upregulation of long non-coding RNA TUG1 correlates with poor prognosis and disease status in osteosarcoma. Tumour Biol. **37**, 4445-4455
52 Li, F., Cao, L., Hang, D., Wang, F. and Wang, Q. (2015) Long non-coding RNA HOTTIP is up-regulated and associated with poor prognosis in patients with osteosarcoma. Int J Clin Exp Pathol. **8**, 11414-11420
53 Marques Howarth, M., Simpson, D., Ngok, S. P., Nieves, B., Chen, R., Siprashvili, Z., Vaka, D., Breese, M. R., Crompton, B. D., Alexe, G., Hawkins, D. S., Jacobson, D., Brunner, A. L., West, R., Mora, J., Stegmaier, K., Khavari, P. and Sweet-Cordero, E. A. (2014) Long noncoding RNA EWSAT1-mediated gene repression facilitates Ewing sarcoma oncogenesis. J Clin Invest. **124**, 5275-5290
54 Wang, Y., Yao, J., Meng, H., Yu, Z., Wang, Z., Yuan, X., Chen, H. and Wang, A. (2015) A novel long non-coding RNA, hypoxia-inducible factor-2α promoter upstream transcript, functions as an inhibitor of osteosarcoma stem cells in vitro. Mol Med Rep. **11**, 2534-2540
55 Min, L., Hong, S., Duan, H., Zhou, Y., Zhang, W., Luo, Y., Shi, R. and Tu, C. (2016) Antidifferentiation Noncoding RNA Regulates the Proliferation of Osteosarcoma Cells. Cancer Biother Radiopharm. **31**, 52-57
56 Ruan, W., Wang, P., Feng, S., Xue, Y. and Li, Y. (2016) Long non-coding RNA small nucleolar RNA host gene 12 (SNHG12) promotes cell proliferation and migration by upregulating angiomotin gene expression in human osteosarcoma cells. Tumour Biol. **37**, 4065-4073
57 Esteller, M. (2011) Non-coding RNAs in human disease. Nat Rev Genet. **12**, 861-874
58 Li, L., Chen, H., Gao, Y., Wang, Y. W., Zhang, G. Q., Pan, S. H., Ji, L., Kong, R., Wang, G., Jia, Y. H., Bai, X. W. and Sun, B. (2016) Long Noncoding RNA MALAT1 Promotes Aggressive Pancreatic Cancer Proliferation and Metastasis via the Stimulation of Autophagy. Mol Cancer Ther. **15**, 2232-2243
59 Meller, V. H., Joshi, S. S. and Deshpande, N. (2015) Modulation of Chromatin by Noncoding RNA. Annu Rev Genet. **49**, 673-695
60 Roberts, T. C., Morris, K. V. and Wood, M. J. (2014) The role of long non-coding RNAs in neurodevelopment, brain function and neurological disease. Philos Trans R Soc Lond B Biol Sci. **369**
61 Mori, K., Arzberger, T., Grässer, F. A., Gijselinck, I., May, S., Rentzsch, K., Weng, S. M., Schludi, M. H., van der Zee, J., Cruts, M., Van Broeckhoven, C., Kremmer, E., Kretzschmar, H. A., Haass, C. and Edbauer,

460 D. (2013) Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into
461 aggregating dipeptide repeat proteins. Acta Neuropathol. **126**, 881-893
462 62 Zu, T., Liu, Y., Bañez-Coronel, M., Reid, T., Pletnikova, O., Lewis, J., Miller, T. M., Harms, M. B.,
463 Falchook, A. E., Subramony, S. H., Ostrow, L. W., Rothstein, J. D., Troncoso, J. C. and Ranum, L. P. (2013)
464 RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. Proc
465 Natl Acad Sci U S A. **110**, E4968-4977
466 63 Lee, D. Y., Moon, J., Lee, S. T., Jung, K. H., Park, D. K., Yoo, J. S., Sunwoo, J. S., Byun, J. I., Shin, J. W.,
467 Jeon, D., Jung, K. Y., Kim, M., Lee, S. K. and Chu, K. (2015) Distinct Expression of Long Non-Coding RNAs in
468 an Alzheimer's Disease Model. J Alzheimers Dis. **45**, 837-849
469 64 Johnson, R., Teh, C. H., Jia, H., Vanisri, R. R., Pandey, T., Lu, Z. H., Buckley, N. J., Stanton, L. W. and
470 Lipovich, L. (2009) Regulation of neural macroRNAs by the transcriptional repressor REST. RNA. **15**, 85-96
471 65 Johnson, R. (2012) Long non-coding RNAs in Huntington's disease neurodegeneration. Neurobiol
472 Dis. **46**, 245-254
473 66 Uchida, S. and Dimmeler, S. (2015) Long noncoding RNAs in cardiovascular diseases. Circ Res. **116**,
474 737-750
475 67 Archer, K., Broskova, Z., Bayoumi, A. S., Teoh, J. P., Davila, A., Tang, Y., Su, H. and Kim, I. M. (2015)
476 Long Non-Coding RNAs as Master Regulators in Cardiovascular Diseases. Int J Mol Sci. **16**, 23651-23667
477 68 Portoso, M., Ragazzini, R., Brenčič, Ž., Moiani, A., Michaud, A., Vassilev, I., Wassef, M., Servant, N.,
478 Sargueil, B. and Margueron, R. (2017) PRC2 is dispensable for HOTAIR-mediated transcriptional repression.
479 EMBO J. **36**, 981-994
480 69 Salvi, J. S. and Mekhail, K. (2015) R-loops highlight the nucleus in ALS. Nucleus. **6**, 23-29
481 70 Groh, M., Lufino, M. M., Wade-Martins, R. and Gromak, N. (2014) R-loops associated with triplet
482 repeat expansions promote gene silencing in Friedreich ataxia and fragile X syndrome. PLoS Genet. **10**,
483 e1004318
484 71 Colak, D., Zaninovic, N., Cohen, M. S., Rosenwaks, Z., Yang, W. Y., Gerhardt, J., Disney, M. D. and
485 Jaffrey, S. R. (2014) Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X
486 syndrome. Science. **343**, 1002-1005
487 72 Sun, Q., Csorba, T., Skourti-Stathaki, K., Proudfoot, N. J. and Dean, C. (2013) R-loop stabilization
488 represses antisense transcription at the Arabidopsis FLC locus. Science. **340**, 619-621
489 73 Nakama, M., Kawakami, K., Kajitani, T., Urano, T. and Murakami, Y. (2012) DNA-RNA hybrid
490 formation mediates RNAi-directed heterochromatin formation. Genes Cells. **17**, 218-233
491 74 Tan, H., Xu, Z. and Jin, P. (2012) Role of noncoding RNAs in trinucleotide repeat neurodegenerative
492 disorders. Exp Neurol. **235**, 469-475
493 75 Mutsuddi, M., Marshall, C. M., Benzow, K. A., Koob, M. D. and Rebay, I. (2004) The spinocerebellar
494 ataxia 8 noncoding RNA causes neurodegeneration and associates with staufen in Drosophila. Curr Biol. **14**,
495 302-308
496 76 Lakhotia, S. C., Mallik, M., Singh, A. K. and Ray, M. (2012) The large noncoding hsrω-n transcripts
497 are essential for thermotolerance and remobilization of hnRNPs, HP1 and RNA polymerase II during
498 recovery from heat shock in Drosophila. Chromosoma. **121**, 49-70
499 77 Prasanth, K. V., Rajendra, T. K., Lal, A. K. and Lakhotia, S. C. (2000) Omega speckles - a novel class of
500 nuclear speckles containing hnRNPs associated with noncoding hsr-omega RNA in Drosophila. J Cell Sci. **113**
501 **Pt 19**, 3485-3497
502 78 Perrimon, N., Lanjuin, A., Arnold, C. and Noll, E. (1996) Zygotic lethal mutations with maternal
503 effect phenotypes in Drosophila melanogaster. II. Loci on the second and third chromosomes identified by
504 P-element-induced mutations. Genetics. **144**, 1681-1692
505 79 Mallik, M. and Lakhotia, S. C. (2009) The developmentally active and stress-inducible noncoding
506 hsromega gene is a novel regulator of apoptosis in Drosophila. Genetics. **183**, 831-852
507 80 Johnson, T. K., Cockerell, F. E. and McKechnie, S. W. (2011) Transcripts from the Drosophila heat-
508 shock gene hsr-omega influence rates of protein synthesis but hardly affect resistance to heat knockdown.
509 Mol Genet Genomics. **285**, 313-323
510 81 Hardiman, K. E., Brewster, R., Khan, S. M., Deo, M. and Bodmer, R. (2002) The bereft gene, a
511 potential target of the neural selector gene cut, contributes to bristle morphogenesis. Genetics. **161**, 231-
512 247

82 Soshnev, A. A., Ishimoto, H., McAllister, B. F., Li, X., Wehling, M. D., Kitamoto, T. and Geyer, P. K. (2011) A conserved long noncoding RNA affects sleep behavior in Drosophila. Genetics. **189**, 455-468
83 Li, M., Wen, S., Guo, X., Bai, B., Gong, Z., Liu, X., Wang, Y., Zhou, Y., Chen, X., Liu, L. and Chen, R. (2012) The novel long non-coding RNA CRG regulates Drosophila locomotor behavior. Nucleic Acids Res. **40**, 11714-11727
84 Wang, K. C. and Chang, H. Y. (2011) Molecular mechanisms of long noncoding RNAs. Mol Cell. **43**, 904-914
85 Deng, X. and Meller, V. H. (2006) roX RNAs are required for increased expression of X-linked genes in Drosophila melanogaster males. Genetics. **174**, 1859-1866
86 Pontier, D. B. and Gribnau, J. (2011) Xist regulation and function explored. Hum Genet. **130**, 223-236
87 Park, Y., Kelley, R. L., Oh, H., Kuroda, M. I. and Meller, V. H. (2002) Extent of chromatin spreading determined by roX RNA recruitment of MSL proteins. Science. **298**, 1620-1623
88 Kelley, R. L., Lee, O. K. and Shim, Y. K. (2008) Transcription rate of noncoding roX1 RNA controls local spreading of the Drosophila MSL chromatin remodeling complex. Mech Dev. **125**, 1009-1019
89 Kelley, R. L. and Kuroda, M. I. (2003) The Drosophila roX1 RNA gene can overcome silent chromatin by recruiting the male-specific lethal dosage compensation complex. Genetics. **164**, 565-574
90 Oh, H., Park, Y. and Kuroda, M. I. (2003) Local spreading of MSL complexes from roX genes on the Drosophila X chromosome. Genes Dev. **17**, 1334-1339
91 Plath, K., Mlynarczyk-Evans, S., Nusinow, D. A. and Panning, B. (2002) Xist RNA and the mechanism of X chromosome inactivation. Annu Rev Genet. **36**, 233-278
92 McHugh, C. A., Chen, C. K., Chow, A., Surka, C. F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A., Sweredoski, M. J., Shishkin, A. A., Su, J., Lander, E. S., Hess, S., Plath, K. and Guttman, M. (2015) The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature. **521**, 232-236
93 Amaral, P. P., Leonardi, T., Han, N., Vire, E., Gascoigne, D. K., Arias-Carrasco, R., Zhang, A., Pluchino, S. and Maracaja-Coutinho, V. (2016) Genomic positional conservation identifies topological anchor point (tap) RNAs linked to developmental loci.
94 Tupy, J. L., Bailey, A. M., Dailey, G., Evans-Holm, M., Siebel, C. W., Misra, S., Celniker, S. E. and Rubin, G. M. (2005) Identification of putative noncoding polyadenylated transcripts in Drosophila melanogaster. Proc Natl Acad Sci U S A. **102**, 5495-5500
95 Shanmugam, M., Molina, C. E., Gao, S., Severac-Bastide, R., Fischmeister, R. and Babu, G. J. (2011) Decreased sarcolipin protein expression and enhanced sarco(endo)plasmic reticulum Ca2+ uptake in human atrial fibrillation. Biochem Biophys Res Commun. **410**, 97-101
96 Anderson, D. M., Makarewich, C. A., Anderson, K. M., Shelton, J. M., Bezprozvannaya, S., Bassel-Duby, R. and Olson, E. N. (2016) Widespread control of calcium signaling by a family of SERCA-inhibiting micropeptides. Sci Signal. **9**, ra119
97 Nelson, B. R., Makarewich, C. A., Anderson, D. M., Winders, B. R., Troupes, C. D., Wu, F., Reese, A. L., McAnally, J. R., Chen, X., Kavalali, E. T., Cannon, S. C., Houser, S. R., Bassel-Duby, R. and Olson, E. N. (2016) A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. Science. **351**, 271-275
98 Pauli, A., Norris, M. L., Valen, E., Chew, G. L., Gagnon, J. A., Zimmerman, S., Mitchell, A., Ma, J., Dubrulle, J., Reyon, D., Tsai, S. Q., Joung, J. K., Saghatelian, A. and Schier, A. F. (2014) Toddler: an embryonic signal that promotes cell movement via Apelin receptors. Science. **343**, 1248636
99 Pegueroles, C. and Gabaldón, T. (2016) Secondary structure impacts patterns of selection in human lncRNAs. BMC Biol. **14**, 60

562 **Figures**

563

564 **Table 1) A table summarising the lncRNAs linked to various kinds of cancer, as covered in this review.**

lncRNA	Associated disease	Reference
MALAT1	Osteosarcoma	[46, 47]
	Non-small cell lung cancer	[36-38]
	Prostate cancer	[39]
	Pancreatic cancer	[58]
HOTAIR	Colorectal cancer	[43, 44]
EWSAT1	Ewing sarcoma	[53]
HOTTIP	Osteosarcoma	[52]
HIF2PUT	Osteosarcoma	[54]
ANCR	Osteosarcoma	[55]
TUSC7	Osteosarcoma	[50]
FGFR3-AS1	Osteosarcoma	[49]
SNHG12	Osteosarcoma	[56]
TUG1	Osteosarcoma	[51]
H19	Wilms tumour	[40]
	Gastric cancer	[41, 42]
LINC00152	Gastric cancer	[45]

565

566

Figure 1) A cartoon depicting the molecular mechanisms by which lncRNAs can function.

a) Some lncRNAs (red), such as *Xist* and *RoX1*, can act to modulate expression of a certain gene by binding to a transcription modifier or chromatin modifier (purple). b) LncRNAs (red) such as *HOTAIR* can act as molecular scaffolds, allowing the assembly of protein complexes (teal, green, dark purple) with genetic regulatory roles e.g. polycomb complex PRC2. c) LncRNAs (red) can act as molecular decoys, to sequester miRNAs (orange) or proteins (purple). d) Alternatively, lncRNAs (red) can act as molecular decoys, occluding or removing transcription factors, proteins, or RNAs (purple) from their functional location. e) LncRNAs (red) can act as a molecular guide, allowing formation of ribonucleoprotein complexes (yellow) to specific target sites. f) It has also been shown that lncRNAs (blue as DNA, red as RNA) can be actively translated into functional smORF peptides (orange) such as the SclA and SclB peptides, which function in regulating calcium transport in cardiac muscle.