Challenges in building barcode reference libraries for marine invertebrates exemplified by the genus *Prionospio* (Annelida: Spionidae)

Martin Mejdell Hektoen†,§, Torbjørn Ekrem†, Torkild Bakken‡

† NTNU University Museum, Trondheim, Norway
§ Åkerblå AS, Trondheim, Norway

Corresponding author: Martin Mejdell Hektoen (koffertfisk@gmail.com)

Received: 22 Feb 2021 | Published: 04 Mar 2021


Abstract

The completeness of reference libraries is often a limiting factor in the effectiveness of biomonitoring using molecular tools. The fact that these libraries are often built upon Sanger sequencing can create a substantial bias due to poor primer fits and unoptimized lab protocols. Some taxa of marine macroinvertebrates are known to be notoriously difficult to sequence using traditional, PCR-based means, and only about 15% of the known bioindicator species world-wide have publicly available sequences for any genetic marker (Aylagas et al. 2014). The Barcode of Life Data System (BOLD) indicates an amplification rate between 46% and 85% of the barcode region of COI for the most commonly used marine invertebrates that indicate pollution in the North East Atlantic (Capitellidae, Cirratulidae, Dorvilleidae, Spionidae and Tubificidae within Annelida, and Thyasiridae within Mollusca). A currently on-going integrative taxonomic study on *Prionospio* Malmgren, 1867 (Spionidae, Annelida) exemplifies the extensive issues of utilizing Sanger sequencing on marine invertebrates. The barcode region of COI was attempted amplified using five primer pairs: three designed to be universal for marine invertebrates (Folmer et al. 1994, Geller et al. 2013, Lobo et al. 2013), one specialized on polychaetes (Carr et al. 2011) and one self-designed. In addition, two DNA polymerases were tested (TaKaRa Ex Taq HS and Qiagen HotStarTaq) and three annealing temperatures. Only five sequences of COI were...
obtained from a total of 255 PCR reactions (2% success rate). Other genetic markers showed better amplification rates: 58% for 16S rDNA, and more than 90% success rates for 28S rDNA and Histone H3. This illustrates the importance of having more than one marker in mind when seeking to complete reference libraries, and the potential effectiveness of a multi-marker approach in molecular biomonitoring surveys. As sequencing costs decrease, utilizing shallow shotgun-based sequencing (genome skimming) on problematic groups such as *Prionospio* to bypass issues regarding unfit primers is also becoming a viable option.

**Keywords**

Reference libraries, Unfit primers, Polychaeta, Spionidae

**Presenting author**

Martin Mejdell Hektoen

**Presented at**

1st DNAQUA International Conference (March 9-11, 2021)

**References**