

# Evidences of somatic retrotransposition in the octopus brain

Volpe M(1)<sup>†</sup>, Petrosino G(2), Simakov O(3), Fiorito G(1), Gustincich S(4,5), Sanges R(1,5)

(1) *Biology and Evolution of Marine Organisms, Stazione Zoologica Anton Dohrn, Napoli, Italy*

(2) *Bioinformatics Core Facility, Institute of Molecular Biology gGmbH (IMB), Mainz, Germany*

(3) *Department of Molecular Evolution and Development, University of Vienna, Vienna, Austria*

(4) *Department of Neuroscience and Brain Technologies, Italian Institute of Technologies (IIT), Genova, Italy*

(5) *Area of Neuroscience, Scuola Internazionale Superiore di Studi Avanzati (SISSA), Trieste, Italy*



<sup>†</sup> Email: [mas.volpe@gmail.com](mailto:mas.volpe@gmail.com)

## Motivation

Transposable elements (TEs) are DNA sequences capable to move from their genomic location to a new one within the genome. They can be divided in classes and families according to the mechanism of transposition and their sequence features [1]. Long interspersed nuclear elements (LINEs) represent a class of retrotransposons which are transcribed into mRNA, are 4-6 Kbp in length and codify for proteins which mediate their retrotransposition [2,3]. They can alter the genomes inducing mutations; however, it is increasingly evident that these elements have played an active role in the evolution since they provide genomic sequences that can be used by the host genome to evolve novelties, for example acting as alternative promoters [4] or enhancers [5]. They were found to be active in the brain of mammals and insects and proposed to be involved in the evolution and development of cognitive abilities [6-8]. Indeed, LINEs retrotransposition has been proposed as a somatic mutational mechanism generating the neuronal genome diversity observed in the brain which could be at the basis of the diversification of neuronal functions [9]. LINEs are present in many copies in their hosts representing the most successful TE in warm-blooded vertebrates [10]. They account for more than 20% of the human genome [11]. Recently, LINE expansion has also been observed in the *Octopus bimaculoides* genome [12]. Octopuses are members of the Mollusca class Cephalopoda and widespread throughout the world's oceans. They present a complex nervous system and are capable to learn, show individual personalities and complex social behaviours [13-15]. For these reasons and for the exceptional complexity of their nervous system, octopuses are considered the most intelligent invertebrates. We sequenced and assembled the *Octopus vulgaris* neural transcriptome discovering a full-length potentially active LINE element. This finding suggested that LINEs could be potentially active in the octopus brain and

led us to search for somatic retrotransposition events.

### Methods

To search for signals of somatic transposition in the octopus, we developed a pipeline to search for full-length TEs presenting an intact open reading frame (ORF) with conserved retrotransposon specific domains. For this purpose, transcriptome of the *Octopus vulgaris* and publicly available *Octopus bimaculoides* genome and transcriptome were annotated for TEs using RepeatMasker [16] and translated in all the six frames collecting the ORFs. The retrotransposon specific domains were detected using both the Conserved Domain Database (CDD)[17] and InterPro [18]. We looked at the transcript abundances of the identified elements by mapping the raw reads on the assembled transcriptome from different tissues using Bowtie [19]. We calculated the age distribution of the selected elements as kimura-2-parameter (K2P)[20] to demonstrate the presence of young fragments for the identified candidates. Young elements, indeed, are more likely to be active with respect to old ones. Finally, to identify non-reference somatic insertions from whole-genome sequencing data of *O. bimaculoides*, we have analysed the candidates performing an insertion-detection analysis using Mobile Element Locator Tool (MELT)[21] on whole genome sequencing data from two different tissues sampled from the same individual: gonads and optic lobe. This final analysis permitted us to find those insertions present in one tissue but absent in the other therefore suggesting the existence of somatic retrotransposition.

### Results

We have found four putative LINE elements in the *O. vulgaris* and seven elements in the *O. bimaculoides* passing stringent search criteria. These elements present a full-length ORF coding for intact and complete retrotransposon specific domains of LINE elements: endonuclease and reverse transcriptase. Age distribution of all the elements is consistent with the possibility that these elements might have been recently active. One element in particular displays a very recent peak of activity showing a very young age for the majority of its annotated fragments in the *O. bimaculoides* genome. This element belongs to the RTE subfamily of LINES and was named RTE-8\_BF6. MELT analysis demonstrates the existence of a significantly higher number of retrotransposition events in the optic lobe with respect to the gonads therefore supporting the hypothesis of the existence of somatic retrotransposition in the octopus brain. The element RTE-8\_BF6 showed to be the candidate with the highest proportion of insertions in the optic lobe with respect to gonads accounting for almost the 92% of the total insertions found. Taken together, our results strongly support the existence of somatic retrotransposition in the *O. bimaculoides* brain.

### References

1. Wicker, T. et al. A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* 8, 973–982 (2007).
2. Feng, Q., Moran, J. V, Kazazian, H. H. & Boeke, J. D. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* 87, 905–16 (1996).
3. Mathias, S. L., Scott, A. F., Kazazian, H. H., Boeke, J. D. & Gabriel, A. Reverse transcriptase encoded by a human transposable element. *Science* 254, 1808–10 (1991).
4. Faulkner, G. J. et al. The regulated retrotransposon transcriptome of mammalian cells. *Nat. Genet.* 41, 563–71 (2009).

5. Makarevitch, I. et al. Transposable Elements Contribute to Activation of Maize Genes in Response to Abiotic Stress. *PLoS Genet.* 11, e1004915 (2015).
6. Baillie, J. K. et al. Somatic retrotransposition alters the genetic landscape of the human brain. *Nature* 479, 534–537 (2011).
7. Perrat, P. N. et al. Transposition-Driven Genomic Heterogeneity in the *Drosophila* Brain. *Science* (80-. ). 340, 91–95 (2013).
8. Coufal, N. G. et al. L1 retrotransposition in human neural progenitor cells. *Nature* 460, 1127–1131 (2009).
9. Singer, T., McConnell, M. J., Marchetto, M. C. N., Coufal, N. G. & Gage, F. H. LINE-1 retrotransposons: mediators of somatic variation in neuronal genomes? *Trends Neurosci.* 33, 345–54 (2010).
10. Abrusán, G., Krambeck, H.-J., Junier, T., Giordano, J. & Warburton, P. E. Biased distributions and decay of long interspersed nuclear elements in the chicken genome. *Genetics* 178, 573–81 (2008).
11. Lander, E. S. et al. Initial sequencing and analysis of the human genome. *Nature* 409, 860–921 (2001).
12. Albertin, C. B. et al. The octopus genome and the evolution of cephalopod neural and morphological novelties. *Nature* 524, 220–224 (2015).
13. Pronk R, W. D. R. and H. R. Video playback demonstrates episodic personality in the gloomy octopus. *J. Exp. Biol.* 213(Pt 7), 1035–41 (2010).
14. C, M. J. A. and A. R. Personalities of Octopuses (*Octopus rubescens*). *J. Comp. Psychol.* Volume 107, 336–340 (1993).
15. Mackie, G. O. *The Anatomy of the Nervous System of Octopus vulgaris*. J. Z. Young. Oxford University Press, New York, 1972. xxxii, 690 pp., illus. 35.25. *Science* (80-. ). 177, 1183–1183 (1972).
16. A.E.A. Smit, R. Hubley & P. Green RepeatMasker at <http://repeatmasker.org>.
17. Marchler-Bauer, A. et al. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.* 45, D200–D203 (2017).
18. Finn, R. D. et al. InterPro in 2017-beyond protein family and domain annotations. *Nucleic Acids Res.* 45, D190–D199 (2017).
19. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, (2009).
20. Kimura, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16, 111–20 (1980).
21. Gardner, E. J. et al. The Mobile Element Locator Tool (MELT): Population-scale mobile element discovery and biology. (2017). doi:10.1101/gr.218032.116