The ice-binding proteins of *Fragilariopsis cylindrus*: a computational exploration of genetic, structural, and evolutionary features.

Johanna C. Winder[1], Amanda Hopes[1], Cock van Oosterhout[1], Thomas Mock[1]

[1]School of Environmental Sciences, University of East Anglia, Norwich, United Kingdom

The genome of the polar diatom *Fragilariopsis cylindrus* encodes over 47 putative ice-binding proteins (IBPs) containing a domain of unknown function 3494 (DUF3494). The biological role of their diversity has only been minimally explored. Here, a computational approach was used to update the catalogue of IBPs from *F. cylindrus*, predict subcellular localisation and structural motif presence, model structures, and compare sequences with other polar algal species. This analysis reveals structural diversity upstream of the DUF3494 domain in many *fc*IBPs. Several are predicted to be membrane bound, with the definitive presence of one nuclear localisation signal and other suggestions of plastid localisation. Structural modeling with AlphaFold reveals variation in the IBPs. Structural features are sometimes associated with surrounding gene architecture: for example, genes encoding nucleic acid binding proteins flank a triple-domain IBP in the genome. Furthermore, comparing IBP sequences between eukaryotic algae provides evidence for multiple independent origins of IBPs in *F. cylindrus* supporting acquisition of these genes via HGT.