



RESEARCH HIGHLIGHT

Hide and Seek: Protein-coding Sequences Inside “Non-coding” RNAs



Daniel Oehler^{1,2,*}, Jan Haas^{1,2,*}

¹ Institute for Cardiomyopathies, Department of Internal Medicine III, University of Heidelberg, 69120 Heidelberg, Germany

² German Centre for Cardiovascular Research (DZHK), Heidelberg/Mannheim, Germany

Received 20 March 2016; accepted 31 March 2016

Available online 9 April 2016

Handled by Andreas Keller

Calcium homeostasis is crucial for muscle contractility

Muscle cells are critically dependent on calcium homeostasis. Without having the right amount of calcium ions just on the spot and coordinated in between muscle cells, no contraction can take place. Therefore, calcium homeostasis is one of the critical regulatory mechanisms in all muscle cells, including skeletal muscle and heart [1,2]. Sarco-endoplasmic reticulum Ca^{2+} adenosine triphosphatase (SERCA) is responsible for the relaxation of muscle cells by pumping Ca^{2+} into the sarcoplasmic reticulum (SR) [3,4]. For the process of muscle contraction, not only excitation via calcium-release but also the quick reestablishing of the baseline calcium gradient is absolutely necessary. Proteins known to be involved in this process include phospholamban (PLN; inhibition of SERCA in phosphorylated state) [5–7], sarcolipin (SLN; uncoupling of SERCA in heat regulation) [7–9], and myoregulin (MLN; inhibition of skeletal muscle SERCA) [10].

DWORF – a peptide “hiding” in a non-coding RNA

Recently Nelson et al. published a paper in *Science* [11], describing their discovery of a so-far unknown small protein, named dwarf open reading frame (DWORF), which could function as an activator of SERCA [11]. DWORF consists of 34 amino acid residues and “is the third smallest full-length protein known to be encoded by the mouse genome” [11]. Astonishingly, DWORF is encoded by a locus within the open reading frame of a previously-annotated long non-coding RNA (lncRNA) gene (“NONMMUG026737” in mice, corresponding to “LOC100507537” in humans). Notably, a similar genomic structure is reported for the gene encoding MLN, which is located within a previously-annotated lncRNA named LINC00948 in humans [10].

To understand the impact of these findings, one must take a look at the background of lncRNAs. lncRNAs are commonly defined as RNAs that have more than 200 nucleotides in length and appear not to have coding potential [12]. The length distinguishes lncRNAs from small regulatory RNAs like microRNAs, while it is clear that there is not always a clear cut-off in discrimination to other RNA subtypes [13]. The biological function of lncRNAs — without going into more detail here — is not always easy to identify and seems to be very diverse [14,15]. In the past years, many newly-described so-called “lncRNA” have been published and it seems that they have biological implication in animal models for different

* Corresponding authors.

E-mail: Daniel.Oehler@med.uni-heidelberg.de (Oehler D), Jan.Haas@med.uni-heidelberg.de (Haas J).

^a ORCID: 0000-0002-0435-4119.

^b ORCID: 0000-0002-8040-8289.

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

<http://dx.doi.org/10.1016/j.gpb.2016.03.002>

1672-0229 © 2016 The Authors. Production and hosting by Elsevier B.V. on behalf of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

diseases, but their precise role mostly remains unclear [14,15]. In the aforementioned paper, the authors searched explicitly for “masked” proteins inside areas previously described as non-coding regions using a recognition software such as PhyloCSF [16]. Doing so, they were able to detect protein-coding information hidden in RNAs, which literally had been “thrown away” before. This is of course not true for every lncRNA. As we understand too little about the functions of lncRNAs in general, sometimes misleading interpretations as described above can occur, making it crucial to understand and divide this heterogeneous group of RNAs for correctly analyzing their functional implications (*e.g.*, regulatory roles).

DWOLF – increasing contractility by improving relaxation

DWOLF is conserved across species, which gives a glimpse at the importance of this protein. Though it is expressed in the heart and in slow-twitch muscle fibers, there is a lack of gene expression of *DWOLF* in fast-twitch muscle. Slow twitch muscle cells contract slower and efficiently use oxygen for enduring contraction purposes, while in contrast fast-twitch muscle cells use preferably anaerobic pathways for short-term contractions [17]. In other words, the DWOLF regulates muscle cells involved in long-term contractility such as the heart.

The authors confirmed the expression of the presumed protein (*e.g.*, by Western blotting) and showed that SERCA and DWOLF not only co-localize, but also interact. In other words, DWOLF activates SERCA using the same binding domain as PLN with similar binding affinities. In tetanic stimulated muscle cells, DWOLF enhanced the relaxation time, which could not be seen with lower stimulation frequencies. Therefore, the authors conclude that DWOLF might be a positive enhancer for re-compensation of muscle contractility after tetanic stimulation by fast “filling” the pool of Ca^{2+} in SR. This activation of SERCA is due to inhibition of PLN, which itself inhibits the SERCA-mediated calcium flow, conclusive to the finding that they act on the same binding site.

By performing further functional analysis, the authors showed that DWOLF is an indirect activator of SERCA. With overexpression of DWOLF, transgenic mice showed an increased influx into SR. Consecutively, the contractility in these mice was significantly higher due to a higher relaxation rate and higher load of Ca^{2+} inside SR.

From the physicians’ point of view, the question remains how this finding could be translated into novel therapies. In their paper, the authors show a downregulation of DWOLF not only in the dilated cardiomyopathy (DCM) mice but also in failing human hearts [11]. Currently, the baseline therapy for patients with a failing heart is to minimize workload and protect it from pathological stress. Combining the findings of the paper, there are two possibilities taking benefit out of these: First, DWOLF could be used as a new biomarker in DCM patients showing the loss of contractility and therefore the severity of the heart failure state in general. This could give physicians the opportunity to optimize therapy before the patient is showing a deteriorated clinical course. Second, DWOLF can be considered as a new drug target. By activating or increasing the amount of DWOLF in heart failing muscle cells, the contractility could be raised by activating SERCA. This would be a new therapeutic way, which would go hand

in hand with the baseline therapy of patients with severe heart failure.

In summary, the search for proteins “hidden in the black space of the genome” will certainly lead to novel surprises as to the complexity of the human heart. Bioinformatics methods are the first stage in this exciting discovery process.

Competing interests

The authors declare that they have no competing interests.

Acknowledgments

DO and JH are supported by the German Centre for Cardiovascular Research (DZHK) and the German Ministry of Education and Research (BMBF).

References

- [1] Bers DM. Calcium fluxes involved in control of cardiac myocyte contraction. *Circ Res* 2000;87:275–81.
- [2] Bravený P. Heart, calcium and time. *Exp Clin Cardiol* 2002;7:3–6.
- [3] Periasamy M, Kalyanasundaram A. SERCA pump isoforms: their role in calcium transport and disease. *Muscle Nerve* 2007;35:430–42.
- [4] Wuytack F, Raeymaekers L, Missiaen L. Molecular physiology of the SERCA and SPCA pumps. *Cell Calcium* 2002;32:279–305.
- [5] Espinoza-Fonseca LM, Autry JM, Ramirez-Salinas GL, Thomas DD. Atomic-level mechanisms for phospholamban regulation of the calcium pump. *Biophys J* 2015;108:1697–708.
- [6] Frank K, Kranias EG. Phospholamban and cardiac contractility. *Ann Med* 2000;32:572–8.
- [7] MacLennan DH, Asahi M, Tupling AR. The regulation of SERCA-type pumps by phospholamban and sarcolipin. *Ann N Y Acad Sci* 2003;986:472–80.
- [8] Fajardo VA, Bombardier E, Vigna C, Devji T, Bloemberg D, Gamu D, et al. Co-expression of SERCA isoforms, phospholamban and sarcolipin in human skeletal muscle fibers. *PLoS One* 2013;8:e84304.
- [9] Shaikh SA, Sahoo SK, Periasamy M. Phospholamban and sarcolipin: Are they functionally redundant or distinct regulators of the Sarco(Endo)Plasmic Reticulum Calcium ATPase? *J Mol Cell Cardiol* 2016;91:81–91.
- [10] Anderson DM, Anderson KM, Chang CL, Makarewich CA, Nelson BR, McAnally JR, et al. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell* 2015;160:595–606.
- [11] Nelson BR, Makarewich CA, Anderson DM, Winders BR, Troupes CD, Wu F, et al. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. *Science* 2016;351:271–5.
- [12] Kung JTY, Colognori D, Lee JT. Long noncoding RNAs: past, present, and future. *Genetics* 2013;193:651–69.
- [13] Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. *Ann Rev Biochem* 2012;81:145–66.
- [14] Cao J. The functional role of long non-coding RNAs and epigenetics. *Biol Proced Online* 2014;16:11.
- [15] Wilusz JE, Sunwoo H, Spector DL. Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev* 2009;23:1494–504.
- [16] Lin MF, Jungreis I, Kellis M. PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. *Bioinformatics* 2011;27:i275–82.
- [17] Schiaffino S, Reggiani C. Fiber types in mammalian skeletal muscles. *Physiol Rev* 2011;91:1447–531.