

# SCIENTIFIC REPORTS



OPEN

## Survey of the green picoalga *Bathycoccus* genomes in the global ocean

Received: 28 April 2016  
Accepted: 03 November 2016  
Published: 30 November 2016

Thomas Vannier<sup>1,2,3</sup>, Jade Leconte<sup>1,2,3</sup>, Yoann Seeleuthner<sup>1,2,3</sup>, Samuel Mondy<sup>1,2,3</sup>, Eric Pelletier<sup>1,2,3</sup>, Jean-Marc Aury<sup>1</sup>, Colomban de Vargas<sup>4</sup>, Michael Sieracki<sup>5</sup>, Daniele Iudicone<sup>6</sup>, Daniel Vaulot<sup>4</sup>, Patrick Wincker<sup>1,2,3</sup> & Olivier Jaillon<sup>1,2,3</sup>

*Bathycoccus* is a cosmopolitan green micro-alga belonging to the Mamiellophyceae, a class of picophytoplankton that contains important contributors to oceanic primary production. A single species of *Bathycoccus* has been described while the existence of two ecotypes has been proposed based on metagenomic data. A genome is available for one strain corresponding to the described phenotype. We report a second genome assembly obtained by a single cell genomics approach corresponding to the second ecotype. The two *Bathycoccus* genomes are divergent enough to be unambiguously distinguishable in whole DNA metagenomic data although they possess identical sequence of the 18S rRNA gene including in the V9 region. Analysis of 122 global ocean whole DNA metagenome samples from the Tara-Oceans expedition reveals that populations of *Bathycoccus* that were previously identified by 18S rRNA V9 metabarcodes are only composed of these two genomes. *Bathycoccus* is relatively abundant and widely distributed in nutrient rich waters. The two genomes rarely co-occur and occupy distinct oceanic niches in particular with respect to depth. Metatranscriptomic data provide evidence for gain or loss of highly expressed genes in some samples, suggesting that the gene repertoire is modulated by environmental conditions.

Phytoplankton, comprising prokaryotes and eukaryotes, contribute to nearly half of the annual global primary production<sup>1</sup>. Picocyanobacteria of the genera *Prochlorococcus* and *Synechococcus* dominate the prokaryotic component<sup>2</sup>. However, small eukaryotes (picoeukaryotes; <2 μm) can be major contributors to primary production<sup>3,4</sup>. In contrast to cyanobacteria, the phylogenetic diversity of eukaryotic phytoplankton is wide, with species belonging to virtually all photosynthetic protist groups<sup>5</sup>. Among them, three genera of green algae belonging to the order Mamiellales (class Mamiellophyceae<sup>6</sup>), *Micromonas*, *Ostreococcus* and *Bathycoccus* are particularly important ecologically because they are found in a wide variety of oceanic ecosystems, from the poles to the tropics<sup>7–12</sup>. The cosmopolitan distribution of these genera raises the questions of their diversity and their adaptation to local environmental conditions. These genera exhibit genetic diversity: for example, there are at least three genetically different clades of *Micromonas* with different habitat preferences<sup>12,13</sup>. One ecotype of *Micromonas* seems to be restricted to polar waters<sup>8,14</sup>. *Ostreococcus* which is the smallest free-living eukaryotic cell known to date with a cell size of 0.8 μm<sup>15</sup> can be differentiated into at least four clades. Two *Ostreococcus* species have been formerly described: *O. tauri* and *O. mediterraneus*<sup>15,16</sup>. Among these *Ostreococcus* clades, different strains seem to be adapted to different light ranges<sup>17</sup>. However, the ecological preferences of *Ostreococcus* strains are probably more complex, implying other environmental parameters such as nutrients and temperature<sup>9</sup>.

The genus *Bathycoccus* was initially isolated at 100 m from the deep chlorophyll maximum (DCM) in the Mediterranean Sea<sup>18</sup> and cells with the same morphology (body scales) had been reported previously from the Atlantic Ocean<sup>19</sup>. *Bathycoccus* has been since found to be widespread in the oceanic environment, in particular in coastal waters<sup>20,21</sup>, and one genome sequence from a coastal strain is available<sup>22</sup>. Metagenomic data have suggested the existence of two *Bathycoccus* ecotypes<sup>10,11,23</sup>, recently named B1 and B2<sup>11</sup>. These two ecotypes have

<sup>1</sup>CEA - Institut de Génomique, GENOSCOPE, 2 rue Gaston Crémieux, 91057 Evry, France. <sup>2</sup>CNRS, UMR 8030, CP5706 Evry, France. <sup>3</sup>Université d'Evry, UMR 8030, CP5706 Evry, France. <sup>4</sup>Sorbonne Universités, UPMC Université Paris 06, CNRS, UMR7144, Station Biologique de Roscoff, 29680 Roscoff, France. <sup>5</sup>National Science Foundation, 4201 Wilson Blvd., Arlington, VA 22230, USA. <sup>6</sup>Stazione Zoologica Anton Dohrn, Villa Comunale, 80121 Naples, Italy. Correspondence and requests for materials should be addressed to P.W. (email: pwincker@genoscope.cns.fr) or O.J. (email: ojaillon@genoscope.cns.fr)