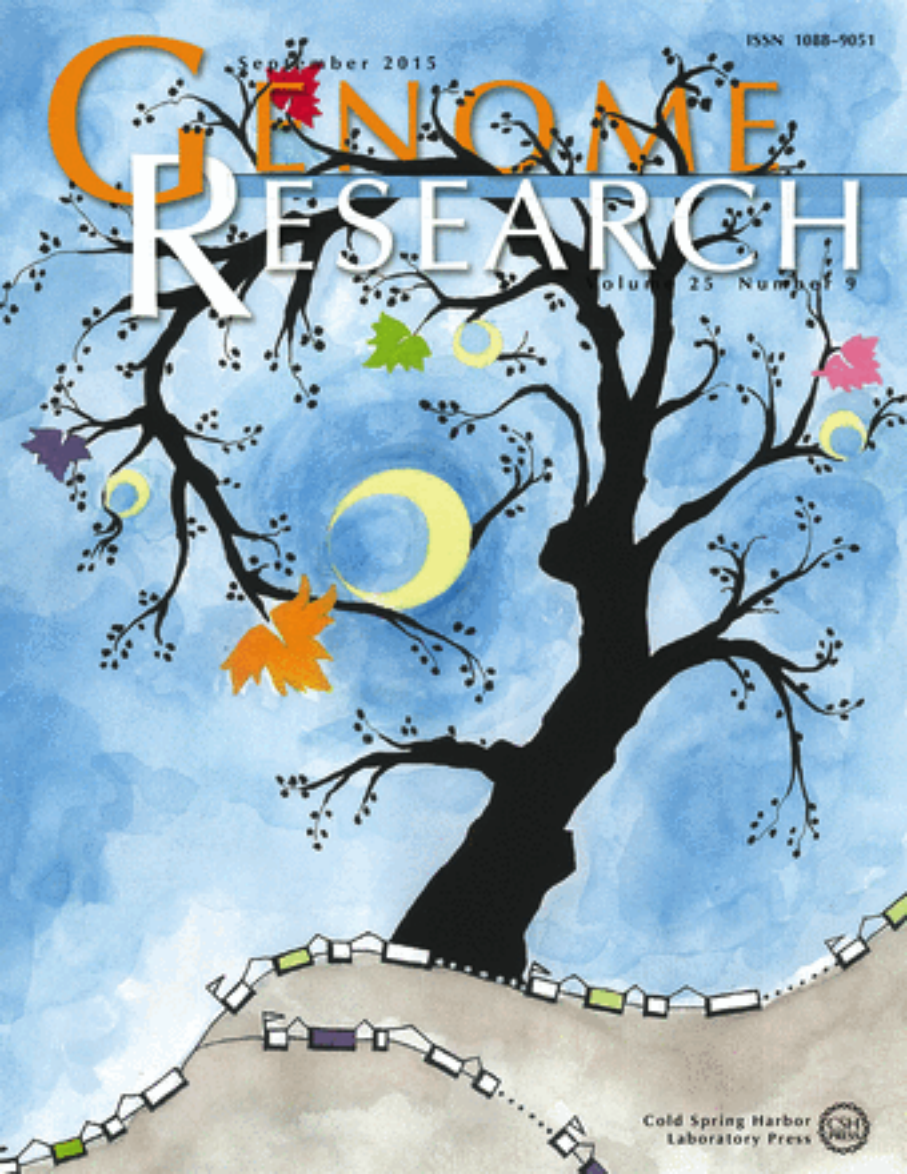
ISSN 1088-9051

September 2015

GENOME RESEARCH

Volume 25 Number 9



Cold Spring Harbor Laboratory Press

Downloaded from genome.cshlp.org on September 4, 2015 - Published by Cold Spring Harbor Laboratory Press

Research

Evolution of selenophosphate synthetases: emergence and relocation of function through independent duplications and recurrent subfunctionalization

Marco Mariotti,[1,2,3,4] Didac Santesmasses,[1,2,3] Salvador Capella-Gutierrez,[1,2] Andrea Mateo,[5] Carme Arnan,[1,2,3] Rory Johnson,[1,2,3] Salvatore D'Aniello,[6] Sun Hee Yim,[4] Vadim N. Gladyshev,[4] Florenci Serras,[5] Montserrat Corominas,[5] Toni Gabaldón,[1,2,7] and Roderic Guigó[1,2,3]

[1]*Bioinformatics and Genomics Programme, Centre for Genomic Regulation (CRG), 08003 Barcelona, Catalonia, Spain;* [2]*Universitat Pompeu Fabra (UPF), 08003 Barcelona, Catalonia, Spain;* [3]*Institut Hospital del Mar d'Investigacions Mèdiques (IMIM), 08003 Barcelona, Catalonia, Spain;* [4]*Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA;* [5]*Departament de Genètica, Facultat de Biologia and Institut de Biomedicina (IBUB) de la Universitat de Barcelona (UB), 08028 Barcelona, Catalonia, Spain;* [6]*Department of Biology and Evolution of Marine Organisms, Stazione Zoologica Anton Dorhn, Villa Comunale, 80121, Napoli, Italy;* [7]*Institució Catalana de Recerca i Estudis Avançats (ICREA), 08010 Barcelona, Catalonia, Spain*

Selenoproteins are proteins that incorporate selenocysteine (Sec), a nonstandard amino acid encoded by UGA, normally a stop codon. Sec synthesis requires the enzyme Selenophosphate synthetase (SPS or SelD), conserved in all prokaryotic and eukaryotic genomes encoding selenoproteins. Here, we study the evolutionary history of *SPS* genes, providing a map of selenoprotein function spanning the whole tree of life. SPS is itself a selenoprotein in many species, although functionally equivalent homologs that replace the Sec site with cysteine (Cys) are common. Many metazoans, however, possess *SPS* genes with substitutions other than Sec or Cys (collectively referred to as *SPS1*). Using complementation assays in fly mutants, we show that these genes share a common function, which appears to be distinct from the synthesis of selenophosphate carried out by the Sec- and Cys- *SPS* genes (termed *SPS2*), and unrelated to Sec synthesis. We show here that *SPS1* genes originated through a number of independent gene duplications from an ancestral metazoan selenoprotein *SPS2* gene that most likely already carried the *SPS1* function. Thus, in *SPS* genes, parallel duplications and subsequent convergent subfunctionalization have resulted in the segregation to different loci of functions initially carried by a single gene. This evolutionary history constitutes a remarkable example of emergence and evolution of gene function, which we have been able to trace thanks to the singular features of *SPS* genes, wherein the amino acid at a single site determines unequivocally protein function and is intertwined to the evolutionary fate of the entire selenoproteome.

[Supplemental material is available for this article.]

Selenoproteins are proteins that incorporate the nonstandard amino acid selenocysteine (Sec or U) in response to the UGA codon. The recoding of UGA, normally a stop codon, to code for Sec is arguably the most outstanding programmed exception to the genetic code. Selenoproteins are found, albeit in small numbers, in organisms across the entire tree of life. Recoding of UGA is mediated by RNA structures within selenoprotein transcripts, the SECIS (SElenoCysteine Insertion Sequence) elements. Sec biosynthesis and insertion also require a dedicated system of *trans*-acting factors that include elements that are common and others that are specific to the three domains of life: bacteria (Kryukov and Gladyshev 2004; Yoshizawa and Böck 2009), archaea (Rother et al. 2001), and eukaryotes (Squires and Berry 2008; Allmang et al. 2009).

The very existence of selenoproteins is puzzling. Sec can apparently be substituted by cysteine (Cys)—as often happens during evolution (Zhang et al. 2006; Chapple and Guigó 2008; Mariotti et al. 2012)—with seemingly a small or null impact on protein function. In fact, selenoproteins may be absent in an entire taxonomic group but present in sister lineages. This can be seen most dramatically within fruit flies: Although *Drosophila melanogaster* and most other flies possess three selenoprotein genes, their relative *Drosophila willistoni* has replaced Sec with Cys in them and lost the capacity to synthesize Sec (Chapple and Guigó 2008; Lobanov et al. 2008). Fungi and plants have also lost this capacity (Lobanov et al. 2009). In other cases, however, such as in *Caenorhabditis elegans*, the entire pathway is maintained only to synthesize a single selenoprotein (Taskov et al. 2005). It appears that selective pressure exists to maintain Sec, at least in vertebrates, since strong purifying selection across Sec sites that prevent mutations to Cys has been reported (Castellano et al. 2009). Sec encoding has been hypothesized to be an ancestral trait, already present in the early genetic code, since a number of selenoprotein families

Corresponding authors: roderic.guigo@crg.cat, marco.mariotti@crg.cat
Article published online before print. Article, supplemental material, and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.190538.115. Freely available online through the *Genome Research* Open Access option.

© 2015 Mariotti et al. This article, published in *Genome Research*, is available under a Creative Commons License (Attribution 4.0 International), as described at http://creativecommons.org/licenses/by/4.0/.

25:1256–1267 Published by Cold Spring Harbor Laboratory Press; ISSN 1088-9051/15; www.genome.org