

Review

<https://doi.org/10.7850/jkso.2021.26.3.238>  
pISSN : 1226-2978 eISSN : 2671-8820

## 돌말류 분석 자료 정도 관리의 현황 및 전망

김혜숙<sup>1</sup> · 김종성<sup>2</sup> · 박진순<sup>3\*</sup><sup>1</sup>한국해양대학교 해양과학기술전문대학원 해양과학기술융합학과 대학원생, <sup>2</sup>서울대학교 자연과학대학 지구환경학부 및 해양연구소 교수, <sup>3</sup>한국해양대학교 해양환경학과 및 해양과학기술전문대학원 해양과학기술융합학과 조교수

# Current Status and Prospects for the Data Quality Control in Terms of Diatom Analysis

HYESUK KIM<sup>1</sup>, JONGSEONG KHIM<sup>2</sup> AND JINSOON PARK<sup>3\*</sup><sup>1</sup>Graduate Student, Department of Convergence on the Ocean Science and Technology, Korea Maritime and Ocean University, Busan 49112, Korea<sup>2</sup>Professor, School of Earth and Environmental Sciences and Research Institute of Oceanography, Seoul National University, Seoul 08826, Korea<sup>3</sup>Assistant Professor, Department of Ocean Science and Department of Convergence on the Ocean Science and Technology, Korea Maritime and Ocean University, Busan 49112, Korea\*Corresponding author: [jpark@kmou.ac.kr](mailto:jpark@kmou.ac.kr)

Editor Eun-Jin Yang

Received 20 June 2021; Revised 24 August 2021; Accepted 25 August 2021

### ABSTRACT

돌말류는 미세조류의 일종으로서 담수 및 해양의 여러 다양한 환경의 수생태계에서 가장 중요한 1차 생산자 중 하나이다. 환경을 대표할 수 있는 생물 지표로서 널리 이용되고 있으며, 여러 지역에서 서로 다른 과학자들에 의한 연구 결과들이 과학적 통일성과 객관성을 갖추도록 하기 위해서는 돌말류 분석 자료의 정도 관리가 매우 중요하다. 현재 돌말류 분석은 형태적인 특징으로 분류하는 형태 기반의 분석과 DNA 서열로 종을 식별하는 DNA 바코드 분석이 사용되고 있다. 형태학 기반의 돌말류 분류는 분석 자료를 해석하는 분류학자들 간의 일관된 종 식별이 요구되는 한편 분자 분류 기반의 돌말류 분류는 신뢰성 있는 참조데이터가 필수적이다. 본 논문에서는 돌말류 분석 자료 정도관리의 국내외 현황을 우선 살펴보았으며, 그에 기반하여 국내 돌말류 분석 자료의 정도 관리에 대한 의견을 또한 제안하였다.

Diatoms are a type of microalgae which inhabits a wide variety of environments as one of the most important primary producers of freshwater and marine ecosystems. They have been widely used as bioindicators which represent the environmental characteristics, thus proper quality control of diatom data is very important to ensure for the researches from many scientists from various different regions to have scientific unity and objectivity. At present, diatom data analysis is primarily based both on morphological features and DNA sequences of given species. Challenge for the morphology-based analysis of diatoms is the consistent species identification among different taxonomists who interpret diatom community, while challenge for the molecular analysis of diatoms to secure reliable reference data. In the present study, we have reviewed the current status of data quality control of diatom analysis in Korea as well as the world. Finally, suggestions for the better data quality control for Korean marine diatoms have been also made.

**Keywords:** Diatom, Data, Quality control, Current status, Prospects

## 1. 서론

돌말류는 단세포 진핵생물로서(Falciatore and Bowler, 2002; Mann *et al.*, 2017), 규산질의 세포벽을 갖는 중요 미세조류 군집이다(Valentin *et al.*, 2019). 이들은 젖은 토양, 담수, 해양에 이르기까지 매우 다양한 수환경에 서식하며(Round *et al.*,

1990; Evans *et al.*, 2007), 광합성이 가능한 엽록소를 가지고 있어 수생태계에서 중요한 1차 생산자 중 하나이다. 돌말류는 전 지구적으로 약 20%의 1차 생산력에 기여할(Falkowski *et al.*, 1998; Mann, 1999) 뿐 아니라, 해양에서는 유기 탄소의 약 40%를 생산하는(Falkowski *et al.*, 1998) 것으로 알려져 있다. 또한 해양에서의 생지화학적 요소 순환에 큰 역할을 하며 (Tréguer *et al.*, 1995; Yool and Tyrrell, 2003; Kim *et al.*, 2020), 다양한 환경에서 서식하기 때문에 환경을 대표할 수 있는 생물 지표로써 이용된다(Underwood and Kromkamp, 1999; Park and Koh, 2012; Im *et al.*, 2020).

돌말류 연구의 역사는 18세기 초, *Tabellaria*속으로 여겨지는 돌말류 종이 삽화와 함께 보고되며 시작되었다(Anonymous, 1703; Mann *et al.*, 2010). 18세기 말 무렵이 되어서야 최초의 돌말류 속인 *Bacillaria*가 명명되어 출판되었고(Gmelin, 1791), 19세기까지 유럽의 학자들에 의해서 많은 수의 속과 종이 제안되었다(Fourtanier and Kocielek, 1999). 20세기에 접어들면서 돌말류가 수질 환경을 잘 반영할 수 있는 생물지표로 알려지면서 유럽과 북미에서 돌말류 연구가 크게 증가하였다. 특히 돌말류를 활용한 하천의 수질 평가 연구가 주목할만하다(Lavoie *et al.*, 2009). 우리나라의 경우 20세기 초가 되어서야 러시아 학자인 Skvortzow(1929)에 의해 돌말류가 최초 보고되었다. 해방 이후 국내 학자들에 의해 돌말류 연구가 꾸준히 진행되었지만, 한국 돌말류 특히 해양 돌말류의 분류와 관련하여는 연구자 및 관련 논문의 수가 많지 않은 것이 현실이다(Park *et al.*, 2014; Kim *et al.*, 2020). 상대적으로 돌말류의 역사가 깊은 유럽과 북미의 연구결과를 무비판적으로 수용한 것 역시 한국 해양 돌말류 연구, 특히 그 중에서도 분류학의 연구가 상대적으로 미진한 이유로 여겨진다.

돌말류 종의 정확한 동정은 돌말류를 활용한 다양한 연구에서 필요한 작업이다. 돌말류의 복잡하고 섬세한 규산질 세포벽은 형태가 매우 다양해 속 또는 종을 구별하는데 유용하게 사용된다. 형태관찰에 있어 주로 광학현미경을 사용하는데, 이는 시료 내 돌말류 군집의 종다양성을 확인할 수 있는 실용적인 방법으로서 18세기부터 현재까지 약 200년간의 긴 시간 동안 활발하게 이용되고 있다(Valentin *et al.*, 2019). 20세기 초반에는 투과전자현미경이, 이어서 20세기 중반에 이르러 주사전자현미경과 같은 시료의 고배율 관찰이 가능한 전자현미경이 연구에 활용되면서(Kawecka and Olech, 1993) 보다 정밀한 돌말류 미세구조의 관찰이 가능하여 졌으며, 이를 통해서 보다 정확한 종의 식별이 가능해졌을 뿐 아니라 돌말류 분류체계를 재정립하는 계기가 되기도 하였다(Round *et al.*, 1990). 그러나 다른 한편으로는 형태분류학의 세계적인 쇠퇴 경향과 맞물려 훈련된 돌말류 분류학자들의 수가 감소하는 것 역시 현실이며, 그에 따라 분류 경험이 부족한 생태학자들의 부정확한 종 동정으로 인해 돌말류 연구에 있어서 과학적 오류가 발생하고 있다(Archibald, 1984). 이러한 이유로 형태분류학자의 현미경을 통한 동정을 거치지 않고서-달리 말해 훈련된 분류학자 없이-돌말류 종을 동정할 수 있는 방법이 고안되었다. 이는 DNA 마커를 활용한 DNA 바코드 분석법으로, 최근 들어 분자생물학적인 돌말류 동정과 관련된 연구가 활발하게 진행되고 있다(Hebert *et al.*, 2003; Sogin *et al.*, 2006; Valentini *et al.*, 2009).

다음 장에서는 돌말류 분류학자들이 종 동정과 분류를 위해 사용하는 형태기반의 분류 및 DNA 바코드 기반의 분류 방법에 대해 논하고자 한다.

## 2. 돌말류 분류학의 두 가지 측면

### 2.1 형태학 기반 돌말류 분류

18세기 초 처음으로 돌말류가 발견된(Mann *et al.*, 2010) 이후 오랜 시간 동안 형태적인 특징에 오롯이 기반하여 돌말류 분류가 수행되었다. 돌말류를 이루는 규산질의 세포벽은 복잡하면서 다양한 형태를 가지고 있어, 종을 식별할 수 있는 형태적 기준으로 이용되었다. 초기의 돌말류 분류 연구(Kützing, 1833, 1844; Ehrenberg, 1838; Smith, 1853, 1856; Gregory, 1857;

Greville, 1857)는 광학 현미경으로 관찰한 개체의 형태적 특징을 글로 기재하는 동시에 관찰자가 종의 특징을 그림으로 표현하여 보고하는 방식으로 수행되었다. 그러나 이러한 연구들은 개체의 그림이 기재의 설명을 모두 담아낼 만큼 세밀하지 못하거나(Mann *et al.*, 2010), 그림에 연구자의 주관적 관점이 반영되거나, 또는 심지어 좋지 않은 인쇄 품질 등으로 인해 분류 및 동정에 있어 신뢰성을 담보하지 못하는 경우가 많았다. 광학현미경의 지속적인 발달, 그리고 20세기 중반 이후 전자현미경의 개발과 활용은 이러한 문제점을 획기적으로 개선하게 되었다. 예를 들어 1980년대의 여러 연구들(Crawford, 1981; Morrow *et al.*, 1981; Cox, 1987)에서는 높은 해상도의 광학현미경 사진들 뿐 아니라 수만 배에 이르는 매우 높은 배율의 주사전자현미경 사진을 쉽게 찾아볼 수 있다. 이러한 현미경 장비의 발달은 돌말류 분류를 크게 발전시키는 계기가 되었고, 19세기 초 약 30개였던 돌말류 속의 수는 20세기 말 무렵에 1000개까지 늘어났다(Fourtanier and Kocielek, 1999). 현재까지도 매년 새로운 돌말류 종이 보고되고 있어 그 수가 증가하고 있다.

형태적 특성을 기반한 전통적인 종의 식별은 경험이 풍부한 분류학자를 필요로 하며(Pečnikar and Buzan, 2014), 독립적으로 종 동정이 가능한 정도의 지식과 경험을 갖추는 데에는 오랜 훈련 기간이 소요된다(Coleman and Radulovici, 2020). 이와 관련하여 최소 석사 학위 이상, 실제로는 많은 경우 박사 학위 이상의 전문적인 교육이 필수적이다. 세계적으로 경험이 풍부한 돌말류 분류학자의 수는 점차 줄어들고 있는데, 이러한 분류학자 수의 감소는 돌말류 분류군에 국한된 것이 아니다. 분류학 전반에서 분류학자의 감소 현상이 나타나고 있는데(Mora *et al.*, 2011; Coleman and Radulovici, 2020), 이는 정확한 종 식별을 위해서 분류학자의 양성이 필요함을 시사한다.

경험이 많고 잘 훈련된 분류학자라 할지라도 종을 구분하는 기준을 달리 사용하거나, 종의 설명을 다르게 해석하는 경우에 종을 오동정 할 수 있다(Mann *et al.*, 2010). 또한 형태 기반 분류법은 특정 생애 단계 혹은 성별을 설명하는 것에만 효과적일 수 있으므로(Valentini *et al.*, 2009) 접합자나 휴면세포 등 특정 단계의 개체들은 분석 과정 중에 무시될 수 있다. 한편 돌말류는 약 200,000종(Mann *et al.*, 2010)이 존재할 것으로 예상되는데, 이 숫자에는 소위 은닉종(cryptic species)역시 포함된다. 은닉종은 형태학적으로 동일하지만 유전적으로는 구별되는 종을 의미하는데(Amato and Montresor, 2008), 형태학적으로는 차이가 없기 때문에 분류 전문가라 할지라도 이를 구분하는 것은 매우 어려운 일이다. 또한 실용적인 관점에서 광학 현미경으로 많은 양의 돌말류 개체를 동정하고 계수하는 것은 높은 수준의 분류학적 지식을 필요로 할 뿐 아니라 매우 긴 시간을 요구하는 작업이기도 하다(Rimet *et al.*, 2016). 이러한 모든 이유들로 인해 분자생물학적 분석을 활용하여 돌말류 자료를 분석하고자 하는 시도가 활발하게 이루어지고 있다.

## 2.2 분자생물학 기반 돌말류 분류

형태학 기반 분류법의 한계를 보완하고, 보다 빠른 시간 내에 종을 식별하는 방법으로 각광받는 것이 DNA 바코드 분석법이다. DNA 바코드 분석은 DNA 염기서열 상의 작은 차이가 종의 생물학적 분리와 일치한다는 전제를 기반하며(Moniz and Kaczmarek, 2009), 세포 내에 존재하는 핵 또는 세포 소기관 게놈의 짧은 영역을 사용해서 종을 구별하는 분석법이다(Hebert *et al.*, 2003; Mann *et al.*, 2010). DNA 서열의 가용성이 늘어나면서 DNA 바코드는 기본 및 응용연구에서 널리 시도되고 있을 뿐 아니라(Evans *et al.*, 2007) 조류(Hebert *et al.*, 2004)와 어류(Ward *et al.*, 2005; Zemlak *et al.*, 2009) 등 많은 분류군에서 이 방법을 적극 활용하고 있다.

DNA 바코드는 형태학적 특성과 달리 DNA 서열 사이의 유전적 거리로 종을 구별하기 때문에 분류학자에게 의존하지 않고서 종을 식별할 수 있는 방법이다(Zimmermann *et al.*, 2015). 돌말류 연구의 경우 분자분류에 주로 사용되는 DNA 마커에는 ribosomal DNA인 SSU (Small subunit) rDNA (Sato *et al.*, 2008; Zimmermann *et al.*, 2011; Luddington *et al.*, 2012),

LSU (Large subunit) rDNA (Hamsher *et al.*, 2011)와 내부 전사 스페이스(Internal Transcribed Spacers; ITS, Meiklejohn *et al.*, 2019), 미토콘드리아의 유전자인 *cox1* (Evans *et al.*, 2007, 2008), 그리고 엽록체의 유전자인 *rbcL*이 사용된다(Hamsher *et al.*, 2011; MacGillivray and Kaczmarska 2011; Kermarrec *et al.*, 2013).

일반적으로 다세포 생물의 경우 작은 조직으로도 DNA 염기서열을 비교적 쉽게 확인할 수 있지만, 단세포인 돌말류의 경우 개체를 죽이지 않고서 DNA 염기서열을 확인하는 것이 불가능하다. 그러므로 돌말류의 DNA 염기서열을 확인하기 위해서는 단종배양이 선행되어야 하는데, 이와 관련하여 몇 가지 문제점이 존재한다. 첫째로 코어자료 분석의 경우이다. 돌말류는 현재뿐 아니라 과거의 환경을 연구하는 데에도 널리 이용되고 있으므로 많은 경우 살아 있는 돌말류가 아닌 죽은 돌말류(화석종, fossil diatoms)가 남긴 세포벽에 기반하여 연구가 진행되게 된다. 이 때에는 DNA를 전혀 확보할 수 없으므로 현미경을 활용한 분석이 유일한 선택일 수밖에 없다. 두 번째로는 살아 있는 돌말류 군집에 있어 배양이 되는 종의 수 보다는 배양이 되지 않는 종의 수가 더 많다는 점이다. 잠재적으로 배양이 가능한 종의 경우에도 해당 서식지에 매우 적은 수의 개체만이 존재하는 경우에는 실제 배양 성공 확률은 매우 낮아지게 된다. 배양이 가능한 종들에 있어서도 기술적인 어려움은 여전히 존재한다. 특히 저서돌말류의 경우 많은 경우 그 크기가 작을 뿐 아니라 기질에 단단하게 부착하고 있어 단종으로 분리하는 것이 쉽지 않다. 일단 단종으로 분리한 후에는 형태적 특징을 기반으로 분리된 종의 정확한 종 동정이 필요한데, 이 과정에서 오동정이 이루어지면 해당 DNA 염기서열이 엉뚱한 종명으로 GenBank와 같은 DNA 염기서열 데이터베이스에 업데이트 될 수 있다. 실제로 그러한 사례가 많이 있으며 이러한 데이터 오류들은 학자들 사이에서 ‘garbage data’로 불리고 있다. 이처럼 DNA 바코드에 기반한 돌말류 분류의 정확성을 위해서는 종 식별에 이용되는 참조데이터의 신뢰성이 중요한데, 이는 다소 역설적인 상황을 야기한다. 형태분류의 한계와 비효율성, 그리고 형태분류학자의 모자란 숫자를 극복하기 위해 분자분류를 도입하였으나, 분자분류의 효율적인 적용을 위해서는 잘 훈련된 형태분류학자의 역할이 필수적인 것이 현실이다.

### 3. 돌말류 분석자료 정돈 관리의 국내외 현황

#### 3.1 국외에서의 형태 분류 자료의 일관성 확보 노력

형태 분류에서 분류학자들 간의 서로 다른 종 식별의 오류를 줄이면서, 분석 자료의 일관성을 갖추기 위해서는 공통적인 방식의 종 식별을 지속적으로 공유해야 한다. 돌말류 연구의 역사가 깊은 유럽의 경우 European Diatom Meeting (EDM) 등의 국제학회 개최를 통해 연구자들 간의 교류가 활발하게 이루어지고 있다. 이를 통해 최신 연구 결과를 공유할 뿐 아니라 분류체계의 통일성을 담보하고 있다. 일례로 2021년 3월 개최된 제 13회 EDM에서는 *Fragilaria* 속에 대한 전문가 워크숍을 가졌으며 해당 속의 모든 종을 문서화하여 논의하였다. 또한 Type slides와 유럽 각지에서 채집된 종을 광학현미경으로 함께 관찰하는 시간을 가졌다. 이처럼 특정 돌말류 속에 대한 해당 속에 대한 문헌을 함께 살펴보고, 이에 대해 논의를 함으로써 보다 정확한 종의 특징을 함께 공유할 수 있다.

마찬가지로 돌말류 연구가 활발한 북미에서는 Society for Freshwater Science (SFS)에서 관련 워크숍을 진행하고 있다. 2015년 5월에 개최한 SFS의 돌말류 분류 인증 워크숍에서는 돌말류 분류와 동정에 있어 종의 특징을 설명하는 방식, 원전의 설명을 얻는 방법, 그리고 돌말류 시료의 광학현미경 관찰 방법 등을 교육하였다. 해당 프로그램을 통해 돌말류 종 관련 정보의 수집과 활용 방법을 공유할 뿐 아니라, 참가자들을 대상으로 한 테스트를 실시하여 통과한 이들에 한해 분류학 인증서를 발급하고 있다. 해당 인증서의 유효기간은 5년으로, 주기적으로 인증 프로그램 수료 및 테스트를 받아야 한다.

이러한 모임이 이뤄질 수 있었던 배경에는 유럽과 북미에서 돌말류를 담수 오염의 생물 지표로 널리 이용하고 있기 때문으

로, 해당 업무의 수행을 위해서는 돌말류의 정확한 종 식별이 필수적이기 때문이다. 돌말류와 관련된 공통된 지식을 공유하고, 규범화된 분류 방법을 지속적으로 유지해야 하는 이유이기도 하다.

### 3.2 국외의 데이터베이스

앞서 언급한 바와 같이 DNA 바코드 연구에서 중요한 것은 본인의 연구 결과로 획득한 자료와 비교할 수 있는 참조 데이터가 필요하다는 것이다. 참조 가능한 데이터는 어디서 구해야 하는가? 이에 대한 해답으로 다음과 같은 웹 기반의 DNA 서열 데이터를 제공하는 미국의 GenBank와 캐나다의 The Barcode of Life Data System (BOLD), 독일의 SILVA 데이터베이스를 예로 들 수 있다(Table 1).

GenBank는 분자생물학을 연구하는 연구자들이라면 모를 수 없을 만큼 매우 중요한 DNA 서열 데이터 베이스 중 하나이다. GenBank는 National Center for Biotechnology Information (NCBI)에서 구축한 데이터베이스로(Benson *et al.*, 2018), 이 데이터베이스는 많은 연구자들이 자신의 연구 데이터와 DNA 서열을 비교하거나(Harris, 2003), 종의 변동 수준을 확인하기 위해 기준에 발표된 데이터를 재분석 할 때(Johns and Avise, 1998; Harris, 2003) 필요한 참조 데이터를 제공한다. 그러나 GenBank는 공개 데이터베이스 시스템이기 때문에 DNA 염기 서열 데이터라면 검증 단계를 거치지 않고, 연구자 누구나 쉽게 DNA 염기 서열을 업로드 할 수 있다는 점에서 문제가 발생할 수 있다. 논문 저자가 생성한 DNA 서열 데이터와 기준에 등록된 GenBank의 모든 데이터가 항상 옳은 것은 아니기 때문이다. 이미 많은 연구들에서 Genbank의 데이터 신뢰성에 대해서 우려하고 있다(Harris, 2003; Wesche *et al.*, 2004; Shen *et al.*, 2013). 또한 GenBank는 형태학적 정보 및 환경 요소에 대한 정보를 제공하지 못한다는 점에서 돌말류 연구의 활용에 한계점을 가진다.

최근에 GenBank 다음으로 DNA 참조 서열 데이터베이스로 많이 언급되는 DNA 서열 데이터베이스는 BOLD이다. BOLD의 주된 목적은 DNA 바코드 시퀀스와 원본 표본간의 지속적인 연결이며, 이러한 연결은 데이터의 저장, 구성 및 요청과 관련된 모든 보안환경을 포함한다(Ratnasingham and Hebert, 2007). BOLD는 DNA 염기서열만 업데이트 되는 GenBank와는 달리, 실제 표본에 대한 메타데이터에 대한 정보(표본의 명명자, 표본이 보관된 기관명, 표본 이미지, DNA 염기서열 등) 제공을 권장하고 있으며(Meiklejohn *et al.*, 2019; Coleman and Radulovici, 2020), 그에 따라 BOLD는 상대적으로 GenBank보다 신뢰성은 더 높다고 하겠다. BOLD에서는 돌말류 개체의 형태를 사진으로 확인 할 수 있지만, 이들의 자세한 묘사가 누락된 점이 아쉽다. 그리고 종에 따라서 채집된 장소 혹은 종의 분포에 대한 정보도 함께 제공하고 있다.

**Table 1.** Overview of the international diatom databases

	photographic data	species description	ecological information	DNA information	note
GenBank (America)	X	X	X	SSU <sup>a</sup> , LSU <sup>b</sup> , COX <sup>c</sup> , etc	lots of DNA sequencing information, yet DNA sequences unverified
BOLD (Canada)	O	X	distribution and/or collection site	SSU <sup>a</sup> , LSU <sup>b</sup> , COX <sup>c</sup> , etc	low number of DNA bases, but with verified information
Algaterra (Germany)	O	O	environment variables	SSU <sup>a</sup> , LSU <sup>b</sup> , COX <sup>c</sup> , etc	specialized in microalgae
SILVA (Germany)	X	X	X	SSU <sup>a</sup> , LSU <sup>b</sup>	specialized in rDNA

<sup>a</sup>SSU : Small subunit rDNA

<sup>b</sup>LSU : Large subunit rDNA

<sup>c</sup>COX : Cytochrome c oxidase subunit

BOLD와 비견할 또 다른 사례로는 독일의 SILVA 데이터베이스가 있다(Table 1). SILVA는 세계 유일의 웹 기반의 ribosomal RNA 데이터베이스로(Yilmaz *et al.*, 2014; Glöckner, 2019), 여러 종류의 DNA 마커들 중 ribosomal DNA인 SSU와 LSU에 대한 정보를 제공한다(Quast *et al.*, 2012). 또한 SILVA는 크게 3개의 대분류군(박테리아, 고세균, 진핵생물)으로 정리되어 있고, 각 분류군에 맞는 다양한 방법의 분류체계(EMBL-EBM/ENA, RDP와 GTDB)를 이용하여 데이터베이스화 되어 있다(Glöckner, 2019). 그러나 종명 및 속명에 기반한 검색을 제공하고 있지 않기에 DNA 염기서열 정보의 접근성은 다소 떨어진다고 볼 수 있다. 다시 말해 해당 종의 분류체계를 명확히 이해하지 못한다면 원하는 DNA 염기서열을 빠르게 찾는 것이 쉽지 않다. 더불어 형태적 정보와 생태적 정보를 제공하지 않는 부분 역시 아쉽다.

앞에서 언급한 3개의 데이터베이스는 동, 식물을 아우르는 다양한 분류군에 대한 DNA 염기 서열을 제공하는 데이터베이스였다면, 마지막으로 소개하는 AlgaTerra는 돌말류를 포함한 담수와 해양의 미세조류에 특화된 데이터베이스이다(Table 1). AlgaTerra는 독일의 Botanischer Garten & Botanisches Museum Berlin (BGBM)에서 개발한 데이터베이스로, 2004년부터 온라인으로 운영되고 있다(Kusber and Jahn, 2007). 미세조류 생물다양성 정보시스템인 AlgaTerra는 2007년을 기준하여 24,650개 이상의 조류의 종명을 포함해 조류의 개념, 형태학적 정보 및 기재, 분자 데이터와 참고문헌 데이터를 함께 제공하는(Kusber and Jahn, 2007) 종합적인 데이터베이스이다. AlgaTerra는 DNA 염기서열뿐만 아니라 형태학적 정보도 모두 제공하고 있어서, 다른 데이터베이스에 비해서 좀 더 신뢰성을 갖는다고 볼 수 있다. 이처럼 잘 구축된 데이터베이스인 AlgaTerra에 한가지 아쉬운 점이 있다면, 해양의 돌말류보다 담수의 돌말류 정보가 더 많다는 점이다. 이는 유럽의 담수 돌말류를 활용한 수질관리 프로그램으로 인해서 돌말류의 관심도 및 중요도가 해양보다는 담수 쪽에 더 큰 비중을 두고 있기 때문으로 여겨진다.

### 3.3 국내의 형태 분류와 데이터베이스의 현황

국내에서의 돌말류 분류 연구는 상대적으로 소수 그룹에서 진행되어 왔다. 이러한 소수 그룹 중 한 곳인 부경대학교의 허성범 교수 연구실에서는 돌말류를 포함하여 미세조류를 전문적으로 분리, 배양하는 기관인 '해양미세조류은행'을 설립한 바 있다. 해양미세조류은행에서는 돌말류를 포함하여 2005년 기준 1,162종(이중 598종이 해양 돌말류였음)의 미세조류를 분리하여 배양하는 데 성공했다([https://www.ibric.org/myboard/read.php?Board=bio\\_plan&id=79](https://www.ibric.org/myboard/read.php?Board=bio_plan&id=79)). 더불어 분리 및 배양으로 확보하여 보유한 미세조류 종에 대한 데이터베이스까지 구축하였다. 이후 해양미세조류은행은 대표인 허성범 교수의 퇴임과 함께 2015년 해양수산부 산하 한국해양과학기술원으로 이전되었다. 한국해양과학기술원으로 이전된 해양미세조류은행의 시료들은 현재 해양시료도서관에서 배양되고 있다.

해양시료도서관(<http://lims.kiost.ac/main/main.do>)은 2012년 6월에 개관하여, 해저 퇴적물, 해양 광물, 해양생물 및 플랑크톤 배양주의 체계적인 수집, 저장 및 관리를 위해서 설립되었다. 해양시료 도서관에서는 시료현황에 대한 데이터베이스를 운영하고 있다. 여기에는 배양 생물의 관리코드, 시료명, 분류군(학명), 채집일자 및 위치, 배양조건에 대한 내용이 기록되어 있다. 해양시료 도서관 데이터베이스의 돌말류 시료 수는 728개로 확인되지만, 이 중에는 중복되는 종이 많고, 종 수준까지 동정된 종이 많지 않다. 그리고 오로지 시료를 관리하기 위한 데이터베이스이기 때문에, 형태학 또는 분자생물학적 정보가 없는 한계를 가지고 있다. 이처럼 현재까지 국내의 돌말류 관련 컬렉션 및 데이터베이스는 엄밀한 검증과 활용보다는 단순한 취합에 방점을 두고 있는 점에서 다소간의 한계를 갖고 있다고 하겠다.

#### 4. 국내 돌말류 분석자료의 정도관리 신뢰성 향상을 위한 제언

우리나라는 돌말류 분석 자료를 검증할 수 있는 경험이 풍부한 분류학자의 수가 부족한 실정이다. 특히 담수 돌말류를 연구하는 분류학자 보다 해양, 그리고 그 중에서도 저서성 돌말류를 분류하는 분류학자의 수가 절대적으로 부족하다. 이는 해양 돌말류 분류학자의 양성이 필요함을 시사한다. 현재 국내에서 돌말류 분류에 대한 교육은 일부 대학의 연구실에서만 가능하다. 돌말류 분류의 입문의 폭이 좁기 때문에, 분류학자를 양성 하기 위해서 북미 SFS의 분류학 인증 프로그램과 같은 교육 프로그램이 필요하다고 여겨진다. 종 식별을 위해 필요한 돌말류의 분류 지식과 현미경의 조작 및 관찰 기술을 함께 교육할 수 있는 프로그램을 통해서 기초적인 분류가 가능한 분류학자를 양성할 수 있다. 또한 분류학자들간의 분석 오차를 줄이기 위해서 유럽의 EDM과 같은 연구진들의 주기적인 모임이 필요하다. 이러한 모임의 목적은 연구진들간의 연구와 분류에 대한 교류를 통해서 분석 오차를 줄이는 것이다. 학술교류와 같은 모임은 공통적인 지식을 함께 공유할 수 있는 가장 직접적인 방법이다. 또한 장기적으로는 미국의 경우에서처럼 인증서를 발급하는 교육프로그램의 개설 역시 고려해 볼 수 있다.

다음으로 분류학자 양성의 기반에 필요한 것은 우리나라도 유럽과 북미와 같은 돌말류를 생물 지표로 활용한 정기적인 수계 모니터링의 적용이다. 국제적으로 환경오염에 대한 염려가 강해지고 있으며, 가까이로는 일본의 후쿠시마 원전 방사능 오염수의 방류가 예정되어 있기 때문에 해양 환경의 모니터링이 중요해지고 있는 상황이다. 이미 국가적으로 정기적인 해양 모니터링이 진행되고 있지만, 해당 모니터링은 돌말류를 활용한 생물 모니터링보다는 화학적인 모니터링에 좀 더 집중하고 있다. 해양 생태계의 위해성 평가를 위해서는 돌말류를 포함한 생물학적 환경 모니터링이 강화되어야 하며, 필요한 부분이다.

마지막으로 가장 중요하면서도 장기적인 노력이 필요한 부분은 돌말류 분류정보 데이터베이스의 구축이다. 이를 위해서는 형태학적 자료 및 DNA 염기서열 자료의 축적 모두가 중요하다. 예를 들어 환경 모니터링의 효율성을 높이기 위해서는 DNA 바코드 분석이 적합할 터이나, 이 때 형태학적으로 검증되지 않은 시료의 DNA 염기서열 자료를 사용하면 분석의 신뢰성은 전혀 담보되지 않을 것이기 때문이다. 따라서 검증된 형태기반 데이터베이스를 기본으로 하여 DNA 염기서열 정보를 추가하는 방법으로 데이터베이스의 구축이 이루어지는 것이 바람직할 것이다. 또한 이러한 종합적인 데이터베이스의 구축을 고려함에 있어, 앞서 언급한 해외 사례들 중에서 BOLD 및 AlgaTerra와 같이 형태적인 메타 정보와 DNA 염기서열을 함께 구축한 데이터베이스들이 좋은 모델이 될 수 있을 것으로 여겨진다.

## 사 사

본 연구는 2021년 해양수산부 재원으로 해양수산과학기술진흥원(과학기술기반 해역이용영향평가 기술개발, 20210427)의 지원을 받아 수행하였습니다.

## 참고문헌(References)

- Amato, A. and M. Montresor, 2008. Morphology, phylogeny, and sexual cycle of *Pseudo-nitzschia mannii* sp. nov. (Bacillariophyceae): a pseudo-cryptic species within the *P. pseudodelicatissima* complex. *Phycologia*, **47**(5): 487-497.
- Anonymous, 1703. Two Letters from a Gentleman in the Country, relating to Mr. Leuwenhoeck's Letter in Transaction, No. 283. Communicated by Mr. C. *Phil Trans*, **23**(288): 1494-1501.
- Archibald, R., 1984. Diatom illustrations—an appeal. *Bacillaria*, **7**: 173-178.
- Benson, D.A., M. Cavanaugh, K. Clark, I. Karsch-Mizrachi, J. Ostell, K.D. Pruitt and E.W. Sayers, 2018. GenBank. *Nucleic Acids Res*, **46**(D1): D41-D47.

- Coleman, C.O. and A. Radulovici, 2020. Challenges for the future of taxonomy: talents, databases and knowledge growth. *Megataxa*, **1**(1): 28-34.
- Cox, E.J., 1987. *Placoneis* Mereschkowsky: the re-evaluation of a diatom genus originally characterized by its chloroplast type. *Diatom Res.*, **2**(2): 145-157.
- Crawford, R.M., 1981. The siliceous components of the diatom cell wall and their morphological variation. In *Silicon and siliceous structures in biological systems* (pp. 129-156). Springer, New York, NY.
- Ehrenberg, C.G., 1838. *Die Infusionsthierchen als vollkommene Organismen. Ein Blick in das tiefere organische Leben der Natur*. Leipzig, Leopold Voss.
- Evans, K.M., A.H. Wortley and D.G. Mann, 2007. An assessment of potential diatom “barcode” genes (*cox1*, *rbcL*, 18S and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). *Protist*, **158**: 349-364.
- Evans, K.M., A.H. Wortley, G.E. Simpson, V.A. Chepurinov and D.G. Mann, 2008. A molecular systematic approach to explore diversity within the *Sellaphora pupula* species complex (Bacillariophyta). *J. Phycol.*, **44**: 215-231.
- Falciatore, A. and C. Bowler, 2002. Revealing the molecular secrets of marine diatoms. *Annu. Rev. Plant Biol.*, **53**(1): 109-130.
- Falkowski, P.G., R.T. Barber and V. Smetacek, 1998. Biogeochemical controls and feedbacks on ocean primary production. *Science*, **281**: 200-6.
- Fourtanier, E. and J.P. Kociolek, 1999. Catalogue of the diatom genera. *Diatom Res.*, **14**(1): 1-190.
- Glöckner, F.O., 2019. The SILVA database project: an ELIXIR core data resource for high-quality ribosomal RNA sequences. *Biodivers. Inf. Sci. Stand.*, **3**: e36125.
- Gmelin, J.F., 1791. *Carolia Linne... Systema Naturae per regna tria naturae secundum classes, ordines, genera species cum characteribus, differentiis, synonymis, locis*. Ed. 13, Tomus I. Pars VI. Vermes Infusoria. pp. 3021-3910. Lipsiae [Leipzig]: Georg Emanuel Beer.
- Gregory, W., 1857. On some new forms of marine Diatomaceae found in the Firth of Clyde and in Loch Fyne. *Trans R Soc Edinb Earth Sci.*, **21**: 473-542.
- Greville, R.K., 1857. Description of some new diatomaceous forms from the West Indies. *Quarterly Journal of Microscopical Science, London*, **5**: 7-12.
- Hamsher, S.E., K.M. Evans, D.G. Mann, A. Poulíckova and G.W. Saunders, 2011. Barcoding diatoms: exploring alternatives to COI-5P. *Protist*, **162**: 405-422.
- Harris, D.J., 2003. Can you bank on GenBank? *Trends Ecol. Evol.*, **18**: 317-319.
- Hebert, P.D., A. Cywinska, S.L. Ball and J.R. Dewaard, 2003. Biological identifications through DNA barcodes. *Proc. Royal Soc. B*, **270**(1512): 313-321.
- Hebert, P.D., M.Y. Stoeckle, T.S. Zemplak and C.M. Francis, 2004. Identification of birds through DNA barcodes. *PLOS Biol.*, **2**(10): e312.
- Im, A., J.S. Khim and J. Park, 2020. Taxonomy and distribution of two small *Tryblionella* (Bacillariophyceae) species from the Northeast Asian tidal flats. *J. Species Res.*, **9**(3): 191-197.
- Johns, G.C. and G.C. Avise, 1998. A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome b gene. *Mol. Biol. Evol.*, **15**: 1481-1490.
- Kawecka, B. and M. Olech, 1993. Diatom communities in the Vanishing and Ornithologist Creek, King George Island, South Shetlands, Antarctica. *Hydrobiologia*, 269-270, 327-333.
- Kermarrec, L., A. Franc, F. Rimet, P. Chaumeil, J.F. Humbert and A. Bouchez, 2013. Next-generation sequencing to inventory taxonomic diversity in eukaryotic communities: a test for freshwater diatoms. *Mol Ecol Resour*, **13**(4): 607-619.
- Kim, H., J.S. Khim and J. Park, 2020. First Report of *Navicula spartinetensis* (Bacillariophyceae) from Korean Tidal Flats Along with Its Distribution in Northeast Asia. *J. Korean Soc. Oceanogr.*, **25**(4): 97-105.

- Kusber, W.H. and R. Jahn, 2007. AlgaTerra Information System: Types data and data types. In Proceedings of the 1st Central European Diatom Meeting, pp. 97-100.
- Kützing, F.T., 1833. Synopsis Diatomacearum oder Versuch einer systematischen Zusammenstellung der Diatomeen. *Linnaea*, **8**: 529-620.
- Kützing, F.T., 1844. Die Kieselschaligen Bacillarien oder Diatomeen. W. Kohne Nordhausen. 152 pp.
- Lavoie, I., P.B. Hamilton, Y.K. Wang, P.J. Dillon and S. Campeau, 2009. A comparison of stream bioassessment in Québec (Canada) using six European and North American diatom-based indices. *Nova Hedwigia*, **135**: 37-56.
- Luddington, I.A., I. Kaczmarska and C. Lovejoy, 2012. Distance and characterbased evaluation of the V4 region of the 18S rRNA gene for the identification of diatoms (Bacillariophyceae). *PLoS ONE*, **7**: e45664.
- MacGillivray, M.L. and I. Kaczmarska, 2011. Survey of the efficacy of a short fragment of the rbcL gene as a supplemental DNA barcode for diatoms. *J. Eukaryot. Microbiol.*, **58**: 529-536.
- Mann, D.G., 1999. The species concept in diatoms. *Phycologia*, **38**: 437-495.
- Mann, D.G., R.M. Crawford and F.E. Round, 2017. "Bacillariophyta." *Handbook of the Protists*, Springer, Cham: 1657.
- Mann, D.G., S. Sato, R. Trobajo, P. Vanormelingen and C. Souffreau, 2010. DNA barcoding for species identification and discovery in diatoms. *Cryptogam., Algal.*, **31**: 557-577.
- Meiklejohn, K.A., N. Damaso and J.M. Robertson, 2019. Assessment of BOLD and GenBank-Their accuracy and reliability for the identification of biological materials. *PLoS ONE*, **14**(6): e0217084.
- Moniz, M.B. and I. Kaczmarska, 2009. Barcoding diatoms: Is there a good marker?. *Mol Ecol Resour*, **9**: 65-74.
- Mora, C., D.P. Tittensor, S. Adl, A.G. Simpson and B. Worm, 2011. How many species are there on Earth and in the ocean?. *PLOS Biol.*, **9**(8): e1001127.
- Morrow, A.C., T.R. Deason and D. Clayton, 1981. A NEW SPECIES OF THE DIATOM GENUS EUNOTIA 1. *J. Phycol.*, **17**(3): 265-270.
- Park, J. and C.-H. Koh, 2012. Taxonomic studies on Korean marine benthic diatoms LM and SEM observations of the diatom genus Amphora from Korean tidal flats. *Ocean Sci. J.*, **47**: 101-112.
- Park, J., B.-O. Kwon, M. Kim, S. Hong, J. Ryu, S.J. Song and J.S. Khim, 2014. Microphytobenthos of Korean tidal flats: A review and analysis on floral distribution and tidal dynamics. *Ocean Coast Manag.*, **102**: 471-482.
- Pečnikar, Ž.F. and E.V. Buzan, 2014. 20 years since the introduction of DNA barcoding: from theory to application. *J. Appl. Genet.*, **55**(1): 43-52.
- Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies and F.O. Glöckner, 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*, **41**: D590-D596.
- Ratnasingham, S. and P.D. Hebert, 2007. BOLD: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Mol. Ecol. Notes*, **7**(3): 355-364.
- Rimet, F., P. Chaumeil, F. Keck, L. Kermarrec, V. Vasselon, M. Kahlert and A. Bouchez, 2016. R-Syst:: diatom: an open-access and curated barcode database for diatoms and freshwater monitoring.
- Round, F.E., R.M., Crawford and D.G. Mann, 1990. *The Diatoms*. Cambridge, New York. 747 pp.
- Sato, S., W.H. Kooistra, T. Watanabe, S. Matsumoto and L.K. Medlin, 2008. A new araphid diatom genus *Psammoneis* gen. nov. (Plagiogrammaeae, Bacillariophyta) with three new species based on SSU and LSU rDNA sequence data and morphology. *Phycologia*, **47**(5): 510-528.
- Shen, Y.Y., X. Chen and R.W. Murphy, 2013. Assessing DNA barcoding as a tool for species identification and data quality control. *PLoS ONE*, **8**(2): e57125.
- Skvortzow, B.W., 1929. Freshwater diatoms from Korea, Japan. *hilipp J. Sci.*, **38**: 283-291.
- Smith, W., 1853. A synopsis of the British Diatomaceae, vol. 1. London, J. van Voorst.
- Smith, W., 1856. A synopsis of the British Diatomaceae, vol. 2. London, J. van Voorst.

- Sogin, M.L., H.G. Morrison, J.A. Huber, D.M. Welch, S.M. Huse, P.R. Neal, J.M. Arrieta and G.J. Herndl, 2006. Microbial diversity in the deep sea and the underexplored “rare biosphere”. *Proc. Natl. Acad. Sci.*, **103**: 12115-12120.
- Tréguer, P., D.M. Nelson, A.J. Van Bennekom, D.J. DeMaster, A. Leynaert and B. Quéguiner, 1995. The silica balance in the world ocean: A reestimate, *Science*, **268**: 375-379.
- Underwood, G.J.C. and J. Kromkamp, 1999. Primary production by phytoplankton and microphytobenthos in estuaries. *Adv. Ecol. Res.*, **29**: 93-153.
- Valentin, V., R. Frédéric, D. Isabelle, M. Olivier, R. Yorick and B. Agnès, 2019. Assessing pollution of aquatic environments with diatoms’ DNA metabarcoding: experience and developments from France Water Framework Directive networks. *Metabarcoding Metagenom.*, **3**: e39646.
- Valentini, A., F. Pompanon and P. Taberlet, 2009. DNA barcoding for ecologists. *Trends Ecol. Evol.*, **24**: 110-117. <https://doi.org/10.1016/j.tree.2008.09.011>.
- Ward, R.D., T.S. Zemlak, B.H. Innes, P.R. Last and P.D. Hebert, 2005. DNA barcoding Australia's fish species. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, **360**(1462): 1847-1857.
- Wesche, P.L., D.J. Gaffney and P.D. Keightley, 2004. DNA sequence error rates in Genbank records estimated using the mouse genome as a reference. *Dna Sequence*, **15**(5-6): 362-364.
- Yilmaz, P., L.W. Parfrey, P. Yarza, J. Gerken, E. Pruesse, C. Quast, T. Schweer, J. Peplies, W. Ludwig and F.O. Glöckner, 2014. The SILVA and “All-species Living Tree Project (LTP)” taxonomic frameworks. *Nucleic Acids Res*, **42**: D643-D648.
- Yool, A. and T. Tyrrell, 2003. Role of diatoms in regulating the ocean’s silicon cycle. *Global Biogeochem Cycles*, **17**: 1103.
- Zemlak, T.S., R.D. Ward, A.D. Connell, B.H. Holmes, and P.D. Hebert, 2009. DNA barcoding reveals overlooked marine fishes. *Mol Ecol Resour*, **9**: 237-242.
- Zimmermann, J., R. Jahn and B. Gemeinholzer, 2011. Barcoding diatoms: evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. *Org. Divers. Evol.*, **11**: 173-192.
- Zimmermann, J., G. Glöckner, R. Jahn, N. Enke and B. Gemeinholzer, 2015. Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. *Mol Ecol Resour*, **15**(3): 526-542.