

ORIGINAL ARTICLE

Reverse transcriptase genes are highly abundant and transcriptionally active in marine plankton assemblages

Magali Lescot¹, Pascal Hingamp¹, Kenji K Kojima², Emilie Villar¹, Sarah Romac^{3,4}, Alaguraj Veluchamy^{5,11}, Martine Boccara⁵, Olivier Jaillon^{6,7,8}, Daniele Iudicone⁹, Chris Bowler⁵, Patrick Wincker^{6,7,8}, Jean-Michel Claverie¹ and Hiroyuki Ogata¹⁰

¹Information Génomique et Structurale, UMR7256, CNRS, Aix-Marseille Université, Institut de Microbiologie de la Méditerranée (FR3479), Parc Scientifique de Luminy, Marseille, France; ²Genetic Information Research Institute, Los Altos, CA, USA; ³CNRS, UMR 7144, team EPEP, Station Biologique de Roscoff, Place Georges Teissier, Roscoff, France; ⁴Sorbonne Universités, UPMC Univ Paris 06, Station Biologique de Roscoff, Place Georges Teissier, FR-Roscoff, France; ⁵Ecole Normale Supérieure, PSL Research University, Institut de Biologie de l'Ecole Normale Supérieure (IBENS), Paris, France; ⁶CEA-Institut de Génomique, GENOSCOPE, Centre National de Séquençage, Evry Cedex, France; ⁷Université d'Evry, Evry Cedex, France; ⁸Centre National de la Recherche Scientifique (CNRS), Evry Cedex, France; ⁹Laboratory of Ecology and Evolution of Plankton, Stazione Zoologica Anton Dohrn, Naples, Italy and ¹⁰Bioinformatics Center, Institute for Chemical Research, Kyoto University, Gokasho, Uji, Kyoto, Japan

Genes encoding reverse transcriptases (RTs) are found in most eukaryotes, often as a component of retrotransposons, as well as in retroviruses and in prokaryotic retroelements. We investigated the abundance, classification and transcriptional status of RTs based on Tara Oceans marine metagenomes and metatranscriptomes encompassing a wide organism size range. Our analyses revealed that RTs predominate large-size fraction metagenomes (> 5 µm), where they reached a maximum of 13.5% of the total gene abundance. Metagenomic RTs were widely distributed across the phylogeny of known RTs, but many belonged to previously uncharacterized clades. Metatranscriptomic RTs showed distinct abundance patterns across samples compared with metagenomic RTs. The relative abundances of viral and bacterial RTs among identified RT sequences were higher in metatranscriptomes than in metagenomes and these sequences were detected in all metatranscriptome size fractions. Overall, these observations suggest an active proliferation of various RT-assisted elements, which could be involved in genome evolution or adaptive processes of plankton assemblage.

The ISME Journal advance online publication, 27 November 2015; doi:10.1038/ismej.2015.192

Introduction

Transposable elements (TEs) have been found in virtually all eukaryotes and 80% of prokaryotes sequenced so far (Hua-Van *et al.*, 2011). They are usually considered as selfish DNA with the capacity to proliferate inside the genome (Doolittle and Sapienza, 1980; Orgel and Crick, 1980). Most organisms have no efficient way of eliminating these

potentially deleterious genetic elements from their genomes, although different mechanisms that silence the activity of TEs are being increasingly revealed (Galagan *et al.*, 2003; Slotkin and Martienssen, 2007; Siomi *et al.*, 2011; Watanabe *et al.*, 2013). In spite of their parasitic nature, TEs also provide beneficial effects on the evolution of their hosts (Dunlap *et al.*, 2006; Slotkin and Martienssen, 2007; Casacuberta and Gonzalez, 2013; Gifford *et al.*, 2013; Riordan and Dupuy, 2013; Bennetzen and Wang, 2014). Characterizing the distribution and classification of TEs is thus important to better evaluate their role in shaping the evolution, structure and function of genomes across the whole tree of life.

TEs are traditionally split into two different classes: Class I TEs (retrotransposons) and Class II TEs (DNA transposons), differing in their mode of transposition (Wicker *et al.*, 2007). Retrotransposons

Correspondence: Dr H Ogata Bioinformatics Center, Institute for Chemical Research, Kyoto University, Gokasho, Uji, Kyoto JP-611-0011, Japan.

E-mail: ogata@kuicr.kyoto-u.ac.jp

¹¹Current address: Biological and Environmental Sciences and Engineering Division, Center for Desert Agriculture, King Abdulah University of Science and Technology, Thuwal, Saudi Arabia. Received 27 April 2015; revised 27 August 2015; accepted 22 September 2015