Environmental DNA (eDNA) evidence of North Sea mollusc transfer across tropical waters through ballast water

Alba Ardura[1, 2*], Anastasija Zaiko[2*], Jose L. Martinez[3], Aurelija Samuiloviene[2, 4], Yaisel Borrell[1], Eva Garcia-Vazquez[1]

* contributed equally to this article

1: Department of Functional Biology, University of Oviedo. C/ Julian Claveria s/n. 33006-Oviedo, Spain.

2: Marine Science and Technology Centre, Klaipeda University. H. Manto 84, LT 92294, Klaipeda, Lithuania.

3: Unit of DNA Analysis, Scientific-Technical Services, University of Oviedo. Edificio Severo Ochoa, Campus del Cristo, 33006-Oviedo, Spain.

4: Department of Biology, Klaipeda University. H. Manto 84, LT 92294, Klaipeda, Lithuania.

Corresponding author: Alba Ardura alarguti@hotmail.com

Abstract

Maritime transport, and in particular ballast water, is considered to be one of the most important pathways of marine biological invasions worldwide. Here we provide the first molecular evidence of potential survival of the European mudsnail, *Peringia ulvae*, in ballast water on cross-latitudinal voyages. Ballast water from the RV Polarstern was sampled at its departure from the North Sea and again in tropical latitudes, DNA extracted and amplicon sequenced employing high-throughput sequencing methodology. Mollusc species were detected by cytochrome oxidase subunit I DNA barcode sequences. The increasing proportion of OTUs that were identified as *P. ulvae* after two weeks of navigation, suggests that this species withstands the harsh conditions in the ballast tank. As such, *P. ulvae* has the potential to reach very distant, new marine areas where it eventually might establish itself as a non-indigenous species. We also discuss the potential of environmental DNA analysis for on-route biodiversity screening, species-specific risk assessments, as well as some current limitations of the approach.

Keywords

eDNA, ballast water, molluscs, *Peringia ulvae*, COI, NGS

Introduction

Shipping is believed to be one of the most important pathways for non-indigenous species transfer across marine regions (Leppäkoski, Gollasch & Olenin 2002). This pathway involves several potential vectors - transport of organisms in ballast waters, ballast tank sediments, hull and sea chest fouling, anchors and anchor chains, etc (Hewitt, Gollasch & Minchin 2009). Ballast water (BW) is recognized as the most significant one of these vectors (Molnar et al. 2008). Approximately 2.2 to 12 billion tons of ballast water is transported across the world oceans annually (Endresen et al. 2004), transferring daily some 7,000 species (Gollasch & David 2011). In a summary of 15 European BW surveys, living specimens of more than 1,000 taxa were found in ballast tanks of vessels arriving in European ports (Gollasch et al. 2002).

Due to the extremely harsh conditions (darkness, temperature changes, salinity pulses , variable turbidity, turbulence and oxygen depletion), numbers of living organisms in ballast tanks decline rapidly after ballasting (Gollasch et al. 2000; Hewitt, Gollasch & Minchin 2009). Nevertheless, there are examples of specimens surviving long intercontinental transfers (Gollasch et al. 2000).

A golden rule for successful invaders is “the more tolerant are the more dangerous” (Sakai et al. 2001; Lee 2002; Madariaga et al. 2014). Therefore migrants surviving long cross-latitudinal voyages within ballast tanks should be of particular concern as potential invaders. Identifying such species is crucial for conducting reliable risk analyses, preventing expansions, and developing efficient control methods (Tsolaki & Diamadopoulos 2010). For many species transported in BW as eggs or larvae, the accurate taxonomic identification is not an easy task however. It is especially complicated in on-route surveys, when samples are collected and analyzed instantly onboard and specific taxonomic expertise is not available. DNA methodologies are very

useful complimentary tools to identify organisms in BW (Darling & Blum 2007; Harvey, Hoy & Rodriguez 2009; Darling & Mahon 2011; Briski et al. 2012). Recently, the development of next generation sequencing (NGS) technologies simplified and speeded up the whole process by allowing the identification of entire communities in water samples, using bulk or environmental DNA (eDNA) analysis. eDNA is extracted directly from environmental samples (e.g. soil or water) (Ficetola et al. 2008). This allows the detection of species from single cells in a sample, such as gamete, secreted feces or mucous and is particularly advantageous for small, rare, and cryptic species or life stages that are difficult to detect otherwise (Ficetola et al. 2008; Valentini, Pompanon & Taberlet 2009; Taberlet et al. 2012; Thomsen et al. 2012). The eDNA approach in a combination with NGS is increasingly exploited in metabarcoding/metagenetic studies aimed at biodiversity research (Hajibabaei et al. 2011; Wood et al. 2013).Many of mollusc invasions have been associated with the unintentional transport of planktonic life stages (e.g. Strayer 2010). Cases in point are the bivalves *Dreissena* spp. (Benson 2013), *Corbicula* spp. (Grigorovich et al. 2003), *Limnoperna fortune* (Ricciardi 1998), *Corbula amurensis* (Carlton et al. 1990) and the gastropods *Crepidula fornicata* (Elliot 2003), *Potamopyrgus antipodarum* (Alonso & Castro-Diaz 2008).

In this study we apply metabarcoding (eDNA) for species identification in BW from the RV Polarstern during the expedition ANT-XXIX/1 in October-December 2012 (from Bremerhaven, Germany to Cape Town, South Africa). We focused on the detection of molluscs that could have survived the harsh BW conditions over the cross-latitudinal transfer, and that hence could become non-indigenous or invasive species. Metabarcoding has been successfully applied for studying the evolution of general biodiversity in BW during this expedition (Zaiko et al. 2015) and as such, we here explore the applicability of eDNA for the taxonomical screening of BW and species-specific risk assessments.

Material and Methods

Collection of water samples and environmental metadata

The aft ballast tank (70 m^3) of the vessel was filled with North Sea water on October 28th, off Bremerhaven. At the time of the BW upload, water temperature and salinity were 13.1 ^{o}C and 34 ppt respectively. Four samples of BW were collected via the water pipe on days 2 and 4 (temperate latitudes) and days 12 and 16 (tropical latitudes) of the cruise (Figure 1).

For each sample, 100 L of BW were pumped through a plankton net (30 cm diameter, 55 µm mesh size). The concentrated material (*ca.* 50 mL) was then vacuum-filtered through a 0.2 µm Nuclepore^{TM} membrane, which was thereafter preserved in 96% ethanol until eDNA extraction. Changes in temperature, pH, and oxygen saturation of the BW were measured using an Ysi Professional Plus Multimeter.

DNA and bioinformatics analyses

DNA was extracted from the filters using the QIAamp DNA Mini Kit (Qiagen) and was quantified with a fluorescence-based quantification method (Picogreen, Invitrogen). In order to validate the findings of mollusc eDNA, the extracted bulk DNA was analyzed with two different NGS platforms: Days 2 and 12 samples - the Ion Personal Genome Machine System (PGM. Life technologies) at the Sequencing unit of Oviedo University (Spain); Days 4 and 16 samples - the Genome Sequencer FLX (Roche 454) at Macrogen (Korea). Universal mini-barcode primers (Meusnier et al. 2008) were used to amplify and sequence a ~140 bp fragment of the mitochondrial cytochrome oxidase subunit I (COI) gene. For the 454 sequencing, a 1/40 of the 454 plate was used for each BW sample. The GS FLX data processing was performed using the Roche GS FLX software (v2.9). The software used tag (barcode) sequences to segregate the reads from each sample, by matching the initial and final bases of the reads to the known tag sequences used in the preparation of the libraries. Zero base errors were allowed in this sorting by tag step. Raw data were then processed using PRINSEQ v0.20.4 (Schmieder & Edwards 2011) for filtering too short and/or too long reads (mode +/- 2SD) and also to eliminate low quality reads (mean ≥ 20). Ambiguous sequences were discarded.

The sequencing and data processing with Ion Torrent was performed as explained in Zaiko et al. (2015). Briefly, libraries were constructed using the kit Ion Plus Fragment Library Kit (Life Technologies) and templates were obtained using the Ion PGM™ Template OT2 200 Kit (Life Technologies). The templates were loaded in a 314 chip and sequenced using the Ion PGM Sequencing 200 Kit v2 (Life Technologies). The yielded sequences were filtered by length (between 130 and 200 bp) and quality (+20). The expected length of the target miniCOI region falls within this length range, so further contig analysis was not necessary for OTU assignment.

Taxonomic classification of the obtained datasets was done by BLAST-aligning sequences against the NCBI nucleotides database (http://www.ncbi.nlm.nih.gov/) using the QIIME platform (Caporaso et al. 2010). The same software was used for prior de-noising and detection of chimeras. Taxonomic criteria were: best hit, max E-value = 0.001, min percent identity = 90.0, which are not sufficiently strict for species assignation, but which allow to retain class, order or family level taxa.

After initial analysis, the dataset of OTUs with their closest reference matches were curated, i.e. species taxonomic information was verified and checked against the World Register of Marine Species (http://www.marinespecies.org/), AlgaeBase (http://www.algaebase.org/) and Encyclopedia of Life (http://eol.org/) databases. OTUs involving non marine organisms were eliminated. The curated sequence dataset was employed for the further analyses.

Since this study did not aim at analyzing biodiversity, we only assigned the sequences of interest to kingdoms and, within Animalia, we calculated the frequency of putative molluscs in temperate and tropical samples. Percentages were employed for this quantification.

Phylogenetic analysis

Sequences identified as mollusc DNA were manually extracted from the NGS output files identified as molluscs in the OTU list, and aligned using the BioEdit software (Hall

1999). Haplotypes were determined with the program DnaSP (Librado & Rozas 2009). Distinguishing between the results from different NGS platforms, we called sequences A and B those obtained from Ion Torrent and 454 respectively.

A reference database of invasive molluscs was constructed from COI gene sequences obtained from the GenBank (Supplementary Table 1). The species selection was made based on recognized invasive capacity of different mollusc taxa from the sequences available in GenBank. The species hereby were included if classified as dangerous/globally invasive for marine habitats in the IUCN ISSG database.

Phylogeneticanalyses were conducted using MEGA version 6 (Tamura et al. 2013). Phylogenetic trees containing the reference and BW sequences obtained in this work were inferred with Maximum Likelihood with the following settings: Tamura Nei model (Tamura & Nei 1993) for nucleotides and JTT Matrix model (Jones-Taylor-Thornton) (Jones, Taylor & Thornton 1992) for amino acids. Robustness of the tree topology was assessed using 1,000 bootstrap replicates.

The Chi-Square statistic was employed to assess the significance of the shift in proportions of the particular haplotypes.

Results

The environmental conditions of the Polarstern BW changed dramatically over the sampling period (Figure 1). The temperature increased by nearly 14 ºC, while oxygen saturation and pH decreased by 84% and 0.5 respectively between the 2nd and the 16th navigation days.

The sequences and putative marine taxa obtained from the analyzed samples using 454 and Ion Torrent platforms are summarized in Table 1. A total number of 16,989 and 22,242 sequences of the expected size (around 150 bp, Meusnier et al. 2008) that BLASTed to marine taxa were obtained from the two temperate samples with Ion Torrent and 454 platforms respectively. From the tropical samples, 3,032 and 11,525 sequences were assigned to marine taxa in A and B datasets respectively, demonstrating a substantial reduction in NGS reads (Table 1). In general, the share of assigned sequences was higher (nearly 100%) in 454 datasets comparing to Ion Torrent ones (62 and 21% in temperate and tropical samples correspondingly). Animalia were clearly the dominant domain in all analysed datasets while Plantae and Chromista were underrepresented in the Ion Torrent datasets. The absolute majority of Mollusca sequences obtained from the samples were BLASTed to *Peringia ulvae* (formerly *Hydrobia ulvae*). The closest match for the OTUs found here from both tropical samples and B-temperate sample was the GenBank reference AF118308 followed by AF118290. In addition 2 OTUs with the closest match with *Lophiotoma leucotropis* (GenBank HQ834093) were found in the B-temperate sample, and 4 were BLASTed to Cephalopods (*Sepia* spp.) in B-temperate and tropical samples.

The BW biota composition shifted between the temperate and the tropical samples, with a particular increase in the proportion of protists in the B dataset (Figure 2). The overall proportion of *Peringia ulvae* within the total number of assigned sequences increased

from 3 to 4% in the B dataset and from 0 to 36% in the A dataset, being by far the most abundant molluscan OTU.

In the manually extracted sequences from the NGS data files, the mollusc sequences found from B temperate sample and BLASTed to *Peringia ulvae* and *Lophiotoma leucotropis* corresponded respectively to eight and one different haplotypes (BWTemperate01-08B, EMBL references HG963478-85 and BWTemperate09; Figure 3). The eight haplotypes BWTemperate 1 to 8 were found approximately in the same proportion within the 748 OTUs BLASTed to *P. ulvae*.

In the B-tropical sample, the 353 mollusc-BLASTed sequences corresponded to one unique haplotype (BWTropical01B, EMBL reference HG963486) with the closest match to *Peringia ulvae*. Exactly the same haplotype was retrieved from the *Peringia*-BLASTed OTUs A-Tropical sample (BWTropical01A, EMBL reference HG963486).. This haplotype was also present in the B-temperate sample, named there as BWTemperate02 (EMBL reference HG963479) (Figure 3). The other haplotypes in the B-temperate sample did not appear in the tropical water sample. A rough quantitative analysis demonstrated an increase of the proportion of this haplotype from approximately 12.6% to 100% of all the *Peringia*-BLASTed OTUs in the B-temperate and B-tropical samples respectively. This increase was statistically significant (contingency Chi-Square = 760.3, $P<<0.001$, for1 degree of freedom). On the other hand, the proportion of this haplotype over the total number of the BW OTUs increased from 0.42% in the B-temperate to 3.03% in the B-tropical sample (contingency Chi-Square = 397.5, $P<<0.001$, for 1 degree of freedom). In the A-tropical sample from the Ion Torrent platform its proportion was much higher (36%) being the unique *Peringia* haplotype as commented above.

To confir BLAST species identification, the mollusc-like sequences from the B dataset were aligned with the reference GenBank COI sequences of invasive molluscs (Supplementary Table 1) plus the GenBank sequences with the closest match (two *Peringia ulvae* and one *Lophiotoma leucotropis*). The resulting phylogenetic trees showed similar topologies (e.g. Figure 3) in which gastropods and bivalves were clustered in separated branches. All the BW sequences that BLASTed as *P. ulvae* clustered closely with the reference *P. ulvae* sequences. The other mollusc-like sequence found in the B-temperate sample, BWTemperate09, clustered in the branch of Gastropods but clearly separated from the *Lophiotoma leucotropis* reference. A closer examination of this sequence revealed that the haplotype contained stop codons (Figure 4). This means that it does not correspond to the true mitochondrial COI coding sequence and could be considered a pseudogene. On the other hand, the eight haplotypes assigned to *Peringia ulvae* (EMBL accession numbers HG963478-85) code for amino acid sequences compatible with the standard COI proteins.

Discussion

The results of this study suggest that the European mudsnail *Peringia ulvae* may successfully cross the oceans in ballast water. We detected the presence of sequences most closely matching with this species in ballast water samples, with the proportion of a particular haplotype increasing over time. This could be explained if such haplotype was less degraded than the rest. Of course, the mere presence of a species specific DNA

does not ensure that the species has been sampled alive. Previous studies have demonstrated that extracellular eDNA molecules can persist in water for several days to weeks (Dejean et al. 2011, Barnes, Turner & Jarde 2014), even if it degrades by the action of environmental factors such as UV, pH and microbial activity (Hall & Ballantyne 2004; Thacker et al. 2006; Pilliod et al. 2013; Barnes et al. 2014). Hence, decay is expected if DNA molecules are not inside the living cells (Levy-Booth et al. 2007; Dejean et al. 2011). So, after 16 days of navigation under increasing temperatures, low oxygen and slightly decreasing pH, it is expected that only living organisms will increase their relative DNA contribution to the BW eDNA pool. In the temperate sample we have found 8 different haplotypes and only one of them was maintained (and even increased its relative proportion) in the tropical sample (Figure 3). These observations indicate that at least one *P. ulvae* haplotype has persisted longer than other organisms during the cross-latitudinal BW transfer. Alternatively, the apparent increase of haplotype BWTemperate02 (=BWTropical01) could be attributed to a difference in sequencing success between the two samples. Yet, its very high proportion in the A-Tropical data rather points to our first hypothesis.

To our knowledge there are no reports of *P. ulvae* out of its native range (North East Atlantic Ocean and Mediterranean Sea) so far. However, due to its biological traits it could exhibit invasive behavior if introduced to other marine ecosystems. Within its native range (e.g. Danish waters) it is known to compete with the sympatric Hydrobiidae *Ecrobia ventrosa* (Gorbushin 1996). In other European regions, *P. ulvae* appears to be tolerant to diverse ecological conditions, inhabiting intertidal zones, whereas its Hydrobiidae competitors *Ecrobia ventrosa* and *Hydrobia neglecta* are confined to the non-tidal lagoons (Barnes 1999). On the other hand, in similar salinity conditions *P. ulvae* would adapt to warm temperatures (up to 30 ºC), better than other Hydrobiidae (Pascual & Drake 2008). These examples indicate the capacity of the species to survive in diverse environments, including the BW conditions. There are some well-known examples of Hydrobiidae being extremely aggressive invaders, e.g. New Zealand mudsnail *Potamopyrgus antipodarum* that have been introduced to many aquatic ecosystems worldwide and induced numerous adverse impacts to the local habitats and communities (Snoeijs 1989; Alonso & Castro-Diaz 2008). Hence, further investigation and risk assessment of the potential invasive capacity of *P. ulvae* is recommended in order to set the adequate management strategy to prevent its spread overseas with the shipping pathway.

The NGS methodology applied here is a promising tool for biodiversity screening and detection of a potential invasive taxa in BW. It meets the efficiency, consistency, and comprehensiveness requirements prescribed for BW surveillance and risk assessment procedures (Helcom 2010, Zaiko et al. 2015). However there are still a few potential limitations that need to be taken into account and ideally ruled out in the future to ensure the robustness of the approach.

For instance, the universal primers employed in our study might not amplify equally well in all taxa present in a sample (Meusnier et al. 2008) or in all samples, so that certain taxa may be overlooked in certain samples. This might explain the “lack” of *Peringia ulvae* in the A-temperate sample. Therefore eDNA analyses need to be validated by taking samples on consecutive days and by using different platforms. The discrepancy between platforms is one of the problems that must be solved for a

generalized use of NGS data for routine monitoring of biological invasions. The DNA fragment targeted here was comparatively short (Meusnier et al. 2008). Although this is an advantage for detecting DNA traces in environmental samples, longer fragments will discriminate better between closely related species and increase the robustness of taxonomic assignment. Additional confirmation of taxonomic assignment from other markers would be desirable when higher taxonomical resolution and identification confidence are required (Kelly et al. 2014).

Another important issue that can potentially compromise NGS results is the availability, the taxonomic coverage and the reliability of the reference sequence databases (Ardura et al. 2013; Pochon et al. 2015). In order to at least partly overcome this limitation, the results of the present study were confirmed by the NCBI-derived reference sequences of selected invasive species and phylogenetic value for ascertaining the taxonomic assignation of eDNA derived sequences. The use of phylogenies for confirming the taxonomical status of ambiguous sequences has been successfully applied in diversity studies (e.g. Moon-van der Staay et al. 2001). However, there is no reference sequence collections provided particularly for the invasive species. For this study, we have compiled a reference database on invasive mollusk species sequences from the publically available sources. Employing such a database we could reasonably reject the idea that mollusc eDNA sequences found in our samples belong to any of the already recognized invasive species, although were somewhat relevant to the New Zealand mudsnail *Potamopyrgus antipodarum*.

The results of this study could be interpreted as indication of the high likelihood of the species survival in ballast water or sediments on cross-regional voyages and therefore used for the species-specific risk assessments required among other within Ballast Water Management Convention (IMO 2004) and for prioritizing species of greatest management concern (Lehtiniemi et al. 2015).

Acknowledgements

We thank the crew of the RV Polarstern for their so valuable collaboration. This study has been supported by the Erasmus Mundus EMBC for BW sampling, the Spanish National Project MINECO CGL2013-42415-R, the Campus of Excellence of the University of Oviedo, the Principality of Asturias Grant GRUPIN14-93, and the Research Council of Lithuania project BALMAN (Development of the ships' ballast water management system to reduce biological invasions. Project # TAP-LLT-14-013). This is a contribution from the Marine Observatory of Asturias. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Ethics statement

Sampling ballast water onboard Polarstern was authorized by the Chief Scientist of the cruise XXIX/1 Dr. Holger Auel (University of Bremen) and the Scientific Coordinator Dr. Reiner Knust. Protected or endangered species were not involved in this study. The Polarstern ballast tank is a closed space and water was not renewed during the studied period. Only ballast water was analyzed; no other water samples were taken, thus

specific permission for other sampling was not needed. We worked with environmental DNA extracted from filtered water samples; therefore sacrifice of individuals was not necessary. Living macroscopic vertebrates needing special treatment were not detected in the water samples analyzed. The work did not require approval by an Institutional Animal Care and Use Committee given the nature of the samples analyzed.

References

ALONSO A., CASTRO-DIEZ P. 2008. What explains the invading success of the aquatic mud snail *Potamopyrgus antipodarum* (Hydrobiidae, Mollusca)? *Hydrobiologia*, 614:107-116.

ARDURA A., PLANES S., GARCIA-VAZQUEZ E. 2013. Applications of DNA barcoding to fish landings: authentication and diversity assessment. *ZooKeys*, 365:49-65.

BARNES R.S.K. 1999. What determines the distribution of coastal Hydrobiid mud snails within North-Western Europe? *Marine Ecology,* 20:97–110.

BARNES M.A., TURNER C.R., JERDE C.L. et al. 2014. Environmental Conditions Influence eDNA Persistence in Aquatic Systems. *Environmental Science Technology*, 48:1819-1827.

BENSON A.J. 2013. Chronological History of Zebra and Quagga Mussels (Dreissenidae) in North America, 1988–2010. Chapter 1. In: *Quagga and zebra mussels: Biology, Impacts and Control*. Second Edition (Eds. Nalepa T., Schlosser D.), pp. 9-31, CRC Press.

BRISKI E., GHABOOLI S., BAILEY S.A., MacISAAC H.J. 2012. Invasion risk posed by macroinvertebrates transported in ships' ballast tanks. *Biological Invasions*, 14:1843-1850.

CAPORASO J.G., KUCZYNSKI J., STOMBAUGH J., BITTINGER K., BUSHMAN F.D. et al. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, 7:335-336.

CARLTON J.T., THOMPSON J.K., SCHEMEL L.E., NICHOLS F.H. 1990. Remarkable invasion of San Francisco Bay (California, USA) by the Asian clam *Potamocorbula amurensis*. I. Introduction and dispersal. *Marine Ecology Progress Series*, 66:81-94.

DARLING J.A., BLUM M.J. 2007. DNA-based methods for monitoring invasive species: a review and prospectus. *Biological Invasions*, 9:751-765.

DARLING J.A., MAHON A.R. 2011. From molecules to management: Adopting DNA-based methods for monitoring biological invasions in aquatic environments. *Environmental Research*, 111:978-988.

DEJEAN T., VALENTINI A., DUPARC A., PELLIER-CUIT S., POMPANON F. et al. 2011. Persistence of environmental DNA in freshwater ecosystems. *PLoS ONE*, 6:e23398

ELLIOT M. 2003. Biological pollutants and biological pollution - an increasing cause for concern. *Marine Pollution Bulletin*, 46:275-280.

ENDRESEN Ø., BEHRENS H.L., BRYNESTAD S., ANDERSEN A.B., SKJONG R. 2004. Challenges in Global Ballast Water Management. *Marine Pollution Bulletin*, 48: 615–623.

FICETOLA G.F., MIAUD C., POMPANON F., TABERLET P. 2008. Species detection using environmental DNA from water samples. *Biology Letters*, 4:423-425.

GOLLASCH S., DAVID M. 2011. Sampling Methodologies and Approaches for Ballast Water Management Compliance Monitoring. *Traffic and Environment*, 33:397-405.

GOLLASCH S., LENZ J., DAMMER M., ANDRES H.G. 2000. Survival of tropical ballast water organisms during a cruise from the Indian Ocean to the North Sea. *Journal of Plankton Research*, 22:923–937.

GOLLASCH S., MacDONALD E., BELSON S., BOTNEN H., CHRISTENSEN J. et al. 2002. *Invasive Aquatic Species of Europe: Distribution, Impacts and Management*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

GORBUSHIN A.M. 1996. The enigma of mud snail shell growth: asymmetrical competition or character displacement? *Oikos*, 77:85–92.

GRIGOROVICH I.A., COLAUTTI R.I., MILLS E.L., HOLECK K., BALLERT A.G. et al. 2003. Ballast-mediated animal introductions in the Laurentian Great Lakes: retrospective and prospective analyses. *Canadian Journal of Fisheries and Aquatic Science*, 60:740-756.

HAJIBABAEI M., SHOKRALLA S., ZHOU X., SINGER G.A., BAIRD D.J. 2011. Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS One*, 6:e17497.

HALL A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41:95-98.

HALL A., BALLANTYNE J. 2004. Characterization of UVC-induced DNA damage in bloodstains: forensic implications. *Analytical and Bioanalytical Chemistry*, 380:72-83.

HARVEY J.B.J., HOY M.S., RODRIGUEZ R.J. 2009. Molecular detection of native and invasive marine invertebrate larvae present in ballast and open water environmental samples collected in Puget Sound. *Journal of Experimental Marine Biology and Ecology*, 369:93–99.

HELCOM. 2010. *HELCOM Ministerial Declaration on the implementation of the HELCOM Baltic Sea Action Plan*. In, p. 9, Moscow

HEWITT C.L., GOLLASCH S., MINCHIN D. 2009. The vessel as a vector - biofouling, ballast water and sediments. In: Rilov G, Crooks JA, eds. *Biological invasions in marine ecosystems*. Ecological studies 204. Berlin Heidelberg: Springer-Verlag. pp. 117-129.

IMO. 2004. *International Convention for the Control and Management of Ships' Ballast Water and Sediments*. IMO.

JONES D.T., TAYLOR W.R., THORNTON J.M. 1992. The rapid generation of mutation data matrices from protein sequences. *Computer Applications in Bioscience*, 8:275-282.

KELLY R.P., PORT J.A., YAMAHARA K.M., MARTONE R.G., LOWELL N., THOMSEN P.F., MACH M.E., BENNET M., PRAHLER E., CALDWELL M.R.,

CROWDER L.B. 2014. Harnessing DNA to improve environmental management. *Science*, 344:1455-1456.

LEE C.E. 2002. Evolutionary genetics of invasive species. *Trends in Ecology and Evolution*, 17:386-391.

LEHTINIEMI M., OJAVEER H., DAVID M., GALIL B., GOLLASCH S., McKENZIE C., MINCHIN D., OCCHIPINTI-AMBROGI A., OLENIN S., PEDERSON J. 2015. Dose of truth - Monitoring marine non-indigenous species to serve legislative requirements. *Marine Policy*, 54:26-35.

LEPPÄKOSKI E., GOLLASCH S., OLENIN S. (eds). 2002. *Invasive aquatic species of Europe - distribution, impact and management*. ISBN 1-4020-0837-6. Dordrecht, Boston, London. Kluwer Academic Publishers. 583 p.

LEVY-BOOTH D.J., CAMPBELL R.G., GULDEN R.H., HART M.M., POWELL J.R. et al. 2007. Cycling of extracellular DNA in the soil environment. *Soil Biology & Biochemistry*, 39:2977–2991.

LIBRADO P., ROZAS J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25:1451–1452.

MADARIAGA D.J., RIVADENEIRA M.M., TALA F., THIEL M. 2014. Environmental tolerance of the two invasive species *Ciona intestinalis* and *Codium fragile*: their invasion potential along a temperate coast. *Biological Invasions*, 16(12):2507-2527.

MEUSNIER I., SINGER G.A.C., LANDRY J.F., HICKEY D.A., HEBERT P.D.N., HAJIBABEI M. 2008. A universal DNA mini-barcode for biodiversity analysis. *BMC Genomics*, 9:214.

MOLNAR J.L., GAMBOA R.L., REVENGA C., SPALDING M.D. 2008. Assessing the global threat of invasive species to marine biodiversity. *Frontiers in Ecology and the Environment*, 6:485-492.

MOON-VAN der STAAY S.Y., DE WACHTER R., VAULOT D. 2001. Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature*, 409:607-610.

PASCUAL E., DRAKE P. 2008. Physiological and behavioral responses of the mud snails *Hydrobia glyca* and *Hydrobia ulvae* to extreme water temperatures and salinities: implications for their spatial distribution within a system of temperate lagoons. *Physiological and Biochemical Zoology*, 81:594-604.

PILLIOD D.S., GOLDBERG C.S., ARKLE R.S., WAITS L.P. 2013. Factors influencing detection of eDNA from a stream-dwelling amphibian. *Molecular Ecology Resources*, 14:109-116.

POCHON X., ZAIKO A., HOPKINS G.A., BANKS J.C., WOOD S.A. 2015. Early detection of eukaryotic communities from marine biofilm using high-throughput sequencing: an assessment of different sampling devices. *Biofouling*, (in press).

RICCIARDI A. 1998. Global range expansion of the Asian Mussel *Limnoperna fortunei* (Mytilidae): another fouling threat to freshwater systems. *Biofouling*, 13: 97-106.

SAKAI A.K., ALLENDORF F.W., HOLT J.S., LODGE D.M., MOLOFSKY J. et al. 2001. The population biology of invasive species. *Annual Review of Ecology, Evolution and Systematics*, 32:305-332.

SCHMIEDER R., EDWARDS R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics,* 27:863-864.

SNOEIJS P. 1989. Effects of increasing water temperatures and flow rates on epilithic fauna in a cooling-water discharge basin. *Journal of Applied Ecology,* 26:935-956.

STRAYER D.L. 2010. Alien species in fresh waters: ecological effects, interactions with other stressors, and prospects for the future. *Freshwater Biology,* 55(s1):152-174.

TABERLET P., COISSAC E., HAJIBABEI M, RIESEBERG L.H. 2012. Environmental DNA. *Molecular Ecology*, 21:1789-1793.

TAMURA K., NEI M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution,* 10:512-526.

TAMURA K., STECHER G., PETERSON D., FILIPSKI A., KUMAR S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution*, 30:2725-2729.

THACKER C.R., OGUZTURUN C., BALL K.M., SYNDERCOMBE COURT D. 2006. An investigation into methods to produce artificially degraded DNA. *International Congress Series*, 1288:592-594.

THOMSEN P.F., KIELGAST J., IVERSEN L.L., MOLLER P.R., RASMUSSEN M., WILLERSLEV E. 2012. Detection of a diverse marine fish fauna using environmental DNA from seawater samples. *PLoS One*, 7:e41732.

TSOLAKI E., DIAMADOPOULOS E. 2010. Technologies for ballast water treatment: a review. *Journal of Chemical Technology and Biotechnology*, 85:19-32.

VALENTINI A., POMPANON F., TABERLET P. 2009. DNA Barcoding for ecologists. *Trends in Ecology and Evolution,* 24:110-117.

WOOD S.A., SMITH K.F., BANKS J.C., TREMBLAY L.A., RHODES L., MOUNTFORT D., CARY S.C., POCHON X. 2013. Molecular genetic tools for environmental monitoring of New Zealand's aquatic habitats, past, present and the future. *New Zealand Journal of Marine and Freshwater Research,* 47:90-119.

ZAIKO A., MARTINEZ J.L., SCHMIDT-PETERSEN J., RIBICIC D., SAMUILOVIENE A., GARCIA-VAZQUEZ E. 2015. Metabarcoding approach for the ballast water surveillance – An advantageous solution or an awkward challenge? *Marine Pollution Bulletin*, doi: 10.1016/j.marpolbul.2015.01.008.

490 **Figure legends**

491 Figure 1. Approximate geographical position of RV Polarstern on the sampling days;
492 dynamics of the physical-chemical conditions (temperature, oxygen saturation and pH)
493 in the ballast tank over the 16 days of the cruise.

494 Figure 2. Proportion of DNA sequences assigned to three eukaryote kingdoms (protists,
495 algae and animals, partitioning *Peringia ulvae* sequence share) found in the ballast
496 water sampled in temperate and tropical latitudes. A and B correspond to datasets
497 obtained from Ion Torrent and 454 NGS platforms, respectively.

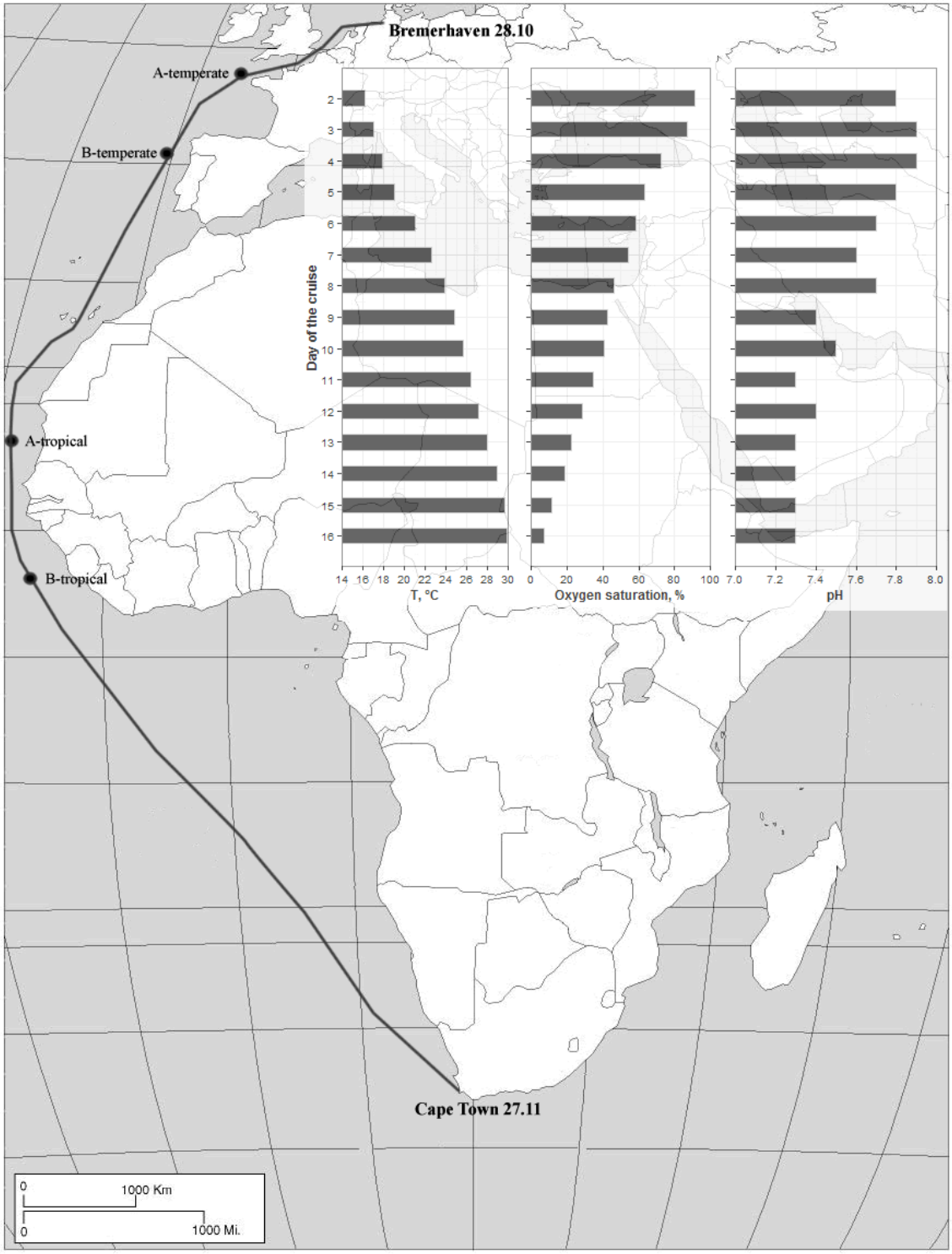
498 Figure 3. Maximum likelihood tree based on NCBI retrieved sequences (with reference
499 numbers) and Ballast Water (BW) COI gene sequences. Bootstrap values in percent.

500 Figure 4. Alignment between inferred amino acid sequences of Temperate09 BW-
501 sample and a *Lophiotoma leucotropis* reference. * represents stop codons.

Table 1. Summary of the NGS results and taxonomic assignment. A and B, Ion Torrent and 454 platforms respectively. Total number of reads obtained for each sample after sequence quality check; total number of sequences assigned to marine taxa OTUs, number of Chromista, Plantae and Animalia, and within them – the number of mollusc and *Peringia ulvae* sequences.

	A-temperate	A-tropical	B-temperate	B-tropical
Reads	27497	14304	22341	11536
Assigned reads	16989	3032	22242	11525
Chromista	423	33	1707	1447
Plantae	2	1	4935	2767
Animalia	16562	2998	15600	7311
Mollusca	2	1102	750	353
Peringia ulvae	0	1100	748	353

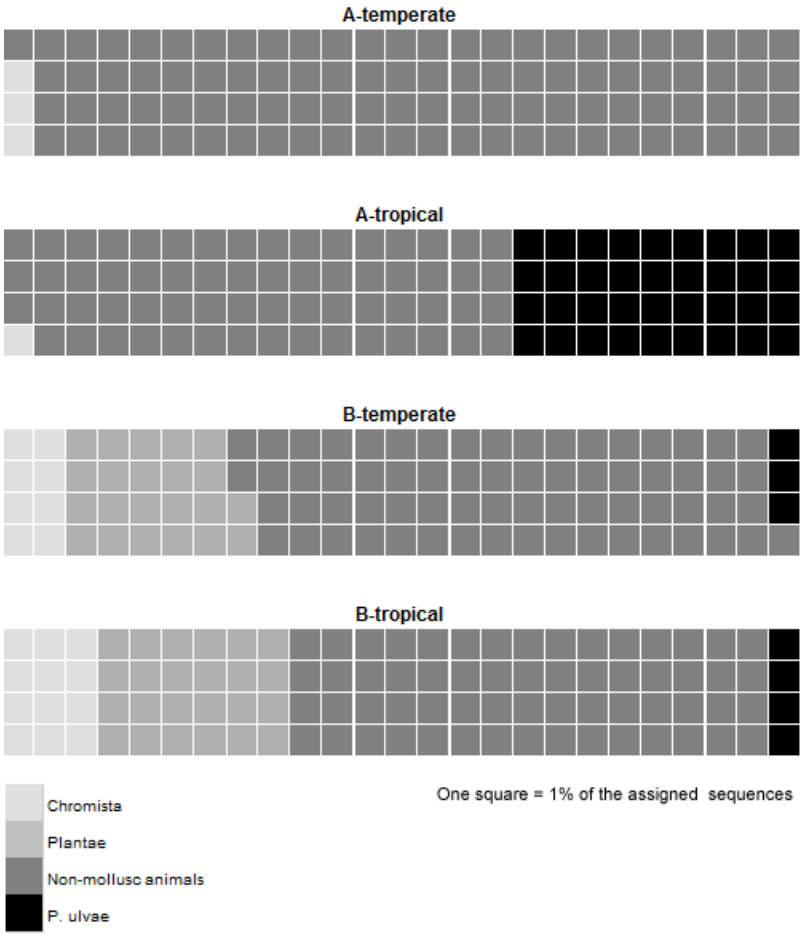
509 Figure 1:



510

511

512 Figure 2

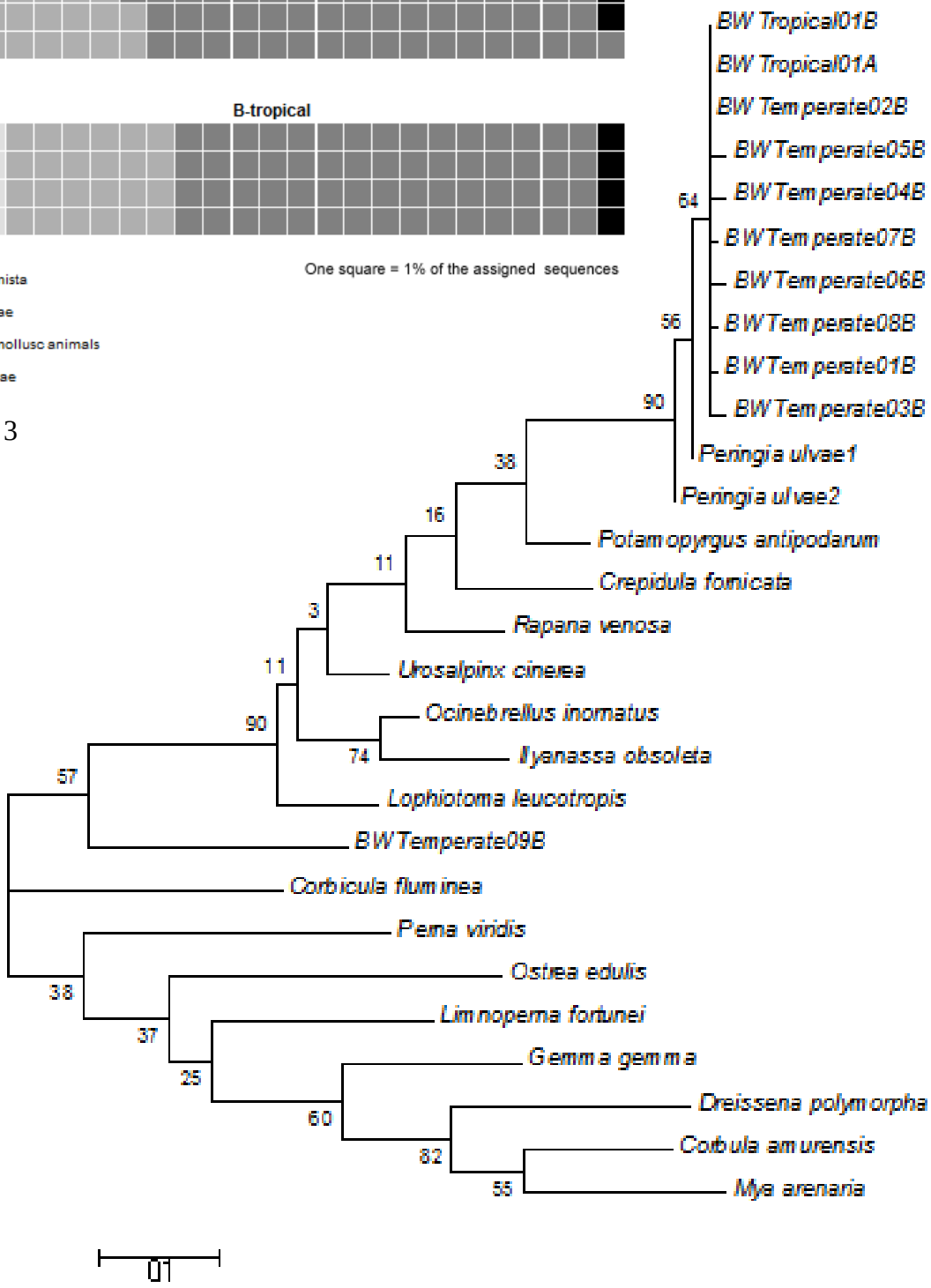


513

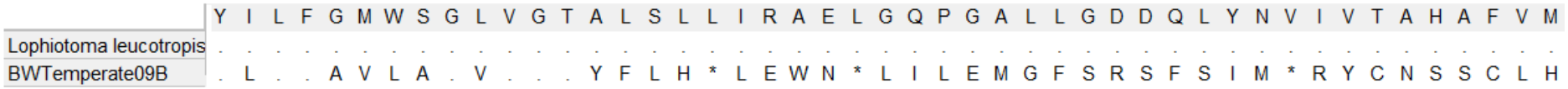
514 Figure 3

515

516



517 Figure 4


Y I L F G M W S G L V G T A L S L L I R A E L G Q P G A L L G D D Q L Y N V I V T A H A F V M
Lophiotoma leucotropis . . . . . .
BWTemperate09B . L . . A V L A . V . . . Y F L H * L E W N * L I L E M G F S R S F S I M * R Y C N S S C L H

518 Supplementary Table 1. Reference sequences of COI gene for different mollusc species
519 and sequences obtained from ballast water in this study. Accession numbers (AN) in the
520 GenBank and EMBL-EBI databases, respectively.

521

Reference species	**GenBank AN**
Corbicula fluminea	EU571247
Corbula amurensis	JQ267796
Crepidula fornicata	AF353129
Dreissena polymorpha	EF414493
Gemma gemma	KC429137
Hydrobia acuta	AF213344
Hydrobia glyca	AF467653
Hydrobia grimmi	GQ505913
Hydrobia knysnaensis	JX970611
Hydrobia neglecta	AF253081
Hydrobia ulvae 1	AF118308
Hydrobia ulvae 2	AF118290
Hydrobia ventrosa	AF118369
Ilyanassa obsoleta	KC759519
Limnoperna fortunei	AB828680
Lophiotoma leucotropis	HQ834093
Mya arenaria	JQ435826
Ocinebrellus inornatus	HM180493
Ostrea edulis	AF120651
Perna viridis	GQ480298
Potamopyrgus antipodarum	AY631101
Rapana venosa	JX503056
Urosalpinx cinerea	FN677423
Ballast water haplotypes	**EMBL-EBI AN**
BW Temperate01	HG963478
BW Temperate02	HG963479
BW Temperate03	HG963480
BW Temperate04	HG963481
BW Temperate05	HG963482
BW TEmperate06	HG963483
BW Temperate07	HG963484
BW Temperate08	HG963485
BW Tropical01	HG963486

522