

Insights into global diatom distribution and diversity in the world's ocean

Shruti Malviya^{a,1}, Eleonora Scalco^b, Stéphane Audic^c, Flora Vincent^a, Alaguraj Veluchamy^{a,2}, Julie Poulain^d, Patrick Wincker^{d,e,f}, Daniele Iudicone^b, Colombar de Vargas^c, Lucie Bittner^{a,3}, Adriana Zingone^b, and Chris Bowler^{a,4}

^aInstitut de Biologie de l'École Normale Supérieure, École Normale Supérieure, Paris Sciences et Lettres Research University, CNRS UMR 8197, INSERM U1024, F-75005 Paris, France; ^bStazione Zoologica Anton Dohrn, 80121 Naples, Italy; ^cCNRS, UMR 7144, Station Biologique de Roscoff, 29680 Roscoff, France; ^dInstitut de Génomique, GENOSCOPE, Commissariat à l'Énergie Atomique et aux Énergies Alternatives, 91057 Évry, France; ^eUMR 8030, CNRS, CP5706, 91057 Évry, France; and ^fUMR 8030, Université d'Évry, CP5706, 91057 Évry, France

Edited by Paul G. Falkowski, Rutgers, The State University of New Jersey, New Brunswick, NJ, and approved January 26, 2016 (received for review May 14, 2015)

Diatoms (Bacillariophyta) constitute one of the most diverse and ecologically important groups of phytoplankton. They are considered to be particularly important in nutrient-rich coastal ecosystems and at high latitudes, but considerably less so in the oligotrophic open ocean. The Tara Oceans circumnavigation collected samples from a wide range of oceanic regions using a standardized sampling procedure. Here, a total of ~12 million diatom V9-18S ribosomal DNA (rDNA) ribotypes, derived from 293 size-fractionated plankton communities collected at 46 sampling sites across the global ocean euphotic zone, have been analyzed to explore diatom global diversity and community composition. We provide a new estimate of diversity of marine planktonic diatoms at 4,748 operational taxonomic units (OTUs). Based on the total assigned ribotypes, *Chaetoceros* was the most abundant and diverse genus, followed by *Fragilariopsis*, *Thalassiosira*, and *Corethron*. We found only a few cosmopolitan ribotypes displaying an even distribution across stations and high abundance, many of which could not be assigned with confidence to any known genus. Three distinct communities from South Pacific, Mediterranean, and Southern Ocean waters were identified that share a substantial percentage of ribotypes within them. Sudden drops in diversity were observed at Cape Agulhas, which separates the Indian and Atlantic Oceans, and across the Drake Passage between the Atlantic and Southern Oceans, indicating the importance of these ocean circulation choke points in constraining diatom distribution and diversity. We also observed high diatom diversity in the open ocean, suggesting that diatoms may be more relevant in these oceanic systems than generally considered.

biodiversity | diatoms | metabarcoding | Tara Oceans | choke points

Diatoms are single-celled photosynthetic eukaryotes deemed to be of global significance in biogeochemical cycles and the functioning of aquatic food webs (1–3). They constitute a large component of aquatic biomass, particularly during conspicuous seasonal phytoplankton blooms, and have been estimated to contribute as much as 20% of the total primary production on Earth (4–6). They are widely distributed in almost all aquatic habitats, except the warmest and most hypersaline environments, and can also occur as endosymbionts in dinoflagellates and foraminifers (7).

Because of their complex evolutionary history (8), diatoms have a “mix-and-match genome” (3) that provides them with a range of potentially useful attributes, such as a rigid silicified cell wall, the presence of vacuoles for nutrient storage, fast responses to changes in ambient light, resting stage formation, proton pump-like rhodopsins, ice-binding proteins, and a urea cycle (9). In general, planktonic diatoms seem well-adapted to regimes of intermittent light and nutrient exposure; however, they are particularly common in nutrient-rich regions encompassing polar as well as upwelling and coastal areas (10), highlighting their success in occupying a wide range of ecological niches and biomes. The quantification of diatom diversity and its variations across space (and time) is thus important for understanding fundamental questions of diatom speciation and

their tight coupling with the global silica and carbon cycles (8, 11), as well as for understanding marine ecosystem resilience to human perturbations.

Estimations of the numbers of diatom species vary widely, from a low of 1,800 planktonic species (12) to a high of 200,000 (13). Most recent estimates range from 12,000 to 30,000 species (14, 15). But such global estimates are confounded by the fact that most studies are focused toward understanding the patterns of diversity in a particular diatom genus at a local or regional scale (e.g., refs. 16–18). Furthermore, as evidenced from the Ocean Biogeographic Information System (OBIS) database, although diatom distributions have been explored extensively in numerous studies, they have predominantly focused on the Northern Hemisphere (19, 20).

Characterization of diatom diversity requires accurate and consistent taxon identification. Morphological analyses alone fail to provide a complete description of diatom diversity so complementary investigations are often performed to provide a uniform means of standardization (e.g., ref. 21). During the past decade, the introduction of DNA sequence analysis to systematics has facilitated the discovery of numerous previously undescribed taxa, revealing distinct species identified by subtle or no morphological variations (e.g., ref. 22). Allozyme electrophoresis (23), DNA fingerprinting

Significance

Diatoms, considered one of the most diverse and ecologically important phytoplanktonic groups, contribute around 20% of global primary productivity. They are particularly abundant in nutrient-rich coastal ecosystems and at high latitudes. Here, we have explored the dataset generated by Tara Oceans from a wide range of oceanic regions to characterize diatom diversity patterns on a global scale. We confirm the dominance of diatoms as a major photosynthetic group and identify the most widespread and diverse genera. We also provide a new estimate of marine planktonic diatom diversity and a global view of their distribution in the world's ocean.

Author contributions: S.M., A.Z., and C.B. designed research; S.M., E.S., F.V., and J.P. performed research; S.A., J.P., P.W., and C.d.V. contributed new reagents/analytic tools; S.M., E.S., F.V., A.V., D.I., L.B., and A.Z. analyzed data; S.M. and C.B. wrote the paper; S.A. and C.d.V. provided the eukaryotic v9-18S rDNA metabarcoding dataset; and J.P. and P.W. provided sequencing of the v9-18S rDNA metabarcoding dataset.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

¹Present address: Biological Oceanography Division, National Institute of Oceanography, Dona Paula, Goa 403 004, India.

²Present address: Biological and Environmental Sciences and Engineering Division, Center for Desert Agriculture, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia.

³Present address: Sorbonne Universités, Université Pierre et Marie Curie (UPMC), CNRS, Institut de Biologie Paris-Seine (IBPS), Evolution Paris Seine, F-75005 Paris, France.

⁴To whom correspondence should be addressed. Email: cbowler@biologie.ens.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1509523113/-DCSupplemental.