GENE REGULATION

The origin of neural microexons

Comparative analysis of microexons across bilaterians identifies a new protein domain associated with the evolutionary origin of microexon inclusion in neural tissues.

Rupert Faraway and Jernej Ule

Microexons have played a central role in the evolution of biology's most impressive innovation: the vertebrate brain. They can be as short as three nucleotides and normally lead to the addition of only one or a few amino acids into the encoded proteins, which have diverse cellular functions, including vesicle transport, cytoskeleton organization, and chromatin regulation[1]. Most microexons are exclusively included in neuronal tissues, where they play a crucial role in development and maintenance of neuronal functions[1,2]. Notably, misregulation of activity-dependent splicing of microexons contributes to autism[1,3,4]. Consistent with their functional importance, neural microexons are evolutionarily conserved, yet their evolutionary origin remains poorly understood.

Writing in *Nature Ecology & Evolution*, Torres-Méndez et al.[5] report that the regulatory program that determines the inclusion of microexons in the neuronal transcriptome is at least 600 million years old. Using tissue-specific RNA sequencing, the authors find increased inclusion of microexons in the transcriptomes of neural tissues in various bilaterians, including the non-vertebrate fruit fly (*Drosophila melanogaster*), centipede (*Strigamia maritima*) and amphioxus (*Branchiostoma lanceolatum*). The authors identify several microexons that are shared by all phyla as well as a multitude of sequence features in neural microexons present in vertebrates and non-vertebrates.

Studies in mammals have previously shown that the neuronal protein serine/arginine repetitive matrix protein 4 (SRRM4, also known as nSR100) has a central role in promoting the neuron-specific inclusion of microexons[1,6]. Torres-Méndez et al. identify a previously unannotated domain in the C terminus of SRRM4, the 'enhancer of microexons' (eMIC), that is essential for its function (Fig. 1). Expression of the eMIC-containing C terminus is sufficient to promote microexon inclusion when transfected into a non-neuronal human HEK293 cell line, while deletion of the eMIC domain abrogates the function of SRRM4. The authors identify the interacting partners of multiple SRRM4 domains via affinity-purification-coupled mass spectrometry. Interestingly, the N-terminal end of SRRM4 interacts with members of the exon-junction complex, which have recently been shown to also play an important role in enhancement of microexon inclusion[7]. The eMIC domain, however, primarily interacts with splicing factor 1 (SF1) and the U2 small nuclear RNA auxiliary factor 2 (U2AF2) proteins, two factors that play a role in the initial steps of spliceosome assembly. In vitro splicing experiments have shown that eMIC-containing peptides enhance the formation of earliest stages of splicing complex formation by promoting the recruitment of SF1 and U2AF2.

There are two vertebrate paralogues of the *Srrm4* gene: the broadly expressed

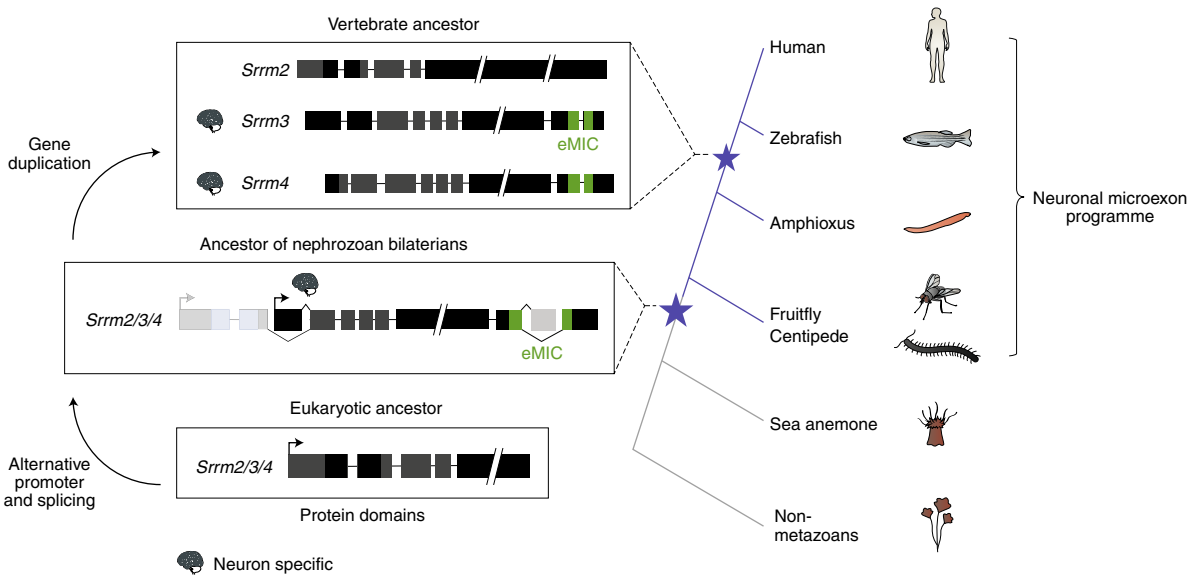

Fig. 1 | The evolutionary origins of neuronal microexons. Torres-Méndez et al. have traced back the evolution of neuronal microexons to a common bilaterian ancestor that existed over 600 million years ago. This neuronal microexon programme is facilitated by the emergence of the previously uncharacterized eMIC domain, the expression of which is restricted to the brain. This tissue specificity is achieved first by neuron-specific alternative splicing and subsequently in vertebrates by gene duplication and specialization.

Srrm2 and the neuron-specific *Srrm3*. Torres-Méndez et al. trace this family back to a single pan-eukaryotic gene referred to as Cwc21 in yeast. The gene lacks the eMIC domain in all non-bilaterians as well as in the most basal, non-nephrozoan bilaterian clade (Xenacoelomorpha), which lacks a brain. The eMIC domain first emerged in the non-vertebrate nephrozoan bilaterians, where it is present in the single gene referred to as *Srrm2/3/4*. Multiple RNA isoforms of the *Srrm2/3/4* gene are produced in non-vertebrates such that the isoforms containing the eMIC domain are expressed only in neuronal tissues. These isoforms are generated through a fascinating combination of alternative promoter usage, splicing and polyadenylation. In vertebrates, however, the domain is primarily expressed in the brain owing to transcriptional control of the *Srrm3* and *Srrm4* genes, both of which contain the eMIC domain (Fig. 1). To demonstrate the universal function of the eMIC domain across evolution, the authors transfected the human HEK293 cell line with the various isoforms of Srrm2/3/4 orthologs from fruit fly and amphioxus, showing that only the isoforms containing the eMIC domain are capable of promoting microexon inclusion. While the *Nematostella* Srrm2 lacks the eMIC domain, the addition of the human eMIC domain from SRRM4 converts it into a protein capable of promoting microexon inclusion.

Torres-Méndez et al. study microexons from a broad evolutionary perspective spanning hundreds of millions of years and employ a detailed biochemical perspective on protein–protein interactions and spliceosomal recruitment. Combined, these approaches reveal not only a much wider role for microexons in the animal kingdom, but also an important lesson about the evolution of gene-regulatory programmes. The eMIC domain appears to have emerged through de novo evolution and is predicted to be intrinsically disordered, which is in line with previous observations that intrinsically disordered proteins and alternatively spliced disordered domains often drive the emergence and inheritance of biological traits[8–10]. It is likely that the neuron-specific program first evolved through regulation of alternative RNA isoforms of *Srrm2/3/4* containing the eMIC domain, which was then consolidated in vertebrates at the level of the genome through duplications of this gene. Ultimately, Torres-Méndez et al. show how small changes in a tissue-specific splicing factor drove the evolution of many tiny microexons, which in turn led to major functional innovations in the brain proteome. ❐

Rupert Faraway[1,2] and Jernej Ule[1,2]*
[1]*The Francis Crick Institute, London, UK.*
[2]*Department for Neuromuscular Diseases, UCL Institute of Neurology, London, UK.*
**e-mail: j.ule@ucl.ac.uk*

Published online: 4 March 2019
https://doi.org/10.1038/s41559-019-0818-1

References

1. Irimia, M. et al. *Cell* **159**, 1511–1523 (2014).
2. Li, Y. I., Sanchez-Pulido, L., Haerty, W. & Ponting, C. P. *Genome Res.* **25**, 1–13 (2015).
3. Quesnel-Vallières, M., Irimia, M., Cordes, S. P. & Blencowe, B. J. *Genes Dev.* **29**, 746–759 (2015).
4. Quesnel-Vallières, M. et al. *Mol. Cell* **64**, 1023–1034 (2016).
5. Torres-Méndez, A. et al. *Nat. Ecol. Evol.* https://doi.org/10.1038/s41559-019-0813-6 (2019).
6. Raj, B. et al. *Mol. Cell* **56**, 90–103 (2014).
7. Gonatopoulos-Pournatzis, T. et al. *Mol. Cell* **72**, 510–524.e12 (2018).
8. Buljan, M. et al. *Mol. Cell* **46**, 871–883 (2012).
9. Ellis, J. D. et al. *Mol. Cell* **46**, 884–892 (2012).
10. Chakrabortee, S. et al. *Cell* **167**, 369–381.e12 (2016).

Competing interests

The authors declare no competing interests.