eman ta zabal zazu

Universidad del País Vasco

Euskal Herriko Unibertsitatea

Doktorego-tesia

Development of new analytical tools for verifying the geographical origin of farmed Mediterranean mussels (*Mytilus galloprovincialis*)

Ane del Rio Lavín

2023

Zuzendariak:

Miguel Ángel Pardo

Elisa Jiménez

Doktorego-tesia

Development of new analytical tools for verifying the geographical origin of farmed Mediterranean mussels (*Mytilus galloprovincialis*)

Autorea

Ane del Rio Lavín

Zuzendariak

Miguel Ángel Pardo

Elisa Jiménez

Saila

Zoologia eta Animalia Zelulen Biologia

Doktorego programa

Itsas Baliabideak eta Ingurugiroa

2023ko Urtarrila

eman ta zabal zazu

Universidad del País Vasco

Euskal Herriko Unibertsitatea

AZTI

MEMBER OF BASQUE RESEARCH & TECHNOLOGY ALLIANCE

The research carried out in this Philosophiae Doctor thesis has been developed in AZTI (Derio, Spain) and it has been financed by SEA-TRACES (EAPA_87/2016) INTERREG Atlantic Area EU project, by the ORIME FEMP Project (00001-1NA2018-33) from the Department of Aquaculture and Fisheries of the Basque Government and by the SEAFOOD-ID Project (PID2020-118012RB-C22) which is funded by MCIN/AEI/10.13039/501100011033 and the European Union. Ane del Rio Lavín was funded by a Doctoral Fellowship ("Ayudas nuevas para el Programa Predoctoral de Formación de Personal Investigador no Doctor") from the Education Department of the Basque Government.

Doktorego-tesi hau AZTIn (Derion) burutu da. Lanak Europako SEA-TRACES (EAPA_87/2016) INTERREG proiektuaren, Eusko Jaurlaritzako Arrantza eta Akuikultura Saileko ORIME FEMP (00001-1NA2018-33) proiektuaren, eta Espainiako Gobernuak (MCIN/AEI/10.13039/501100011033) eta Europar batasunak finatzatutako SEAFOOD-ID (PID2020-118012RB-C22) proiektuaren diru-laguntzak jaso ditu. Ane del Rio Lavín Eusko Jaurlaritzaren ikertzaile ez-doktoreen prestakuntzarako doktoratu aurreko laguntzaren onuraduna izan da.

Recommended citation:
Del Rio-Lavín, A. (2023) Development of new analytical tools for verifying the geographical origin of farmed Mediterranean mussels (*Mytilus galloprovincialis*). PhD Thesis. Department of Zoology and Animal Cell Biology, University of the Basque Country.

Cover design was performed using Canva (www.canva.com). Illustrations and images that open some of the sections were also taken from this source.

Etxekoei

"Above all, don't fear difficult moments.
The best comes from them."

Rita Levi-Montalcini

AURKIBIDEA - CONTENTS

LABURDUREN ZERRENDA - LIST OF ACRONYMS

δ13C	Carbon stable isotope ratio / Karbono isotopo egonkorren ratioa
δ15N	Nitrogen stable isotope ratio / Nitrogeno isotopo egonkorren ratioa
AFLP	Amplified Fragment Length Polymorphisms / Anplifikatutako-zatien luzerako polimorfismoak
AGES	Ría de Arousa location (Galicia, Spain) / Arousako itsasadarra kokalekua (Galizia, Espainia)
Al	Aluminum / Aluminio
ANOVA	Analysis of variance / Bariantza-analisia
AOOF	Almeria-Oran Oceanographic Front / Almeria-Oran Fronte Ozeanografikoa
As	Arsenic / Artseniko
ASC	Aquaculture Stewardship Council
ATL	Atlantic area / Eremu atlantikoa
B	Boron / Boro
Ba	Barium /Bario
bp	Base pair / Base-pare
C	Carbon / Karbono
Cd	Cadmium / Kadmio
Ce	Cerium / Zerio
CI	Confidence Interval / Konfiantza-tartea
Co	Cobalt / Kobalto
CO_2	Carbon dioxide / Karbono dioxido
COCL	Coliumo location (Biobío, Chile) / Coliumo kokalekua (Biobío, Txile)
Cr	Chromium / Kromo
Cu	Copper / Kobre
DEES	Delta del Ebro location (Catalonia, Spain) / Ebroko Delta kokalekua (Katalunia, Espainia)
DFA	Discriminant Function Analysis / Bereizketa-funtzioaren analisia
DG	Digestive gland / Digestio-guruina
DGGE	Denaturing Gradient Gel Electrophoresis / Gradiente desnaturalizatzaile gel elektroforesia
DNA	Deoxyribonucleic Acid / Azido Desoxirribonukleiko
dNTPs	Deoxynucleotide triphosphates / Desoxinukleotido-trifosfatoak
EU / EB	European Union / Europar Batasuna
Fe	Iron / Burdina
GEIT	Goro location (Emilia-Romagna, Italy) / Goro kokalekua (Emilia-Romagna, Italia)

H_2O_2	Hydrogen peroxide / Hidrogeno peroxido
He	Helium / Helio
HNO_3	Nitric acid / Azido nitriko
HRM	High Resolution Melting / Bereizmen handiko urtzea
IAEA	International Atomic Energy Agency / Energia Atomikoaren Nazioarteko Agentzia
ICPMS	Inductively Coupled Plasma Mass Spectrometry / Induktiboki akoplatutako plasma iturridun masa-espektrometroa
LBFR	Loquemeau location (Brittany, France) / Loquemeau kokalekua (Bretainia, Frantzia)
LBTN	Bizerte lagoon location (Bizerte, Tunisia) / Bizertako urmaela kokalekua (Bizerta, Tunisia)
Li	Lithium / Litio
MAF	Minimum Allele Frequency / Alelo txikienaren frekuentzia
MDS	Multidimensional Scaling / Mailaketa multidimentsionala
MED	Mediterranean area / Mediterraneoko eremua
MEES	Mendexa location (Basque Country, Spain) / Mendexa kokalekua (Euskal Autonomia Erkidegoa, Espainia)
Mg	Magnesium / Magnesio
$MgCl_2$	Magnesium chloride / Magnesio kloruro
Mn	Manganese / Manganeso
MSC	Marine Stewardship Council
MUES	Mutriku location (Basque Country, Spain) / Mutriku kokalekua (Euskal Autonomia Erkidegoa, Espainia)
N	Nitrogen / Nitrogeno
NCBI	National Center for Biotechnology Information/ Bioteknologiaren Informaziorako Zentro Nazionala
NGS	Next-Generation Sequencing / Sekuentziazio masiboa
Ni	Nickel / Nikel
NMDS	Non-Metric Multidimensional Scaling / Mailaketa Multidimentsional ez-metrikoa
NOx	Nitrogen oxides / Nitrogeno-oxidoak
OTU	Operational Taxonomical Unit
Pb	Lead / Berun
PBPT	Porto da Baleeira location (Algarve, Portugal) / Porto da Baleeira kokalekua (Algarve, Portugal)
PCA	Principal Component Analysis / Osagai nagusien bidezko analisia
PCR	Polymerase Chain Reaction / Polimerasaren kate-erreakzioa
PDO	Protected Designation of Origin / Jatorri-deitura babestua
PERMANOVA	Permutational Multivariate Analysis of Variance / Bariantzaren permutazio-bidezko aldagai anitzeko analisia

PGI	Protected Geographical Indication / Adierazpen geografiko babestua
qPCR	Real-time Polymerase Chain Reaction / Denbora errealeko kate-erreakzioa
RAD-seq	Restriction Site-Associated DNA Sequencing / Murriztegunerari loturiko DNAren sekuentziazioa
RF	Random Forest
RFLP	Restriction Fragment Length Polymorphism / Murriztapen-zatien luzerako polimorfismoa
rRNA	Ribosomal Ribonucleic Acid / Azido Erribonukleiko erribosomiko
SD	Standard Deviation / Desbideratze estandarra
Se	Selenium / Selenioa
SEP	South-eastern Pacific area / Ozeano Bareko hego-ekialdeko eremua
SGES	Ría de Betanzos-Sada location (Galicia, Spain) / Betanzos-Sada-ko itsasadarra (Galizia, Espainia)
SIRA	Stable Isotope Ratio Analysis / Isotopo egonkorren analisia
SNP	Single Nucleotide Polymorphisms / Nukleotido bakarreko polimorfismoak
TBCL	Caleta Tumbes location (Biobío, Chile) / Caleta Tumbes kokalekua (Biobío, Txile)
TEF	Trace Element Fingerprinting / Aztarna-elementuen analisia
THL	Training, Holdout & Leave-One-Out
Ti	Titanium / Titanio
Tl	Thallium / Talio
T_m	Melting Temperature / Urtze-tenperatura
V	Vanadium / Banadio
Y	Yttrium / Itrio
Zn	Zinc / Zink

EKOIZPEN ZIENTIFIKOA - SCIENTIFIC PRODUCTION

Argitalpenak - Publications

del Rio-Lavín, A., Jiménez, E., and Pardo, M. A. (2021). SYBR-Green real-time PCR assay with melting curve analysis for the rapid identification of *Mytilus* species in food samples. *Food Control*, 130: 108257. https://doi.org/10.1016/j.foodcont.2021.108257

del Rio-Lavín, A., Díaz-Arce, N., Larraín, M. A., Araneda, C., Rodríguez-Ezpeleta, N., Jiménez, E., and Pardo, M. A. (2022). Population structure and geographic origin assignment of *Mytilus galloprovincialis* mussels using SNPs. *Aquaculture*, 550. https://doi.org/10.1016/j.aquaculture.2021.737836

del Rio-Lavín, A., Weber, J., Molkentin, J., Jiménez, E., Artetxe-Arrate, I., and Pardo, M. A. (2022). Stable isotope and trace element analysis for tracing the geographical origin of the Mediterranean mussel (*Mytilus galloprovincialis*) in food authentication. *Food Control*, 139. https://doi.org/10.1016/j.foodcont.2022.109069

Kongresuak - Conferences

del Rio-Lavín, A., Jiménez, E., Pardo, M.A. (2019). Novel and rapid method for the identification of European mussel species based on real time PCR melting curve analysis. Aquaculture Europe 2019, Berlin (Germany), 7-10th October 2019. [Poster]

del Rio-Lavín, A., Jiménez, E., Pardo, M.A. (2019). New molecular tool for tracing geographical origin of farmed mussels. 49th WEFTA Conference, Faroe Islands (Kingdom of Denmark), 14-18th October 2019. [Oral communication]

del Rio-Lavín, A., Jiménez, E., Pardo, M.A. (2020). Development of new molecular tools for tracing the geographical origin of farmed mussels (*Mytilus galloprovincialis*). International Postgraduate Course Research in Marine Environment & Resources RiMER-2020, Donostia-San Sebastián (Spain), 3-7th February 2020. [Oral communication]

del Rio-Lavín, A., Weber, J., Molkentin, J., Jiménez, E, and Pardo, M. A. (2021). Tracing the geographical origin of the Mediterranean mussel (*Mytilus galloprovincialis*) in food authentication using stable isotope and trace element analysis. Aquaculture Europe 2020, Online, 12-15th April 2021. [Oral communication]

del Rio-Lavín, A., Jiménez, E., Muñoz-Colmenero M., Velasco A., Sotelo C.G., Lee, R-S, Longa, M.A., Larraín, M.A., Araneda, C., Pardo, M.A. (2021). Next generation sequencing and trace element fingerprinting approaches for tracing provenance of mussels. World Fisheries Congress 2021, Online-Adelaide (Australia), 20-24th September 2021. [Oral communication]

del Rio-Lavín, A., Jiménez, E., Pardo, M.A. (2021). Geographic origin assignment of *Mytilus galloprovincialis* mussels using reduced SNP panels. ISLAS 2021, Online-Vigo (Spain), 20-21st October 2021. [Oral communication]

del Rio-Lavín, A., Jiménez, E., Pardo, M.A. (2022). Development of new molecular tools for tracing the geographical origin of farmed mussels (*Mytilus galloprovincialis*). International Postgraduate Course Research in Marine Environment & Resources RiMER-2022, Donostia-San Sebastián (Spain), 31-4th February 2022. [Oral communication]

del Rio-Lavín, A., Monchy, S., Jiménez, E., Pardo, M.A. (2022). Marcadores microbiológicos como herramienta para determinar el origen geográfico del mejillón Mediterráneo (*Mytilus gallprovincialis*) cultivado. Mikrobiogune, II Basque Microbiology Meeting, Bilbao (Spain), 13th December 2022. [Oral communication]

LABURPENA

Merkatuen globalizazioa eta itsas produktuen eskari hazkuntza dela medio, hornidura kateak gero eta konplexuagoak bilakatzen ari dira, trazabilitate ahaleginak erronka handia bihurtuz. Testuinguru honetan, arrain eta itsaskien etiketan ageri den informazioa zuzena dela egiaztatzeko tresna analitiko zehatzak garatzea ezinbestekoa da. Jatorri geografikoa egiaztatzea funtsezkoa da elikagaien trazabilitate eta segurtasun legeak betetzen direla ziurtatzeko, itsas baliabideen kudeaketa eraginkor eta jasangarria bermatzeko, bai eta agintari, ekoizle eta kontsumitzaile arduratsuen egungo eta etorkizuneko itxaropenak betetzeko ere. Hori dela eta, tesi hau *Mytilus galloprovincialis* muskuiluen jatorri geografikoa egiaztatzeko tresna analitikoen garapenera bideratuta dago, produktu fresko zein prozesatu hauen integritatea bermatzeko. Helburu hau lortzeko, tesia lau kapitulutan banatu da.

Itsas produktuen etiketan ageri den informazioa zuzena dela bermatzeko, jatorri geografikoa egiaztatu aurretik, arrain edo itsaskiaren espeziea egiaztatzea da egin beharreko lehengo pausua. Hori dela eta, **1. kapituluan**, Europar Batasunean merkaturatzen diren *Mytilus* espezie garrantzitsuenak identifikatzeko denbora errealeko PCR metodo azkar, fidagarri eta errentagarri bat garatu da. Muskuilu espezie ezberdinen bisuko proteina itsasgarriak kodetzen dituen genea anplifikatuz, eta DNA produktu horien disoziazio-kurbak aztertuz, muskuilu fresko, izoztu edo kontserben espeziea % 100eko espezifikotasunaz, sentikortasunaz eta zehaztasunaz identifikatu da. Gainera, metodo hau termoziklagailu ezberdinetan eta *software* berezien beharrizanik gabe erabil daitekeela erakutsi dute emaitzek. Hortaz, kapitulu honetan deskribaturiko metodologia muskuilu espeziearen autentifikaziorako erabilgarria da eta gainera, erraz erabil daiteke elikagaien segurtasuna eta identitatea bermatzen duten kontrol laborategietan, ohiko analisi azkar gisa.

Indibiduoen espeziea egiaztatu ondoren, jatorria egiaztatzeko tresna analitikoen erabilera ikertu da. Horretarako, proposatutako lehenengo metodologia markatzaile genetikoen erabileran oinarritzen da (**2. kapitulua**). Kapitulu honetan, RAD-seq

teknikaren bidez, muskuiluen genoman zehar sakabanatutako milaka nukleotido bakarreko polimorfismo (SNP) markatzaile identifikatu dira, eta hauek erabiliz, *M. galloprovincialis* bibalbioen populazio egitura argitu da bere jatorrizko banaketa eremuan zehar. Gainera, aurkitu berri diren markatzaile genetiko hauen jatorria esleitzeko ahalmena aztertu da. Emaitzek Atlantikoko eta Mediterraneoko *M. galloprovincialis* populazioak nabarmen bereizten direla erakutsi dute. Aldi berean, nahiz eta antzinako populazio beretik etorri, Mediterraneoko muskuilu natiboen eta sartu berri diren Ozeano Bareko hego-ekialdeko muskuiluen artean ezberdintasun esanguratsuak daudela ikusi da. Gainera, ahalmen diskriminatzaile handiena agertu duten 10-25 SNP erabiliz, testatutako indibiduoen % 90-100a Atlantikoko edo Mediterraneoko/Hego-ekialdeko jatorri eremuetara era egokian esleitu dira.

3. kapituluan, *M. galloprovincialis* muskuiluen jatorri geografikoa egiaztatzeko, aztarna elementuen eta isotopo egonkorren konposizioa, era independentean zein konbinatuan, erabiltzearen potentziala aztertu da. Analisirako, laginak Mediterraneo itsasoko, Europako kostalde Atlantikoko eta Ozeano Bareko hego-ekialdeko zortzi eskualde ezberdinetatik hartu dira. Emaitzen arabera, alde nabarmenak daude akuiluktura-instalazio ezberdinetan hazitako muskuiluen konposizio kimikoen artean, eta bereziki eskualde geografiko desberdinetakoak direnen artean. Ikaste automatikoan oinarritutako *Random Forest* sailkapen-metodoa erabiliz, aztarna kimikoen (isotopo egonkor eta aztarna-kimikoen) datuen konbinazioak, jatorria zehaztasun handiz aurreikusteko gaitasunik handiena erakutsi du; testatutako indibiduoen % 97a behar bezala esleitu dira, jatorriaren bereizketarako elementu garrantzitsuenak $\delta^{13}C$, $\delta^{15}N$, Pb, Ba, Mn, eta Al izanik.

4. kapituluan, *M. galloprovincialis* muskuilu freskoen digestio-guruinean dauden bakterio-komunitateen profilak aztertu dira, hazkuntza jatorria bereizteko trazabilitate tresna gisa erabil ote daitezkeen ikertzeko. Horretarako, 2019an, Espainiako hiru eskualde ezberdinetan (Galizian, Euskal Autonomia Erkidegoan eta Katalunian) kokatutako bost akuikultura-instalazio ezberdinetatik laginak bildu ziren urtaroka, eta haien bakterio-komunitatearen konposizioa aztertu zen 16S rRNA genearen V3-V4 gune aldakorraren anplifikazio eta sekuentziazioaren bitartez. Emaitzek bakterio-

komunitateen profilak jatorriaren eta urtaroaren arabera nabarmen aldatzen direla erakutsi dute, jatorri geografikoaren eragina urtaroen aldakortasuna gainditzen duelarik. Trazabilitate tresna honen egonkortasuna eta potentziala aztertzeko, 2020ko udazkenean EAEn bildutako 20 indibiduo berriren bakterio profila (laginak era independentean prozesatu, sekuentziatu eta analizatu ostean lortua), 2019an lortutako profilekin alderatu zen. Emaitzek, EAEn bi urtez jarraian bildutako muskuiluak elkarrekin biltzen direla erakutsi dute, baita uzta-denboraldiarekin bat datozela ere.

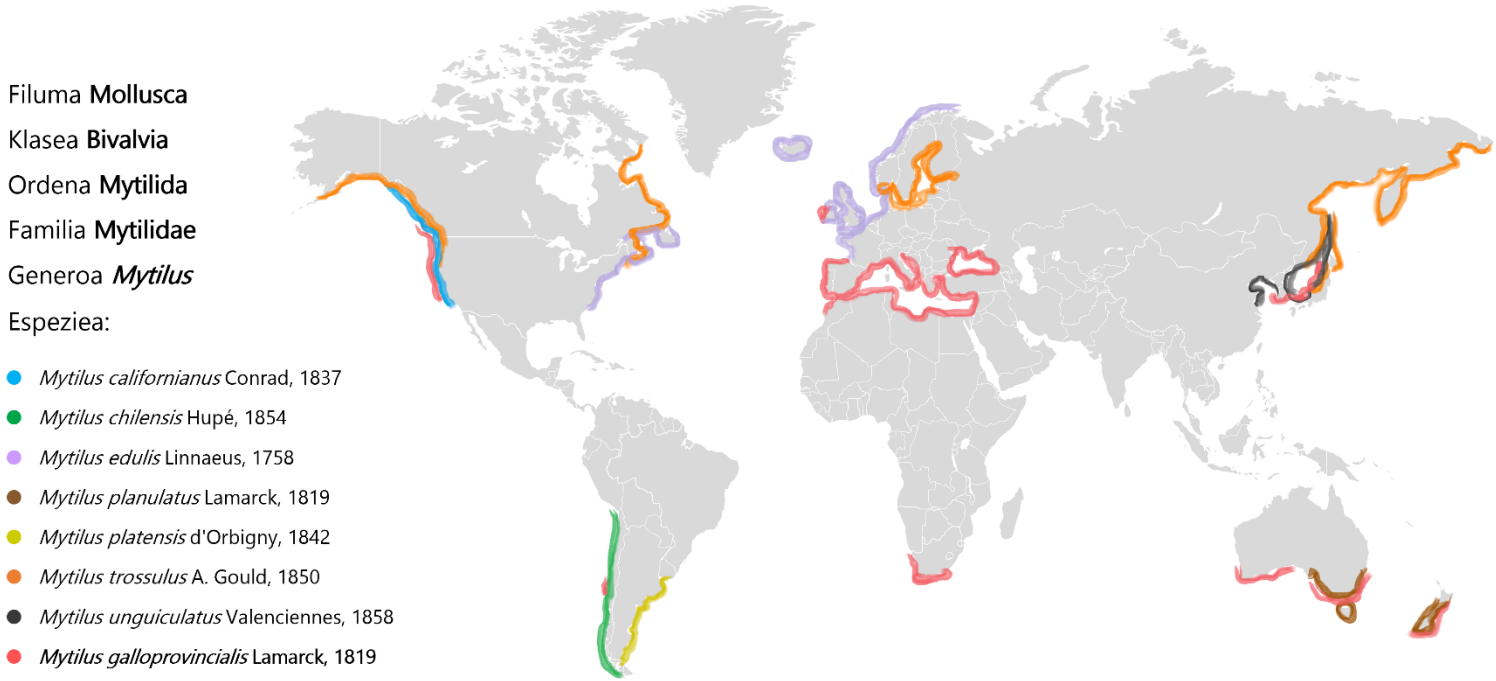
Lan honetan aztertutako tresnak hazitako *M. galloprovincialis* muskuiluen jatorri geografikoa egiaztatzeko egokiak direla frogatu da, eta beraz, funtsezko zeregina izan dezaketela merkataritza globalizatuaren gardentasuna eta segurtasuna bermatzeko. Halaber, azpimarratzekoa da momentuz ez dagoela tresna perfekturik, denek agertzen baitituzte mugak. Beraz, aztertu nahi den espeziearen eta produktu motaren arabera (bizirik edo prozesatuta), tresna bakoitzaren egokitasuna aztertu beharko da.

SARRERA OROKORRA

1. *Mytilus galloprovincialis*: taxonomia eta banaketa-eremua

Mytilus generoko muskuiluak munduan zehar oso hedatuta dauden bibalbio janari-filtratzaileak dira (Gosling, 2003). Ipar eta hego hemisferioko ur hotz eta epeletako itsasertz arrokatsuetan ageri dira (Seed and Suchanek, 1992; Hilbish et al., 2000) (**1. irudia**). Bibalbio hauek ingeniari ekologiko gisa (Arribas et al., 2014), kutsaduraren biomarkatzaile gisa (Hagger et al., 2008; Garmendia et al., 2010; Grbin et al., 2019) eta akuikultura espezie gisa (FAO, 2022b) duten garrantziagatik itsas organismo aztertuenen artean daude. Egun, zortzi espeziek osatzen dute *Mytilus* generoa: *M. californianus* (Conrad, 1837), *M. chilensis* (Hupé, 1854), *M. edulis* (Linnaeus, 1758), *M. trossulus* (Gould, 1850), *M. planulatus* (Lamarck, 1819), *M. platensis* (d'Orbigny, 1842), *M. unguiculatus* (Valenciennes, 1858) and *M. galloprovincialis* (Lamarck, 1819) (World Register of Marine Species, 2022). Aipatutako espezieetatik, hiruk Europan dute jatorria: *M. edulis*, *M. galloprovincialis* and *M. trossulus.* Hauek hirurak "*Mytilus edulis* konplexua" osotzen dute (McDonald et al., 1991). Espezie bakoitzak eskualde biogeografiko propioa du, baina trantsizio eremuetan (maila geografikoa gainjartzen den tokietan) hibridoak sor daitezke aldameneko espezieen artean (Hilbish et al., 2002; Gosling et al., 2008; Zbawicka et al., 2014).

Mytilus galloprovincialis (**1. kutxa**) muskuiluaren jatorrizko banaketa-eremua Mediterraneo itsasoa, Itsaso Beltza eta Ozeano Atlantikoko ekialdeko kosta (Irlandatik Marokoraino) da (Gosling, 1992). Hala ere, gizakiak eragindako ekintzengatik, bai istripuz (itsas garraioen bidez) zein berariaz (akuikultura ekoizpenerako), eta ingurumen-baldintza desberdinetara bizkor egokitzeko duen gaitasunagatik, mundu osoan zehar hedatu da. Gaur egun, Japonian, Hong Kongen, Ipar Amerikako mendebaldeko kostaldean (Kaliforniako hegoaldetik (AEB) Columbia Britainiarrera (Kanada)), Txilen, Australian, Hego Afrikan, Norvegian, Groenlandian, Txinako iparraldean eta Errusiako ekialdean topatu da (Lee and Morton, 1985; McDonald et al., 1991; Geller et al., 1994; Apte et al., 2000; Anderson et al., 2002; Elliott et al., 2008; Wilkins et al., 2008; Gardner et al., 2021). Hori dela eta, *M. galloprovincialis* munduko espezie inbaditzaile arrakastatsuenen zerrendan ageri da (Global Invasive Species Database, 2022).

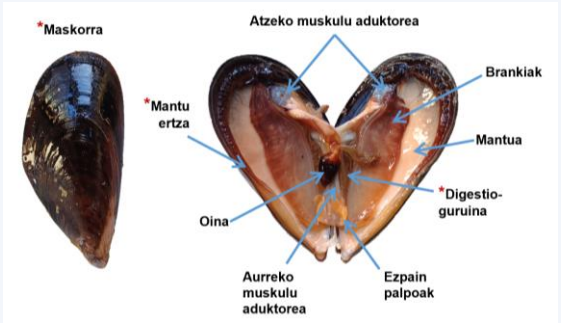


1. Irudia. *Mytilus* generoko muskuiluen taxonomia eta jatorri-eremua. Irudi eraldatua; iturria: Gaitán-Espitia et al. (2016).

1Kutxa

Mytilus galloprovincialis espeziaren datu-txartela

Morfologia eta anatomia *Gorputza kaltzio karbonatozko bi kuskuko maskor batez babestua du. Maskorraren egitura luzanga eta zertxobait triangeluarra da, eta kanpotik kolore urdin-beltza agertzen du:*



Tamaina *5-12 cm*

Tenperatura *10-32 °C (Optimoa 17 °C)*

Gazitasuna *10-38 ppt*

Dieta *Janari-iragazlea da. Bere elikadura fitoplankton, zooplankton eta partikula organiko inerteetan oinarritzen da.*

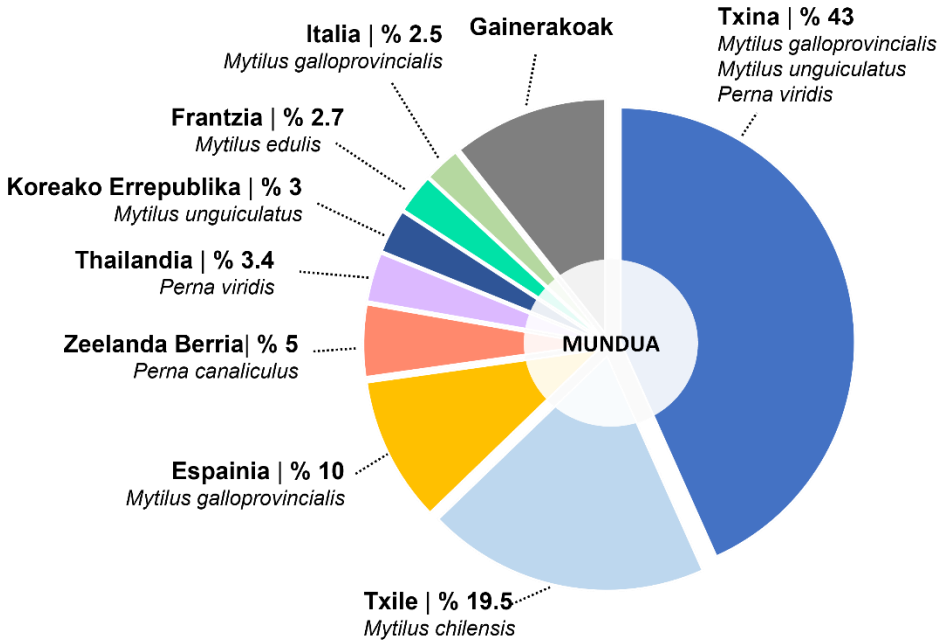
Ugalketa *Kanpo-ernalketaren bitartez ugaltzen dira, hau da, arrek eta emeek euren gametoak uretara askatzen dituzte eta bertan gertatzen da ernalketa, zigotoa eratuz. Prozesu hau urtean behin, birritan edo gehiagotan gerta daiteke. Galizian, esaterako, ugalketa prozesua urte osoan zehar gertatzen da, udaberrian eta udazken hasieran nabarmenagoa izanik. Garapen enbrionarioaren ondoren, lehen larba sortuko da, trokofora izenekoa. Larbak izen ezberdinak hartuko ditu garapenean zehar: beliger, unboduna eta pedibeliger. Ur mugimenduei esker, larben sakabanaketa gertatzen da. Larba-bizitza finkaketa prozesuarekin amaitzen da, larba substratuari itsastean. Momentu horretan metamorfosia gertatzen da, hau da, egitura eta bizimodu aldaketa (bizimodu pelagikotik bentonikora).*

Garrantzia
- *Ingeniari ekologiko gisa*
- *Kutsadura biomarkatzaile gisa*
- *Akuikultura espezie gisa*

Informazio iturria: (Figueras, 2007; OESA, 2017)

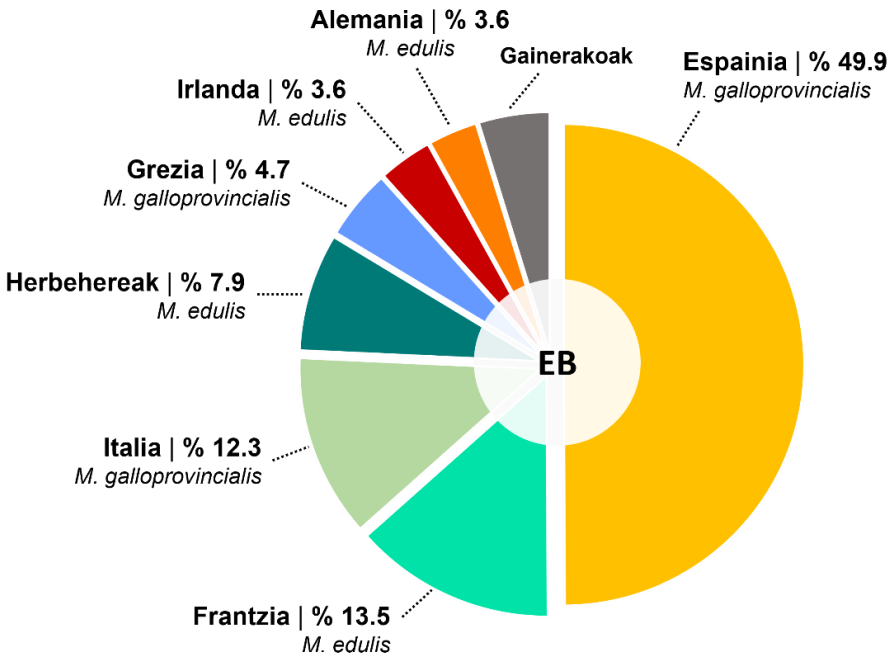
2. Muskuiluen akuikultura ekoizpena eta merkataritza

Akuikulturak azken hamarkadan mundu mailan izan duen garapena izugarria izan da, eta horrek ekarpen handia egin dio itsas jakien hornidurari. 2020an, 88 milioi tona (265 mila milioi dolar) arrain eta itsaski ekoitzi ziren akuikultura bidez, munduko ekoizpen osoaren % 49a (FAO, 2022b). Hazitako animalien artean, muskuiluek garrantzi handia dute. 2020an muskuiluen ekoizpena 2,05 milioi tonara iritsi zen, 3,3 milioi dolarreko balioarekin, duela hamar urte ekoitzitako kopuruaren bikoitza baino gehiago (FAO, 2022a). Urte horretan, munduko muskuilu ekoizle nagusiak Txina (% 43), Txile (% 19,5), Espainia (% 10), Zeelanda Berria (% 5), Koreako Errepublika, Frantzia eta Italia izan ziren (**2. irudia**), Europak ekoizpen osoaren % 21a bere gain hartuz.

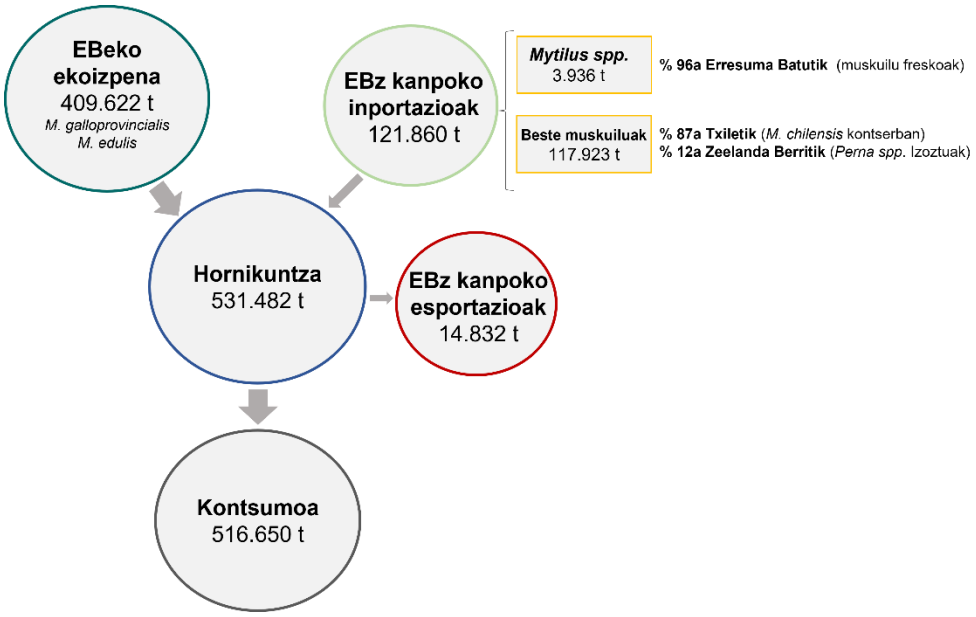


2. Irudia. Munduko muskuilu akuikultura-ekoizpena herrialdeka. Datu iturria: FAO (2022a).

Europar Batasunean (EBn), muskuiluek akuikultura-ekoizpenaren bolumen osoaren % 37,45a hartzen dute, *M. galloprovincialis* eta *M. edulis* izanik gehien ekoizten eta kontsumitzen diren espezieak. 2020an ekoizpena 409.622 tonakoa izan zen, 439 milioi euroko balioarekin, hau da, akuikulturan hazitako molusku-uzta osoaren % 76,2a (FAO, 2022a). Urte honetan, ekoizle nagusiak Espainia (% 49,9), Frantzia (% 13,5), Italia (% 12,3) eta Herbehereak (% 7,9) izan ziren (**3. irudia**). Muskuiluen merkataritzari dagokionez (EUMOFA, 2022a) (**4. irudia**), EBn ekoizten diren *Mytilus* muskuiluen inportazio eta esportazioen % 96a bertako herrialdeen artekoa da. Europar batasunetik kanpoko merkataritzan beste muskuilu espezie batzuk salerosten dira, hala nola, *M. chilensis* edo *Perna spp*. 2020an EBz kanpoko inportazioak 47.000 tonara iritsi ziren; horietatik % 84a (39.400 tona) Txiletik etorritako kontserbako muskuiluak izan ziren, eta % 9a Erresuma Batutik etorritako muskuilu freskoak. Espainiak, Frantziak eta Italiak inportatutako muskuilu hauen % 73a jaso zuten. EBz kanpoko esportazioei dagokionez, 2020an 8.600 tona banatu ziren, % 57a muskuilu freskotan eta % 35a kontserban. Helmuga nagusiak Erresuma Batua (% 48), Suitza (% 16) eta AEB (% 8) izan ziren.



3. Irudia. EBko muskuilu akuikultura-ekoizpena herrialdeka.



4. irudia. EBeko muskuiluen hornikuntza, 2020an, tonatan (t). EUMOFAko konbertsio-faktoreak erabili dira inportazio eta esportazio produktu prosezatuen pisu garbia musluilu freskoen pisura bihurtzeko. Irudi eraldatua; jatorria (EUMOFA, 2022b).

Ikusi dugunez, *Mytilus galloprovincialis* espezieak garrantzi handia du akuikulturan, hauen hazkuntza, kaiola flotagarrietan, "longline" sistemetan edo substratu gaineko hazkuntza sistemetan egiten delarik (Smaal, 2002). Espainia EBeko *M. galloprovincialis* ekoizle nagusia da, urtean 200.000 tona inguru haziz (OESA, 2017; EUMOFA, 2022a). Muskuilu horien % 95a baino gehiago Galizian ekoizten da, Iberiar penintsularen ipar-mendebaldeko kostaldean. Sektore honek tradizio luzea du Galizian eta bertako ekonomian nabarmenki eragiten du hainbat lanpostu sortuz tokiko biztanleen artean. Ekoizpenaren gainerako % 5a Katalunian (Ebro deltako bi arroetan, Fangar eta Alfacs-en), Andaluzian, Valentziako Erkidegoan, Balear Uharteetan eta, berriki, Euskal Herrian egiten da (OESA, 2017). Horretaz gain, Espainia Europa mailan saltzen diren prozesatutako muskuiluen ekoizle nagusia ere bada, 2019an 14.406 tona muskuilu kontserban ekoitzi zituelarik (EUMOFA, 2022b). Espezie hau oso preziatua da bai zapore eta bai nutrizio aldetik. Elikadura-iturri garrantzitsu bat da, kalitate handiko proteinak, mineralak, bitaminak eta omega-3 gantz-azido esentzialak ematen dituelako (Wright et al., 2018).

3. Itsas produktuen jatorria egiaztatzearen garrantzia

Trazabilitatea produktu bati buruzko edozein informazio eskuratzeko gaitasuna da, bere bizi ziklo osoan zehar eta identifikazio erregistratuen bidez. Hau, paperezko erregistro tradizionalen bidez edo azkenengo urteetan garatzen ari diren sistema digitalen bidez (adibidez, bloke-katea edo "*Blockchain*" teknologia) egiten da (Olsen and Borit, 2013; Galvez et al., 2018). Kasu gehienetan, trazabilitate-sistema hauei esker, arrantza jatorriei buruz, haiekin burututako eraldaketa eta prozesuei buruz, edota hornidura-katean zehar egiten dituzten mugimenduei buruzko informazio fidagarria lor daiteke. Hau da, itsasoko produktuen funtsezko ezaugarri askori buruzko informazioa ematen dute (Schröder, 2007; Olsen and Borit, 2013).

Hala ere, egun, merkatuen globalizazioa eta itsas produktuen eskari hazkuntza dela medio, hornidura kateak gero eta konplexuagoak bilakatzen ari dira, trazabilitate ahaleginak zailduz (Leal et al., 2015; Fox et al., 2018). Nazioarteko arrain eta itsaski hornidura kateak oso konplexuak dira; espezie asko merkaturatzen dira eta bitartekari askok parte hartzen dute, besteak beste, arrantzaleak eta akuikultura-ekoizleak, eraldatzaileak, handizkariak, banatzaileak eta txikizkariak (Fox et al., 2018). Honen ondorioz, arrain eta itsaskiek bide luzea egiten dute hartzen direnetik gure etxeko sukaldeetara heltzen direnera arte. Egoera honetan, itsas produktuen etiketan ageri den informazioa zuzena dela egiaztatzeko tresna analitiko zehatzak garatzea ezinbestekoa da, trazabilitate-sistemen bidez bildutako datuak balioztatuko dituztenak.

Itsas produktuen etiketatze okerrak ondorio larriak izan ditzake ekonomian, ingurumenean eta giza osasunean. Orain arte, itsas produktuen benetakotasuna bermatzeko ahalegin gehienak etiketan ageri den espeziaren izena egiaztatzera bideratu dira, jatorria egiaztapena albo batera utzita. Jatorriaren nahitaezko etiketatze okerra, jatorri iruzurra, irabazi ekonomiko bat lortzeko egiten da, askotan arraina edo itsaskia kalitate eta balio txikiagoko produktu batekin ordezkatuz, edota jasangarritasun eta etika erregulazio gutxiago dituen toki batetik hartuz. Nolanahi ere, jatorriaren etiketatze okerrak, engainu ekonomikoa ez ezik, segurtasuna, jasangarritasuna eta kontsumitzailearen konfiantza kolokan jartzen ditu (Lindley, 2020). Ingurumenari

dagokionez, jatorri iruzurrak arriskuan dauden edo gehiegi ustiatuta dauden espezieen merkataritzari ateak ireki diezazkioke, arrantzaren iraunkortasuna eta itsas ekosistemen jasangarritasuna kaltetuz (Jacquet and Pauly, 2008; Agnew et al., 2009). Kontsumitzaileei dagokionez, etiketatze okerrak uko egin diezaioke jaten dituzten itsaskien jatorriari buruzko baliozko aukera bat egiteko eskubideari (Kelly et al., 2005; Fonner and Sylvia, 2015). Gerta daiteke kontsumitzaileek tokiko produktuak erosi nahi izatea bertako akuikulturari laguntzeko edo garraioari lotutako CO_2 aztarnak gutxitzeko. Era berean, jatorria gaizki etiketatzeak ondorioak ekar ditzake elikagaien segurtasun arloan. Patogenoekin, biotoxinekin edo kutsatzaile kimikoekin kutsatuta egon daitezkeen eremu mugatuetan sortutako itsaskiak kontsumitzea, arriskutsua izan daiteke gizakien osasunerako (Feldhusen, 2000; Guo et al., 2007). Jatorria egiaztatzea bereziki garrantzitsua da bibalbioen kasuan, hainbat infekzio-agerraldirekin lotu direlako. Janari-filtratzaileak izanik, patogenoak (adib., *E. coli*, *Salmonella* spp. edo *Vibrio* spp.) metatzeko eta kontzentratzeko gai dira, eta gainera, maiz gordinik edo azpi-egosita kontsumitzen dira, arriskua areagotuz (Potasman et al., 2002; Rodriguez-Manzano et al., 2014; Hassard et al., 2017; Romalde et al., 2018).

Faktore guzti hauek trazabilitatea erregulatzeko prozedura eta araudien garapena bultzatu dute. Itsas produktuen jatorriari buruzko informazioa eskaintzearen beharra nazioarteko legeetan islatzen da. Europar batasunean, Kontseiluaren 1379/2013 (EE) Erregelamenduak, kontseiluaren 1184/2006 (EE) eta 1224/2009 (EE) Erregelamenduak aldatzen dituena, eta Kontseiluaren 104/2000 (EE) Erregelamendua indargabetzen duena, arrantza eta akuikulturako produktuen etiketek adierazi beharrekoa arautzen du. Lege honen arabera, espeziaren merkataritza-izena eta izen zientifikoa, ekoizpen metodoa eta produktua harrapatu edo hazi den eremua adieraztea nahitaezkoa da (EU, 2013b).

Nahitaezko lege hauetaz gain, jatorri geografikoa aitortzen duten borondatezko etiketak ere badaude. Alde batetik, jatorri-deitura babestuak eta adierazpen geografiko babestuak daude. Bi adierazpen geografiko hauek produktu bat jatorriz eremu geografiko jakin batekoa dela ziurtatzen dute, haren kalitate edo ezaugarriek jatorriarekin harreman estua dutelarik. Honen ondorioz, produktu hauek balio gehigarria

hartzen dute. Muskuiluen kasuan, *M. galloprovincialis* espezieari dagokionez, Espainian, Galiziako muskuiluak ("Mexillón de Galicia"), eta Italian, Scardovari-ko muskuiluak ("Cozza di Scardovary") Europar Batasunaren jatorri-deitura babestuaren zigilua dute (EU, 2007; EU, 2013a). Bestetik, MSC eta ASC zigiluak daude. Borondatezko zigilu hauek, produktuen jasangarritasuna bermatzeaz gain, zaintza-katearen ziurtagiria ere badute. Ziurtagiri honek hornidura katearen maila bakoitza ziurtatzen du, jatorritik azken produktura arte egin duen bidea bermatuz.

Beraz, nahitaezko zein borondatezko etiketetan agertzen den jatorriari buruzko informazioa zuzena dela bermatzeko eta lehen aipatutako gaizki etiketatzearen ondorioak ekiditeko, premiazkoa da metodo zientifiko zehatz eta sentikorrak garatzea, kontrol-agintariek eta erakunde arautzaileek jatorri geografikoa egiaztatzeko aukera izan dezaten.

3.1. Jatorria egiaztatzeko tresnak

Arrain eta itsaskien jatorria arakatzeko edo egiaztatzeko metodologia ugari proposatu eta aplikatu dira, bai era independentean zein konbinatuan (Leal et al., 2015; Gopi et al., 2019a; Cusa et al., 2021). Metodologia hauek aztarna-elementuen analisian (Ricardo et al., 2015; Bennion et al., 2019; Gopi et al., 2019b), isotopo egonkorren analisian (Kim et al., 2015; Bernatchez et al., 2017; Gopi et al., 2019b; Zhao et al., 2019; Doubleday et al., 2022), gantz-azidoen analisian (Ricardo et al., 2017c; Fonseca et al., 2022), espektroskopian (Liu et al., 2015; Ghidini et al., 2019) eta DNAren analisian (Ogden, 2011; Nielsen et al., 2012; Larraín et al., 2014; Ogden and Linacre, 2015; Rodríguez-Ezpeleta et al., 2019) oinarritzen dira besteak beste. Konkretuki, lan honetan DNAren, aztarna elementuen eta bakterio-komunitateen analisian oinarritutako tresna analitikoak garatuko dira jatorria egiaztatzeko.

3.1.1. Markatzaile genetikoak

Markatzaile genetikoak askotan erabili izan dira arrain eta itsaski espezieak identifikatzeko, aldiz, ez hainbeste haien jatorri geografikoa identifikatzeko. Honen arrazoia da, jatorria egiaztatzeko, oinarrizko datu genetiko egokiak behar direla, analisi

estatistikoak konplexuagoak direla, eta emaitzak interpretatzeko espeziearen eboluzio eta biologiaren ikuspegi sakonagoa behar dela (Nielsen, 2016). Jatorri geografikoaren identifikazio genetikoa indibiduo bat bere jatorrizko populazio genetiko edo ebolutiboari esleitzeko gaitasunean oinarritzen da, eta horretarako, bere jatorrizko populazioa gainerako populazioetatik aski desberdina izan behar du (Ogden and Linacre, 2015). Beraz, jatorria identifikatzeko, lehenengo urratsa espezie horren populazioaren egitura identifikatzea da. Espezie jakin batek ez baldin badu populazio-bereizketarik azaltzen bere eremu geografikoan zehar, orduan teknika genetikoak ez dira eraginkorrak izango. Aldiz, markatzaile genetikoren bat intereseko populazioak banatzeko gai baldin bada, orduan jatorriaren esleipen probabilitatea estimatu daiteke. Markatzaile genetikoak lortzeko erabiltzen diren laborategi eta konputazio teknikak nabarmen garatu dira azken hamarkadan. Hori dela eta, ugaritu egin dira eskura ditugun markatzaile genetiko motak: markatzaile klasikoetatik hasita (adib. murriztapen-zatien luzerako polimorfismoak, anplifikatutako-zatien luzerako polimorfismoak, mikrosateliteak eta polimorfismo mitokondrialak), markatzaile berriagoetara (adib. SNPak) (Cuéllar-Pinzón et al., 2016). Ildo berean, errendimendu handiko sekuentziazio teknologien garapenak (adib. RAD-seq), genoma osoan milaka SNP markatzaile identifikatzea ahalbidetu dute, bereizmen handiko datu genomikoak emanez. Arrain eta itsaskien trazabilitateari dagokionez, zenbat organismoekin egindako ikerketek SNP markatzaileak jatorri geografikoa esleitzeko baliagarriak direla frogatu dute, hala nola, Atlantikoko izokinaren, Atlantikoko hegaluzearen, Bizkaiko golkoko antxoaren, Atlantikoko bakailaoaren, Atlantikoko sardinzarraren, Europako mihi-arrainaren eta legatzaren jatorria esleitzeko (Nielsen et al., 2012; Gilbey et al., 2016; Montes et al., 2017).

3.1.2. Aztarna-elementuak

Moluskuen maskorrak, ornogabeen estatolitoak edo arrainen otolitoak bezalako egitura biogeniko gogorretan pilatzen diren aztarna-elementuak markatzaile natural gisa erabil daitezke indibiduoen populazioak bereizteko (Leal et al., 2015). Egitura hauek indibiduoekin batera hazten dira, eta inguruko urek eta ingurumen baldintzek eragin zuzena dute haien konposizioan (Becker et al., 2005; Strasser et al., 2008). Hau da, egitura hauetatik lortutako aztarna-elementuek indibiduo jakin baten ingurune

geografikoa islatzen dute jaiotzetik harrapatzen direnera arte, eta hortaz, jatorri geografikoa bereizteko erabil daitezke. Aztarna-elementuen analisia, induktiboki akoplatutako plasma iturridun masa-espektrometroa (ICPMS) eta analisi estatistikoetan oinarritutako sailkapen metodoak erabiliz, arrain eta itsaskien jatorri geografikoa esleitzeko metodo eraginkorra dela frogatu da zeinbait kasutan (Broadaway and Hannigan, 2012; Ricardo et al., 2015; Artetxe-Arrate et al., 2019).

3.1.3. Bakterio-komunitateak

Bakterio-komunitateen analisiek organismo jakin baten dauden bakterioen profil orokorra erakusten dute. Bakterio profil hau eremu geografikoaren berariazkoa dela oinarritzat hartuz, markatzaile biologiko gisa erabil daiteke jatorria arakatzeko. Arrain eta itsaskiei loturiko bakterio-komunitateen analisia, bai eta lortutako profila jatorri geografikoari esleitzea ere, biologia molekularreko teknika tradizionalak erabiliz aztertu da, hala nola, 16S rRNA genean oinarritutako PCR-DGGE teknikaren bidez (Le Nguyen et al., 2008; Montet et al., 2008; Léopold et al., 2010). Azken aldian, NGS metodoen garapenarekin, bereizmen handiagoko analisi azkarragoak eta errentagarriagoak erabiltzen ari dira bakterioen konposizioa aztertzeko. Jatorri geografikoa egiaztatzeari dagokionez, ikerketa gutxi batzuek baino ez dute bakterioen komunitateen analisiek duten potentziala aztertu; hauen artean daude, akuikultura bidez ekoitzitako lupiarekin, (*Dicentrarchus labras*) (Pimentel et al., 2017), *Ruditapes philippinarum* txirlekin (Milan et al., 2019) eta *Mya arenaria* txirlekin (Liu et al., 2020) egindako ikerketak.

HIPOTESIA ETA HELBURUAK

Hipotesia

"Markatzaile genetikoak, aztarna-elementuen konposizioa eta bakterio-komunitateen profila hazitako M. galloprovincialis muskuiluen jatorri geografikoa egiaztatzeko erabil daitezke.

Helburuak

Lan honen helburu nagusia *Mytilus galloprovincialis* muskuiluen jatorri geografikoa egiaztatzeko tresna analitikoak garatzea da, itsas produktuen integritatea bermatzeko. Helburu nagusi hau lortzeko, honako helburu zehatz hauek landuko dira:

1. Europan merkaturatzen diren *Mytilus* espezie garrantzitsuenak (*M. galloprovincialis, M. edulis eta M. chilensis*) identifikatzeko DNAn oinarritutako tresna azkar, fidagarri eta errentagarria garatzea, bai muskuilu fresko zein prozesatuetan erabiltzeko. (1. kapitulua)
2. RAD-seq teknikaren bidez lortutako SNP markatzaileek *M. galloprovincialis* muskuiluen jatorria egiaztatzeko duten ahalmena aztertzea, lehenbizi haien populazio egitura zehaztuz eta ondoren, aukeratutako SNPen jatorria esleitzeko ahalmena aztertuz (2. kapitula).
3. *M. galloprovincialis* muskuiluen jatorri geografikoa egiaztatzeko, aztarna-elementuen eta isotopo egonkorren konposizioa, era independentean zein konbinatuan, erabiltzearen potentziala ebaluatzea (3. kapitula).
4. *M. galloprovincialis* muskuiluen digestio-guruinean dauden bakterio-komunitateen konposizioa eta egonkortasuna aztertzea, bai eta jatorri geografikoa egiaztatzeko tresna gisa duen potentziala aztertzea ere (4. kapitulua).

1. KAPITULUA

Mytilus espezieen identifikazioa *SYBR-Green* bidezko denbora errealeko PCR probaren disoziazio-kurben analisiarekin

Kapitulu hau honela argitaratu zen:

del Rio-Lavín, A., Jiménez, E., and Pardo, M. A. (2021). SYBR-Green real-time PCR assay with melting curve analysis for the rapid identification of *Mytilus* species in food samples. *Food Control*, 130: 108257.
https://doi.org/10.1016/j.foodcont.2021.108257

Laburpena

Mytilus generoko muskuiluek garrantzi handia dute Europar Batasuneko akuikultura ekoizpenean, *M. galloprovincialis*, *M. edulis* eta *M. chilensis* izanik bertan merkaturatzen diren espezie nagusiak. Hortaz, bibalbio hauen espeziea egiaztatzea funtsezkoa da etiketetan ageri den informazioa zuzena dela bermatzeko, eta trazabilitatea eta gardentasuna indartzeko. Lan honetan, 174 muskuilu lagin erabili dira EBn merkaturatzen diren *Mytilus* espezie garrantzitsuenak identifikatzeko gai den denbora errealeko PCR metodo azkar, fidagarri eta errentagarri bat garatzeko. Muskuiluen bisuko proteina itsasgarriak kodetzen dituen genea anplifikatuz, eta PCR produktu horien disoziazio-kurbak aztertuz, muskuilu fresko, izoztu edo kontserben espeziea % 100eko espezifikotasunaz, sentikortasunaz eta zehaztasunaz identifikatu da. Gainera, metodo hau PCR termoziklagailu ezberdinetan eta *software* berezien beharrizanik gabe erabil daitekeela erakutsi dute emaitzek. Hortaz, kapitulu honetan deskribaturiko metodologia, *Mytilus* generoko muskuiluen espeziea egiaztatzeko erabilgarria da eta gainera, erraz inplementa daiteke elikagaien segurtasuna eta identitatea bermatzen duten kontrol laborategietan, ohiko analisi azkar gisa.

1. Sarrera

Muskuiluak munduan zehar oso hedatuta daude eta akuikultura espezie gisa garrantzi handia dute. Egun kontsumitzen diren muskuiluen % 94a akuikulturak hornitzen du. Munduko muskuilu ekoizpenak goranzko joera erakutsi du 2007tik 2016ra, Europar Batasuna izanik ekoizle nagusienetariko bat (EUMOFA, 2019). EBn bi muskuilu espezie hazten dira gehien bat: *M. galloprovincialis* (Espainia eta Italia ekoizle nagusiak izanik) eta *M. edulis* (Herbehereak eta Frantzia ekoizle nagusiak izanik) (EUMOFA, 2019). Bi espezie horien eremu geografikoa gainjartzen denean, hibridoak sortzen direla ikusi da (Hilbish et al., 2002; Beaumont et al., 2004; Michalek et al., 2016). Itsas produktu hauen eskari hazkuntza dela medio, Txiletik genero bereko *M. chilensis* muskuiluak inportatzen dira. Espainia EBeko *M. galloprovincialis* ekoizle nagusia da, urtean 200.000 tona inguru haziz, eta aldi berean, kontsumitzaile nagusienetariko bat ere bada. Hortaz, aurretiaz aipatutako hiru *Mytilus* espezieak topa daitezke Espainiako merkatuetan, bai freskoan bai kontserban.

Arrain eta itsaskien ezaugarri morfologikoak eraldatzen edo kentzen direnean, hauen autentifikazioa zaildu egiten da. Askotan, muskuiluak maskorrik gabe merkaturatzen dira, eta horrek begi-bistako espeziearen identifikazioa oztopatzen du. Europar Batasunean, arrantza eta akuikulturako produktuen merkaturatzeari dagokion Kontseiluaren 1379/2013 (EE) Erregelamendua oso zorrotza da. Lege honen arabera, espeziaren merkataritza-izena eta izen zientifikoa, ekoizpen metodoa eta produktua harrapatu edo hazi den eremua adieraztea nahitaezkoa da (EU, 2013b). Beraz, arrain eta itsaskien espeziearen identifikazioa ezinbestekoa da aipatutako legedia betetzen dela ziurtatzeko, etiketa iruzurra saihesteko eta trazabilitatea bermatzeko.

Mytilus generoko espezieak identifikatzeko, PCR bidezko DNA markatzaile espezifiko baten anplifikazioan oinarritutako metodoak erabiltzen dira gehien bat (Santaclara et al., 2006; Fernández-Tajes et al., 2011; Brooks and Farmen, 2013; Meistertzheim et al., 2017; Zhang et al., 2019). Orain arte erabilitako markatzaile molekularren artean, muskuiluen bisuko proteina itsasgarriak kodetzen dituen genea izan da erabiliena. Gene hau PCR

bidez anplifikatu eta ondoren, AFLP (Inoue et al., 1995; Brooks and Farmen, 2013) edo RFLP (Santaclara et al., 2006; Fernández-Tajes et al., 2011) teknikak erabiliz aztertu izan da. Analisia azkartzeko, denbora errealeko anplifikazio metodoak ere garatu dira, bai *TaqMan*^{TM} zunden erabileran (Dias et al., 2008) zein HRM teknikaren erabileran (Jilberto et al., 2017) oinarritu direnak.

Azkenengo bi metodo hauek analisi denbora nabarmen murrizten duten arren, desabantaila batzuk ere badituzte. Lehenengo teknikaren kasuan, zundak erabili behar direnez analisia garestitzen da, eta bigarren teknikaren kasuan, tresna eta *software* espezifikoak behar dira analisia burutu ahal izateko. Muga hauek gainditzeko, lan honetan, *SYBR-Green* bidezko denbora errealeko PCR proba, anplifikatutako produktuen disoziazio-kurben (edo urtze-kurben) analisiarekin, erabiltzea proposatzen da. Metodo honek, zunda fluoreszenteak erabili ordez, kate bikoitzeko DNAri lotzen zaion *SYBR-Green* molekula tindatzailea erabiltzen du. Gainera, denbora errealeko termoziklagailu arruntetan eta *software* berezien beharrizanik gabe erabil daiteke, kostuak murrizten direlarik. Teknika hau hainbat espezie identifikatzeko egokia dela frogatu da (Hsu et al., 2003; Barcenas et al., 2005; Yu et al., 2005; Berry and Sarre, 2007; Winder et al., 2011; Castigliego et al., 2015).

Lan honetan, *SYBR-Green* bidezko denbora errealeko PCR metodo azkar, fidagarri eta errentagarri bat garatu da, anplifikatutako produktuen disoziazio-kurbak aztertuz, Europar Batasunean merkaturatzen diren *Mytilus* espezie garrantzitsuenak identifikatzea ahalbidetzen duena, bai muskuilu freskoa zein prozesatua izan.

2. Materialak eta metodoak

2.1. Laginketa eta DNAren erauzketa

Mytilus generoko 174 erreferentzia lagin bildu ziren: 89 *M. galloprovincialis*, 26 *M. edulis*, 24 *M. chilensis*, 7 *M. trossulus* eta 28 *M. galloprovincialis* x *M. edulis* hibrido. Erreferentziazko lagin hauen espeziea PCR-RFLP Me15-16 *AciI* metodoaren bitartez identifikatu zen, aurrerago zehazten den bezala (Santaclara et al., 2006). Kontrol

negatibo gisa, *Mytilus* generokoak ez ziren honako bibalbio hauek hautatu ziren: *Callista chione*, *Cerastorema edule*, *Crassostrea gigas*, *Ensis macha*, *Ostrea edulis*, *Perna viridis*, *Ruditapes decussatus*, *Tagelus dombeii* and *Venerupis philippinarum* (lau indibiduo espezie bakoitzeko). Horrez gain, 16 muskuilu produktu prozesatu (ozpinetan, oliba oliotan, espeziekin edo saltsan egosita) erosi ziren Euskal Autonomia Erkidegoko merkatuetan. Produktu bakoitzeko bi indibiduo aztertu ziren.

DNAren erauzketarako *Wizard® Genmic DNA Purification Kit*-a (Promega, Madison, WI, USA) erabili zen, 50 mg mantu ertzetik abiatuta eta fabrikatzailearen jarraibideei jarraituz. Prozesatutako laginen gehiegizko olioa paper batez kendu zen erauzketa hasi baino lehen. Erauzitako DNA *NanoDrop 1000* (Thermo Fisher Scientific, Waltham, MA, USA) espektrofotometroa erabiliz kuantifikatu zen. DNA lagin guztiak -20 °C-tan gorde ziren ondorengo analisietarako.

2.2. *Mytilus* espezieen identifikazioa PCR-RFLP Me15-16 *AciI* metodoa erabiliz

Erreferentziazko 174 muskuilu laginen espeziearen identifikazioa PCR-RFLP Me15-16 *AciI* metodoa erabiliz egin zen Santaclara et al. (2006)-ek deskribatu bezala. Laburki, PCRa egiteko 50 µl-ko nahasketa bat prestatu zen: 20 ng DNA, 0,2 mM dNTP, 2mM $MgCl_2$, 1 µM hasle bakoitzeko eta 1 U *Taq* DNA polimerasa (Illustra™, GE Healthcare, Chicago, IL, USA) erabiliz. PCRaren baldintzak honako hauek izan ziren: 5 minutuko hasierako desnaturalizazioa 95 °C-tan; 35 ziklo 30 s 95 °C-tan, 30 s 56 °C-tan eta 30 s 72 °C-tan; eta bukatzeko, 5 minutuko luzapena 72 °C-tan. Anplifikatutako produktuak % 1,5eko agarosa-gelean (p/v) aztertu ziren. *M. galloprovincialis* eta *M. chilensis* espezieak bereizteko, anplifikatutako produktuak *GFX PCR DNA eta Gel Band Purification Kit*-aren (Illustra™, GE Healthcare, Chicago, IL, USA) bitartez purifikatu ziren dimeroak eta hasleak kentzeko. Purifikatutako DNA produktuen digestioa, *AciI* entzimaren (Thermo Fisher Scientific, Waltham, MA, USA) 10 unitate eta DNAren 5 µl erabiliz egin zen. Nahasketa hau (20 µl guztira) gauez 37 °C-tan inkubatu zen. Digeritutako laginak *DNA 1000 chip* (Agilent Technologies, Santa Clara, CA, USA) batean kargatu ziren eta *Bioanalyzer* (Agilent Technologies, Santa Clara, CA, USA) plataforma erabiliz aztertu ziren.

2.3. *Mytilus* espezieen identifikazioa *SYBR-Green* bidezko denbora errealeko PCR produktuen disoziazio-kurben analisiarekin

Denbora errealeko PCR saiakuntza hiru termoziklagailu ezberdin erabiliz burutu zen, haien artean egon zitezkeen aldeak ebaluatzeko: *LightCycler® 480 Instrument II* (Roche Molecular Systems, Pleasanton, CA, USA), *Step One™* (Applied Biosystems, Waltham, MA, USA) eta *CFX96 Touch DeepWell™* (BioRad, Hercules, CA, USA). Horretaz gain, bi PCR *Master Mix* (PCR nahasketa komertziala) alderatu ziren: *Brilliant III Ultra-Fast SYBR-Green QPCR Master Mix* (Agilent Technologies, Santa Clara, CA, USA) eta *LightCycler® 480 SYBR-Green I Master* (Roche Molecular Systems, Pleasanton, CA, USA).

PCR erreakziorako nahasteak (10 µL) honako hau zuen: 5 µL *Master Mix* A (*Brilliant III Ultra-Fast SYBR-Green QPCR Master Mix*) edo *Master Mix* B (*LightCycler® 480 SYBR-Green I Master*), 20 ng DNA eta Dias et al. (2008) ikerketan deskribatutako hasleen 450 nM. Erreakzio baldintzak honako hauek izan ziren: 3 minutuko hasierako desnaturalizazioa 95 °C-tan eta 45 ziklo 5 s 95 °C-tan eta 10 s 60 °C-tan. PCR produktuen disoziazio-kurbak termoziklagailu berdinarekin neurtu ziren, 65 °C-tik 95 °C-ra tenperatura gradiente bat eginez eta 0.1 °C bakoitzeko fluoreszentzia neurtuz. PCR analisian DNArik gabeko kontrol negatibo bat eta espezie bakoitzaren erreferentziazko kontrol positibo bat ere neurtu ziren. Lortutako disoziazio-kurben grafikoak fluoreszentziaren balio logaritmikoek tenperaturarekiko zuten aldaketa erakutsi zuten.

2.4. Datuen analisia

Denbora errealeko PCR analisiaren bidez lortutako disoziazio-kurben grafikoek seinale fluoreszenteen balio logaritmikoak tenperaturarekiko zuten aldaketa erakutsi zuten. Disoziazio-kurbaren piko bakoitzak *Mytilus* muskuilu bakoitzaren produktu anplifikatuaren urtze-tenperatura (T_m) adierazten zuen. Espezie bakoitzaren T_m heina honela kalkulatu zen: T_m-aren batez besteko balioa (y) ± di aldiz desbideratze estandarra (T_m = y ± 2SD). Metodo honen efizientzia eta errendimendua honako metrika hauek erabiliz ebaluatu zen: sentikortasuna, espezifikotasuna, sendotasuna eta zehaztasuna. Sentikortasuna, honela kalkulatu zen: era egokian identifikatu ziren espezie jakin bateko indibiduo kopurua, zati espezie horren indibiduo kopuru osoa. Espezifikotasuna honela

kalkulatu zen: era egokian baztertu ziren espezie jakin bateko indibiduo kopurua, zati espezie horretakoak ez ziren indibiduo kopuru osoa. Sendotasuna, bi *Master Mix* ezberdin eta hiru termoziklagailu ezberdin erabiliz ebaluatu zen. Zehaztasuna, errepikagarritasunarekin eta erreproduzigarritasunarekin lotua, aldatu gabeko baldintzetan egindako neurketa errepikatuek emaitza berberak erakusten zituzten kasuak aztertuz neurtu zen. Errepikagarritasuna honela kalkulatu zen: laborategiko analista berak, ekipo bera erabiliz eta lagin positibo bera 10 aldiz neurtu ondoren, lortutako benetako emaitza positiboen kopurua zenbatuz. Erreproduzigarritasuna honela kalkulatu zen: laborategiko bi analista ezberdinek, lagin positibo bera 10 aldiz neurtu ondoren, lortutako benetako emaitza positiboen kopurua zenbatuz (ISO/20813, 2019).

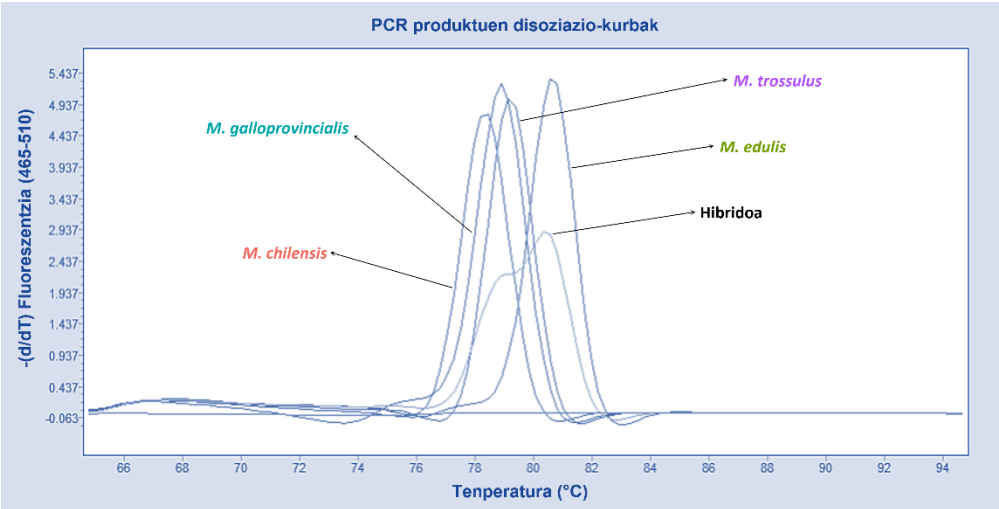
3. Emaitzak eta eztabaida

3.1. *SYBR-Green* bidezko denbora errealeko PCR metodoaren garapena, anplifikatutako produktuen disoziazio-kurben analisiarekin

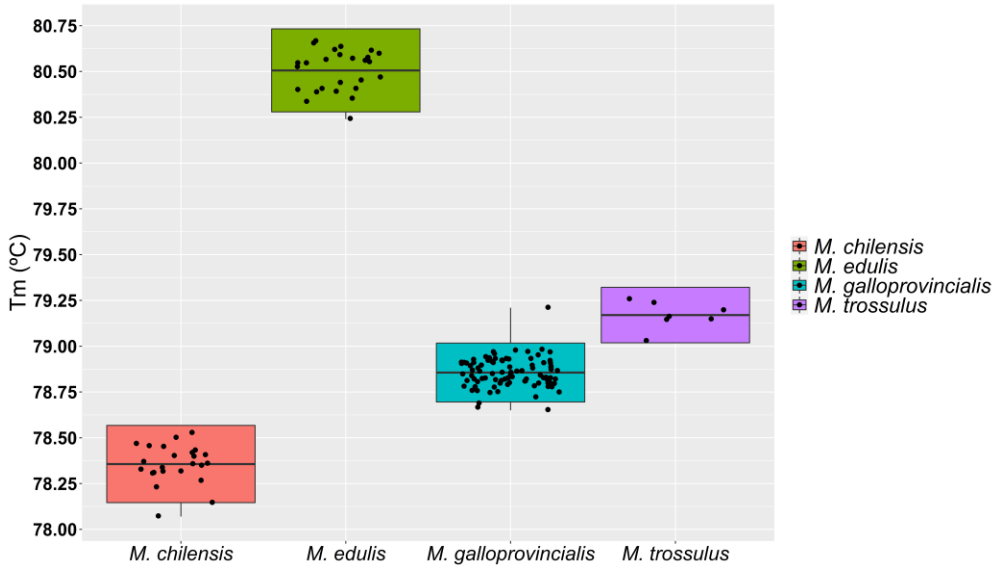
Denbora errealeko PCR produktuen disoziazio-kurben analisia, Dias et al. (2008) ikerketan deskribatutako hasle unibertsalak erabiliz eta muskuiluen proteina itsaskorrak kodetzen dituen genea anplifikatuz (Inoue et al., 1995), arrakastaz erabili da Europan merkaturatzen diren *Mytilus* espezie garrantzitsuenak identifikatzeko. Proteina itsaskorren gene nuklear hau askotan erabili izan da *Mytilus* generoko espezieak identifikatzeko; hala nola, PCR arrunta erabiliz, anplifikatuak AFLP (Inoue et al., 1995) edo RFLP (Santaclara et al., 2006) bidez aztertuz, edota denbora errealeko PCR teknikak erabiliz, *TaqMan*^{TM} zunden bidez (Dias et al., 2008) edo HRM bidez (Jilberto et al., 2017). Hortaz, markatzaile hau *Mytilus* generoko espezieak era azkar eta fidagarri batean identifikatzeko erabiliena da. Markatzaile bakarra erabiltzeak (bat baino gehiago erabili ordez) bere mugak ere baditu. Izan ere, markatzaile bakar hau ez da gai introgresioa detektatzeko, ezta hego hemisferioko *M. chilensis* muskuiluak Australiako eta Zeelanda Berriko *M. galloprovincialis* muskuiluetatik bereizteko ere. Muga horiek gorabehera, eta kontuan hartuta gure helburua Europan merkaturatzen diren muskuilu produktuak azkar

identifikatzea dela, markatzaile bakar baten erabilera nahiko egokia da (Larrain et al., 2019; Araneda et al., 2021).

Ikerketa honetan lortutako PCR produktuen disoziazio-kurbek, aztertutako espezie bakoitzaren T_m heinan piko espezifiko bat erakutsi zuten. Aldiz, *M. galloprovincialis* x *M. edulis* hibridoen kasuan, espezie bakoitzaren T_m-arekin lotura zuten bi piko ikusi ziren (**1.1 irudia**). Aztertutako espezie bakoitzarentzat T_m balio konkretuak lortu ziren, 89 lagin *M. galloprovincialis* (T_m=78.85 °C ± 0.080), 26 lagin *M. edulis* (T_m=80.49 °C ± 0.12), 24 lagin *M. chilensis* (T_m=78.35 °C ± 0.11) eta 7 lagin *M. trossulus* (T_m=79.17 °C ± 0.076) aztertu ostean. Metodoak % 100eko sentikortasuna agertu zuen; izan ere, PCR-RFLP Me15/16 *Acil* metodo konbentzionala erabiliz lortutako emaitzekin alderatuz (Santaclara et al., 2006), lagin guztiak (n = 174) ondo identifikatuta zeuden (**1.2 irudia**). Espezifikotasun arrakastatsua ere lortu zen, *Mytilus* generokoak ez ziren 36 indibiduoak era egokian anplifikatu ez zirelako (datuak ez dira ageri).



1.1 irudia. Espezie bakoitzarentzat lortutako disoziazio-kurbak. Fluoreszentziaren balio logaritmikoak tenperaturarekiko duten aldaketa erakusten dute. Piko bakoitzak *Mytilus* muskuilu bakoitzaren produktu anplifikatuaren T_m-a adierazten du. Grafika hau A *Master Mix*-a eta *LightCycler® 480 Instrument II* termoziklagailua erabiliz lortu zen. Hibridoa *M. galloprovincialis* x *M. edulis*-ri dagokio.



1.2 irudia. Espezie bakoitzarentzat lortutako T_m balioak (n = 146). *M. galloprovincialis* (n=89): 78.85 °C ± 0.080 | *M. edulis* (n=26): 80.49 °C ± 0.12 | *M. chilensis* (n=24): 78.35 °C ± 0.13 | *M. trossulus* (n=7): 79.24 °C ± 0.086. T_m heina bakoitza honela kalkulatu zen: T_m-aren batez besteko balioa (y) ± di aldiz desbideratze estandarra (T_m = y ± 2SD). Grafika hau A *Master Mix*-a eta *LightCycler® 480 Instrument II* termoziklagailua erabiliz lortu zen.

Denbora errealeko PCRaren bitartez lortutako produktuen disoziazio-kurbaren analisia, mikroorganismoak (Winder et al., 2011), ugaztunak (Berry and Sarre, 2007), arrainak (Castigliego et al., 2015), intsektuak (Barcenas et al., 2005; Yu et al., 2005) edota onddoak (Hsu et al., 2003) identifikatzeko tresna erabilgarria dela frogatu dute zenbait ikerketek. Lan honetan, *Mytilus* espezieak identifikatzeko tresna azkar eta errentagarri bat garatu dugu, denbora errealeko PCR produktuen urtze tenperaturaren araberakoa. Bi pausutan egiten diren PCR konbentzionaletan oinarritutako teknikak bibalbio hauek identifikatzeko erabili izan dira; PCR produktuak AFLP teknikaren (Inoue et al., 1995) edo RFLP teknikaren (Santaclara et al., 2006) bidez aztertuz. Ondoren, pausu bakarreko denbora errealeko PCR teknikak garatu dira, eta horrek, analisi denbora aurreztea ahalbidetu dute. Denbora errealeko analisi hauetan *TaqMan™* zundak (Dias et al., 2008) erabili ordez, lan honetan proposatutako *SYBR-Green* koloratzailea erabiltzen baldin bada, analisia askoz merkeagoa da.

Anplifikatutako DNA produktuen disoziazio-kurben analisiak, HRM analisiaren oinarrizko printzipioak partekatzen ditu. HRM analisia aplikazio nahiko berria da, eta arrakastaz erabili da *Mytilus* espezieen identifikaziorako (Jilberto et al., 2017). HRM analisirako erabiltzen diren tindatzaileak (adib. *EvaGreen*® eta *SensiFast*^{TM}), *SYBR-Green* tindatzailea ez bezala, kate bikoitzeko DNA guztiz bete arte lotzen dira, eta hortaz, nukleotido bakarreko polimorfismoak bezalako aldaketa txikiak identifikatzea ahalbidetzen du. Hala ere, HRM analisiak desabantaila bat du; metodo honek *software* berezia behar du dentsitate handiko fluoreszentzia datuak batu ahal izateko, aldiz, lan honetan proposatutako teknika denbora errealeko edozein tekmoziklagailu erabiliz aplikatu daiteke (Herrmann et al., 2007). Ikerketa honetan lortutako emaitzek argi utzi dute Europan merkaturatzen diren *Mytilus* espezie nagusiak modu eraginkorrean bereizi daitezkeela, HRM metodoa erabili gabe. Lehenik eta behin, *M. galloprovincialis* eta *M. edulis* espezieak zehaztasun handiz desberdindu genituen, izan ere, anplifikatutako DNA zatien GC edukia, luzera eta sekuentzia bera, disoziazio-kurba espezifikoak sortu zituen espezie bakoitzerako (**1.1 irudia**). Bigarrenik, anplifikatutako *M. galloprovincialis* eta *M. chilensis* espezieen DNA zatiak base bakarrean (G/T) (Jilberto et al., 2017) desberdindu arren, bereizi ahal izan genituen, izan ere, PCR produktua hain laburra zenez (126 bp) purina bat pirimidina batekin ordezkatzeak eragin handia izan zuen urtze tenperaturan. Azkenik, *M. galloprovincialis* eta *M. trossulus* espezieen artean ezberdintasunak ikusi genituen arren, haien T_m balioak lar antzekoak ziren elikagai prozesatuen espezie bereizketa zuzena bermatzeko. Espezie batzuen arteko T_m balioen aldea txikia izan daitezkeenez, identifikazio zuzena bermatzeko, analisi bakoitzean kontrol positibo moduan espezie bakoitzaren erreferentziazko DNA sartzea gomendatzen da.

Metodo honen sendotasuna bi *Master Mix* ezberdin erabiliz ebaluatu zen. *LightCycler® 480 Instrument II* termoziklatzailea erabiliz, A *Master Mix*-a B *Master Mix*-a baino egokiagoa zela frogatu zen. Izan ere, T_m balioen arteko diferentzia (hau da, lortutako T_m altuenaren eta baxuenaren arteko diferentzia) handiagoa zen A *Master Mix*-aren kasuan (2,15eko diferentzia) B *Master Mix*-aren kasuan baino (1,67ko diferentzia) (**1.1 taula**). Behin A *Master Mix*-a aukeratuta, denbora errealeko termoziklagailu

ezberdinak testatu ziren. Emaitzen arabera, *M galloprovincialis*, *M. edulis* eta *M. chilensis* espezieak era egokian identifika daitezke termoziklatzaile ezberdinak erabiliz (**1.2 taula**). Oztopo bakar bat antzeman zen; *Step One™* plataformaren bidez *M. trossulus* espeziearen identifikazioa ez zen era egokian ebatzi. Muga hau, termoziklagailu bakoitzarentzat fabrikatzaileak gomendatutako *Master Mix*-a erabiliz konpondu daiteke. Laburbilduz, lan honetan garatutako denbora errealeko metodoa plataforma ezberdinetan inplementa daiteke; eta kasu batzuetan, metodoaren erabilpena optimizatzeko, beste *Master Mix* bat aukeratzea bezalako aldaketa txiki batzuk egin beharko dira, antzeko lanetan deskribatu den moduan (Graziano et al., 2017).

1.1 taula. *Mytilus* espezieak identifikatzeko denbora errealeko PCR metodoaren sendotasuna bi *Mastex Mix* ezberdin erabiliz.

Espeziea	**A *Master Mix***		**B *Master Mix***	
	T_m (Batez bestekoa)	**SD**	**T_m (Batez bestekoa)**	**SD**
M. galloprovincialis	78.85	0.08	78.09	0.01
M. edulis	80.50	0.11	79.31	0.15
M. trossulus	79.17	0.07	78.01	0.03
M. chilensis	78.35	0.10	77.64	0.01

PCR guztiak *LightCycler® 480 Instrument II* termoziklagailua erabiliz burutu ziren. T_m balio bakoitza lagin bera lau aldiz neurtuz lortu ziren, eta ondoren espezie bereko T_m balioen batezbestekoa kalkulatu zen. SD: desbideratze estandarra.

1.2 taula. *Mytilus* espezieak identifikatzeko denbora errealeko PCR metodoaren sendotasuna hiru termoziklagailu ezberdin erabiliz.

Espeziea	***LightCycler® 480 Instrument II***		***CFX96 Touch DeepWell™***		***Step One™***	
	T_m (Batez bestekoa)	**SD**	**T_m (Batez bestekoa)**	**SD**	**T_m (Batez Bestekoa)**	**SD**
M. galloprovincialis	78.86	0.08	77.86	0.07	78.20*	0.355
M. edulis	80.50	0.11	79.44	0.08	79.20	0.354
M. trossulus	79.24	0.07	78.18	0.02	78.20*	0.355
M. chilensis	78.36	0.10	77.38	0.04	76.94	0.005

PCR erreakzio guztiak A *Master Mix*-a erabiliz burutu ziren. T_m balio bakoitza lagin bera lau aldiz neurtuz lortu ziren, eta ondoren espezie bereko T_m balioen batezbestekoa kalkulatu zen. SD: desbideratze estandarra.* Ondo identifikatu ez direnak.

Azkenik, metodo honen zehaztasuna ere aztertu zen. Errepikagarritasuna, *M. galloprovincialis* muskuiluaren erreferentziazko lagin bera 10 aldiz aztertu ondoren kalkulatu zen. Lagin guztiak zuzen esleitu zitzaizkion espezie honi. Erreproduzigarritasuna, bi analista ezberdinek lagin bera 10 aldiz aztertu ondoren kalkulatu zen. Era berean, lagin guztiak behar bezala identifikatu ziren % 100eko tasarekin. Zehaztasun balio hauek, aurretiaz Jilberto et al. (2017) HRM bidez *Mytilus* espezieen identifikaziorako egindako analisietan lortutako balioen antzekoak dira.

3.2. Kontserbako muskuilu produktuen analisia

Ikerketa honen helburu nagusia, Europan merkaturatzen diren *Mytilus* generoko itsas produktuen espezie nagusiak identifikatzeko metodo azkar eta zehatz bat garatzea da. Hauen artean kontserba muskuiluak daude, zeinek normalean anplifikazioa inhibi dezaketen osagai gehigarriak dituzten, metodoaren efizientzian eraginez. Gainera, kontserben esterilizazio-prozesuak DNAren zatiketa eragiten du, 200 bp edo gutxiagoko zatiak sortuz (Pardo and Pérez-Villareal, 2004; Lopez and Pardo, 2005; Chapela et al., 2007). Beraz, garatutako denbora errealeko metodoaren erabilera balioztatzeko, ezinbestekoa da honelako laginak analizatzea. Lan honetan aztertu den DNA zatiaren tamaina aski txikia da kontserben analisia era egokian burutu ahal izateko (Santaclara et al., 2006; Dias et al., 2008; Jilberto et al., 2017).

Aurretiaz, ikerketa gutxi batzuk baino ez dituzte supermerkatu eta denda lokaletan erositako produktu prozesatuak aztertu (Santaclara et al., 2006; Fernández-Tajes et al., 2011). Gure kasuan, marka eta saltsa desberdinetako 16 muskuilu produktu erosi genituen tokiko merkatuetan, Espainian saltzen diren produktu mota gehienak barne hartuz (**1.3 taula**). Lagin guztiak era egokian anplifikatu ziren denbora errealeko PCRa erabiliz, eta ondoren, DNA produktuen disoziazio-kurbak aztertuz identifikatu ziren. Kasu guztietan, lortutako T_m balioak *M. galloprovincialis* espeziearen heinaren barruan zeuden. Lortutako emaitzek, metodo hau prozesatutako *Mytilus* generoko muskuilu produktuen espezie identifikaziorako egokia dela balioztatu dute, marka eta erabilitako saltsa edozein izanda ere.

Europar batasuneko etiketatze araudiak, itsas produktuek espeziearen izen zientifiko osoa agertu behar dutela dio, hau da, generoa eta espeziea, latinezko nomenklatura binomiala erabiliz (EU, 2013b). Araudi honi jarraituz, aztertutako muskuilu produktu guztiak behar bezala etiketatuta zeudela frogatu genuen; *M. galloprovincialis* izen zientifikoa agertzen zuten eta hortaz, ez zen iruzurrik antzeman Espainiako merkatuan. Espainia, Txiletik datozen muskuilu izoztu edo prozesatuen inportatzaile nagusia da, kontserba industriarako lehengai gisa erabiltzen dituelako (FAO, 2019). Hortaz, testuinguru honetan, bai muskuiluen espeziea zein jatorri geografikoa egiaztatzeko autentifikazio tresnak garatzea ezinbestekoa da, tokiko industria eta azken kontsumitzailea babesteko.

1.3 taula. *Mytilus* espezieak identifikatzeko denbora errealeko PCR metodoaren emaitzak 16 muskuilu kontserba aztertu ostean.

Laginak	Muskuilu kontserba mota	T_m (Batez bestekoa)	SD	Identifikatutako *Mytilus* espeziea
C01	Eskabetxean	78.79	0.080	*M. galloprovincialis*
C02	Eskabetxean eta oliba oliotan (1. marka)	78.66	0.050	*M. galloprovincialis*
C03	Eskabetxean eta oliba oliotan (2. marka)	78.83	0.030	*M. galloprovincialis*
C04	Eskabetxean eta oliba oliotan (3. marka)	78.65	0.078	*M. galloprovincialis*
C05	Eskabetxean eta oliba oliotan (4. marka)	78.69	0.080	*M. galloprovincialis*
C06	Eskabetxean algekin eta oliba oliotan	78.68	0.055	*M. galloprovincialis*
C07	Eskabetxean frijituak eta oliba oliotan (1. marka)	78.66	0.081	*M. galloprovincialis*
C08	Eskabetxean frijituak eta oliba oliotan (2. marka)	78.71	0.050	*M. galloprovincialis*
C09	Eskabetxean frijituak eta oliba oliotan (3. marka)	78.70	0.026	*M. galloprovincialis*
C10	“Cabrales” gazta saltsan frijituak	78.74	0.070	*M. galloprovincialis*
C11	“Albariño” ardo saltsan frijituak	78.74	0.010	*M. galloprovincialis*
C12	Baratxuriarekin	78.76	0.061	*M. galloprovincialis*
C13	Oliba olioan frijituak	78.67	0.090	*M. galloprovincialis*
C14	Piperbeltzarekin	78.71	0.049	*M. galloprovincialis*
C15	Eskabetxean piperrautsarekin	78.68	0.071	*M. galloprovincialis*
C16	Ur gazitan	78.68	0.066	*M. galloprovincialis*

PCR guztiak A *Master Mix*-a eta *LightCycler® 480 Instrument II* termoziklagailua erabiliz burutu ziren. T_m balio bakoitza lagin bera lau aldiz neurtuz lortu ziren, eta ondoren espezie bereko T_m balioen batezbestekoa kalkulatu zen. SD: desbideratze estandarra.

4. Ondorioak

Lan honetan deskribaturiko *SYBR-Green* bidezko denbora errealeko PCR saiakuntza azkarrak, anplifikatutako produktuen disoziazio-kurben analisiarekin, Europan merkaturatzen diren *Mytilus* espezie garrantzitsuenen (*M. galloprovincialis*, *M. edulis* eta *M. chilensis*) identifikazioa ahalbidetzen du, % 100eko sentikortasun eta espezifikotasuna lortuz. Gainera, metodo hau muskuilu fresko, izoztu zein kontserba produktuen espeziea era egokian identifikatzeko erabili daitekeela frogatu da, ohiko termoziklagailu ezberdinetan eta *software* berezien beharrizanik gabe erabil daitekeena. Hortaz, kapitulu honetan deskribaturiko metodologia muskuilu espeziearen autentifikaziorako erabilgarria da eta gainera, erraz erabil daiteke elikagaien segurtasuna eta identitatea bermatzen duten kontrol laborategietan, ohiko analisi azkar gisa.

Itsas produktuen trazabilitatea eta etiketen gardentasuna bermatzeko, eta kontsumitzaileen konfiantza areagotzeko, DNAren erabileran oinarritzen diren teknika hauen potentziala aztertzea ezinbestekoa da, ez bakarrik espeziearen identifikaziorako, baizik eta jatorri geografikoa egiaztatzeko ere.

CHAPTER 2

Population structure and geographic origin assignment of *Mytilus galloprovincialis* mussels using SNPs

Published as:

del Rio-Lavín, A., Díaz-Arce, N., Larraín, M. A., Araneda, C., Rodríguez-Ezpeleta, N., Jiménez, E., and Pardo, M. A. (2022). Population structure and geographic origin assignment of *Mytilus galloprovincialis* mussels using SNPs. *Aquaculture*, 550. https://doi.org/10.1016/j.aquaculture.2021.737836

All the Supplementary Information of this manuscript can be found in **Appendix A**.

Laburpena

Arrain eta itsaskien trazabilitatea bermatzea berebiziko garrantzia du mundu osoko kontrol agintari eta itsas-industriarentzat, haien helburua iruzurra saihestea, itsas baliabideak babestea eta kontsumitzaileen segurtasuna mantentzea delako. Markatzaile genetikoek, indibiduo bat bere jatorrizko populazio genetiko edo ebolutiboari esleitzeko gaitasuna dute, eta hortaz, itsas produktuen jatorri geografikoa arakatzeko erabil daitezke. Kapitulu honetan, RAD-seq teknikaren bidez, muskuiluen genoman zehar sakabanatutako milaka SNP markatzaile identifikatu dira, eta hauek erabiliz, *M. galloprovincialis* bibalbioen populazio egitura argitu da, bere jatorrizko banaketa eremuan zehar (Atlantikoan eta Mediterraneoan) eta inbaditu berri duen eremuan (Ozeano Bareko hego-ekialdean) zehar. Gainera, aurkitu berri diren markatzaile genetiko hauen jatorria esleitzeko ahalmena aztertu da. Emaitzek Atlantikoko eta Mediterraneoko *M. galloprovincialis* populazioak nabarmen bereizten direla erakutsi dute. Aldi berean, nahiz eta antzinako populazio beretik etorri, Mediterraneoko muskuilu natiboen eta sartu berri diren Ozeano Bareko hego-ekialdeko muskuiluen artean ezberdintasun esanguratsuak daudela ikusi da. Gainera, ahalmen diskriminatzaile handiena agertu duten 10-25 SNP erabiliz, testatutako indibiduoen % 90-100a Atlantikoko edo Mediterraneoko/Hego-ekialdeko jatorri eremuetara era egokian esleitu dira. Lan honetan, markatzaile genetikoak erabiliz jatorria egiaztatzeko tresna analitiko bat garatzea posiblea dela frogatu da; akuikultura sektorearen kudeaketan, itsas produktuen trazabilitatean edota elikagaien segurtasunean erabilera izan dezakeena.

1. Sarrera

Arrantza eta akuikultura sektoreek berebiziko garrantzia dute munduko elikagai hornikuntzan eta enpleguaren garapenean. Merkatuen globalizazioa eta itsas produktuen eskari hazkuntza dela medio, hornidura kateak gero eta konplexuagoak bilakatzen ari dira, trazabilitate ahaleginak erronka handia bihurtuz. Testuinguru honetan, arrain eta itsaskien trazabilitatea bermatzeko tresna analitikoen garapena ezinbestekoa da. Itsas produktuen jatorri geografikoa egiaztatzea funtsezkoa da iruzurra saihesteko, arrantzaren eta akuikulturaren kudeaketa eraginkor eta jasangarria bermatzeko, eta elikagaien segurtasunari buruzko arauak betetzen direla ziurtatzeko. Trazabilitate tresna hauek eskura izatea garrantzitsua da bai kontrol-agintarientzat, baita haien produktuak babestu eta kontsumitzaileen konfiantza sustatu nahi duten ekoizleentzat ere.

Europar batasunean, Kontseiluaren 1379/2013 (EE) Erregelamenduak, kontseiluaren 1184/2006 (EE) eta 1224/2009 (EE) Erregelamenduak aldatzen dituena, eta Kontseiluaren 104/2000 (EE) Erregelamendua indargabetzen duena, arrantza eta akuikulturako produktuen etiketek adierazi beharrekoa arautzen du. Lege honen arabera, espeziaren merkataritza-izena eta izen zientifikoa, ekoizpen metodoa eta produktua harrapatu edo hazi den jatorri-eremua adieraztea nahitaezkoa da (EU, 2013b). Jatorria bermatzea bereziki garrantzitsua da muskuiluen kasuan, nazioarteko merkataritzan duten garrantziagatik. Muskuiluak prozesatzen direnean haien kontserbatze-denbora luzatu egiten da, eta horrek, itsas produktu honek ekoizleengandik kontsumitzaileengana egiten duten distantzia handitzea ahalbidetzen du, trazabilitatea nabarmenki zailduz.

Bibalbio hau proteina iturri garrantzitsua da eta mundu osoan aurki daiteke produktu fresko, izoztu zein prozesatu gisa. Munduko muskuilu ekoizpenak gora egin du azken hamarkadan; 2018an, konkretuki, 2,11 milioi tona ekoitzi ziren 4,519 milioi dolarreko balioa lortuz (FAO, 2020a). Akuikultura ekoizpen honetarako erabiltzen diren muskuilu haziak ingurunetik jasotzen dira, eta hortaz, populazioak naturaltzat har daitezke.

Europan zehar hiru *Mytilus* espezie bizi dira: *M. galloprovincialis*, *M. edulis* and *M. trossulus.* EBn muskuiluek akuikultura-ekoizpenaren bolumen osoaren % 34a hartzen dute *M. galloprovincialis* eta *M. edulis* izanik gehien ekoizten eta kontsumitzen diren espezieak (EUMOFA, 2019). Espainia EBeko *M. galloprovincialis* ekoizle nagusia da, urtean 200.000 tona inguru haziz (OESA, 2017; EUMOFA, 2022a). Muskuilu horien % 97a baino gehiago Galizian ekoizten da, Iberiar penintsularen ipar-mendebaldeko kostaldean.

M. galloprovincialis muskuiluaren jatorrizko banaketa-eremua Mediterraneo itsasoa, Itsaso Beltza eta Ozeano Atlantikoko ekialdeko kosta (Irlandatik Marokoraino) da (Gosling, 1992). Hala ere, gizakiak eragindako ekintzengatik, bai istripuz (itsas garraioen bidez) zein berariaz (akuikultura ekoizpenerako), eta ingurumen-baldintza desberdinetara bizkor egokitzeko duen gaitasunagatik, mundu osoan zehar hedatu da (Westfall and Gardner, 2010; Gardner et al., 2016; Larrain et al., 2018; Zardi et al., 2018; Zbawicka et al., 2018b). Ipar hemisferioko *M. galloprovincialis* muskuiluak beste herrialde bat inbaditzeak, ondorio larriak eragin ahal ditu bertako muskuilu espezie natiboen ekoizpenean. Zenbait ikerketek, *Mytilus* espezie honen presentzia aitortu dute Txile erdialdeko kostaldean (BíoBio eskualdean), inbasio iturria duela gutxiko jarduera antropogenikoak (hala nola, itsas garraioa eta akuikultura) direla aditzera emanez (Toro et al., 2005; Westfall and Gardner, 2010; Tarifeño et al., 2012; Astorga et al., 2015; Larrain et al., 2018). Hala, espezie honen jatorri geografikoa identifikatzea funtsezkoa inbasioen aurreko kudeaketa estrategia egokiak garatu eta ezarri ahal izateko. Honek, gertatutako inbasioa arin identifikatzea eta bertako espezie natiboak babestea ahalbidetuko luke.

SNPak bezalako markatzaile genetikoak arrain eta itsaskien jatorria arakatzeko potentzial handia dutela frogatu da (Martinsohn and Ogden, 2009). Jatorri geografikoaren identifikazio genetikoa, indibiduo bat bere jatorrizko populazio genetiko edo ebolutiboari esleitzeko gaitasunean oinarritzen da, eta horretarako, bere jatorrizko populazioa gainerako populazioetatik aski desberdina izan behar du. Hortaz, jatorriaren esleipenerako erabiliko diren SNP markatzaileak populazioen arteko bereizketa mailaren arabera aukeratuko dira (Ogden and Linacre, 2015). *Mytilus* muskuiluak erabiliz, SNP

markatzaileetan oinarritutako zenbait ikerketa burutu dira, bai genero honetako espezieak bereizteko bai espezie bereko populazioak bereizteko ere. SNP hauen aurkikuntzarako, datu-base publikoetan eskuragarri dauden sekuentziak (Zbawicka et al., 2012; Fraisse et al., 2015; Wenne, 2018), edo berrikiago, RAD-seq teknika (Araneda et al., 2016) erabili dira. Beste markatzaile batzuekin alderatuta, SNPek abantaila ugari dituzte; batetik, genomaren estaldura handiagoa lortzen da, bai eskualde neutroetan zein hautespen-eremuetan zehar (Morin et al., 2004; Helyar et al., 2011); eta bestetik, oso zatituta dagoen DNA aztertzeko erabil daitezke, prozesatutako itsas produktuen genotipazio azkarra eta eraginkorra ahalbidetuz (Scarano and Rao, 2014).

Ikerketa gutxi batzuk baino ez dute *M. galloprovincialis* muskuiluen dibertsitate genetikoa eta populazio egitura sakonki aztertu bere jatorrizko banaketa-eremuan zehar. SNPak erabiliz, Paterno et al. (2019)-ek Mediterraneo itsasoko eta Itsaso Beltzeko *M. galloprovincialis* populazioen egitura genetikoa aztertu zuten. Ikerketa honetan bi itsasoetako muskuilu populazioak nabarmen bereizten zirela ikusi zen; eta aldiz, mendebaldeko eta ekialdeko Mediterraneoko indibiduoen artean ez zen desberdintasunik aurkitu. Hala ere, espezie honen jatorrizko banaketa-eremu osoa aztertzeko, Atlantikoko indibiduoak aztertzea falta izan zitzaion ikerketa honi. Iberiar Penintsulako kostaldean, zenbait ikerketek deskribatu dute *M. galloprovincialis* muskuiluen populazio egitura; alozimak (Quesada et al., 1995b), markatzaile mitokondrialak (Quesada et al., 1995a; Kijewski et al., 2011) eta mikrosateliteak (Diz and Presa, 2008) erabiliz. Ikerketa horiek, Iberiar Penintsulan bi *M. galloprovincialis* populazio nagusi bereizten direla erakutsi dute, Atlantikokoa eta Mediterraneokoa. Hala ere, erabilitako markatzaile mota eta kopurua mugatua da, eta jatorriaren esleipen probabilitatea kalkulatzea falta zaie.

Lan honetan, hazitako *M. galloprovincialis* muskuiluen populazio genetikoa aztertuko dugu, bederatzi jatorri ezberdinetatik batutako laginak erabiliz eta RAD-seq bidez lortutako milaka SNP markatzaileetan oinarrituz. Gure helburuak honako hauek dira: (1) *M. galloprovincialis* muskuiluen populazio egitura zehaztea, bere jatorrizko banaketa-eremuan (Atlantikoan eta Mediterraneoan) eta inbaditu berri duen eremuan

(Ozeano Bareko hego-ekialdean); eta (2) SNPetan oinarritutako jatorria geografikoa esleitzeko tresna analitiko bat garatzea eta ebaluatzea.

2. Materials and Methods

2.1. Sample collection, DNA extraction and RAD-seq libraries preparation and sequencing

A total of 222 adult samples of *Mytilus galloprovincialis* (5-9 cm) were collected from nine different mussel farms from autumn 2018 to winter 2019 (**Table 2.1**, **Figure 2.1**). From each mussel, a piece of mantle border tissue was excised and immediately stored in RNAlater® (Sigma Aldrich, Burlington, MA, USA). Genomic DNA extraction was performed using the Wizard® Genomic DNA Purification Kit (Promega, Madison, WI, USA), starting from 50 mg of tissue and following manufacturer's instructions. Extracted DNA was eluted in sterile Milli-Q water (Millipore, Burlington, MA, USA) and its concentration was determined by means of the Quant-iT dsDNA HS assay kit using a Qubit® 1.0 Fluorometer (Life Technologies, Carlsbad, CA, USA). DNA integrity was assessed by electrophoresis on agarose gel 0.8% (w/v). Species identification was performed by a real-time PCR assay targeting the polyphenolic adhesive protein gene using primers previously described by Dias et al. (2008), with a SYBR-Green post-PCR melting curve analysis (del Rio-Lavín et al., 2021). This methodology shows specific peaks in dissociation curves with unique T_m values for each species.

RAD-seq libraries were prepared following the method described by Etter et al. (2011). Briefly, 500 ng of starting DNA was digested with the *SbfI* restriction enzyme (Thermo Fisher Scientific, Waltham, MA, USA) and ligated to modified Illumina P1 adapters containing 5 bp unique barcodes. Pools of DNA from 32 individuals were sheared using the Covaris® M220 Focused-ultrasonicator™ Instrument (Life Technologies, Carlsbad, CA, USA) and size selected to 300-500 bp by AMPure XP Beads (Beckman Coulter, Brea, CA, USA) (Bronner et al., 2014). After Illumina P2 adaptor ligation, each library was amplified using 12 PCR cycles. Each pool was paired-end sequenced (125 bp) on an Illumina HiSeq2500 (Illumina Inc., San Diego, CA, USA).

Table 2.1. Mussel samples used for population genetic analyses. MED = Mediterranean area; ATL = Atlantic area; SEP = South-eastern Pacific area.

Area	Country/Region	Sampling point	Code	Latitude	Longitude	Nº indiv.	Nº indiv. genotyped	Nº indiv. training set	Nº indiv. holdout set
ATL	Spain / Basque country	Mutriku	MUES	43º 18'40.2'' N	2º 22'35.2'' W	25	20	10	10
	Spain / Basque country	Mendexa	MEES	43º21'20.2'' N	2º 26'53.4'' W	25	22	9	13
	Spain / Galicia	Ría de Betanzos-Sada	SGES	43º 23' 45.0'' N	8º 17' 40.0'' W	23	10	4	6
	Spain / Galicia	Ría de Arousa	AGES	42º 29' 36.0'' N	8º 55' 51.0'' W	24	18	10	8
	Portugal / Algarve	Porto da Baleeira	PBPT	37°00'41.8" N	8°55'48.7" W	25	24	14	10
						122	**94**	**47**	**47**
MED	Spain/ Catalonia	Delta del Ebro	DEES	40°33'16.3" N	0°32'27.8" E	25	22	10 (9)	12 (13)
	Tunisia / Bizerte	Bizerte lagoon	LBTN	37°11'48.0" N	9°51'23.0" E	25	10	4 (7)	6 (3)
						50	**32**	**14 (16)**	**18 (16)**
SEP	Chile / Biobío	Coliumo	COCL	36º22'16.4'' S	72º57'25.2'' W	25	18	10 (9)	8 (9)
	Chile / Biobío	Caleta Tumbes	TBCL	36º32'14.0'' S	72º57'32.0'' W	25	23	13 (12)	10 (11)
						50	**41**	**23 (21)**	**18 (20)**
					Dataset 1	**222**	**167**	**84**	**83**
					Dataset 2	**100**	**73**	**37**	**36**

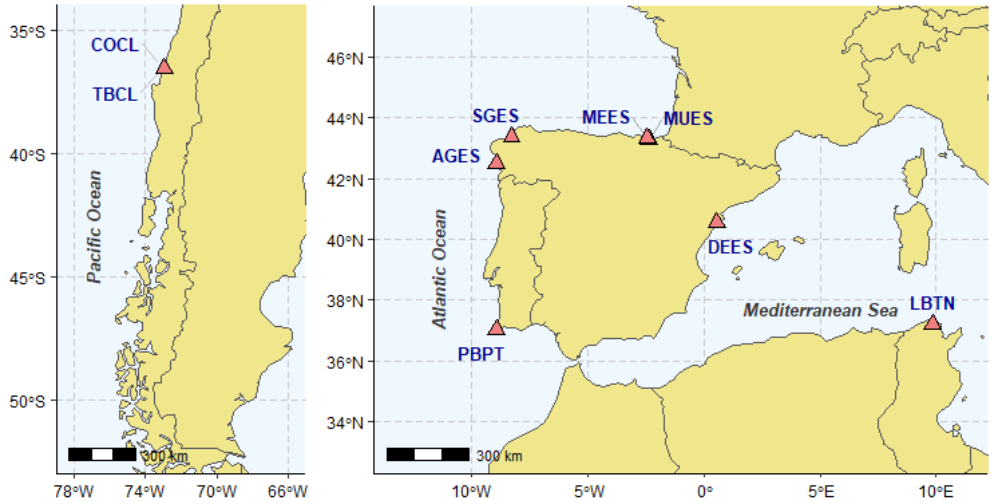


Figure 2.1. Map showing the geographic location where the samples used for this study were collected. See **Table 2.1** for definitions of abbreviations.

2.2. RAD-tag assembly and SNP calling

Raw reads were analysed using Stacks version 2.3b (Catchen et al., 2013). Detailed procedures followed in this study are shown in **Figure S1**. Demultiplexing and quality filtering were performed using the "process_radtags" module, truncate ng all reads to 80 nucleotides to avoid the low-quality bases at the end of the reads. PCR duplicates were removed by applying the "clone_filter" module to the reads whose forward and reverse pairs passed filtering. Only samples with at least 500,000 quality passing forward reads were kept (**Table S2.1**). Stacks per individual were *de novo* assembled using the "ustacks" module with a minimum depth coverage to create a stack (m) of three and a maximum nucleotide mismatches (M) allowed between stacks of four. Only samples with between 6000 and 25000 RAD loci and a minimum coverage of 10X were kept. A catalog of RAD loci was generated using the "cstacks" module, allowing a maximum number of mismatches between sample tags (n) of five and using a population map with the nine mussel production points. Stacks parameters (m, M and n) were selected following the method of *de novo* parameter optimization based on Paris et al. (2017). Matches of individual RAD loci to the catalog were searched using the "sstacks" module, data were stored by locus and SNPs were called using information from paired-end reads using the "tsv2bam" and "gstacks" modules.

From the generated catalog, genotype tables including all individuals (Dataset 1) or only the Mediterranean and South-eastern Pacific individuals (Dataset 2) were generated as follows. SNPs present in RAD loci found in at least 70% of the individuals across populations were selected and exported to PLINK, Variant Call Format (VCF) and Genepop formats using the "populations" module. Only SNPs located on the forward reads were considered. Using PLINK version 1.09b (Purcell et al., 2007), SNPs with a Minimum Allele Frequency (MAF) lower than 0.05, a genotyping rate lower than 0.9 and failing the Hardy-Weinberg equilibrium test at $P<0.05$ in two or more of the sampling points were excluded. Only samples with a genotyping rate above 0.8 were retained. Genotype tables were exported to Structure and Genepop formats using PDGSpider 2.1.1.5 (Lischer and Excoffier, 2012).

2.3. Genetic diversity and population structure analysis

For population structure, the following analyses were performed for each dataset including only the first SNP per tag. Principal Component Analysis (PCA) was performed using the adegenet R package with no previous population assignment of samples (Jombart and Ahmed, 2011). ADMIXTURE (Alexander et al., 2009) was run setting 5000 bootstraps for each K value (number of potential ancestral populations, that ranged from 1 to the number of presumed populations +1). A first ADMIXTURE run was launched for each value of K to check the number of steps necessary to reach the default 0.001 likelihood value during the run. This information was used to set the "-c" parameter (steps to be fulfilled in each bootstrapped run) to assure convergence. The K value with the lowest associated error was identified using ADMIXTURE's cross-validation procedure. Genetic differentiation between sampling areas was analysed through the total and pairwise F_{ST} (fixation index) values with a 95% confidence interval (CI) estimated running 10,000 bootstraps with the R package diveRsity (Keenan et al., 2013). Loci potentially influenced by positive selection in each dataset were identified using the R package pcadapt (Luu et al., 2016). A screeplot representing the percentage of variance explained by each PC was used to choose the number of principal components (K), and SNPs with p-values (adjusted following Benjamini and Hochberg (1995)) below 0.05 were classified as genomic outliers.

2.4. Origin assignment informative loci selection

The selection of the most informative loci for geographic origin assignment was performed based on F_{ST} values calculated by GENEPOP 4.7.3 (Raymond and Rousset, 1995; Rousset, 2008) following Weir and Cockerham, 1984. Considering the genetic clusters obtained in this study, two different assignment analysis were performed in order to match individuals to one of the three ocean areas of interest: (1) Atlantic *vs* Mediterranean/South-eastern Pacific individuals (using Dataset 1) and (2) Mediterranean *vs* South-eastern Pacific individuals (using Dataset 2). Genetic assignment was performed following the "Training, Holdout & Leave-one-out" (THL) method (Anderson, 2010) in order to avoid high-grading bias. Briefly, for each dataset, individuals were randomly divided in two groups (training and holdout) with equally distributed number of samples from each location (**Table 2.1**). Training samples were used to identify the loci with the highest discriminative power to assign samples their areas of origin based on F_{ST} values. Next, using both training and holdout samples as reference, assignment scores for each holdout individual was calculated with Geneclass2 (Piry et al., 2004) using Rannala and Mountain (1997) criterion (0.05 threshold) and following leave-one-out procedure. Assignment power of the most discriminant SNPs was assessed by calculating percentages of correctly and incorrectly assigned holdout individuals to their origin. The assignment rate was calculated using two levels of assignment probability thresholds: 80% and 90%, and samples below these values were considered unassigned.

3. Results

All the 222 mussel samples analysed in this study were identified as *M. galloprovincialis*, all showing the dissociation curve peak corresponding to a melting temperature of 78.856 ± 0.16. Three samples from Galicia and five from the Basque Country, excluded from the analysis, were identified as *M. galloprovincialis* x *M. edulis* hybrids, meaning that hybrids were found in four of the nine production points.

3.1. RAD-tag assembly and SNP calling

The total number of read pairs retained after quality filtering was 288,996,135 (79%), with an average of 1,301,784 reads per individual. After PCR duplicate removal, with an average of 27.6% clones per individual, 201,648,224 were used for SNP identification. The mean number of RAD tags obtained per sample was 13,976. After individual and SNP filtering, the final datasets resulted in 167 samples and 959 SNPs (Dataset 1) and 1,506 SNPs and 73 samples (Dataset 2) (**Table S2.2**).

3.2. Population Structure and Genetic Clustering

PCA and ADMIXTURE analyses showed that mussels from the Atlantic area (Basque Country, Galicia and Algarve) are genetically distinct from those from the Mediterranean Sea (Ebro Delta and Bizerte) and South-eastern Pacific (Biobío) areas (**Figure 2.2A and 2.3A**). ADMIXTURE analysis identified K=2 as the most likely number of ancestral populations whose proportions differentiate two main *M. galloprovincialis* populations within the Iberian coast. Individuals from the Atlantic Ocean were clearly separated from the Mediterranean Sea and South-eastern Pacific populations by the first principal component explaining 5,55% of variance in allele frequency in the PCA analysis. Interestingly, the Mediterranean samples grouped closer to the South-eastern Pacific samples, than to the Atlantic samples. Although originated from the same ancestral populations, samples from Mediterranean and South-eastern Pacific locations have diverged and are genetically differentiated (**Figure 2.2B and 2.3B**).

Total F_{ST} estimations revealed significant differences between the three above mentioned groups. The highest F_{ST} values were obtained between the Atlantic and the other two areas: 0.0959 (95% CI: 0.0882, 0.1043) with respect to the Mediterranean and 0.0723 (95% CI: 0.0660, 0.0795) with respect to the South-eastern Pacific. As expected, a lower value was obtained for Mediterranean with respect to the South-eastern Pacific F_{ST} statistic (0.0134) (95% CI: 0.0070, 0.0207). This genetic differentiation was also reflected in the pairwise F_{ST} estimations between the nine mussel sampling points (**Table 2.2**). Although F_{ST} values between the Easten Pacific COCL location and the Mediterranean populations were not significantly different from zero at 95% CI when

using Dataset 1, a clear and significant genetic differentiation was observed when using the specific dataset developed for these two areas (Dataset 2) (**Table S2.3**). Outlier analyses run by pcadapt detected 40 and 9 SNPs showing clear evidence of positive selection (pvalue<0.05) for Dataset 1 and 2 respectively.

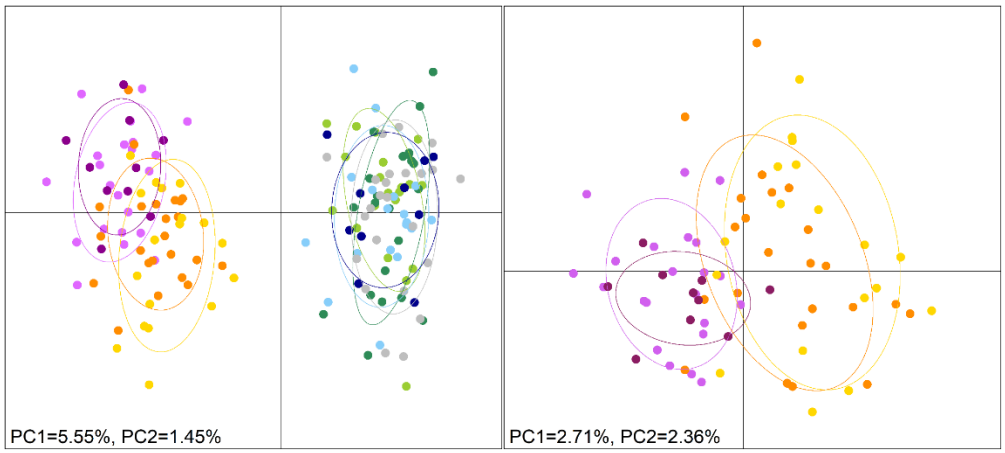


Figure 2.2. Principal Component analysis (PCA) of mussel samples. **(A)** Individuals from all areas (Dataset 1): Atlantic (● MUES |● MEES|● SGES|● AGES |● PBPT), Mediterranean (● DEES|● LBTN) and South-eastern Pacific (● COCL|● TBCL). **(B)** Mediterranean (● DEES|● LBTN) and South-eastern Pacific (● COCL|● TBCL) individuals (Dataset 2). Each point represents one sample and colours denote its sampling site. Ovals represent 95% inertia ellipses.

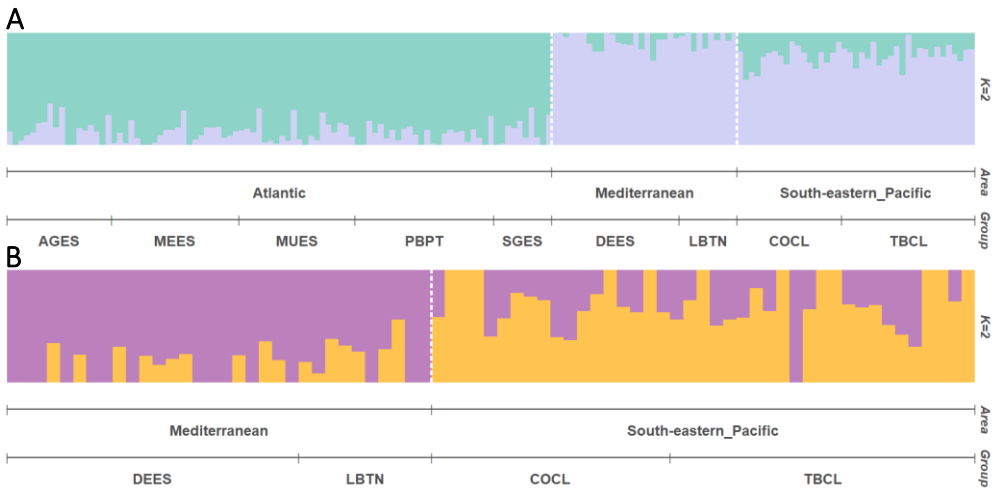


Figure 2.3. Graphical representation of ADMIXTURE clustering approach, where each bar represents an individual and each colour its inferred membership to each of the two potential ancestral populations (K=2). (A) Dataset 1 including all areas: Atlantic (MUES|MEES|SGES|AGES|PBPT), Mediterranean (DEES|LBTN) and South-eastern Pacific (COCL|TBCL). (B) Dataset 2 including only the Mediterranean (DEES|LBTN) and South-eastern Pacific individuals (COCL|TBCL). See **Figure S2.3** for K=3.

Table 2.2. Pairwise F_{ST} values matrix for all mussel farming points (using Dataset 1): Atlantic (MUES|MEES|SGES|AGES|PBPT), Mediterranean (DEES|LBTN) and Eastern Pacific (COCL|TBCL). Values significantly different from zero at 95% CI are in bold.

	MUES	MEES	SGES	AGES	PBPT	DEES	LBTN	COCL
MEES	-0.0003							
SGES	0.0006	0.0020						
AGES	-0.0018	0.0006	-0.0021					
PBPT	0.0035	0.0020	0.0044	0.0023				
DEES	**0.0934**	**0.1005**	**0.1025**	**0.0935**	**0.1030**			
LBTN	**0.0873**	**0.0966**	**0.1016**	**0.0953**	**0.1020**	0.0047		
COCL	**0.0615**	**0.0664**	**0.0694**	**0.0656**	**0.0711**	0.0127	0.0139	
TBCL	**0.0712**	**0.0781**	**0.0773**	**0.0739**	**0.0823**	**0.0138**	**0.0171**	-0.0017

3.3. Traceability tool development

With the aim of developing a traceability SNP panel with the highest accuracy and lowest number of SNPs, the 100 most discriminant SNPs were selected using half of the individuals (**Table S2.4**). The remaining half, excluded from the SNP selection, were assigned using Geneclass2. The assignment was performed combining both, neutral and outlier loci information; being under positive selection 18 out of 100 SNPs for dataset 1 and 1 out of 100 SNPs for dataset 2 (**Table S2.4**). For Dataset 1, the percentage of correctly assigned samples when using 10 to 25 SNPs ranked based on F_{ST} criterion, resulted in between 90-100% of assignment for Atlantic and Mediterranean/South-eastern Pacific samples (**Figure 2.4A, Figure S2.2)**. For Dataset 2, assignment success to the region of origin using a panel of 75 SNPs ranked according to de F_{ST} value resulted in 87.5% and 90% correct assignment of the Mediterranean and South-eastern Pacific origin samples, respectively (**Figure 2.4B, Figure S2.2).**

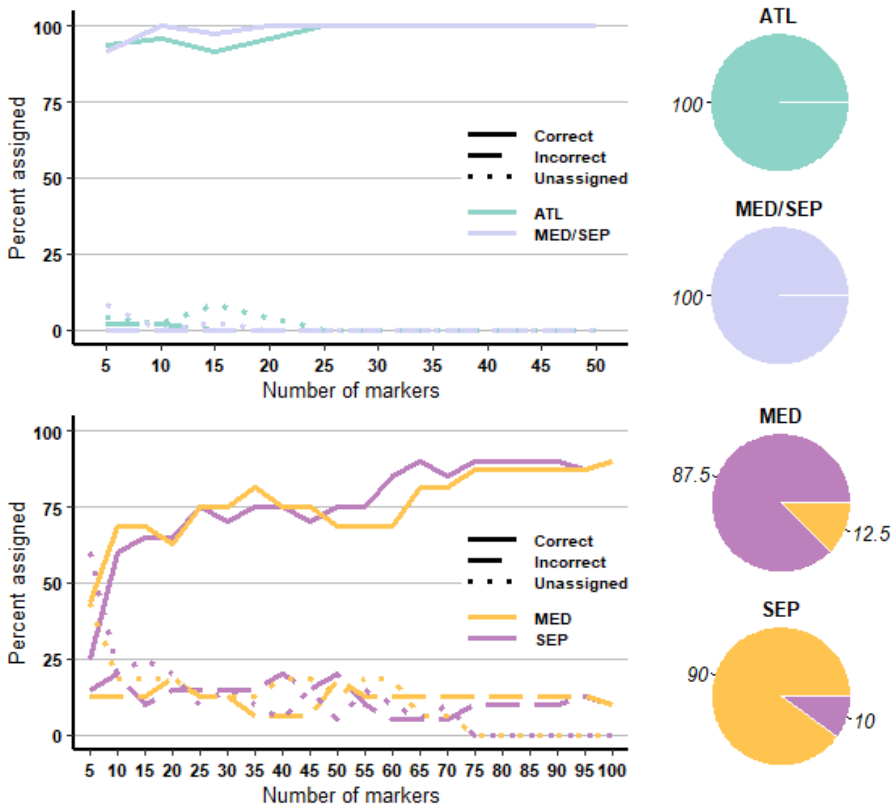


Figure 2.4. Origin assignment of individuals following THL method. (Left) Progression of the percentage of correctly, incorrectly and unassigned Atlantic (ATL), Mediterranean (MED) and South-eastern Pacific (SEP) samples as the number of markers (SNPs) increases. (Right) Percentages of samples assigned to their origin using a 25 SNP panel for Dataset 1 (A) and a 75 SNP panel for Dataset 2 (B). Markers were selected and ranked according to their F_{ST} value. The assignment threshold used was 80% in both cases, see **Figure S2.2** for 90% assignment scores.

4. Discussion

4.1. Population structure of *M. galloprovincialis*

The analysis of the *M. galloprovincialis* from nine locations performed here has provided a robust population genetic structure of this species in part of its both native and introduced range. This has allowed the clear differentiation between two mussel lineages within *M. galloprovincialis*, the Atlantic and the Mediterranean, with significant

differences observed between the ancestral Mediterranean native and the more recently introduced mussel in the South-eastern Pacific area.

The genetic differentiation between Atlantic and Mediterranean populations is concordant with previous studies performed along the Iberian Peninsula using allozymes (Quesada et al., 1995b), mitochondrial markers (Quesada et al., 1995a; Kijewski et al., 2011) and microsatellites (Diz and Presa, 2008). These studies place the Iberian phylogeographic break between Almeria and Alicante, which corresponds with the position of the well-defined Almeria-Oran oceanographic front (AOOF). The concordance between these four different marker types strongly supports genetic differentiation between the two Iberian *M. galloprovincialis* populations and proves the temporal stability of this structure, allowing the development of accurate geographic origin verification tools. The differentiation between these two lineages using SNP markers has also being proved when analysing a set of reference Northern hemisphere *M. galloprovincialis* individuals from Oristano (Italy) and Punta Camarinal (Spain) (Larrain et al., 2018; Zbawicka et al., 2018b; Zbawicka et al., 2019; Popovic et al., 2020). The use of thousands of SNP markers throughout the genome, has increased the resolution in population structure revealing high levels of genetic differentiation (F_{ST}=0.0867). Among Atlantic samples, our findings contradict the large proportion of significant pairwise F_{ST} values found when using six microsatellites (Diz and Presa, 2008), since no significant differences were depicted. This discrepancy could be explained by the higher number of genetic markers used in our study, which might provide a better resolution of the genetic differentiation.

We found that Mediterranean samples grouped closer to the South-eastern Pacific samples than to the Atlantic samples, which is explained by a relatively recent human-mediated introduction of Mediterranean *M. galloprovincialis* mussels on the coast of Chile (Westfall and Gardner, 2010; Tarifeño et al., 2012; Larrain et al., 2018). The presence of Mediterranean populations in South America has also recently been described by Lins et al. (2021), being genetically more similar to the Black Sea and South-eastern Mediterranean populations. Nevertheless, although originated from Mediterranean ancestral population, our analyses support significant differentiation

between current Mediterranean and South American populations. This indicates that these populations have diverged over the years and are now genetically differentiated. Interestingly, Larrain et al. (2018) found an F_{ST} value of zero between Mediterranean (from Oristano, Italy) and South-eastern Pacific samples (from BioBío, Chile), potentially due to the fact that this previous study used lower number of SNP markers. The thousands of SNP markers developed here specifically for *M. galloprovincialis* mussels may have uncovered an existing genetic differentiation between these two populations. Within the Mediterranean samples, the complete lack of differentiation found agrees with the genetic structure observed between *M. galloprovincialis* mussels from Western Mediterranean by Paterno et al. (2019), where no differences were found between samples from the Western and Central Mediterranean.

4.2. SNP markers for *M. galloprovincialis* geographical origin assignment

Once the overall picture of *M. galloprovincialis* population structure has been settled, the geographic origin assignment power of SNP markers presented here paves the way for an accurate and cost-effective traceability tool development. Due to the high genetic differentiation observed between Atlantic and Mediterranean *M. galloprovincialis*, our study has shown that a correct assignment between 90-100% can be obtained when using 10 to 25 SNPs. As in our study, small SNP panels with high assignment power have previously shown to be useful for genetic traceability systems of other marine species. Proof of this are the small panels of SNP markers (n = 10–30) that were developed for the traceability of European commercial fish (cod, herring, sole and hake) with 93-100% of correct assignment (Nielsen et al., 2012; Ogden and Linacre, 2015). In this sense, Montes et al. (2017) also detailed the use of a SNPs-based test to determine the provenance of anchovies with an overall assignment percentage of 94.3% of the individuals analysed.

As expected, our SNP traceability panel was less effective distinguishing between mussels from the Mediterranean and South-eastern Pacific areas, which is consistent with the patterns observed in the population structure analysis. A possible explanation could be the insufficient accumulation of genetic differences, either by genetic drift or

adaptation to the local environment between these two populations, due to the relatively recent introduction of the population in the South-eastern Pacific. Nevertheless, the assignment score for the Mediterranean and South-eastern Pacific samples could be improved by increasing the number of samples used for SNP selection.

Previous origin assignment analysis of mussels in South America, where Atlantic and Mediterranean *M. galloprovincialis* samples were used as reference to identify species within *Mytilus* genus, observed lower assignment rates between these two European locations using higher number of SNPs (Larrain et al., 2018; Zbawicka et al., 2018b). Larrain et al. (2018) used a 49-SNP panel to assign the Atlantic and Mediterranean populations with an 86.2% and 93.1% assignment success, respectively, suggesting mixture of individuals between these two European locations. Zbawicka et al. (2018b) used a 51-SNP panel to assign Atlantic and Mediterranean samples with an 89.65% and 96.55% assignment success. In both studies, only two sampling sites were selected around the Iberian Peninsula, having 29 samples each of them. This lower assignment power can be explained by the fact that the SNPs selected for the panels aimed to identify the different *Mytilus* species located in the Southern Hemisphere and not to differentiate lineages within Northern Hemisphere *M. galloprovincialis*. As shown in our study, the use of a specific SNP panel developed precisely for origin assignment of *M. galloprovincialis,* shows a better performance in assigning the geographical origin than using SNPs shared by different species.

In summary, SNP panels represent a promising traceability tool for food authorities due to their easy laboratory transfer, their potential assignment power and their applicability in processed seafood products. However, it is worth mentioning that the spatial resolution of DNA based methods is often limited, since it is not possible to differentiate samples from different farming locations within the same geographical area. In these cases, a multidisciplinary approach combining DNA based techniques with other techniques that might reflect individuals' surroundings, such as isotope or trace element analysis, would improve the resolution determining the geographic location. In addition, the origin assignment markers presented here could also help to determine the source of invasion of this species in non-native areas and to understand the subsequent

impacts in local shellfish production. Consequently, this will allow the development of management strategies to mitigate future invasions, ensuring the conservation of native biodiversity.

Our research supports the possibility of the development of an accurate and cost-effective origin assignment tool based on a reduced panel of SNPs which can be applied to fresh, frozen or processed mussel products. Developing genetic tools to assign seafood to their provenance will protect marine resources, guarantee sustainable aquaculture management, enforce current food safety regulations and prevent commercial fraud.

CHAPTER 3

Stable isotope and trace element analysis for tracing the geographical origin of the Mediterranean mussel (*Mytilus galloprovincialis*) in food authentication

Published as:

del Rio-Lavín, A., Weber, J., Molkentin, J., Jiménez, E., Artetxe-Arrate, I., and Pardo, M. A. (2022). Stable isotope and trace element analysis for tracing the geographical origin of the Mediterranean mussel (*Mytilus galloprovincialis*) in food authentication. *Food Control*, 139. https://doi.org/10.1016/j.foodcont.2022.109069

* This study was conducted in collaboration with stable isotope experts of the Max Rubner-Institute (Jan Weber and Joachim Molkentin). Ane del Rio-Lavín was in charge of trace element analysis.

All the Supplementary Information of this manuscript can be found in **Appendix B**.

Laburpena

Arrain eta itsaskien jatorri geografikoa bermatzea berebiziko garrantzia du mundu osoko kontrol agintari eta itsas-industriarentzat; haien helburua, iruzurra saihestea, itsas baliabideak babestea eta kontsumitzaileen segurtasuna mantentzea delako. Hortaz, kapitulu honetan, *M. galloprovincialis* muskuiluen jatorri geografikoa egiaztatzeko, aztarna elementuen eta isotopo egonkorren konposizioa, era independentean zein konbinatuan, erabiltzearen potentziala aztertu da. Analisirako, laginak Mediterraneo itsasoko, Europako kostalde Atlantikoko eta Ozeano Bareko hego-ekialdeko zortzi eskualde ezberdinetatik hartu dira. Konkretuki, muskuilu ehunaren karbono eta nitrogeno isotopoak (n=179) eta muskuilu maskorren 16 aztarna-elementu (n=100) neurtu dira. Emaitzen arabera, alde nabarmenak daude akuiluktura-instalazio ezberdinetan hazitako muskuiluen konposizio kimikoen artean, eta bereziki eskualde geografiko desberdinetakoak direnen artean. Ikaste automatikoan oinarritutako *Random Forest* sailkapen-metodoa erabiliz, aztarna kimikoen (isotopo egonkor eta aztarna-kimikoen) datuen konbinazioak, jatorria zehaztasun handiz aurreikusteko gaitasunik handiena erakutsi du; testatutako indibiduoen % 97a behar bezala esleituz dira, jatorriaren bereizketarako elementu garrantzitsuenak $\delta^{13}C$, $\delta^{15}N$, Pb, Ba, Mn, eta Al izanik. Lan honetan lortutako emaitzek, isotopo egonkorren eta aztarna elementuen analisian oinarritutako tresna analitikoak itsas produktuen jatorri geografikoa egiaztatzeko egokiak direla frogatu da; elikagai hauen trazabilitatea eta segurtasuna bermatuz.

1. Introduction

Global production of aquatic resources, including fisheries and aquaculture, reached a new peak in 2018, with a production of about 178.5 million tonnes of fish, molluscs, crustaceans and others, with an estimated total value of US$164.1 billion U.S. dollars (FAO, 2020b). Within the fishery sector, mussels are an economically important product which is traded globally. The declaration of mussels has strict requirements and standards particularly when entering the EU market, e.g., regarding the indication of origin or production method. In 2018, 17.7 million tonnes (USD 34.6 billion) of molluscs, mainly bivalves, were harvested, with consistent growth in production due to persistently high prices for bivalves in the past (FAO, 2020b). Mussels of the species *Mytilus spp*. are globally almost exclusively farmed in aquacultures in China, the EU and Chile (EUMOFA, 2019). Within the widespread genus *Mytilus*, three species occur exclusively in the Northern Hemisphere (*M*. *trossulus*, *M*. *edulis*, *M*. *californianus*) and two in the Southern Hemisphere (*M*. *chilensis* and *M*. *platensis*), while *M*. *galloprovincialis* occurs on both hemispheres (McDonald et al., 1991; Gaitán-Espitia et al., 2016). The latter species, the Mediterranean mussel, is mainly endemic in Mediterranean Sea and Atlantic Ocean waters, while it is considered as an invasive species off North and South America, Australia, New Zealand, South Africa and East Asian coasts (Hilbish et al., 2010; Westfall and Gardner, 2013; Larrain et al., 2018; Zardi et al., 2018; Zbawicka et al., 2018a; Gardner et al., 2021). Due to the overlapping habitats, natural inter-species hybridisation challenges the genetic authentication of *Mytilus* species for a valid declaration on the food market (Kijewski et al., 2009).

Considering the economic relevance of this extensively farmed species and its importance in the international trade, verification of its geographical origin is necessary for labelling, traceability and food safety purposes (Hixson, 2014; Luque and Donlan, 2019). Species identification is mandatory in the EU, as stated in Council Regulation (EC) No 1379/2013 (EU, 2013b). This includes the labelling of the commercial and scientific name, the production method and the geographical origin of the product. Furthermore, it should be noted that mussels from the region of Galicia (“Mexillón de Galicia”, Spain)

and Scardovari ("Cozza di Scardovi", Italy) have been authorized with the PDO in 2007 and 2013, respectively, due to their local history in mussel cultivation (EU, 2007; EU, 2013a).

Mislabelling of the geographical origin is a common fraud, affecting the credibility of traders and deceiving consumers (Luque and Donlan, 2019). Innovative, accurate and sensitive analytical methods and approaches are essential to verify the compliance with the regulations in food authenticity. Methods used for seafood traceability include genetic markers (Nielsen et al., 2012; Ogden and Linacre, 2015; Giusti et al., 2020; Kappel and Schröder, 2020), fatty acid profiling (Ricardo et al., 2017a; Cherifi et al., 2018) and near infrared spectroscopy (Ghidini et al., 2019). In addition, stable isotopes and elemental fingerprints have been recognized as particularly useful origin discriminators (Li et al., 2016).

SIRA has been established as a potent tool to trace the local provenance of bivalves and other aquatic species including fish, shrimps and sea cucumber (Gustafson et al., 2007; Deudero et al., 2009; Suh and Shin, 2013; Ortea and Gallardo, 2015; Zhao et al., 2019; Kang et al., 2020). Bivalves are sessile filter-feeders that assimilate in their tissues the isotopic composition from feeding sources such as phytoplankton, zooplankton, suspended particulate organic matter and other microorganisms from the water column (Purchon, 1968). Stable isotopes of carbon (δ13C) and nitrogen (δ15N) are well known geochemical tracers to study food web structures and ecosystem connectivity in aquatic environments (Peterson and Fry, 1987; Trueman and St John Glew, 2019). δ13C values of consumers reflect the source of energy that is produced by primary producers (e.g., phytoplankton, zooplankton, or particulate organic matter), while δ15N values reflect the relative trophic position of the organism within the food web (Deniro and Epstein, 1981; Post, 2002). Although δ13C and δ15N composition of organisms primarily reflects their food absorbed (Deniro and Epstein, 1978; Deniro and Epstein, 1981), it has to be considered that isotope signatures of the mussels' food can be also influenced by local abiotic factors that may vary seasonally, such as sea surface temperature and dissolved inorganic carbon in the case of carbon isotope signatures or changing nutrient conditions

(e.g. eutrophic *vs* oligotrophic system) in the case of nitrogen isotope signatures (Briant et al., 2018).

In addition to considering seasonal variations in isotopic values, the selection of the sample material is critical, as different mussel tissues show different isotopic enrichment. While muscle tissue represents the average isotopic composition of the diet consumed over a longer period, the digestive gland and gonad, which have a high metabolic activity, primarily tend to reflect the recent diet (Paulet et al., 2006; Ezgeta-Balic et al., 2014). $\delta^{13}C$ and $\delta^{15}N$ in mussel tissue have already been used to trace the origin of *M. galloprovincialis* from the Balearic Islands and Yessow scallops (*Patinopecten yessoensis*) from the Yellow Sea (Deudero et al., 2009; Zhao et al., 2019) , anthropogenic pollution and sewage (Babaranti et al., 2019; Ozdilek et al., 2019), trophic level in the food web (Briant et al., 2018), and seasonal variation in the primary diet (Page and Lastra, 2003; Ezgeta-Balic et al., 2014). Accordingly, analysing $\delta^{13}C$ and $\delta^{15}N$ in the muscle of *M. galloprovincialis* while considering seasonal variation seems a promising approach for origin identification of internationally traded bivalves.

Over the last decades, trace elements have attracted much attention as natural fingerprints to trace the geographical origin of bivalves, both to identify population connectivity and to reconstruct past environmental changes (Becker et al., 2005; Carson, 2010; Gomes et al., 2016; Marali et al., 2017; Prendergast et al., 2017; Honig et al., 2020). As surrounding water chemistry and environmental conditions influence the incorporation of different elements into bivalve shells, variations in the concentrations of selected elements can be used as natural markers to discriminate among groups of mussels grown under different environments (Becker et al., 2005; Strasser et al., 2008). Regarding traceability, several studies have demonstrated the capability of ICP-MS based on TEF of bivalve shells to assign samples to their origin(Dunphy et al., 2015; Ricardo et al., 2015; Ricardo et al., 2017b; Bennion et al., 2019; Morrison et al., 2019; Bennion et al., 2021). However, these studies have limited their sampling sites to smaller spatial scales within the same country and none of them has focused on tracing the harvest location of the extensively worldwide distributed *M. galloprovincialis*.

The present study aims to assess the potential of using chemical signatures to trace the geographical origin of *M. galloprovincialis* mussels collected from ten commercial farms in the Mediterranean Sea, as well as on the European Atlantic coast and the Chilean Pacific coast. In addition to the two independent methodical approaches SIRA and TEF, we present a novel seafood traceability approach that combines organic stable isotope with inorganic trace element analysis.

2. Materials and methods

2.1. Sampling and preparation

M. galloprovincialis samples for SIRA and TEF were collected between September 2018 and September 2019, with exception of Tunisian samples that were collected in January 2018. Collected mussel samples were immediately frozen (≤ -20 °C) until further preparation. For this study, we defined the sampling seasons as follows: spring (April to May), summer (June to August), autumn (September to November) and winter (December to March). This division refers to the seasons in the Northern Hemisphere. Samples were obtained from 10 different locations (**Figure 3.1**) in 6 different countries namely Spain, Portugal, France, Italy, Tunisia, and Chile (**Table 3.1**). The species identification of each individuum was verified genetically by a SYBR-Green post-PCR melting curve analysis (Del Rio-Lavín, Jiménez, & Pardo, 2021). The sample set of the study consisted of a total of 179 soft tissue samples for SIRA and 100 shell samples for TEF; of which 64 individuals had both, the soft tissue and the shell, analysed (**Table 3.1**).

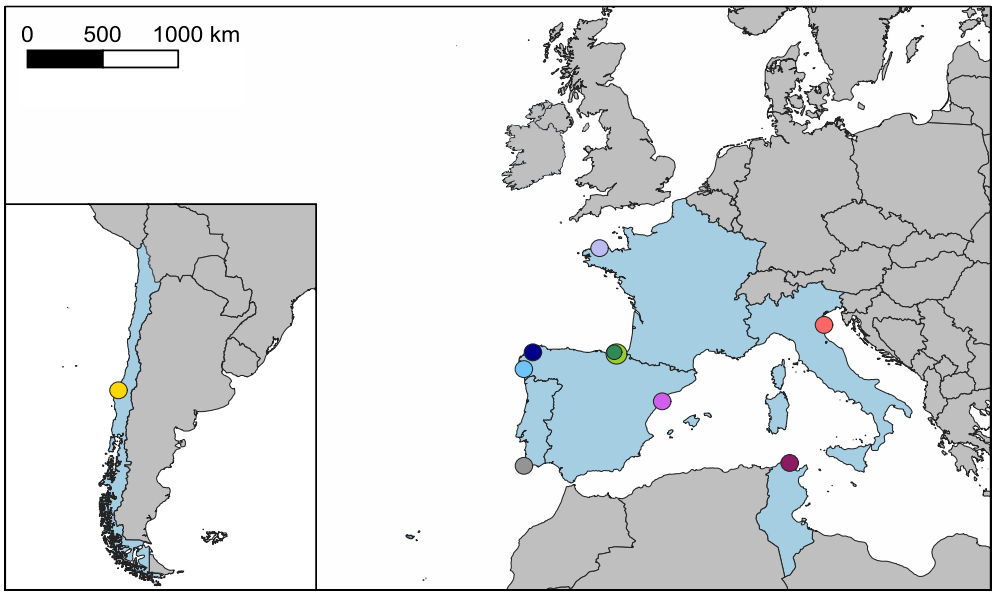


Figure 3.1. Sampling locations of *M. galloprovincialis* and colour code; Atlantic Ocean area (● LBFR, ● PBPT, Galicia region (● AGES, ● SGES), Basque Country region (● MEES, ● MUES)); the Mediterranean Sea area (● DEES, ● GEIT, ● LBTN); and the South Eastern Pacific Ocean area (● COCL). See **Table 3.1** for definitions of abbreviations

Table 3.1. Samples used for TEF and SIRA analysis. MED = Mediterranean Sea area; ATL = Atlantic Ocean area; SEP = South Eastern Pacific Ocean area. *5 samples per season. ** Autumn missing.

Country	Area	Region	Sampling location	Acronym	n (TEF)	n (SIRA)	Seasonal sampling (SIRA)*	n (TEF + SIRA)
France	ATL	Brittany	Loquemeau	● LBFR	10	17	yes**	8
Portugal	ATL	Algarve	Porto da Baleeira	● PBPT	10	20	yes	6
Spain	ATL	Galicia (GAES)	Ría de Arousa	● AGES	10	20	yes	5 + 5
			Ría de Betanzos-Sada	● SGES	10	20	yes	
Spain	ATL	Basque Country (BQES)	Mendexa	● MEES	10	20	yes	5 + 5
			Mutriku	● MUES	10	20	yes	
Spain	MED	Catalonia	Delta del Ebro	● DEES	10	20	yes	5
Italy	MED	Emilia-Romagna	Goro	● GEIT	10	10	no	10
Tunisia	MED	Bizerte	Bizerte lagoon	● LBTN	10	12	no	10
Chile	SEP	Biobío	Coliumo	● COCL	10	20	yes	5
Total					*100*	*179*		*64*

Mussel samples for SIRA (n=179) were processed by removing the digestive gland from the tissue. The remaining soft tissue was cut into fine pieces of a few millimetres, mixed with purified water 1:2 (v/v) and chopped to a homogenous slurry using an Ultra-Turrax disperser (IKA, Königswinter, Germany). Finally, the slurry was lyophilized. Alternatively, the cut mussel tissue was directly dried in an oven set at 60 °C for 24h. Both methods worked equally well and showed no influence on the result. All dried tissue samples were mortared and stored at -20 °C until further analysis. Due to their low absolute lipid content with limited variations (Cherifi et al., 2018), lipids were not extracted from the tissue samples prior to SIRA (Deudero et al., 2009). Although compared to the defatted fraction a depletion in $\delta^{13}C$ by -7.4‰ was analysed in the lipid fraction, the overall seasonal variation of lipids comprising 3% of total dry matter (10 -13%) will just account for a maximum variation in $\delta^{13}C$ of 0.2‰ in the composite samples. Thus, omitting the lipid correction had no significant effect on the variability of carbon results, but provides a simpler method for application in practice.

For trace element analysis, mussel shells (n=100) were processed following a methodology similar to the one previously described by (Ricardo et al., 2017b; Bennion et al., 2019; Morrison et al., 2019; Bennion et al., 2021). Briefly, all laboratory equipment used for sample preparation was cleaned using a 2-5% solution of DECON 90 (Decon™, King Of Prussia, PA, USA) for 2h, washed with running water, soaked in 10% of HNO_3 overnight and rinsed with MilliQ (Millipore, Burlington, MA, USA) water. The valves were separated, and the organic tissue was removed using blades and tweezers. In order to remove any remaining organic matter and the periostracum, one valve per sample was soaked in 20ml high-purity H_2O_2 (30% w/v) (Thermo Fisher Scientific, Waltham, MA, USA) overnight and rinsed in MilliQ (Millipore, Burlington, MA, USA) water three times.

2.2. Stable isotope ratio analysis and calibration

For sequentially δ13C and δ15N measurements, aliquots of 1.08 mg of dried tissue were weighed into tin capsules, using three replicates per individuum to check the repeatability of measurements as well as the homogeneity of samples. The wrapped samples were transferred into a Flash EA 1112 elemental analyser using an AS200

autosampler (Thermo Fisher Scientific, Waltham, MA, USA). The EA was equipped with a first reactor packed with chromium (III) oxide and silvered cobaltous oxide, where C and N were oxidized, supported by an oxygen pulse for 3 s, to CO_2 and NOx at 1020 °C. Subsequently, the gases were transferred into a second reactor packed with copper wire, where NO_x was reduced to N_2 at 680 °C, followed by a water trap packed with magnesium perchlorate. The remaining gases N_2 and CO_2 were separated on a packed GC column at 45 °C at a continuous flow rate of 90 ml min^{-1} and transferred to a Delta Plus XL isotope-ratio mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) via a ConFlo III interface (Thermo Fisher Scientific, Waltham, MA, USA). The isotope ratios are given in δ-notation expressed in ‰ referring to VPDB (Vienna Pee Dee Belemnite) and AIR, which define the zero point of the scale for $\delta^{13}C$ and $\delta^{15}N$, respectively, and calculated according to the following equation using the software Isodat 2.0 (Thermo Fisher Scientific, Waltham, MA, USA):

$$\delta X\ [‰] = [(R_{sample}/R_{standard}) - 1]$$

where X is ^{13}C or ^{15}N and R is the corresponding $^{13}C/^{12}C$ or $^{15}N/^{14}N$ ratio.

The laboratory working standards, sucrose ($\delta^{13}C$) and urea ($\delta^{15}N$) (Merck, Darmstadt, Germany), were calibrated by the following international secondary standards: NBS 22 ($\delta^{13}C_{VPDB}$ = -30.031‰), IAEA-CH-3 ($\delta^{13}C_{VPDB}$ = -24.724‰), and IAEA-CH-6 ($\delta^{13}C_{VPDB}$ = -10.449‰) for carbon and IAEA-N1 ($\delta^{15}N_{Air}$ = +0.4‰) and IAEA-N2 ($\delta^{15}N_{Air}$ = +20.3‰) for nitrogen. The working standards were analysed in each sequence to validate the repeatability and to calibrate the monitoring gases carbon dioxide and nitrogen (Air Liquide, Paris, France). The standard deviation of consecutive monitoring gas analyses (n=9) was ≤0.05‰, while triplicate sample measurements yielded a standard deviation of ≤0.15‰ for $\delta^{13}C$ and $\delta^{15}N$ (median 0.04‰, respectively).

2.3. Digestion and trace element analysis

From each shell, two fragments of approximately 10 mm2 were selected and ground together: one fragment from the most recently formed edge and the other fragment located halfway between the umbo and the edge. Fragments were powdered using a

ball mill MM 200 (Retsch, Hahn, Germany) with an agate grinding jar and digested in a close microwave device Mars 5 (CEM, Vertex, Barcelona, Spain) equipped with Teflon™ vessels and temperature controllers. The optimized microwave-assisted digestion procedure was carried out to achieve a shorter digestion time and dilution factor. Samples (around 200 mg) were directly placed into Teflon™ vessels, 10 ml of a concentrated HNO_3 (65%) (Agilent Technologies, Santa Clara, CA, USA) was added to each flask and kept for 10 min at room temperature. Subsequently, the samples were digested following a one-stage digestion program (180 °C, 15 min). After digestion, the vessels were left to cool down and the resulting solution was diluted by the addition of ultrapure water obtained from Milli-Q Direct Water Purification System to reach a final volume of 50 ml.

The quantitative analysis of 16 trace elements (^{11}B, ^{27}Al, ^{47}Ti, ^{51}V, ^{52}Cr, ^{55}Mn, ^{56}Fe, ^{59}Co, ^{60}Ni, ^{63}Cu, ^{66}Zn, ^{75}As, ^{78}Se, ^{111}Cd, ^{137}Ba, ^{208}Pb) was performed using an ICP-MS (7700x, Agilent Technologies, Santa Clara, CA, USA) with a MicroMist micro-uptake glass concentric nebulizer (Glass Expansion, West Melbourne, Victoria, Australia). In order to reduce MO^+ formation in the plasma, the spray chamber was Peltier cooled at 2 °C. A standard quartz torch with 2.5 mm internal diameter injector was used. Finally, standard nickel cones (sample and skimmer) were used. The optimization of the ICP-MS conditions was achieved by adjusting the torch position and tuning for reducing oxide and doubly charged ion formation with a standard tuning solution containing 1.0 µg l^{-1} of ^{7}Li, ^{24}Mg, ^{59}Co, ^{89}Y, ^{140}Ce and ^{205}Tl in 1.0% HNO_3. This equipment includes a collision cell (He gas, ORS3 system, Agilent Technologies, Santa Clara, CA, USA) to discriminate spectral interferences for all the trace metals considered here with high performance. Alongside the samples, the European Reference Material ERM®-CE278K (powdered *M. edulis* mussel tissue) certified under the responsibility of the Institute for Reference Materials and Measurements (IRMM) of the European Commission's Joint Research Centre was analysed. Before data analysis, elements which were not successfully recovered in the Certified Reference Materials were removed from all analyses (^{56}Fe and ^{78}Se were removed).

2.4. Statistical analysis

Basic statistical treatments of SIRA data were performed using JMP 15.1 (SAS Institute, USA). One-way ANOVA with post-hoc Tukey's HSD test was used to identify significant differences between the mean stable isotope values of individuals from different locations and on ocean area scale.

Trace element statistical analyses were performed using R software version 4.0.4 (R Core Team, 2021) and the statistical significance was determined at the alpha level of 0.05. Normality and homoscedasticity of each element individually were tested using Shapiro-Wilks (Strasser et al., 2008) and Fligner-Killeen tests ("*stats*" package) respectively. Since trace element data did not meet parametric assumptions, Kruskal-Wallis test ("*stats*" package) was used to determine whether individual elemental concentrations varied among harvesting sites. When significant differences were found, post hoc pairwise comparisons were applied to identify the source of differences between harvest sites using Pairwise Wilcoxon Rank Sum Test ("*stats*" package). Statistical significance was determined based on adjusted p-values after the Benjamini-Hochberg correction (Benjamini and Hochberg, 1995).

To test whether isotope and element fingerprinting could be used to successfully assign samples to their harvesting locations, the machine learning RF classification method was used. RF uses a bootstrap aggregation algorithm to construct a multitude of decision trees (from randomly selected subset of training data) and combines the output of all the trees to obtain the most accurate prediction. This method was chosen as it has proven to have considerable advantages over many other multivariate classification analysis (Breiman, 2001) , as it reduces overfitting problems in decision trees and also reduces the variance, resulting in an improvement of the accuracy. Additionally, RF classification does not require the strict multivariate distributional and normality assumptions (Breiman, 2001) and performs best in the presence of skewness (Jones et al., 2017). In order to increase the number of individuals per location, data from sampling locations within the same region were combined and treated as a single group. RF was implemented using the R package "randomForest"(Liaw and Wiener, 2002) with the

number of variables available for splitting at each tree node (mtry) limited to the square root of the number of variables in the analysis and using the Gini index (Cutler et al., 2007) to obtain the best elemental combination. This index measures the node impurity of the trees, and shows the probability of a new record being incorrectly classified at a given node in a decision tree based on the training data. So, the higher Mean Decrease in Gini index, the more important the variable is in classifying the data. In all cases, data were randomly split into a training dataset (75%) and a testing dataset (25%), and this procedure was randomly repeated 1000 times to avoid sampling effects. Each time, the rate of classification success (i.e., percentage of correct assignment of mussels to their harvesting region) was calculated, and mean accuracy values were extracted. Cohen´s Kappa (κ) statistic was also calculated, which is a method that accounts for the agreement occurring just by chance. Values of κ range from 0 to 1, where 0 indicates that the RF resulted in no improvement over chance, and 1 indicates perfect agreement (Titus et al., 1984). RF was performed for trace element (n=100) and stable isotope (n=179) data individually, as well as for the combination of both measures coming from the same individual, to increase the number of predictor variables included. The resultant combined dataset was limited by the number of individuals for which both types of data were available (n=64) (**Table 3.1**).

3. Results and discussion

3.1. Stable isotope ratio analysis

3.1.1. Spatial variability

SIRA results of mean δ13C and δ15N values for the individual locations are summarized in **Table 3.2**. The lowest average δ13C was found in the Spanish location DEES (-21.79±0.51‰), while the highest average carbon isotope ratio occurred also in Spain at AGES (-18.11±0.54‰). The highest average δ15N was observed in Chile (COCL) with +11.77±0.60‰, while the lowest average of +4.47±0.40‰ occurred in Portugal (PBPT). The average δ13C of Mediterranean sites (MED: -21.03±0.93‰) is significantly more negative than that of Atlantic sites (ATL: -19.53±1.45‰; $p < 0.001$) and the average

of South East Pacific site (SEP: -19.86±0.87‰) ($p < 0.001$), while no significant differentiation between ATL and SEP could be observed ($p < 0.283$). The mean δ15N means differ significantly between all three marine areas (MED: +7.07±1.74‰, ATL: +5.41±0.92‰, SEP: +11.77±0.60‰; $p < 0.0001$). Similarly, the present data show a great distinction between European and non-European samples, which is mainly driven by the significantly elevated δ15N values of Chilean and Tunisian samples ($p < 0.0001$). Basically, our isotopic data obtained for *M. galloprovincialis* tissue are consistent with literature data except for the elevated δ15N values in individuals from Chile and Tunisia (Page and Lastra, 2003; Deudero et al., 2009; Ezgeta-Balic et al., 2014; Rumolo et al., 2017; Briant et al., 2018; Ozdilek et al., 2019). However, the higher δ15N level in Chilean and Tunisian individuals is in a comparable range with other mussel species from certain regions at the French coast (Dubois et al., 2007b; Briant et al., 2018). Overall, the δ13C values in this study are consistent with marine sources of organic matter (phytoplankton, algae and organic detritus) as the primary diet, being generally lower than those resulting from terrestrial sources (Simenstad and Wissmar, 1985). Similarly, marine specific nitrogen fixation processes were confirmed to be reflected in tissue $\delta^{15}N$ of the analysed mussels. The $\delta^{15}N$ is strongly influenced by the marine phytoplankton community, with N_2-fixing diazotrophs producing low $\delta^{15}N$ values (-2 to 0‰) by incorporating atmospheric N_2 as well as denitrifying organisms assimilating NO_3^- with enriched $\delta^{15}N$.

Table 3.2. $\delta^{13}C$ and $\delta^{15}N$ all-season means ± SD in *M. galloprovincialis* tissue. Within a column, letters identify significant differences ($p < 0.05$) between locations. MED = Mediterranean Sea area; ATL = Atlantic Ocean area; SEP = South Eastern Pacific Ocean area.

Location	Country	Area	$\delta^{13}C$		$\delta^{15}N$	
LBFR	France	ATL	-20.02±0.39	c	+7.05±0.68	c
PBPT	Portugal	ATL	-19.00±0.75	b	+4.47±0.39	e
AGES	Spain	ATL	-18.11±0.54	a	+5.97±0.38	d
SGES	Spain	ATL	-18.57±0.58	ab	+6.16±0.57	d
MEES	Spain	ATL	-21.16±0.74	de	+4.87±0.88	e
MUES	Spain	ATL	-20.81±1.16	d	+5.57±0.94	d
DEES	Spain	MED	-21.79±0.51	e	+5.88±0.58	d
GEIT	Italy	MED	-20.37±0.19	cd	+5.79±0.42	d
LBTN	Tunisia	MED	-21.77±0.52	e	+10.17±0.17	b
COCL	Chile	SEP	-19.86±0.87	c	+11.77±0.60	a

A bivariate plot of $\delta^{13}C$ *vs* $\delta^{15}N$ means and ranges at different locations (all-seasons) is illustrated in **Figure 3.2**. The overall variation of -22.89 to -17.14‰ for $\delta^{13}C$ and +3.49 to +12.50‰ for $\delta^{15}N$ shows a considerable range within just one species, which is, however, in good agreement with literature data of mussels from various regions. $\delta^{13}C$ values in mussel tissue of the species *M. edulis* range from -22.0 to -17.1‰ and from -23.0 to -17.3 for *M. galloprovincialis,* while $\delta^{15}N$ values for these species range from +4.3 to +12.5‰ and +3.2 to +4.9‰, respectively(Dubois et al., 2007a; Dubois et al., 2007b; Riera, 2007; Deudero et al., 2009; Briant et al., 2018). **Figure 3.2** reveals that the origin of *M. galloprovincialis* can be identified just to a limited extent by SIRA. Especially the ranges of several European locations overlap and do not show significant differences in stable isotopic composition (cp. also **Table 3.2**). Hence, it can be concluded that the primary food sources and other environmental factors are largely comparable along the European coasts. Nevertheless, some regional specifics can be identified.

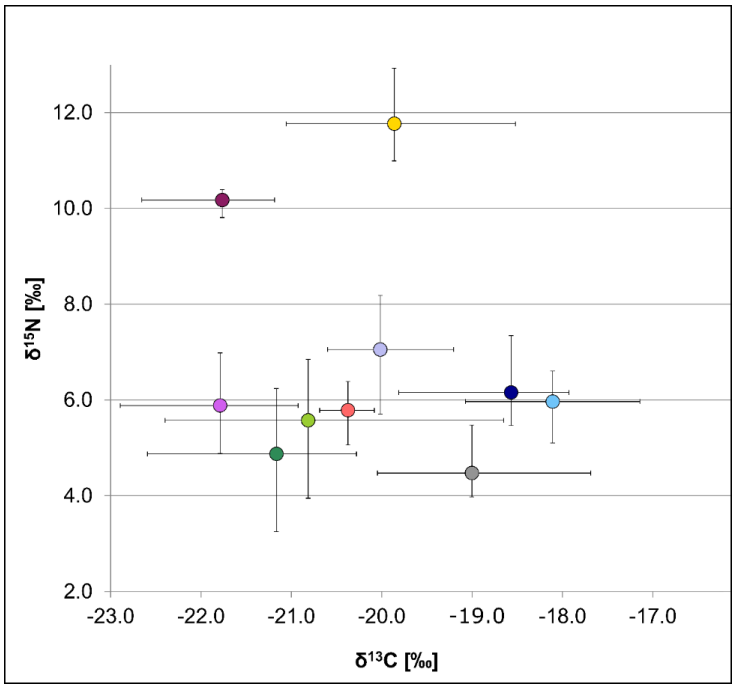


Figure 3.2. Bivariate plot of $\delta^{13}C$ and $\delta^{15}N$ mean values with min/max ranges in *M. galloprovincialis* samples from different locations: Atlantic Ocean area (● LBFR, ● PBPT, Galicia region (● AGES, ● SGES), Basque Country region (● MEES, ● MUES)); the Mediterranean Sea area (● DEES, ● GEIT, ● LBTN); and the South Eastern Pacific area (● COCL). See **Table 3.1** for definitions of abbreviations.

Within the Spanish *M. galloprovincialis* samples, comparable $\delta^{13}C$ and $\delta^{15}N$ values could be detected between the Galician samples AGES and SGES, and the samples MUES and MEES from the Basque Country, suggesting similar nutritional and nutrient conditions in the same region. Thus, based on the regionally consistent isotopic composition, the origin of PDO-protected *M. galloprovincialis* from Galicia can be verified by SIRA from most other regions, especially since unknown or potentially mislabelled mussels of other origins will mostly show different delta values (**Figure 3.2**). However, samples from the Basque region (MUES, MEES) have a wide range of isotopic values, with especially $\delta^{13}C$ overlapping with almost all other European sites, which makes it difficult to assign samples to this origin unequivocally.

The $\delta^{13}C$ values in organic tissue of aquatic organisms are influenced by their primary diet. However, the diet of aquatic organisms can be influenced by anthropogenic pollution such as sewage, which may result in a potential interference in isotope values of mussels form coastal sites close to anthropogenic sources such as harbours or industries (Ozdilek et al., 2019). Hence, the high variations in $\delta^{13}C$ at MUES, MEES (both Spain) and to a lesser extent at PBPT (Portugal) may reflect an anthropogenic overprint, as these sampling locations are nearby harbours, large coastal cities or estuaries. As discussed below, these fluctuations could also include a seasonal variability of the environment including ocean currents.

Comparable to the elevated $\delta^{15}N$ in Chilean mussels found in this study, several authors identified enhanced $\delta^{15}N$ values also in different marine species at the coasts of South America (Carrera and Gallardo, 2017; Espinoza et al., 2017). At the western coast of South America, the productive upwelling system of the Humboldt current causes ^{15}N enriched isotopic values, while beyond that, N_2-fixation by diazotrophs appears to have minor influence (Somes et al., 2010). The significantly elevated $\delta^{15}N$ values in the Tunisian samples are in contrast to an earlier study of mussel species from the Mediterranean Sea (Briant et al., 2018). In this study, the authors identified depleted $\delta^{15}N$ values in bivalve samples from the French coast bordering the Mediterranean Sea compared to individuals from the Atlantic Ocean and English Channel, probably due to the oligotrophic regime in Mediterranean Sea where N_2-fixation occurs (Liénart et al.,

2017; Briant et al., 2018). However, we did not find significantly depleted δ15N levels in individuals from the remaining Mediterranean coasts locations (DEES, GEIT) compared to individuals from locations bordering the Atlantic (PBPT, AGES, SGES, MEES and MUES) or the English Channel (LBFR) (**Figure 3.2**). Accordingly, the elevated δ15N from the Bizerte lagoon in Tunisia (LBTN) may rather be attributed to anthropogenic pollutions. Several studies identified various signs of chemical contamination from agriculture, industry, shipping and urbanisation in the lagoons' sediments, but also in organisms thriving in the Bizerte lagoon (Ben Ameur et al., 2013; Barhoumi et al., 2014; Barhoumi et al., 2016).

3.1.2. Seasonal variability

Seasonal fluctuations in carbon and nitrogen isotopic values of different organisms and food sources have been reported frequently, depending not only on the seasonal climate variability, but also on diet and feeding strategies (McCutchan and Lewis, 2001; Bahar et al., 2008; Molkentin and Giesemann, 2010; Ezgeta-Balic et al., 2014; Zhao et al., 2019). As shown above, an exact geographical differentiation of the mussels is not feasible due to the high variation in $\delta^{13}C$ and $\delta^{15}N$ occurring at individual locations (**Figure 3.2**). However, when considering the stable isotopic data on a seasonal basis, a more accurate discrimination between the different locations was achieved (**Figure 3.3**). There is a considerable improvement in the distinction among the European sampling locations associated with a marginal overlap in the seasonal isotopic ranges. Nevertheless, no distinct trend in isotopic variation over the year could be observed for both elements.

It seems that the seasonal variation in isotopic values at individual locations is more affected by $\delta^{13}C$ than by $\delta^{15}N$. Highest seasonal variation in in $\delta^{13}C$ is observed at MUES, with a difference of 3.75‰ (spring *vs* autumn), while the maximum $\delta^{15}N$ variation, 2.98‰, is found at MEES (winter *vs* autumn). In comparison, seasonal variations in AGES and SGES are rather limited, which is again favourable for the PDO detection in Galicia.

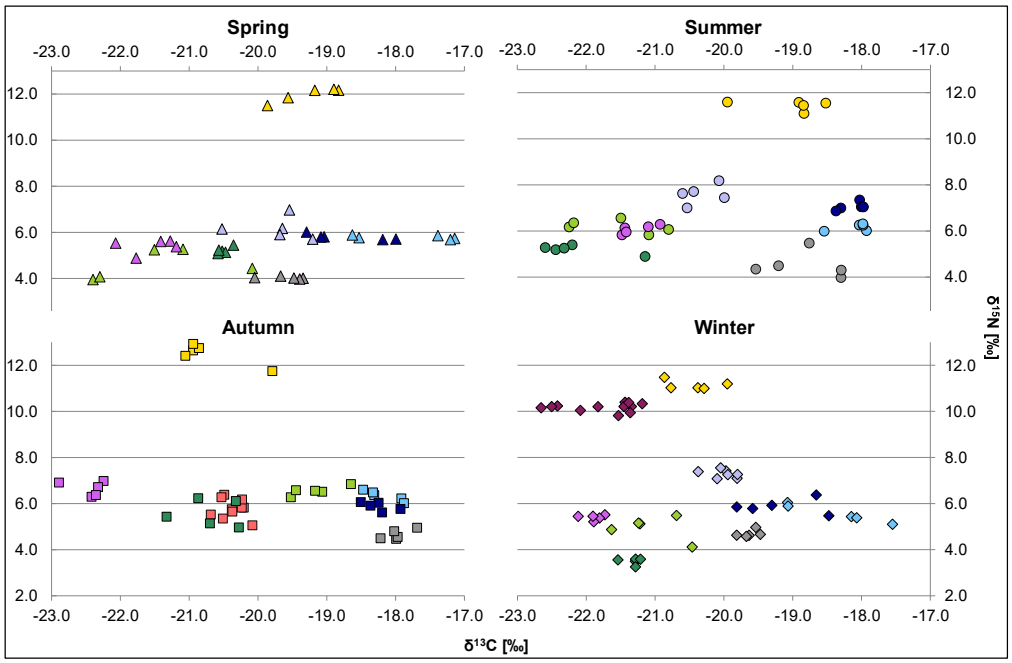


Figure 3.3. Seasonal distribution of individual $\delta^{13}C$ and $\delta^{15}N$ values in *M. galloprovincialis* (n = 179). Each point represents an individual mussel belonging to one of the harvesting locations in the Atlantic Ocean area (● LBFR, ● PBPT, Galicia region (● AGES, ● SGES), Basque Country region (● MEES, ● MUES)); the Mediterranean Sea area (● DEES, ● GEIT, ● LBTN); and the South Eastern Pacific area (● COCL). See **Table 3.1** for definitions of abbreviations.

Several studies have detected significant seasonal differences in stable isotope composition of aquatic organisms due to seasonal food preferences, food availability or climatic factors (Sant'Ana et al., 2010; Chaguri et al., 2015; Zhao et al., 2019). Potential environmental factors influencing the isotopic mussel composition may be seen at the Ebro delta (DEES). Samples from this location are highly influenced by the seasonal variations of the Ebro river discharge and varying water conditions over the year, which in turn varies depending on the climate (Ibañez et al., 1997). Additionally, seasonal variations of the thriving phytoplankton community within the Ebro delta may be reflected in the mussels' stable isotopic composition. Phytoplankton biomarkers from October were reported to be more depleted in $\delta^{13}C$ in comparison to samples from June (Gómez-Gutiérrez et al., 2011), which is in accordance with the lowest $\delta^{13}C$ values found in DEES individuals of *M. galloprovincialis* from autumn.

The Chilean coast is strongly affected by seasonal disturbances within the Humboldt Current (Thiel et al., 2007). In contrast to the northern Chilean coast, where there is lower productivity associated with more stable conditions, a highly productive seasonal upwelling occurs in south-central Chile, where our samples were collected from (Thomas et al., 2001; Montecino and Lange, 2009). Both $\delta^{13}C$ and $\delta^{15}N$ values from COCL are subjected to fluctuations throughout the year, being more pronounced in carbon isotopes (**Figure 3.3**). Since the Humboldt Current is weaker in the Chilean summer, which corresponds to the Northern Hemisphere winter description in **Figure 3.3**, $\delta^{15}N$ of the mussels is correspondingly lower then, due to the lack of the Humboldt increasing effect (Espinoza et al., 2017). Inter-annual scale events as the "El Niño" in the Southern Hemisphere have significant impact of the local nitrogen and carbon cycle, which in turn have an impact on the local phytoplankton and zooplankton community, which is the primary food of mussels (Montecino and Lange, 2009). Since the mussels used in this study were harvested between Northern Hemisphere winter 2018/19 and autumn 2019, and a weak "El Niño" occurred during the same period (NOAA, 2021), this sample set may not be fully representative neither for non-"El Niño" years, nor for strong "El Niño" years.

In summary, the SIRA results show that the isotopic composition of mussels of the species *M. galloprovincialis* from different locations is highly variable. However, the application of isotopic data in source authentication is limited by their seasonal variation. The pronounced seasonal variation in stable isotopic values within individual sampling locations results in overlapping ranges, making an accurate determination of origin challenging. Therefore, the application of stable isotope analysis to samples of unknown origin allows the exclusion of individual locations rather than the exact determination of a specific origin. However, knowledge of the sampling season can improve the resolution, thus allowing a more accurate designation of the origin to be made.

3.2. Trace element analysis

Concentrations of individual elements detected in *M. galloprovincialis* shells varied differently between harvest locations when analysing them individually (**Table S3.2**). Results show that among the fourteen elements considered in the final analysis, Ba, Mn, Pb and Al (Kruskal Wallis Test; p-value(Ba)= $1.471e^{-12}$, p-value(Mn)= $1.544e^{-10}$, p-value(Pb)= $7.373e^{-13}$, p-value(Al)= $2.862e^{-07}$) were the trace elements that varied most between sites, meaning these elements were the main drivers of the observed differentiation. Conversely, no significant differences were depicted for Ti, Co and As (Kruskal Wallis Test; p-value $p>0.05$) nor for Cr and Cu concentrations when performing pairwise comparisons between harvesting sites (Pairwise Wilcoxon Rank Sum Tests, $p>0.05$) (**Table S3.2**). The incorporation of these elements into the calcite structures like mussel shells can be influenced by a number of interacting factors, such as elemental concentrations in seawater, metabolic efficiency of the individual, or environmental conditions (temperature, salinity, pH and dissolved oxygen among others) (Klein et al., 1996; Poulain et al., 2015; Zhao et al., 2020). For example, higher concentrations of Ba and Mn have previously been related with terrestrial runoff and it has been seen that their uptake could be affected by temperature (influencing Ba) and dissolved oxygen (affecting both Ba and Mn) (Carson, 2010; Broadaway and Hannigan, 2012), whereas heavy metals like Pb and Al have been associated to anthropogenic impacts (Carson et al., 2013). Although the understanding of factors contributing to the variability of elements can be helpful, the combination of these factors affecting each of the coastal ecosystems generates a unique pattern that might allow an accurate identification of the geographic origin, which is the focus of the present study.

3.3. Origin assignment: a combined approach

For the joint evaluation of SIRA and TEF data, locations within the same region were grouped together into BQES for Basque Country (in the case of MEES and MUES) or GAES for Galicia (in the case of AGES and SGES) regions (**Table 3.1**). The random forest model showed different classification success assigning individuals to their harvest region depending on the variables analysed. When the model was run based on trace element

fingerprints of shells, mussels were assigned to their local harvest origin with an overall accuracy of 87.8% (κ=0.93) (**Table S3.1 and Figure 3.4B**). According to the Gini index, the most important elements for discriminating individuals between regions were Pb, Ba, Mn, and Al, which was expected as they were the elements that differed the most between origins in the statistical analysis. Shell elemental analyses have previously shown to be useful to assign bivalve species to their geographical origin using different classification approaches. Using discriminant function analysis (DFA), Becker et al. (2005) correctly assigned 56% of *M. californianus* and *M. galloprovincialis* mussels to six sampling points located in Southern California (reaching 90% of success when grouping samples in two regions) and Dunphy et al. (2011) correctly assigned 69% of *P. canaliculus* individuals to six sampling points in northern New Zealand (reaching 63-100% when grouping samples into two regions). Performing an analysis of similarity, Ricardo et al. (2015) obtained 92% of classification success assigning *C. edule* individuals to six sampling points in Portugal. Recently, by applying the machine learning RF classification, Bennion et al. (2019) correctly assigned 90% of *M. edulis* individuals to four sampling points in the west coast of Ireland and Morrison et al. (2019) achieved 97.5% of success assigning *P. maximus* to three sampling points in Ireland. These two last studies reached 100% of correctly assigned individuals to their origin when analysing elemental signature of both, tissue and shells. On this basis, in this present study, we have proved the potential of TEF of shells at higher spatial scales, correctly classifying 87.8% of *M. galloprovincialis* individuals to 8 different regions located in 6 different European countries. Using the stable isotopic composition of the tissue alone, results showed an overall classification success of 80.5% (κ=0.81) to their harvest region (**Table S3.1 and Figure 3.4A**). Mussels from LBTN and COCL show the highest classification accuracy, which is consistent with the pattern observed in the bivariate plot of $\delta^{13}C$ and $\delta^{15}N$ (**Figure 3.2**).

To increase the likelihood of classification success, a combined dataset of trace element concentrations of shells and stable isotope ratios of muscle was used. The classification success of the RF analysis ranged from 76% to 100% (κ =0.98) depending on the sampling region. Individuals collected from LBFR (France), BQES (Basque Country,

Spain), GAES (Galicia, Spain), GEIT (Italy) and LBTN (Tunisia) displayed the highest percentages of correct classification (>96%) (**Table 3.4 and Figure 3.4C)**. The most important predictor variables included both stable isotope ratios ($\delta^{13}C$ and $\delta^{15}N$) as well as the previously mentioned four trace elements (Pb, Ba, Mn and Al). Indeed, the combination of these six variables proved to be the most successful approach, obtaining an overall classification accuracy of 97% (κ =0.98) This means that the number of variables analysed in future studies could be potentially reduced, which may result in a lower cost of the analysis and a reduced computational time.

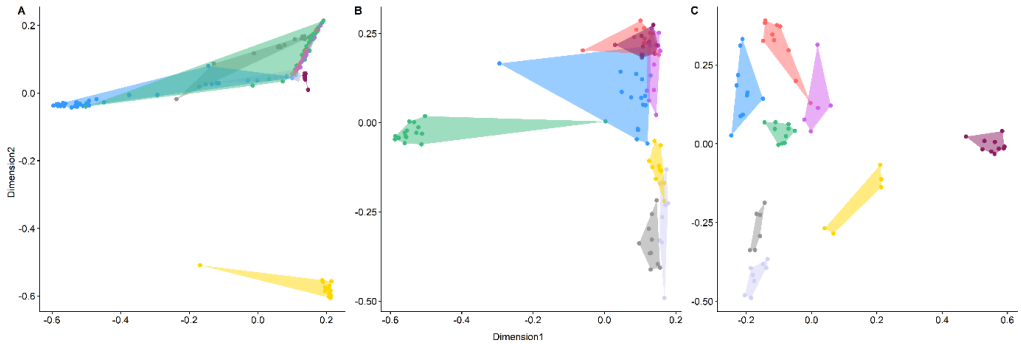


Figure 3.4. Multidimensional scaling (MDS) of proximity scores from each of the RF classifications based on (A) $\delta^{13}C$ and $\delta^{15}N$ values of soft tissues (n=179), (B) fourteen trace element concentrations of shells (n=100) and (C) the combination of both chemical signatures (n=64). Each point represents an individual *M. galloprovincialis* mussel belonging to one of the harvest locations in the Atlantic Ocean area (● LBFR, ● PBPT, Galicia region (● GAES), Basque Country region (● BQES)); the Mediterranean Sea area (● DEES, ● GEIT, ● LBTN); and the South Eastern Pacific area (● COCL). See **Table 3.1** for definitions of abbreviations.

The origin traceability tool presented here has proven to be useful to trace the origin of farmed Mediterranean mussels at the broadest geographical scale investigated to date, approaching a real usage scenario. Using chemical signatures for a species with numerous harvest areas like *M. galloprovincialis* can be a challenge, as the higher the number of harvest locations, the increased the likelihood of overlapping geochemical fingerprints. Nevertheless, the combination of stable isotopes and trace element fingerprints has demonstrated to be a successful approach, correctly assigning 97% of the individuals to their harvest region based on RF model.

Table 3.4. Random forest classification success rates (%) in assigning farmed *M. galloprovincialis* mussels to their harvest region based on a combination of all trace element and stable isotope ratios. Assignment results using only the most important predictor variables for the assignment ($\delta^{13}C$, $\delta^{15}N$, Pb, Ba, Mn, and Al) are shown in brackets. See **Table 3.1** for definitions of abbreviations.

		ATL				MED			SEP
		LBFR	**PBPT**	**GAES**	**BQES**	**DEES**	**GEIT**	**LBTN**	**COCL**
ATL	**LBFR**	**99** (100)	8 (2)	0	0	0	0	0	6 (2)
	PBPT	1 (0)	**92** (98)	0	0	0	0	0	0
	GAES	0	0	**96** (95)	0	0	0	0	0
	BQES	0	0	0	**97** (99)	2 (2)	0	0	0
MED	**DEES**	0	0	0	3 (1)	**76** (84)	0 (1)	0	0
	GEIT	0	0	4 (5)	0	22 (13)	**100** (99)	0	0
	LBTN	0	0	0	0	1 (1)	0	**100** (100)	2 (2)
SEP	**COCL**	0	0	0	0	0	0	0	**92** (95)

4. Conclusions

This study reveals that the combination of stable isotope ratio analysis of tissue and trace element analysis of shells is an effective technique for determining the geographical origin of *M. galloprovincialis* mussels cultivated in ten different farms covering a large part of the geographical distribution. The separate use of each individual approach does not appear to be sufficient to fully discriminate the provenance, although certain locations can be excluded with certainty. This is not, however, the case of the PDO-protected *M. galloprovincialis* from Galicia, which could be well verified using SIRA alone. Overall, the present study reinforces the potential of using chemical signatures to trace seafood to their geographic origin. However, before this approach can be applied as a traceability tool by food control authorities, a reference collection of microchemical signatures should be first developed, which may require an intensive fieldwork at both temporal and spatial scales (Ricardo et al., 2017b; Bennion et al., 2019). Besides, this methodology must be tested with a larger number of samples within each region and from additional farming locations, such as the United Kingdom, Ireland, and the eastern Mediterranean area, where this species is also actively cultivated. Our research paves the way for a reliable traceability tool development that will enforce food safety regulations and prevent commercial fraud.

CHAPTER 4

Gut microbiota fingerprinting as a potential tool for tracing the geographical origin of farmed mussels (*Mytilus galloprovincialis*).

Under review as:

del Rio-Lavín, A., Monchy, S., Jiménez, E., and Pardo, M. A. (2022). Gut microbiota fingerprinting as a potential tool for tracing the geographical origin of farmed mussels (*Mytilus galloprovincialis*). *Frontiers in Microbiology*.

All the Supplementary Information of this manuscript can be found in **Appendix C**.

Laburpena

Arrain eta itsaskien jatorri geografikoa bermatzea berebiziko garrantzia du, iruzurra saihestea, itsas baliabideak babestea eta kontsumitzaileen segurtasuna mantentzea helburu duten mundu osoko kontrol agintari eta itsas-industriarentzat. Hortaz, kapitulu honetan, *M. galloprovincialis* muskuilu freskoen digestio-guruinean dagoen bakterio-komunitatearen profila aztertu da, hazkuntza jatorria bereizteko trazabilitate tresna gisa erabil ote zitekeen ikertzeko. Horretarako, 2019an, Espainiako hiru eskualde ezberdinetan (Galizian, Euskal Autonomia Erkidegoan eta Katalunian) kokatutako bost akuikultura-instalazio ezberdinetatik laginak bildu ziren urtaroka, eta haien bakterio-komunitatearen konposizioa aztertu zen 16S rRNA genearen V3-V4 gune aldakorraren anplifikazio eta sekuentziazioaren bitartez. Emaitzek bakterio-komunitateen profilak jatorriaren eta urtaroaren arabera nabarmen aldatzen direla erakutsi dute, jatorri geografikoaren eragina urtaroen aldakortasuna gainditzen duelarik. Trazabilitate tresna honen egonkortasuna eta potentziala aztertzeko, 2020ko udazkenean EAEn bildutako 20 indibiduo berriren bakterio profila (laginak era idependentean prozesatu, sekuentziatu eta analizatu ostean lortua), 2019an lortutako profilekin alderatu zen. Emaitzek, EAEn bi urtez jarraian bildutako muskuiluak elkarrekin biltzen direla erakutsi dute, baita uzta-denboraldiarekin bat datozela ere. Lan honetan ikertutako tresna analitikoa erabiliz muskuilu freskoen jatorri geografikoa egiaztatzea posiblea dela frogatu da; eta hortaz, elikagai hauen trazabilitatea eta segurtasuna bermatzeko erabili liteke.

1. Introduction

Fisheries and aquaculture production is at record high and provides nutritious food and employment options to many countries around the world. Global production of aquatic animals reached 214 million tonnes in 2020, valued in 406 billion USD (FAO, 2022b). As the demand for fish and fish products increases, so does the awareness of the need to ensure transparency and traceability along the food chain. In the EU, the Council Regulation (EC) Nº 1379/2013 on the common organization of the markets in fishery and aquaculture products, requires that seafood labelling indicates the commercial designation and scientific name of the species, the production method, and the area where the product was caught or farmed (EU, 2013b). Thus, tracing the geographic origin of seafood represents a major goal for control authorities to address current legislation, but also to prevent commercial fraud and to ensure sustainable fisheries and aquaculture management. Traceability is also essential to protect public health by preventing hazardous products from reaching the marketplace. Consuming seafood, and specially shellfish, from restricted areas which may be contaminated with biotoxins, bacterial pathogens or chemical pollutants, poses a potential health risk for consumers. Likewise, being able to guarantee the provenance is also relevant for producers who aim at certifying their products and promote consumers' confidence. In this context, a global need for suitable geographic origin traceability analytical tools has emerged.

Several experimental approaches have been proposed to validate the geographic origin of seafood, which include techniques based on DNA markers, fatty acids, trace elemental profiling and stable isotope analysis (Leal et al., 2015; El Sheikha and Montet, 2016; Gopi et al., 2019a). Another methodology that has also been considered is the analysis of the microbial communities associated to seafood and its linkage to a particular geographic area. In this regard, the potential of analysing the bacterial composition by 16S rRNA targeted NGS has been explored to trace the geographic origin of cultured seabass (*Dicentrarchus labras*) (Pimentel et al., 2017), Manila clams (*Ruditapes philippinarum*) (Milan et al., 2019) and soft-shell clams (*Mya arenaria*) (Liu et al., 2020).

In this study, we wanted to assess the applicability of this bacterial profiling approach to differentiate the origin of the highly commercialized *M. galloprovincialis* mussel. In the EU, mussels represent a major aquaculture species accounting for 34% of the total production. Spain is the main producer of *M. galloprovincialis* in the EU, reaching 247,897 tonnes in 2020 for a market value of 147 million euro. More than 97% of the Spanish mussels' production is carried out in Galicia, a region located in the north-western coast of the Iberian Peninsula, and whose mussels fall within the domain of PDO "Mexillón de Galicia". Thus, this geographic area represents the perfect scenario to evaluate the suitability of using bacterial communities present in mussels' digestive gland to validate their origin. Indeed, the microbiota of filter-feeding bivalves has been shown to be influenced by the geographic location, but also by season, temperature, salinity and other environmental conditions (King et al., 2012; Lokmer and Wegner, 2015; Pierce et al., 2015; Lokmer et al., 2016; Li et al., 2019; Pierce and Evan Warda, 2019).

Therefore, the present study aimed to assess whether the geographic origin influences the bacterial composition present in the digestive gland of *M. galloprovincialis* mussels in such way that could be used as a traceability approach. First, the effect of geographical origin and seasonal variations has been examined by analysing the digestive gland microbiota of mussels collected seasonally from five different farms in Spain using 16S rRNA targeted NGS. Then, the potential and stability of this traceability approach has been evaluated by attempting to trace the geographical origin of new samples collected a year after, which had been independently processed and analysed.

2. Materials and methods

2.1. Sampling and DNA extraction

Mussel samples were collected seasonally from five different farms in Spain during 2019: Mutriku (MUES; Basque country; rafts within a small port), Mendexa (MEES; Basque country; offshore longlines), Ría de Arousa (AGES; Galicia; rafts in estuarine inlets), Ría de Betanzos-Sada (SGES; Galicia; rafts in estuarine inlets) and Delta Ebro (DEES; Catalonia; offshore longlines) (**Figure 4.1**). A total of 160 *M. galloprovincialis* adult

mussels (n=8 per location/season) were collected and immediately transported to the laboratory. To further evaluate the stability of the bacterial fingerprint, 20 additional mussel samples were collected from the Basque country farms a year later (September 2020). In the laboratory, mussels were scrubbed to remove epibionts and gently washed with artificial seawater to remove part of the non-resident microbiota. For each mussel, the digestive gland (DG) was removed under sterile conditions and immediately stored at −80 °C. Bacterial DNA extraction was performed using the DNeasy PowerSoil Kit (Qiagen, Hilden, Germany), starting from 50 mg of the DG and following manufacturer's instructions. Extracted DNA concentration was determined by means of the Quant-iT dsDNA HS assay kit using a Qubit® 1.0 Fluorometer (Life Technologies, Carlsbad, CA, USA).

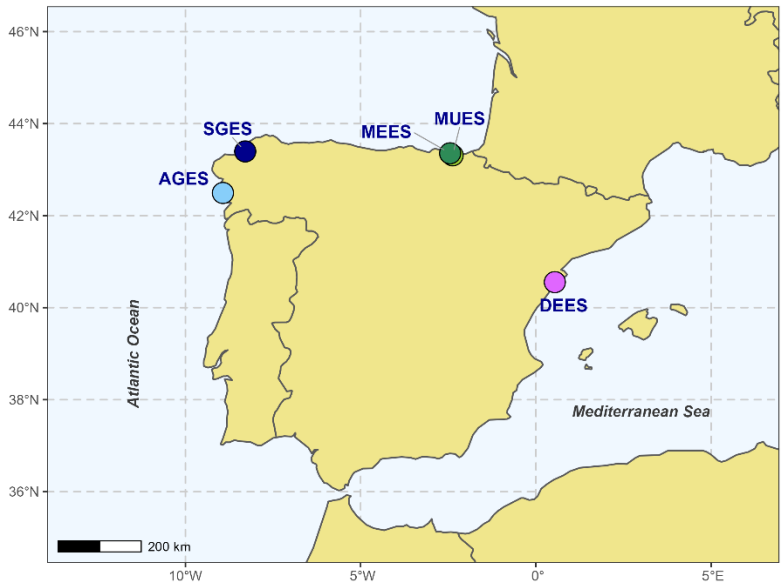


Figure 4.1. Map showing the geographic location where the samples used for this study were collected. Samples were collected from three different regions: Galicia [Ría de Arousa (● AGES) and Ría de Betanzos-Sada (● SGES)], Catalonia [Delta Ebro (● DEES)] and Basque Country [Mendexa (● MEES) and Mutriku (● MUES)].

2.2. Library preparation and Next Generation Sequencing

The 16S amplicon library was prepared according to a standardized Illumina protocol (Illumina, 2013). Briefly, the V3-V4 region of bacterial 16S rRNA gene was amplified by

using universal primers 341F (5'- CCTACGGGNGGCWGCAG-3') and 805R (5'-GACTACHVGGGTATCTAAT CC-3') (Klindworth et al., 2013) tagged with 33 bp overhang adapters. These PCR reactions (15 µl) contained an initial concentration of 0.2 µM of each primer, 7.5 µl of Phusion High-Fidelity PCR Master Mix with HF Buffer (1X) (Thermo Fisher Scientific, Waltham, MA, USA) and 20 ng of template DNA. Triplicate PCR reactions were pooled and purified using AMPure XP beads (Beckman Coulter, Brea, CA, USA) according to the manufacturer's instructions. Secondary PCR reactions were performed in order to add Nextera XT Illumina sequencing indices and adapters. After purification, 16S amplicon libraries were pooled together at equimolar concentration and subjected to Illumina MiSeq paired-end sequencing (Illumina Inc., San Diego, CA, USA). The 16S amplicon library of mussel samples collected in 2020 was prepared and sequenced in a different batch. All sequencing data has been submitted to the NCBI sequence read archive database (SRA accession: PRJNA899701).

2.3. Sequence processing

Sequences were analysed using Mothur software v.1.44.3 (Schloss et al., 2009) following the standard operating procedure (http://www.mothur.org/wiki/MiSeq_SOP). Sequences were quality filtered, de-replicated to unique sequences and aligned against the SILVA database (Quast et al., 2013). Suspected chimeras were removed using Uchime software (Edgar et al., 2011). Resulting sequences (median length of 460 bp) were clustered into Operational Taxonomical Unit (OTU) using 97% similarity threshold. For taxonomic affiliation, OTU sequences were searched against the latest available SILVA database (Release 138.1) using the Wang approach (Wang et al., 2007) as implemented in Mothur. OTUs corresponding to chloroplasts were removed from the data set, and samples were rarefied (by random picking, based on the sample having the lowest read number) for comparison. Sequences of mussel samples collected in 2020 were processed independently following the same procedure.

2.4. Bioinformatics and statistical analyses

Alpha diversity estimators (observed OTUs, Shannon, Simpson and Berger Parker) were calculated using the Past 3.26 software (Hammer et al., 2001) for each sample. All

other statistical analyses were performed in R (R Core Team, 2021). Differences in alpha diversity patterns between sampling points and seasons were analysed using ANOVA and Tukey's honest significance test of ANOVA (pairwise comparisons) for normally distributed metrics and Kruskal-Wallis and Wilcoxon rank sum test (pairwise comparisons) for non-parametric metrics ("stats" package). Normality was tested using Shapiro-Wilk test.

To assess beta diversity, sample-specific read abundances of OTU were subjected to Hellinger-transformation using the function 'decostand' (method = 'hellinger', which consist of dividing the total reads for each OTU by the total number of reads in the corresponding sample and taking the square root of the quotient). All subsequent analysis were based only on standardized data. The variation in microbiota composition between samples based on Bray-Curtis dissimilarities (at OTU level) was visualized by hierarchical cluster analysis using 'hclust' function ("stats" package) and by non-metric multidimensional scaling (NMDS) using 'metamds' function ("vegan" package). Differences in beta diversity between sampling points and seasons were tested using a permutational multivariate analysis of variance (PERMANOVA) using 'adonis' function ("vegan" package). Statistical significance was determined based on adjusted p-values after the Benjamini-Hochberg correction (Benjamini & Hochberg, 1995) and pairwise comparisons were performed using 'pairwise.adonis' function ("vegan" package).

For taxon distribution analysis, relative abundances were calculated at each taxonomic rank for all datasets. Individuals were grouped by location and season, and the top 20 taxa of each subset visualized using the R "ggplot2" package (Wickham 2009). Based on the relative abundance, taxa significantly different between harvesting regions was determined by Kruskal-Wallis and Wilcoxon rank sum test (pairwise comparisons) with Benjamini-Hochberg FDR correction of the p-value. Finally, to evaluate the stability and the potential of this approach to trace the origin of new samples, a taxonomy based hierarchical cluster analysis was performed combining 2019 and 2020 bacterial profiling datasets.

3. Results

3.1. Gut microbial community structure of mussels from different geographic origins and harvesting seasons

Overall, a total of 7,366,326 bacterial sequences were obtained by Illumina MiSeq high-throughput sequencing targeting the V3-V4 region of the 16S rRNA gene, with an average of 43,515 reads per mussel gut sample. Each sample was rarefied to 17,057 bacterial reads, resulting in a total of 49,270 OTUs (21,374 OTUs with counts ≥2), with an average of 730, 932, 743, 503 and 376 OTUs for locations AGES, SGES, MUES, MEES and DEES respectively.

Alpha diversity analysis of gut microbiota revealed that richness, diversity (Shannon and Simpson) and dominance (Berger Parker) estimators were significantly (Kruskal-Wallis, $p<0.05$) influenced by the sampling location of mussels and the seasonality (**Figure 4.2** and **Supplementary material 4.1**). Higher bacterial richness and diversity, and lower dominance were observed in mussel farms located in estuarine inlets (AGES and SGES) and ports (MUES), compared to mussel farms located in offshore longlines (MUES, DEES), with significant differences between those inshore and offshore locations (Wilcoxon rank sum test, $p<0.05$). Regarding seasonality, winter period displayed the highest differences in mussel gut microbial alpha diversity between farming locations, whereas autumn had the lowest.

Regarding beta diversity, hierarchical clustering dendrograms and NMDS ordination based on Bray-Curtis dissimilarity matrix revealed that gut bacterial sequences were grouped first by mussel origin and then by season (**Figure 4.3**). PERMANOVA results confirmed this observation, with farm geographical origin (by site, five levels) ($p=0.000999$, $R2=0.144$, $F=7.982$) and harvesting season ($p=0.000999$, $R2=0.061$, $F=4.505$), as well as the interaction between these two factors ($p=0.000999$, $R2=0.171$, $F=3.143$), contributing significantly to the variation observed among the microbial communities.

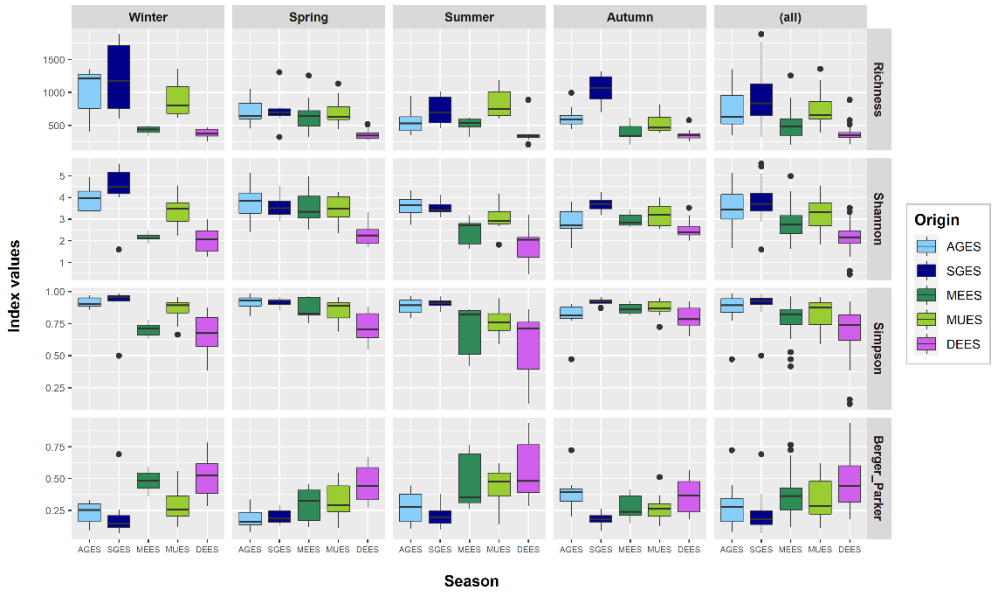


Figure 4.2. Alpha diversity of *M. galloprovincialis* DG microbiota. The box plot shows mean values and standard deviation of the richness, diversity (Simpson and Shannon) and dominance (Berger-Parker) estimators for bacterial communities according to the season and harvest locations: in Galician region (● AGES, ● SGES), Catalonia region (● DEES) and Basque Country region (● MEES, ● MUES).

The geographic origin effect was also significant across all four seasons ($p<0.05$) when each season was analysed individually. Results showed a clear separation between the microbial community structure of mussels collected from Galician (AGES, SGES) and Catalonian farms (DEES). However, Basque Country samples (MEES, MUES) grouped distinctly depending on the season. The microbiota of mussels collected from the Basque Country region in autumn were closer to the Catalonian fingerprint, whereas winter and spring Basque country microbial structures were closer to those from Galicia. Significant differences were still depicted when performing the statistical analysis at a regional scale (Galicia, Catalonia and Basque Country) instead of by farm ($p=0.000999$). Regarding seasonality, multiple pairwise comparisons between seasons performed separately for each sampling site, showed that the DG bacterial communities differed significantly between all four seasons (pairwise PERMANOVA, $p<0.001$) when the five mussel farms were analysed individually.

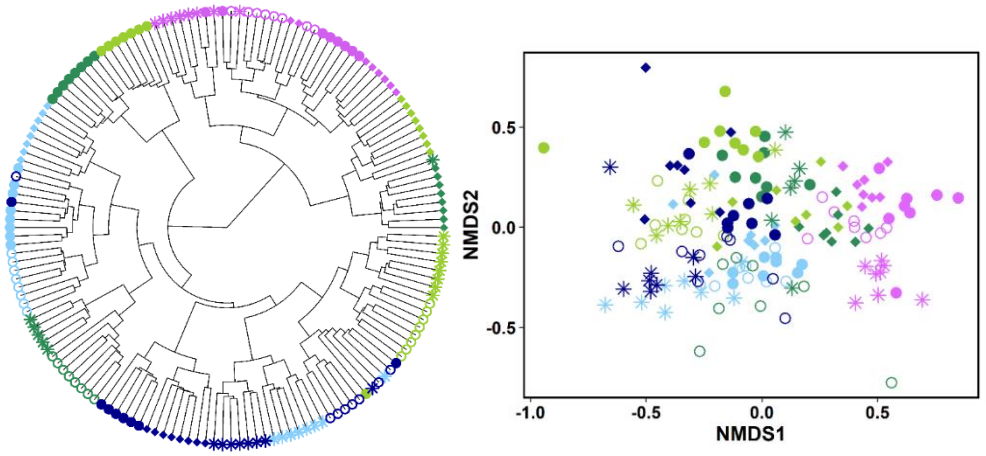


Figure 4.3. Hierarchical clustering dendrogram (left) and NMDS ordination (right) based on Bray-Curtis dissimilarities at OTU level of farmed mussel gut microbiota. Each symbol represents an individual *M. galloprovincialis* mussel; shapes of symbols correspond to different harvesting seasons - winter (✳), spring (○), summer (●) and autumn (◆) - and colours correspond to different harvest locations in Galician region (● AGES, ● SGES), Catalonia region (● DEES) and Basque Country region (● MEES, ● MUES).

3.2. Potential of using mussel microbiota as geographic indicator

3.2.1. Identification of taxa that differ between origins

Once settled that mussel DG microbiota differed depending on their geographical region, farming site and harvesting season, we wanted to identify which taxa had a significantly different representation between the geographical regions and could therefore act as potential origin indicator.

Overall, the bacterial community of mussel DG consisted of 12 phylum, 55 families and 67 genus (> 0.1% relative abundance). The seven most abundant bacterial phyla, which collectively accounted for 91.28% of all reads, were: *Firmicutes* (43.22%), *Proteobacteria* (22.38%), *Bacteroidota* (11.12%), *Cyanobacteria* (4.67%), *Verrucomicrobiota* (4.02%), *Fusobacteriota* (3.48%) and *Planctomycetota* (2.36%).

At family level, bacterial community composition of mussel DG differed between geographic locations and seasons (**Figure 4.4)**. According to Wilcoxon rank sum test analysis based on bacterial relative read abundance, several taxa displayed significantly

different representation between the three geographical regions. Overall, mussels from the Catalonian region showed significantly lower relative abundances of *Vibrionaceae*, *Rhodobacteraceae* and *Halieaceae* (phylum *Proteobacteria*) and *Flavobacteriaceae* (phylum *Bacteroidetes*), with respect to the other two regions. Mussel farms from the Galician region showed higher proportion of *Rubritaleaceae* (phylum *Verrucomicrobiota*) and *Pseudomonadales* (phylum *Proteobacteria*).

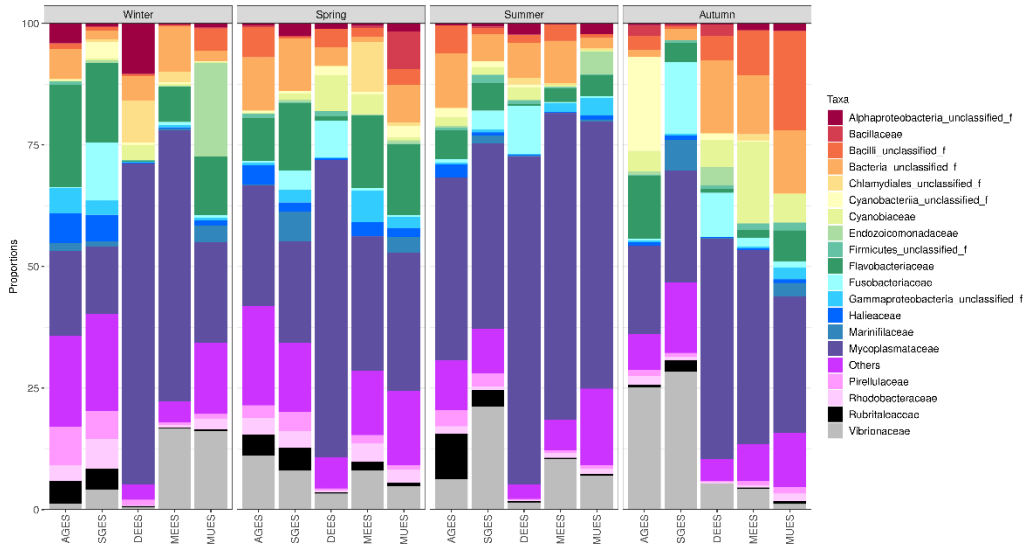


Figure 4.4. Relative abundance of bacterial communities, at family level, of mussel DG harvested from five different farms: in Galician region (AGES, SGES), Catalonia region (DEES) and Basque Country region (MEES, MUES). Taxa not within the 20 most abundant families were pooled together as "Other".

At genus level, mussels from the Catalonian region showed significantly lower relative abundances of *Polaribacter* (*Flavobacteriaceae*), *Lutimonas* (*Flavobacteriaceae*), Rubritalea (Rubritaleaceae), Rubripirellula (Pirellulaceae) genus with respect to the other two regions (**Figure 4.5)**. Galician region showed a significantly higher proportion of *Persicirhabdus* (*Rubritaleaceae*) and *Candidatus endoecteinascidia* (*Piscirickettsiaceae*). However, when applying these analyses on farming site **(Supplementary material 4.2)**, predominant genus also differed within the same region.

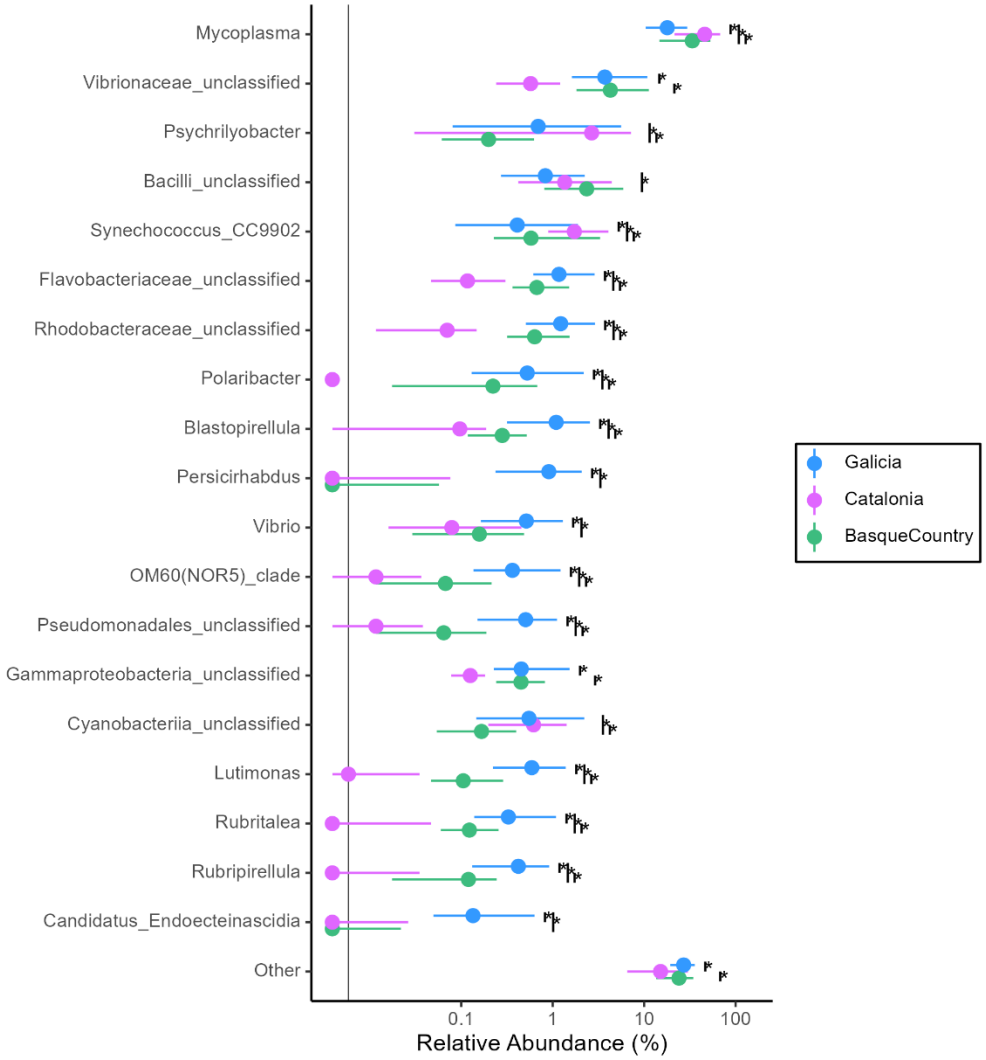


Figure 4.5. Relative abundance of genus that differed significantly (Kruskall Wallis, $p<0.05$) between the three mussel farming regions. Points are median values; lines represent the interquartile range and the black vertical line is the limit of detection. Pairwise comparisons were calculated by Wilcoxon rank sum test with Benjamini-Hochberg correction and significant differences are marked with asterisks (*). Taxa with a median relative abundance < 0.5% for all the locations were grouped in "Others". Colours correspond to different harvest regions: Galician region (● AGES + SGES), Basque Country region (● MEES + MUES) and Catalonia region (● DEES).

3.2.2. Stability of the bacterial community composition

Another critical factor, that would determine whether the bacterial community composition of mussels could be used as a reliable origin traceability tool, is the stability

of the fingerprint associated to each farm. To assess whether the microbial composition changed every year, the bacterial community structure of samples collected in 2019 was compared to those collected, sequenced and analysed independently a year later (2020) in the Basque Country at the exact same locations (MEES and MUES).

First, we evaluated the consistency and robustness of potential bacterial candidates which showed a significantly different representation by origin and could act as geographic origin indicators. Although some genera, such as *Persicirhabdus* and *Lutimonas,* showed similar proportions in 2020 (**Supplementary material 4.3**), others such as *Rubritalea* and *Polaribacter* showed significantly lower values a year later, closer to the values expected for Catalonia. This means that all the potential markers were not entirely consistent one year later and thus, focusing on specific differences (a small group of taxa or of OTUs that have significantly different representation between the geographical regions) might not be a suitable and robust approach for origin traceability.

Then, in order to circumvent this issue, we hypothesized whether analysing the overall community structure, instead of specific differences, will be more stable and thus, more suitable as a traceability approach. In this case, a hierarchical clustering based on the overall relative abundance of the bacterial family taxa was performed, by pooling farming site replicates (**Figure 4.6**). Results revealed that mussels collected in the Basque Country in two consecutive years clustered together, and even matching the season of harvesting.

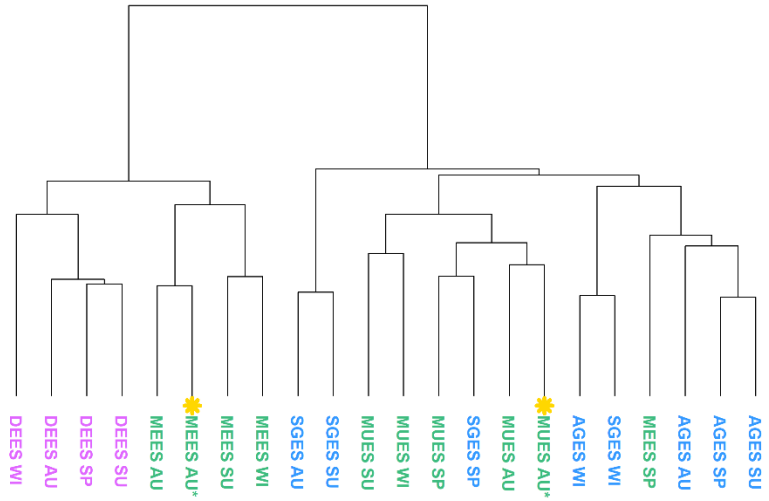


Figure 4.6. Hierarchical cluster analysis of Bray-Curtis dissimilarity matrices based on average relative abundance of mussel gut microbiota at family level. Individuals collected a year later, in 2020, are marked with a yellow asterisk. Colours correspond to different harvest regions: Galician region (● AGES + SGES), Basque Country region (● MEES + MUES) and Catalonia region (● DEES).

4. Discussion

4.1. Influence of geographic origin, farm location and season on the digestive microbiota of mussels

The present study investigated the bacterial communities associated to the digestive gland of *M. galloprovincialis* mussels collected seasonally from five different farms located in three regions of Spain. We found that microbial community structure of mussels was significantly different between harvesting locations. Overall, the bacterial composition we found in *Mytilus galloprovincialis* digestive gland, which was primarily composed of the phyla *Firmicutes*, *Proteobacteria* and *Bacteroidota*, is in concordance with the few previous microbiota studies performed on this species (Cappello et al., 2015; Vezzulli et al., 2018; Li et al., 2019; Musella et al., 2020). However, none of the previously mentioned studies has analysed the effect of geographical origin on microbial composition as considered here. Still, the influence of the geographic origin in the microbiota of bivalves has previously been described for other bivalves such as oysters

and clams (King et al., 2012; Lokmer et al., 2016; Pierce and Evan Warda, 2019; Liu et al., 2020). Several factors might have contributed to these changes in the bacterial fingerprints of mussels collected from different geographic locations. Previous studies have demonstrated that bivalve microbiota is influenced by a variety of environmental parameters (i.e. temperature, etc), diet and health condition, which account for the differences between shellfish (Li et al., 2019; Lokmer et al., 2016; Lokmer & Wegner, 2015; Melissa L. Pierce & Ward, 2018; Melissa L. Pierce et al., 2015).

The present study also showed that bacterial communities present in the digestive gland of mussels varied seasonally. When individuals from each farm were compared seasonally, it was possible to confirm that bacterial fingerprints associated to the digestive gland changed significantly. This seasonal variation could be depicted at all taxonomic levels (from OTU to phylum level). These results are in concordance with previous studies showing temporal variability in bivalves microbiota (Pierce et al., 2015; Lokmer et al., 2016; Pierce and Evan Warda, 2019). Overall, we found that the composition of the microbiota was distinctive between harvesting locations and seasons, with the effect prompted by the origin exceeding the seasonal variability, which is relevant for this approach to be used as an origin traceability tool.

4.2. Is the analysis of the microbiota an accurate tool to trace mussel farming origin?

The present study also assessed the potential of using the bacterial community analysis as an approach to discriminate the geographical origin. Indeed, results show that it was possible to discriminate the origin of farmed mussels collected from three regions of the Iberian Peninsula. The differentiation between Catalonian and Galician mussels was very consistent (**Figure 4.2**), with samples collected from each of the locations during 2019 clustering in their corresponding group. The bacterial fingerprint of individuals collected within the region of the Basque Country showed a different grouping pattern depending on the taxonomic level. Regarding the stability of the bacterial fingerprints, results have shown that mussel samples collected in two consecutive years clustered together matching even the corresponding season of harvesting (**Figure 4.5**). Besides,

our study also proved that the data obtained on individuals processed and analysed in different batches could be combined at different taxonomic levels. This would allow people to enrich the database with their own mussel's bacterial fingerprint and directly compare it for geographic origin assignment, without the need of reanalysing all the data.

The present study also allowed to recognize key genera that were potentially relevant to discriminate the geographic origin of the mussels being surveyed, such as, low proportions of *Polaribacter, Lutimonas, Rubritalea* and *Rubripirellula* in Catalonian region and higher levels of *Persicirhabdus* and *Candidatus endoecteinascidia* in Galicia. Still, our proposal is to rely on the overall community structure recorded for each production area, using the whole bacterial fingerprint instead of a few markers.

As previously suggested by Liu et al. (2020), for this approach to be used as a traceability tool, the development of a database with the reference bacterial fingerprint for each location of interest is necessary. In this way, the bacterial profile of new samples can be compared to this database to verify the origin claimed by the producer or trader. It is also important to highlight that the database must be periodically verified and updated, to enable a successful classification over time. Based on the results obtained in this study, the time window during which the bacterial fingerprint is likely to remain stable is at least a year, which enables the possibility to rely on previously recorded fingerprints without needing to constantly update the database. A well-designed strategy to maintain the database updated and facilitate the origin classification, could also provide useful information regarding the microbiological safety of the seafood products, allowing to detect potential viral and bacterial pathogens and thus, preventing contaminated products from entering markets.

4.3. Limitations of the bacterial profiling-based origin traceability approach and potential next steps

Although our results have promising implications for traceability, some limitations were identified. Mussels used in this study were fresh; however, shellfish are usually subjected to depuration, processing and preservation techniques prior to

commercialization. These procedures will likely impact the sensitivity of this approach for tracing the geographic origin using 16S rRNA amplicon sequencing of mussel microbiota. Previous studies have reported lower bacterial diversity and richness in depurated (Rubiolo et al., 2018; Vezzulli et al., 2018) and retail (Liu et al., 2020) bivalves. To fully understand the effect of these techniques in the bacterial fingerprint and to assess the usefulness of this approach in such cases, further studies are required. Besides, this type of bivalves usually has distinct trades since they can be commercialized as fresh, frozen or processed mussel products. Thus, this single methodological approach may not be sufficient to reliably trace the whole international trade. In these cases, a multidisciplinary approach combining different traceability techniques might be the best option.

Overall, the present study reinforces the potential use of bacterial profiles to trace seafood to their geographic origin. In the future, it would be interesting to test the approach with a larger number of samples from additional regions where this species is actively cultivated, and also, to combine the bacterial profiles obtained with novel machine learning pipelines (as proposed by Milan et al. (2019)) to progress towards estimating the assignment probability.

EZTABAIDA OROKORRA

Arrain eta itsaskien eskaria gero eta handiagoa da mundu osoan, eta horrek hornikuntza-kateen konplexutasuna areagotzen du, trazabilitate ahaleginak zailduz. Itsas produktuen jatorri geografikoa egiaztatzea funtsezkoa da elikagaien trazabilitate eta segurtasun legeak betetzen direla ziurtatzeko eta itsas baliabideen kudeaketa eraginkor eta jasangarria bermatzeko, baita agintari, ekoizle eta kontsumitzaile arduratsuen egungo eta etorkizuneko itxaropenak betetzeko ere. Testuinguru honetan, itsas-produktuen etiketan ageri den informazioa zuzena dela egiaztatzeko tresna analitikoen garapenak berebiziko garrantzia du.

Lan hau, hazitako *Mytilus galloprovincialis* muskuiluen jatorri geografikoa egiaztatzeko tresna analitikoen garapenera bideratuta dago. Horretarako, lehenik eta behin, ikerketarako erabili diren indibiduoen espeziea egiaztatu da, 1. kapituluan garatutako PCR metodologia azkarraren bidez. Ondoren, bibalbio hauen jatorria egiaztatzeko, markatzaile genetikoetan (2. kapitulua), aztarna-elementuen konposizioan (3. kapitulua) eta bakterio-komunitateen profilan (4. kapitulua) oinarritutako tresna analitikoek duten erabilera aztertu da. Lortutako emaitzek argi utzi dute muskuilu hauen jatorria egiaztatzea posiblea dela, produktu mota bakoitzerako tresna egokia hautatzen denean.

1. Jatorriaren trazabilitaterako tresnak: indarguneak eta mugak

Azken urteotan, gero eta metodologia berri gehiago garatzen ari dira arrain eta itsaskien jatorria egiaztatzeko (Leal et al., 2015; Gopi et al., 2019a). Lan honetan zehar aztertutako tresna genetiko, kimiko eta mikrobiologikoek arrakasta maila ezberdinak erakutsi dituzte, bakoitzak bere zailtasun eta aukerak aurkeztuz.

Aztertutako tresnen artean, SNP markatzaileek *M. galloprovincialis* muskuiluen egitura genetikoa argitu dute eta haien jatorri geografikoa esleitzeko gaitasun handia erakutsi dute. SNP markatzaileak erabiltzeak abantaila ugari ditu (Bernatchez et al., 2017). Lehenik eta behin, SNPetan oinarritutako tresna hau erraz inplementa daiteke edozein laborategitan; eta nahiz eta ikerketa taldea desberdina izan eta SNPak genotipatzeko plataforma ezberdina erabili, emaitzak arazo barik aldera daitezke (Ogden and Linacre, 2015; Nielsen, 2016). Bigarrenik, markatzaile hauek sentikortasun eta

espezifikotasun handia dute, eta beste bi tresnek ez bezala, itsas produktu prozesatuetan aplika daitezke (Scarano and Rao, 2014). Honi esker, hartu berri diren muskuilu freskoetan zein merkatuan dauden hainbat produktu prozesatuetan (adib., produktu izoztuetan, kontserbatan, atmosfera aldatua duten produktuetan edota jateko prest dauden produktuetan) erabil daitezke. Azkenik, SNP markatzaileak egonkorragoak dira denboran zehar; hau da, populazioen aztarna genetikoek ez dituzte aldaketak jasaten epe laburrean (Nielsen et al., 2012).

Jakina, jatorria egiaztatzeko metodo honek muga batzuk ere baditu. Lehenengoa, SNP markatzaileetan oinarritutako metodoak ez dira geografikoki hurbil dauden espezie bereko indibiduoak bereizteko gai (Nielsen, 2016). Larba planktonikoak dituzten itsas espezieen kasuan (hau da, muskuiluen kasuan), fluxu genetikorako langa egon ezean, geografikoki gertu dauden indibiduoen arteko bereizketa genetikoa ezinezkoa da. Hori dela eta, markatzaile genetikoak erabiliz, gertuko akuikultura-instalazioetan hazitako muskuiluen jatorria desberdintzea ez da posiblea izango (Leal et al., 2015). Lan honen 2. kapituluan lortutako emaitzek, Atlantikoko eta Mediterraneoko *M. galloprovincialis* leinuen arteko bereizketa argia erakusten dute, AOOF fluxu genetikorako langaren ondorioz izan daitekeena (Diz and Presa, 2008). Aldiz, Atlantikoan batutako muskuiluen artean (Galizia eta Euskadiko muskuiluen artean esaterako) ez da bereizketa genetikorik antzeman. Atlantiko barneko homogeneotasun honek adierazten du bertan ez dagoela muskuiluen fluxu genetikoa eteten duen langarik. Bigarren muga, muskuilu haziaren jatorriarekin lotura du. Espainian *M. galloprovincialis* muskuiluak hazteko, hazia ingurune naturaletik bertatik jasotzen dute, beraz ez dago arazorik. Aldiz, hazia beste eremu geografiko batetik lekualdatuko balitz edo atzerriko hazitegi batetik hartuko balitz, ezin izango litzateke azterketa genetikoak erabili jatorria zehazteko.

Muga hauek gainditzeko, herentzia prozesuetan oinarritutako tresna genetikoak erabili ordez, indibiduoen ingurune ekologikoa islatzen duten markatzaileetan oinarrituriko tresnak ere aztertu ziren. Aztarna kimikoetan, eta bereziki aztarna-elementuetan, oinarritutako tresnek populazioak bereizteko aukera ematen dute, moluskuen maskorrak bezalako egitura biogenikoetan dauden elementuen konposizioa aztertuz. Hainbat arrazoi daude aztarna-elementuen konposizioa jatorria esleitzeko

markatzaile natural gisa erabiltzeko. Lehenik eta behin, geografikoki gertu dauden populazioak bereizteko metodo zehatza eta fidagarria da. Tesiaren 3. kapituluan frogatu denez, gertuko akuikultura-instalazioetan hazitako muskuiluen jatorria zehaztasun handiz desberdintzea lortu da. Hau, Ricardo et al. (2015) eta Bennion et al. (2019) egindako ikerketekin bat dator, non aztarna-elementuen konposizioa erabili zen bata bestearengandik 6 kilometrotara hazitako bibalbioen jatorria esleitzeko. Bigarrenik, egitura biogeniko hauek inerteak dira; hau da, metatu berri den materiala ez da berriro xurgatzen ezta eraldatzen ere, eta hortaz, ingurumen-baldintza aldakorren erregistro kimiko iraunkorra ematen du (Artetxe-Arrate et al., 2019).

Jatorriaren egiaztapenerako aztarna-elementuen konposizioa analizatzearen muga nagusia egitura hauek (maskorrak gure kasuan) aztertu nahi den produktuan presente egon behar direla da. Hori dela eta, maskorrik gabeko muskuilu prozesatuetan (adib., produktu izoztu edo kontserbatan) ezin izango da tresna hau erabili (Leal et al., 2015). Gainera, ingurumen-baldintzek (hala nola, tenperatura aldaketek, euriteek edota gazitasun aldaketek) aztarna-elementuen konposizioan eragin zuzena dute (Klein et al., 1996; Poulain et al., 2015; Zhao et al., 2020), eta horrek tresna horren erabileran oztopatu dezake. Aurreko ikerketek frogatu dutenez, aztarna-elementuak egonkor mantentzen diren denbora-tartea sei hilabete eta urtebete bitartekoa da (Ricardo et al., 2017b). Aldakortasun hau dela eta, aztertu nahi diren ekoizpen-guneetan hazten diren muskuiluen aztarna-elementuekin erreferentzia datu-base bat sortu behar da. Datu-base hau aldizka eguneratu behar da, denboran zehar laginak arrakastaz sailkatu ahal izateko. Aldizkako eguneratzearen kostu ekonomikoei aurre egiteko, laginketa hau, uraren kalitatea neurtzeko edota patogenoen eta alga kaltegarrien presentzia kontrolatzeko egiten diren ohiko laginketekin batera egin daiteke.

Azkenik, bakterioen 16S rRNA genearen V3-V4 gune aldakorraren anplifikazio eta sekuentziazioaren bidez lortutako bakterio-komunitateen profilak jatorria arakatzeko biomarkatzaile gisa erabili ahal direla frogatu da; izan ere, lortutako profilak akuikultura-instalazio bakoitzaren jatorriaren adierazgarri dira (4. kapitulua). Jatorria egiaztatzeko tresna gisa erabiltzeaz gain, bakterio-komunitateen profilak itsas produktuen segurtasun mikrobiologikoari buruzko informazio erabilgarria eman lezake.

Hala ere, aztarna-elementuekin gertatzen den moduan, lortutako profilak denborarekin aldatuz doaz eta hortaz, intereseko jatorrietatik batutako muskuiluen bakterio-komunitateen profilekin erreferentziazko datu-base bat eratzea beharrezkoa da. Horrela, datu basea eguneratua mantenduz, denboran zehar lortzen diren lagin berrien jatorri esleipen egokia egin ahal izango da. Tesi honetan, bakterio-komunitateen profilak urtebete behintzat egonkor mantentzen direla frogatu da. Izan ere, urte bat beranduago batutako laginak esperotako jatorria erakutsi dute, eta gainera, urtaroarekin bat datozela ere ikusi da. Bibalbio hauen mikrobiota markatzaile gisa erabiltzeak beste eragozpen bat ere izan dezake: bakterio-komunitateen profila alda daiteke arazketa eta biltegiratze prozesuen ondorioz (Rubiolo et al., 2018; Liu et al., 2020). Seguruenik, bi prozesu hauek eragin handia izango dute jatorria esleitzeko tresna honen erabileran, eta hortaz, ikerketa gehiago egin beharko lirateke efektua guztiz ondo ulertzeko. Horretaz gain, bakterio-komunitateen profilaren bidezko jatorri esleipenaren probabilitatea estimatzeko, interesgarria izango litzateke jatorri ikaste automatikoan oinarritutako *Random Forest* bezalako sailkapen-metodoa erabiltzea, Milan et al. (2019)-ek *Ruditapes philippinarum* txirlekin egin zuen moduan.

Laburbilduz, jatorri geografikoaren egiaztapena oso konplexua da, eta momentuz, ez dago tresna perfekturik, denek agertzen baitituzte mugak. Beraz, aztertu nahi den espeziearen eta produktu motaren arabera (bizirik edo prozesatuta), tresna bakoitzaren egokitasuna aztertu beharko da.

2. Moluskuen trazabilitatearen garrantzia: egungo aplikazioak eta hurrengo urratsak

Lan honetan aztertutako tresnak molusku bibalbioen jatorria egiaztatzeko egokiak direla frogatu da, eta beraz, funtsezko zeregina izan dezakete merkataritza globalizatuaren gardentasuna eta segurtasuna bermatzeko. Azken urteetan, arrain eta itsaskien trazabilitatea asko hobetu da, baina oraindik ere espezie edo jatorriarekin lotutako iruzurra dago. Honek, ekonomia eta kontsumitzaileen informazio-eskubideak arriskuan jartzeaz gain, osasun publikoan ere kalte izugarria eragin dezake (Fox et al.,

2018). Bibalbioak nutriente-dentsitate handiena duten itsas produktuen artean sailkatuta dauden arren, patogeno edo toxikoekin kutsatutako uretan hazten baldin badira edo behar bezala prestatu gabe (gordin edo gutxi eginda) jaten baldin badira, kontsumitzaileen osasuna arriskuan jar dezakete. Hortaz, hazkuntza jatorri geografiko zehatza egiaztatzeko tresnak espezifikoak edukitzea funtsezkoa da osasun publikoa babesteko.

Aldi berean, trazabilitate tresna hauek oso baliotsuak dira itsas produktuen ekoizleentzat, haien produktuen jatorri izendapenak behar bezala egiaztatzea ahalbidetzen dutelako. Jatorriak, kontsumitzaileek produktu baten kalitateari eta zaporeari buruz duten pertzepzioan eragiten du (Fonner and Sylvia, 2015; Fox et al., 2018). Hau da, gerta daiteke kontsumitzaileak toki jakin batzuetatik etorritako itsas produktuengatik gehiago ordaintzeko prest egotea, eta hortaz, jatorria bermatzeak produktu bati balio erantsi handiagoa lortzeko aukera ematea. Lan honetan lortutako jatorrien bereizketa egokiak, jatorri aitorpenak behar bezala egiaztatzeko bidea eraikitzen du.

Oro har, lan honetan aurkeztutako emaitzek agerian utzi dute jatorri geografikoa tresna analitikoen bidez egiaztatzearen konplexutasuna. Gaur egun ez dago trazabilitate tresna perfekturik, testuinguru zehatzaren eta itsas produktu motaren arabera tresna bakoitzaren egokitasuna aztertu behako da. Bien bitartean, badirudi, elkarren osagarri diren tresnen konbinazioa dela zehaztasuna eta fidagarritasuna handitzeko biderik onena. Zenbait ikerlariek arrain eta itsaskien jatorria egiaztatzeko tresnak konbinatzearen onurak frogatu dituzte (Gopi et al., 2019c; Brophy et al., 2020; Cusa et al., 2021). Halaber, azpimarratzekoa da ere, orain arte, ikerketa gutxi batzuk baino ez dutela jatorria esleitzeko probabilitatea zehaztu (Cusa et al., 2021), (lan honetan markatzaile genetiko eta kimikoekin egin den bezala), eta probabilitate hau da, hain zuzen ere, kontrol-agintariek eta erakunde arautzaileek bilatzen dutena. Zalantzarik gabe, etorkizunean jatorria egiztatzeko tresna analitikoen garapenean aurrerapauso handiak emango dira, itsas espezie gehiagoren jatorria modu fidagarri eta zehatzagoan esleituz. Nolanahi ere, teknika hauek inplementatu ahal izateko, komunitate zientifikoak estuki lan egin beharko du erakunde arautzaileekin, gobernu-organoekin eta itsas-

industriarekin; tresna analitikoen potentziala erakusteko eta egungo eta etorkizuneko trazabilitate erronkei aurre egiteko.

ONDORIOAK ETA TESIA

Ondorioak

Tesi honen helburuak kontuan hartuta, honako hau ondorioztatu da:

1. Lan honetan garatutako denbora errealeko PCR saiakuntza azkarrak Europan merkaturatzen diren *Mytilus* espezie garrantzitsuenen (*M. galloprovincialis, M. edulis* eta *M. chilensis*) identifikazioa ahalbidetzen du, produktu izoztuak eta kontserbak barne hartuz.
2. Milaka SNP markatzaile nuklearretan oinarritutako populazio egituraren analisiek, Atlantikoko eta Mediterraneoko *M. galloprovincialis* populazioak nabarmen bereizten direla erakutsi dute, eta aldi berean, nahiz eta antzinako populazio beretik etorri, Mediterraneoko muskuilu natiboen eta sartu berri diren Ozeano Bareko hego-ekialdeko muskuiluen artean ere ezberdintasun esanguratsuak daudela erakutsi dute.
3. Diskriminazio ahalmen handiena agertu duten SNPak erabiliz, muskuiluak Atlantikoko edo Mediterraneoko/Hego-ekialdeko jatorri eremuetara era egokian esleitzeko tresna analitiko bat garatu da; hain zuzen, 10 SNP erabiliz % 90eko jatorri esleipen egokia lortu da eta 25 SNP erabiliz % 100ekoa. Honek, muskuilu fresko, izoztu edo eraldatuetan aplikagarria den trazabilitate tresna zehatz eta errentagarri baten garapenerako bidea irekitzen du.
4. Muskuiluen ehunean dauden isotopo egonkorren konposizioa eta muskuiluen maskorrean dauden aztarna-elementuen konposizioa konbinatuz, hamar akuikultura-instalazio ezberdinetan hazitako *M. galloprovincialis* muskuiluen jatorri geografikoa esleitzea posiblea da.
5. *M. galloprovincialis* muskuilu freskoen digestio-guruinari lotutako bakterio-komunitateen profila haien hazkuntza eremua bereizteko erabil daiteke, jatorri geografikoaren eragina urtaroen aldakortasuna gainditzen duelarik.
6. Bakterio-komunitateen profilaren egonkortasunari dagokionez, urtaroaren arabera aldatu arren, urte bateko aldearekin batutako, prozesatutako eta

sekuentziatutako laginak haien jatorri geografikoko laginekin taldekatu dira, bilketa urtaroarekin ere bat eginez. Honek indartu egiten du arrain eta itsaskien jatorria arakatzeko bakterio-komunitateen profilek duten erabilera.

7. Lan honetan aurkeztutako emaitzek agerian utzi dute jatorri geografikoa tresna analitikoen bidez egiaztatzearen konplexutasuna. Momentuz, jatorri esleipenaren zehaztasuna eta fidagarritasuna handitzeko, elkarren osagarri diren tresnak konbinatzea da hurbilketarik onena.

Lan honetan lortutako emaitzei esker, planteatutako hipotesia landu ahal izan da, tesia honako hau izanik:

Tesia

"Eremu geografiko ezberdinetan hazitako Mytilus galloprovincialis muskuiluek aztarna genetiko, kimiko eta mikrobiologiko ezberdinak dituzte, eta horiei esker, elikagai honen jatorria egiaztatzeko tresna analitiko berriak garatu daitezke, arrain eta itsaskien hornidura-katearen gardentasuna indartuz eta kontsumitzaileen konfiantza areagotuz.

REFERENCES

Agnew, D. J., Pearce, J., Pramod, G., Peatman, T., Watson, R., Beddington, J. R., and Pitcher, T. J. (2009). Estimating the Worldwide Extent of Illegal Fishing. *PLOS ONE*, 4: e4570. https://doi.org/10.1371/journal.pone.0004570

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19: 1655-1664. https://doi.org/10.1101/gr.094052.109

Anderson, A. S., Bilodeau, A. L., Gilg, M., and Hilbish, T. J. (2002). Routes of introduction of the Mediterranean mussel (*Mytilus galloprovincialis*) to Puget Sound and Hood Canal. *Journal of Shellfish Research*, 21: 75-79.

Anderson, E. C. (2010). Assessing the power of informative subsets of loci for population assignment: standard methods are upwardly biased. *Molecular Ecology Resources*, 10: 701-710. https://doi.org/10.1111/j.1755-0998.2010.02846.x

Apte, S., Holland, B. S., Godwin, L. S., and Gardner, J. P. A. (2000). Jumping Ship: A Stepping Stone Event Mediating Transfer of a Non-indigenous Species Via a Potentially Unsuitable Environment. *Biological Invasions*, 2: 75-79. https://doi.org/10.1023/A:1010024818644

Araneda, C., Larrain, M. A., Hecht, B., and Narum, S. (2016). Adaptive genetic variation distinguishes Chilean blue mussels (*Mytilus chilensis*) from different marine environments. *Ecology and Evolution*, 6: 3632-3644. https://doi.org/10.1002/ece3.2110

Araneda, C., Pardo, M. Á., Jiménez, E., Longa, Á., Lee, R.-S., Segura, C., Zbawicka, M*., et al.* (2021). A comment on Giusti et al. (2020) "Mussels (*Mytilus spp*.) products authentication: A case study on the Italian market confirms issues in species identification and arises concern on commercial names attribution, Food Control Volume 118, December 2020, 107, 379". *Food Control*, 121: 107626. https://doi.org/10.1016/j.foodcont.2020.107626

Arribas, L., Donnarumma, L., Palomo, M., and Scrosati, R. (2014). Intertidal mussels as ecosystem engineers: Their associated invertebrate biodiversity under contrasting wave exposures. *Marine Biodiversity*, 44: 203-211. https://doi.org/10.1007/s12526-014-0201-z

Artetxe-Arrate, I., Fraile, I., Crook, D. A., Zudaire, I., Arrizabalaga, H., Greig, A., and Murua, H. (2019). Otolith microchemistry: a useful tool for investigating stock structure of yellowfin tuna (*Thunnus albacares*) in the Indian Ocean. *Marine and Freshwater Research*, 70: 1708. https://doi.org/10.1071/mf19067

Astorga, M. P., Cardenas, L., and Vargas, J. (2015). Phylogenetic Approaches to Delimit Genetic Lineages of the *Mytilus* Complex of South America: How Many Species are there? *Journal of Shellfish Research*, 34: 919-930, 912. https://doi.org/10.2983/035.034.0322

Babaranti, O., Horn, S., Jowett, T., and Frew, R. (2019). Isotopic signatures in *Mytilus galloprovincialis* and *Ulva latuca* as bioindicators for assessing discharged sewage effluent in coastal waters along Otago Peninsula, New Zealand. *Geology, Ecology, and Landscapes*, 3: 53-64. https://doi.org/10.1080/24749508.2018.1485079

Bahar, B., Schmidt, O., Moloney, A. P., Scrimgeour, C. M., Begley, I. S., and Monahan, F. J. (2008). Seasonal variation in the C, N and S stable isotope composition of retail organic and conventional Irish beef. *Food Chemistry*, 106: 1299-1305. https://doi.org/10.1016/j.foodchem.2007.07.053

Barcenas, N. M., Unruh, T. R., and Neven, L. G. (2005). DNA Diagnostics to Identify Internal Feeders (Lepidoptera: Tortricidae) of Pome Fruits of Quarantine Importance. *Journal of Economic Entomology*, 98: 299-306. https://doi.org/10.1093/jee/98.2.299

Barhoumi, B., El Megdiche, Y., Clérandeau, C., Ameur, W. B., Mekni, S., Bouabdallah, S., Derouiche, A.*, et al.* (2016). Occurrence of polycyclic aromatic hydrocarbons (PAHs) in mussel (*Mytilus galloprovincialis*) and eel (*Anguilla anguilla*) from Bizerte lagoon, Tunisia, and associated human health risk assessment. *Continental Shelf Research*, 124: 104-116. https://doi.org/10.1016/j.csr.2016.05.012

Barhoumi, B., LeMenach, K., Devier, M.-H., Ameur, W. B., Etcheber, H., Budzinski, H., Cachot, J.*, et al.* (2014). Polycyclic aromatic hydrocarbons (PAHs) in surface sediments from the Bizerte Lagoon, Tunisia: levels, sources, and toxicological significance. *Environmental Monitoring and Assessment*, 186: 2653-2669. https://doi.org/10.1007/s10661-013-3569-5

Beaumont, A. R., Turner, G., Wood, A. R., and Skibinski, D. O. F. (2004). Hybridisations between *Mytilus edulis* and *Mytilus galloprovincialis* and performance of pure species and hybrid veliger larvae at different temperatures. *Journal of Experimental Marine Biology and Ecology*, 302: 177-188. https://doi.org/10.1016/j.jembe.2003.10.009

Becker, B. J., Fodrie, F. J., McMillan, P. A., and Levin, L. A. (2005). Spatial and temporal variation in trace elemental fingerprints of mytilid mussel shells: A precursor to invertebrate larval tracking. *Limnology and Oceanography*, 50: 48-61. https://doi.org/10.4319/lo.2005.50.1.0048

Ben Ameur, W., El Megdiche, Y., Eljarrat, E., Ben Hassine, S., Badreddine, B., Souad, T., Bèchir, H.*, et al.* (2013). Organochlorine and organobromine compounds in a benthic fish (*Solea solea*) from Bizerte Lagoon (northern Tunisia): Implications for human exposure. *Ecotoxicology and Environmental Safety*, 88: 55-64. https://doi.org/10.1016/j.ecoenv.2012.10.021

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society*, 57: 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bennion, M., Morrison, L., Brophy, D., Carlsson, J., Abrahantes, J. C., and Graham, C. T. (2019). Trace element fingerprinting of blue mussel (*Mytilus edulis*) shells and soft tissues successfully reveals harvesting locations. *The Science of the total environment*, 685: 50-58. https://doi.org/10.1016/j.scitotenv.2019.05.233

Bennion, M., Morrison, L., Shelley, R., and Graham, C. (2021). Trace elemental fingerprinting of shells and soft tissues can identify the time of blue mussel (*Mytilus edulis*) harvesting. *Food Control*, 121: 107515. https://doi.org/10.1016/j.foodcont.2020.107515

Bernatchez, L., Wellenreuther, M., Araneda, C., Ashton, D. T., Barth, J. M. I., Beacham, T. D., Maes, G. E.*, et al.* (2017). Harnessing the Power of Genomics to Secure the Future of Seafood. *Trends in ecology & evolution*, 32: 665-680. https://doi.org/10.1016/j.tree.2017.06.010

Berry, O., and Sarre, S. (2007). Gel-free species identification using melt-curve analysis. *Molecular Ecology Notes*, 7: 1-4. https://doi.org/10.1111/j.1471-8286.2006.01541.x

Breiman, L. (2001). Random Forests. *Machine Learning*, 45: 5-32. https://doi.org/10.1023/a:1010933404324

Briant, N., Savoye, N., Chouvelon, T., David, V., Rodriguez, S., Charlier, K., Sonke, J. E.*, et al.* (2018). Carbon and nitrogen elemental and isotopic ratios of filter-feeding bivalves

along the French coasts: An assessment of specific, geographic, seasonal and multi-decadal variations. *Science of the Total Environment*, 613: 196-207. https://doi.org/10.1016/j.scitotenv.2017.08.281

Broadaway, B. J., and Hannigan, R. E. (2012). Elemental Fingerprints used to Identify Essential Habitats: Nantucket Bay Scallop. *Journal of Shellfish Research*, 31: 671-676, 676. https://doi.org/10.2983/035.031.0310

Bronner, I. F., Quail, M. A., Turner, D. J., and Swerdlow, H. (2014). Improved Protocols for Illumina Sequencing. *Current Protocols in Human Genetics*, 80: 1-42. https://doi.org/10.1002/0471142905.hg1802s80

Brooks, S. J., and Farmen, E. (2013). The Distribution of the Mussel *Mytilus* Species Along the Norwegian Coast. *Journal of Shellfish Research*, 32: 265-270. https://doi.org/10.2983/035.032.0203

Brophy, D., Rodriguez-Ezpeleta, N., Fraile, I., and Arrizabalaga, H. (2020). Combining genetic markers with stable isotopes in otoliths reveals complexity in the stock structure of Atlantic bluefin tuna (*Thunnus thynnus*). *Scientific Reports*, 10: 14675. https://doi.org/10.1038/s41598-020-71355-6

Cappello, S., Volta, A., Santisi, S., Lucrezia, Genovese, Maricchiolo, G., Bonsignore, M.*, et al.* (2015). Study of Bacterial Communities in Mussel *Mytilus galloprovincialis* by a Combination of 16 s crDNA and 16S rDNA Sequencing. *International Journal of Microbiology and Application*, 2: 18-24.

Carrera, M., and Gallardo, J. M. (2017). Determination of the Geographical Origin of All Commercial Hake Species by Stable Isotope Ratio (SIR) Analysis. *Journal of Agricultural and Food Chemistry*, 65: 1070-1077. https://doi.org/10.1021/acs.jafc.6b04972

Carson, H. S. (2010). Population connectivity of the Olympia oyster in Southern California. *Limnology and Oceanography*, 55: 134-148. https://doi.org/10.4319/lo.2010.55.1.0134

Carson, H. S., López-Duarte, P. C., Cook, G. S., Fodrie, F. J., Becker, B. J., DiBacco, C., and Levin, L. A. (2013). Temporal, spatial, and interspecific variation in geochemical signatures within fish otoliths, bivalve larval shells, and crustacean larvae. *Marine Ecology Progress Series*, 473: 133-148. https://doi.org/10.3354/meps10078

Castigliego, L., Armani, A., Tinacci, L., Gianfaldoni, D., and Guidi, A. (2015). Two alternative multiplex PCRs for the identification of the seven species of Anglerfish (*Lophius* spp.) using an end-point or a melting curve analysis real-time protocol. *Food Chemistry*, 166: 1-9. https://doi.org/10.1016/j.foodchem.2014.06.014

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., and Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*, 22: 3124-3140. https://doi.org/10.1111/mec.12354

Chaguri, M. P., Maulvault, A. L., Nunes, M. L., Santiago, D. A., Denadai, J. C., Fogaça, F. H., Sant'Ana, L. S.*, et al.* (2015). Different tools to trace geographic origin and seasonality of croaker (*Micropogonias furnieri*). *LWT - Food Science and Technology*, 61: 194-200. https://doi.org/10.1016/j.lwt.2014.11.006

Chapela, M. J., Sotelo, C. G., Pérez-Martín, R. I., Pardo, M. Á., Pérez-Villareal, B., Gilardi, P., and Riese, J. (2007). Comparison of DNA extraction methods from muscle of canned tuna for species identification. *Food Control*, 18: 1211-1215. https://doi.org/10.1016/j.foodcont.2006.07.016

Cherifi, H., Chebil Ajjabi, L., and Sadok, S. (2018). Nutritional value of the Tunisian mussel *Mytilus galloprovincialis* with a special emphasis on lipid quality. *Food Chemistry*, 268: 307-314. https://doi.org/10.1016/j.foodchem.2018.06.075

Cuéllar-Pinzón, J., Presa, P., Hawkins, S. J., and Pita, A. (2016). Genetic markers in marine fisheries: Types, tasks and trends. *Fisheries Research*, 173: 194-205. https://doi.org/10.1016/j.fishres.2015.10.019

Cusa, M., St John Glew, K., Trueman, C., Mariani, S., Buckley, L., Neat, F., and Longo, C. (2021). A future for seafood point-of-origin testing using DNA and stable isotope signatures. *Reviews in Fish Biology and Fisheries*. https://doi.org/10.1007/s11160-021-09680-w

Cutler, D. R., Edwards Jr., T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., and Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology*, 88: 2783-2792. https://doi.org/10.1890/07-0539.1

del Rio-Lavín, A., Jiménez, E., and Pardo, M. Á. (2021). SYBR-Green real-time PCR assay with melting curve analysis for the rapid identification of *Mytilus* species in food samples. *Food Control*, 130: 108257. https://doi.org/10.1016/j.foodcont.2021.108257

Deniro, M. J., and Epstein, S. (1978). Influence of diet on the distribution of carbon isotopes in animals. *Geochimica et Cosmochimica Acta*, 42: 495-506. https://doi.org/10.1016/0016-7037(78)90199-0

Deniro, M. J., and Epstein, S. (1981). Influence of diet on the distribution of nitrogen isotopes in animals. *Geochimica et Cosmochimica Acta*, 45: 341-351. https://doi.org/10.1016/0016-7037(81)90244-1

Deudero, S., Cabanellas, M., Blanco, A., and Tejada, S. (2009). Stable isotope fractionation in the digestive gland, muscle and gills tissues of the marine mussel *Mytilus galloprovincialis*. *Journal of Experimental Marine Biology and Ecology*, 368: 181-188. https://doi.org/10.1016/j.jembe.2008.10.008

Dias, P. J., Sollelis, L., Cook, E. J., Piertney, S. B., Davies, I. M., and Snow, M. (2008). Development of a real-time PCR assay for detection of *Mytilus* species specific alleles: Application to a sampling survey in Scotland. *Journal of Experimental Marine Biology and Ecology*, 367: 253-258. https://doi.org/10.1016/j.jembe.2008.10.011

Diz, A. P., and Presa, P. (2008). Regional patterns of microsatellite variation in *Mytilus galloprovincialis* from the Iberian Peninsula. *Marine Biology*, 154: 277-286. https://doi.org/10.1007/s00227-008-0921-3

Doubleday, Z. A., Martino, J. C., and Trueman, C. (2022). Harnessing universal chemical markers to trace the provenance of marine animals. *Ecological Indicators*, 144. https://doi.org/10.1016/j.ecolind.2022.109481

Dubois, S., Jean-Louis, B., Bertrand, B., and Lefebvre, S. (2007a). Isotope trophic-step fractionation of suspension-feeding species: Implications for food partitioning in coastal ecosystems. *Journal of Experimental Marine Biology and Ecology*, 351: 121-128. https://doi.org/10.1016/j.jembe.2007.06.020

Dubois, S., Orvain, F., Marín Leal, J. C., Ropert, M., and Lefebvre, S. (2007b). Small-scale spatial variability of food partitioning between cultivated oysters and associated suspension-feeding species, as revealed by stable isotopes. *Marine Ecology Progress Series*, 336: 151-160. https://doi.org/10.3354/meps336151

Dunphy, B., Silva, C., and Gardner, J. P. A. (2015). Testing techniques for tracing the provenance of green-lipped mussel spat washed up on Ninety Mile Beach. *New Zealand Aquatic Environment and Biodiversity Report* 164. https://doi.org/10.13140/rg.2.1.3379.3525

Dunphy, B. J., Millet, M. A., and Jeffs, A. G. (2011). Elemental signatures in the shells of early juvenile green-lipped mussels (*Perna canaliculus*) and their potential use for larval tracking. *Aquaculture*, 311: 187-192. https://doi.org/10.1016/j.aquaculture.2010.12.016

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, 27: 2194-2200. https://doi.org/10.1093/bioinformatics/btr381

El Sheikha, A. F., and Montet, D. (2016). How to Determine the Geographical Origin of Seafood? *Critical reviews in food science and nutrition*, 56: 306-317. https://doi.org/10.1080/10408398.2012.745478

Elliott, J., Holmes, K., Chambers, R., Leon, K., and Wimberger, P. (2008). Differences in morphology and habitat use among the native mussel *Mytilus trossulus*, the non-native *M. galloprovincialis*, and their hybrids in Puget Sound, Washington. *Marine Biology*, 156: 39-53. https://doi.org/10.1007/s00227-008-1063-3

Espinoza, P., Lorrain, A., Ménard, F., Cherel, Y., Tremblay-Boyer, L., Argüelles, J., Tafur, R.*, et al.* (2017). Trophic structure in the northern Humboldt Current system: new perspectives from stable isotope analysis. *Marine Biology*, 164: 86. https://doi.org/10.1007/s00227-017-3119-8

Etter, P. D., Bassham, S., Hohenlohe, P. A., Johnson, E. A., and Cresko, W. A. (2011). SNP Discovery and Genotyping for Evolutionary Genetics Using RAD Sequencing. *In* Molecular Methods for Evolutionary Genetics, pp. 157-178. Ed. by V. Orgogozo, and M. V. Rockman. https://doi.org/10.1007/978-1-61779-228-1_9

EU (2007). Commission Regulation (EC) No 1050/2007 of 12 September 2007 registering certain names in the Register of protected designations of origin and protected geographical indications (Mejillón de Galicia or Mexillón de Galicia (PDO) — Café de Colombia (PGI) — Castagna Cuneo (PGI) — Asparago Bianco di Bassano (PDO)). http://data.europa.eu/eli/reg/2007/1050/oj

EU (2013a). Commission Implementing Regulation (EU) No 1200/2013 of 25 November 2013 entering a name in the register of protected designations of origin and protected geographical indications [Cozza di Scardovari (PDO)]. http://data.europa.eu/eli/reg_impl/2013/1200/oj

EU (2013b). Regulation (EU) No 1379/2013 of 11 December 2013 on the Common Organisation of the Markets in fishery and aquaculture products, amending Council Regulations (EC)No 1184/2006 and (EC) No 1224/2009 and repealing Council Regulation (EC) No 104/2000. 1-21. https://data.europa.eu/eli/reg/2013/1379/oj

EUMOFA. (2019). The EU Fish Market. https://doi.org/10.2771/168390

EUMOFA. (2022a). European Market Observatory for Fisheries and Aquaculture Products. Aquaculture data division. Yearly ad-hoc queries [online]. [Cited Wednesday, October 5th 2022]. https://www.eumofa.eu/en/ad-hoc-queries3

EUMOFA. (2022b). European Market Observatory for Fisheries and Aquaculture Products. Species Profile: Mussel. https://www.eumofa.eu/documents/20178/137160/Mussel_31-1.pdf

Ezgeta-Balic, D., Lojen, S., Dolenec, T., Rozic, P. Z., Dolenec, M., Najdek, M., and Peharda, M. (2014). Seasonal differences of stable isotope composition and lipid content in four bivalve species from the Adriatic Sea. *Marine Biology Research*, 10: 625-634. https://doi.org/10.1080/17451000.2013.833338

FAO. (2019). Bivalves report steady demand and high prices. http://fao.org/in-action/globefish/market-reports/resource-detail/en/c/1207892/

FAO. (2020a). FAO yearbook. Fishery and Aquaculture Statistics 2018/ FAO annuaire. Statistiques des pêches et de l'aquaculture 2018/ FAO anuario. Estadísticas de pesca y acuicultura 2018. http://www.fao.org/3/ca5495t/ca5495t.pdf

FAO. (2020b). The State of World Fisheries and Aquaculture 2020. Sustainability in action. https://doi.org/10.4060/ca9229en

FAO. (2022a). Global Aquaculture Production. Fisheries and Aquaculture Division [online]. [Cited Wednesday, October 5th 2022]. https://www.fao.org/fishery/en/collection/aquaculture

FAO. (2022b). The State of World Fisheries and Aquaculture 2022. Towards Blue Transformation. https://doi.org/10.4060/cc0461en

Feldhusen, F. (2000). The role of seafood in bacterial foodborne diseases. *Microbes and Infection*, 2: 1651-1660. https://doi.org/10.1016/s1286-4579(00)01321-6

Fernández-Tajes, J., Longa, A., García-Gil, J., Chiu, Y.-W., Huang, Y.-S., Méndez, J., and Lee, R.-S. (2011). Alternative PCR–RFLP methods for mussel *Mytilus* species identification. *European Food Research and Technology*, 233: 791-796. https://doi.org/10.1007/s00217-011-1574-x

Figueras, A. J. (2007). Biología y cultivo del mejillón (*Mytilus galloprovincialis*) en Galicia*, Consejo Superior de Investigaciones Científicas*.

Fonner, R., and Sylvia, G. (2015). Willingness to Pay for Multiple Seafood Labels in a Niche Market. *Marine Resource Economics*, 30: 51-70. https://doi.org/10.1086/679466

Fonseca, V. F., Duarte, I. A., Matos, A. R., Reis-Santos, P., and Duarte, B. (2022). Fatty acid profiles as natural tracers of provenance and lipid quality indicators in illegally sourced fish and bivalves. *Food Control*, 134: 108735. https://doi.org/10.1016/j.foodcont.2021.108735

Fox, M., Mitchell, M., Dean, M., Elliott, C., and Campbell, K. (2018). The seafood supply chain from a fraudulent perspective. *Food Security*, 10: 939-963. https://doi.org/10.1007/s12571-018-0826-z

Fraisse, C., Belkhir, K., Welch, J. J., and Bierne, N. (2015). Local interspecies introgression is the main cause of extreme levels of intraspecific differentiation in mussels. *Molecular Ecology*, 25: 269-286. https://doi.org/10.1111/mec.13299

Gaitán-Espitia, J. D., Quintero-Galvis, J. F., Mesas, A., and D'Elía, G. (2016). Mitogenomics of southern hemisphere blue mussels (Bivalvia: Pteriomorphia): Insights into the evolutionary characteristics of the *Mytilus* edulis complex. *Scientific Reports*, 6: 26853. https://doi.org/10.1038/srep26853

Galvez, J. F., Mejuto, J. C., and Simal-Gandara, J. (2018). Future challenges on the use of blockchain for food traceability analysis. *Trends in Analytical Chemistry*, 107: 222-232. https://doi.org/10.1016/j.trac.2018.08.011

Gardner, J. P., Zbawicka, M., Westfall, K. M., and Wenne, R. (2016). Invasive blue mussels threaten regional scale genetic diversity in mainland and remote offshore locations: the need for baseline data and enhanced protection in the Southern Ocean. *Global Change Biology*, 22: 3182-3195. https://doi.org/10.1111/gcb.13332

Gardner, J. P. A., Oyarzún, P. A., Toro, J. E., Wenne, R., and Zbawicka, M. (2021). Phylogeography of Southern Hemisphere Blue Mussels of the Genus *Mytilus*: Evolution, Biosecurity, Aquaculture and Food Labelling. pp. 139-228. https://doi.org/10.1201/9781003138846-3

Garmendia, L., Soto, M., Cajaraville, M. P., and Marigómez, I. (2010). Seasonality in cell and tissue-level biomarkers in *Mytilus galloprovincialis*: relevance for long-term pollution monitoring. *Aquatic Biology*, 9: 203-219. https://doi.org/10.3354/ab00245

Geller, J. B., Carlton, J. T., and Powers, D. A. (1994). PCR-based detection of mtDNA haplotypes of native and invading mussels on the northeastern Pacific coast: latitudinal pattern of invasion. *Marine Biology*, 119: 243-249. https://doi.org/10.1007/BF00349563

Ghidini, S., Varrà, M. O., and Zanardi, E. (2019). Approaching Authenticity Issues in Fish and Seafood Products by Qualitative Spectroscopy and Chemometrics. *Molecules*, 24: 1812. https://doi.org/10.3390/molecules24091812

Gilbey, J., Cauwelier, E., Coulson, M. W., Stradmeyer, L., Sampayo, J. N., Armstrong, A., Verspoor, E.*, et al.* (2016). Accuracy of Assignment of Atlantic Salmon (*Salmo salar* L.) to Rivers and Regions in Scotland and Northeast England Based on Single Nucleotide Polymorphism (SNP) Markers. *Plos One*, 11: e0164327. https://doi.org/10.1371/journal.pone.0164327

Giusti, A., Tosi, F., Tinacci, L., Guardone, L., Corti, I., Arcangeli, G., and Armani, A. (2020). Mussels (*Mytilus spp*.) products authentication: A case study on the Italian market confirms issues in species identification and arises concern on commercial names attribution. *Food Control*, 118: 107379. https://doi.org/10.1016/j.foodcont.2020.107379

Global Invasive Species Database. (2022). Species profile: *Mytilus galloprovincialis*. [Cited Wednesday, October 5th]. http://www.iucngisd.org/gisd/speciesname/Mytilus+galloprovincialis

Gomes, I., Peteiro, L. G., Albuquerque, R., Nolasco, R., Dubert, J., Swearer, S. E., and Queiroga, H. (2016). Wandering mussels: using natural tags to identify connectivity patterns among Marine Protected Areas. *Marine Ecology Progress Series*, 552: 159-176. https://doi.org/10.3354/meps11753

Gómez-Gutiérrez, A., Tolosa, I., Jover, E., Bayona, J. M., and Albaigés, J. (2011). Origin and vertical dynamics of particulate organic matter in a salt-wedge estuary, the Ebro Delta, inferred from lipid molecular markers and compound-specific isotope carbon analysis. *Marine Chemistry*, 126: 269-280. https://doi.org/10.1016/j.marchem.2011.06.006

Gopi, K., Mazumder, D., Sammut, J., and Saintilan, N. (2019a). Determining the provenance and authenticity of seafood: A review of current methodologies. *Trends in Food Science & Technology*, 91: 294-304. https://doi.org/10.1016/j.tifs.2019.07.010

Gopi, K., Mazumder, D., Sammut, J., Saintilan, N., Crawford, J., and Gadd, P. (2019b). Combined use of stable isotope analysis and elemental profiling to determine provenance of black tiger prawns (*Penaeus monodon*). *Food Control*, 95: 242-248. https://doi.org/10.1016/j.foodcont.2018.08.012

Gopi, K., Mazumder, D., Sammut, J., Saintilan, N., Crawford, J., and Gadd, P. (2019c). Isotopic and elemental profiling to trace the geographic origins of farmed and wild-caught Asian seabass (Lates calcarifer). *Aquaculture*, 502: 56-62. https://doi.org/10.1016/j.aquaculture.2018.12.012

Gosling, E. (2003). Bivalve Molluscs: Biology, Ecology and Culture*, Wiley-Blackwell*. https://doi.org/10.1002/9780470995532

Gosling, E., Doherty, S., and Howley, N. (2008). Genetic characterization of hybrid mussel (*Mytilus*) populations on Irish coasts. *Journal of the Marine Biological Association of the United Kingdom*, 88: 341-346. https://doi.org/10.1017/s0025315408000957

Gosling, E. M. (1992). Systematics and geographic distribution of *Mytilus*. *Developments in aquaculture and fisheries science*, 25: 1-20.

Graziano, S., Gullì, M., and Marmiroli, N. (2017). Development and validation of a SYBR-Green I Real-Time PCR test to detect bivalves including *Mytilus* species in foods. *International Journal of Food Science & Technology*, 52: 1567-1575. https://doi.org/10.1111/ijfs.13429

Grbin, D., Sabolic, I., Klobucar, G., Dennis, S. R., Srut, M., Bakaric, R., Bakovic, V.*, et al.* (2019). Biomarker response of Mediterranean mussels *Mytilus galloprovincialis* regarding environmental conditions, pollution impact and seasonal effects. *The Science of the total environment*, 694: 133470-133470. https://doi.org/10.1016/j.scitotenv.2019.07.276

Guo, J. Y., Zeng, E. Y., Wu, F. C., Meng, X. Z., Mai, B. X., and Luo, X. J. (2007). Organochlorine pesticides in seafood products from southern China and health risk assessment. *Environmental Toxicology and Chemistry*, 26: 1109-1115. https://doi.org/10.1897/06-446r.1

Gustafson, L., Showers, W., Kwak, T., Levine, J., and Stoskopf, M. (2007). Temporal and spatial variability in stable isotope compositions of a freshwater mussel: implications for biomonitoring and ecological studies. *Oecologia*, 152: 140-150. https://doi.org/10.1007/s00442-006-0633-7

Hagger, J. A., Jones, M. B., Lowe, D., Leonard, D. R., Owen, R., and Galloway, T. S. (2008). Application of biomarkers for improving risk assessments of chemicals under the Water Framework Directive: a case study. *Marine Pollution Bulletin*, 56: 1111-1118. https://doi.org/10.1016/j.marpolbul.2008.03.040

Hammer, O., Harper, D., and Ryan, P. (2001). PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica*, 4: 1-9.

Hassard, F., Sharp, J. H., Taft, H., LeVay, L., Harris, J. P., McDonald, J. E., Tuson, K.*, et al.* (2017). Critical Review on the Public Health Impact of Norovirus Contamination in Shellfish and the Environment: A UK Perspective. *Food and Environmental Virology*, 9: 123-141. https://doi.org/10.1007/s12560-017-9279-3

Helyar, S. J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M. I., Ogden, R., Limborg, M. T., Cariani, A.*, et al.* (2011). Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. *Molecular ecology resources*, 11 Suppl 1: 123-136. https://doi.org/10.1111/j.1755-0998.2010.02943.x

Herrmann, M. G., Durtschi, J. D., Wittwer, C. T., and Voelkerding, K. V. (2007). Expanded instrument comparison of amplicon DNA melting analysis for mutation scanning and genotyping. *Clinical Chemistry*, 53: 1544-1548. https://doi.org/10.1373/clinchem.2007.088120

Hilbish, T., Carson, E., Plante, J., Weaver, L., and Gilg, M. (2002). Distribution of *Mytilus edulis*, *M. galloprovincialis*, and their hybrids in open-coast populations of mussels in southwestern England. *Marine Biology*, 140: 137-142. https://doi.org/10.1007/s002270100631

Hilbish, T., Mullinax, A., Dolven, S., Meyer, A., Koehn, R., and Rawson, P. (2000). Origin of the antitropical distribution pattern in marine mussels (*Mytilus spp*.): routes and timing of transequatorial migration. *Marine Biology*, 136: 69–77. https://doi.org/10.1007/s002270050010

Hilbish, T. J., Brannock, P. M., Jones, K. R., Smith, A. B., Bullock, B. N., and Wethey, D. S. (2010). Historical changes in the distributions of invasive and endemic marine invertebrates are contrary to global warming predictions: the effects of decadal climate oscillations. *Journal of Biogeography*, 37: 423-431. https://doi.org/10.1111/j.1365-2699.2009.02218.x

Hixson, S. M. (2014). Fish Nutrition and Current Issues in Aquaculture: The Balance in Providing Safe and Nutritious Seafood, in an Environmentally Sustainable Manner. *Journal of Aquaculture Research and Development*, 5: 1-10. https://doi.org/10.4172/2155-9546.1000234

Honig, A., Etter, R., Pepperman, K., Morello, S., and Hannigan, R. (2020). Site and age discrimination using trace element fingerprints in the blue mussel, *Mytilus edulis*. *Journal of Experimental Marine Biology and Ecology*, 522. https://doi.org/10.1016/j.jembe.2019.151249

Hsu, M. C., Chen, K. W., Lo, H. J., Chen, Y. C., Liao, M. H., Lin, Y. H., and Li, S. Y. (2003). Species identification of medically important fungi by use of real-time LightCycler PCR. *Journal of Medical Microbiology*, 52. https://doi.org/10.1099/jmm.0.05302-0

Ibañez, C., Pont, D., and Prat, N. (1997). Characterization of the Ebre and Rhone estuaries: A basis for defining and classifying salt-wedge estuaries. *Limnology and Oceanography*, 42: 89-101. https://doi.org/10.4319/lo.1997.42.1.0089

Illumina (2013). 16S Metagenomic Sequencing Library Preparation: Preparing 16S Ribosomal RNA Gene Amplicons for the Illumina MiSeq System.

Inoue, K., Waite, J. H., Matsuoka, M., Odo, S., and Harayama, S. (1995). Interspecific variations in adhesive protein sequences of *Mytilus edulis*, *M. galloprovincialis*, and *M. trossulus*. *Biological Bulletin*, 189: 370-375. https://doi.org/10.2307/1542155

ISO/20813 (2019). Molecular biomarker analysis — Methods of analysis for the detection and identification of animal species in foods and food products (nucleic acid-based methods) — General requirements and definitions. https://iso.org/standard/69297.html

Jacquet, J. L., and Pauly, D. (2008). Trade secrets: Renaming and mislabeling of seafood. *Marine Policy*, 32: 309-318. https://doi.org/10.1016/j.marpol.2007.06.007

Jilberto, F., Araneda, C., and Larrain, M. A. (2017). High resolution melting analysis for identification of commercially-important *Mytilus* species. *Food Chem*, 229: 716-720. https://doi.org/10.1016/j.foodchem.2017.02.109

Jombart, T., and Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27: 3070-3071. https://doi.org/10.1093/bioinformatics/btr521

Jones, C. M., Palmer, M., and Schaffler, J. J. (2017). Beyond Zar: the use and abuse of classification statistics for otolith chemistry. *Journal of Fish Biology*, 90: 492-504. https://doi.org/10.1111/jfb.13051

Kang, X., Zhao, Y., Shang, D., Zhai, Y., Ning, J., Ding, H., and Sheng, X. (2020). Identification of the geographical origins of sea cucumbers in China: The application of stable isotope ratios and compositions of C, N, O and H. *Food Control*, 111: 107036. https://doi.org/10.1016/j.foodcont.2019.107036

Kappel, K., and Schröder, U. (2020). Difficulties in DNA barcoding-based authentication of snapper products due to ambiguous nucleotide sequences in public databases. *Food Control*, 118: 107348. https://doi.org/10.1016/j.foodcont.2020.107348

Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., Prodöhl, P. A., and O'Hara, R. B. (2013). diveRsity: An Rpackage for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, 4: 782-788. https://doi.org/10.1111/2041-210x.12067

Kelly, S., Heaton, K., and Hoogewerff, J. (2005). Tracing the geographical origin of food: The application of multi-element and multi-isotope analysis. *Trends in Food Science & Technology*, 16: 555-567. https://doi.org/10.1016/j.tifs.2005.08.008

Kijewski, T., Śmietanka, B., Zbawicka, M., Gosling, E., Hummel, H., and Wenne, R. (2011). Distribution of *Mytilus* taxa in European coastal areas as inferred from molecular markers. *Journal of Sea Research*, 65: 224-234. https://doi.org/10.1016/j.seares.2010.10.004

Kijewski, T., Wijsman, J. W. M., Hummel, H., and Wenne, R. (2009). Genetic composition of cultured and wild mussels *Mytilus* from The Netherlands and transfers from Ireland and Great Britain. *Aquaculture*, 287: 292-296. https://doi.org/10.1016/j.aquaculture.2008.10.048

Kim, H., Suresh Kumar, K., and Shin, K.-H. (2015). Applicability of stable C and N isotope analysis in inferring the geographical origin and authentication of commercial fish (Mackerel, Yellow Croaker and Pollock). *Food Chemistry*, 172: 523-527. https://doi.org/10.1016/j.foodchem.2014.09.058

King, G. M., Judd, C., Kuske, C. R., and Smith, C. (2012). Analysis of Stomach and Gut Microbiomes of the Eastern Oyster (*Crassostrea virginica*) from Coastal Louisiana, USA. *Plos One*, 7. https://doi.org/10.1371/journal.pone.0051475

Klein, R. T., Lohmann, K., and Thayer, C. W. (1996). Bivalve skeletons record sea-surface temperature and δ18O via Mg/Ca and 18O/16O ratios. *Geology*, 24: 415-418. https://doi.org/10.1130/0091-7613

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., and Glockner, F. O. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res*, 41: e1. 10.1093/nar/gks808

Larraín, M. A., Díaz, N. F., Lamas, C., Uribe, C., and Araneda, C. (2014). Traceability of mussel (*Mytilus chilensis*) in southern Chile using microsatellite molecular markers and assignment algorithms. Exploratory survey. *Food Research International*, 62: 104-110. https://doi.org/10.1016/j.foodres.2014.02.016

Larrain, M. A., Gonzalez, P., Perez, C., and Araneda, C. (2019). Comparison between single and multi-locus approaches for specimen identification in *Mytilus* mussels. *Scientific Reports*, 9: 19714. https://doi.org/10.1038/s41598-019-55855-8

Larrain, M. A., Zbawicka, M., Araneda, C., Gardner, J. P. A., and Wenne, R. (2018). Native and invasive taxa on the Pacific coast of South America: Impacts on aquaculture, traceability and biodiversity of blue mussels (*Mytilus spp*.). *Evolutionary Applications*, 11: 298-311. https://doi.org/10.1111/eva.12553

Le Nguyen, D. D., Ngoc, H. H., Dijoux, D., Loiseau, G., and Montet, D. (2008). Determination of fish origin by using 16S rDNA fingerprinting of bacterial communities by PCR-DGGE: An application on Pangasius fish from Viet Nam. *Food Control*, 19: 454-460. https://doi.org/10.1016/j.foodcont.2007.05.006

Leal, M. C., Pimentel, T., Ricardo, F., Rosa, R., and Calado, R. (2015). Seafood traceability: current needs, available tools, and biotechnological challenges for origin certification. *Trends in biotechnology*, 33: 331-336. https://doi.org/10.1016/j.tibtech.2015.03.003

Lee, S. Y., and Morton, B. (1985). The introduction of the Mediterranean mussel *Mytilus galloprovincialis* into Hong Kong. *Malacological review*, 18: 107-109.

Léopold, T. N., Maiwore, J., Hadjia, M. B., Loiseau, G., Montet, D., and Mbofung, C. (2010). Study of the microbial diversity of *Oreochromis niloticus* of three lakes of Cameroon by PCR-DGGE: Application to the determination of the geographical origin. *Food Control*, 21: 673–678. https://doi.org/10.1016/j.foodcont.2009.10.006

Li, L., Boyd, C. E., and Sun, Z. (2016). Authentication of fishery and aquaculture products by multi-element and stable isotope analysis. *Food chemistry*, 194: 1238-1244. https://doi.org/10.1016/j.foodchem.2015.08.123

Li, Y. F., Xu, J. K., Chen, Y. W., Ding, W. Y., Shao, A. Q., Liang, X., Zhu, Y. T.*, et al.* (2019). Characterization of Gut Microbiome in the Mussel *Mytilus galloprovincialis* in Response to Thermal Stress. *Frontiers in physiology*, 10: 1086. https://doi.org/10.3389/fphys.2019.01086

Liaw, A., and Wiener, M. (2002). Classification and Regression by RandomForest. *Forest*, 23.

Liénart, C., Savoye, N., Bozec, Y., Breton, E., Conan, P., David, V., Feunteun, E.*, et al.* (2017). Dynamics of particulate organic matter composition in coastal systems: A spatio-temporal study at multi-systems scale. *Progress in Oceanography*, 156: 221-239. https://doi.org/10.1016/j.pocean.2017.03.001

Lindley, J. (2020). Food Security Amidst Crime: Harm of Illegal Fishing and Fish Fraud on Sustainable Oceans. *In* The Palgrave Handbook of Climate Resilient Societies, pp. 1-19. *Springer International Publishing*, Cham. 10.1007/978-3-030-32811-5_127-1

Lins, D. M., Zbawicka, M., Wenne, R., Poćwierz-Kotus, A., Molina, J. R. A., Alves, L. P., and Rocha, R. M. (2021). Ecology and genetics of *Mytilus galloprovincialis*: A threat to bivalve aquaculture in southern Brazil. *Aquaculture*, 540: 736753. https://doi.org/10.1016/j.aquaculture.2021.736753

Lischer, H. E., and Excoffier, L. (2012). PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, 28: 298-299. https://doi.org/10.1093/bioinformatics/btr642

Liu, X., Teixeira, J. S., Ner, S., Ma, K. V., Petronella, N., Banerjee, S., and Ronholm, J. (2020). Exploring the Potential of the Microbiome as a Marker of the Geographic Origin of Fresh Seafood. *Frontiers in microbiology*, 11: 696. https://doi.org/10.3389/fmicb.2020.00696

Liu, Y., Ma, D.-h., Wang, X.-c., Liu, L.-p., Fan, Y.-x., and Cao, J.-x. (2015). Prediction of chemical composition and geographical origin traceability of Chinese export tilapia fillets products by near infrared reflectance spectroscopy. *LWT - Food Science and Technology*, 60: 1214-1218. https://doi.org/10.1016/j.lwt.2014.09.009

Lokmer, A., Goedknegt, M. A., Thieltges, D. W., Fiorentino, D., Kuenzel, S., Baines, J. F., and Wegner, K. M. (2016). Spatial and Temporal Dynamics of Pacific Oyster Hemolymph Microbiota across Multiple Scales. *Frontiers in Microbiology*, 7. https://doi.org/10.3389/fmicb.2016.01367

Lokmer, A., and Wegner, K. M. (2015). Hemolymph microbiome of Pacific oysters in response to temperature, temperature stress and infection. *The ISME Journal*, 9: 670-682. https://doi.org/10.1038/ismej.2014.160

Lopez, I., and Pardo, M. A. (2005). Application of relative quantification TaqMan real-time polymerase chain reaction technology for the identification and quantification of *Thunnus alalunga* and *Thunnus albacares*. *Journal of Agricultural and Food Chemistry*, 53: 4554-4560. https://doi.org/10.1021/jf0500841

Luque, G. M., and Donlan, C. J. (2019). The characterization of seafood mislabeling: A global meta-analysis. *Biological Conservation*, 236: 556-570. https://doi.org/10.1016/j.biocon.2019.04.006

Luu, K., Bazin, E., and Blum, M. G. B. (2016). Pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, 17. https://doi.org/10.1101/056135

Marali, S., Schöne, B. R., Mertz-Kraus, R., Griffin, S. M., Wanamaker, A. D., Matras, U., and Butler, P. G. (2017). Ba/Ca ratios in shells of *Arctica islandica* - Potential environmental proxy and crossdating tool. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 465: 347-361. https://doi.org/10.1016/j.palaeo.2015.12.018

Martinsohn, J. T., and Ogden, R. (2009). FishPopTrace—Developing SNP-based population genetic assignment methods to investigate illegal fishing. *Forensic Science International: Genetics Supplement Series*, 2: 294-296. https://doi.org/10.1016/j.fsigss.2009.08.108

McCutchan, J. H., and Lewis, W. M. (2001). Seasonal variation in stable isotope ratios of stream algae. *SIL Proceedings, 1922-2010*, 27: 3304-3307. https://doi.org/10.1080/03680770.1998.11902437

McDonald, J. H., Seed, R. B., and Koehn, R. K. (1991). Allozymes and morphometric characters of three species of *Mytilus* in the Northern and Southern Hemispheres. *Marine Biology*, 111: 323-333.

Meistertzheim, A.-L., Héritier, L., and Lejart, M. (2017). High-Resolution Melting of 18S rDNA sequences (18S-HRM) for discrimination of bivalve's species at early juvenile stage: application to a spat survey. *Marine Biology*, 164. 10.1007/s00227-017-3162-5

Michalek, K., Ventura, A., and Sanders, T. (2016). *Mytilus* hybridisation and impact on aquaculture: A minireview. *Marine genomics*, 27: 3-7. https://doi.org/10.1016/j.margen.2016.04.008

Milan, M., Maroso, F., Dalla Rovere, G., Carraro, L., Ferraresso, S., Patarnello, T., Bargelloni, L*., et al.* (2019). Tracing seafood at high spatial resolution using NGS-generated data and machine learning: Comparing microbiome versus SNPs. *Food Chemistry*, 286: 413-420. https://doi.org/10.1016/j.foodchem.2019.02.037

Molkentin, J., and Giesemann, A. (2010). Follow-up of stable isotope analysis of organic versus conventional milk. *Analytical and Bioanalytical Chemistry*, 398: 1493-1500. https://doi.org/10.1007/s00216-010-3995-y

Montecino, V., and Lange, C. B. (2009). The Humboldt Current System: Ecosystem components and processes, fisheries, and sediment studies. *Progress in Oceanography*, 83: 65-79. https://doi.org/10.1016/j.pocean.2009.07.041

Montes, I., Laconcha, U., Iriondo, M., Manzano, C., Arrizabalaga, H., and Estonba, A. (2017). Reduced Single Nucleotide Polymorphism Panels for Assigning Atlantic Albacore and Bay of Biscay Anchovy Individuals to Their Geographic Origin: Toward Sustainable Fishery Management. *Journal of Agricultural and Food Chemistry*, 65: 4351-4358. https://doi.org/10.1021/acs.jafc.7b00619

Montet, D., Le Nguyen, D. D., El Sheikha, A., Condur, A., Metayer, I., and Loiseau, G. (2008). Application of PCR-DGGE in determining food origin : cases studies of fish and fruits. pp. 11-22. Ed. by W. M. Eds by Richard Glass. *Aspects of Applied Biology*.

Morin, P. A., Luikart, G., Wayne, R. K., and group, S. w. (2004). SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution*, 19: 208-216. https://doi.org/10.1016/j.tree.2004.01.009

Morrison, L., Bennion, M., Gill, S., and Graham, C. T. (2019). Spatio-temporal trace element fingerprinting of king scallops (*Pecten maximus*) reveals harvesting period and location. *Science of the Total Environment*, 697. https://doi.org/10.1016/j.scitotenv.2019.134121

Musella, M., Wathsala, R., Tavella, T., Rampelli, S., Barone, M., Palladino, G., Biagi, E*., et al.* (2020). Tissue-scale microbiota of the Mediterranean mussel (*Mytilus*

galloprovincialis) and its relationship with the environment. *The Science of the total environment*, 717: 137209. https://doi.org/10.1016/j.scitotenv.2020.137209

Nielsen, E. E. (2016). Chapter 8 - Population or Point-of-Origin Identification. *In* Seafood Authenticity and Traceability, pp. 149-169. Ed. by A. M. Naaum, and R. H. Hanner. *Academic Press*, San Diego. https://doi.org/10.1016/B978-0-12-801592-6.00008-5

Nielsen, E. E., Cariani, A., Mac Aoidh, E., Maes, G. E., Milano, I., Ogden, R., Taylor, M.*, et al.* (2012). Gene-associated markers provide tools for tackling illegal fishing and false eco-certification. *Nature communications*, 3: 851. https://doi.org/10.1038/ncomms1845

NOAA. (2021). National Oceanic and Atmospheric Administration: Cold and warm episodes by season. https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php

OESA. (2017). Cultivo del mejillón (*Mytilus galloprovincialis*). 104 pp. https://www.observatorio-acuicultura.es/recursos/publicaciones/8-cultivo-del-mejillon-mytilus-galloprovincialis

Ogden, R. (2011). Unlocking the potential of genomic technologies for wildlife forensics. *Molecular ecology resources*, 11 Suppl 1: 109-116. https://doi.org/10.1111/j.1755-0998.2010.02954.x

Ogden, R., and Linacre, A. (2015). Wildlife forensic science: A review of genetic geographic origin assignment. *Forensic science international. Genetics*, 18: 152-159. https://doi.org/10.1016/j.fsigen.2015.02.008

Olsen, P., and Borit, M. (2013). How to define traceability. *Trends in Food Science & Technology*, 29: 142-150. https://doi.org/10.1016/j.tifs.2012.10.003

Ortea, I., and Gallardo, J. M. (2015). Investigation of production method, geographical origin and species authentication in commercially relevant shrimps using stable isotope ratio and/or multi-element analyses combined with chemometrics: an exploratory analysis. *Food Chemistry*, 170: 145-153. https://doi.org/10.1016/j.foodchem.2014.08.049

Ozdilek, S. Y., Demir, N., and Gurkan, S. E. (2019). Assessment of pollution biomarker and stable isotope data in *Mytilus galloprovincialis* tissues. *Environmental monitoring and assessment*, 191: 60. https://doi.org/10.1007/s10661-019-7189-6

Page, H., and Lastra, M. (2003). Diet of intertidal bivalves in the Ría de Arosa (NW Spain): Evidence from stable C and N isotope analysis. *Marine Biology*, 143: 519-532. https://doi.org/10.1007/s00227-003-1102-z

Pardo, M. A., and Pérez-Villareal, B. (2004). Identification of commercial canned tuna species by restriction site analysis of mitochondrial DNA products obtained by nested primer PCR. *Food Chemistry*, 86: 143-150. https://doi.org/10.1016/j.foodchem.2003.09.024

Paris, J. R., Stevens, J. R., and Catchen, J. M. (2017). Lost in parameter space: a road map for stacks. *Methods in Ecology and Evolution*, 8: 1360-1373. https://doi.org/10.1111/2041-210x.12775

Paterno, M., Bat, L., Souissi, J. B., Boscari, E., Chassanite, A., Congiu, L., Guarnieri, G.*, et al.* (2019). A Genome-Wide Approach to the Phylogeography of the Mussel *Mytilus galloprovincialis* in the Adriatic and the Black Seas. *Frontiers in Marine Science*, 6. https://doi.org/10.3389/fmars.2019.00566

Paulet, Y. M., Lorrain, A., Richard, J., and Pouvreau, S. (2006). Experimental shift in diet δ13C: A potential tool for ecophysiological studies in marine bivalves. *Organic Geochemistry*, 37: 1359-1370. http://doi.org/10.1016/j.orggeochem.2006.01.008

Peterson, B. J., and Fry, B. (1987). Stable isotopes in ecosystem studies. *Annual review of ecology and systematics*, 18: 293-320.

Pierce, M. L., and Evan Warda, J. E. (2019). Gut Microbiomes of the Eastern Oyster (*Crassostrea virginica*) and the Blue Mussel (*Mytilus edulis*): Temporal Variation and the Influence of Marine Aggregate-Associated Microbial Communities. *American Society For Microbiology*, 4: e00730-00719. https://doi.org/10.1128/mSphere.00730-19

Pierce, M. L., Ward, J. E., Holohan, B. A., Zhao, X., and Hicks, R. E. (2015). The influence of site and season on the gut and pallial fluid microbial communities of the eastern oyster, *Crassostrea virginica* (Bivalvia, Ostreidae): community-level physiological profiling and genetic structure. *Hydrobiologia*, 765: 97-113. 10.1007/s10750-015-2405-z

Pimentel, T., Marcelino, J., Ricardo, F., Soares, A., and Calado, R. (2017). Bacterial communities 16S rDNA fingerprinting as a potential tracing tool for cultured seabass *Dicentrarchus labrax*. *Scientific Reports*, 7. 10.1038/s41598-017-11552-y

Piry, S., Alapetite, A., Cornuet, J. M., Paetkau, D., Baudouin, L., and Estoup, A. (2004). GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity*, 95: 536-539. https://doi.org/10.1093/jhered/esh074

Popovic, I., Matias, A. M. A., Bierne, N., and Riginos, C. (2020). Twin introductions by independent invader mussel lineages are both associated with recent admixture with a native congener in Australia. *Evol Appl*, 13: 515-532. 10.1111/eva.12857

Post, D. M. (2002). Using Stable Isotopes to Estimate Trophic Position: Models, Methods and Assumptions. *Ecology*, 83: 703-718. https://doi.org/10.1890/0012-9658(2002)083[0703:USITET]2.0.CO;2

Potasman, I., Paz, A., and Odeh, M. (2002). Infectious Outbreaks Associated with Bivalve Shellfish Consumption: A Worldwide Perspective. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 35: 921-928. https://doi.org/10.1086/342330

Poulain, C., Gillikin, D. P., Thébault, J., Munaron, J. M., Bohn, M., Robert, R., Paulet, Y. M.*, et al.* (2015). An evaluation of Mg/Ca, Sr/Ca, and Ba/Ca ratios as environmental proxies in aragonite bivalve shells. *Chemical Geology*, 396: 42-50. https://doi.org/10.1016/j.chemgeo.2014.12.019

Prendergast, A. L., Versteegh, E. A. A., and Schöne, B. R. (2017). New research on the development of high-resolution palaeoenvironmental proxies from geochemical properties of biogenic carbonates. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 484: 1-6. https://doi.org/10.1016/j.palaeo.2017.05.032

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J.*, et al.* (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics*, 81: 559-575. https://doi.org/10.1086/519795

Purchon, R. D. (1968). The biology of the mollusca. International series of monographs in pure and applied biology Zoology*., Pergamon Press*, Oxford.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J.*, et al.* (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*, 41: D590-D596. https://doi.org/10.1093/nar/gks1219

Quesada, H., Beynon, C. M., and Skibinski, D. O. (1995a). A mitochondrial DNA discontinuity in the mussel *Mytilus galloprovincialis Lmk*: pleistocene vicariance biogeography and secondary intergradation. *Molecular biology and evolution*, 12: 521-524. https://doi.org/10.1093/oxfordjournals.molbev.a040227

Quesada, H., Zapata, C., and Alvarez, G. (1995b). A multilocus allozyme discontinuity in the mussel *Mytilus galloprovincialis*: the interaction of ecological and life-history factors. *Marine Ecology Progress Series*, 116: 99-115. https://doi.org/10.3354/meps116099

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rannala, B., and Mountain, J. L. (1997). Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America*, 94: 9197-9201. https://doi.org/10.1073/pnas.94.17.9197

Raymond, M., and Rousset, F. (1995). GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. *Journal of Heredity*, 86: 248-249. https://doi.org/10.1093/oxfordjournals.jhered.a111573

Ricardo, F., Genio, L., Costa Leal, M., Albuquerque, R., Queiroga, H., Rosa, R., and Calado, R. (2015). Trace element fingerprinting of cockle (*Cerastoderma edule*) shells can reveal harvesting location in adjacent areas. *Scientific Reports*, 5: 11932. https://doi.org/10.1038/srep11932

Ricardo, F., Maciel, E., Domingues, R., and Calado, R. (2017a). Spatio-temporal variability in the fatty acid profile of the adductor muscle of the common cockle *Cerastoderma edule* and its relevance for tracing geographic origin. *Food Control*, 81: 173-180. https://doi.org/10.1016/j.foodcont.2017.06.005

Ricardo, F., Pimentel, T., Genio, L., and Calado, R. (2017b). Spatio-temporal variability of trace elements fingerprints in cockle (*Cerastoderma edule*) shells and its relevance for tracing geographic origin. *Scientific Reports*, 7. https://doi.org/10.1038/s41598-017-03381-w

Ricardo, F., Pimentel, T., Maciel, E., Moreira, A. S. P., Domingues, M. R., and Calado, R. (2017c). Fatty acid dynamics of the adductor muscle of live cockles (*Cerastoderma edule*) during their shelf-life and its relevance for traceability of geographic origin. *Food Control*, 77: 192-198. https://doi.org/10.1016/j.foodcont.2017.01.012

Riera, P. (2007). Trophic subsidies of *Crassostrea gigas*, *Mytilus edulis* and *Crepidula fornicata* in the Bay of Mont Saint Michel (France): A δ13C and δ15N investigation. *Estuarine, Coastal and Shelf Science*, 72: 33-41. https://doi.org/10.1016/j.ecss.2006.10.002

Rodriguez-Manzano, J., Hundesa, A., Calgua, B., Carratala, A., Maluquer de Motes, C., Rusinol, M., Moresco, V*., et al.* (2014). Adenovirus and Norovirus Contaminants in Commercially Distributed Shellfish. *Food and Environmental Virology*, 6: 31-41. https://doi.org/10.1007/s12560-013-9133-1

Rodríguez-Ezpeleta, N., Díaz-Arce, N., Walter, J. F., Richardson, D. E., Rooker, J. R., Nøttestad, L., Hanke, A. R*., et al.* (2019). Determining natal origin for improved management of Atlantic bluefin tuna. *Frontiers in Ecology and the Environment*, 17: 439-444. https://doi.org/10.1002/fee.2090

Romalde, J. L., Rivadulla, E., Varela, M. F., and Barja, J. L. (2018). An overview of 20 years of studies on the prevalence of human enteric viruses in shellfish from Galicia, Spain. *Journal of applied microbiology*, 124: 943-957. https://doi.org/10.1111/jam.13614

Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, 8: 103-106. https://doi.org/10.1111/j.1471-8286.2007.01931.x

Rubiolo, J. A., Lozano-Leon, A., Rodriguez-Souto, R., Fol Rodriguez, N., Vieytes, M. R., and Botana, L. M. (2018). The impact of depuration on mussel hepatopancreas bacteriome composition and predicted metagenome. *Antonie van Leeuwenhoek*, 111: 1117-1129. https://doi.org/10.1007/s10482-018-1015-y

Rumolo, P., Carannante, M., Gherardi, S., Tamburrino, S., Vallefuoco, M., Migliaccio, R., D'Argenzio, C.*, et al.* (2017). Measuring Stable Carbon and Nitrogen Isotopes in *Mytilus galloprovincialis* to Elucidate the Sources of Organic Matter in Three Different Nearshore Marine Environments. *Journal of Coastal Research*, 33: 1172-1181. https://doi.org/10.2112/jcoastres-d-16-00079.1

Sant'Ana, L. S., Ducatti, C., and Ramires, D. G. (2010). Seasonal variations in chemical composition and stable isotopes of farmed and wild Brazilian freshwater fish. *Food Chemistry*, 122: 74-77. https://doi.org/10.1016/j.foodchem.2010.02.016

Santaclara, F. J., Espineira, M., Cabado, G., Aldasoro, A., Gonzalez-Lavin, N., and Vieites, J., M. (2006). Development of a method for the genetic identification of mussel species belonging to *Mytilus*, *Perna*, *Aulacomya*, and other genera. *Journal of Agricultural and Food Chemistry*, 54: 8461-8470. https://doi.org/10.1021/jf061400u

Scarano, D., and Rao, R. (2014). DNA Markers for Food Products Authentication. *Diversity*, 6: 579-596. https://doi.org/10.3390/d6030579

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A.*, et al.* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology*, 75: 7537-7541. https://doi.org/10.1128/AEM.01541-09

Schröder, U. (2007). Challenges in the Traceability of Seafood. *Journal of Consumer Protection and Food Safety*, 3: 45-48. https://doi.org/10.1007/s00003-007-0302-8

Seed, R., and Suchanek, T. (1992). The Mussel *Mytilus*: Ecology, Physiology, Genetics and Culture. *Development in Aquaculture Fisheries Science*, 25: 87-157.

Simenstad, C., and Wissmar, R. C. (1985). $\delta^{13}C$ evidence of the origins and fates of organic carbon in estuarine and near-shore food webs. *Marine Ecology*, 22: 141-152. https://doi.org/10.3354/meps022141

Smaal, A. C. (2002). European mussel cultivation along the Atlantic coast: Production status, problems and perspectives. *Hydrobiologia*, 484. https://doi.org/10.1023/A:1021352904712

Somes, C., Schmittner, A., Galbraith, E., Lehmann, M., Altabet, M., Montoya, J., Letelier, R.*, et al.* (2010). Simulating the global distribution of nitrogen isotopes in the ocean. *Global Biogeochemical Cycles*, 24: GB4019. https://doi.org/10.1029/2009gb003767

Strasser, C. A., Mullineaux, L. S., and Thorrold, S. R. (2008). Temperature and salinity effects on elemental uptake in the shells of larval and juvenile softshell clams *Mya arenaria*. *Marine Ecology Progress Series*, 370: 155-169. https://doi.org/10.3354/meps07658

Suh, Y. J., and Shin, K.-H. (2013). Size-related and seasonal diet of the manila clam (*Ruditapes philippinarum*), as determined using dual stable isotopes. *Estuarine, Coastal and Shelf Science*, 135: 94-105. https://doi.org/10.1016/j.ecss.2013.06.029

Tarifeño, E., Galleguillos, R., Llanos-Rivera, A., Arriagada, D., Ferrada, S., Canales-Aguirre, C. B., and Seguel, M. (2012). Erroneous identification of the mussel, *Mytilus galloprovincialis* (Lamarck 1819) as the specie, *Mytilus chilensis* (Hupe 1854) in the

Bay of Concepcion, Chile. *Gayana (Concepción)*, 76: 167-172. https://doi.org/10.4067/s0717-65382012000300010

Thiel, M., Macaya, E. C., Acuña, E., Arntz, W. E., Bastias, H., Brokordt, K., Camus, P. A., *et al.* (2007). The humboldt current system of northern and central Chile: Oceanographic processes, ecological interactions and socioeconomic feedback. *Oceanography and Marine Biology*, 45: 195-344. https://doi.org/10.1201/9781420050943.ch6

Thomas, A. C., Carr, M., and Strub, P. (2001). Chlorophyll variability in eastern boundary currents. *Geophysical Research Letters*, 28. https://doi.org/10.1029/2001GL013368

Titus, K., Mosher, J. A., and Williams, B. K. (1984). Chance-corrected Classification for Use in Discriminant Analysis: Ecological Applications. *The American Midland Naturalist*, 111: 1-7. https://doi.org/10.2307/2425535

Toro, J. E., Ojeda, J. A., Vergara, A. M., Castro, G. C., and Alcapan, A. C. (2005). Molecular characterization of the Chilean blue mussel (*Mytilus chilensis* Hupe 1854) demonstrates evidence for the occurrence of *Mytilus galloprovincialis* in southern Chile. *Journal of Shellfish Research*, 24: 1117-1121. https://doi.org/10.2983/0730-8000(2005)24[1117:MCOTCB]2.0.CO;2

Trueman, C. N., and St John Glew, K. (2019). Chapter 6 - Isotopic Tracking of Marine Animal Movement. *In* Tracking Animal Migration with Stable Isotopes (Second Edition), pp. 137-172. Ed. by K. A. Hobson, and L. I. Wassenaar. *Academic Press*. https://doi.org/10.1016/B978-0-12-814723-8.00006-4

Vezzulli, L., Stagnaro, L., Grande, C., Tassistro, G., Canesi, L., and Pruzzo, C. (2018). Comparative 16SrDNA Gene-Based Microbiota Profiles of the Pacific Oyster (*Crassostrea gigas*) and the Mediterranean Mussel (*Mytilus galloprovincialis*) from a Shellfish Farm (Ligurian Sea, Italy). *Microbial Ecology*, 75: 495-504. https://doi.org/10.1007/s00248-017-1051-6

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and environmental microbiology*, 73: 5261-5267. https://doi.org/10.1128/AEM.00062-07

Wenne, R. (2018). Single nucleotide polymorphism markers with applications in aquaculture and assessment of its impact on natural populations. *Aquatic Living Resources*, 31: 2. https://doi.org/10.1051/alr/2017043

Westfall, K. M., and Gardner, J. P. A. (2010). Genetic diversity of Southern hemisphere blue mussels (Bivalvia: Mytilidae) and the identification of non-indigenous taxa. *Biological Journal of the Linnean Society*, 101: 898-909. https://doi.org/10.1111/j.1095-8312.2010.01549.x

Westfall, K. M., and Gardner, J. P. A. (2013). Interlineage *Mytilus galloprovincialis Lmk*. 1819 hybridization yields inconsistent genetic outcomes in the Southern hemisphere. *Biological Invasions*, 15: 1493-1506. https://doi.org/10.1007/s10530-012-0385-8

Wilkins, N. P., Fujino, K., and Gosling, E. M. (2008). The mediterranean mussel *Mytilus galloprovincialis Lmk*. in Japan. *Biological Journal of the Linnean Society*, 20: 365-374. https://doi.org/10.1111/j.1095-8312.1983.tb01597.x

Winder, L., Phillips, C., Richards, N., Ochoa-Corona, F., Hardwick, S., Vink, C. J., and Goldson, S. (2011). Evaluation of DNA melting analysis as a tool for species identification. *Methods in Ecology and Evolution*, 2: 312-320. https://doi.org/10.1111/j.2041-210X.2010.00079.x

World Register of Marine Species. (2022). MolluscaBase. *Mytilus* Linnaeus, 1758. [Date accessed 2022-10-31]. https://www.marinespecies.org/aphia.php?p=taxdetails&id=138228

Wright, A. C., Fan, Y., and Baker, G. L. (2018). Nutritional Value and Food Safety of Bivalve Molluscan Shellfish. *Journal of Shellfish Research*, 37: 695-708, 614.

Yu, D., Chen, Z., Zhang, R., and Yin, W. (2005). Real-time qualitative PCR for the inspection and identification of *Batrocera philippinensis* and *Batrocera occipitalis* (Diptera: Tephritidae) using SYBR Green assay. *The Raffles Bulletin of Zoology*, 53: 73-78.

Zardi, G. I., McQuaid, C. D., Jacinto, R., Lourenço, C. R., Serrão, E. A., and Nicastro, K. R. (2018). Re-assessing the origins of the invasive mussel *Mytilus galloprovincialis* in southern Africa. *Marine and Freshwater Research*, 69: 607. https://doi.org/10.1071/mf17132

Zbawicka, M., Drywa, A., Śmietanka, B., and Wenne, R. (2012). Identification and validation of novel SNP markers in European populations of marine *Mytilus* mussels. *Marine Biology*, 159: 1347-1362. https://doi.org/10.1007/s00227-012-1915-8

Zbawicka, M., Gardner, J. P. A., and Wenne, R. (2019). Cryptic diversity in smooth-shelled mussels on Southern Ocean islands: connectivity, hybridisation and a marine invasion. *Frontiers in Zoology*, 16: 32. https://doi.org/10.1186/s12983-019-0332-y

Zbawicka, M., Sanko, T., Strand, J., and Wenne, R. (2014). New SNP markers reveal largely concordant clinal variation across the hybrid zone between *Mytilus spp*. in the Baltic Sea. *Aquatic Biology*, 21: 25-36. https://doi.org/10.3354/ab00566

Zbawicka, M., Trucco, M. I., and Wenne, R. (2018a). Single nucleotide polymorphisms in native South American Atlantic coast populations of smooth shelled mussels: hybridization with invasive European *Mytilus galloprovincialis*. *Genet Sel Evol*, 50: 5. https://doi.org/10.1186/s12711-018-0376-z

Zbawicka, M., Trucco, M. I., and Wenne, R. (2018b). Single nucleotide polymorphisms in native South American Atlantic coast populations of smooth shelled mussels: hybridization with invasive European *Mytilus galloprovincialis*. *Genetics Selection Evolution*, 50. https://doi.org/10.1186/s12711-018-0376-z

Zhang, W., Li, R., Chen, X., Wang, C., Gu, Z., Mu, C., Song, W.*, et al.* (2019). Molecular identification reveals hybrids of *Mytilus coruscus* × *Mytilus galloprovincialis* in mussel hatcheries of China. *Aquaculture International*, 28: 85-93. https://doi.org/10.1007/s10499-019-00445-8

Zhao, L. Q., Milano, S., Tanaka, K., Liang, J., Deng, Y. W., Yang, F., Walliser, E. O.*, et al.* (2020). Trace elemental alterations of bivalve shells following transgenerational exposure to ocean acidification: Implications for geographical traceability and environmental reconstruction. *Science of the Total Environment*, 705. https://doi.org/10.1016/j.scitotenv.2019.135501

Zhao, X., Liu, Y., Wang, G., Tao, W., Lou, Y., Li, N., and Liu, Y. (2019). Tracing the geographical origins of Yesso scallop (*Patinopecten yessoensis*) by using compound-specific isotope analysis: An approach for overcoming the seasonal effect. *Food Control*, 102: 38-45. https://doi.org/10.1016/j.foodcont.2019.03.016

APPENDIX A

Figure S2.1. Detailed schematic view of the work-flow done in this study. First, thousands of SNPs were discovered and filtered on 222 mussel samples using RAD-seq. Then, population structure and traceability analysis were done.

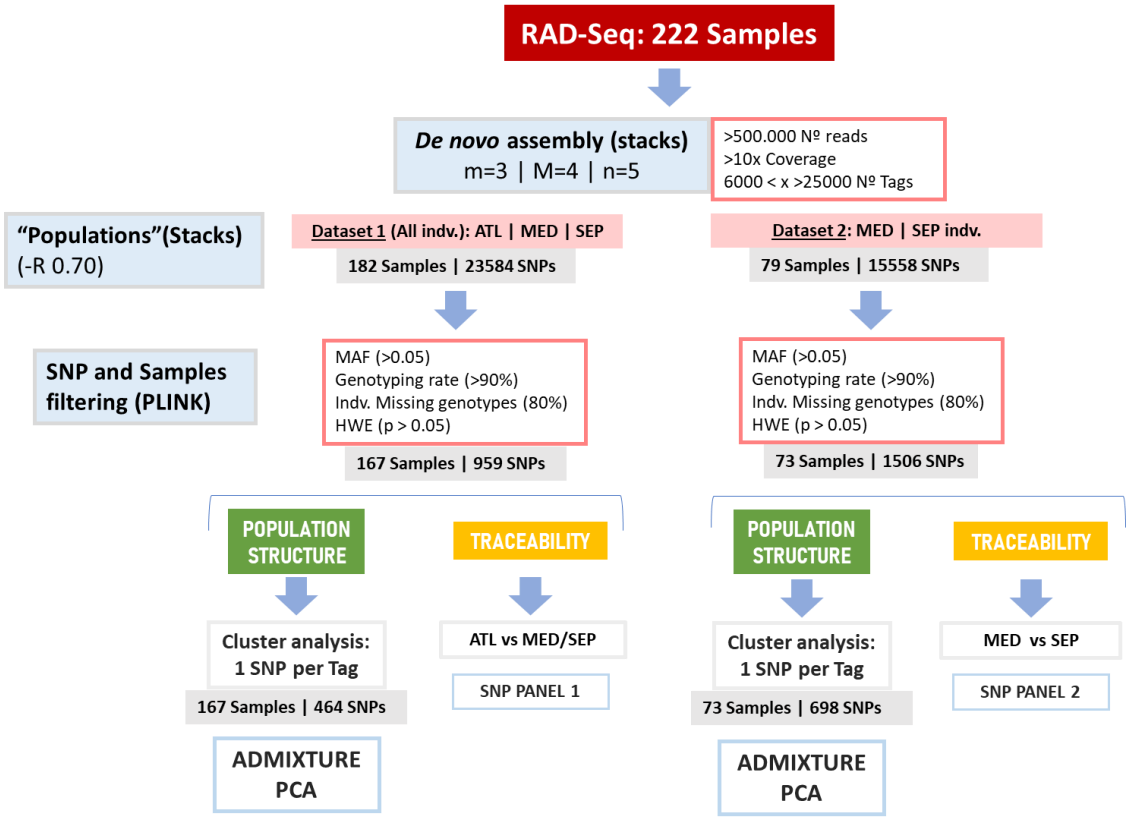


Figure S2.2. Origin assignment of individuals following THL method. (a) Progression of the percentage of correctly, incorrectly and unassigned Atlantic (ATL), Mediterranean (MED) and South-eastern Pacific (SEP) samples as the number of markers (SNPs) increases with an assignment threshold of 90%.

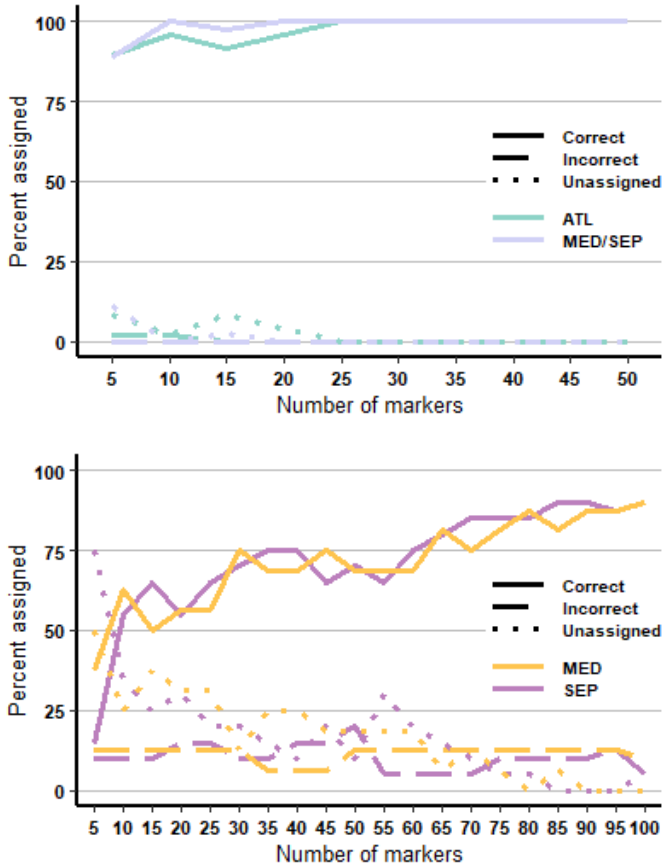
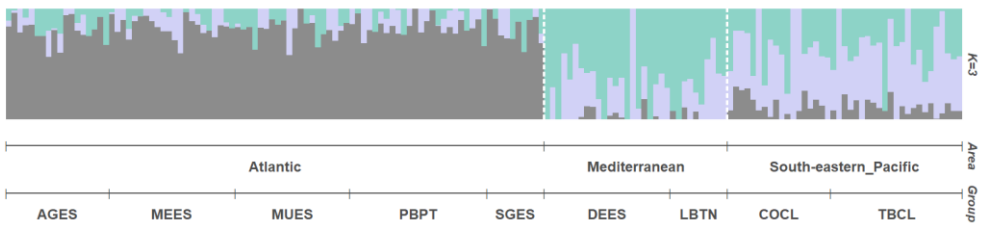


Figure S2.3. Graphical representation of ADMIXTURE clustering approach, where each bar represents an individual and each colour its inferred membership to each of the potential ancestral populations (K=3). (A) Dataset 1 including all areas: Atlantic (MUES|MEES|SGES|AGES |PBPT), Mediterranean (DEES|LBTN) and South-eastern Pacific (COCL|TBCL). (B) Dataset 2 including only the Mediterranean (DEES|LBTN) and South-eastern Pacific individuals (COCL|TBCL).



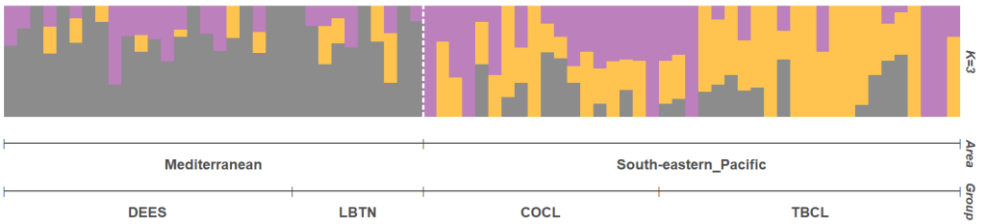


Table S2.1. Samples processed with RADSeq. Only the 182 samples from the 222 original samples that passed the Stacks quality filters are shown.

Sample	Location	Nº Tot Reads (process_ radtags)	Nº p. Reads (clone_filter)	% Remv. Clone Reads	Coverages	Tags (ustacks)	SNPs (ustacks)
PBPT_392	**Algarve**	5076786	1189177	**28.20%**	**89.55**	10366	4291
PBPT_393	**Algarve**	4044936	1017075	**22.56%**	**95.24**	9059	3755
PBPT_395	**Algarve**	2438848	794023	**21.37%**	**48.77**	13009	5038
PBPT_396	**Algarve**	2565864	787512	**22.95%**	**53.4**	11665	4561
PBPT_397	**Algarve**	3149866	829538	**22.79%**	**67.03**	9917	4023
PBPT_398	**Algarve**	3890548	1042733	**23.76%**	**63.34**	11465	4865
PBPT_399	**Algarve**	5033808	1311416	**32.76%**	**91.88**	11392	3825
PBPT_400	**Algarve**	5975144	1317016	**34.94%**	**95.52**	10897	4073
PBPT_401	**Algarve**	2137766	541723	**22.33%**	**41.21**	10256	3988
PBPT_402	**Algarve**	4539428	1040057	**23.94%**	**83.2**	10327	3657
PBPT_403	**Algarve**	4339019	1135448	**27.20%**	**76.95**	11494	4338
PBPT_404	**Algarve**	4290582	1341511	**26.69%**	**106.96**	10439	3614
PBPT_405	**Algarve**	3825244	1043543	**27.46%**	**87.46**	9800	3831
PBPT_406	**Algarve**	4605500	1416486	**28.99%**	**107.5**	10732	3913
PBPT_407	**Algarve**	5008234	1327133	**25.06%**	**51.21**	17491	5965
PBPT_408	**Algarve**	5123475	1273641	**26.07%**	**86.37**	11500	4117
PBPT_409	**Algarve**	4828978	1048483	**35.14%**	**55.24**	14189	4945
PBPT_410	**Algarve**	5500180	1347509	**35.72%**	**90.6**	11469	4499
PBPT_411	**Algarve**	3430824	629931	**32.54%**	**25.66**	12995	4759
PBPT_412	**Algarve**	3275373	864301	**31.55%**	**30.64**	20917	7358
PBPT_413	**Algarve**	3821709	960461	**33.23%**	**48.46**	15104	5212
PBPT_414	**Algarve**	5328508	1271723	**32.53%**	**63.69**	14654	5192
PBPT_415	**Algarve**	5008306	1309433	**33.89%**	**84.29**	12213	4320
PBPT_416	**Algarve**	4903033	1239606	**35.47%**	**92.11**	10939	4220
MEES_242	**Basque C.**	2366453	525078	**24.38%**	**30.61**	11897	4430
MEES_243	**Basque C.**	3537453	853199	**25.93%**	**52.23**	12674	4605
MEES_245	**Basque C.**	4658669	1242767	**26.94%**	**61.92**	13976	4725
MEES_246	**Basque C.**	3532914	914533	**25.32%**	**47.79**	15255	5481
MEES_247	**Basque C.**	5033866	1426457	**27.03%**	**59.66**	17192	5859
MEES_249	**Basque C.**	4209427	959527	**26.42%**	**69.21**	10923	4212
MEES_251	**Basque C.**	4535427	1186712	**27.08%**	**94.83**	10393	3719
MEES_252	**Basque C.**	2128824	672319	**22.40%**	**27.91**	19088	6447
MEES_253	**Basque C.**	4447828	941124	**22.27%**	**35.52**	20188	7088
MEES_255	**Basque C.**	2629192	548175	**20.10%**	**36**	10317	4326

MEES_257	**Basque C.**	6003018	1491087	**29.48%**	**112.74**	10543	4179
MEES_261	**Basque C.**	3431588	812516	**24.08%**	**56.25**	10344	4249
MEES_262	**Basque C.**	1466463	427133	**18.16%**	**16.58**	19889	6824
MEES_264	**Basque C.**	3604647	1104573	**24.57%**	**88.38**	9949	4114
MEES_268	**Basque C.**	4181578	1093416	**22.54%**	**44.85**	19251	7084
MEES_269	**Basque C.**	3112673	767195	**23.67%**	**23.77**	22559	8398
MEES_271	**Basque C.**	2821467	721644	**19.97%**	**32.73**	17025	6125
MEES_273	**Basque C.**	1194606	330749	**18.54%**	**13.33**	15339	5400
MEES_274	**Basque C.**	1626622	396218	**17.72%**	**28.22**	10698	4129
MEES_276	**Basque C.**	6367932	1405710	**24.86%**	**57.28**	15317	5850
MEES_279	**Basque C.**	4519576	854699	**22.22%**	**40.11**	14022	5357
MEES_280	**Basque C.**	5809026	1348610	**23.99%**	**96.95**	11062	4249
MEES_283	**Basque C.**	7133295	1776703	**25.91%**	**124.65**	11401	4122
MUES_203	**Basque C.**	2811992	589233	**24.01%**	**43.25**	10262	4090
MUES_205	**Basque C.**	4672434	1476083	**29.38%**	**57.83**	18061	7442
MUES_207	**Basque C.**	4647581	1201051	**29.50%**	**98.85**	10183	3958
MUES_208	**Basque C.**	998147	142289	**20.36%**	**12.19**	6665	2018
MUES_209	**Basque C.**	2343322	623443	**25.14%**	**47.39**	10397	3983
MUES_211	**Basque C.**	1243622	267733	**21.37%**	**22.65**	8681	3194
MUES_214	**Basque C.**	1546297	465872	**21.48%**	**34.12**	10192	3885
MUES_215	**Basque C.**	8274557	2444309	**31.48%**	**97.39**	19863	6507
MUES_216	**Basque C.**	4644428	1003710	**24.34%**	**75.07**	10615	4011
MUES_217	**Basque C.**	580148	188337	**18.51%**	**14.23**	7279	2611
MUES_218	**Basque C.**	1499803	437863	**23.09%**	**37.54**	9072	3548
MUES_220	**Basque C.**	1801063	520975	**20.30%**	**39.71**	9705	3823
MUES_221	**Basque C.**	2872767	813519	**22.07%**	**68.91**	9694	3928
MUES_222	**Basque C.**	3905079	1274891	**24.93%**	**101.63**	10215	4016
MUES_223	**Basque C.**	5092698	1392098	**24.98%**	**104.61**	10875	3980
MUES_224	**Basque C.**	5131080	1193966	**26.69%**	**74.82**	11485	4655
MUES_225	**Basque C.**	3516755	1096663	**26.05%**	**48.88**	18143	6188
MUES_228	**Basque C.**	3961444	969150	**34.87%**	**37.26**	15948	6627
MUES_230	**Basque C.**	3218700	722806	**26.92%**	**60.23**	9716	3708
MUES_231	**Basque C.**	2106803	422775	**25.55%**	**30.74**	9866	3788
MUES_232	**Basque C.**	2220780	563554	**24.82%**	**36.93**	10491	3632
MUES_234	**Basque C.**	4372284	1076073	**28.01%**	**92.62**	9629	3647
MUES_236	**Basque C.**	14602549	4250456	**35.24%**	**174.11**	16297	4307
MUES_239	**Basque C.**	4595005	1118768	**28.75%**	**63.19**	13771	4506
COCL_476	**BioBío**	4300022	933916	**43.50%**	**34.6**	21661	6818
COCL_477	**BioBío**	1602702	463562	**27.39%**	**40.22**	9237	3677
COCL_481	**BioBío**	1786698	560532	**24.98%**	**21.41**	19898	6984
COCL_482	**BioBío**	1933173	614277	**24.86%**	**23.97**	19245	6873
COCL_483	**BioBío**	4838590	1218300	**44.55%**	**49.47**	19834	6440
COCL_485	**BioBío**	3734175	1272596	**29.05%**	**89.84**	11908	4172
COCL_487	**BioBío**	4293696	1055084	**46.51%**	**40.96**	21120	6917
COCL_488	**BioBío**	3090198	962221	**29.22%**	**71.58**	10902	4099
COCL_489	**BioBío**	3482892	797955	**45.75%**	**39.22**	16233	5183
COCL_490	**BioBío**	4392205	1107772	**46.42%**	**41.41**	20966	7171
COCL_493	**BioBío**	6744346	1527296	**47.98%**	**69.36**	16646	5015
COCL_494	**BioBío**	2140534	612151	**26.18%**	**24.73**	18521	6383
COCL_495	**BioBío**	1730823	560582	**25.30%**	**28.89**	14934	5233
COCL_496	**BioBío**	3061101	889055	**33.62%**	**55.11**	12613	4957
COCL_497	**BioBío**	1929633	601031	**25.68%**	**50.21**	9618	3936
COCL_498	**BioBío**	1934552	651722	**27.25%**	**56.47**	9559	3892

COCL_499	**BioBío**	1998765	592566	**26.38%**	**32.25**	14081	5155
COCL_500	**BioBío**	2296323	585315	**32.00%**	**44.62**	9664	3999
TBCL_451	**BioBío**	1469661	505096	**21.11%**	**44.44**	9468	3772
TBCL_452	**BioBío**	1818186	670103	**19.32%**	**47.84**	11456	4174
TBCL_453	**BioBío**	2346459	822182	**21.96%**	**30.81**	21670	7633
TBCL_454	**BioBío**	2378957	760812	**23.03%**	**66.67**	9444	3857
TBCL_455	**BioBío**	1892631	676038	**22.12%**	**33.79**	15910	5793
TBCL_456	**BioBío**	2326966	820996	**21.48%**	**27.75**	24220	8124
TBCL_457	**BioBío**	2408368	836572	**22.31%**	**40.46**	17101	5823
TBCL_458	**BioBío**	2736666	1008432	**22.44%**	**41.17**	20208	6984
TBCL_459	**BioBío**	2454619	845776	**21.86%**	**64.65**	11088	4247
TBCL_460	**BioBío**	2287035	821498	**23.62%**	**76.89**	9103	3669
TBCL_461	**BioBío**	1808420	631793	**21.33%**	**38.31**	13652	5034
TBCL_462	**BioBío**	1899030	674253	**20.56%**	**30.57**	18045	6321
TBCL_463	**BioBío**	7421246	2459512	**30.84%**	**116.77**	16985	5256
TBCL_464	**BioBío**	3116468	1050171	**24.17%**	**82.82**	10739	4036
TBCL_465	**BioBío**	2278639	798351	**20.63%**	**35.02**	18690	6229
TBCL_466	**BioBío**	2874081	1040986	**20.57%**	**44.97**	19439	6830
TBCL_467	**BioBío**	5965295	1913493	**29.85%**	**91.97**	17276	5298
TBCL_468	**BioBío**	1895194	633493	**22.42%**	**48.12**	10918	4241
TBCL_469	**BioBío**	3677087	1282880	**25.18%**	**45.48**	23396	7637
TBCL_470	**BioBío**	2020598	717422	**23.17%**	**45.13**	13224	4557
TBCL_471	**BioBío**	3510663	1178440	**26.16%**	**64.27**	15350	4951
TBCL_472	**BioBío**	2561141	913386	**24.82%**	**47.39**	16110	5325
TBCL_473	**BioBío**	3143634	1083438	**25.84%**	**62.71**	14634	5015
TBCL_475	**BioBío**	2257935	762781	**21.33%**	**27.87**	21921	7507
DEES_293	**Catalonia**	8059690	1992804	**28.84%**	**76.09**	19235	5924
DEES_294	**Catalonia**	4224198	1142882	**27.43%**	**89.63**	10576	4008
DEES_295	**Catalonia**	3090298	931636	**26.45%**	**65.99**	10925	4075
DEES_296	**Catalonia**	7598809	2155660	**34.50%**	**94.52**	17869	5381
DEES_297	**Catalonia**	4424996	969769	**24.16%**	**42.41**	15904	5507
DEES_299	**Catalonia**	5230017	1265067	**26.07%**	**58.51**	15671	5261
DEES_300	**Catalonia**	4262358	1357074	**27.54%**	**100.61**	11253	4340
DEES_301	**Catalonia**	3745587	1088394	**26.61%**	**69.74**	13078	4899
DEES_302	**Catalonia**	3489525	965862	**24.22%**	**69.57**	10235	4074
DEES_303	**Catalonia**	4531243	1120913	**23.94%**	**40.66**	20703	7593
DEES_304	**Catalonia**	4574323	1175078	**22.14%**	**35.71**	19085	6947
DEES_305	**Catalonia**	1867058	478825	**18.52%**	**40.52**	9110	3886
DEES_306	**Catalonia**	4434095	1105225	**24.02%**	**77.82**	11016	4210
DEES_307	**Catalonia**	4424587	1250214	**25.57%**	**40.37**	23520	8686
DEES_310	**Catalonia**	4722881	1196742	**38.28%**	**75.09**	13030	4624
DEES_311	**Catalonia**	3940111	869083	**32.80%**	**54.02**	12191	4363
DEES_312	**Catalonia**	5437071	1322049	**35.03%**	**42.9**	22259	7318
DEES_313	**Catalonia**	4594845	969020	**31.10%**	**37.44**	18775	6232
DEES_314	**Catalonia**	1211973	332293	**27.86%**	**26.46**	9392	3498
DEES_315	**Catalonia**	4763636	1037826	**32.73%**	**32.31**	23320	7795
DEES_316	**Catalonia**	5124568	1231098	**34.65%**	**56.34**	15281	5369
DEES_332	**Catalonia**	3687325	733031	**33.31%**	**47.51**	11290	4318
AGES_343	**Galicia**	3533972	1037302	**24.64%**	**82.63**	10350	4030
AGES_344	**Galicia**	3129636	908989	**23.91%**	**67.67**	11225	4424
AGES_345	**Galicia**	2660302	901133	**23.47%**	**80.43**	9472	3542
AGES_347	**Galicia**	2557168	819845	**25.73%**	**44.5**	15434	5401
AGES_348	**Galicia**	3362831	1096665	**24.94%**	**81.86**	11438	4207

AGES_349	**Galicia**	3405203	1191146	**25.95%**	**101.95**	9854	3524
AGES_350	**Galicia**	2673425	867875	**22.74%**	**68.25**	10829	4197
AGES_351	**Galicia**	2931543	873816	**27.81%**	**59.32**	12014	4535
AGES_352	**Galicia**	2883643	946070	**23.54%**	**88.58**	9272	3842
AGES_353	**Galicia**	2343779	838496	**21.54%**	**77.58**	9276	3989
AGES_354	**Galicia**	2296208	726043	**22.22%**	**60.94**	9668	3852
AGES_355	**Galicia**	2134775	703869	**20.75%**	**57.1**	10429	4163
AGES_356	**Galicia**	2448028	844600	**19.71%**	**36.68**	18757	6707
AGES_357	**Galicia**	2584089	843375	**22.67%**	**56.97**	12417	4620
AGES_359	**Galicia**	3270757	1139199	**22.32%**	**39.42**	23846	8276
AGES_362	**Galicia**	3667951	1066268	**26.25%**	**74.44**	11646	4688
AGES_364	**Galicia**	3671465	1245359	**26.12%**	**69.5**	14538	5165
AGES_366	**Galicia**	3325813	1022947	**25.26%**	**53.75**	14999	5367
SGES_367	**Galicia**	5473339	1237950	**43.57%**	**60.12**	16384	5328
SGES_368	**Galicia**	2766042	772347	**39.46%**	**58.84**	10493	3621
SGES_370	**Galicia**	1860846	443421	**42.79%**	**34.62**	8958	3554
SGES_371	**Galicia**	1116003	303185	**38.19%**	**28.24**	8175	2933
SGES_374	**Galicia**	662796	180373	**28.40%**	**14.22**	7285	2988
SGES_375	**Galicia**	1304206	362099	**33.15%**	**37.61**	7321	2613
SGES_377	**Galicia**	703437	196842	**27.79%**	**19.47**	6470	2256
SGES_379	**Galicia**	2179564	570282	**30.80%**	**46.8**	9397	3585
SGES_381	**Galicia**	990379	238150	**28.69%**	**23.11**	7152	2763
SGES_384	**Galicia**	2921009	832895	**33.38%**	**62.72**	10634	4197
SGES_386	**Galicia**	821346	170387	**40.84%**	**14.73**	6006	1993
SGES_387	**Galicia**	1274546	331230	**38.62%**	**28.07**	8601	3018
SGES_388	**Galicia**	1807352	435069	**40.03%**	**30.06**	9891	3288
SGES_390	**Galicia**	2340109	553186	**41.84%**	**42.7**	9362	3471
LBTN_418	**Bizerte**	5594613	1197688	**44.10%**	**65.89**	14571	4667
LBTN_419	**Bizerte**	3000926	457298	**38.90%**	**16.97**	13932	4240
LBTN_421	**Bizerte**	5721051	1380368	**22.94%**	**18.43**	20135	9792
LBTN_422	**Bizerte**	3048769	423415	**41.53%**	**15.33**	14477	4713
LBTN_425	**Bizerte**	4969442	752390	**18.59%**	**13.5**	14408	5229
LBTN_427	**Bizerte**	4006477	625744	**20.17%**	**25.5**	11575	4418
LBTN_429	**Bizerte**	3232220	575823	**21.59%**	**17.63**	10482	4048
LBTN_430	**Bizerte**	4391075	611260	**19.30%**	**18.42**	10695	3812
LBTN_431	**Bizerte**	5040792	946069	**22.54%**	**25.07**	14268	5707
LBTN_433	**Bizerte**	3817173	560157	**20.00%**	**14.41**	13604	5001
LBTN_434	**Bizerte**	3162190	510465	**19.53%**	**12.83**	6480	2478
LBTN_437	**Bizerte**	2794409	337863	**18.66%**	**10.13**	8329	2958
LBTN_439	**Bizerte**	4795385	707797	**19.96%**	**20.46**	10164	3712
LBTN_442	**Bizerte**	4445158	583263	**18.60%**	**16.04**	11028	3959
LBTN_446	**Bizerte**	3916268	659249	**18.31%**	**13.67**	11099	4011

Table S2.2. Datasets before and after SNP filtering including all individuals (dataset 1) or only the Mediterranean and Eastern Pacific individuals (dataset 2). ATL = Atlantic area (94 Samples); SEP = South-eastern Pacific area (41 Samples), MED = Mediterranean area (32 Samples).

	AFTER "POPULATIONS"	AFTER SNP FILTERING
DATASET 1: ATL + MED + SEP	23,584 SNPs 182 Samples 2,226 Tags	959 SNPs 167 Samples 464 Tags
DATASET 2: MED + SEP	15,558 SNPs 79 Samples 2,412 Tags	1,506 SNPs 73 Samples 698 Tags

Table S2.3. Pairwise F_{ST} values calculated using the 100 most informative SNPs for geographic origin assignment of Mediterranean (DEES|LBTN) *vs* South-eastern Pacific individuals (COCL|TBCL) using Dataset 2. 95% CIs are showed in parentheses.

	DEES	LBTN	COCL
LBTN	0.0105 (-0.0153, 0.0437)		
COCL	**0.0793 (0.0467, 0.1232)**	**0.097 (0.0491, 0.1511)**	
TBCL	**0.0786 (0.0494, 0.1113)**	**0.0976 (0.0563, 0.1442)**	0.0076 (-0.0141, 0.038)

Table S2.4. Ranking of the 100 most informative SNPs for geographic origin assignment based on F_{ST} values. Two different list were generated to match individuals to one of the three ocean areas of interest: (A) Atlantic *vs* Mediterranean/South-eastern Pacific individuals (using Dataset 1) and (B) Mediterranean *vs* South-eastern Pacific individuals (using Dataset 2). MAF=Minor (alternative) allele frequency. (*) SNPs under positive selection identified by pcadapt (p-value < 0.05).

(A) ATL *vs* MED/SEP				
SNP name	**F_{ST} value**	**Reference allele**	**Alternative allele**	**MAF**
185180_16(*)	0.6896	C	A	0.4536
107603_28	0.5826	C	T	0.4872
9174_7(*)	0.5512	G	A	0.3491
4084_54	0.5386	T	C	0.3245
6271_47	0.5355	G	A	0.1923
2764_76(*)	0.5266	C	T	0.4222
207941_36(*)	0.5184	G	A	0.4145
6534_16(*)	0.514	T	G	0.3841
4084_68	0.5129	C	T	0.3245
2764_45(*)	0.511	T	G	0.4162
1771_65	0.5042	G	T	0.4243
207941_22(*)	0.4827	T	C	0.4145
207941_23(*)	0.4827	T	C	0.4013
6414_39(*)	0.4715	T	C	0.4783

1771_74	0.4628	C	T	0.471
1249_69(*)	0.4589	A	G	0.4737
288018_69	0.4581	A	G	0.2715
107_27(*)	0.4449	T	A	0.3097
1771_37	0.4403	T	G	0.4487
7974_48	0.4322	A	T	0.3758
9012_11	0.4229	G	A	0.4308
264595_42	0.3841	T	A	0.3693
8283_77(*)	0.3719	C	T	0.3994
687_66	0.3075	C	A	0.3258
687_58	0.2943	T	C	0.3301
3804_43	0.2935	G	T	0.3067
230756_25	0.2926	G	T	0.1589
230756_58	0.2926	G	T	0.1589
8296_24(*)	0.2892	C	G	0.3662
82_62	0.2887	A	G	0.2893
82_50	0.2867	C	T	0.2862
229314_57	0.2726	C	T	0.2222
2374_39	0.2721	G	A	0.4844
8793_58	0.2713	C	T	0.2484
4563_6	0.2702	A	G	0.4224
2471_63(*)	0.2651	A	T	0.3302
2471_69(*)	0.2651	C	T	0.3313
2471_64	0.2649	T	A	0.3418
2471_10(*)	0.2597	T	C	0.3313
3214_24	0.2593	T	A	0.2342
42_77	0.2578	C	A	0.3519
8976_8	0.2562	T	C	0.4295
6762_12	0.2516	G	A	0.2941
8793_36	0.2493	C	T	0.2484
209411_74	0.2487	C	T	0.3464
128977_21(*)	0.2464	T	C	0.465
4654_71	0.2251	A	T	0.3084
9394_36	0.2244	T	G	0.4367
4835_31	0.2225	T	A	0.1786
283850_26	0.2199	C	A	0.3791
9394_27	0.2198	T	C	0.4487
4710_52	0.2197	G	T	0.3933
9205_30	0.2175	G	T	0.3113
9205_31	0.2171	C	T	0.3125
3268_6	0.2166	A	G	0.2006
3268_52	0.2166	C	T	0.2006
3268_64	0.2166	G	T	0.2006
559_7	0.2164	C	A	0.2467
2854_47	0.2142	G	A	0.1402
185180_77	0.2132	A	G	0.1871
377874_50	0.2118	G	A	0.4837
150815_52	0.2046	C	T	0.1487
150815_62	0.2046	T	C	0.1487
5242_31	0.2021	A	T	0.4735
232115_13	0.2004	C	G	0.3247
2854_24	0.1992	C	A	0.1463
2213_26	0.1971	A	C	0.2969
108726_23(*)	0.1962	T	C	0.325
288018_19	0.1953	T	A	0.4106
134251_49	0.1948	A	T	0.1536
479_17	0.1914	T	A	0.2256
1074_8	0.1874	G	A	0.1551

260762_22	0.1796	C	A	0.3914
3123_41	0.1784	C	T	0.1442
1801_55	0.1784	T	C	0.06442
5859_42	0.1778	C	A	0.1472
9394_18	0.1758	A	G	0.4615
244387_77	0.1756	T	G	0.08278
7773_47	0.1731	C	T	0.1516
6271_16	0.1729	G	T	0.4647
2793_48	0.1712	C	T	0.4412
4704_55	0.1709	A	G	0.118
2478_72	0.1699	T	C	0.3374
2273_33	0.1679	C	T	0.4129
4588_78	0.1677	A	T	0.1012
7773_26	0.1649	C	T	0.2484
3069_76	0.1641	C	A	0.06587
3069_77	0.1641	A	T	0.06587
199247_26	0.1609	T	A	0.3494
4704_22	0.1601	T	A	0.1149
4541_66	0.1591	G	A	0.2532
107_21	0.1588	C	A	0.4777
9394_75	0.1572	A	G	0.4513
7096_67	0.1567	A	C	0.1646
4024_69	0.1531	T	A	0.2
232115_22	0.1518	T	G	0.3149
232115_28	0.1518	T	C	0.3117
2076_50	0.1484	A	T	0.05556
129799_74	0.1477	C	T	0.09936
4068_53	0.1475	C	T	0.1603

(B) MED *vs* SEP

SNP name	F_{ST} value	Reference allele	Alternative allele	MAF
1874_19	0.3396	A	T	0.4338
8162_38	0.3232	A	C	0.125
9568_11	0.3136	G	T	0.2424
6850_27	0.3106	A	T	0.2121
1800_32	0.3058	A	T	0.2609
4654_74	0.2837	G	A	0.2603
3154_57	0.2673	A	T	0.4348
11280_75	0.2373	G	A	0.07463
6251_19	0.2297	G	A	0.1515
313595_49	0.2284	A	G	0.2164
16972_34	0.2068	T	A	0.2836
8940_27	0.2058	T	C	0.2123
14870_48	0.1987	C	T	0.1544
260762_22	0.1936	A	C	0.4485
5738_73	0.193	T	C	0.07534
3436_46	0.1895	A	T	0.07042
6189_63	0.189	A	T	0.3824
3486_46	0.1824	A	G	0.07246
3486_27	0.1824	T	A	0.07246
227193_61	0.1799	A	T	0.3409
227193_43	0.1799	T	A	0.3409
227193_40	0.1799	G	A	0.3258
227193_17	0.1799	T	C	0.3409
2129_77	0.1764	T	G	0.06061

2349_45	0.1744	A	G	0.09091
6251_58	0.1665	C	A	0.1591
5165_37	0.1652	A	T	0.08333
5165_30	0.1652	G	A	0.08333
9735_62	0.1642	A	T	0.06061
740_19	0.1624	A	C	0.3043
1240_65	0.1553	A	C	0.2121
2120_62	0.1524	C	A	0.05556
345768_30	0.1516	A	G	0.3456
266621_48	0.1508	A	G	0.375
266621_47	0.1508	G	T	0.375
6189_47	0.1499	T	A	0.1103
500_69	0.1491	G	C	0.1714
537_19	0.1491	T	C	0.08696
108726_23(*)	0.1487	T	C	0.125
5211_6	0.1464	G	T	0.1818
7837_8	0.1455	A	G	0.06944
232589_6	0.1436	G	T	0.1522
2508_36	0.1429	G	A	0.2687
3578_21	0.1406	G	A	0.05714
3578_15	0.1406	A	T	0.05714
232032_11	0.1398	A	T	0.1304
5065_51	0.1394	A	T	0.07857
547_55	0.138	T	C	0.1438
7974_20	0.136	C	A	0.07746
5073_26	0.1342	A	G	0.4929
241_68	0.1332	A	C	0.08571
2471_64	0.1328	T	A	0.4
4404_27	0.1323	C	T	0.125
4589_17	0.1323	C	T	0.1176
2649_48	0.1322	T	C	0.2887
2649_47	0.1322	T	A	0.2887
253050_25	0.1313	G	T	0.08088
317210_22	0.1307	T	A	0.1194
229617_25	0.1305	G	A	0.06716
4398_13	0.13	T	C	0.05479
261328_44	0.1293	T	C	0.1397
302471_17	0.127	A	T	0.07463
285853_19	0.1257	C	T	0.2388
2129_65	0.1246	G	A	0.07576
6114_72	0.1233	C	A	0.07746
2273_33	0.1218	T	C	0.3714
6261_22	0.1216	C	A	0.2803
2471_69	0.1215	T	C	0.4
2471_63	0.1215	A	T	0.3731
2471_10	0.1215	C	T	0.4
829_23	0.1213	C	T	0.194
883_78	0.1191	G	C	0.1268
883_76	0.1191	C	T	0.1458
883_66	0.1191	T	A	0.1268
9596_18	0.1189	T	A	0.1571
337_49	0.117	G	A	0.05797
1926_66	0.1159	A	G	0.1
1926_6	0.1159	T	A	0.1
1926_53	0.1159	C	T	0.1
1926_39	0.1159	G	T	0.1
1926_36	0.1159	A	T	0.1
313629_72	0.1158	A	G	0.08333

164720_36	0.1135	C	A	0.0597
164720_33	0.1135	G	T	0.0597
285853_9	0.1127	A	G	0.05224
396097_35	0.1118	A	T	0.1439
101612_72	0.111	T	C	0.2647
2280_32	0.1105	A	G	0.5
9596_24	0.1098	G	T	0.403
4875_72	0.1088	T	C	0.08333
4024_52	0.1079	T	C	0.1159
7157_64	0.1074	A	C	0.05797
399703_18	0.107	C	T	0.2206
6083_55	0.1069	A	C	0.06944
8620_68	0.1069	G	A	0.05479
65261_78	0.1069	G	A	0.0625
965_10	0.1034	T	G	0.4783
6538_27	0.1028	G	A	0.4779
930_20	0.1015	C	T	0.1164
8793_36	0.1012	C	T	0.09286

APPENDIX B

Table S3.1. Random forest classification success rates (%) in assigning farmed *M. galloprovincialis* mussels to their harvesting region based on (A) TEF and (B) SIRA. ATL = Atlantic Ocean area; MED = Mediterranean Sea area; SEP = South Eastern Pacific Ocean area. κ= Cohen´s Kappa statistic.

(A) TEF

Overall Accuracy: 87.76 (κ=0.93)

		ATL				MED			SEP
		LBFR	**PBPT**	**GAES**	**BQES**	**DEES**	**GEIT**	**LBTN**	**COCL**
ATL	**LBFR**	**89**	4	0	0	0	0	0	1
	PBPT	9	**94**	2	0	0	0	0	0
	GAES	0	2	**85**	6	12	10	5	1
	BQES	0	0	5	**94**	0	1	0	0
MED	**DEES**	2	0	4	0	**80**	9	6	0
	GEIT	0	0	2	0	8	**72**	1	0
	LBTN	0	0	1	0	0	8	**89**	0
SEP	**COCL**	0	0	0	0	0	0	0	**98**

(B) SIRA

Overall Accuracy: 80.45 (κ=0.81)

		ATL				MED			SEP
		LBFR	**PBPT**	**GAES**	**BQES**	**DEES**	**GEIT**	**LBTN**	**COCL**
ATL	**LBFR**	**75**	0	5	2	0	10	0	0
	PBPT	0	**89**	2	3	0	1	0	0
	GAES	13	7	**92**	6	0	0	0	0
	BQES	7	4	1	**65**	34	35	0	0
MED	**DEES**	0	0	0	17	**66**	0	0	0
	GEIT	6	0	0	7	0	**54**	0	0
	LBTN	0	0	0	0	0	0	**100**	0
SEP	**COCL**	0	0	0	0	0	0	0	**100**

Table S3.2. Trace element concentrations (µg g-1; mean ± SD) by location of *M. galloprovincialis* shells. P-values for Kruskal-Wallis test determining whether the individual elemental concentrations varied among harvesting sites are in blue and statistically significant values are marked with asterisks (*). Within each column, red letters identify significant differences (p<0.05) for BH adjusted Pairwise Wilcoxon Rank Sum Test comparisons, sharing the letter when no significant differences were identified. MED = Mediterranean Sea area; ATL = Atlantic Ocean area; SEP = South Eastern Pacific Ocean area.

Sample ID	Country	^{60}Ni p-value= 0.0005736*		^{63}Cu p-value= 0.0007473*		^{66}Zn p-value=2.891e-05*		^{75}As p-value= 0.07435		^{111}Cd * p-value=6.009e-05*		^{137}Ba p-value= 1.471e-12*		^{208}Pb p-value=7.373e-13*	
LBFR	France	0.22±0.02	bcd	0.03±0.07	a	2.47±0.79	ab	0.02±0.01	-	0.00±0.00	bc	0.96±0.31	f	0.10±0.04	e
PBPT	Portugal	0.17±0.08	acd	0.76±2.06	a	2.78±1.92	abc	0.02±0.01	-	0.00±0.00	bc	0.83±0.15	f	0.22±0.04	b
AGES	Spain	0.15±0.04	a	0.00±0.00	a	1.69±0.73	ab	0.02±0.01	-	0.33±0.38	a	2.79±0.76	a	0.25±0.08	abc
SGES	Spain	0.48±0.38	abcd	0.00±0.00	a	4.20±3.17	bc	0.04±0.02	-	0.01±0.02	abc	1.09±0.22	bc	0.30±0.05	ac
MEES	Spain	0.15±0.05	ad	0.07±0.19	a	3.88±1.71	c	0.02±0.01	-	0.01±0.01	ac	1.34±0.29	b	0.85±0.21	f
MUES	Spain	0.24±0.15	abcd	1.03±2.88	a	4.54±3.18	bc	0.03±0.03	-	0.29±0.32	abc	1.20±0.21	c	0.49±0.07	g
DEES	Spain	0.24±0.05	bc	0.00±0.00	a	1.35±0.82	a	0.02±0.01	-	0.00±0.00	b	1.98±0.71	d	0.29±0.03	a
GEIT	Italy	0.27±0.08	b	0.03±0.09	a	3.10±1.79	bc	0.04±0.02	-	0.00±0.00	bc	4.68±1.05	e	0.24±0.07	bc
LBTN	Tunisia	0.24±0.05	bc	0.59±1.38	a	5.12±1.22	c	0.02±0.01	-	0.00±0.00	bc	2.21±0.51	ad	0.19±0.02	b
COCL	Chile	0.24±0.08	abcd	0.06±0.16	a	2.22±1.17	ab	0.02±0.01	-	0.00±0.00	bc	1.41±0.41	bc	0.01±0.02	d

Sample ID	Country	^{11}B p-value= 5.96e-06*		^{27}Al p-value= 2.862e-07*		^{47}Ti p-value= 0.3056		^{51}V p-value= 0.000129*		^{52}Cr p-value= 0.02558*		^{55}Mn p-value= 1.544e-10*		^{59}Co p-value= 0.05545	
LBFR	France	8.68±2.02	a	7.35±1.66	ab	0.01±0.03	-	0.02±0.01	a	0.14±0.04	a	1.82±0.44	d	0.04±0.01	-
PBPT	Portugal	8.33±1.74	a	12.08±6.94	bc	0.02±0.05	-	0.03±0.01	a	0.15±0.11	a	1.24±0.26	e	0.04±0.01	-
AGES	Spain	9.02±2.52	a	10.00±3.83	abc	0.08±0.12	-	0.03±0.01	a	0.23±0.11	a	3.56±1.10	a	0.03±0.01	-
SGES	Spain	12.1±2.78	ab	21.74±5.00	d	0.16±0.21	-	0.03±0.01	a	1.02±0.92	a	3.46±0.76	ac	0.04±0.01	-
MEES	Spain	10.53±2.55	ab	39.64±21.66	e	0.05±0.10	-	0.03±0.01	a	0.13±0.08	a	2.18±0.42	d	0.03±0.01	-
MUES	Spain	10.15±2.52	ab	14.57±10.16	abc	0.01±0.03	-	0.04±0.06	a	0.41±0.33	a	5.11±1.03	b	0.04±0.01	-
DEES	Spain	8.49±2.65	a	6.33±2.48	a	0.15±0.39	-	0.01±0.01	b	0.16±0.06	a	3.46±0.99	a	0.03±0.01	-
GEIT	Italy	9.93±2.71	a	11.39±6.37	abc	0.09±0.15	-	0.03±0.01	a	0.2±0.14	a	4.84±1.34	bc	0.04±0.01	-
LBTN	Tunisia	13.6±1.65	b	16.25±9.41	cd	0.01±0.02	-	0.03±0.01	a	0.23±0.1	a	5.28±1.79	b	0.04±0.01	-
COCL	Chile	7.89±2.46	a	9.18±3.88	abc	0.04±0.04	-	0.03±0.01	a	0.26±0.15	a	3.37±0.78	a	0.04±0.01	-

APPENDIX C

Supplementary material 1. Alpha diversity estimators' significance. Differences in alpha diversity patterns between sampling points and seasons were analysed using ANOVA and Tukey's honest significance test of ANOVA (pairwise comparisons) for normally distributed metrics and Kruskal-Wallis and Wilcoxon rank sum test (pairwise comparisons) for non-parametric metrics ("stats" package). Normality was tested using Shapiro-Wilk test.

(A) RICHNESS

 # OVERALL

p-value (Kruskal-Wallis) = 4.182e-15

RICHNESS	AGES	DEES	MEES	MUES
DEES	1.7E-08	-	-	-
MEES	0.00117	0.00117	-	-
MUES	0.50197	5.2E-09	0.00012	-
SGES	0.02264	4.30E-09	1.30E-06	0.043

#BY SEASON

p-value (Kruskal-Wallis) = 6.084e-05

WINTER_RI	AGES	DEES	MEES	MUES
DEES	0.00548	-	-	-
MEES	0.01955	0.32634	-	-
MUES	0.56324	0.00078	0.00167	-
SGES	0.56324	7.80E-04	1.67E-03	0.41026

p-value (Kruskal-Wallis) = 0.004789

SPRING_RI	AGES	DEES	MEES	MUES
DEES	0.0031	-	-	-
MEES	0.8785	0.0175	-	-
MUES	0.8785	0.0031	0.8785	-
SGES	0.8785	9.80E-03	7.65E-01	0.8423

p-value (Kruskal-Wallis) = 0.000927

SUMMER_RI	AGES	DEES	MEES	MUES
DEES	0.0295	-	-	-
MEES	0.7209	0.1083	-	-
MUES	0.0295	0.0098	0.0098	-
SGES	0.2007	9.80E-03	1.19E-01	0.3647

p-value (Kruskal-Wallis) = 2.965e-05

AUTUMN_RI	AGES	DEES	MEES	MUES
DEES	0.00272	-	-	-
MEES	0.02109	0.95913	-	-
MUES	0.26055	0.01166	0.0812	-
SGES	0.00373	7.80E-04	7.80E-04	0.00104

(B) SIMPSON

 # OVERALL

p-value (Kruskal-Wallis) = 8.312e-11

SIMPSON	AGES	DEES	MEES	MUES
DEES	0.0000005	-	-	-
MEES	0.00632	0.02712	-	-
MUES	0.12565	0.00067	0.16742	-
SGES	0.03009	2.50E-08	9.40E-06	0.00067

#BY SEASON

p-value (Kruskal-Wallis) = 0.0005552

WINTER_SI	AGES	DEES	MEES	MUES
DEES	0.0031	-	-	-
MEES	0.0033	0.8518	-	-
MUES	0.4909	0.026	0.0488	-
SGES	0.2148	2.33E-02	4.00E-02	0.0542

p-value (Kruskal-Wallis) = 0.008224

SPRING_SI	AGES	DEES	MEES	MUES
DEES	0.0148	-	-	-
MEES	0.6317	0.0703	-	-
MUES	0.3897	0.0689	0.7171	-
SGES	0.8785	6.20E-03	4.92E-01	0.3908

p-value (Kruskal-Wallis) = 0.0006533

SUMMER_SI	AGES	DEES	MEES	MUES
DEES	0.0054	-	-	-
MEES	0.0764	0.4103	-	-
MUES	0.0758	0.335	0.7209	-
SGES	0.7209	3.10E-03	6.20E-03	0.0175

p-value (Kruskal-Wallis) = 0.009546

AUTUMN_SI	AGES	DEES	MEES	MUES
DEES	0.645	-	-	-
MEES	0.229	0.217	-	-
MUES	0.244	0.217	0.645	-
SGES	0.011	5.20E-02	5.20E-02	0.217

(C) SHANON

 # OVERALL

p-value (ANOVA) =2.75e-15

SHANNON	AGES	DEES	MEES	MUES
DEES	0	-	-	-
MEES	0.001687	0.0065621	-	-
MUES	0.5391619	0.0000003	0.1433091	-
SGES	0.6704277	0.00E+00	7.80E-06	0.0369624

#BY SEASON

p-value (ANOVA)= 1.54e-06

WINTER_SH	AGES	DEES	MEES	MUES
DEES	0.0002332			
MEES	0.0014856	0.9984562		
MUES	0.5166713	0.0181682	0.0641591	
SGES	0.852068	1.25E-05	1.04E-04	0.0957991

p-value (ANOVA) = 0.00229

SPRING_SH	AGES	DEES	MEES	MUES
DEES	0.0026824			
MEES	0.9831159	0.0114574		
MUES	0.8984416	0.0282978	0.9963894	
SGES	0.9833616	1.14E-02	1.00E+00	0.9963061

p-value (ANOVA) = 7.63e-06

SUMMER_SH	AGES	DEES	MEES	MUES
DEES	0.0000299			
MEES	0.0100658	0.2795541		
MUES	0.4319339	0.004603	0.393416	
SGES	0.9992872	6.06E-05	1.85E-02	0.5752336

p-value (ANOVA) = 0.000991

AUTUMN_SH	AGES	DEES	MEES	MUES
DEES	0.760739			
MEES	0.9944254	0.5184627		
MUES	0.5960827	0.0845782	0.8261007	
SGES	0.0148622	5.57E-04	3.99E-02	0.3297857

(D) BERGER PARKER

 # OVERALL

p-value (Kruskal-Wallis) = 3.457e-09

BERGER-PARKER	AGES	DEES	MEES	MUES
DEES	0.000085	-	-	-
MEES	0.0211	0.0497	-	-
MUES	0.1654	0.0064	0.3902	-
SGES	0.0288	5.00E-08	3.50E-05	0.0011

#BY SEASON

p-value (Kruskal-Wallis) = 0.001341

WINTER_BP	AGES	DEES	MEES	MUES
DEES	0.0093	-	-	-
MEES	0.0067	0.662	-	-
MUES	0.5615	0.0369	0.071	-
SGES	0.4103	2.33E-02	4.00E-02	0.1863

p-value (Kruskal-Wallis) = 0.008423

SPRING_BP	AGES	DEES	MEES	MUES
DEES	0.0093	-	-	-
MEES	0.2436	0.2436	-	-
MUES	0.2436	0.2436	0.7209	-
SGES	0.6375	6.20E-03	2.44E-01	0.2436

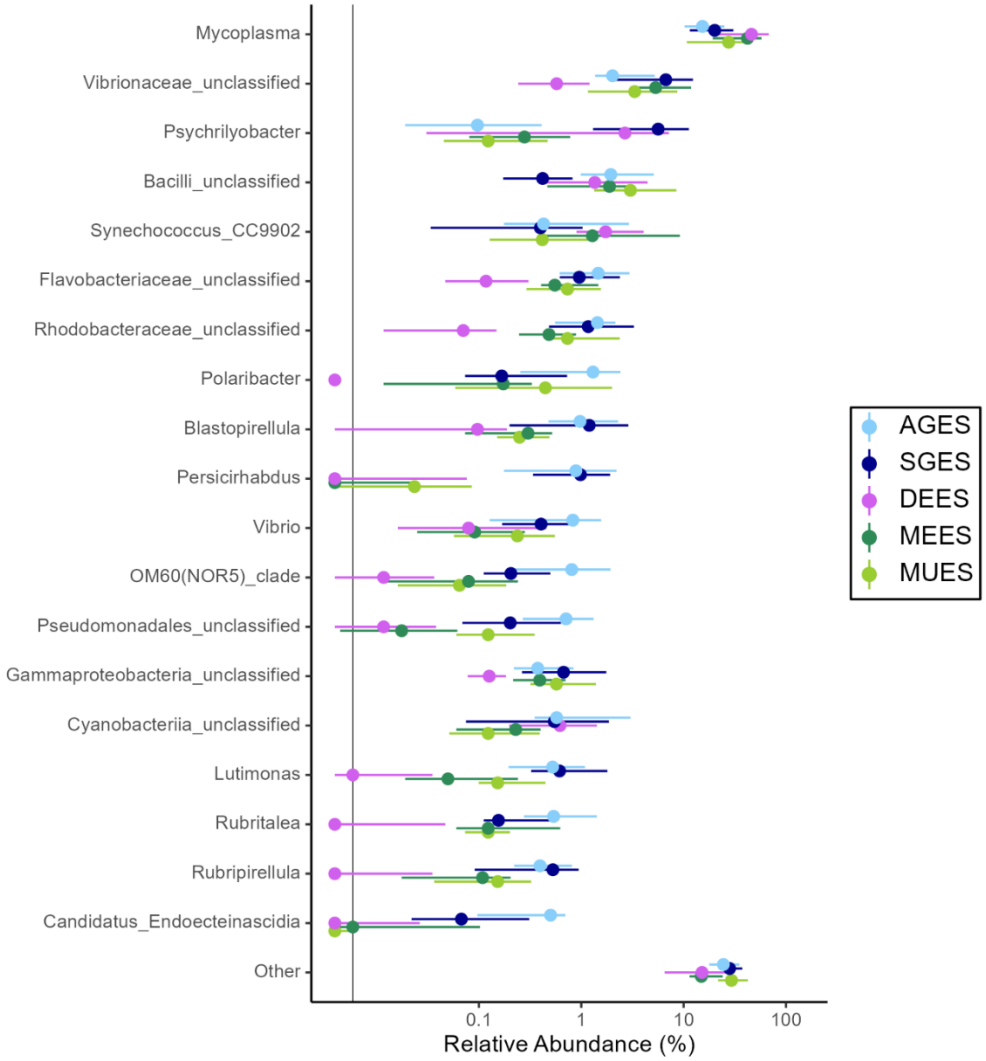
p-value (Kruskal-Wallis) = 0.002974

SUMMER_BP	AGES	DEES	MEES	MUES
DEES	0.052	-	-	-
MEES	0.175	0.552	-	-
MUES	0.076	0.562	0.878	-
SGES	0.335	1.10E-02	1.50E-02	0.035

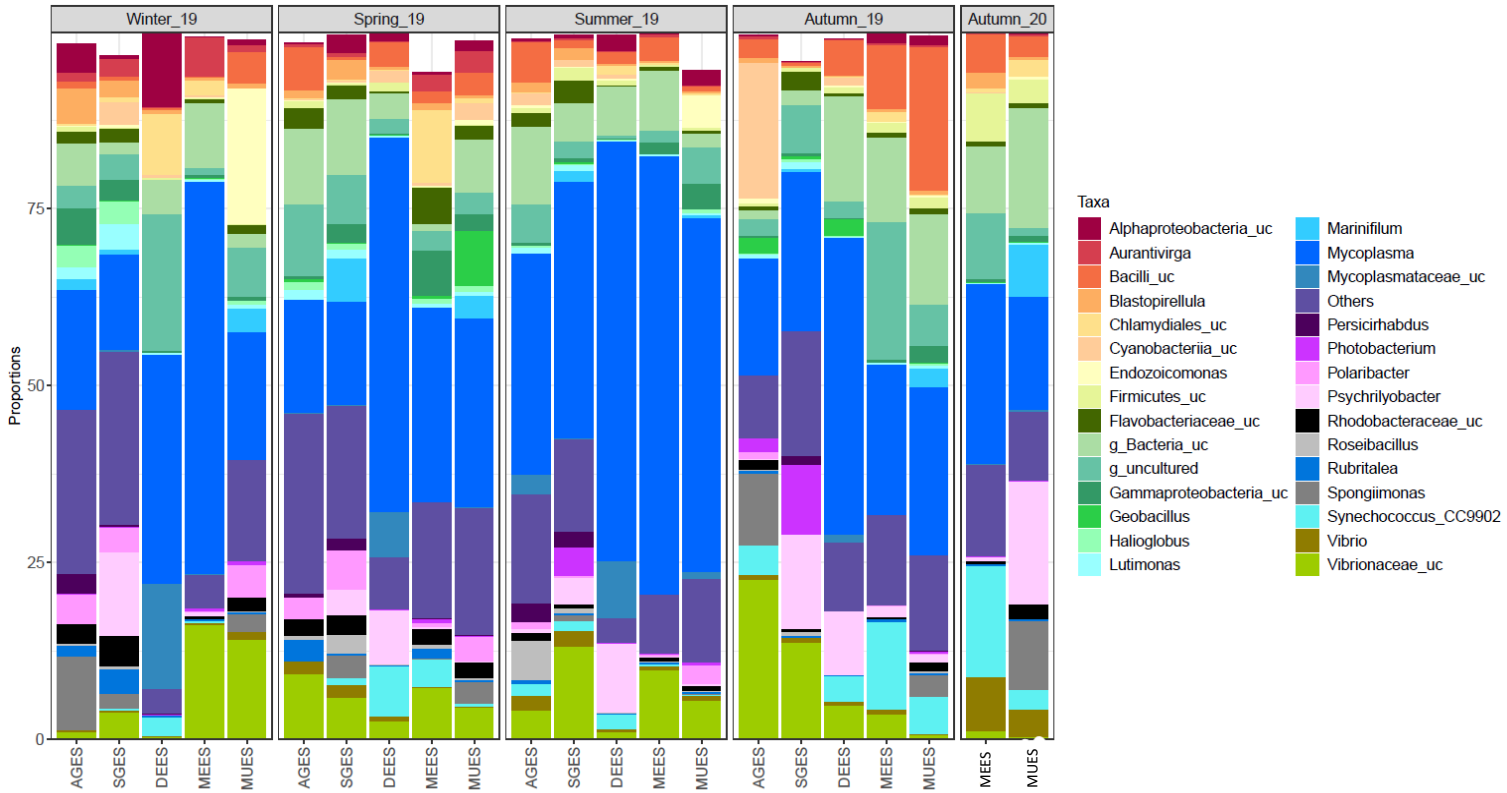
p-value (Kruskal-Wallis) = 0.01375

AUTUMN_BP	AGES	DEES	MEES	MUES
DEES	1	-	-	-
MEES	0.21	0.296	-	-
MUES	0.207	0.348	1	-
SGES	0.019	9.00E-02	2.07E-01	0.26

Supplementary figure 4.2. Relative abundance of genus that differed significantly (Kruskall Wallis, p<0.05) between the five mussel farms. Points are median values; lines represent the interquartile range and the black vertical line is the limit of detection. Taxa with a median relative abundance < 0.5% for all the locations were grouped in "Others". Colours correspond to different harvest locations in Galician region (● AGES, ● SGES), Catalonia region (● DEES) and Basque Country region (● MEES, ● MUES).



Supplementary material 4.3. Relative abundances at genus level; 2019 *vs* 2020. **(A).** Relative abundance of bacterial communities, at genus level, of mussel DG harvested from five different farms during 2019 and 2020: in Galician region (AGES, SGES), Catalonia region (DEES) and Basque Country region (MEES, MUES). Taxa not within the 20 most abundant families were pooled together as "Other"



Supplementary material 4.3. Relative abundances at genus level; 2019 *vs* 2020. **(b). Left.** Relative abundance of genus that differed significantly (Kruskall Wallis, p<0.05) between the mussel farms in 2019 and 2020. Points are median values; lines represent the interquartile range and the black vertical line is the limit of detection. Taxa with a median relative abundance < 0.5% for all the locations were grouped in "Others".