Marine Genomics 45 (2019) 11–15

Contents lists available at ScienceDirect

Marine Genomics

journal homepage: www.elsevier.com/locate/margen

Population genomic footprints of environmental pollution pressure in natural populations of the Mediterranean mussel

Ângela M. Ribeiro[a], Carlos A. Canchaya[c], Fernando Penaloza[d], Juan Galindo[c,e], Rute R. da Fonseca[a,b,*]

[a] CIIMAR, Interdisciplinary Centre of Marine and Environmental Research, University of Porto, Portugal
[b] Center for Macroecology, Evolution and Climate, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark
[c] Department of Biochemistry, Genetics and Immunology, University of Vigo, Vigo, Spain
[d] Unidad de Investigación Biomédica en Cáncer, Instituto Nacional de Cancerología-Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Mexico City, Mexico
[e] UVigo Marine Research Centre (UVIGO-MRC), University of Vigo, Vigo, Spain

A B S T R A C T

Bivalve molluscs of the genus *Mytilus* are considered a model organism in ecotoxicology and are known to be well adapted to marine ecosystems affected by multiple anthropogenic factors, including pollution. In order to assess whether pollution interferes with the reproductive success of *Mytilus* and affects the diversity within and between populations, we sequenced the transcriptomes of 72 individuals from 9 populations of *Mytilus galloprovincialis* collected along a ca. 130-km north-south transect on the Western coast of the Iberian Peninsula. We found that polluted areas are acting as a barrier to gene flow, potentially because of the detrimental effect of anthropogenic chemicals on larvae carried from more pristine environments. Furthermore, we observed an increase in genetic diversity in populations from polluted site, which could be indicative of higher mutagenicity driven by the environment. We propose that a microevolutionary perspective is essential for a full characterization of human activities on the dispersal of *M. galloprovincialis* and that it should be incorporated into management, and conservation plans and policies in the context of the effects of pollution on populations.

1. Introduction

Marine ecosystem are affected by multiple anthropogenic factors, perhaps the most relevant being pollution, i.e. introduction of substances that are extraneous to the ecosystem and have harmful effects. Pollution causes stress to individuals affecting their reproductive success, therefore shaping the demographic trends of their corresponding populations and metapopulation dynamics. It is thus evident that ecological effects of pollutants can influence the evolutionary trend of the affected populations. Nevertheless, the microevolutionary perspective of pollution effects on populations has been overlooked by ecotoxicologists.

The life-history features of Bivalve molluscs of the genus *Mytilus* made them one of the model organism in ecotoxicology [e.g. (Azevedo et al. 2015; Bellas et al. 2011; Fernandes et al. 2009; Gorbi et al. 2008; Lima et al. 2007; Martins et al. 2014; Zouiten et al. 2016)]. International organizations, such as the Convention for the Protection of the Marine Environment of the North East Atlantic (the 'OSPAR Convention'), have recommend the use of *Mytilus* as a biomonitor i.e. an indicator of chemical pollution. A better understanding of what affects the diversity in *Mytilus* sp. also has direct economic relevance since their production represents 30% of the total aquaculture value in the European Union (Robert et al., 2013). This activity is highly dependent on the natural populations and its demographic dynamics, because the spats (young mussels) are collected from the wild and then farmed in closed conditions (Robert et al. 2013).

Mytilus galloprovincialis, the Mediterranean mussel, ranges from the Black Sea, continues through the Mediterranean basin and Iberian coast, until the Great Britain, occupying both pristine and highly polluted sites. Pollution causes stress to individuals in such a way that can affect their reproductive success and therefore the demographic trend of not only populations exposed to pollutants but also of populations that have a demographic relationship with them. At the population genetic level, responses to chemical contamination can be divided into four categories (Bickham 2011): 1) change in allele frequencies due to mutagenic effect of contaminants; 2) loss of genetic diversity across the genome caused by a reduction in the effective population size; 3) change in directionality and rates of gene flow among populations due to alteration in dispersal behaviour; 4) changes in the frequency of particular genotypes due to selective advantages: individuals with traits

* Corresponding author at: Center for Macroecology, Evolution and Climate, Natural History Museum of Denmark, University of Copenhagen, Universitetsparken 15, 2100 Copenhagen, Denmark.
E-mail address: rfonseca@snm.ku.dk (R.R. da Fonseca).

https://doi.org/10.1016/j.margen.2018.10.009
Received 11 August 2018; Received in revised form 9 October 2018; Accepted 30 October 2018
Available online 14 November 2018
1874-7787/

that promote survival and reproductive success in such stressful environments are favoured.

Examining anthropogenic effects on evolutionary processes can be challenging: natural and historic processes need to be separated from contemporary anthropogenic ones. RNA-Seq has been successfully used to assess diversity in marine species [e.g. (De Wit and Palumbi 2013; Lamichhaney et al. 2012; Ribeiro et al. 2017)] and is an adequate choice for a non-model organism (da Fonseca et al. 2016). Therefore, to determine the causal effect of pollution on the genetic diversity of mussels, we compared the transcriptomes of 72 individuals from nine populations (inhabiting three polluted areas intercalating six clean areas) found along a 130-km north-south transect in the Western coast of the Iberian Peninsula. The diversity results were then used to test for the effect of pollution while explicitly controlling for the effect of natural environmental variation such as sea water surface temperature, salinity and dissolved oxygen.

2. Methods

2.1. Selection of populations and environmental conditions

Mediterranean mussels measuring 35–50 mm were collected manually in February of 2013 from intertidal rocks in three regions: Porto, Viana do Castelo and Vigo (Fig. 1; Fig. S1). Each region comprises three sampling sites: one polluted and two clean sites located upstream and downstream of the polluted site (at similar geographical distances to control for potential confounding effects of demographic processes such as dispersal). The collection was performed in the same calendar week to minimize the effect of any possible environmental variation and physiological stage. The mussels were kept in seawater during transfer to the laboratory and dissected within 12 h. Tissue (mantle and branchia) was sampled from eight individuals per site ($n = 72$). Branchia and mantle were minced and preserved in RNALater (Qiagen) then stored until further used. To characterize the environment at the sampling sites, six abiotic variables (salinity, mean sea surface temperature, pH, dissolved O_2, nitrates, phosphates) were used, at a spatial resolution of 5 arcmin (9.2 km), as available at Bio-ORACLE database (Tyberghein et al. 2012). A PCA was performed in R 3.0 (R development team 2014) to describe the abiotic variation among sites and later account for its potential confounding effects (Fig. S6).

2.2. RNA extraction, sequencing and de novo transcriptome assembly

Equal quantities of tissue (mantle and branchia) from four individuals per site were pooled to provide material for RNA extraction. The experimental design, from tissue sampling to library construction is illustrated in Fig. 1C. Total RNA was extracted using the Qiagen RNeasy columns (Qiagen) following a procedure optimized by us based on the manufacturer's protocol. The integrity and quantity of the RNA was measured on a Qubit fluorometer (Invitrogen) and on the Agilent Bioanalyzer 2100 (Agilent). The Illumina TruSeq Kit v.2.0 was used to isolate the mRNA and prepare cDNA libraries for sequencing, following the standard protocol. Compatible indexes were assigned to individual libraries to allow for multiplexing on four lanes of 100 bp paired-end technology in an Illumina HiSeq 2000 cell flow. The sequencing of the cDNA libraries was done in the Sequencing Center of the University of Copenhagen in Denmark. We evaluated the quality of the raw data with FastQC v0.10.0 (Patel and Jain 2012). The raw data is available for download as part of BioProject PRJNA484309. After removing indexes and adaptors with CutAdapt (Martin 2011), we trimmed the reads with the FASTX-toolkit (http://hannonlab.cshl.edu/fastx_toolkit) removing bases with a Phred-scale quality score lower than 25.

We used the sample for which we generated the highest amount of reads (PR2) to build a reference transcriptome with Trinity (Haas et al. 2013). This software was used with the default settings including a fixed k-mer size of 25 as suggested by the authors. Furthermore, we estimated the completeness of the assembly as measured with CEGMA (Parra et al. 2007). When analysing transcriptome data, the coverage of contigs is uneven due to differences in gene expression, and several isoforms per gene can be represented in a single sample (Martin and Wang 2011). Additionally, chimeric contigs that can be misinterpreted as isoforms. For population genomics studies, only a single isoform per gene is desired, so we mapped the reads back to de novo transcriptome to select the most abundant isoform within each gene. In some instances, the de novo assembly algorithm fails to combine adjacent contigs from the same exon (Martin and Wang 2011) and thus creates as excessively larger number of components (genes) that are redundant. To overcome this redundancy we used CAP3 (Huang and Madan 1999) to merge components with overlap length of at least 100 bases and percent identity ≥ 99. This Transcriptome Shotgun Assembly project has been deposited at DDBJ/ENA/GenBank under the accession GGUW00000000. The version described in this paper is the first version, GGUW01000000. Reads from all libraries were mapped using TopHat2 (Kim et al. 2013) using only bases with a quality larger than the Phred-scale quality score 33, providing information about the method used to produce the library (fr-unstranded) to help identify splice junctions, while allowing only for two mismatches with the reference transcriptome. Levels of differentiation between populations were calculated using F_{ST} and departures from neutrality were estimated with Tajima's D (Tajima 1989) using approaches implemented in Popoolation2 (Kofler et al. 2011) (use all SNPs for which at least 6 reads supported the minor allele for both population simultaneously, and with a coverage depth ranging between 20 and 200 reads).

3. Results and discussion

The de novo assembled transcriptome included 205,267 possible transcripts. Choosing the longest transcript per component resulted in a total of 46,547 transcripts being used as the final reference dataset. The N50 of this dataset is 1090 bp, with a total assembly length of 109,339,607 bp and a high CEGMA completeness of 92.34%. After adaptor and quality trimming, we retained and average of 125.8 million paired- end reads per pool for the population genetic analyses (depth of coverage per pool in Table 1).

3.1. Genetic variability and population structure in populations of Mytilus galloprovincialis

It is often assumed that marine invertebrates with pelagic larvae have high dispersal; however, compelling studies revealed that *M. galloprovincialis* tends to recruit within 5 km (McQuaid and Phillips 2000) to 30 km (Becker et al. 2007) from their natal populations. The quantification of genetic differentiation can provide insights into its larvae dispersal and connectivity between populations. Knowledge of these patterns is fundamental for understanding the evolution and ecology of species and for coastal resource management. We chose to sample Mediterranean mussel populations from the Atlantic coast of Iberia because of three main reasons: i) it is an intermediate location in the native distribution of this species and far from their major phylogeographical break (Diz and Presa 2008; Luis et al. 2011) (Fig. S2); ii) geographically distant from the known area of introgression with *Mytilus edulis* (Bierne et al. 2003); iii) there are consecutive areas characterized by low and high pollution intensity, within biologically relevant distances.

F_{ST} values between populations were generally low, ranging between 0.029 and 0.061. We found no evidence of isolation by distance, as geographical distance between study sites does not correlate with genetic distances (Mantel test = −0.009, $p = 0.471$; Fig. S3) which suggests that 130 km might not be enough to detect population structure. We did find a correlation between the sea surface temperature with genetic differentiation (Mantel test = 0.3502, $p = 0.038$; Fig. S4) suggesting that populations inhabiting similar ecological conditions are

Fig. 1. Location of the nine sampling sites along the Western coast of the Iberian Peninsula. Sites depicted in red for polluted sites and blue for clean sites. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

more closely related than populations in different ecological settings.

Next, we compared allele frequencies between polluted sites and their adjacent southern clean sampling sites. Considering that the oceanic currents in Western Iberia are in the direction from North to South, we would expect that there would similar allele frequencies differences between locations that are within similar distances. Instead, we find that the differences between the northern clean populations and those of their adjacent polluted sites are larger than those of the polluted sites and their southern clean neighbours ($p < 2.2e^{-16}$; Fig. 2), which could results from the negative effect of pollution on larvae survival. Furthermore, genetic diversity in polluted areas, measured as the average number of pairwise differences, was shown to be significantly higher than in clean areas (Fig. S5), which could be a consequence of the mutagenic effect of pollution.

4. Conclusions

The efforts to incorporate the effects of pollution into management and conservation plans and policies reflects the awareness that humans have serious impacts on marine ecosystems at the ecological scale. The ecological effects of pollutants certainly influence the evolutionary trend of populations, and should be incorporated in marine conservation plans, but its microevolutionary impact is usually overlooked by ecotoxicologists. Here we show that polluted areas could be acting as effective barriers to dispersal, behaving as ecological sinks, and reducing gene flow between populations. Furthermore, polluted areas effectively harbour populations with higher genetic diversity, potentially resulting from the detrimental mutagenic effects of anthropogenic chemicals. Thus, our study suggests a role for human impact in the

Table 1
Sampling locations and sequencing statistics (more detailed information in Table S2).

#	Location	Short id	Type	Replicate	Id	Number of reads	Average depth	Number of covered sites
1	Cabo Home	CH	clean	1	CH1	9,086,697	14.422	28,473,007
				2	CH2	49,754	2.354	1,485,840
2	Vigo	VG	polluted	1	VG1	19,478	1.914	310,892
				2	VG2	33,083,322	35.058	34,500,341
3	Pedra Rubia	PR	clean	1	PR1	24,450,095	29.768	34,638,383
				2	PR2	81,240,601	72.462	45,505,478
4	Gelfa	GL	clean	2	GL2	8,432,331	15.19	29,558,000
5	Viana	VN	polluted	1	VN1	19,483,504	20.689	30,392,919
				2	VN2	7,073,777	14.347	25,704,716
6	Sao Bartolomeu	SB	clean	1	SB1	1,392,837	8.634	11,566,527
				2	SB2	6,206,869	12.229	27,394,790
7	Vila Cha	VC	clean	1	VC1	4,621,955	11.898	21,341,011
				2	VC2	915,570	7.046	9,634,047
8	Matosinhos	MT	polluted	1	MT1	4,231,619	10.823	22,072,464
				2	MT2	10,190,920	15.888	29,015,489
9	Valadares	VL	clean	1	VL1	24,841,399	28.376	26,085,891
				2	VL2	26,245,409	29.444	30,687,606

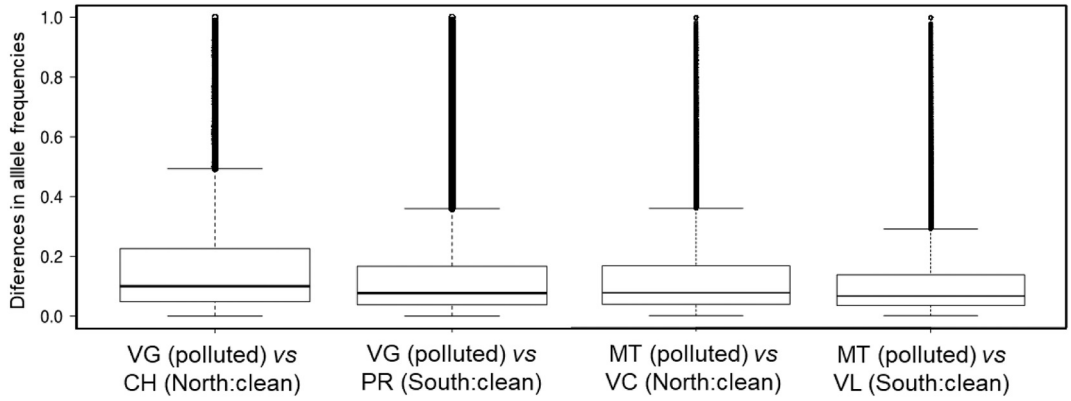

Fig. 2. Differences in allele frequencies between selected populations. There are larger differences in allele frequencies between a Northern population from a clean site and the adjacent polluted site, that between the latter and its Southern clean site, indicative of an effect of pollution on larvae survival. Kruskal-Wallis (VG comparisons) = 6762.876, df = 1, $p < 2.2e^{-16}$; Kruskal-Wallis (MT comparisons) = 2574.928, df = 1, $p < 2.2e^{-16}$.

dispersal patterns of *M. galloprovincialis*.

Authors' contributions

RRF conceived and coordinated the project with input from AMR. AMR collected the samples and performed the lab work with input from JG. AMR and RRF analyzed the data with contributions from CAC and FP. AMR designed and performed the statistical analysis. AMR and RF wrote the paper with critical input from all authors, who read and approved the final manuscript.

Acknowledgements

This work was supported by: the Portuguese Science Foundation (PTDC/MAR/115347/2009, COMPETE-FCOMP-01-012; FEDER-015453), and Marie Curie (FP7-PEOPLE-2010-IEF, Proposal 272927). R.R.F. thanks the Danish National Research Foundation for its support of the Center for Macroecology, Evolution, and Climate (grant DNRF96) and the Villum Fonden for the Young Investigator Grant VKR023446.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.margen.2018.10.009.

References

Azevedo, C.C., Guzmán-Guillén, R., Martins, J.C., Osório, H., Vasconcelos, V., da Fonseca, R.R., Campos, A., 2015. Proteomic profiling of gill GSTs in Mytilus galloprovincialis from the North of Portugal and Galicia evidences variations at protein isoform level with a possible relation with water quality. Mar. Environ. Res. 110, 152–161. https://doi.org/10.1016/j.marenvres.2015.08.008.

Becker, B.J., Levin, L.A., Fodrie, F.J., McMillan, P.A., 2007. Complex larval connectivity patterns among marine invertebrate populations. Proc. Natl. Acad. Sci. U. S. A. 104, 3267–3272. https://doi.org/10.1073/pnas.0611651104.

Bellas, J., González-Quijano, A., Vaamonde, A., Fumega, J., Soriano, J.A., González, J.J., 2011. PCBs in wild mussels (Mytilus galloprovincialis) from the N-NW Spanish coast: current levels and long-term trends during the period 1991–2009. Chemosphere 85, 533–541. https://doi.org/10.1016/j.chemosphere.2011.08.017.

Bickham, J.W., 2011. The four cornerstones of evolutionary toxicology. Ecotoxicology 20, 497–502. https://doi.org/10.1007/s10646-011-0636-y.

Bierne, N., Borsa, P., Daguin, C., Jollivet, D., Viard, F., Bonhomme, F., David, P., 2003. Introgression patterns in the mosaic hybrid zone between Mytilus edulis and M. galloprovincialis. Mol. Ecol. 12, 447–461.

De Wit, P., Palumbi, S.R., 2013. Transcriptome-wide polymorphisms of red abalone (*Haliotis rufescens*) reveal patterns of gene flow and local adaptation. Mol. Ecol. 22, 2884–2897. https://doi.org/10.1111/mec.12081.

Diz, A.P., Presa, P., 2008. Regional patterns of microsatellite variation in Mytilus galloprovincialis from the Iberian Peninsula. Mar. Biol. 154, 277–286. https://doi.org/10.1007/s00227-008-0921-3.

Fernandes, S., Welker, M., Vasconcelos, V.M., 2009. Changes in the GST activity of the mussel *Mytilus galloprovincialis* during exposure and depuration of microcystins. J. Exp. Zool. A Ecol. Genet. Physiol. 311A, 226–230. https://doi.org/10.1002/jez.524.

da Fonseca, R.R., Albrechtsen, A., Themudo, G.E., Ramos-Madrigal, J., Sibbesen, J.A., Maretty, L., Zepeda-Mendoza, M.L.L., Campos, P.F., Heller, R., Pereira, R.J., 2016. Next-generation biology: sequencing and data analysis approaches for non-model organisms. Mar. Genomics 30, 3–13. https://doi.org/10.1016/j.margen.2016.04.012.

Gorbi, S., Virno Lamberti, C., Notti, A., Benedetti, M., Fattorini, D., Moltedo, G., Regoli,

F., 2008. An ecotoxicological protocol with caged mussels, Mytilus galloprovincialis, for monitoring the impact of an offshore platform in the Adriatic sea. Mar. Environ. Res. 65, 34–49. https://doi.org/10.1016/J.MARENVRES.2007.07.006.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., Leduc, R.D., Friedman, N., Regev, A., 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8, 1494–1512. https://doi.org/10.1038/nprot.2013.084.

Huang, X., Madan, A., 1999. CAP3: a DNA sequence assembly program. Genome Res. 9, 868–877.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., Salzberg, S.L., 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14, R36. https://doi.org/10.1186/gb-2013-14-4-r36.

Kofler, R., Pandey, R.V., Schlotterer, C., 2011. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics 27, 3435–3436. https://doi.org/10.1093/bioinformatics/btr589.

Lamichhaney, S., Martinez Barrio, A., Rafati, N., Sundström, G., Rubin, C.-J., Gilbert, E.R., Berglund, J., Wetterbom, A., Laikre, L., Webster, M.T., Grabherr, M., Ryman, N., Andersson, L., 2012. Population-scale sequencing reveals genetic differentiation due to local adaptation in Atlantic herring. Proc. Natl. Acad. Sci. U. S. A. 109, 19345–19350. https://doi.org/10.1073/pnas.1216128109.

Lima, I., Moreira, S.M., Osten, J.R.-V., Soares, A.M.V.M., Guilhermino, L., 2007. Biochemical responses of the marine mussel Mytilus galloprovincialis to petrochemical environmental contamination along the North-western coast of Portugal. Chemosphere 66, 1230–1242. https://doi.org/10.1016/J.CHEMOSPHERE.2006.07.057.

Luis, J.R., Comesaña, A.S., Sanjuan, A., 2011. mtDNA differentiation in the mussel Mytilus galloprovincialis Lmk. On the Iberian Peninsula coast: first results. Mar. Ecol. 32, 102–106. https://doi.org/10.1111/j.1439-0485.2011.00430.x.

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. 17, 10–12.

Martin, J.A., Wang, Z., 2011. Next-generation transcriptome assembly. Nat. Rev. Genet. 12, 671–682. https://doi.org/10.1038/nrg3068.

Martins, J.C., Campos, A., Osório, H., da Fonseca, R., Vasconcelos, V., 2014. Proteomic profiling of cytosolic glutathione transferases from three bivalve species: Corbicula fluminea, Mytilus galloprovincialis and Anodonta cygnea. Int. J. Mol. Sci. 15, 1887–1900. https://doi.org/10.3390/ijms15021887.

McQuaid, C., Phillips, T., 2000. Limited wind-driven dispersal of intertidal mussel larvae: in situ evidence from the plankton and the spread of the invasive species Mytilus galloprovincialis in South Africa. Mar. Ecol. Prog. Ser. 201, 211–220. https://doi.org/10.3354/meps201211.

Parra, G., Bradnam, K., Korf, I., 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics 23, 1061–1067. https://doi.org/10.1093/bioinformatics/btm071.

Patel, R.K., Jain, M., 2012. NGS QC toolkit: a toolkit for quality control of next generation sequencing data. PLoS One 7, e30619. https://doi.org/10.1371/journal.pone.0030619.

Robert, R., Sanchez, J.L., Perez-Paralle, L., Ponis, E., Kamermans, P., O'Mahoney, M., 2013. A glimpse on the mollusc industry in Europe. Aquaculture 38, 5–11 (ISSN 1018-9661).

Ribeiro, Â.M., Foote, A.D., Kupczok, A., Frazão, B., Limborg, M.T., Piñeiro, R., Abalde, S., Rocha, S., da Fonseca, R.R., 2017. Marine genomics: news and views. Mar. Genomics 31, 1–8. https://doi.org/10.1016/j.margen.2016.09.002.

Robert, R., Sanchez, J.L., Perez-Paralle, L., Ponis, E., Kamermans, P., O'Mahoney, M., 2013. A glimpse on the mollusc industry in. Europe. Aquac. Soc. 38, 5–11.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F., De Clerck, O., 2012. Bio-ORACLE: a global environmental dataset for marine species distribution modelling. Glob. Ecol. Biogeogr. 21, 272–281. https://doi.org/10.1111/j.1466-8238.2011.00656.x.

Zouiten, A., Beltifa, A., Van Loco, J., Mansour, H. Ben, Reyns, T., 2016. Ecotoxicological potential of antibiotic pollution–industrial wastewater: bioavailability, biomarkers, and occurrence in Mytilus galloprovincialis. Environ. Sci. Pollut. Res. 23, 15343–15350. https://doi.org/10.1007/s11356-016-6713-2.