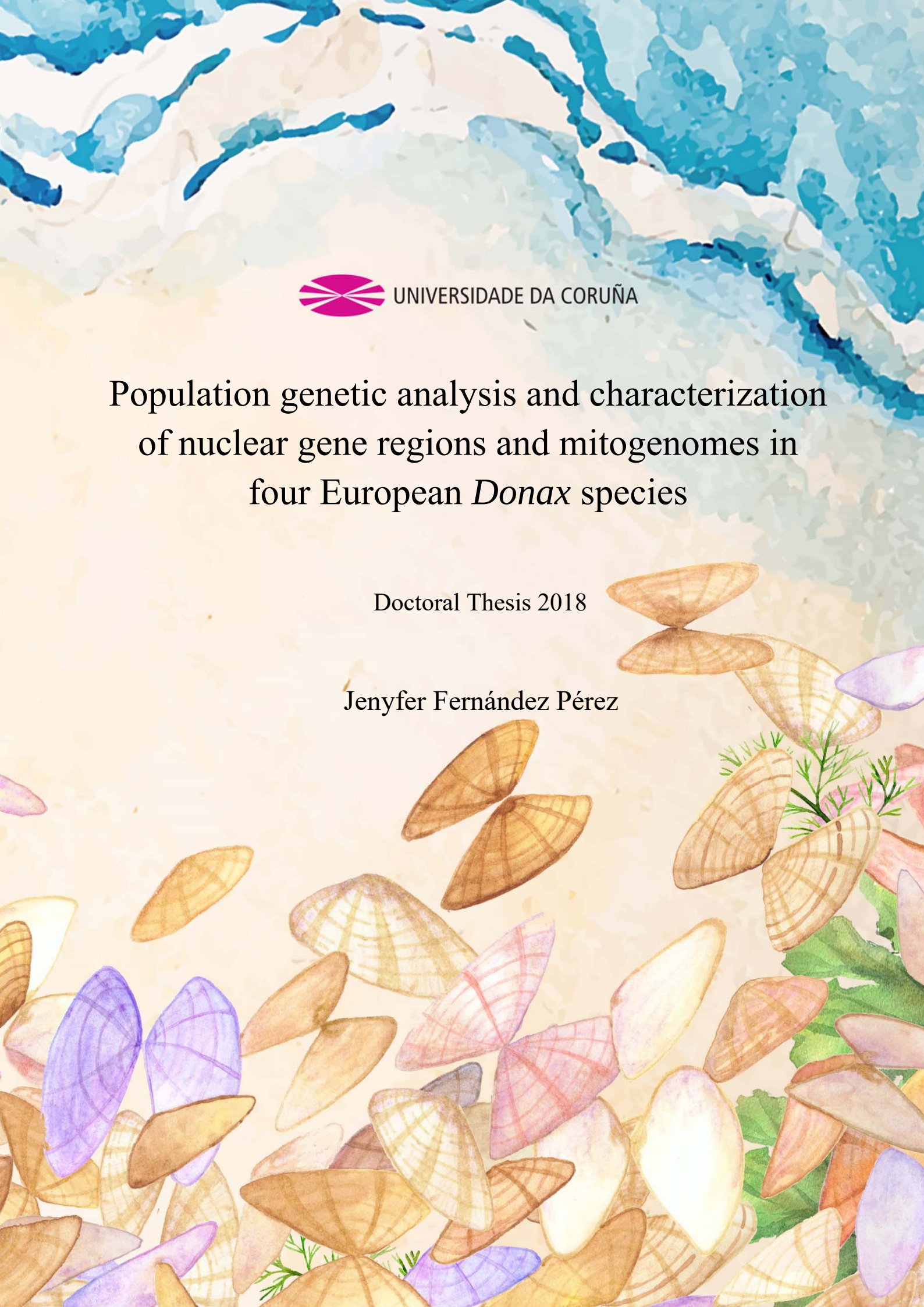


UNIVERSIDADE DA CORUÑA

Population genetic analysis and characterization of nuclear gene regions and mitogenomes in four European *Donax* species

Doctoral Thesis 2018

Jenyfer Fernández Pérez

Population genetic analysis and characterization of nuclear gene regions and mitogenomes in four European *Donax* species

Jenyfer Fernández Pérez

Doctoral Thesis

2018

Directora/Tutora: Josefina Méndez Felpeto

Programa Oficial de Doctorado en Biología Celular y Molecular

UNIVERSIDADE DA CORUÑA

Population genetic analysis and characterization of nuclear gene regions and mitogenomes in four European *Donax* species

DEPARTAMENTO DE BIOLOGÍA

Memoria que para optar al Título de Doctora con Mención Internacional presenta

Jenyfer Fernández Pérez

2018

Directora: Dra. Josefina Méndez Felpeto

UNIVERSIDADE DA CORUÑA

UNIVERSIDADE DA CORUÑA

Dña. Josefina Méndez Felpeto, Doctora en Biología y Catedrática de Universidad del Área de Genética del Departamento de Biología de la Universidade da Coruña,

INFORMA:

Que el trabajo titulado "Population genetic analysis and characterization of nuclear gene regions and mitogenomes in four European *Donax* species" presentado por Dña. Jenyfer Fernández Pérez para optar al Título de Doctora en Biología con Mención Internacional, ha sido realizado bajo mi dirección y, considerándolo finalizado, autorizo su presentación y defensa.

A Coruña, 29 de Mayo de 2018

Fdo. Josefina Méndez Felpeto

Financiación de la investigación

La realización de esta tesis ha sido posible gracias a la financiación obtenida a través del proyecto "Contribución genética para la recuperación de los bancos naturales de coquina (*Donax* spp.) en Galicia", dirigido por la Dra. Josefina Méndez Felpeto y concedido por el Ministerio de Economía, Industria y Competitividad del Gobierno de España (AGL2016-75288-R). La autora de este trabajo ha disfrutado de un contrato temporal con cargo a dicho proyecto (marzo 2017 – julio 2018).

Estancias

Parte del trabajo ha sido realizado en:

– *Interdisciplinary Centre of Marine and Environmental Research* (CIIMAR) (Universidade do Porto, Portugal) durante una estancia de tres meses en el año 2016, dentro del programa de "*Axudas para estadías predoutorais INDITEX-UDC 2016*", y bajo la supervisión de la Dra. Elsa Froufe.

– Departamento de Genética de la Universidad de Granada durante una breve estancia en el año 2016, bajo la supervisión del Dr. Juan Pedro Martínez Camacho.

Publicaciones

Por el momento, parte de este trabajo ha sido publicado en forma de cuatro artículos científicos y dos capítulos de libro:

– Fernández-Pérez, J., Nantón, A., Méndez, J., 2018. An alternative method for rapid and specific authentication of four European *Donax* species, including *D. trunculus*, a commercially-important bivalve. Eur Food Res. Technol. DOI: 10.1007/s00217-018-3093-5

– Fernández-Pérez, J., Nantón, A., Arias-Pérez, A., Freire, R., Martínez-Patiño, D., Méndez, J., 2018. Mitochondrial DNA analyses of *Donax trunculus* (Mollusca: Bivalvia) population

structure in the Iberian Peninsula, a bivalve with high commercial importance. Aquatic Conserv.: Mar. Freshw. Ecosyst. DOI: 10.1002/aqc.2929

– Fernández-Pérez, J., Nantón, A., Ruiz-Ruano, F.J., Camacho, J.P.M., Méndez, J., 2017. First complete female mitochondrial genome in four bivalve species genus *Donax* and their relationships within the Veneroida order. PLoS ONE. 19(9),e0184464.

– Fernández-Pérez, J., Froufe, E., Nantón, A., Gaspar, M.B., Méndez, J., 2017. Genetic diversity and population genetic analysis of *Donax vittatus* (Mollusca: Bivalvia) and phylogeny of the genus with mitochondrial and nuclear markers. Estuar. Coast. Shelf Sci. 197,126- 135.

– Fernández-Pérez, J., Nantón, A., Rodríguez, L., Febrero, F., Nóvoa, S., Ojea, J., Martínez-Patiño, D., Méndez, J., 2016. Evaluación genética de la coquina *Donax trunculus* en tres localidades gallegas mediante el marcador mitocondrial 16S. En: Foro Rec. Mar. Ac. Rías Gal. Rey-Méndez, M., Fernández Casal, J., Lodeiros, C., Guerra, A. (eds) 18,233-241. ISBN: 978-84-608-7689-2.

– Nantón, A., Fernández-Pérez, J., Arias-Pérez, A., Freire, R., Méndez, J., 2015. Estudio de la diversidad genética y estructura poblacional de la coquina *Donax trunculus* de la Península Ibérica mediante marcadores de ADN mitocondrial. En: Foro Rec. Mar. Ac. Rías Gal. Rey-Méndez, M., Fernández Casal, J., Lodeiros, C., Guerra, A. (eds) 17,163-172. ISBN: 978-84-606-7810-6.

A mis padres y a mi hermano

A Adrián

Agradecimientos

En primer lugar, me gustaría mostrar mi agradecimiento más sincero a la Dra. Josefina Méndez Felpeto, por darme la gran oportunidad de formar parte de su equipo y haberme permitido realizar este trabajo. Gracias también por permitirme trabajar en un marco de confianza y libertad fundamentales para la realización de este trabajo y para mi formación profesional, por inculcarme siempre un espíritu ganador, por su tiempo y por compartir conmigo su conocimiento y experiencia.

A la Dra. Elsa Froufe por acogerme durante mi estancia en el *Interdisciplinary Centre of Marine and Environmental Research* (CIIMAR) de la Universidade do Porto (Portugal). Gran parte de este trabajo se debe a ella. Gracias también a todos mis compañeros por el cariño y el buen trato recibido, en especial a André, Cláudia, David y Vanessa. De todos vosotros no sólo he ganado conocimiento, sino valores y vivencias que siempre me acompañarán y me harán recordaros con gran cariño.

Al Dr. Juan Pedro Martínez Camacho por abrirme las puertas de su laboratorio, así como por su disposición y amabilidad en todo momento. Gracias también a Paquillo por ser mi consejero y maestro del apasionante mundo de la *NGS* y, sobre todo, por su disposición y paciencia desde el primer momento. Vuestra colaboración ha sido esencial para el desarrollo de parte de esta Tesis.

También quiero agradecer a todos los participantes de los muestreos, tanto de la Xunta de Galicia como del Instituto Português do Mar e da Atmosfera de Portugal, que hicieron posible la obtención de una parte importante de los ejemplares empleados en este trabajo. La mitad de esta tesis no hubiera sido posible sin la colaboración de Miguel Gaspar, Susana y Tea, gracias por el valiosísimo material.

Una especial mención a Alberto y Ruth, por su colaboración en parte de este trabajo y por estar siempre dispuestos a ayudarme.

Gracias a todos mis compañeros de Genética, tanto a los que se mudaron al CICA, como a los que continúan. En especial quisiera agradecer a Ana, por ser mi madrina de poyata y la guía de mis primeros pasos en el laboratorio, por enseñarme todo sobre la coquina, por su amistad y por su

infinita paciencia y perseverancia. A Anita, por su apoyo y por sus valiosas sugerencias. A Luisa, por su ayuda desinteresada, por todas las correcciones de esta tesis, por sus acertados consejos, su constante ánimo y, sobre todo, por contagiarnos a todos con su buen humor. A Rosa, por su inestimable ayuda. A Vero, por preguntarme siempre por cómo me va y por los ánimos brindados en esta última etapa. A Zeltia, por su ánimo y simpatía. Muchas gracias por los buenos momentos compartidos, especialmente por vuestra motivación, apoyo y ayuda prestada en esta etapa final. Resulta fácil sacar un proyecto adelante cuando tu entorno se compone de personas como vosotras.

Gracias también a Noe, a la que no sólo se debe el diseño de la portada de este trabajo, si no también muchos buenos momentos vividos en A Coruña.

A mis amigos y familia por interesarse siempre por mi trabajo y por su constante ánimo.

A Adrián por recordarme que las metas están para ser alcanzadas, por mantenerme con ánimos, por estar siempre pendiente de mí y por hacer que esta etapa fuese más fácil a su lado.

Gracias a mi hermano por su apoyo incondicional durante el transcurso de todos estos años y por estar siempre ahí cuando necesitaba un consejo o unas palabras de ánimo. Él es la persona que me hace mirar hacia adelante sin miedo a nada.

Finalmente, quisiera agradecer de una manera muy especial a mis padres, a quienes les debo todo. Ellos han sido los promotores de mis sueños y los que han permitido que se hagan realidad. A ellos les dedico esta tesis en reconocimiento a todos los esfuerzos y sacrificios que han hecho por mí. Gracias por todo lo que me habéis enseñado a lo largo de la vida.

Quisiera también dejar escrita la dedicatoria de este trabajo a tres personas muy especiales que quise y quiero con toda mi alma, y que me mostraron en su propio ejemplo lo que significa ser una gran persona, Abuelita, Josefa y Miguel.

Muchas gracias.

Table of Contents

Abbreviations and Symbols

AFLPs	Amplified Fragment Length Polymorphisms
Ala	Alanine
AMOVA	Analysis of molecular variance
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
atp6	Atpase subunit 6
atp8	Atpase subunit 8
BI	Bayesian Inference
bp	Base pair
Cl	Chlorine
cob	Cytochrome b
COI	Cytochrome c oxidase subunit I
cox1	Cytochrome c oxidase subunit I
cox2	Cytochrome c oxidase subunit II
cox3	Cytochrome c oxidase subunit III
Cys	Cysteine
Cytb	Cytochrome b
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphates
DUI	Double Uniparental Inheritance
E	East
EDTA	Ethylenediaminetetraacetic acid
ETS	External transcribed spacer
F	Female
FAO	Food and Agriculture Organization of the United Nations
F_{IS}	Inbreeding coefficient

Gb	Giga base pairs
Gln	Glutamine
Glu	Glutamic acid
Gly	Glycine
h	Number of haplotypes
Hd	Haplotype diversity
H_e	Expected heterozygosity
H_o	Observed heterozygosity
His	Histidine
HWE	Hardy-Weinberg equilibrium
Ile	Isoleucine
ITS	Internal transcribed spacer
kb	Kilobases
kg	Kilogram
Leu	Leucine
Lys	Lysine
M	Male
MCMC	Markov chain Monte Carlo
MES	Minimum effective size
Met	Methionine
min	Minute
mg	Milligram
ML	Maximun Likelihood
ml	Millilitre
mM	Millimolar
MNCN	Museo Nacional de Ciencias Naturales
mt	mitochondrial

mtDNA	Mitochondrial DNA
mtgenome	Mitochondrial genome
N	Number of individuals
N	North
N_A	Number of alleles per locus
NAF	Null-allele frequency
nad1	Subunit 1 of the NADH dehydrogenase
nad2	Subunit 2 of the NADH dehydrogenase
nad3	Subunit 3 of the NADH dehydrogenase
nad4	Subunit 4 of the NADH dehydrogenase
nad4l	Subunit 4l of the NADH dehydrogenase
nad5	Subunit 5 of the NADH dehydrogenase
nad6	Subunit 6 of the NADH dehydrogenase
NCR	Non-coding region
nDNA	Nuclear DNA
ng	Nanogram
NGS	Next-generation Sequencing
Nm	Number of migrants
nt	Nucleotide
NTS	Non-transcribed spacer
P	*P*-value
PCG	protein-coding gene
PCR	Polymerase chain reaction
ph	Private haplotypes
Phe	Phenylalanine
Pro	Proline
RAPDs	Random Amplified Polymorphic DNA

rDNA	Ribosomal DNA
RFLPs	Restriction Fragment Length Polymorphisms
rpm	Revolutions per minute
rRNA	Ribosomal RNA
rrnL	Large-subunit ribosomal RNA
rrnS	Small-subunit ribosomal RNA
***R*s**	Allelic richness
s	Second
Ser	Serine
SDS	Standard Deviation Score
SMI	Strict Maternal Inheritance
SNPs	Single Nucleotide Polymorphisms
SSRs	Simple Sequence Repeats
STRs	Short Tandem Repeats
Thr	Threonine
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer RNA
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine
W	West
~	Above/Approximately
ºC	Degree Celsius
€	Euro
µg	Microgram
µl	Microliter
µM	Micromolar

π	Nucleotide diversity
%	Percentage
®	Registered
™	Trade Mark

List of tables and figures

Introduction

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Short abstracts

Abstract

The wedge clam *Donax trunculus* is an important shellfish resource in Europe that is in regression in Galician natural beds (Northwest of Spain). Nowadays in Galicia, only two markets located in Arousa and Cedeira sell this species, demonstrating the decline of this resource. In order to conserve, manage and recover this resource effectively, the following genetic analyses have been carried out: First, mitochondrial DNA markers (16S and Cytb) were used to analyse the population diversity and population structure of natural beds from Iberian Peninsula. Furthermore, the existence of *Donax vittatus* in the Atlantic coast of Iberian Peninsula, as another species to be conserved and/or managed, the development of molecular markers was needed to characterize its natural beds and to carry out studies similar to those conducted for *D. trunculus*. Likewise, given the existence of other *Donax* species (*Donax semistriatus* and *Donax variegatus*) in this region, the 5S rDNA, the internal transcribed spacers (ITS1 and ITS2) and the mitogenomes were characterized in these four species, in order to describe species-specific molecular markers for their identification and to construct the phylogenetic relationships within the order Veneroida.

Resumen

La coquina *Donax trunculus* es un importante recurso marisquero en Europa que se encuentra en regresión en los bancos naturales gallegos (noroeste de España). Actualmente, en Galicia, sólo las cofradías de Arousa y Cedeira comercializan esta especie, poniendo de manifiesto su declive. Con el propósito de conservar, gestionar y recuperar este recurso de manera eficaz, se han realizado los siguientes análisis genéticos: En primer lugar, la evaluación genético-poblacional de bancos naturales de la Península Ibérica mediante marcadores de ADN mitocondrial (16S y Citb) para conocer la diversidad genética y la estructura poblacional. Además, ante la presencia de *Donax vittatus* en la costa atlántica de la Península Ibérica, como otra especie a conservar y/o gestionar, fue necesario el desarrollo de marcadores moleculares para caracterizar sus bancos naturales y llevar a cabo estudios similares a los realizados para *D. trunculus*. Así mismo, dada la existencia de

otras especies del género *Donax* (*Donax semistriatus* y *Donax variegatus*) en el área de distribución a analizar, se caracterizaron el ADNr 5S, los espaciadores transcritos internos (ITS1 e ITS2), y los genomas mitocondriales en las cuatro especies objeto de estudio con el fin de obtener marcadores especie-específicos para su identificación y construir las relaciones filogenéticas dentro del orden Veneroida.

Resumo

A cadelucha *Donax trunculus* é un importante recurso marisqueiro en Europa que se atopa en regresión nos bancos naturais galegos (noroeste de España). Actualmente, en Galicia, só as cofrarías de Arousa e Cedeira comercializan esta especie, poñendo de manifesto o seu declive. Co propósito de conservar, xestionar e recuperar este recurso de maneira eficaz, realizáronse as seguintes análises xenéticas: En primeiro lugar, a avaliación xenético-poboacional de bancos naturais da Península Ibérica mediante marcadores de ADN mitocondrial (16S e Citb) para coñecer a diversidade xenética e a estrutura poboacional. Ademais, ante a presenza de *Donax vittatus* na costa atlántica da Península Ibérica, como outra especie para conservar e/ou xestionar, foi necesario o desenvolvemento de marcadores moleculares para caracterizar os seus bancos naturais e levar a cabo estudos similares aos realizados para *D. trunculus*. Así mesmo, dada a existencia doutras especies do xénero *Donax* (*Donax semistriatus* e *Donax variegatus*) na área de distribución a analizar, caracterizáronse o ADNr 5S, os espaciadores transcritos internos (ITS1 e ITS2), e os xenomas mitocondriales nas catro especies obxecto de estudo co fin de obter marcadores especie-específicos para súa identificación e construír as relacións filoxenéticas dentro da orde Veneroida.

Extended abstract

Introducción

La diversidad biológica del planeta está siendo debilitada paulatinamente como consecuencia directa o indirecta de las acciones del ser humano. Aunque se desconoce el número exacto, existe gran cantidad de especies que han desaparecido, mientras que otras muchas han sufrido un drástico descenso en su densidad, lo que ha provocado que estén en riesgo de extinción. En este contexto, la intervención del ser humano es necesaria para mejorar la gestión de los recursos naturales, así como para asegurar su mantenimiento.

En el campo de la acuicultura marina, los moluscos bivalvos constituyen uno de los grupos más importantes, tanto desde un punto de vista productivo como económico. En Galicia, su importancia económica y social es evidente al emplear de forma directa a miles de personas. De acuerdo con los datos presentados por la Xunta de Galicia, la producción de moluscos bivalvos en lonja en el año 2017 fue superior a 9.700 toneladas, representando un valor económico del orden de 81 millones de euros (datos procedentes de la Consellería do Mar, Xunta de Galicia). Sin embargo, la FAO estima que hasta 2030 habrá un estancamiento en la extracción de los bancos marinos mundiales (FAO, 2002). Por este motivo, los estudios encaminados a mejorar el manejo, conservación y explotación de estos recursos naturales, junto con el desarrollo y optimización de prácticas de acuicultura para nuevas especies, o especies ya explotadas, son extraordinariamente beneficiosos. Estos estudios pasan necesariamente por incrementar el conocimiento que se tiene de la biología de estas especies y, en este sentido, el estudio de la diversidad y diferenciación genética juega un papel clave para lograr una gestión adecuada. La conservación y el uso sostenible de los stocks productivos necesitan conocer los recursos genéticos disponibles para ser utilizados en la acuicultura y, al mismo tiempo, recopilar información de los bancos naturales (FAO, 1997, 2008). Esta información es necesaria para evitar la endogamia, mantener la integridad de los stocks y minimizar la transferencia de stocks genéticamente diferenciados (FAO, 1997). Diversos factores como las características biológicas de la especie; factores demográficos; aislamiento geográfico por barreras físicas; causas genéticas; factores físicos como los cambios bruscos en gradientes de temperatura y salinidad, o corrientes hidrodinámicas que favorecen la

retención larvaria; y factores históricos como los cambios climáticos y glaciaciones del Pleistoceno; entre otros, pueden provocar la subdivisión poblacional (Infante y Manchado, 2007; García-Merchán *et al*., 2012). Esto hace que sea importante examinar genéticamente las poblaciones naturales para conocer su grado de diferenciación y, de este modo, poder seleccionar la más apropiada para proporcionar los reproductores de criadero, así como generar información de referencia que permita gestionar la semilla producida y evaluar el impacto de las prácticas de repoblación (Ward, 2006). No obstante, y a pesar de su importancia comercial, los estudios genéticos en muchas especies de bivalvos son escasos, como es el caso de la coquina.

La coquina es un molusco bivalvo de gran interés comercial que presenta elevados índices de cotización en lonja. En Europa se han identificado diferentes especies de coquina pertenecientes al género *Donax,* y algunas de ellas son explotadas comercialmente en varios países como España (Ramón *et al*., 2005; Molares *et al*., 2008), Francia (Thébaud *et al*., 2005), Italia (Zeichen *et al*., 2002), Portugal (Gaspar *et al*., 1999; Chícharo *et al*., 2002; Pereira *et al*., 2012) y Turquía (Deval, 2009; Özden et al., 2009). En Galicia, la especie *Donax trunculus* (Linnaeus, 1758) presenta un elevado valor comercial, llegando a generar ingresos más altos que otras especies consideradas como principales. Atendiendo a las estadísticas de la Plataforma Tecnológica de la Pesca de la Xunta de Galicia (Consellería do Mar, Xunta de Galicia), en el año 2017 fue el bivalvo que alcanzó un mayor precio medio (42,79 €/kg), muy superior al cotizado en otras regiones de España y del sur de Portugal. No obstante, los bancos de la costa gallega están en regresión, debido a un posible agotamiento de este recurso marisquero. Esto se pone de manifiesto en el escaso número de lonjas que comercializan esta especie que, en la actualidad, son únicamente las lonjas de Cedeira (A Coruña) y Arousa (Pontevedra). Esta situación conlleva a que exista gran interés en la realización de estudios que permitan su recuperación y, en el futuro, la gestión sostenible de este bivalvo. Sin embargo, es conocido que las técnicas de repoblación pueden tener efectos deletéreos en la composición genética de las poblaciones naturales, como la pérdida de adaptación a condiciones locales o la pérdida de variabilidad (Ward, 2006). En consecuencia, para desarrollar una acuicultura sostenible y programas de repoblación adecuados, debe evaluarse tanto

la estructura como la variación genética de los bancos naturales (Frankham *et al*., 2002; Allendorf *et al.*, 2013). Los análisis de diversidad y diferenciación genética permiten obtener datos que contribuyen a la gestión sostenible, pues en el caso de encontrarse diferenciación genética entre los individuos de las localidades analizadas, ésta debería considerarse tanto en los planes de extracción como de repoblación, ya que excesos en las capturas podrían llevar a la extinción de bancos, y la repoblación tendría que respetar la diferenciación existente para no aumentar el riesgo de provocar cambios genéticos no intencionados o reducir la adaptación a las condiciones locales. Los individuos utilizados como reproductores en el criadero deben proceder preferentemente de la población a repoblar y, en su defecto, de otra genéticamente similar (Ward, 2006). Es por todo ello que se hacen necesarios los estudios genético-poblacionales en las especies objeto de recuperación.

Recientemente, estos estudios basados en marcadores microsatélite han proporcionado evidencias de que, en la Península Ibérica, *D. trunculus* muestra una clara diferenciación poblacional a lo largo de la transición Atlántico-Mediterráneo con diferentes grupos: el Océano Atlántico, el Mar de Alborán y el Mar Mediterráneo (Marie *et al*., 2016; Nantón *et al*., 2017; Rico *et al*., 2017). Sin embargo, la utilización de diferentes tipos de marcadores moleculares ha mostrado discrepancias en los patrones de diversidad en otras especies (Johnson *et al*., 2003; Brito, 2007; Hoffman *et al*., 2009; Toews y Brelsford, 2012; Ferchaud *et al*., 2015; Martínez *et al*., 2015). Por esta razón, es importante ampliar el espectro de marcadores con el fin de dilucidar los patrones de diversidad genética y el grado de diferenciación poblacional.

Para contribuir a la recuperación y gestión sostenible de una especie, además de los análisis genético-poblacionales, es necesario conocer las diferentes especies de un mismo género que coexisten en la misma región geográfica. Entre las diferentes especies de coquina pertenecientes al género *Donax*, cinco de ellas se encuentran a lo largo del litoral europeo: *D. trunculus* (Atlántica y Mediterránea), *Donax semiestriatus* (Atlántica y Mediterránea), *Donax variegatus* (Atlántica y Mediterránea), *Donax venustus* (Atlántica y Mediterránea), y *Donax vittatus* (Atlántica) (Ansell y Lagardère, 1980; Bayed y Guillou, 1985; Salas-Casanova, 1987; Salas-

Casanova y Hergueta, 1990; Rufino *et al*., 2010). Aunque *D. venustus* es prácticamente inexistente en la Península Ibérica, ya que solamente se ha encontrado un individuo en la costa del sur de Portugal entre los años 2000 y 2006 (Rufino *et al*., 2010), varias especies del género coexisten en simpatría en la península. Por ejemplo, *D. trunculus* y *D. vittatus* viven simpátricamente en las mismas playas de Portugal (Gaspar *et al*., 2002; Rufino *et al*., 2010). Además, aunque algunos estudios indican la existencia de *D. vittatus* en épocas anteriores (Bejega *et al*., 2010), no se había detectado su presencia en Galicia en las últimas décadas. No obstante, en un estudio multidisciplinar en el que ha colaborado nuestro grupo de investigación, se ha detectado la presencia de *D. vittatus* en la costa gallega, concretamente en las Rías Baixas (Praia América, A Lanzada y Corrubedo), mediante marcadores moleculares desarrollados en nuestro laboratorio (Pereira *et al.*, 2012; Nantón *et al*., 2015). Además de la coexistencia de estas dos especies, parece que éstas muestran una distribución vertical diferente: *D. trunculus* en el intermareal y *D. vittatus* en el sublitoral, como ya se ha descrito anteriormente en la costa atlántica francesa (Ansell y Lagardère, 1980).

Tradicionalmente, la identificación de moluscos bivalvos se ha realizado en base a criterios morfológicos. Sin embargo, la elevada variabilidad de este tipo de caracteres hace que el uso de marcadores moleculares resulte más adecuado para identificar de forma inequívoca a las diferentes especies. La mayor parte de estos marcadores están basados en la amplificación de ADN, nuclear (ADNn) o mitocondrial (ADNmt), mediante la reacción en cadena de la polimerasa (PCR) debido a la simplicidad, especificidad y sensibilidad que presenta esta técnica para amplificar regiones polimórficas de ADN genómico. Dentro de las diferentes regiones de ADN, destaca el ADN ribosómico (ADNr), ya que sus características hacen que estas secuencias resulten candidatas atractivas para la búsqueda de marcadores especie-específicos (Fernández *et al*., 2001; Cross y Rebordinos, 2006; Fernández-Tajes y Méndez, 2007). En moluscos bivalvos, los genes ribosomales representan las entidades genéticas más estudiadas tanto a nivel citogenético como molecular. De manera general, el ADNr se ha utilizado para conocer la estructura y organización de los genomas (Insua *et al*., 2001), establecer relaciones filogenéticas

(Insua *et al*., 2003; Jansen *et al*., 2006), diferenciar especies y/o evaluar la variabilidad genética de las poblaciones (López-Piñón *et al*., 2002; Freire *et al*., 2005; Fernández-Tajes y Méndez, 2007; Vierna *et al*., 2010; Pereira *et al*., 2012).

El hecho de que en Galicia *D. vittatus* esté presente, al menos, en las Rías Baixas, plantea la necesidad de realizar un estudio exhaustivo de su presencia en el resto de los bancos naturales de la costa gallega y desarrollar estudios genéticos similares a los realizados para *D. trunculus*. Para llevar a cabo la evaluación genética se requiere el desarrollo de marcadores moleculares especie-específicos para *D. vittatus*. Los estudios genéticos realizados en este trabajo, contribuirán a mejorar los programas de gestión y conservación de especies del género *Donax* en esta región. Así, si se confirma la presencia de bancos en los que coexisten al menos dos especies diferentes, sería necesario desarrollar estrategias diferentes para incrementar su producción y contribuir a su recuperación. En este contexto, destaca la utilidad de incluir en el estudio muestras procedentes de Portugal, puesto que, en el caso de considerarse la posibilidad de recuperar bancos naturales mediante repoblación, podría valorarse la opción de emplear individuos de las costas portuguesas, donde este bivalvo es muy abundante. Para ello, debe corroborarse previamente la ausencia de diferenciación genética entre los individuos procedentes de diferentes localidades.

El interés comercial que las especies del género *Donax* presentan para Galicia, plantea la necesidad de conocer los genomas de estos moluscos bivalvos para, posteriormente, aplicar el conocimiento de sus características genéticas a programas de explotación adecuados. Los datos obtenidos en este trabajo permitirán caracterizar diferentes regiones mitocondriales y nucleares de las diferentes especies de coquina presentes en Europa, con el fin de identificar marcadores moleculares que permitan diferenciarlas, y evaluar la variabilidad genética y diferenciación poblacional para el apoyo de estrategias de conservación de los bancos naturales.

Objetivos

En Galicia, las especies objeto de este trabajo son consideradas como recurso para las cofradías debido a sus elevados índices de cotización en lonja. No obstante, en la actualidad sólo las cofradías de Arousa y Cedeira comercializan este molusco bivalvo, lo cual pone de manifiesto el agotamiento del mismo, careciendo actualmente de datos resolutivos que justifiquen este hecho (explotación incontrolada, degradación de los espacios naturales por la acción del hombre, cambio climático, contaminación química, etc.). Ante la necesidad de aportar conocimientos científicos que ayuden a identificar el problema y permitir la puesta en marcha de medidas reguladoras y de control, con el fin de conservar y proteger estos recursos e incrementar su producción de manera sostenible, nos hemos planteado los siguientes objetivos específicos:

1. Evaluar la diversidad genética y la diferenciación poblacional en la coquina *Donax trunculus* de la Península Ibérica mediante marcadores de ADN mitocondrial.

2. Caracterizar marcadores de ADN mitocondrial y nuclear en *Donax vittatus*, incluyendo el desarrollo de *loci* microsatélite mediante *Next Generation Sequencing* (*NGS*), y evaluar genéticamente los bancos naturales de la Península Ibérica.

3. Caracterizar el ADNr 5S y los espaciadores transcritos internos, ITS1 e ITS2, y desarrollar nuevos marcadores para diferenciar especies del género *Donax*.

4. Determinar, por primera vez, las secuencias completas del genoma mitocondrial femenino en las cuatro especies del género *Donax* presentes en la Península Ibérica para compararlas con las de otros bivalvos marinos y establecer las relaciones filogenéticas dentro del orden Veneroida.

Capítulo 1

En este capítulo se estudió la variación genética de los marcadores mitocondriales 16S y citocromo b (Cytb) en individuos *D. trunculus* procedentes de 18 localidades de la Península

Ibérica (prestando especial atención al área noroccidental, Galicia, donde las poblaciones naturales han disminuido drásticamente) para comprobar si los patrones de diversidad genética, estructura poblacional e inferencia demográfica deducidos previamente con microsatélites son compatibles con los inferidos por estos nuevos marcadores. Además, dado que *D. trunculus* muestra doble herencia uniparental (DUI: *Double Uniparental Inheritance*), se analizaron ambos tipos de ADNmt (F y M).

Los resultados mostraron un alto nivel de diversidad genética y estructura genética significativa a lo largo de la Península Ibérica. Sin embargo, los dos tipos de ADNmt revelaron diferentes patrones geográficos. Los valores F_{ST} por pares obtenidos con las secuencias tipo F, mostraron dos grupos diferentes: el grupo Mediterráneo, que incluye las localidades situadas en el mar Mediterráneo (Gandía, Delta del Ebro, Maresme), y el grupo Atlántico, incluyendo todas las localidades atlánticas y la localidad de Fuengirola, ubicada en la costa del mar de Alborán. Mientras que los valores F_{ST} por pares para las secuencias tipo M, mostraron tres grupos: el grupo mediterráneo (Gandía, Delta del Ebro, Maresme), el grupo Atlántico (todas las localidades en la costa del Océano Atlántico) y el grupo del mar de Alborán (Fuengirola). Estos resultados ponen de manifiesto que las secuencias del genoma M son más adecuadas para investigar la estructura genética de las poblaciones de *D. trunculus*. Desde el punto de vista de la conservación, los tres grupos genéticamente divergentes confirmados (Océano Atlántico, Mar de Alborán y Mar Mediterráneo) deben considerarse como unidades de manejo diferentes. Esta diferenciación genética debe tenerse en cuenta en las estrategias de repoblación para evitar poner en riesgo la diversidad genética de la especie. En este capítulo, además, se proponen varias estrategias para la conservación y manejo de este bivalvo con alta importancia ecológica y comercial.

Capítulo 2

En el capítulo 2 se analizó, por primera vez, la diversidad genética y la estructura poblacional de *D. vittatus* a lo largo de toda la costa atlántica de la Península Ibérica. Para ello, se utilizaron cuatro genes mitocondriales (COI, Cytb, 16S F y M) y tres nucleares (H3, 18S y 28S). Estos

mismos marcadores moleculares también se secuenciaron en *D. semistriatus* y *D. variegatus* para abordar las relaciones filogenéticas de las especies del género *Donax* comunes a lo largo de las costas europeas.

Los resultados mostraron alta diversidad haplotípica en combinación con una baja diversidad nucleotídica y redes en forma de estrella con un haplotipo predominante, lo que indica una expansión poblacional reciente en las localidades de *D. vittatus* analizadas. Además, los análisis de diferenciación poblacional, realizados con el marcador mitocondrial COI, indicaron ausencia de estructura genética significativa en las localidades de la costa atlántica de la Península Ibérica.

En este capítulo sugerimos que *D. vittatus* podría ser un recurso alternativo potencialmente explotable como complemento de la explotación de *D. trunculus*, cuyas reservas naturales han disminuido drásticamente en algunas áreas. Además se presenta, por primera vez, la evidencia de DUI en las especies *D. vittatus* y *D. semistriatus.*

En conjunto, el presente capítulo proporciona datos útiles para taxonomía, ecología, conservación de la biodiversidad y también gestión de los recursos, mostrando que *D. vittatus* es una especie de interés para la acuicultura y que puede explotarse como complemento de *D. trunculus*.

Capítulo 3

En este capítulo se desarrollaron 15 *loci* microsatélite para *D. vittatus* mediante *NGS*, con el fin de revelar su estructura poblacional con un mayor número de marcadores y dilucidar las unidades de manejo apropiadas a lo largo de la costa atlántica de la Península Ibérica. Los 15 microsatélite se amplificaron en siete localidades (Corrubedo, A Lanzada, Praia América, Mira-Vagueira, Lisboa, Setúbal y Sines), seis de las cuales se estudiaron en el capítulo anterior con el marcador mitocondrial COI.

Todos los *loci* resultaron polimórficos y las localidades mostraron valores similares de riqueza alélica y de heterocigosidad observada y esperada (10,750-15,201; 0,813-0,845; y 0,389-0,460; respectivamente). Además, se detectó un déficit significativo de heterocigotos para todas las

localidades, siendo los alelos nulos el factor más probable a contribuir en estos déficits. En cuanto a los análisis de diferenciación poblacional, incluyendo la estimación del F_{ST} global, los valores F_{ST} por pares (tanto con la corrección de alelos nulos como sin ella), el número de migrantes y el AMOVA, indicaron ausencia de estructura genética en *D. vittatus* en la región estudiada, confirmando los resultados del capítulo anterior basados en el gen COI.

Estos datos son útiles para ayudar al desarrollo de la pesquería y los planes de conservación de esta especie. Además, los nuevos marcadores microsatélite desarrollados permitirán llevar a cabo otros estudios genético-poblacionales a lo largo de la distribución geográfica de la especie, así como análisis genéticos necesarios para las actividades de acuicultura.

Capítulo 4

En este capítulo, se analizaron las unidades de repetición completas del ADNr 5S y del espaciador transcrito interno (ITS) en las cuatro especies del género *Donax* presentes en Europa. Tras amplificar, clonar y secuenciar varias unidades 5S e ITS, se describieron sus características básicas y su variación genética.

El tamaño de la unidad de repetición del ADNr 5S presentó poca variación entre especies, excepto en *D. trunculus* que difirió del resto de las especies *Donax* en 170-210 pb. El tamaño inferido para la región codificante resultó ser en todos los casos 120 pb, y mostró regiones de control interno (ICRs) implicadas en la transcripción; mientras que, la longitud de la región espaciadora no transcrita (NTS) osciló entre 157-165 pb en *D. trunculus* y entre 335-367 pb en las otras tres especies. También se analizó el grado de conservación de las regiones reguladoras de la transcripción, revelando una región conservada similar a una caja TATA localizada aguas arriba.

En cuanto a las secuencias ITS, las cuatro especies *Donax* mostraron ligeras diferencias de tamaño entre los clones debido a la variación existente en el ITS1 y el ITS2, excepto *D.*

variegatus que no mostró diferencias de tamaño en el ITS2. La longitud total de las secuencias ITS osciló entre 814 y 1014 pb.

Éste no es solo un trabajo de investigación básica, donde se aportan nuevos datos y conocimientos acerca de las especies *Donax*, de las que no hay mucha información genética, sino que también ofrece una aplicación viable en acuicultura que es tratada en el siguiente capítulo.

Capítulo 5

En este capítulo, se desarrolló una PCR *multiplex* de las dos regiones de ADNr caracterizadas en el capítulo anterior (5S e ITS) para identificar las cuatro especies del género *Donax* presentes en Europa: *D. semistriatus*, *D. trunculus*, *D. variegatus* y *D. vittatus*. Los perfiles electroforéticos especie-específicos permitieron su identificación de una manera rápida, simple, eficiente y de bajo coste, ofreciendo una aplicación factible tanto en laboratorios no especializados, como en laboratorios de inspección alimentaria, particularmente cuando se analiza un elevado número de muestras. Este método, además de ofrecer la precisión suficiente para satisfacer los requisitos de autentificación de las especies *Donax*, también permitirá caracterizar los reproductores utilizados en el criadero para obtener su correcta identificación taxonómica.

Capítulo 6

En el capítulo 6, se secuenciaron y caracterizaron los genomas mitocondriales femeninos completos de las cuatro especies de coquina comunes en la Península Ibérica (*D. semistriatus*, *D. trunculus*, *D. variegatus* y *D. vittatus*), siendo los primeros representantes de la familia Donacidae analizados al respecto.

Los genomas mitocondriales de las cuatro especies resultaron ser moléculas cerradas compuestas por 37 genes: 13 genes codificantes de proteínas (incluyendo el gen atp8), 2 genes que codifican para los ARNs ribosómicos (ARNr) y 22 genes que lo hacen para ARNs de transferencia (ARNt), todos ubicados en la misma cadena. Los cuatro genomas mitocondriales varían en tamaño, desde

17.044 pb (*D. semistriatus*) hasta 17.365 pb (*D. trunculus*), y en contenido A+T, desde 58,9% (*D. trunculus*) a 63,5% (*D. vittatus*). Las diferencias de longitud se deben principalmente a la variación de tamaño de la región no codificante más larga (NCR), la cual se identificó en cada una de las cuatro especies *Donax* entre los genes cob y cox2, y que posiblemente se trate de la Región de Control. Además, se encontró el gen atp8 en ocho especies del orden Veneroida, mejorando los estudios previos en los que se indicaba su ausencia.

En cuanto a los análisis filogenéticos del orden Veneroida, llevados a cabo por Inferencia Bayesiana, tanto en MrBayes como en Phylobayes, y bajo Máxima Verosimilitud, mostraron árboles congruentes con alto soporte en la mayoría de las ramas, e indicaron que las cuatro especies de *Donax* forman un solo clado como un grupo hermano de otros bivalvos dentro de la superfamilia Tellinoidea. Sin embargo, aunque Tellinoidea es en realidad monofilética, ninguna de sus familias lo es. También se comparó la disposición génica de los cuatro genomas mitocondriales con la de otras especies relacionadas pertenecientes a la superfamilia Tellinoidea. Esta comparación fue realizada previamente por Yuan *et al.* (2012a), sin tener en cuenta el gen atp8 y sin incluir la especie *Macoma balthica* y las especies del género *Donax*. Sus resultados apoyaron la conclusión de que las comparaciones de la disposición de los genes del ADNmt son, en cierta medida, una herramienta útil para los estudios filogenéticos. En este trabajo, siete de los diez genomas mitocondriales de especies de la superfamilia Tellinoidea disponibles hasta la fecha (incluidos los genomas mitocondriales de *M. balthica*, *Moerella iridescens*, *Solecurtus divaricatus* y de las cuatro especies *Donax* analizadas en este capítulo) muestran una disposición génica idéntica, y el genoma mitocondrial de *Soletellina diphos* sólo difiere en la falta de un ARNt-Phe. El gen atp8 muestra la misma ubicación dentro del genoma mitocondrial de estas ocho especies de la superfamilia Tellinoidea, específicamente entre ARNt-Met y ARNt-Ser1.

En este capítulo, no sólo se ha incrementado el número de secuencias de genomas mitocondriales dentro del orden Veneroida, sino que también se han ilustrado las relaciones filogenéticas entre las especies *Donax* y su posición dentro de este orden. Los resultados demuestran que la secuenciación de los genomas mitocondriales completos proporciona información muy valiosa

para el análisis filogenético en bivalvos. Además, este trabajo proporciona información básica y de gran interés para la comunidad científica que puede utilizarse para su aplicación en acuicultura. De hecho, las secuencias de ADNmt aportadas en este capítulo añaden marcadores genéticos significativamente útiles para i) ayudar a diferenciar estas especies comerciales morfológicamente similares, ii) detectar y evitar el fraude, iii) proteger los derechos del consumidor y lograr otros objetivos de calidad, como el certificado de origen y iv) utilizar en estudios de genética de poblaciones para la gestión sostenible de estas especies. Sin embargo, esta posible aplicabilidad requiere un trabajo más amplio, donde los diferentes marcadores sean probados en un mayor número de individuos, no sólo en individuos frescos sino también en procesados, envasados o congelados, y tanto en machos como en hembras.

Introduction

Biology of the genus *Donax*

The wedge clams of the genus *Donax* are bivalve molluscs belonging to the Bivalvia class, within the order Veneroida and the family Donacidae. *Donax* species are an important constituent of the macrofauna of open sandy beaches in temperate, tropical and subtropical zones, being the dominant organisms in this type of environment (Ansell, 1983). They are usually found in exposed areas such as the intertidal or shallow sublittoral of clean beaches. Their success can be attributed to their burial and migration ability in response to physical conditions such as waves, tides and currents (Trueman, 1971; McLachlan et al., 1979; Leber, 1982; Ansell, 1983; Ellers, 1995; Gaspar et al., 1999; Rufino et al., 2010). They are more abundant in eutrophic areas, with high production of plankton, since they are suspension feeders (Wade, 1964; Mouëza and Chessel, 1976). *Donax* species show little variability in their life cycle, being all gonochoric, with equal sex ratio, planktonic larval development and with primary sedimentation occurring in early development stages. Most species have a lifespan of one to four years, with rapid growth and early maturity, although growth, reproduction and sedimentation show seasonal responses depending on environmental conditions (Ansell, 1972; Gaspar et al., 1999; Gaspar et al., 2002).

The shell is calcareous, wedge-shaped, small to medium size, with equal and not gaping valves and weak sculpture. The umbo is a third of the distance from the hind end of the shell, the ligament is external, and the hinge has two cardinal teeth per valve (Brown and McLachlan, 1990) (Figure 1).

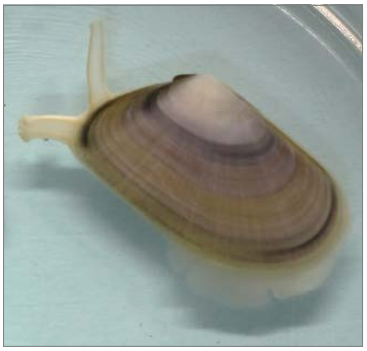


Figure 1. External appearance of the wedge clam *Donax trunculus*.

Five *Donax* species occur along the European littoral: *Donax vittatus* (Da Costa, 1778) belonging to the Mediterranean boreal fauna and extending to the coast of Norway; *Donax semistriatus* (Poli, 1775) with a limited distribution to the Mediterranean and Black Seas and the Portuguese coast (Rufino et al., 2010); *Donax trunculus* (Linnaeus, 1758) and *Donax variegatus* (Gmelin, 1791), warm water Atlantic-Mediterranean species that reach the Black Sea and the Atlantic coast of France, and the southwest coasts of France and England, respectively (Ansell and Lagardère, 1980; Bayed and Guillou, 1985; Salas-Casanova, 1987); and *Donax venustus* (Poli, 1775), distributed throughout the Mediterranean and Black Seas. However, *D. venustus* is practically non-existent in the Iberian Peninsula as just one single individual has been found between the years 2000 and 2006 along the south coast of Portugal (Rufino et al., 2010).

Thus, in the littoral of the Iberian Peninsula, several species of the genus can coexist in sympatry. For instance, *D. trunculus* and *D. vittatus* live sympatrically in the same beaches of Portugal (Gaspar et al., 2002; Rufino et al., 2010). Moreover, although Bejega et al. (2010) recorded the existence of *D. vittatus* in earlier times in Galicia (northwest of Spain), its presence in this area has not been detected in the last decades. However, in a multidisciplinary study in which our research team has collaborated, the presence of *D. vittatus* has been detected in the Galician coast (specifically in three beaches of the Rías Baixas: Praia América, A Lanzada and Corrubedo).

Where more than one species of *Donax* occurs within an area, there is generally little habitat overlap between species. Three main factors are important in the separation of species in such instances; bathymetric depth, sediment composition and wave energy, although these are also clearly interrelated. Other factors such as tolerance of extreme temperatures and air exposure may also be involved. Typically, geographically-sympatric species of *Donax* occupy three distinct habitats: the intertidal zone of moderately-exposed or exposed beaches with clean well-sorted sand, the shallow sublittoral zone of such beaches, and an offshore zone where coarse sediments occur under the influence of strong currents (Ansell, 1983). Thus, in areas where *D. vittatus* and *D. trunculus* coexist, as is the case of Galicia, Ile d'Oléron (France) and Portugal, *D. trunculus* is found in the intertidal zone, while *D. vittatus* is found mainly in the shallow sublittoral, up to a

depth of 5-6 m (Lagardère, 1972; Rufino et al., 2010). In the case of southern Portuguese coast, where the four species of the genus coexist, these show sequential maximum occurrences, from shallower to deeper distribution: *D. trunculus*, *D. vittatus*, *D. semistriatus* and *D. variegatus* (Rufino et al., 2010).

Importance of *Donax* species

Few species of the genus *Donax* are commercially exploited, but some are consumed locally or used as fishing bait. *D. trunculus* is exploited in many countries surrounding the Mediterranean Sea and Atlantic Ocean, including France (Thébaud et al., 2005), Italy (Zeichen et al., 2002), Portugal (Gaspar et al., 1999; Chícharo et al., 2002; Pereira et al., 2012), Turkey (Deval, 2009; Özden et al., 2009), and Spain (Ramón et al., 2005; Molares et al., 2008).

Only in the Iberian Peninsula, the recorded captures from 1999 to 2016 equal 10,643 tonnes, with a maximum production of 1,042 tonnes in 2005 followed by an incessant decline reaching only 195 tonnes in 2016 (FAO-FIGIS, 2018). Although these data only reflect production since fishermen were obliged to declare their captures (Gaspar et al., 1999), the species has been subjected to intense exploitation over the last decades and, currently, some *D. trunculus* populations seem to be at high long-term risk of extinction (Marie et al., 2016). Furthermore, this species constitutes an important shellfish resource due to its high economic value. In Galicia (northwest of Spain), *D. trunculus* is a species with a high contribution rate, being the bivalve with the greatest commercial value in markets over the last few years. For instance, in 2017 it reached the highest average commercial value among bivalves (42.79 €/kg) but was one of the least-sold species (171.10 kg) in the markets of the Galician Community (data from http://www.pescadegalicia.gal/estadisticas/ Consellería do Mar, Xunta de Galicia) (Figure 2). Although its natural beds have started to show signs of decline, demand for wedge clams has continued to increase. In fact, the level of captures has decreased in the last seventeen years from ~17 tonnes (2001) to 171.10 kg (2017) (Consellería do Mar, Xunta de Galicia) (Figure 3). This

decline could be related to the exploitation to which this species has been subjected, which might have caused the regression of its natural beds in the Galician coast. Moreover, this decrease has caused a drastic drop in the number of Galician fishermen's associations that exploit the resource and, in the last fourteen years, the extraction has been practically monopolised by the association of Cedeira (Figure 3). At present, only the fishermen's association of Arousa (108.05 kg) and Cedeira (365.39 kg) commercialise this bivalve mollusc (Data from Consellería do Mar, Xunta de Galicia, May 24th, 2018).

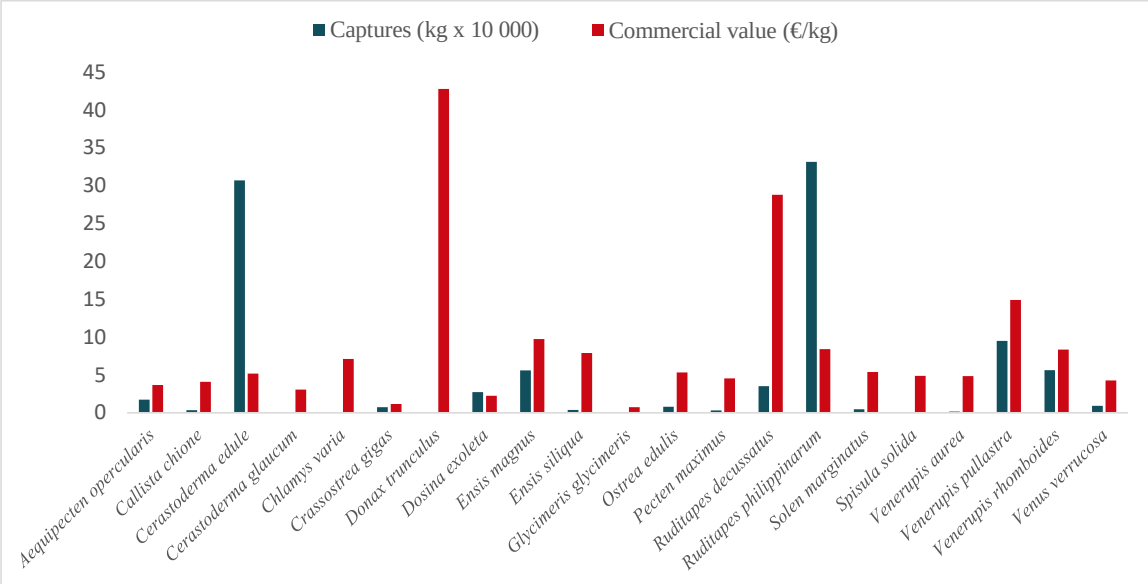


Figure 2. Level of captures (kg x 100 000) and average commercial value (€/kg) of the bivalve molluscs sold in Galician fishermen's associations during the year 2017. Data from "Plataforma Tecnológica de Pesca de la Xunta de Galicia".

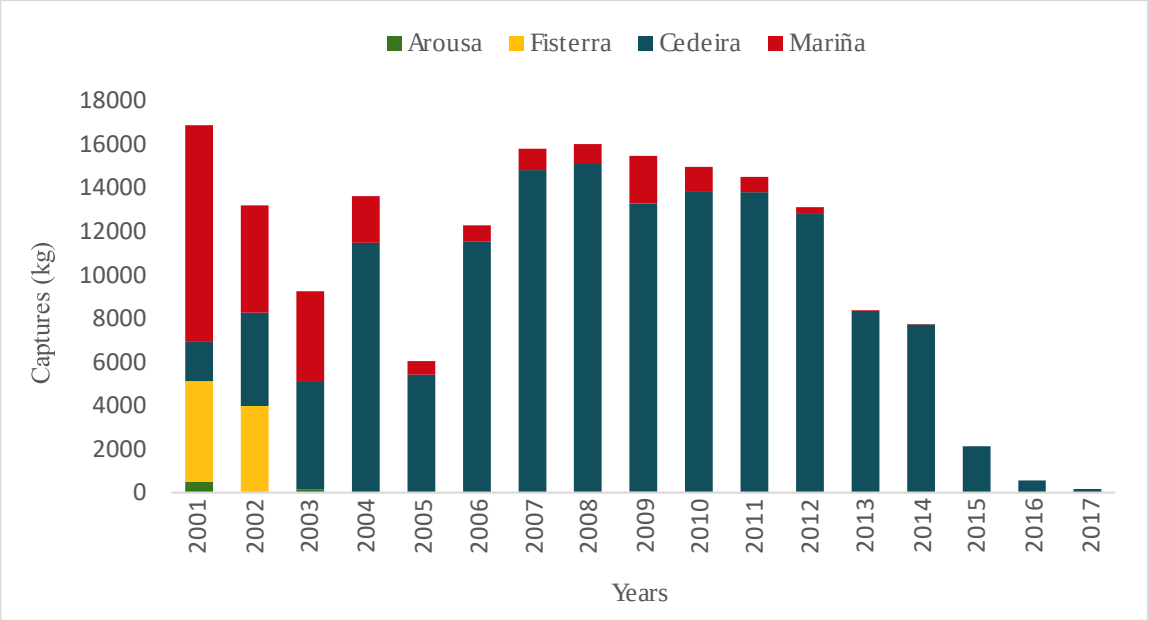


Figure 3. Captures of the wedge clam *D. trunculus* according to the first sale data registered by the "Plataforma Tecnológica de Pesca de la Xunta de Galicia". Each colour represents a fishermen's association (Arousa, Fisterra, Cedeira and Mariña).

Although *Donax* species can be distinguished by their morphological traits, this is not a straightforward task for consumers since the shell of these species is very similar in size, shape and colour (Figure 4). Even within each species, the pattern and coloration of shells are very variable, with specimens in yellow, tan and violet tones. Therefore, it is likely to find other *Donax* species with lower economic value being sold as *D. trunculus*. Nevertheless, despite difficulty in species differentiation, overexploitation and economic importance for the shellfish sector, relatively few genetic resources are available for this species and the whole genus.

Figure 4. External appearance of the four *Donax* species from the Iberian Peninsula: *D. semistriatus* (1), *D. trunculus* (2), *D. variegatus* (3) and *D. vittatus* (4).

Genetic studies in European *Donax* clams

With respect to the genus *Donax*, two methods for the molecular identification of European *Donax* species have been developed. The first protocol is based on the amplification of the nuclear marker 5S rDNA and allows for the differentiation of *D. trunculus* and *D. variegatus* from the remaining *Donax* species (Pereira et al., 2012). The second is based on PCR-RFLP (Polymerase Chain Reaction - Restriction Fragment Length Polymorphism) analysis of a cytochrome c oxidase subunit I (COI) fragment, with which *D. semistriatus*, *D. trunculus*, *D. variegatus* and *D. vittatus* have been identified (Nantón et al., 2015). Lastly, a work that deals with the description and study of the karyotype of *D. vittatus* and compares it with the karyotype of *D. trunculus* has recently been published (García-Souto et al., 2017).

Regarding *D. trunculus*, most of the works have focused on studying the karyotype of the species (Cornet and Soulard, 1990; González-Tizón et al., 1998; Martínez et al., 2002), as well as on the analysis of mobile elements and satellite DNA (Plohl and Cornudella, 1996, 1997; Plohl et al., 2002; Petrović and Plohl, 2005; Petrović et al., 2009; Satovic and Plohl, 2013). In addition, Theologidis et al. (2008) have studied the mode of inheritance of mitochondrial DNA. Recently, microsatellite markers have also been developed (Nantón et al., 2014; Rico et al., 2017) and population genetic analyses based on this type of molecular markers have been carried out (Marie et al., 2016; Nantón et al., 2017; Rico et al., 2017).

Genetic aspects of the conservation and management of fishery resources

Biodiversity is rapidly decreasing, to a large extent due to anthropogenic activity that compromise the organisms' ability to adapt to environmental changes (Saura and Faria, 2011). Among the most common threats to the marine environment stand out overexploitation, uncontrolled harvesting, water pollution, habitat degradation and introduction of exotic species (Saura and Faria, 2011). Human harvest of animals in the wild is often intense and occurs in terrestrial and

aquatic habitats throughout the world. Harvest has the potential to cause three types of genetic change: alteration of population subdivision, loss of genetic variation and selective genetic changes. For instance, individuals with commercially desirable phenotypic traits have more probability to be extracted from natural beds. If the phenotype has a genetic base (i.e. it is heritable), genetic changes will happen in the harvested natural population (Allendorf et al., 2008). Such changes are likely to reduce the frequency of attractive phenotypes and can cause a decrease in different factors, such as fertility, larval viability and growth rate. This could generate both short-term effects, reducing the productivity of exploited populations and the recovering capacity of overexploited populations, and long-term consequences, diminishing the ability of populations to evolve and increasing the extinction risk of species (Olsen et al., 2004; Walsh et al., 2006). To sustain the productivity of harvested populations, it is crucial to incorporate genetic considerations into management (Allendorf et al., 2008).

Early conservation efforts have most often focussed on the effects of abiotic habitat factors on species and on the complex relationships among species with the intention of giving detailed descriptions of the species´ habitat requirements. However, more recent conservation approaches have shown that ecological studies can greatly benefit from their combination with genetic studies (Geist, 2010). Fisheries genetics is the application of genetic principles and methods to fisheries biology and management, focusing on solving problems dealing on how to administrate fisheries resources. In this field the main areas of interest are stock identification and traceability.

Population genetic studies

Genetic analysis plays important roles in fisheries management, helping to identify stock resources and providing potentially useful management suggestions. Correct delineation of marine species boundaries is essential for conservation (Avise, 1998; Sweijd et al., 2000). To ensure long-term sustainability of wild stocks, management programmes should recognize two levels of genetic diversity; within and among localities.

On the one hand, maintenance of genetic diversity is a main goal in conservation biology. In the short-term, genetic variation is relevant to the preservation of reproductive success and to population productivity, while in the long-term, it contributes to adaptive potential (Frankham et al., 2002, Allendorf et al., 2008). Different factors affect genetic diversity and their combination determines levels of variation in populations. Mutation and migration contribute to increase the variation, while balancing selection maintains it, and genetic drift and directional selection reduce the variation (Frankham et al., 2002). The balance between the forces altering genetic diversity depends on population sizes. In large populations, selection has an important effect, while in small populations, genetic drift is the most relevant factor (Frankham et al., 2002).

On the other hand, genetic differentiation among populations and correct delineation of marine species boundaries are also essential to species conservation and management (Avise, 1998; Sweijd et al., 2000) because these aspects provide information about the genetic exchange among them and their evolutionary history (Allendorf et al., 2013). In many marine organisms, different factors contribute to the absence of genetic differentiation among their populations, including large population sizes and high gene flow, due to a planktonic life stage that can result in movement over large distances and the lack of geographical barriers to larval movement (Riginos and Nachman, 2001; Launey et al., 2002). However, an increasing number of studies that demonstrate the existence of population subdivision in marine species exist. Although the predominant mechanisms leading to population differentiation are not always clear, several factors may be important either singly or in combination, such as abrupt changes of temperature and salinity, oceanographic currents that promote larval retention, local adaptation, or historical factors, as climate change or Pleistocene glaciations (Riginos and Nachman, 2001; Infante and Manchado, 2007). In this context, evaluation of population differentiation is fundamental to assess the genetic risk associated with restocking programmes (Ward, 2006) and translocation of individuals between natural populations (Johnson, 2000). It is necessary to have a good understanding of population structure before carrying out any restocking or stock enhancement projects (Cross, 2000), since the introduction of genotypes that are not very representative of the

natural population can have negative effects. Genotype introductions from genetically similar populations are likely to have smaller negative effects than introductions from genetically dissimilar populations. The purpose of restocking and stock enhancement programmes is generally to increase the biomass of the target species. Ideally, this should be achieved with little or no adverse effect on the local gene pool (Ward, 2006). Therefore, individuals used as breeding stock in the hatchery should preferably come from the population to be repopulated, but where this is not feasible (when it is not possible to take broodstock from the stock to be enhanced or the local population is very small and suspected to be highly inbred), the most genetically similar population should be used (Ward, 2006).

For these reasons, obtaining basic data on population genetics using molecular markers is valuable for designing restocking programmes and managing strategies that minimise harmful genetic changes caused by exploitation of wild populations.

Species identification

Nowadays, there is a growing requirement to develop techniques that allow consumers to know what they purchase, guaranteeing the correct identification and traceability of seafood products in order to ensure the composition and safety of commercial marine products, as well as avoiding commercial fraud, in which one seafood species is illegally substituted for another of lower economic value (Céspedes et al., 1999; Civera, 2003; Rasmussen and Morrissey, 2008). European Union labelling regulations specify that seafood products must be correctly labelled with their commercial and scientific names to guarantee their traceability and correct identification throughout the value chain (EC 104/2000; EC 2065/2001; EU 1379/2013). The recent Art. 35 of the 1379/2013 EU Regulation on Common Market Organization requires that consumers receive precise information about the seafood they purchase, such as scientific and commercial names, production method, catch area and fishing gears used (D´Amico et al., 2016). In addition to its negative commercial effects, incorrect species identification can reduce the effectiveness of

management and conservation programmes of marine species, designed to protect the oceanic habitats and the endangered species (Civera, 2003).

Traditionally, identification of bivalve mollusc species has been based on morphological criteria. However, this process requires a meticulous analysis of features that often show great variability since they are subject to environmental influences. In consequence, correct species identification based on morphological features is unreliable and may be difficult for non-specialists. Moreover, the authentication of commercial food species from canned or frozen individuals (without shells) or larval and juvenile stages is impossible to carry out due to these forms are morphologically unclassifiable.

For these reasons, molecular and biochemical techniques for species authentication are needed, since they could allow for their identification regardless of their life stage (larva, juvenile or adult) or state (fresh or processed). The first markers designed for this purpose in aquaculture genetics were based on proteins (Buroker et al., 1979a,b; May et al., 1980; Seeb and Seeb, 1986). Nevertheless, protein analysis has a number of disadvantages when compared with DNA-based markers, such as their smaller variability, that in many cases prevents species differentiation. Moreover, enzymatic degradation during the processing of products or the absence of enzymes in the tissues to be analysed may also result in difficulty of identification (Cross and Rebordinos, 2006). The use of molecular methods based on DNA can overcome these difficulties because DNA is a very stable and long-lived biological molecule present in all tissues of all organisms. Furthermore, the introduction of PCR has simplified previous molecular methods that were tedious, expensive and time-consuming. For species identification, both nuclear and mitochondrial DNA markers can be used, but ribosomal DNA (rDNA) is particularly useful, given that it has both slowly and rapidly evolving regions. The first are useful for comparing distantly related species and the second regions are appropriate for differentiating more closely related groups.

Molecular markers

All organisms are subject to mutation as a result of normal cellular operations or interactions with the environment, leading to genetic variation or polymorphism. Genetically inherited variation, either in the form of an identifiable phenotypic character or a distinguishable genetic mutation, which can be inferred through molecular techniques, is known as a genetic marker (Liu and Cordes, 2004). Markers based on the analysis of DNA or proteins are called molecular markers and are the tools used to reveal and quantify genetic variation. The advance of biochemistry and molecular biology has allowed for the development of different types of molecular markers.

Allozymes, allelic variants of proteins produced by a single locus (Liu and Cordes, 2004), stand out within the group of protein-based markers. Despite the fact that they are among the earliest markers used in aquaculture (May et al., 1980; Seeb and Seeb, 1986), provide a high degree of sensitivity and can be used successfully with very small quantities of material (Hu et al., 1992), allozymes have several disadvantages, such as their differential expression depending on the tissue and stage of development and that most of useful proteins are denatured during the manufacturing of processed products and rendered useless for analysis (Sweijd et al., 2000). Due to these and other limitations, allozymes have been largely replaced by DNA markers.

Regarding DNA-based markers, they can be categorized into two classes, nuclear DNA (nDNA) and mitochondrial DNA (mtDNA) markers. Depending on the selected genomic region and the type of variation examined, there are different types of DNA markers such as Random Amplified Polymorphic DNA (RAPDs), in which fragments of DNA are amplified using one single primer of random sequence; Restriction Fragment Length Polymorphisms (RFLPs), originated as a consequence of nucleotide changes in the recognition sequence of a restriction enzyme; Amplified Fragment Length Polymorphisms (AFLPs), that combine characteristics of the two previous markers and are obtained by the selective amplification of fragments obtained after DNA digestion with restriction enzymes; Microsatellites also known as Short Tandem Repeats (STRs) or Simple Sequence Repeats (SSRs), stretches of short DNA sequence in which a motif from 2 to 6 bp is tandemly repeated; and Single Nucleotide Polymorphisms (SNPs) caused by changes in

one single nucleotide base in a specific position of the genome (Chambers and MacAvoy, 2000; Liu and Cordes, 2004; Rasmussen and Morrissey, 2008).

Mitochondrial DNA markers

Most metazoan mtDNAs are typically closed circular molecules ranging in size from 14 to 18 kb that encode 37 genes: 13 protein-coding genes (PCGs), 22 transfer RNA (tRNA) genes and two ribosomal RNA (rRNA) genes (Wolstenholme, 1992). In addition, at least one extensive non-coding sequence is present, which contains elements that control the initiation of replication and transcription (Hoffman et al., 1992). Size variation in mtDNA is usually due to differences in the non-coding sequence (Brown, 1985; Harrison, 1989), but it is occasionally due to duplications of other portions of genome (Moritz and Brown, 1986, 1987; Zevering et al., 1991).

mtDNA sequences are one of the most popular molecular markers and have been extensively used for studying genetic diversity, population structure, and phylogeography, defining stocks and identifying species (Moritz et al., 1987; Arnaud-Haond et al., 2003; Katsares et al., 2008; Luis et al, 2011; Martínez et al., 2015). Furthermore, the advance of molecular techniques, such as Next-generation Sequencing (NGS), and bioinformatic programmes have made it quick, easy and cheap to sequence and assemble entire mitochondrial genomes, so that there have recently been significant increases in the number of complete mitochondrial sequences available. The complete mitochondrial genome is regarded as the marker of choice for the reconstruction of phylogenetic relationships at several taxonomic levels, from populations to phyla, and has been widely used for the resolution of taxonomic controversies (Gissi et al., 2008). It has proved to be useful as a tool for population genetic and phylogenetic studies not only because complete mitochondrial genomes are often more informative than single genes, but also because they reveal some genome-level details, such as the rearrangement of genes, that are valuable information for studies of evolutionary relationships among species (Boore and Brown, 1998; Rokas and Holland, 2000; Gissi et al., 2008; Shen et al., 2009). Besides its use as a phylogenetic marker, mtDNA represents a “full” genome and several features, for example, genome size, gene content, gene order,

compositional features, non-coding sequences and secondary structure of the encoded RNA, can be quite easily investigated in the small mitochondrial genome (Gissi et al., 2008).

Furthermore, mitochondrial genome has several valuable features that make it unique, including its fast evolutionary rate (higher than that of single-copy nDNA), the presence of genes or regions evolving at different rates, and a relatively conserved gene content and organization (Gissi et al., 2008; Shen et al., 2009). The mitochondrial genome is also haploid and generally presents clonal inheritance without genetic recombination, which implies that the entire genome behaves as a single locus with multiple alleles. In addition, its effective size is four times lower than that of nuclear markers, making it more sensitive to the effect of genetic drift (Ferguson et al., 1995; Ballard and Whitlock, 2004; Zink and Barrowclough, 2008). All these features make this molecule a reliable and easy to use marker.

Strict maternal inheritance (SMI) is considered to be the paradigm of mtDNA transmission in animal species (Birky, 2001). However, an extreme exception is found in some bivalve molluscs, which possess an unusual and unique system known as Doubly Uniparental Inheritance (DUI) (Hoeh et al., 1991, 1996; Skibinski et al., 1994a,b; Zouros et al., 1994a,b; Liu et al., 1996). Species showing DUI display two distinct sex-associated mitochondrial lineages: the maternal or female type (F type) that is transmitted through the eggs to all offspring, and the paternal or male type (M type) that is present in sperm and enters all eggs at the time of fertilization but is only retained and transmitted through male offspring. In adults, the F type is predominant in all tissues of both sexes except in the male gonad, where the M type prevails. Although some exceptions have been documented, adult females are essentially homoplasmic and adult males are heteroplasmic (Figure 5) (reviewed in Breton et al., 2007; Passamonti and Ghiselli, 2009; Zouros, 2013).

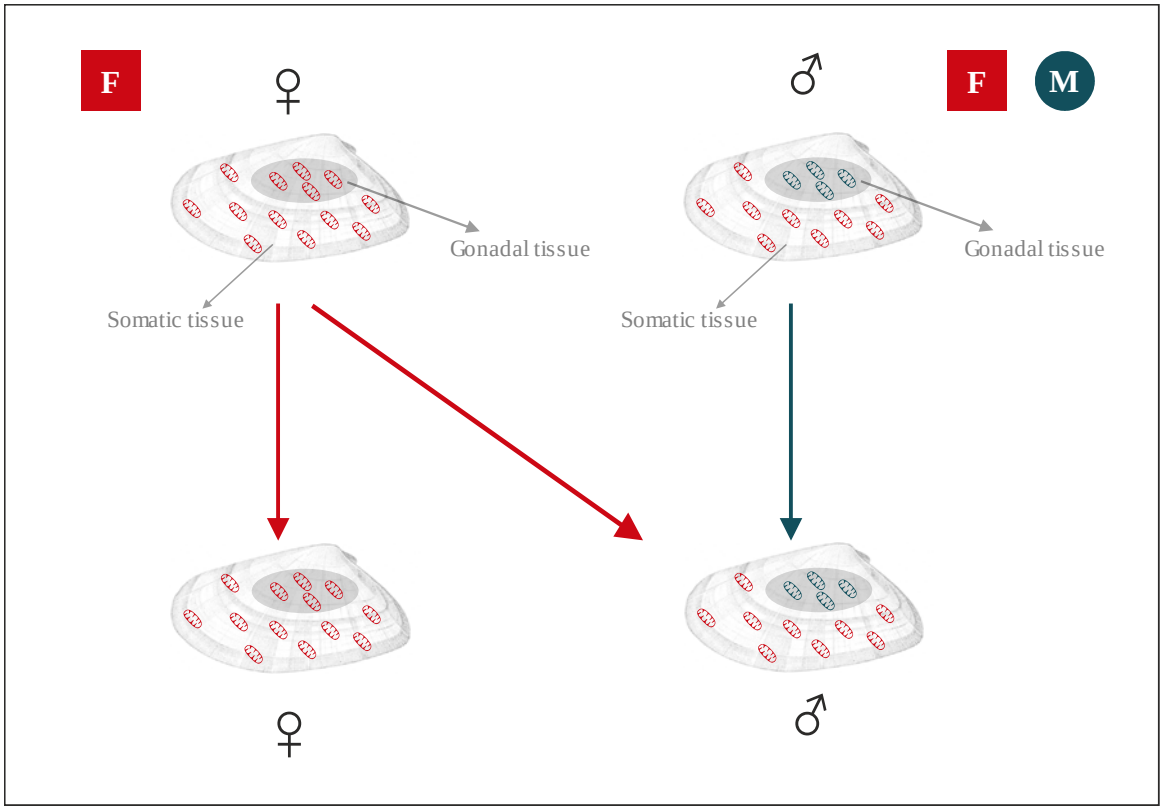


Figure 5. Diagram of the DUI system in *D. trunculus*. Distribution of mitochondrial types by tissue and sex. F type is represented in red and M type in blue.

To date, most reported DUI species belong to freshwater bivalves of the order Unionoida (families Hyriidae, Margaritiferidae, Unionidae) (Liu et al., 1996; Curole and Kocher, 2002; 2005; Hoeh et al., 2002), marine mussels of the order Mytiloida (family Mytilidae) (Fisher and Skibinski, 1990; Skibinski et al., 1994a; Zouros et al., 1994a,b; Passamonti, 2007; Vargas et al., 2015), clams of the order Veneroida (families Articidae, Donacidae, Mactridae, Solenidae, Veneridae) including the wedge clam *D. trunculus* (Passamonti and Scali, 2001; Theologidis et al., 2008; Plazzi, 2015; Plazzi et al., 2015; Dégletagne et al., 2016), and protobranchs of the order Nuculanoida (family Nuculanidae), an order belonging to the most basal protobranch bivalve lineage (Boyle and Etter, 2013; Gusman et al., 2016).

The combination of both types of mtDNA (F and M types) can provide information on the evolutionary history and phylogeographic structure of the species and it can be also a valuable tool for genetic evaluation of populations and species conservation (Passamonti and Scali, 2001;

Breton et al., 2009; Huang et al., 2013). For instance, the M type mtDNA can be used as a new marker to estimate genetic diversity so that its results can be compared with those of the F type mtDNA. For this reason, the analysis of both types of mtDNA can be relevant to the study of genetic diversity and population structure in species with this mode of inheritance.

Although the mitochondrial genome of some bivalves has evolutionary properties different from those traditionally observed, there are several reasons that justify its use as a molecular marker. These include its small size, very few duplications, no introns, very short intergenic regions (Gissi et al., 2008) and its relative ease to amplify, because it appears in multiple copies in the cell and is abundant in animal tissues (Galtier et al., 2009). Likewise, the existence of variable regions flanked by conserved ones, suitable for primers design, means that they can be easily amplified by PCR. For these reasons, mtDNA constitutes one of the simplest, most convenient and cheapest solutions to genetically study a new species (Galtier et al., 2009).

Nuclear DNA markers

Microsatellites

Microsatellites, also known as STRs or SSRs, are short tandem repeat DNA sequences that range in size from 2 to 6 bp, which are abundantly distributed across genomes and demonstrate high levels of allele polymorphism (Chambers and MacAvoy, 2000; Chistiakov et al., 2006). They are present in all prokaryotic and eukaryotic genomes analysed to date, being located anywhere in the genome, both in protein-coding and non-coding regions, although they are more abundant in the latter. Nevertheless, the type and sequence of repeated motifs as well as the length and distribution of microsatellites vary among different taxonomic groups (Tóth et al., 2000; Ross et al., 2003; Chistiakov et al., 2006). Although microsatellites are considered selectively neutral markers, it has been shown that they can play a role in in various biological processes, including chromatin organization, regulation of DNA recombination, transcription and translation, gene expression and cell cycle dynamics (Li et al., 2002).

The key feature of microsatellites as molecular markers is their hypermutability and hence their hypervariability in species and populations. Microsatellite mutation rate is estimated at 10^{-2}-10^{-6} per locus per generation, which is several orders of magnitude higher than that of regular nonrepetitive DNA (10^{-9}) (Ellegren, 2000). This instability gives rise to the high polymorphism of these markers, which manifests mainly as variations in the number of repeating units.

Two models, polymerase slippage during DNA replication and unequal recombination, have been suggested to explain the high mutation rate of microsatellites (Schlötterer, 2000; Li et al., 2002; Chistiakov et al., 2006). The first model is considered the main mechanism and imply an incorrect match between the template strand and the nascent strand, resulting in the newly synthesized strand having a different number of repeats from the template strand once DNA replication is complete. If the error is formed in the template strand, a lower number of repetitions is generated, whereas if it is formed in the nascent strand, the number of microsatellite repeats increases. In the second model, unequal recombination could potentially change the length of the microsatellite by unequal crossing over or gene conversion. The repetitive sequence of microsatellites can lead to the formation of a hairpin, so that during recombination there will be an unequal exchange of DNA fragments, resulting in a loss of repeat units in one DNA molecule and a gain in the other.

Microsatellites are codominant markers of relatively small size, that can be easily amplified by PCR. These features coupled with their high degree of polymorphism, their mendelian inheritance pattern, and their abundance in the genome, make them very powerful genetic markers. They have proven to be an extremely valuable tool in a wide range of fundamental and applied fields of biology, including forensics, genetic mapping, population genetics and conservation and management of biological resources (Chistiakov et al., 2006). For instance, in the field of fisheries and aquaculture, microsatellites are useful for characterization of genetic stocks and broodstock selection and they are one of the most used molecular markers in genetic diversity and population structure studies. Over the past few decades these types of markers have been developed in different species of bivalve molluscs such as clams (Yasuda et al., 2007; An et al., 2009; Pereira et al., 2010; Chacón et al., 2013; Borrel et al., 2014), cockles (Martínez et al., 2009), mussels

(Presa et al., 2002; Gardeström et al., 2008; Li et al., 2011a), oysters (Launey et al., 2002; Li et al., 2003), razor clams (Arias-Pérez et al., 2012, Nie et al., 2017), and scallops (Arias et al., 2010; 2011). Recently, within the genus *Donax* 20 microsatellite loci have been identified for *Donax deltoides*, an endemic species from Australia (Miller et al., 2012), and 19 (Nantón et al., 2014) and 16 (Rico et al., 2017) microsatellite markers in the wedge clam *D. trunculus*.

The main disadvantage of microsatellites is that they must be isolated de novo from most species being examined for the first time, since it is necessary to know the sequences that flank the microsatellite motifs for primers design. This is due to fact that microsatellites are usually found in non-coding regions where nucleotide substitution rate is higher than in the coding region. Therefore, the strategy of designing universal primers matching conserved sequences, that is very effective for mtDNA, is more problematic for microsatellites (Zane et al., 2002).

The task of microsatellite markers development is a laborious process that involves a considerable investment of effort and time, especially in organisms with scarce information about their genomic sequences. In some cases, interspecific amplifications have been successfully carried out (Primmer and Ellegren, 1998, An et al., 2009, Nie et al., 2017), but the potential for transfer between species is limited (Barbará et al., 2007) due to accumulation of mutations in flanking regions that prevent primer hybridization. For de novo development of microsatellite markers in a species, different methodologies have been described. The most traditional method for microsatellite isolation consists in the construction of partial libraries, screening of recombinant clones by Southern blot hybridization and sequencing of positive clones (Rassmann et al., 1991). However, the number of positive clones (containing microsatellites), that can be obtained by this approach, usually ranges from 12% to less than 0.04% (Zane et al., 2002), so that this strategy is effective only in species with a high frequency of microsatellites and when only a relatively low number of microsatellites is needed. Although statistical power depends not only on the number of scored loci but also on other factors such as sample size and the level of polymorphism of each locus, the use of a limited number of loci might fail to provide sufficient information (Zane et al., 2002). As mentioned above, although relatively simple, especially for microsatellite-rich

genomes, this approach requires considerable experience, can be expensive and can turn out to be extremely tedious and inefficient for species with low microsatellite frequencies. Therefore, several alternative strategies have been devised to reduce the time invested in microsatellite isolation and to significantly increase yield.

One of the most widely used methods to isolate microsatellites is the construction of enriched libraries. In this case, unlike the partial genomic libraries described above, the fragments used to build the library are already rich in microsatellite motifs. Other microsatellite isolation strategies are based on the hybridization and cloning of bands obtained from other molecular markers such as AFLPs, RAPDs, or ISSRs (Inter Simple Sequence Repeats) (Lunt et al., 1999; Lian et al., 2001), so the construction of a library and screening of the fragments are avoided. However, the efficiency of these methods is also usually low.

Microsatellites can also be identified from sequences available on DNA databases, mainly in databases of Expressed Sequence Tags (ESTs) which are sequenced randomly from libraries of complementary DNA (cDNA). Loci identified from ESTs have high informative potential, especially if they are located in genes of known sequence, since they will allow investigating the correlation between the number of repetitions and the functionality of the gene. Microsatellite expansions or contractions within genes can affect the corresponding gene product and lead to phenotypic changes (Li et al., 2004). In addition, several studies have shown that loci identified from ESTs have a greater success of interspecific amplification, since sequences flaking the microsatellite are more likely to be conserved (Gupta et al., 2003; Coulibaly et al., 2005). This methodology avoids the construction of libraries and the screening of amplified fragments, considerably reducing the cost. For this reason, it has been widely used in different organisms, as is the case of some bivalve molluscs (Li et al., 2011a, b). The critical limitation of this method is the availability of EST sequences on databases, so it is a difficult strategy to apply in organisms little studied at genomic level or without genetic information available. Other limitation of this method is that primers designed for a microsatellite sequence may not amplify from genomic DNA, since ESTs are obtained from cDNA and primers can hybridize in the intron-exon

boundaries (Hu et al., 2004). In addition, this method ignores microsatellites located in introns and microsatellites identified from ESTs tend to show lower polymorphism rate than those isolated from genomic DNA (Chabane et al., 2005).

Recent developments in NGS technologies and bioinformatic analysis have provided new opportunities to discover high quality microsatellite markers, with a good representativeness of loci across the genome (Martin et al., 2010), especially in non-model organisms, for which genomic information is scarce (Shendure and Ji, 2008; Zhang et al., 2011). Moreover, this methodology has proved to be cost/time efficient and has allowed isolating microsatellite loci from different marine organisms (Plough and Marko, 2014; Yu et al., 2014), including bivalve molluscs such as *Aequipecten tehuelchus* (Domínguez-Contreras et al., 2017), *Anodonta anatina* (Lopes-Lima et al., 2016), *Bathymodiolus manusensis* (Schultz et al., 2011), *Lutraria rhynchaena* (Thai et al., 2016), *Mytilus coruscus* (An and Lee, 2012; Kang et al., 2013), *Perna viridis* (Lukehurst et al., 2017), *Potomida littoralis* (Froufe et al., 2013), and within the genus *Donax* in *D. deltoides* (Miller et al., 2012) and *D. trunculus* (Rico et al., 2017). The main limitation of this methodology is the small size of the sequenced fragments, which can cause the loss of microsatellite motifs, since they are likely to be located near one of the ends of the sequence, preventing the design of primers for their flanking regions (Abdelkrim et al., 2009).

Once obtained sequences with microsatellite motifs, specific primers that hybridise in the regions flanking the microsatellite are designed and PCR conditions to achieve amplification of each locus are optimised (Zane et al., 2002). Yield of the process of genotyping of multiple microsatellite loci, can be considerably increased by the development of multiplex PCRs that allow the amplification of several loci simultaneously in a single reaction. The coamplified loci can be differentiated by the size of their alleles or, in case allele size ranges overlap, by different fluorochromes associated with their respective primers.

Ribosomal DNA

Ribosomal genes and their associated spacers are known as ribosomal DNA (rDNA). In higher eukaryotes, rDNA comprises two different multigene families (Long and Dawid, 1980), including the major 45S rDNA family, encoding 18S, 5.8S, and 28S rRNA, and the minor 5S rDNA family, encoding 5S rRNA, each composed of hundreds to thousands of copies and organised in tandem repeats (Figures 6 and 7). The secondary structure of these genes, the different evolutionary rates among different regions, and the organization of genes in tandem, make rDNA an attractive candidate for species and population characterization, phylogenetic studies and evolutionary relationships and genomic structure analysis (Jansen et al., 2006).

The 5S rDNA consists of a highly conserved 120 bp coding sequence clustered in long direct tandem arrays and separated by variable non-transcribed flanking DNA sequences know as non-transcribed spacers or NTSs (Figure 6). Although the 5S rRNA gene is highly conserved, even among unrelated species, NTSs are variable, both in length and sequence (Rebordinos et al., 2013). These variations are due to base substitutions, deletions, insertions, and pseudogenes, and they have been used for evolutionary studies and as species-specific or population-specific markers in several bivalve species (Freire et al., 2005; Fernández-Tajes and Méndez, 2007; López-Piñón et al., 2008). Moreover, NTS regions seem to be subject to rapid evolution, which makes them important for studies concerning the organization and evolution of the 5S multigene family and also as markers for tracing recent evolutionary events (Vierna et al., 2009, 2013; Freire et al., 2010a; Perina et al., 2011; Rebordinos et al., 2013).

Although the 5S rDNA is characterized by being organised in tandem repeats, intense dynamism has been observed in this multigene family and organization of the 5S rDNA in the genome shows great plasticity. It has been found arranged i) in clusters composed of similar or distinct tandemly arranged copies (Shippen-Lenzta and Vezza, 1988); ii) in clusters linked to other gene families such as small nuclear RNAs (snRNAs) (Nilsen et al., 1989; Pelliccia et al., 2001; Cross and Rebordinos, 2005; Manchado et al., 2006a; Freire et al., 2010a; Vierna et al., 2011; Vizoso et al.,

2011), or PCGs such as histones (Eirín-López et al., 2004); iii) dispersed throughout the genome (Wood et al., 2002); and iv) both in clusters and dispersed (Little and Braaten, 1989).

The 5S rRNA gene is transcribed by RNA polymerase III whose promoter, known as the internal control region (ICR), is located within the coding region (Bogenhagen et al., 1980; Sakonju et al., 1980). One of the promoters analysed in more detail is the ICR of the 5S rDNA from somatic cells of *Xenopus laevis* (Pieler et al., 1987) that is formed by three elements: box A, between nucleotides +50 and +64, with high affinity for the TFIIIC transcription factor; the intermediate element (IE), between +67 and +72; and box C, between +80 and +97; the last ones with affinity for the TFIIIA transcription factor (for a review, see Paule and White, 2000). Nevertheless, it is becoming quite clear that the internal promoter alone is not self-sufficient to carry on the transcription (Vizoso et al., 2011). In fact, 5S rDNA is known to contain upstream transcriptional regulatory regions in several taxa. Although non-transcribed DNA sequences, such as the NTS, seem to have no value for the genome, TATA-like sequences located in the NTS have been observed in several fish species (Pendás et al., 1995; Murakami and Fujitani, 1998), suggesting that this element may influence control of the 5S rDNA genes. Furthermore, a TATA-like motif located at around nucleotides -30 to -25 is essential for efficient in vitro transcription in *Neurospora crassa* (Tyler, 1987), *Drosophila melanogaster* (Sharp and Garcia, 1988), and *Caenorhabditis elegans* and *Caenorhabditis briggsae* (Nelson et al., 1998) and it has been demonstrated that a TATA sequence located in the NTS plays an important role in the regulation of 5S rRNA gene expression in several mammals (Nielsen et al., 1993; Suzuki et al., 1996). Finally, it has recently been proposed that this region, together with RNA polymerase II-like transcriptional factors, could be involved in RNA polymerase III transcription (Raha et al., 2010). Therefore, and given the versatility of 5S rDNA transcription, it is of great interest to analyse this multigene family in different lineages and identify conserved motifs that may have a function in transcription.

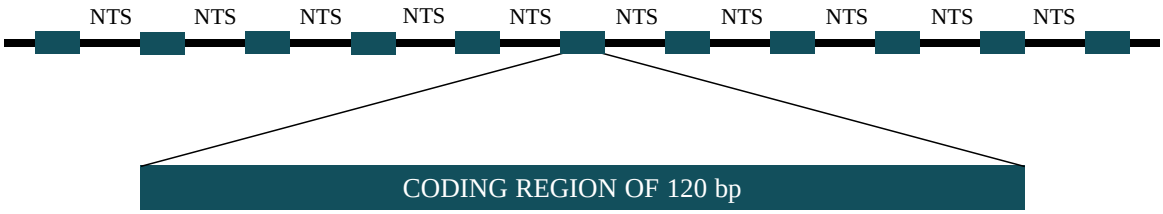


Figure 6. General structure of 5S rDNA in eukaryotes.

The major 45S rDNA multigene family contains tandemly repeated transcriptional units separated by intergenic spacers. Within each transcriptional unit, the internal transcribed spacer 1 (ITS1) separates the small subunit (18S) rDNA from the 5.8S, while the internal transcribed spacer 2 (ITS2) separates the 5.8S from the large subunit (28S) rDNA (Figure 7). Coding regions show little sequence divergence among closely related species, whereas spacer regions may exhibit higher variability as mutations occur at a relatively rapid rate in internal transcribed spacers. Therefore, the first can be useful in comparing distantly related species, while ITS can be appropriate for differentiating more closely related groups.

The ITS region (ITS1, 5.8S and ITS2) can be easily amplified by PCR, using primers annealing in the conserved flanking genes, and constitutes one of the most widely applied molecular markers in phylogenetics and species differentiation. The 5.8S gene shows a slow rate of evolutionary change, but the level of sequence variation of the spacers is higher (Hillis and Dixon, 1991) and they can be used to infer phylogenetic relationships from populations to families and even higher taxonomic levels (González et al., 1990; Coleman and Vacquier, 2002; Insua et al., 2003; Peay et al., 2008). In bivalve molluscs, a variety of methods, such as PCR amplification, PCR-RFLP analysis or sequencing of the ITS region or one of its spacers, have been used to differentiate related species, to distinguish hybrids and to explore phylogenetic relationships among bivalve species (Fernández et al., 2001; López-Piñón et al., 2002; Insua et al., 2003; He et al., 2005; Masaoka and Kobayashi, 2005; Wood et al., 2007; Freire et al., 2010b; Hurtado et al., 2011).

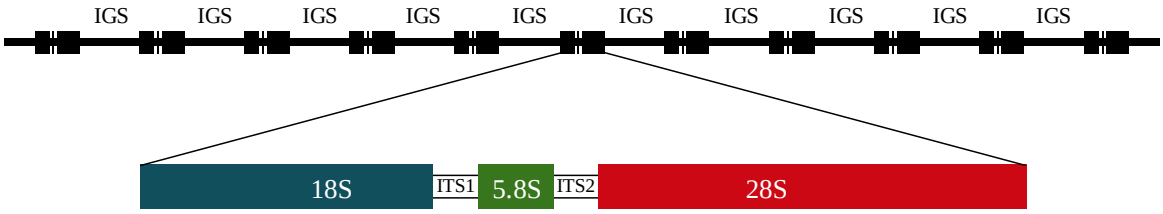


Figure 7. General structure of 45S rDNA in eukaryotes.

To date, several reports on the characterization of 5S rDNA and the ITS in various mollusc species, including bivalves, gastropods, and cephalopods have been published (Insua et al., 2001; Manchado et al., 2006a; Vizoso et al., 2011). Bivalve molluscs stand out for being one of the most widely studied group of organisms regarding 5S rDNA and the ITS region, showing high levels of gene organization as well as a vast diversity of gene arrangements. Molecular organization of the ITS region and 5S rDNA has been studied in cockles (Insua et al., 1999; Freire et al., 2005, 2010a,b), mussels (Insua et al., 2001), oysters (Cross and Rebordinos, 2005, 2006), scallops (Insua et al., 1998, 1999, 2003, 2006; Huang et al., 2007; López-Piñón et al., 2008), razor clams (Fernández-Tajes and Méndez, 2009; Vierna et al., 2009, 2010) and Veneroida clams (Cheng et al., 2006), but it has never been studied in *Donax* species, of which there is not much genetic information available.

Additionally, due to the conserved character of the coding regions and the variability in the sequence of the spacer regions, the use of the 5S rDNA and the ITS region has been reported to be useful for the discrimination of several fish and bivalve mollusc species with commercial value. On the one hand, the 5S rDNA has been used as a marker for the differentiation of several marine species, such as anchovies (Chairi and Rebordinos, 2014); cockles (Freire et al., 2005); oysters (Cross and Rebordinos, 2006); perch, grouper and wreck fish (Asensio et al., 2001); razor clams (Fernández-Tajes and Méndez, 2007); sole species (Manchado et al., 2006b); sole and Greenland halibut (Céspedes et al., 1999); salmons, rainbow trout and Atlantic pomfret (Carrera et al., 2000); and wedge clams (Pereira et al., 2012). On the other hand, ITS regions have also been commonly used to distinguish several bivalve species. ITS size and RFLPs identify some *Mytilus* mussels

(Heath et al., 1995; Toro, 1998), Pectinidae scallops (López-Piñón et al., 2002) and Veneridae clams (Fernández et al., 2001; Hurtado et al., 2011); ITS1-RFLPs discriminate *Choromytilus* and *Perna* mussels (Santaclara et al., 2006), and several freshwater mussels (White et al., 1994); ITS2 size was used to differentiate *Mytilus edulis* and two *Modiolus* species (Dixon et al., 1995); and ITS2 sequence was used to differentiate five pearl oyster species (*Pinctada chemnitzi*, *Pinctada margaritifera*; *Pinctada martensi*, *Pinctada maxima*, *Pinctada nigra*, and *Pteria penguin*) (He et al., 2005).

Goals

Goals

In Galicia, species under study are considered as a resource for the fishermen's associations due to their high market price indexes. Nevertheless, at present only the fishermen's association of Arousa and Cedeira commercialise this bivalve mollusc, which shows the exhaustion of the resource, without decisive data that can justify this fact (uncontrolled harvesting, degradation of the natural spaces by the action of the man, climate change, chemical pollution, etc.). Given the need to provide scientific knowledge to help identify the problem and allow the implementation of regulatory and control measures, in order to conserve and protect these resources and to increase its production in a sustainable way, the following specific goals have been proposed:

1. To evaluate the genetic diversity and population differentiation in the wedge clam *Donax trunculus* of the Iberian Peninsula using mitochondrial DNA markers.

2. To characterize mitochondrial and nuclear DNA markers in *Donax vittatus*, including the development of microsatellite loci through Next Generation Sequencing (NGS), and genetically evaluate the natural beds of the Iberian Peninsula.

3. To characterize the 5S rDNA region and internal transcribed spacers, ITS1 and ITS2, and develop new markers to differentiate species of the genus *Donax*.

4. To determine, for the first time, the complete sequences of the female mitochondrial genome in the four *Donax* species present in the Iberian Peninsula to compare them with those of other marine bivalves and establish the phylogenetic relationships within the Veneroida order.

Chapter 1

Mitochondrial DNA analyses of *Donax trunculus* (Mollusca: Bivalvia) population structure in the Iberian Peninsula, a bivalve with high commercial importance

Jenyfer Fernández-Pérez[1], Ana Nantón[1], Alberto Arias-Pérez[1], Ruth Freire[1], Dorotea Martínez-Patiño[2], Josefina Méndez[1]

1 Grupo Xenomar, Departamento de Bioloxía, Facultade de Ciencias and CICA (Centro de Investigacións Científicas Avanzadas), Universidade da Coruña, Campus de A Zapateira, 15071 A Coruña, Spain

2 Xunta de Galicia. Centro de Cultivos Mariños de Ribadeo (CIMA). Consellería do Medio Rural e do Mar. Muelle de Porcillán s/n, 27700 Ribadeo, Lugo, Spain

Accepted for publication in *Aquatic Conservation: Marine and Freshwater Ecosystems*
Article DOI: 10.1002/aqc.2929

Abstract

1. The wedge clam, *Donax trunculus* (Linnaeus, 1758), is an ecologically and economically important species found in the warm waters of the Atlantic and Mediterranean areas. *D. trunculus* has been suffering from a severe population decline in some wild stocks of the Iberian Peninsula and genetic monitoring of wild populations can provide valuable information to support conservation measures and management policies.

2. In this study, mitochondrial 16S rDNA (16S) and Cytochrome b (Cytb) gene variation was examined at 18 localities across the Iberian Peninsula to test whether the patterns of genetic diversity, population structure, and demographic inferences, previously deduced by using microsatellites, are supported by these molecular markers.

3. Samples (N=557) were collected from 18 localities of the Iberian Peninsula and attention was paid to the North-Western Atlantic area, where wild stocks have diminished greatly. AMOVA, F_{ST} and demographic inferences were estimated based on mitochondrial DNA. Since *D. trunculus* shows Doubly Uniparental Inheritance, both types of mtDNA (F and M) were analysed.

4. A high level of genetic diversity was detected using the haplotype diversity dataset, showing different geographic patterns.

5. F_{ST} values for the M type revealed significant differentiation between populations on both sides of the Almeria-Oran front, and between the locality from the Alboran Sea (FU) and localities from the Atlantic Ocean. However, the F type did not detect the differentiation of FU. This makes the M genome sequences more suitable for investigating genetic structure of *D. trunculus* populations.

6. Management and recovering strategies of this important resource should consider the genetic structure observed.

7. Several strategies for conservation and management of this commercial and ecologically important bivalve are proposed.

Introduction

The wedge clam, *Donax trunculus* (Linnaeus, 1758), is an Atlantic-Mediterranean warm-temperate species that has been recorded from Senegal to France in the Atlantic Ocean (Tebble, 1966), and has also been found in the Mediterranean and the Black Sea (Bayed and Guillou, 1985). This bivalve mollusc is commercially exploited in several European countries, including: France (Thébaud et al., 2005), Italy (Zeichen et al., 2002), Turkey (Deval et al., 2009), Portugal (Gaspar et al., 1999; Chícharo et al., 2002) and Spain (Ramón et al., 2005; Molares et al., 2008), and constitutes an important fishery resource due to its high economical value. However, in Galicia (North-West coast of Spain) this species has been subjected to intense harvesting and wild stocks have decreased dramatically in this region. In fact, the level of captures has decreased within the last 16 years from ~17 tonnes (2001) to 557.20 kg (2016) (Consellería do Mar, Xunta de Galicia) and the exploitation of this species is now concentrated into one single locality, Vilarrube. Furthermore, in some southern areas of the Iberian Peninsula, *D. trunculus* natural beds seem to be at high, long-term risk of extinction (Marie et al., 2016). Since wild stocks in some areas have decreased greatly, it is necessary to use different strategies to increase the production of exhausted natural banks. In bivalves, a common practice consists of releasing seed obtained in breeding facilities (hatcheries). This strategy has been used in different bivalves (e.g. Zhang and Yan, 2006; Da Costa and Martínez-Patiño, 2009; Da Costa et al., 2012; Borrell et al., 2014) and for *D. trunculus*, it has been showed that there is a real possibility of obtaining seed in hatcheries and releasing it in the natural beds to reach commercial size (Louzán et al., 2016). Furthermore, the translocation of adult individuals from some areas is other feasible option. Nevertheless, it is known that restocking programmes and translocation of individuals may have deleterious effects on the genetic composition of wild populations. Consequently, the evaluation of genetic differentiation is essential to assess the state of the resource and to contribute to its sustainable management (Ward, 2006). Moreover, obtaining basic data on population genetics is valuable for designing restocking programs and other managing strategies. In this way, these analyses could be

useful for choosing the most suitable broodstock to obtain seed in hatcheries and to release in natural beds.

To date, population genetic analyses based on microsatellite markers have provided evidence that in the Iberia Peninsula, *D. trunculus* displays a genetic differentiation along the Atlantic-Mediterranean transition with different groups: the Atlantic Ocean, the Alboran Sea and the Mediterranean Sea (Marie et al., 2016; Nantón et al., 2017; Rico et al., 2017). However, the use of different marker types has shown discrepancies in diversity patterns in other species (Johnson et al., 2003; Brito, 2007; Hoffman et al., 2009; Toews and Brelsford, 2012; Ferchaud et al., 2015; Martínez et al., 2015). For this reason, it is important to broaden the spectrum of markers to elucidate the patterns of genetic diversity and amount of population differentiation.

Mitochondrial DNA (mtDNA) is a popular genetic marker because it is relatively easy to amplify, and its content is strongly conserved across animals, with very few duplications, no introns, and short intergenic regions and maternal inheritance (Gissi et al., 2008). Though, an extreme exception to the typical maternal inheritance of mtDNA in Metazoa is found in some bivalve lineages, which possess a peculiar system termed “Doubly Uniparental Inheritance” (DUI; for reviews see Breton et al., 2007; Passamonti and Ghiselli, 2009; Zouros, 2013). To date, the vast majority of species with DUI that have been reported belong to the Mytiloida, Nuculanoida, Unionoida and Veneroida orders (Gusman et al., 2016), including species of the genus *Donax* (Theologidis et al., 2008; Fernández-Pérez et al., 2017a), such as *D. trunculus*. In species with DUI, two types of mtDNA exist, one that is transmitted through the eggs to both female and male offspring (the maternal or F type genome) and another that is transmitted through the sperm only to male offspring (the paternal or M type genome). In adult males, the F type predominates in the somatic tissues and the M type in the gonads (Stewart et al., 1995; Sutherland et al., 1998). Both types of mtDNA can provide complementary information on the evolutionary history and phylogeographic structure of species. For this reason, mtDNA in species with DUI could be a useful tool for genetic evaluation and conservation (Passamonti and Scali, 2001; Breton et al., 2009). However, the presence of two lineages of mitochondrial genome, that evolve

independently within the same species, is an important fact to consider when mtDNA is used for population genetic studies because it can interfere with the interpretation of results (Ladoukakis et al., 2002). In mussels, where this mode of inheritance is widely studied, it has been found that the M genome usually evolves faster than the F genome (Skibinski et al., 1994b; Rawson and Hilbish, 1995; Stewart et al., 1995). In theory, this would make the sequences obtained from the M genome more appropriate to investigate the genetic structure of populations (Śmietanka et al., 2013). However, when the F and M sequences were studied, no genetic structure was detected in some species, such as the mussel *Mytilus californianus* (Ort and Pogson, 2007), or paradoxically, the F genome provided a greater geographical resolution, as in the case of *Mytilus galloprovincialis* and *Mytilus edulis* (Śmietanka et al., 2009). For this reason, the analysis of both types of mtDNA can be relevant for studying the genetic diversity and population structure in species with this mode of inheritance.

In this study, we used two mitochondrial genes, 16S rDNA (16S) and Cytochrome b (Cytb), that were extensively employed in previous population genetic studies of marine organisms (Baharum and Nurdalila, 2012; Chiu et al., 2013; Maggi and González-Wangüemert, 2015; Martínez et al., 2015; Lim et al., 2016; Shen et al., 2016a; Fernández-Pérez et al., 2017a). The use of different regions of DNA, which might be differently affected by evolutionary forces, provides a more complete picture of the current and past processes that influence the patterns of genetic variation and offers valuable data for the management of species.

The objectives of the study were to characterise the genetic diversity within localities and to verify the population structure and demographic patterns previously inferred based on nuclear markers (Marie et al., 2016; Nantón et al., 2017). The results can aid in the design of fishery management and conservation strategies for *D. trunculus* inhabiting the Atlantic and Mediterranean coasts, and the recovery of the natural beds in the north-western Atlantic area.

Materials and methods

Sample collection and DNA extraction

D. trunculus samples were collected from natural beds on the Iberian Peninsula; 14 from the North-eastern coast of the Atlantic Ocean and four from the Western Mediterranean coast (Figure 1, Table 1). Taxonomic identification of the specimens was based on Pereira et al. (2012) protocol. Gender determination was performed on each individual by microscopic examination of gametogenic tissue from the visceral mass, based on the presence of eggs or sperm. Total genomic DNA was extracted from ethanol-preserved foot material using a Chelex-100 (Sigma-Aldrich, USA) protocol based on Walsh et al. (1991). In males, gonadal tissue was used to analyse gene sequences from the male-specific (M type) mitochondrial lineage.

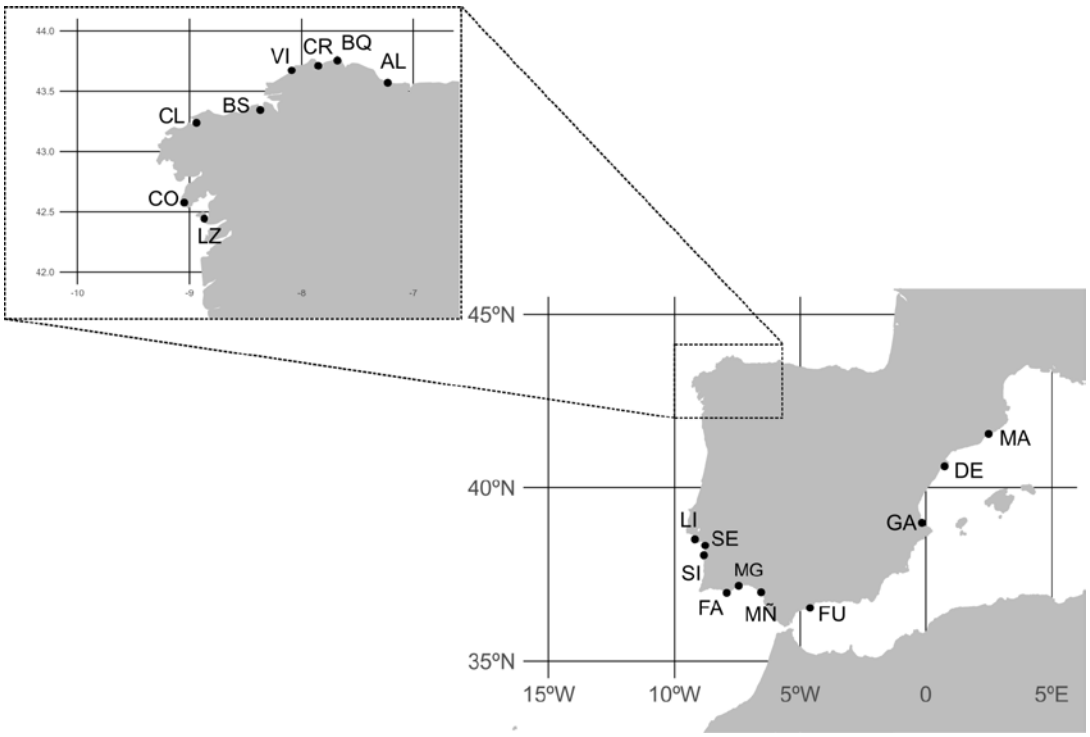


Figure 1. Map showing the location of the sampled *Donax trunculus* localities along the Iberian Peninsula (AL: Altar, BQ: O Barqueiro, CR: Cariño, VI: Vilarrube, BS: Bastiagueiro, CL: Corme-Laxe, CO: Corrubedo, LZ: A Lanzada, LI: Lisboa, SE: Setúbal, SI: Sines, FA: Faro, MG: Monte Gordo, MÑ: Matalascañas, FU: Fuengirola, GA: Gandía, DE: Delta del Ebro, MA: Maresme).

Table 1. Sampling details of *D. trunculus*

Code	Locality	Country	Sample size	Latitude	Longitude
AL	Altar	Spain	30	43°34´N	07°14´W
BQ	O Barqueiro	Spain	30	43°45´N	07°41´W
CR	Cariño	Spain	31	43°42´N	07°51´W
VI	Vilarrube	Spain	39	43°38´N	08°05´W
BS	Bastiagueiro	Spain	32	43°20´N	08°22´W
CL	Corme-Laxe	Spain	29	43°14´N	08°56´W
CO	Corrubedo	Spain	30	42°34´N	09°03´W
LZ	A Lanzada	Spain	30	42°27´N	08°52´W
LI	Lisboa	Portugal	30	38°31´N	09°11´W
SE	Setúbal	Portugal	33	38°20´N	08°47´W
SI	Sines	Portugal	32	38°03´N	08°50´W
FA	Faro	Portugal	27	36°58´N	07°55´W
MG	Monte Gordo	Portugal	29	37°10´N	07°27´W
MÑ	Matalascañas	Spain	31	36°59´N	06°33´W
FU	Fuengirola	Spain	32	36°32´N	04°37´W
GA	Gandía	Spain	31	38°59´N	00°09´W
DE	Delta del Ebro	Spain	34	40°37´N	00°44´E
MA	Maresme	Spain	30	41°33´N	02°30´E

16S rDNA and Cytochrome b genes sequencing

A 447 bp fragment of the 16S F type was amplified with the universal primer 16AR (Kocher et al., 1989) and reverse primer DT16S-R395 (Theologidis et al., 2008), while a 413 bp fragment of the 16S M type was amplified with the specific primers DT16SM-F and DT16SM-R (Fernández-Pérez et al., 2017a). A 432 bp fragment of the Cytb gene (F and M sequences) was amplified with the specific primers DTCYTB-F and DTCYTB-R (Theologidis et al., 2008) (see Table 2 for detailed information on primers).

PCR reactions were carried out in a final volume of 25 µl, containing ~100 ng DNA, 0.75 U of *Taq* DNA polymerase (Roche Diagnostics, Basel, Switzerland), 0.25 mM of each dNTP, 0.6 µM of each primer, 1.5 mM of $MgCl_2$ for the 16S F type, 2 mM for the 16S M type and 2.5 mM for Cytb gene, and the buffer recommended by the polymerase suppliers. PCR cycling profile consisted of 2 min at 94 °C, 35 cycles of 1 min at 94 °C, 1 min and 20 s at 49 ºC (16S) or 50 °C

(Cytb), 1 min at 72 °C, followed by 5 min at 72 °C. PCR products were migrated on 2% agarose gel electrophoresis. Gels were stained by immersion in a 0.5 µg/ml ethidium bromide solution for 30 min, visualized and recorded on a transilluminator Gel Doc XR Systems (Bio-Rad, Barcelona, Spain). PCR products were sequenced on an ABI PRISM 3130xl (Applied Biosystems, Foster City, CA, USA) at the Molecular Biology Unit of the University of A Coruña (Spain). All haplotypes were deposited in GenBank (accession numbers: MF983875 - MF984107).

Table 2. PCR primers used to amplify fragments of 16S rDNA and Cytb from *D. trunculus* mtDNA.

Locus	Primer pair	Forward primer (5´to 3´)	Reverse primer (5´to 3´)	Reference
16S F type	16SAR/ DT16S-R395	CGCCTGTTTATCAAAAACAT	CYYTAATYCAACATCGAGGTC	Kocher et al. 1989/Theologidis et al. 2008
16S M type	DT16SM-F/DT16SM-R	TGATTTAAACGGCTGCAGCTAA	GCGAACAGTCCCACTTCTCT	Fernández-Pérez et al. 2017a
Cytb	DTCYTB-F/ DTCYTB-R	TGTCGTATTGGGGGGCTACTG	GTACCACTCTGGCTGAATGTG	Theologidis et al. 2008

Data analysis

Sequence data were aligned via MAFFT (Katoh and Toh, 2008) using the standard FFT-NS-i algorithm and manually checked using the BioEdit v.7.2.5 sequence editor (Hall, 1999). For single genes (16S and Cytb) and combined gene data (16S F type + Cytb F type, and 16S M type + Cytb M type), a number of genetic diversity indices were calculated for each locality with DnaSP v.5.10.01 (Librado and Rozas, 2009): number of haplotypes (h), haplotype diversity (Hd), and nucleotide diversity (π). Observed values of genetic diversity were compared using a resampling procedure implemented in R (R Core Team, 2014). The minimum number of individuals analysed per gene was used as sample size with 10,000 resamples. The comparison was performed using mean values of haplotype and nucleotide diversity and 95% confidence intervals (ICs 95%). Given that males contain both mitochondrial lineages (F and M types), the frequency distribution of haplotypes from the F lineage among females and males was tested by simple χ^2 tests of homogeneity with the statistical package SPSS 16.0 (SPSS Inc.). In these analyses, only the most frequent haplotypes were used.

For combined gene data, pairwise F_{ST} values were estimated with Arlequin v.3.5.1.2 (Excoffier and Lischer, 2010), based on nucleotide distance matrix. These distances were calculated according to the Jukes and Cantor method (1969), which assumes that all nucleotide substitutions have the same probability. F_{ST}'s null distribution was obtained by bootstrapping (1,000 replicates), and the corresponding *p*-values were adjusted using sequential Bonferroni correction (Rice, 1989). Population structure was examined by Analyses of Molecular Variance (AMOVA) using the program Arlequin v.3.5.1.2 (Excoffier and Lischer, 2010). Haplotype networks were constructed with TCS v.1.21 (Clement et al., 2000) and edited using tcsBU (TCS Beautifier) (Santos et al., 2016). Tajima´s D (Tajima, 1989), Fu´s F_S (Fu, 1997) and Ramos-Onsis and Rozas´s R_2 tests (Ramos-Onsis and Rozas, 2002) were performed in DnaSP v.5.10.01 (Librado and Rozas, 2009) to analyse the possibility of population expansion (1,000 replicates). Graphical outputs were used to illustrate demographic model fit.

Results

Genetic diversity

For the 16S gene, an alignment of 386 bp of the F type from 557 individuals showed 59 haplotypes (38 of which were private) defined by 42 variable sites (GenBank accession numbers: MF983875 - MF983933). Haplotype 1 (H1) was the most frequent (global frequency = 52.96%), found in all studied localities, followed by Haplotype 3 (H3) (12.93%) and Haplotype 8 (H8) (15.44%) in 16 and 17 localities, respectively. Haplotype 47 (H47) was only found in Mediterranean localities (FU, GA, DE and MA) but with an inferior global frequency, 1.08%. Globally, high Hd (0.679) and low π (0.0046) were detected (Table 3). Hd ranged from 0.393 (DE) to 0.722 (CL) and π from 0.0017 (DE) to 0.0058 (BS). For the 16S gene M type, an alignment of 358 bp from 286 individuals showed 15 variable sites and 16 haplotypes (12 private) were defined (GenBank accession numbers: MF983934 - MF983949). Haplotype 1 (H1) (82.46%) was the most ubiquitous and was shared among 16 localities, excluding DE and MA, followed by Haplotype 12 (H12) (11.19%) and Haplotype 13 (H13) (1.12%) only found in three localities (GA, DE and MA). Altogether, low Hd (0.308) and π (0.0017) were detected (Table 3). Excluding AL, BQ, CR, BS, CL, LI, SI and FA with no variation, Hd ranged from 0.118 (VI) to 0.528 (GA) and π from 0.0003 (VI) to 0.0034 (GA).

For the Cytb gene, alignment of 329 bp of the F type from 464 individuals showed 77 variable positions, and 94 haplotypes (63 of which were private) were defined (GenBank accession numbers: MF983950 - MF984043). Haplotype 1 (H1) was the most frequent (52.16%), found in all localities, followed by Haplotype 11 (H11) (7.11%) found in 13 localities, all of them situated in the Atlantic coast. Haplotypes 31 and 33 (H31 and H33) were only found in localities from the Mediterranean coast (GA, DE and MA) and their global frequencies were lower than 3%. Globally, high Hd (0.720) and low π (0.0044) were detected (Table 3). Hd ranged from 0.468 (FA) to 0.908 (MA) and π from 0.0019 (CO and FA) to 0.0081 (MA). For the Cytb gene M type, an alignment of 326 bp from 255 individuals showed 65 variable positions and 64 haplotypes (48

private) were defined (GenBank accession numbers: MF984044 – MF984107). Haplotype 1 (H1) was the most ubiquitous (57.25%), found at all localities excluding MA, followed by Haplotype 21 (H21) (4.71%), only found at Mediterranean localities. Altogether, Hd was high (0.669) and π (0.015) was very much higher than that observed for the F type (Table 3). Hd ranged from 0.182 (MÑ) to 0.933 (GA) and π from 0.0006 (MÑ) to 0.0332 (GA).

Regarding concatenated sequences, for F and M types, an alignment of 715 bp from 442 individuals displayed 142 haplotypes (102 of which were private), and an alignment of 684 bp from 221 individuals exhibited 63 haplotypes (46 private), respectively. For both types of mtDNA (F and M), Haplotype 1 (H1) was the most frequent (global frequency = 28.96% and 55.20%, respectively), found in 16 studied localities, excluding DE and MA. For the F type, Hd ranged from 0.770 (SE) to 0.957 (MA) with an average of 0.902, while π ranged from 0.0032 (FA) to 0.0051 (LI) with an average of 0.0046 (Table 3). For the M type, Hd ranged from 0.295 (CL) to 0.972 (GA) with an average of 0.693, while π ranged from 0.0006 (CR) to 0.0184 (GA) with an average of 0.0074 (Table 3). Using a resampling procedure, the mean value of haplotype diversity and its confidence interval were calculated (Supporting Information: S1 Figure). For haplotype diversity, CIs overlap in all cases, suggesting similar diversity levels. However, for nucleotide diversity with M sequences, CI95% of GA locality only overlaps with CIs of MG and MA localities (Supporting Information: S2 Figure).

When χ^2 tests of homogeneity were carried out to investigate whether the F lineage haplotypes found among males have the same frequency distribution as those found among females, none were significant (16S: d.f. = 2, $\chi^2 = 3.838$, $p = 0.147$; Cytb: d.f. = 7, $\chi^2 = 4.803$, $p = 0.684$; Combined gene; d.f. = 10, $\chi^2 = 14.017$, $p = 0.172$). Thus, there is not enough evidence to reject the null hypothesis that maternally transmitted genomes (F type) have an equal probability of being found among female and male individuals.

Table 3. Genetic variability across populations for 16S and Cytb. N = Number of individuals; h = number of haplotypes; Hd = haplotype diversity; and π = nucleotide diversity. Locality names follow Figure 1 and Table 1.

		16S				Cytb				Combined gene			
Locality	**Type**	**N**	**h**	**Hd**	**π**	**N**	**h**	**Hd**	**π**	**N**	**h**	**Hd**	**π**
AL	F	30	8	0.685	0.0051	24	10	0.710	0.0052	24	16	0.917	0.0050
	M	16	1	0.000	0.0000	15	5	0.629	0.0023	15	5	0.629	0.0011
BQ	F	30	9	0.593	0.0041	24	11	0.714	0.0035	24	15	0.909	0.0040
	M	16	1	0.000	0.0000	16	5	0.667	0.0032	16	5	0.667	0.0015
CR	F	31	8	0.660	0.0053	29	7	0.567	0.0028	28	12	0.836	0.0044
	M	16	1	0.000	0.0000	16	4	0.350	0.0012	16	4	0.350	0.0006
VI	F	39	12	0.591	0.0040	32	12	0.716	0.0044	30	15	0.860	0.0045
	M	17	2	0.118	0.0003	24	10	0.667	0.0033	15	8	0.733	0.0020
BS	F	32	9	0.720	0.0058	26	9	0.622	0.0025	25	14	0.890	0.0045
	M	13	1	0.000	0.0000	12	6	0.682	0.0031	12	6	0.682	0.0015
CL	F	29	10	0.722	0.0045	28	10	0.593	0.0037	27	15	0.883	0.0042
	M	14	1	0.000	0.0000	13	3	0.295	0.0014	13	3	0.295	0.0007
CO	F	30	6	0.653	0.0051	13	5	0.539	0.0019	13	8	0.897	0.0043
	M	14	3	0.275	0.0008	17	6	0.515	0.0025	13	5	0.538	0.0011
LZ	F	30	7	0.462	0.0032	29	12	0.759	0.0040	27	16	0.886	0.0036
	M	13	2	0.154	0.0004	15	6	0.571	0.0028	11	5	0.618	0.0018
LI	F	30	5	0.582	0.0050	20	10	0.837	0.0051	20	12	0.937	0.0051
	M	15	1	0.000	0.0000	13	4	0.423	0.0023	8	4	0.643	0.0017

Table 3. Continuation.

SE	F	33	5	0.521	0.0042	31	10	0.650	0.0035	30	12	0.770	0.0039
	M	15	2	0.133	0.0004	12	5	0.576	0.0026	12	6	0.682	0.0015
SI	F	32	9	0.607	0.0035	20	9	0.653	0.0036	19	14	0.912	0.0037
	M	15	1	0.000	0.0000	14	5	0.506	0.0018	13	4	0.423	0.0007
FA	F	27	8	0.630	0.0044	22	5	0.468	0.0019	21	11	0.814	0.0032
	M	16	1	0.000	0.0000	16	7	0.625	0.0026	14	6	0.604	0.0012
MG	F	29	6	0.468	0.0039	26	15	0.828	0.0050	25	17	0.903	0.0048
	M	14	3	0.275	0.0008	15	9	0.848	0.0048	12	8	0.848	0.0026
MÑ	F	30	9	0.515	0.0024	31	11	0.710	0.0039	30	18	0.894	0.0031
	M	13	3	0.295	0.0009	11	2	0.182	0.0006	11	4	0.491	0.0008
FU	F	32	7	0.623	0.0046	20	8	0.590	0.0043	20	11	0.837	0.0048
	M	17	3	0.228	0.0007	15	5	0.810	0.0044	15	6	0.848	0.0023
GA	F	31	8	0.680	0.0036	29	13	0.810	0.0049	27	15	0.923	0.0041
	M	14	3	0.582	0.0034	10	7	0.933	0.0332	9	8	0.972	0.0184
DE	F	32	7	0.393	0.0017	34	17	0.877	0.0063	30	18	0.917	0.0039
	M	15	2	0.133	0.0004	10	7	0.867	0.0048	9	6	0.833	0.0026
MA	F	30	10	0.641	0.0025	26	13	0.908	0.0081	22	16	0.957	0.0047
	M	15	5	0.476	0.0015	11	6	0.727	0.0190	7	6	0.952	0.0045
TOTAL	F	557	59	0.679	0.0046	464	94	0.720	0.0044	442	142	0.902	0.0046
	M	268	16	0.308	0.0017	255	64	0.669	0.0150	221	63	0.693	0.0074

Population structure

The combined genes showed significant differentiation ($P < 0.001$) between localities (global F_{ST} for F and M types was 0.0214 and 0.0857, respectively). For the F type, 37 out of the 153 pairwise combinations were significant after sequential Bonferroni correction (Table 4), a lower proportion than revealed for the M type, with 54 significant combinations. The highest F_{ST} values were found in pairwise combinations involving FU, GA, DE or MA, four localities from the Mediterranean coast that were implicated in the total number of significant pairwise combinations, but only GA, DE and MA, both in F and M types, showed genetic differentiation with all other localities, except from each other.

According to F_{ST} pairwise results and the superposition of haplotype frequencies on a geographic map (Figures 2 and 3), samples were pooled into three groups: one involving all localities from the Atlantic coast, called Atlantic group; another involving three localities from the Mediterranean coast (GA, DE and MA), called Mediterranean group; and a third group comprising FU. The AMOVA results based on these groups showed significant differences among the groups (5.38% and 22.42% of the total variation for the F and the M type, respectively) and most of the variation corresponded to differences between individuals within groups (94.63% for the F type and 77.72% for the M type) (Table 5).

Table 4. Pairwise F_{ST} values for F type (above diagonal) and M type (below diagonal) based on 16S and Cytb concatenated sequences of 18 localities of *D. trunculus*. Significant values are followed by an asterisk: (*) $p = 0.05$. (**) significant after sequential Bonferroni correction.

Localities	AL	BQ	CR	VI	BS	CL	CO	LZ	LI	SE	SI	FA	MG	MN	FU	GA	DE	MA
AL		-0.019	-0.008	-0.019	-0.018	-0.021	0.020	-0.006	-0.001	-0.022	-0.018	-0.018	-0.024	0.040*	-0.005	0.114*	0.232**	0.220**
BQ	-0.006		0.015	-0.020	-0.009	-0.020	0.058	-0.016	0.024	-0.016	-0.022	-0.024	-0.010	0.023	0.021	0.154**	0.283**	0.271**
CR	0.037	0.055		0.012	-0.022	0.006	-0.017	0.046	-0.003	-0.004	0.036	0.016	-0.004	0.101*	-0.027	0.082*	0.191**	0.185**
VI	0.019	0.015	0.002		-0.004	-0.011	0.053	-0.013	0.018	-0.019	-0.015	-0.022	-0.015	0.020	0.023	0.156**	0.279**	0.267**
BS	-0.030	-0.021	0.010	-0.012		-0.003	-0.011	0.015	-0.002	-0.015	0.011	-0.003	-0.010	0.070*	-0.021	0.089*	0.202**	0.195**
CL	0.042	0.022	0.002	-0.016	0.002		0.062	-0.005	0.031	-0.013	-0.017	-0.017	-0.014	0.028	0.015	0.130**	0.260**	0.246**
CO	-0.003	0.022	0.005	-0.003	-0.014	0.000		0.109*	-0.013	0.038	0.099*	0.081	0.028	0.196*	-0.026	0.076*	0.161**	0.162**
LZ	0.017	0.044	0.026	0.013	-0.010	0.029	0.008		0.048	-0.008	-0.026	-0.023	0.000	-0.005	0.058*	0.190**	0.325**	0.312**
LI	0.081	0.018	0.090	-0.002	0.036	-0.005	0.048	0.042		0.008	0.046	0.030	0.002	0.119*	0.002	0.129*	0.234**	0.227**
SE	0.023	0.039	-0.006	-0.012	0.000	0.002	0.001	0.004	0.036		-0.010	-0.021	-0.017	0.031	0.003	0.129**	0.254**	0.243**
SI	0.042	0.022	0.002	-0.005	0.002	-0.026	0.000	0.012	0.021	0.002		-0.026	-0.008	0.002	0.040	0.174**	0.310**	0.293**
FA	0.048	0.017	0.022	0.000	0.013	-0.019	0.013	0.020	-0.016	0.013	-0.036		-0.019	0.009	0.030	0.171*	0.307**	0.296**
MG	0.007	-0.006	0.030*	0.005	-0.030	0.006	-0.001	0.015	-0.003	0.006	0.006	0.006		0.040*	0.002	0.121*	0.239**	0.229**
MÑ	0.038	0.039	0.006	-0.010	-0.002	0.001	-0.002	0.020	0.055	-0.002	0.001	0.011	0.008		0.120*	0.236**	0.369**	0.351**
FU	0.336**	0.313**	0.369**	0.268**	0.294**	0.342**	0.315**	0.280*	0.272*	0.294**	0.342**	0.317**	0.255**	0.322*		0.042	0.139**	0.139**
GA	0.285**	0.273**	0.302**	0.262**	0.238**	0.262**	0.262**	0.226**	0.171*	0.245**	0.265**	0.266**	0.225**	0.236**	0.175*		0.018	0.023
DE	0.946**	0.939**	0.957**	0.932**	0.936**	0.950**	0.943**	0.929**	0.924**	0.936**	0.950**	0.943**	0.918**	0.944**	0.927**	0.623*		-0.004
MA	0.958**	0.950**	0.967**	0.944**	0.949**	0.962**	0.956**	0.945**	0.943**	0.950**	0.962**	0.955**	0.933**	0.958**	0.939**	0.663*	-0.007	

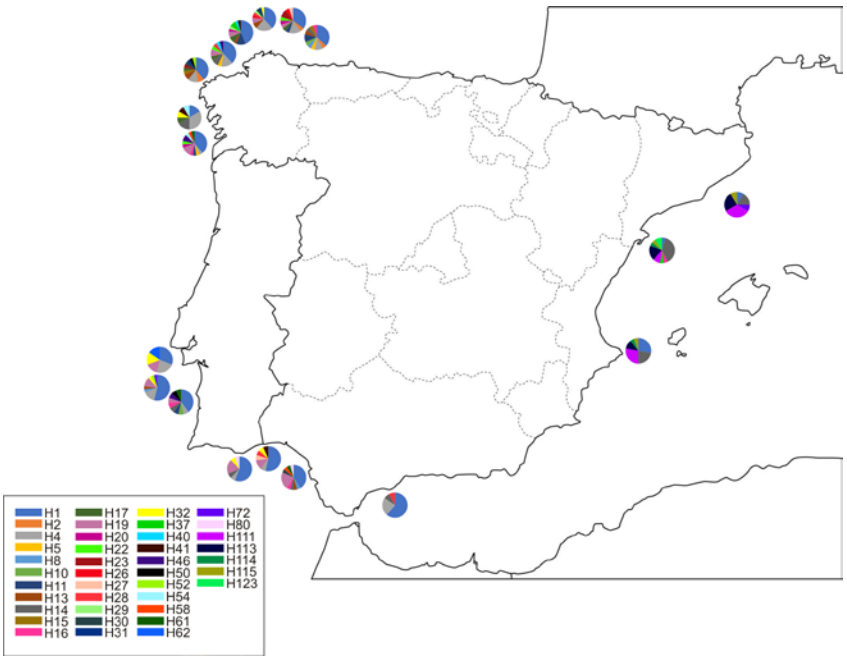


Figure 2. Haplotype distribution across localities for the F type. Each *colour* represents a haplotype. Sectors in pie charts are proportional to haplotype frequency. Singleton haplotypes were not considered for graphic representation.

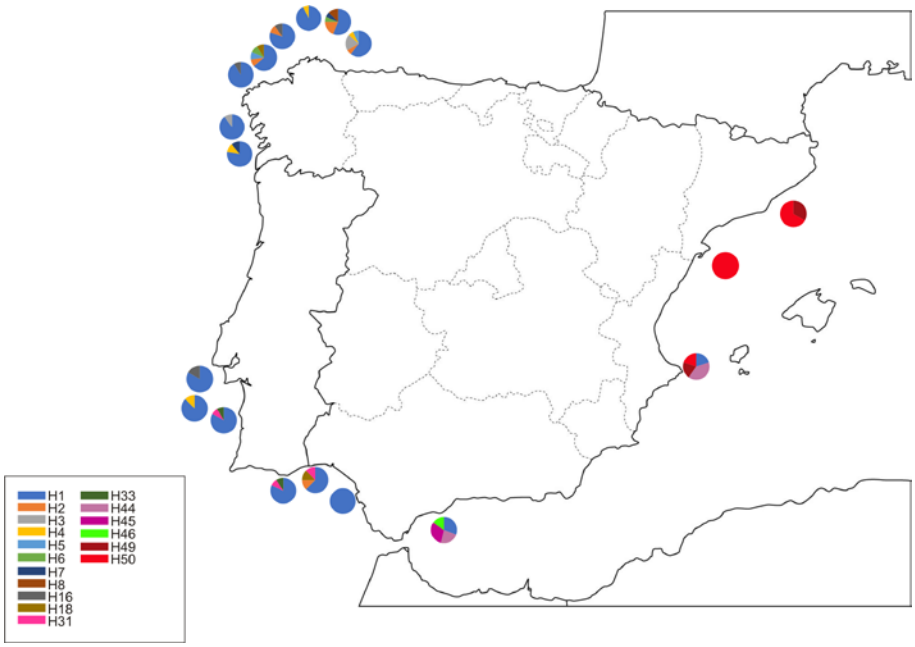


Figure 3. Haplotype distribution across localities for the M type. Each *colour* represents a haplotype. Sectors in pie charts are proportional to haplotype frequency. Singleton haplotypes were not considered for graphic representation.

Table 5. AMOVA results using the three geographical groups as subpopulations.

		Source of variation		
		Among groups	Among populations within groups	Within populations
Combined F sequences	Percentage variation	5.38	-0.01	94.63
	Fixation index	*F*ct=0.05382*	*F*sc=-0.00009	*F*st=0.05374*
Combined M sequences	Percentage variation	22.42	-0.13	77.72
	Fixation index	*F*ct=0.22416*	*F*sc=-0.00172	*F*st=0.22282*

* Significant at 5% level

Demographic inference

The statistical parsimony networks for the F and M types (excluding unique haplotypes) showed a similar star-like shape, typical of an expansion scenario, centred on the most frequent haplotype (H1 in both cases, found in 16 localities, excluding DE and MA). In the haplotype network for the F type one additional and less-defined star-like pattern with the main Mediterranean haplotype located in central position can be distinguished (Figure 4). Likewise, in the haplotype network for the M type, it highlights the existence of two haplotypes (H49 and H50), separated from the core and only present in localities belonging to the Mediterranean group (GA, DE and MA); and three haplotypes (H44, H45 and H46) present in FU and GA localities (Figure 4).

Only the Atlantic group for both types of mtDNA (F and M) and the Mediterranean group for the F type yielded negative and significant values showed both by Tajima´s D and Fu´s F_S statistics and a statistically significant Ramos-Onsis´s R_2 value (Table 6), all indicative of population expansion.

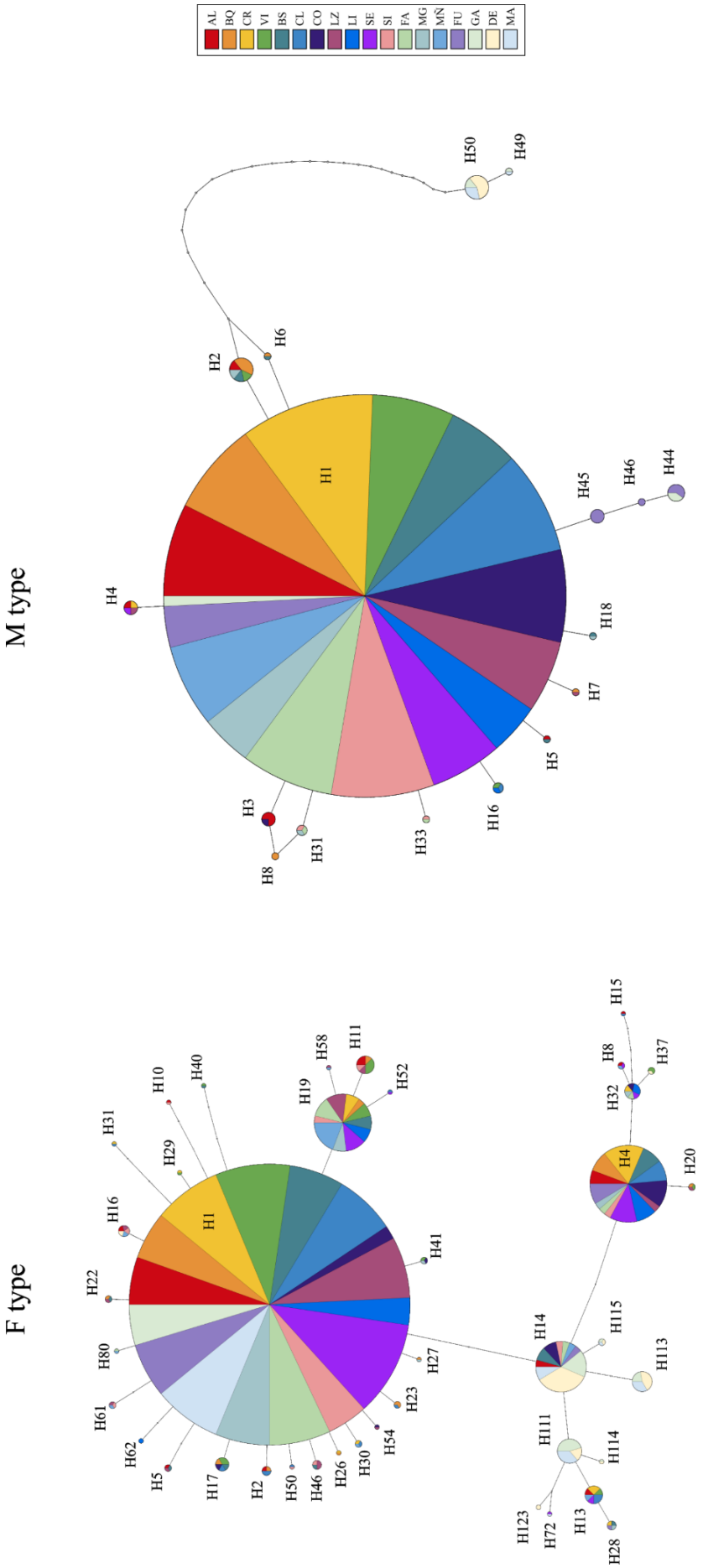


Figure 4. Haplotype (TCS) networks (excluding singleton haplotypes) for each mtDNA type showing the phylogenetic relationship between haplotypes.

Each *circle* represents a haplotype, and its size is proportional to its frequency. *White dots* indicate hypothetical intermediate haplotypes. *Colours* denote the geographic origin of each haplotype. Sample name abbreviations are as in Table 1.

Table 6. Neutrality tests for regional groups. Significant values are followed by an asterisk: (*) $p < 0.05$ (**) $p < 0.01$.

Locus	Region	F_S	D	R_2
Combined F sequences	Atlantic	-157.163**	-2.343**	0.016**
	Mediterranean	-37.383**	-2.191**	0.029**
	FU	-3.276*	-1.366	0.075
Combined M sequences	Atlantic	-74.97**	-2.617**	0.013**
	Mediterranean	-1.757	0.034	0.128
	FU	-1.495	0.053	0.146

Discussion

Genetic diversity

This study uses for the first time in *D. trunculus*, mitochondrial markers to analyse genetic variation and population differentiation in samples collected along the coast of the Iberian Peninsula.

Globally, both markers analysed revealed high genetic diversity in *D. trunculus* in line with previous studies based on microsatellites (Marie et al., 2016; Nantón et al., 2017). Based on the 16S gene, the F type revealed abundant haplotype diversity (0.679) and low levels of nucleotide diversity (0.0046). Meanwhile, the M type showed lower diversity values (haplotype diversity 0.308; nucleotide diversity 0.0017) so that the maternal type is more diverse than the paternal type. This difference between the two types of mtDNA takes its extreme form in AL, BQ, CR, BS, CL, LI, SI and FA samples, where the M type did not show polymorphism. This result, greater diversity detected in the F type, has been reported in other works with DUI species such as *M. galloprovincialis* (Ladoukakis et al., 2002), *Musculista senhousia* (Passamonti, 2007), *Perumytilus purpuratus* (Vargas et al., 2015) and *Donax vittatus* (Fernández-Pérez et al., 2017a). Differences in variability between the two mitochondrial lineages could respond to the older evolutionary history of the F genome than the M genome (Vargas et al., 2015). Nevertheless, this difference may also be due to the lower frequency with which the M genome is observed,

meaning that its diversity may be underestimated (Vargas et al., 2015). An alternative explanation could be that DUI has evolved to protect mtDNA in females, while selection could be relaxed in males (Passamonti, 2007; Zouros, 2013) or that the M genome is under relaxed selection because its function is restricted to the male germ line whereas the role of F genome is to support the female germ line and the soma of both males and females. Another explanation could be that higher variability of the 16S F type could result from populations having experienced a heavy bottleneck in the past. If males were much fewer than females in the founding population this would likely cause lower variability in the M genome (Passamonti, 2007; Guerra et al., 2014).

In contrast, for the Cytb gene, haplotype diversity values observed for the F (0.720) and M (0.669) types were similar. This similarity between the two types of mtDNA has been previously observed for other mitochondrial genes in different species of mussel (Rawson and Hilbish, 1995; Ort and Pogson, 2007; Śmietanka et al., 2013; Zouros, 2013). These values obtained for *D. trunculus* are in line with those observed for the same gene in some species of clams such as *Artica islandica* (0.678; Dahlgren et al., 2000) and *Venerupis pullastra* (0.620; Pereira, 2013) and in the cockle *Cerastoderma edule* (0.799; Martínez et al., 2015), species in which DUI has not been described. To date, given the lack of studies of genetic diversity for Cytb gene in species with DUI, it is not possible to compare the values obtained for the M type with the values observed for this mitochondrial lineage in other species. Regarding nucleotide diversity, the total value obtained for the F type (0.0044) is similar to reports of *V. pullastra* (0.0040; Pereira, 2013), but slightly lower than that found in *C. edule* (0.0071; Martínez et al., 2015). With respect to the observed value for the M type in this work (0.0150), it was three times higher than that obtained for the F type. This result agrees with previous studies in other bivalves, where the M type showed to have two to four times higher nucleotide diversity values than the F type when analysing different mitochondrial regions (Hoeh et al., 1996; Passamonti and Scali, 2001; Passamonti et al., 2003; Ort and Pogson, 2007; Śmietanka et al., 2013). The existence of a higher level of polymorphism and a higher rate of evolution in the M genome opposite to the F has been reported in previous studies with mussels (Skibinski et al., 1994b; Rawson and Hilbish, 1995;

Stewart et al., 1995; Hoeh et al., 1996; Stewart et al., 1996; Ladoukakis et al., 2002). Different mechanisms have been proposed that could contribute to this pattern, such as different effective sizes between genomes, differences in selection pressures, or faster M mitochondrial replication rates during early development of male embryos (Skibinski et al., 1994b; Stewart et al., 1996; Hasegawa et al., 1998; Ballard, 2000a,b). The first hypothesis is based on the larger effective population size of the F genome. Nevertheless, this explanation ignores the fact that a mutation that occurs in a F molecule during its residence in a male individual will not pass on to the next generation, thus males are evolutionary dead ends for the F genome. The second hypothesis implies that the M genome is under strong directional selection imposed by a presumed specialized high-performance sperm role (Skibinksi et al., 1994b). However, there is also the contrary assumption that the M genome is under relaxed selection as discussed in the case of the 16S gene. For instance, there were many more sites at which mutation would be nearly neutral or slightly deleterious when occurring in the M genome, but highly deleterious when occurring in the F genome. This would be related to the third and last hypothesis, higher evolutionary rate of the M molecule.

However, the haplotype diversity based on a single gene (16S and Cytb) was not in line with the result from the combined genes. Globally, a high level of genetic diversity (0.902 and 0.693 for the F and M types, respectively) was detected using the combined dataset. These results demonstrate the importance of analysing various regions to have a broader view of the genetic variability in the localities. Discrepancies in diversity levels provided by different mitochondrial markers were also observed in other studies of marine invertebrates (Yuan et al., 2009; Zitari-Chatti et al., 2009) where the 16S rRNA gene is more conserved than the Cytb gene (Hixson and Brown, 1986; Baharum and Nurdalila, 2012).

Population genetic structure

Data analyses with 16S and Cytb sequences combined, including global and pairwise F_{ST}, indicated significant genetic structure in *D. trunculus* along the Iberian Peninsula. The results

obtained with F combined sequences only showed two different groups: the Mediterranean group, including localities situated in the Mediterranean Sea (GA, DE, MA) and the Atlantic group, including all Atlantic localities and FU located on the coast of the Alboran Sea, the westernmost part of the Mediterranean Sea. However, pairwise F_{ST} using the M combined sequences displayed three groups: the Mediterranean group (GA, DE, MA), the Atlantic group (all localities on the coast of the Atlantic Ocean), and the Alboran group (FU). The population structure proposed in this study (three groups: Atlantic, Mediterranean and Alboran) was observed previously using microsatellite markers (Marie et al., 2016; Nantón et al., 2017). The existence of two independently evolving mitochondrial lineages should be considered when population structure studies are carried out in species with DUI. In these studies, differences between the results obtained with each mitochondrial genome are common, maybe because the M genome typically evolves faster, potentially allowing for insight into more recent events in a population's history (Śmietanka et al., 2013). Moreover, several authors have proposed that the M lineage shows higher differentiation than the F lineage (Liu et al., 1996; Ladoukakis et al., 2002) as is clearly the case in our study, where the M genome provided greater resolution and differentiated the Atlantic samples from both the Mediterranean and the Alboran Sea. In fact, AMOVA analysis revealed the existence of a higher percentage of variation attributed to differences among groups for the M type than for the F type (22.42% and 5.38%, respectively) (Table 5). This is in line with the results obtained in previous studies in *M. galloprovincialis*, where the F type exhibits higher degrees of diversity within localities, while the M type produces higher degrees of differentiation among populations (Ladoukakis et al., 2002).

The genetic structure detected for *D. trunculus* in the Iberian Peninsula could be related to the existence of geographical barriers in this region and the Almeria-Oran front and the Strait of Gibraltar could be the reason for this geographical confinement. The Almeria-Oran front constitutes an oceanographic discontinuity which not only has shown a significant effect on the genetic structure of this wedge clam (Marie et al., 2016; Nantón et al., 2017), but also on different species of bivalves such as the mussels *M. edulis* and *M. galloprovincialis* (Sanjuan et al., 1994;

Quesada et al., 1995; Sanjuan et al., 1996, 1997; Diz and Presa, 2008; Luis et al., 2011) and in the scallops *Pecten jacobaeus* and *Pecten maximus* (Ríos et al., 2002), or in other marine organisms such as *Dicentrarchus labrax* (Naciri et al., 1999), *Meganyctiphanes norvegica* (Zane et al., 2000), *Serranus cabrilla* (Schunter et al., 2011) or *Liocarcinus depurator* (García-Merchán et al., 2012). The Alboran Sea, located west of the Mediterranean Sea, borders the Atlantic Ocean along the Strait of Gibraltar. This region is a transition zone between two basins with different oceanographic characteristics where Atlantic and Mediterranean species coexist. The Almeria-Oran front is an area of strong currents where abrupt temperature and salinity changes occur (Tintore et al., 1988). This area constitutes an oceanographic barrier to dispersal of planktonic larvae and therefore can prevent gene flow between the Atlantic and Mediterranean populations, thereby contributing to their differentiation. The location of FU, within the Alboran Sea and limited in the east by the Almeria-Oran front and in the west by the Strait of Gibraltar, could contribute to the observed genetic differentiation. In addition, the differentiation between localities on both sides of the Almeria-Oran front and between FU and localities from the Atlantic Ocean, is reinforced by previous results obtained with samples from FU or its vicinity (Marie et al., 2016; Nantón et al., 2017), which also indicated the existence of biogeographical barriers to gene flow of *D. trunculus* in this area.

The genetic homogeneity of the Atlantic localities detected in this study is in accordance with previous results in other marine species that live on the European Atlantic coast (Sanjuan et al., 1996; Duran et al., 2004; Triantafyllidis et al., 2005; Couceiro et al., 2007; Xavier et al., 2012; Fernández-Pérez et al., 2017a). Oceanographic characteristics of the Atlantic coast, among which a mixing of water bodies (Bowden, 1975), favour larval exchange, reducing any differentiation among populations of the Atlantic coast of the Iberian Peninsula. Most *D. trunculus* dispersal takes place during its larval planktonic phase, thus the genetic homogeneity detected in this region may be the result both of ocean currents and demographic processes that likely play leading roles in determining connectivity within this group of localities. In summer, coinciding with the *D. trunculus* spawning season (Gaspar et al., 1999), water flow along the Iberian Atlantic coast is

ruled by the Iberian Poleward Current system, with water flowing from the west coast of Portugal and Spain up to near the French coast (Frouin et al., 1990; Haynes and Barton, 1990). At the southern limit of the Iberian region, water circulation is governed by the eastward branch of the Azores Current, which enters the Gulf of Cádiz and turns southwards towards north-west Africa (Martins et al., 2002). Commonly, the mean flow of the surface waters is southward along the western coast of the Iberian Peninsula, but seasonal wind can result in both northward and southward flows (Martins et al., 2002). Also, the lack of genetic structure in this group could be influenced by recent population expansion. Such an event is reflected by the networks analysis, which revealed a star-like pattern, characteristic of exponentially growing populations after a bottleneck (Slatkin and Hudson, 1991), and by the significant results obtained in neutrality tests. Furthermore, the 16S and Cytb data clearly indicated that all sampled individuals were probably the result of recent divergence from a common ancestral haplotype (H1 in all cases).

The Mediterranean group (GA, DE and MA) also showed genetic homogeneity. Neutrality tests for the F type genes suggested the existence of population expansion. However, neutrality tests for the M type were not significant, indicating the existence of a constant population size over a long period of time. Moreover, in the case of the three localities included in the Mediterranean group, AMOVA analyses with both mtDNA types did not detect significant differentiation among populations within this group (Fsc = -0.0009 and Fsc = -0.00172 for the F and the M type, respectively) and pairwise F_{ST} for M and F types did not show significant differentiation after Bonferroni among GA, DE and MA localities.

Implications for conservation management

Some marine resources, such as the wedge-clam *D. trunculus*, have shown great fluctuations in time and short- and long-term alterations have been generated. These variations have produced a decrease in the commercialization of the resources and, consequently, a decline in the socioeconomic conditions of the extractive fishery sector in Spain, and especially in Galicia. To increase the production of exhausted natural bivalve beds in a sustainable way several practices

are possible. One option is the introduction of adult individuals from other genetically similar areas where this species is abundant (Camara and Vadopalas, 2009). Repopulation could be a way to maintain exploitation of natural stocks and promote the development of the fishing sector. Other possible and common action for restocking and enhancement of bivalve species is the release of seed obtained in breeding facilities or hatcheries (Camara and Vadopalas, 2009). In the case of *D. trunculus*, it has been recently demonstrated that there is a real possibility of obtaining seed in hatcheries for release in natural beds to reach commercial size (Louzán et al., 2016). Furthermore, other measure should be implemented to relieve some fishing pressure on this species and apply appropriate mating schemes for the restocking programme (Miller et al., 2003). In the face of the collapse of a fishery, while the object species recovers, a suitable strategy is the exploitation of a similar species. Such exploitation should be carried under appropriate management practises to avoid transferring the problem to the new species. As it was indicated in a previous study by our group, *D. vittatus* could be a potential alternative exploitable resource for the *D. trunculus* fisheries (Fernández-Pérez et al., 2017a). However, it must be seen as an important but temporary emergency measure to rescue and maintain genetically unique populations of *D. trunculus* and their variability until their natural bank can be restored.

In this way, the conservation and restoration of natural clam beds requires an intense genetic evaluation using several molecular markers to assess the risk of different strategies such as restocking programmes or the translocation of individuals from some areas to others. This study corroborates the fact that *D. trunculus* displays a considerable level of genetic diversity with population differentiation along the Iberian Peninsula. However, in this case, the F type mtDNA did not detect the differentiation of localities from the Alboran Sea (FU), in contrast to the M type and previous studies based on microsatellites. The characteristics of different molecular markers could explain the difference observed, maybe because M genome evolves typically faster, potentially allowing for insight into more recent events in population history (Śmietanka et al., 2013), as microsatellite markers do. This reason makes the M genome sequences more suitable to investigate the genetic structure in *D. trunculus* populations. From a conservation point of view,

the three confirmed genetically divergent groups, Atlantic Ocean, Alboran Sea and Mediterranean Sea, should be considered as different management units. This genetic differentiation should be taken into account in relation to restocking strategies to avoid putting at risk the genetic diversity of the species.

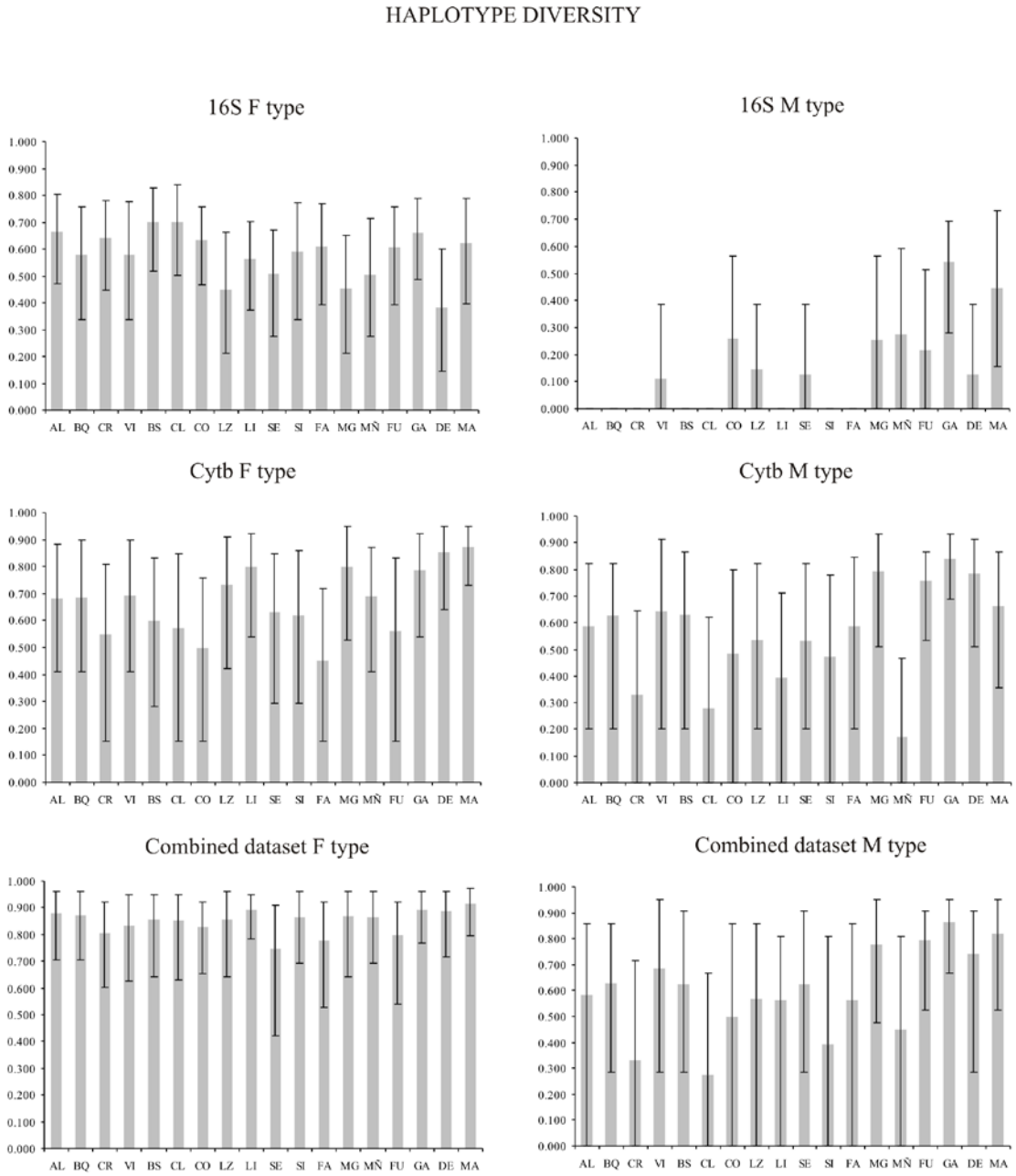
Genetic analysis plays an important role in fisheries management, helping identify stock resources and providing potentially useful management suggestions. Maintenance of genetic diversity in exploited populations is a main goal in conservation biology. In the short-term, genetic variation is relevant to the preservation of reproductive success and to population productivity, while in the long-term, it contributes to adaptive potential (Frankham et al., 2002; Allendorf et al., 2008). Genetic differentiation among populations and correct delineation of marine species boundaries are also essential for species conservation and management (Avise, 1998; Sweijd et al., 2000) because these aspects provide information about the genetic exchange among them and their evolutionary history (Allendorf et al., 2013). Moreover, evaluation of population differentiation is fundamental to assess the genetic risk associated with restocking programmes (Ward, 2006) and translocation of individuals between natural populations (Johnson, 2000). Restocking programmes must consider the genetic structure of the species, since the introduction of genotypes that are not very representative of the natural population can have negative effects. Therefore, individuals used as breeding stock in the hatchery should preferably come from the population to be repopulated, but where this is not feasible (when it is not possible to take broodstock from the stock to be enhanced or the local population is very small and suspected to be highly inbred), the most genetically similar population should be used (Ward, 2006). In this study, genetics data revealed three independent management units (i.e. Atlantic Ocean, Alboran Sea and Mediterranean Sea). Therefore, these three management units need separate management strategies, especially those relevant to translocation and propagation, and an effort to avoid homogenization should be made by managing each unit separately and restricting transfers of seed across units.

This study provided valuable information on the genetic structure of 18 localities across Iberian Peninsula of the wedge clam *D. trunculus*. The data can be useful to the sustainable management of wild stocks as well as in promoting successful restocking actions based on aquaculture production. Localities formed three groups that we suggest should be treated as separate management units. In addition, transfers of seed or broodstock across these units should be restricted.

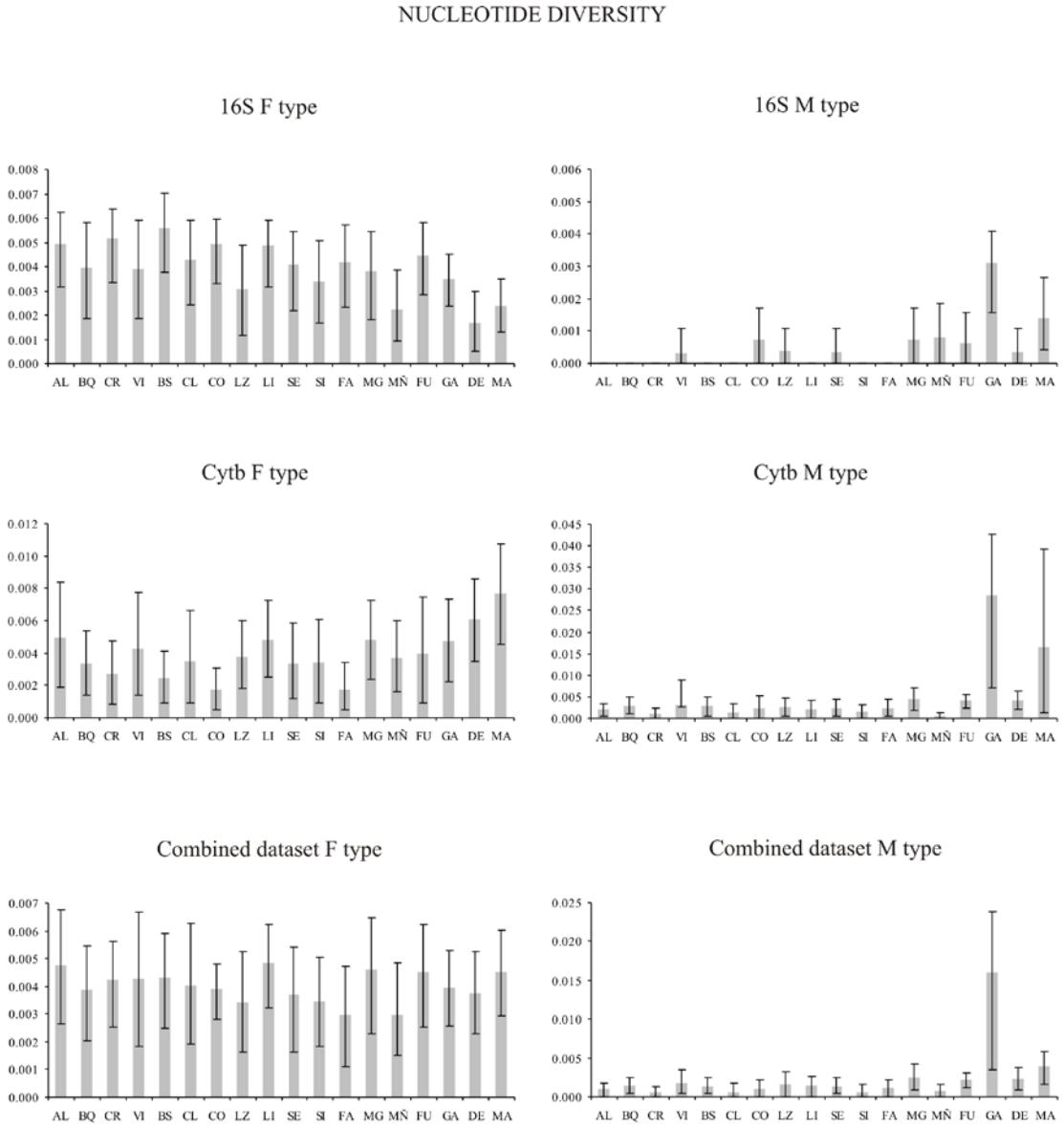
Acknowledgements

We would like to thank Dra. S. Nóvoa from Centro de Cultivos Marinos de Ribadeo – CIMA (Xunta de Galicia) and Dr. M. B. Gaspar from Instituto Português do Mar e da Atmosfera – IPMA (Portugal) for providing specimens, and Dra. L. Martínez and K.M. Hein for her constructive and valuable comments. The authors wish to thank Professor John Baxter for helpful remarks and suggestions that improved the quality of the manuscript. We also thank two anonymous referees for useful comments during manuscript review. This work was supported by the Ministerio de Economía y Competitividad (Spain) through project AGL2016-75288-R AEI/FEDER, UE.

Supporting information



S1 Figure. Confidence intervals of haplotype diversity means for 16S F type, 16S M type, Cytb F type, Cytb M type, combined F type sequences and combined M type sequences.



S2 Figure. Confidence intervals of nucleotide diversity means for 16S F type, 16S M type, Cytb F type, Cytb M type, combined F type sequences and combined M type sequences.

Chapter 2

Genetic diversity and population genetic analysis of *Donax vittatus* (Mollusca: Bivalvia) and phylogeny of the genus with mitochondrial and nuclear markers

Jenyfer Fernández-Pérez[1], Elsa Froufe[2], Ana Nantón[1], Miguel B. Gaspar[3,4], Josefina Méndez[1]

[1] Grupo Xenomar, Departamento de Bioloxía, Facultade de Ciencias and CICA (Centro de Investigacións Científicas Avanzadas), Universidade da Coruña, Campus de A Zapateira, 15071 A Coruña, Spain

[2] CIIMAR/CIMAR - Interdisciplinary Centre of Marine and Environmental Research, University of Porto, Terminal de Cruzeiros do Porto de Leixões, Av. General Norton de Matos s/n, 4450-208 Matosinhos, Portugal

[3] Instituto Português do Mar e da Atmosfera, I.P., Av. 5 de Outubro s/n, 8700-305 Olhão, Portugal

[4] Centro de Ciências do Mar (CCMAR), Universidade do Algarve, Portugal

Published in *Estuarine, Coastal and Shelf Science*, 197: 126-135 (2017).

https://doi.org/10.1016/j.ecss.2017.08.032

Abstract

In this study, the genetic diversity of *Donax vittatus* across the Iberian Peninsula was investigated using four mitochondrial (COI, Cytb, 16S F and M types) and three nuclear (H3, 18S and 28S) genes. These same molecular markers were also sequenced in *Donax semistriatus* and *Donax variegatus* to address the phylogenetic relationships of the species of the genus *Donax* common along the European coasts. Our results showed high haplotype diversity in combination with a low nucleotide diversity and star-shaped networks with a predominant haplotype, indicating a recent population expansion for the examined sampling sites of *D. vittatus*. Furthermore, analyses of population differentiation performed with COI mitochondrial marker, including global F_{ST} estimation and pairwise F_{ST} values, indicated the non-existence of significant genetic structure in *D. vittatus* of Atlantic Iberian populations. Because these localities show a high genetic similarity, we suggest that *D. vittatus* could be a potentially alternative exploitable resource, as complement to the *D. trunculus* fisheries, whose natural stocks have decreased dramatically in some areas. Furthermore, we present for the first time, evidence of DUI in the clams *D. vittatus* and *D. semistriatus.*

Introduction

The wedge-clams of the genus *Donax* are widespread on exposed sandy beaches of tropical and temperate coasts (Ansell, 1983). Five *Donax* species occur along the European littoral: *Donax vittatus* belonging to the Mediterranean boreal fauna, extending to the coast of Norway; *Donax semistriatus* with a limited distribution to the Mediterranean and Black Seas, although also appearing in the Portuguese coast (Rufino et al., 2010); *Donax trunculus* and *Donax variegatus*, Atlantic-Mediterranean species of warm waters, which reach the Black Sea and Atlantic coast of France, and the southwest coasts of France and England, respectively (Ansell and Lagardère, 1980; Bayed and Guillou, 1985; Salas-Casanova, 1987); and *Donax venustus* with a distribution throughout the Mediterranean and Black Seas. However, *D. venustus* is practically non-existent in

the Iberian Peninsula as a single individual has been found between the years 2000 and 2006 along the south coast of Portugal (Rufino et al., 2010). Thus, on the Atlantic coast of the Iberian Peninsula, several species of the genus can coexist in sympatry. For instance, *D. trunculus* and *D. vittatus* live sympatrically in the same beaches of Portugal (Gaspar et al., 2002; Rufino et al., 2010). Moreover, although Bejega et al. (2010) recorded the existence of *D. vittatus* in earlier times in Galicia (northwest of Spain), its presence has not been detected in the last decades. However, in a multidisciplinary study in which our research team has collaborated, we detected the presence of *D. vittatus* on the Galician coast (in Rías Baixas: Praia América, A Lanzada and Corrubedo) using molecular markers developed in our laboratory (Pereira et al., 2012; Nantón et al., 2015).

D. trunculus is commercially exploited in France (Thébaud et al., 2005), Italy (Zeichen et al., 2002), Turkey (Deval, 2009; Özden et al., 2009), Portugal (Gaspar et al., 1999; Chícharo et al., 2002), and Spain (Ramón et al., 2005; Molares et al., 2008), and constitutes an important shellfish resource due to its high economical value. In Galicia, *D. trunculus* is the bivalve with greater commercial value in the markets during the last years (42.37 €/kg in the year 2017), and its value has increased with the consequent increase in its fishing pressure. Although natural wedge-clam beds have started to show signs of overexploitation, demand has continued to increase. Wild stocks have decreased dramatically in this region and are currently restricted to specific areas of the north coast (Martínez-Patiño et al., 2003). Furthermore, some *D. trunculus* localities of the Iberian Peninsula seem to be at high long-term risk of extinction (Marie et al., 2016).

In the face of the collapse of a fishery, a possible action is the assessment of a similar species as an alternative resource to the exploited one, allowing to increase the number of captures without increasing the pressure on the natural beds. Among the several bivalve species susceptible of being commercially exploited, *D. vittatus* stands out. Its vast area of distribution (along most of the Atlantic coast of the Iberian Peninsula), high yields and population structure suitable for commercial exploitation (Gaspar et al., 2004) make this species a good complement to the fishery of *D. trunculus.* In this line, studies have been carried out in Portugal to propose *D. vittatus* as an alternative species to the white clam *Spisula solida* (Maia and Pimenta, 2007). So, altogether, the

increased captures and the increasing commercial value of the wedge clam *D. trunculus* have inclusively raised some expectations about the potential of *D. vittatus* as a new species for molluscan aquaculture, both for commercial production and for stock enhancement in the northwest of Iberian Peninsula. In this context, genetic analyses about species are essential for establishing successful stock enhancement or restocking operations (Bell et al., 2005; Gaffney, 2006; Wenne et al., 2007).

Mitochondrial DNA (mtDNA) is a useful marker to investigate genetic diversity and it has been used to analyse polymorphism and define stocks in many bivalve marine species (Arnaud-Haond et al., 2003; Katsares et al., 2008; Luis et al., 2011; Martínez et al., 2015), due to technical simplicity, abundance of selectively neutral mutations, low rate of recombination and uniparental inheritance. However, a major peculiar feature of mtDNA in some bivalve species is the presence of an alternative pattern of mitochondrial inheritance, known as Doubly Uniparental Inheritance (DUI) (Breton et al., 2007; Passamonti and Ghiselli, 2009; Zouros et al., 2013 for reviews), that involves two separate lineages of mtDNAs. One mtDNA (called F) is transmitted from the mother to the complete offspring, whereas the other one (M) is transmitted from the father to sons. Usually, females have the F mtDNA in all their tissues (homoplasmic), whereas males contain the F mtDNA in their somatic tissues, and the M mtDNA in their gametes (heteroplasmic) (Skibinski et al., 1994; Zouros et al., 1994a). The presence of two mtDNA lineages that evolve independently within the same species is an important fact to consider in population genetic studies because they can interfere with the interpretation of the results (Ladoukakis et al., 2002). In addition, mtDNA loci also contain the signatures of historical demographic events that can play a major role in establishing contemporary biogeographic patterns and population structure.

Nevertheless, mtDNA must be used in combination with nuclear markers to accurately identify populations for conservation (Moritz, 1994). Association of mitochondrial and nuclear sequences should allow a more reliable and detailed scene of historical and present-day population structure of *D. vittatus*. In addition, in population genetics, the analysis of several loci has become a requirement, as studies based on nucleotide variation at a single locus provide insufficient

information of genetic patterns. In this sense, the selection of loci with different evolutionary histories, such as mitochondrial vs nuclear genes, is important to differentiate among factors affecting genetic variation (e.g. variation due to a selective sweep or a population bottleneck). In fact, the level of mutation is expected to vary among different genomic regions. For instance, in phylogenetics, more conserved genes are usually employed to resolve internal nodes in the phylogeny, whereas those more variable are, in general, able to resolve the external nodes (Vierna et al., 2012).

In this study, we used information from four mitochondrial (Cytochrome c oxidase subunit I, COI; Cytochrome b, Cytb; and 16S rDNA F and M types) and three nuclear (H3, 18S rDNA 28S rDNA) genes from 140 individuals, collected in 6 localities sampled across the *D. vittatus* distribution in Iberian Peninsula. Sequences of these seven genes were combined and analysed to evaluate the genetic diversity, population differentiation and demographic patterns of *D. vittatus*, in order to discuss the conservation implications of these species in the study area. Additionally, we aim to clarify and establish the phylogenetic relationships of genus *Donax* species´ common along the European coasts. This is the first study based on mitochondrial and nuclear markers on the population genetic diversity and structure of *D. vittatus*.

Materials and methods

Sample collection and DNA extraction

D. vittatus specimens were collected from 6 localities in the northeast coast of the Atlantic Ocean (Figure 1 and Table 1). In order to investigate the phylogenetic relationships between *Donax* species from de Iberian Peninsula, we also collected 8 individuals of *D. semistriatus* and 8 individuals of *D. variegatus* from the southern Portuguese coast (Figure 1 and Table 1). Gender determination was performed on each individual, when possible, by microscopic examination of gametogenic tissue from the visceral mass, and was based on the presence of eggs or sperm. Total genomic DNA was extracted from ethanol-preserved foot using a Chelex-100 (Sigma-Aldrich, USA) protocol based on Walsh et al. (1991). In males, gonadal tissue was sampled to sequence genes from the male-specific (M type) mitochondrial lineage and its genomic DNA was extracted using a standard high-salt protocol (Sambrook et al., 1989). Specimens were taxonomically identified using Pereira et al. (2012) and Nantón et al. (2015) protocols.

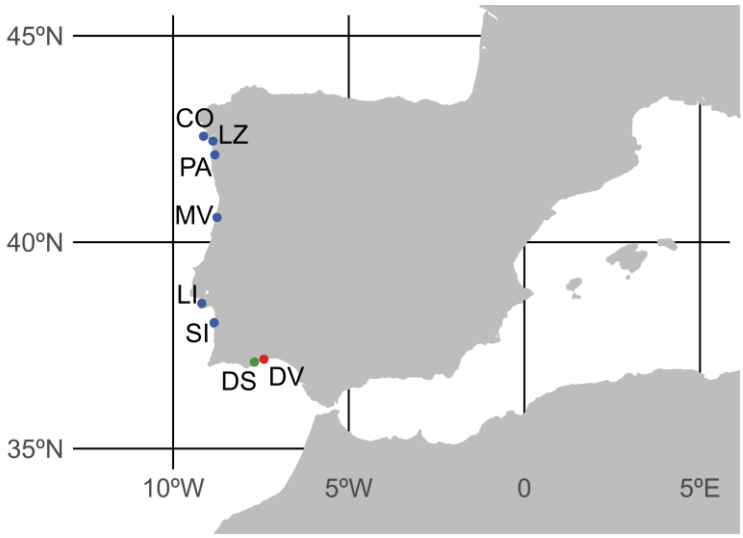


Figure 1. Map showing the location of the sampled *Donax* localities. Blue: *D. vittatus* (CO: Corrubedo, LZ: A Lanzada, PA: Praia América, MV: Mira-Vagueira, LI: Lisboa, SI: Sines); green: *D. semistriatus* (DS); and red: *D. variegatus* (DV).

Table 1. Sampling details for *D. vittatus*, *D. semistriatus* and *D. variegatus*.

Code	Species	Locality	Region	Country	Sample size	Latitude	Longitude
CO	*D. vittatus*	Corrubedo	A Coruña	Spain	25	42°34´N	09°03´W
LZ	*D. vittatus*	A Lanzada	Pontevedra	Spain	25	42°27´N	08°52´W
PA	*D. vittatus*	Praia América	Pontevedra	Spain	20	42°07´N	08°49´W
MV	*D. vittatus*	Mira-Vagueira	Beira Litoral	Portugal	25	40°36´N	08°45´W
LI	*D. vittatus*	Lisboa	Estremadura	Portugal	22	38°31´N	09°11´W
SI	*D. vittatus*	Sines	Baixo Alentejo	Portugal	22	38°03´N	08°50´W
DS	*D. semistriatus*	Monte Gordo	Algarve	Portugal	8	37°10´N	07°30´W
DV	*D. variegatus*	Monte Gordo	Algarve	Portugal	8	37°06´N	07°36´W

Amplification and sequencing

A total of 138 *D. vittatus* specimens, 8 *D. semistriatus* and 8 *D. variegatus* were amplified for the mtDNA cytochrome c oxidase subunit I gene (COI; ca. 550 bp fragment) with COIDonax-F and COIDonax-R primers (Nantón et al., 2015); the mtDNA cytochrome b oxidase (Cytb; ca. 400 bp fragment), with UCYTB144F and UCYTB272R primers (Merritt et al., 1998); and the mtDNA 16Sr DNA (16S rRNA; ca. 447 bp fragment), with the universal primer 16AR (Kocher et al., 1989) and reverse primer DT16S-R395 (Theologidis et al., 2008). These primers produced a 447 bp fragment of the 16S rDNA from the DNA extracted from the foot of both female and male specimens (F type), but not from male gonads (M type). So, a new set of primers (DT16SM-F and DT16SM-R) was designed using Primer3 software v.0.4.0 (http://bioinfo.ut.ee/primer3-0.4.0/) based on M type sequences from *D. trunculus* retrieved from Theologidis et al. (2008). These primers produced a 413 bp fragment of the 16S rDNA M type. The nuclear H3 gene (H3; ca. 350 bp), with H3F and H3R primers (Colgan et al., 1998); the nuclear 18S rRNA (18S; ca. 1700 bp fragment), with primers described by Sogin (1990); and the nuclear 28S rRNA (28S; ca. 700 bp), with 28S NLF796 (Sørensen et al., 2015) and LSU1600R (Williams et al., 2003) primers (see Table 2 for detailed information on primers).

Table 2. PCR primers.

Locus	Primer pair	Forward primer (5´to 3´)	Reverse primer (5´to 3´)	Reference
COI	COIDonax-F/COIDonax-R	GAGTDATAATRCGKACTGARYTRATAC	GTGAATRAATAAHACNGGRTCYCC	Nantón et al. 2015
Cytb	UCYTB144F/UCYTB272R	TGAGSNCARATGTCNTWYTG	GCRAANAGRAARTACCAYTC	Merritt et al. 1998
16S F type	16SAR/ DT16S-R395	CGCCTGTTTATCAAAAACAT	CYYTAATYCAACATCGAGGTC	Kocher et al.1989/Theologidis et al. 2008
16S M type	DT16SM-F/DT16SM-R	TGATTTAAACGGCTGCAGCTAA	GCGAACAGTCCCACTTCTCT	This study
H3	H3F/H3R	ATGGCTCTGACCAAGCAGACVGC	ATATCCTTRGGCATRATRGTGAC	Colgan et al. 1998
18S	18SF/18SR	CAACCTGGTTGATCCTGCCAGT	CTGATCCTTCTGCAGGTTCACCTAC	Sogin 1990
28S	NLF796/LSU1600R	GTCTTGAAACACGGACCAAGG	AGCGCCATCCATTTTCAG	Sørensen et al. 2015/Williams et al. 2003

PCR reactions were carried out in a final volume of 25 µl containing 100 ng of genomic DNA, 1 U of Taq DNA polymerase (NZYTech), 0.1 mM of each dNTP, 10 µM of each primer, 3 mM (H3 and 28S) or 2 mM (16S, 18S and Cytb) of $MgCl_2$, and the buffer recommended by polymerase suppliers. PCR cycling profile consisted of 3 min at 94 °C, 38 cycles of 30 s at 94 °C, 1 min at 45 ºC (H3 and Cytb), 49 ºC (16S), 55 °C (18S), or 58 °C (28S), 1 min at 72 °C, followed by 10 min at 72 °C. COI PCR conditions are described in Nantón et al. (2015). PCR products were migrated on 2% agarose gel electrophoresis. The gels were stained by immersion in 0.5 µg/ml ethidium bromide solution for 30 min, visualized and recorded on a transilluminator Gel Doc XR Systems (Bio-Rad, Barcelona, Spain). Amplified PCR products were sequenced on a 3730 DNA Analyzer sequencer (Applied Biosystems) at Secugen (Madrid, Spain). All haplotypes were deposited in GenBank (accession numbers MF668310 - MF668528).

Identification of mitochondrial F and M types

Given that Doubly Uniparental Inheritance (DUI) has been described in *D. trunculus* (Thelogidis et al., 2008), we tried to identify F and M types in mitochondrial genes in the three species under study. For this, COI, Cytb and 16S amplification products obtained from DNA extracted from muscle and gonadal tissue were sequenced in 10 males and 10 females of *D. vittatus* and in all specimens of *D. semistriatus*. In the case of the COI and Cytb genes, the sequences obtained did not show differences that could be associated with sex or tissue, so both genes were amplified only using muscle tissue in all individuals analysed. Nevertheless, for 16S gene, sequences obtained from DNA extracted from female gonad and muscle from both sexes showed a high number of differences with the sequences obtained from male gonad. When verifying the identity of the sequences, it was observed that the sequences amplified from female gonad and muscle of both sexes showed a high identity (86%) to the sequence of the 16S F type described by Theologidis et al. (2008) for *D. trunculus*. However, the sequences obtained from male gonad of *D. vittatus* and *D. semistriatus* had a high percentage of identity (81%) with the sequence described by these authors for M type. Because of this, the population analyses were performed separately in the F and

M types for the 16S gene, the first being amplified from DNA extracted from male and female muscle, and the second from male gonad. On the other hand, population analyses for COI and Cytb genes, were made with the sequences resulting from their amplification from muscle tissue in all individuals analysed.

Data analyses: diversity, differentiation and demographic analyses

For diversity and differentiation analyses, all individual gene sequence data were aligned via MAFFT (Katoh and Toh, 2008) using the standard FFT-NS-1 algorithm and manually checked using the BioEdit v.7.2.5 sequence editor (Hall, 1999). For each gene, haplotype networks were constructed using TCS 1.21 (Clement et al., 2000) with a threshold of 95%. A number of genetic diversity indexes (number of haplotypes (h), haplotype diversity (Hd) and nucleotide diversity (π)) were also calculated for each gene fragment using the software DnaSP v5.10.01 (Librado and Rozas, 2009). Genetic distances (*p*-uncorrected) within and between lineages were calculated with MEGA 7.0.20 (Kumar et al., 2016). In order to estimate the hierarchical distribution of mtDNA genetic differentiation (COI), within and among populations, an analysis of molecular variance (AMOVA) was performed using Arlequin v.3.5.1.2 (Excoffier and Lischer, 2010) with 10,000 permutations. Pairwise F_{ST} values were estimated also with Arlequin v.3.5.1.2 (Excoffier and Lischer, 2010), taking into account the haplotype frequencies. F_{ST}'s null distribution was obtained by bootstrapping (1,000 replicates), and the corresponding *p*-values were adjusted using sequential Bonferroni correction (Rice, 1989). In order to test for molecular signatures of demographic expansion pairwise mismatch distribution analyses (Rogers and Harpending, 1992) as well as Tajima´s D (Tajima, 1989), Fu´s F_S (Fu, 1997) and Ramos-Onsis and Rozas´s R_2 tests (Ramos-Onsis and Rozas, 2002) were performed in DnaSP v.5.10.01 (Librado and Rozas, 2009) to analyse the possibility of population expansion (1000 replicates). Graphical outputs were used to illustrate demographic model fit. The model used for expected values was the constant population size.

Phylogenetic analyses

For the phylogenetic analyses, sequence data (except the 16S M type sequences) were aligned in MAFFT (Katoh and Toh, 2008) using the standard FFT-NS-1 algorithm. A final concatenated alignment, including 137 sequences from *D. vittatus*, 8 from *D. semistriatus*, 8 from *D. variegatus* and three from outgroups (*D. trunculus*, *Meretrix meretrix* and *Venerupis philippinarum*; Table 3), was then analysed using Bayesian Inference (BI) and Maximum Likelihood (ML) methods. The best-fit models of nucleotide substitution were selected using jModelTest v.2.1.8 (Darriba et al., 2012) under the corrected Akaike Information Criterion for each gene partition and for each codon positions for the coding genes (i.e. COI, Cytb and H3 genes). The optimal chosen methods were: F81+I for the first Cytb codon positions; JC for the first H3 codon positions; GTR for the third H3 codon positions; GTR + G for 16S, 18S, 28S and the second codon positions of the COI, Cytb and H3; GTR+I for the first COI and Cytb codon positions; and HKY + G was selected for the third COI codon position. For the ML phylogenetic analyses, sequences were analyzed in RaxML 8.0.0 (Stamatakis, 2014) where GTR + G + I model was assumed for each partition with 1000 bootstrap replicates. BI analyses were run using MrBayes v.3.2.6 (Ronquist et al., 2012) with sequences also partitioned according to each region, using the previously selected models. Two independent Markov chain Monte Carlo (MCMC) runs, each comprising four linked chains (one cold and three heated; as default settings), were performed for 5,000,000 generations, sampling every 1,000 generations. The first 25% trees were discarded as burn-in.

Table 3. GenBank numbers outgroups.

Species	COI	Cytb	16S	H3	18S	28S
D. trunculus	KC429143.1	EF417550.1	EF417546.1	KC429226.1	AJ309018.1	KC429503.1
M. meretrix	DQ399402	GQ463598.1	JN969955.1	FJ429106.1	EF426291.1	DQ399408.1
V. philippinarum	KR078003.1	GQ443286.1	AM085107.1	EF670667.1	EU660737	AM779742.1

Results and discussion

Identification of mitochondrial F and M types

In this study, we present for the first time, evidence of DUI in the clams *D. vittatus* and *D. semistriatus.*

16S gene was successfully amplified in both types of mtDNA (F and M) using specific primers newly designed for each type. However, in the case of COI and Cytb genes the presence of the two types could not be detected, despite using specific M type primers for Cytb (Theologidis et al., 2008). In the case of *D. variegatus*, we could not determine the presence of DUI, because the only 8 available samples were female individuals. Therefore, all amplifications were performed from DNA extractions from somatic tissue and the sequences assumed to correspond to type F. However, it is expected that this species also presents DUI.

The detection of DUI can be complex. The high divergence between F and M types often does not allow the same pair of primers to amplify from both genomes; also the F genome may interfere with the amplification of the M in preparations from male gonads that are usually contaminated from somatic cells (Zouros, 2013). In addition, it is possible that in some males the paternal genome is in a low proportion (Ladoukakis et al., 2002). However, in this study, it seems unlikely that both contamination during tissue dissection and DNA extraction as the low proportion of the M genome have had a relevant effect, since in the case of the 16S gene both F and M types have been amplified satisfactorily. Nucleotide sequences of both types are another important factor for their detection, i.e. the differences between the two types may be so small as to make it difficult to identify them (Theologidis et al., 2008). This homogeneity between the sequences may be due to the existence of mtDNA recombination, a phenomenon described in the mussels *Mytilus galloprovincialis* (Ladoukakis and Zouros, 2001) and *Mytilus trossulus* (Burzyński et al., 2003). In contrast, if the gene sequences in the two types of mtDNA are highly divergent, primers that amplify the sequence in one type may fail in the other (Theologidis et al., 2008). In fact, high levels of divergence have been detected between the two mtDNA types (~10 – 30%) in some species which could complicate

their detection and amplification (Passamonti et al., 2003). In this study, it was not possible to detect the two types of mtDNA for COI and Cytb genes. Although no specific mitochondrial lineage primers exist for the amplification of COI of *Donax* species, in the case of the Cytb gene, specific primers for type M in *D. trunculus* were used (Theologidis et al., 2008), but no amplicons were obtained.

Genetic diversity

Overall diversity estimates for each locality and for each fragment are shown in Table 4. As expected, COI, Cytb and 28S were more polymorphic than 16S, 18S and H3. Regarding mitochondrial markers, for COI and Cytb genes, an alignment of 553 bp from 138 *D. vittatus* individuals showed 60 haplotypes (54 of which were private), and an alignment of 388 bp from 132 individuals, exhibited 71 (62 private) haplotypes, respectively. Globally, high Hd (COI: 0.879; Cytb: 0.939) and low π (COI: 0.0036; Cytb: 0.0069) were detected. Regarding 16S gene fragment, an alignment of 411 bp for the F type from 133 individuals resulted in 21 haplotypes (16 private) and an alignment of 395 bp for the M type from 41males exhibited only 2 haplotypes (1 private). It is noted that there are no diversity values calculated for 16S type M gene in both CO and LZ localities, because the specimens were not mature and could not be sexed. Altogether, low Hd (16S F type: 0.434; 16S M type: 0.049) and low π (16S F type: 0.0014;16S M type:0.0003) were detected (Table 4). However, our results show higher haplotype diversity in the F lineage compared with the M lineage. This difference between the two types of mtDNA takes its extreme form in PA, MV and LI localities, where there is no polymorphism for the M type. This result is in line with the variability levels obtained in previous studies in *M. galloprovincialis* (Ladoukakis et al., 2002), *Musculista senhousia* (Passamonti, 2007), and *Perumytilus purpuratus* (Vargas et al., 2015), other species with DUI in which greater diversity values were detected in the F type than in M type when analysing the same mitochondrial region, which may be explained by an older evolutionary history of the F genome than the M genome (appearing at a more recent date). Nevertheless, this difference may also be due to the lower frequency with which the M genome is observed, meaning that its

diversity may be underestimated (Vargas et al., 2015). Furthermore, several authors suggest that the M genome is under relaxed selection due to its limited function in the male germline, whereas the F genome´s role is to support the female germline and the somatic cell line of both sexes (Passamonti, 2007; Zouros, 2013). Moreover, if in the founding population, males were much fewer than the females, this would likely result in lower variability in the M type (Passamonti, 2007; Guerra et al., 2014). This scenario seems appropriate to explain observed difference between the two types of mtDNA in *D. vittatus*, as this species could be exposed to changing conditions during Last Glacial Maximum (LGM) in its distribution range in Europe as occurred in other species of bivalve molluscs (Krakau et al., 2012; Martínez et al., 2015). Nevertheless, it has to be mentioned that to better test for this scenario, a wide population sample would be needed. See below the haplotype networks details which help interpret these results.

In respect to nuclear markers, for 18S gene fragment, an alignment of 1707 bp from 135 *D. vittatus* individuals, revealed that all sequences were identical, showing 1 haplotype, so that Hd and π values are equal to zero. For 28S and H3 genes, an alignment of 717 bp from 135 individuals presented 11 haplotypes (3 private), and an alignment of 338 bp from 133 individuals showed 8 haplotypes (5 private), respectively. Globally, high Hd (28S: 0.780; H3: 0.525) and low π (28S: 0.0026; H3: 0.0017) were detected (Table 4).

Table 4. Genetic variability across *D. vittatus* populations. N = Number of individuals; h = number of haplotypes; ph = private haplotypes; Hd = haplotype diversity; and π = nucleotide diversity. Locality names follow Figure 1 and Table 1.

	CO	LZ	PA	MV	LI	SI	TOTAL
COI							
N	25	25	20	24	22	22	138
h	14	16	8	11	14	14	60
ph	9	11	6	8	10	10	54
Hd	0.860	0.933	0.758	0.822	0.900	0.922	0.879
π	0.0036	0.0039	0.0023	0.0030	0.0047	0.0036	0.0036
Cytb							
N	22	25	19	22	22	22	132
h	16	20	14	13	15	14	71
ph	11	15	9	9	11	7	62
Hd	0.948	0.970	0.965	0.870	0.939	0.874	0.939
π	0.0088	0.0082	0.0053	0.0052	0.0082	0.0052	0.0069
16S F type							
N	22	25	19	23	22	22	133
h	4	6	5	6	6	7	21
ph	-	4	3	2	4	3	16
Hd	0.260	0.367	0.462	0.458	0.537	0.541	0.434
π	0.0007	0.0016	0.0015	0.0017	0.0017	0.0015	0.0014
16S M type							
N	-	-	6	11	13	11	41
h	-	-	1	1	1	2	2
ph	-	-	-	-	-	1	1
Hd	-	-	-	-	-	0.182	0.049
π	-	-	-	-	-	0.0009	0.0003
18S							
N	23	25	19	24	22	22	135
h	1	1	1	1	1	1	1
ph	-	-	-	-	-	-	-
Hd	-	-	-	-	-	-	-
π	-	-	-	-	-	-	-
28S							
N	23	24	20	23	22	21	133
h	6	6	5	7	7	8	11
ph	-	-	-	1	2	-	3
Hd	0.731	0.670	0.784	0.854	0.784	0.767	0.780
π	0.0024	0.0025	0.0029	0.0026	0.0026	0.0022	0.0026
H3							
N	23	25	19	24	22	22	135
h	5	3	4	2	2	3	8
ph	2	1	2	-	-	-	5
Hd	0.605	0.508	0.530	0.580	0.492	0.545	0.525
π	0.0021	0.0016	0.0018	0.0015	0.0015	0.0017	0.0017

Population genetic structure and demographic analyses

Haplotype networks based on mtDNA and nuclear sequences are presented in Figure 2. Similar networks were recovered for all the mitochondrial genes, which showed a star-like topology, typical of an expansion scenario, centred at the most frequent haplotype, except the 16S M type, with only two haplotypes recovered. The networks obtained for Cytb and 16S F show one high frequency central haplotype and several additional haplotypes connected by few-step mutations, while the COI network revealed two central haplotypes, both connected with several additional low frequency haplotypes. In more detail, the most common haplotype represents 22% (29 of 132 individuals for Cytb), 24.6% (34 of 138 individuals for COI), 75.2% (100 of 133 individuals for 16S F type) and 97.6% (40 of 41 individuals for 16S M type) of the individuals sampled from all localities. Therefore, it was assumed that the dominant haplotype found is most likely the starting point of all other haplotypes. These star-like patterns (Figure 2) indicate an expanding population because new haplotypes mostly arise from the most abundant haplotype. In addition, these networks topologies associated with high haplotype diversity and low nucleotide diversity, as shown in Table 4, are often an indication of a recent population expansion in the history of the species. The low differentiation and frequency of private haplotypes also points to a population expansion from a small number of individuals after a genetic bottleneck or a founder event, such as a recolonization from a glacial refugia after the last European Ice Age (Allcock and Strugnell, 2012).

On the other hand, the nuclear gene networks (except to 18S sequences which show 1 haplotype), showed that the most common haplotype represents 66.7% (90 of 135 individuals for H3) and 39.1% (52 of 133 individuals for 28S) of the individuals sampled from all localities.

COI gene showed non-significant differentiation ($p > 0.001$) among localities (global F_{ST} was 0.01418). The highest F_{ST} values were found in the pairwise combinations involving CO, PA, LI and SI localities (Table 5). However, only one (PA-LI), out of the 15 pairwise combinations, was significant after sequential Bonferroni correction (Table 5). Therefore, the analyses performed with the mitochondrial marker COI, including global F_{ST} estimation and pairwise F_{ST}, indicated the absence of significant genetic structure in *D. vittatus* along the Atlantic coast of Iberian Peninsula.

The genetic homogeneity detected in this region has also been documented for other marine invertebrates with high dispersal potential along northeast Atlantic coast: the mussel *M. galloprovincialis* (Sanjuan et al., 1996), the sea urchin *Paracentrotus lividus* (Duran et al., 2004), the lobster *Homarus gammarus* (Triantafyllidis et al., 2005), the netted dog whelk *Nassarius reticulatus* (Couceiro et al., 2007) and the isopods *Stenosoma lancifer* and *Stenosoma acuminatum* (Xavier et al., 2012). Furthermore, a large population group with a very limited structure, ranging from the Iberian Peninsula to the British Isles, and genetically differentiated from a more northern group, has been observed in other marine invertebrates: the flat oyster *Ostrea edulis* (Launey et al., 2002), the green crab *Carcinus maenas* (Roman and Palumbi, 2004), the common cockle *Cerastoderma edule* (Krakau et al., 2012; Martínez et al., 2015) and the great scallop *Pecten maximus* (Morvezen et al., 2016).

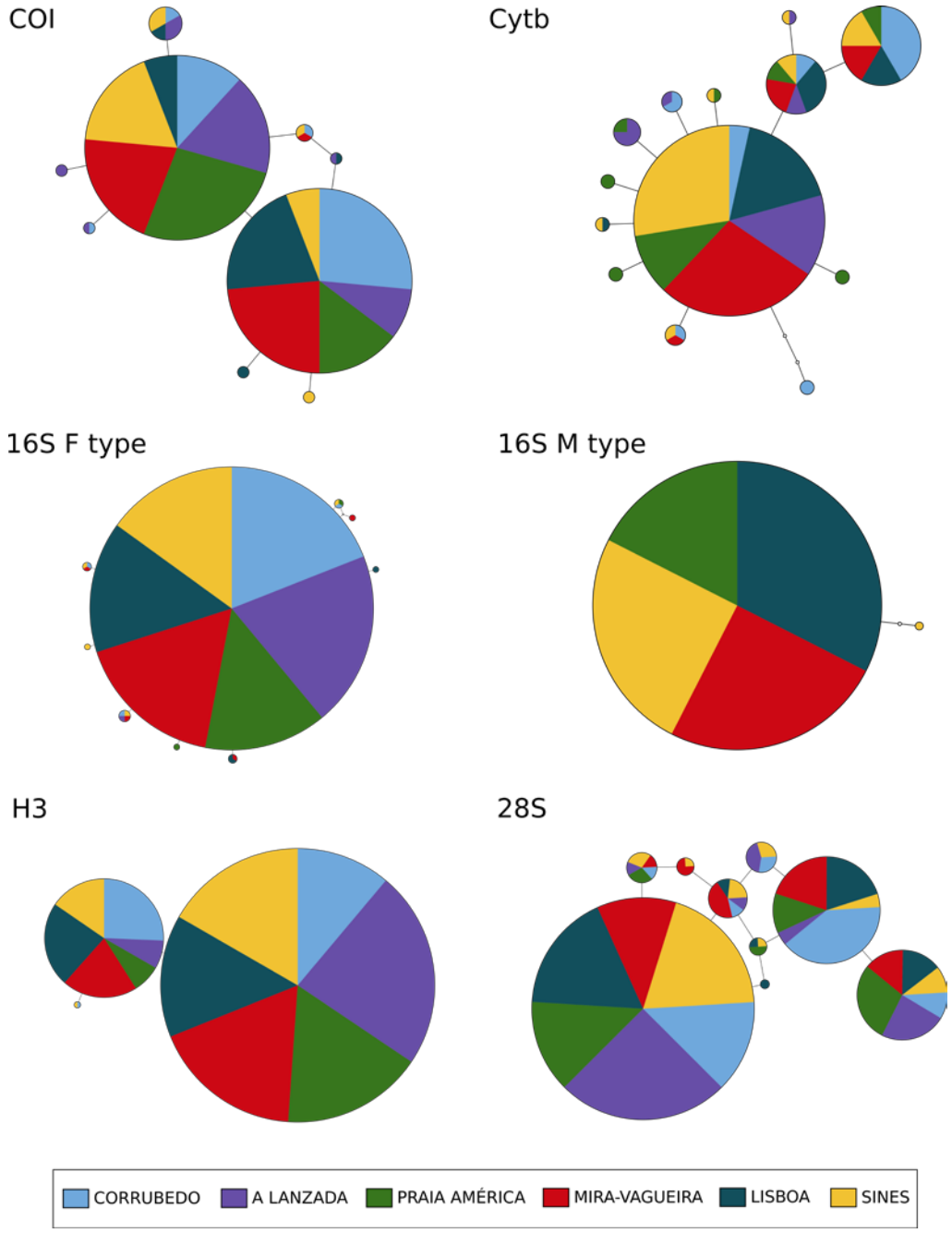


Figure 2. Haplotype (TCS) networks (excluding singleton haplotypes) for each gene showing the relationship of *D. vittatus* individuals sequenced. Each *circle* represents a haplotype, and its size is proportional to its frequency. *White dots* indicate hypothetical intermediate haplotypes. *Colours* denote the geographic origin of each haplotype.

Table 5. Pairwise F_{ST} values for COI. Significant values are followed by an asterisk: (*) $p = 0.05$; (**) significant after sequential Bonferroni correction.

	CO	LZ	PA	MV	LI	SI
CO	-					
LZ	0.01853	-				
PA	0.03383*	0.01778	-			
MV	-0.01153	0.01347	-0.00639	-		
LI	-0.01269	0.01878	0.05670**	0.00708	-	
SI	0.02981*	-0.01241	0.01652	0.01826	0.03279*	-

Given that most *D. vittatus* dispersal takes place during its larval plancktonic phase, lack of genetic structure observed in this study may be the result of both demographic processes and ocean currents which could play a leading role in increasing the connectivity between these localities. Usually, the mean flow on the surface is southward along western coast of the Iberian Peninsula, but seasonal winds can result in both northward and southward flows (Martins et al., 2002). In summer, coinciding with *D. vittatus* spawning season (Ansell, 1972), currents along Iberia are mainly southward, while during the winter there is a predominant poleward current along the west coasts of Portugal and Spain up to near the French coast (Frouin et al., 1990; Haynes and Barton, 1990). In addition, the genetic homogeneity detected in this study could also be the result of a recent population expansion, reflected in the network analyses, as well as in the mismatch distributions, which showed similar unimodal patterns, supporting scenarios of demographic expansion in the six localities (Figure 3). These results are in accordance with significant results obtained in the neutrality tests (Table 6): negative and significant values showed by both Tajima´s D and Fu´s F_S statistics and the statistically significant Ramos-Onsis´s R_2 value, all indicative of population expansion. Furthermore, mitochondrial data clearly indicated that all individuals sampled were probably the result of recent divergence from a common ancestral haplotype (central and most common haplotypes). Similarly, the presence of a large number of singleton haplotypes and one or two predominant haplotypes is a common pattern in other marine organisms associated with Pleistocene glaciations (Baker et al., 2008; Sotelo et al., 2009). Recovery from a bottleneck or selective sweep could also cause genetic patterns similar to those generated by a phenomenon of

population expansion. However, no bottleneck signal was detected in this group of localities, and it is likely that selective sweep has occurred on multiple unlinked mtDNA and nuclear markers (Strasser and Barber, 2009).

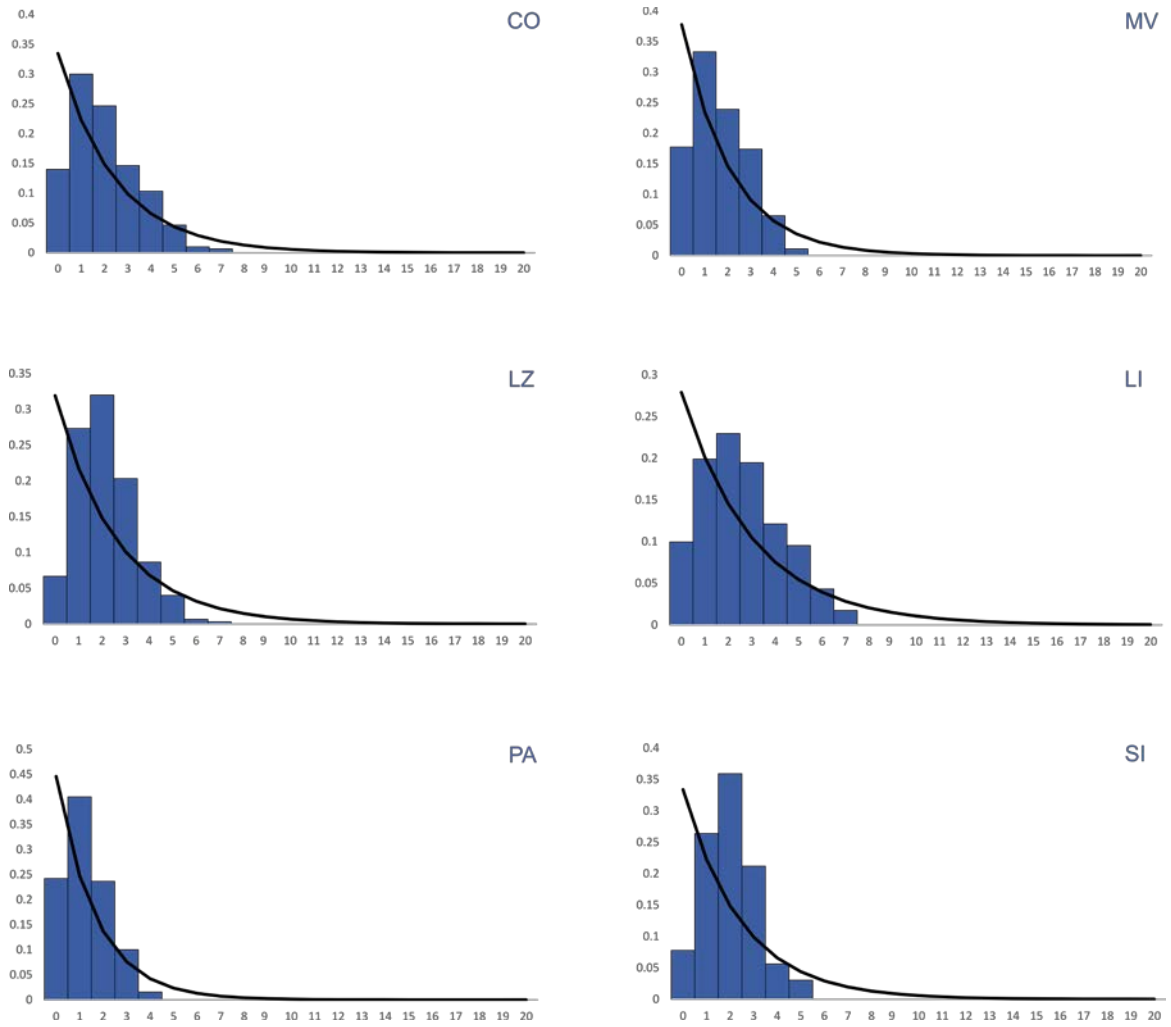


Figure 3. Mismatch-distributions for COI in the 6 localities of *D. vittatus*. Black curves show the expected distribution of mutations according to the null of demographic expansion. The number of pairwise differences and their frequencies is shown on the horizontal and vertical axes, respectively.

According to the F_{ST} pairwise results (Table 5) and networks (Figure 2), samples were afterwards pooled into one group and the results of AMOVA (based on this group) showed that most of the variation corresponds to differences within localities (98.58%), and once again displayed absence of population structure (Table 7).

Table 6. Neutrality test for COI. Significant values are followed by an asterisk: (*) $p < 0.05$ (**) $p < 0.01$

Locality	Species	*Fs*	*D*	R_2	SDS	Raggedness index
CO	*D. vittatus*	-9.792**	-1.750*	0.0563*	0.00222	0.04493
LZ	*D. vittatus*	-13.122**	-1.870*	0.0495**	0.00609	0.07542
PA	*D. vittatus*	-4.056*	-1.755*	0.0777	0.00347	0.08105
MV	*D. vittatus*	-6.310*	-1.963**	0.0556**	0.00187	0.05229
LI	*D. vittatus*	-8.727**	-1.966**	0.0533**	0.00100	0.02179
SI	*D. vittatus*	-10.866**	-1.718**	0.0589**	0.01250	0.09127
TOTAL	*D. vittatus*	-97.596**	-24.749**	0.0176**	0.00092	0.05318

Table 7. AMOVA results using the COI gene data.

Source of variation	d.f.	Sum of squares	Variance components	% of variation	*F* index
Among populations	5	2.888	0.00624 Va	1.42	F_{ST} = 0.01418
Within populations	132	57.315	0.43420 Vb	98.58	
Total	137	60.203	0.44045		

Finally, the integration of nuclear markers allows a better understanding of possible causes (demography or selection) of the observed results of mitochondrial genetic variation, because demographic events are expected to affect similarly the whole genome, whereas selection will have locus-specific effect (Andolfatto, 2001). The comparison of the levels of genetic diversity between molecular markers suggest that, for *D. vittatus*, demographic expansion may be a plausible explanation. In fact, both mitochondrial and nuclear markers show high genetic similarity.

Phylogenetic relationships of the genus *Donax*

The final combined data set of the six gene fragments included 4,099 aligned positions for 138 samples of *D. vittatus* individuals, 8 *D. semistriatus*, 8 *D. variegatus* and the 3 sequences used as outgroups, comprising 552 bp of COI (only F type COI sequences were considered), 388 bp of Cytb 415 bp of 16S, 328 bp of H3, 1696 bp of 18S and 720 bp of 28S.

The concatenated tree topologies resulting from the single tree recovered from the ML and BI approaches were congruent, with high support in the nodes, and produced topologically identical trees. The results of the BI phylogenetic analyses are shown in Figure 4. The results showed a well-resolved phylogeny where the four species of the genus *Donax* from the Iberian Peninsula form a single clade. *D. trunculus* is the basal clustering with a group formed by all the others, where *D. variegatus* is the sister clade of (*D. vittatus* + *D. semistriatus*). *D vittatus* individuals, collected along the whole Atlantic cost of the Iberian Peninsula, cluster together without once again showing evidence of population structure, and join with high support with *D. semistriatus* individuals collected in the south of the Portuguese coast (Figure 1 and Table 1).

Measures of genetic diversity for the mitochondrial COI gene among the analysed species are summarized in Table 8. Mean genetic distance (*p*-uncorrected) ranged from 1.3% between *D. vittatus* and *D. semistriatus* and 4.2% between *D. vittatus* and *D. variegatus* (Table 8). Levels of haplotype diversity were higher in *D. semistriatus* when compared to the remaining species (Table 8).

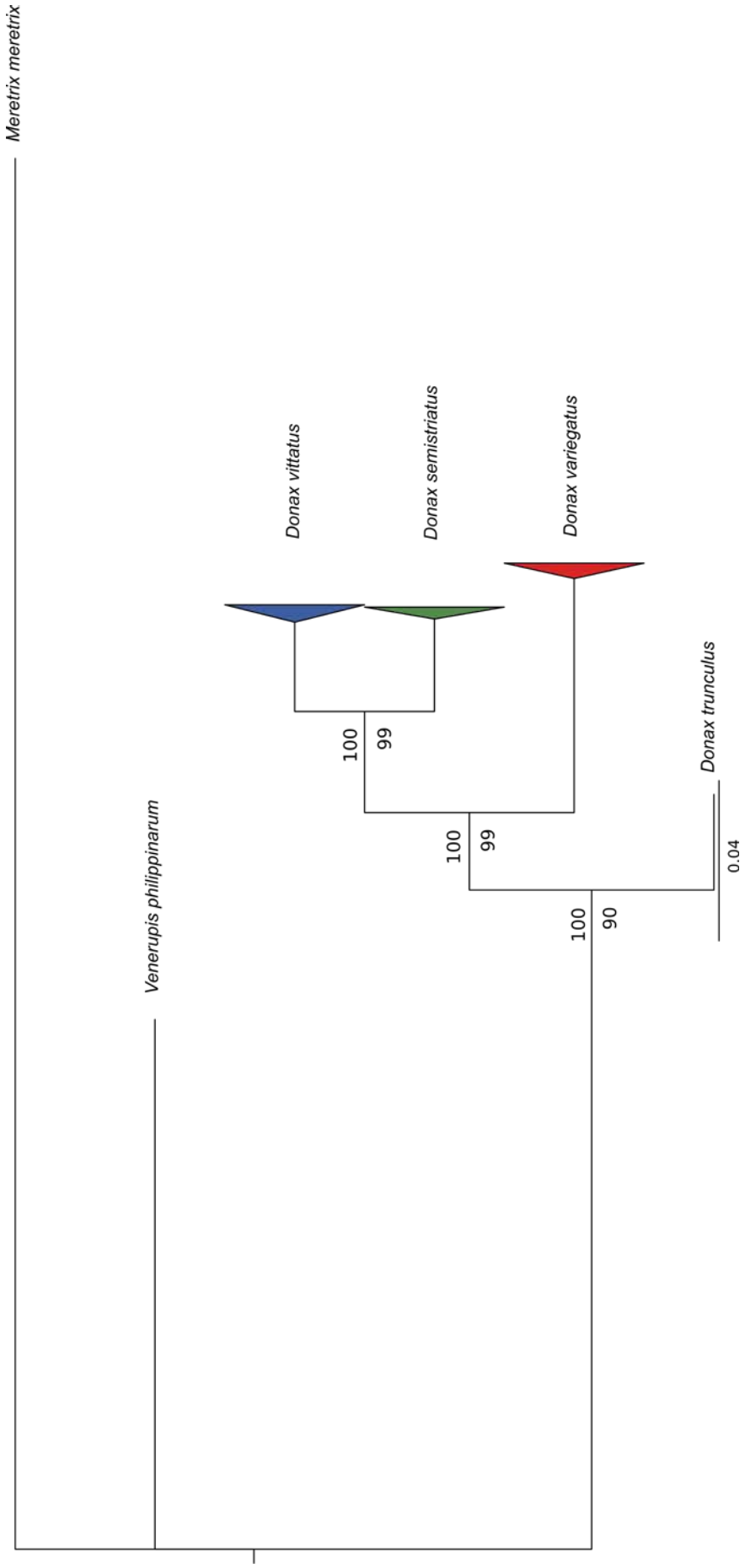


Figure 4. Phylogenetic tree obtained by Bayesian Inference (BI) analysis of *Donax* individuals (COI + Cytb + 16S + H3 + 18S + 28S). For the major nodes support values (%) are given as Bayesian posterior probability above nodes and Bootstrap support (ML) below nodes. Outgroups: *D. trunculus*, *M. meretrix* and *V. philippinarum*.

Table 8. Mean divergence for the COI data set, between the three *Donax* species depicted in the phylogeny (Figure 4). N = Number of individuals; h = number of haplotypes; Hd = haplotype diversity; and π = nucleotide diversity.

	D. vittatus	***D. semistriatus***	***D. variegatus***	**N**	**h**	**Hd**	**π**
D. vittatus				138	60	0.879 ± 0.020	0.0036
D. semistriatus	0.013			8	8	1.000 ± 0.063	0.0045
D. variegatus	0.042	0.041		8	5	0.875 ± 0.108	0.0032

Management implications

Overall, the results for mitochondrial and nuclear analyses showed high haplotype diversity in combination with a low nucleotide diversity and star-shaped networks with a predominant haplotype, indicating a recent population expansion for the examined sampling sites. Furthermore, our analyses performed with the mitochondrial marker COI, including global F_{ST} estimation and the pairwise F_{ST}, indicate the non-existence of significant genetic structure in *D. vittatus* of Atlantic Iberian populations.

Because these localities show high genetic similarity, we suggest that *D. vittatus* could be a potentially alternative exploitable resource, as complement to the *D. trunculus* fisheries, whose natural stocks have decreased dramatically in some areas of the Iberian Peninsula.

Conclusions

This study presents for the first time the population genetic diversity and structure of *D. vittatus* along the whole Atlantic coast of the Iberian Peninsula, based on different mitochondrial and nuclear gene regions, the first evidence of DUI in the clams *D. vittatus* and *D. semistriatus*, and its phylogenetic relationship with the other species of the genus *Donax* present in European coasts. Altogether, the present study provides useful data for taxonomy, ecology, biodiversity conservation and also resource management, as *D. vittatus* is a species of interest to aquaculture for food production and here shown that could be used as a complement of the fishery of *D. trunculus.*

Future research should include the use of faster evolving markers to study the present population dynamics, such as microsatellite markers, that can provide information on more recent evolutionary episodes and thus reveal some finer-grain population structure.

Acknowledgements

The authors wish to thank the two anonymous reviewers for helpful remarks and suggestions that improved the quality of the manuscript. This work was supported by the Ministerio de Economía y Competitividad (Spain) through project AGL2016-75288-R AEI/FEDER, UE and under the Framework of the Structured Program of R&D&I INNOVMAR - Innovation and Sustainability in the Management and Exploitation of Marine Resources (Reference NORTE-01-0145-FEDER-000035), namely within the Research Line ECOSERVICES, supported by the Northern Regional Operational Programme (NORTE2020), through the European Regional Development Fund (ERDF). JFP acknowledges financial support for her pre-doctoral stay at Interdisciplinary Centre of Marine and Environmental Research (CIIMAR) awarded by INDITEX-UDC 2016. EF was supported by the Portuguese Foundation for Science and Technology (FCT) under grant SFRH/BPD/108445/2015.

Chapter 3

Fifteen novel microsatellite loci, developed using Next-generation sequencing, reveal the lack of genetic structure in *Donax vittatus* from Iberian Peninsula

Jenyfer Fernández-Pérez[1], Ana Nantón[1], Alberto Arias-Pérez[1], Josefina Méndez[1]

[1] Grupo Xenomar, Departamento de Bioloxía, Facultade de Ciencias and Centro de Investigaciones Científicas Avanzadas (CICA), Universidade da Coruña, Campus de A Zapateira, 15071 A Coruña, Spain

(In Preparation)

Abstract

The wedge clam *Donax trunculus* is an important shellfish species in several southern European countries. Nevertheless, production of this species has suffered a sharp decrease, especially in Galicia (Northwest of Spain). In the face of the collapse of a fishery, while the object species recovers, a suitable strategy is the exploitation of a similar species, and *Donax vittatus* could potentially be an alternative exploitable resource. Such exploitation should be carried under appropriate management practises to avoid transferring the problem to the new species. Given the need to reveal the population structure and to elucidate appropriate management units of *D. vittatus* along the Atlantic coast of the Iberian Peninsula, fifteen variable microsatellite markers were reported and tested in seven localities (six of them previously studied with a mitochondrial COI marker). All localities displayed similar genetic diversity values (allelic richness: 10.75-15.20; expected heterozygosity: 0.813-0.845; and expected heterozygosity: 0.389-0.460). Significant deficits of heterozygotes were recorded for all populations, probably due to null alleles. Analyses of population differentiation, including global F_{ST}, pairwise F_{ST} with and without null allele correction, number of migrants and Analyses of Molecular Variance, indicated lack of genetic structure in *D. vittatus* along Atlantic coast of the Iberian Peninsula, confirming the previous results based on COI gene. These data are useful to assist the fishery development and the conservation plans of this species. Furthermore, the panel of developed microsatellites will allow to undertake further genetic population studies throughout the geographic distribution of the species, as well as other genetic analysis needed for aquaculture activities.

Introduction

The wedge clams of the genus *Donax* (Bivalvia: Donacidae) are an important constituent of the macrofauna of open sandy beaches in temperate, tropical and subtropical zones, being the dominant organisms in this type of environment (Ansell, 1983). They inhabit sandy beaches exposed to tidal rhythms, characterised by intense wave action and sediment instability (Brown and McLachlan,

1990). In the European littoral, four *Donax* species can live sympatrically in some areas (Salas et al., 2001; Rufino et al., 2010): *Donax semistriatus* (Poli, 1775), *Donax trunculus* (Linnaeus, 1758), *Donax variegatus* (Gmelin, 1791) and *Donax vittatus* (Da Costa, 1778). Of the four species, *D. trunculus* populations are capable of reaching sufficiently high densities to support large commercial fleets (Gaspar et al., 1999). In fact, it plays an important role in coastal food webs and it is commercially fished in several countries including France, Italy, Portugal, Spain and Turkey (Gaspar et al., 1999; Zeichen et al., 2002; Ramón et al., 2005; Thébaud et al., 2005; Molares et al., 2008; Özden et al., 2009; Pereira et al., 2012), being considered one of the most valuable and culinary-appreciated resources. In the Atlantic coast of the Iberian Peninsula *D. trunculus* constitutes an important fishing resource due to its high economic value (Gaspar et al., 1999; Fernández-Pérez et al., 2017a,b). For instance, in Galicia (northwest of Spain), the main producing area of bivalve molluscs in Spain, it is a highly prized species, being the bivalve with the greatest commercial value in markets over the last few years (Consellería do Mar, Xunta de Galicia). Nevertheless, European production of this species has dropped, going from 1,353 tonnes in 2005 to only 818 tonnes in 2016 (FAO-FIGIS, 2018). Its decline is especially evident in Galicia where *D. trunculus* captures per year have gone from 16.9 tonnes in 2001 to 171.1 kg in 2017 (data from Consellería do Mar, Xunta de Galicia) and nowadays the exploitation is concentrated into a few beds. Over-fishing could have caused the large-scale decrease in wild stocks and captures observed in this area but other factors might be involved such as habitat loss as a consequence of sand movements after severe storms, construction of sea promenades, type of substrate or pollution.

The decline of *D. trunculus* natural beds could be mitigated through the development of an alternative fishery or aquaculture for another species, and *D. vittatus* could be that species. Several reasons justify the selection of this species: coexistence in sympatry with *D. trunculus* in the more southern European Atlantic coast including the Iberian Peninsula (Gaspar et al., 2002; Rufino et al., 2010; Fernández-Pérez et al., 2017a), high yields, population structure suitable for commercial exploitation (Gaspar et al., 2004) and biological characteristics analogous to *D. trunculus* such as similar pattern of seasonal recruitment and growth (Ansell and Lagardère, 1980). However, the

exploitation of *D. vittatus* should be carried out under appropriate management practises. In this context, understanding genetic connectivity among populations allows to determine the appropriate units and spatial scale for fisheries conservation and management (Waples and Gaggiotti, 2006; Funk et al., 2012). In a previous study, we examined in *D. vittatus* several mitochondrial (COI, Cytb and 16S) and nuclear (H3, 18S and 28S) markers in six Atlantic localities of the Iberian Peninsula (Fernández-Pérez et al., 2017a), reporting estimates of genetic diversity and demographic features. Results showed high haplotype diversity in combination with a low nucleotide diversity and star-shaped networks with a predominant haplotype, indicating a recent population expansion for the examined sampling sites. In the same study, genetic population structure was assessed using mitochondrial COI marker. F_{ST} estimations and Analysis Molecular of Variance (AMOVA) did not detect significant genetic structure among the examined localities.

Mitochondrial gene sequences have been a standard molecular marker for inferring genetic population structure, but the use of a range of molecular markers is desirable to get more realistic vision. In effect, discrepancies between mitochondrial DNA (mtDNA) and nuclear DNA (nDNA) markers regarding the genetic differentiation patterns are not unusual (Johnson et al., 2003; Brito, 2007; Hoffman et al., 2009; Toews and Brelsford, 2012; Ferchaud et al., 2015; Martínez et al., 2015). Furthermore, contrasting results throughout markers, especially those markers that have different inheritance modes and evolve at different pace, such as mtDNA and nDNA markers, may provide new insights that could not be obtained with either type of data alone (Brito, 2007). Among the nuclear markers, the microsatellites, short tandem repeat DNA sequences with unit length of 2-6 bp (Chambers and MacAvoy, 2000), are the most popular in genetic studies related to fisheries and aquaculture. Microsatellites are codominant markers of relatively small size, that can be easily amplified by PCR. These features coupled with their high polymorphism, their mendelian inheritance pattern, and their abundance in the genome, make them very powerful genetic markers. These markers have been widely used in genetic studies in several bivalve species, including the wedge clam *D. trunculus* (Marie et al., 2016; Nantón et al., 2017; Rico et al., 2017), the cockle *Cerastoderma edule* (Martínez et al., 2015), the clam *Ruditapes decussatus* (Borrell et al., 2014;

Arias-Pérez et al., 2016) and oysters *Crassostrea gigas* (Liu et al., 2017; Zhang et al., 2018) and *Ostrea edulis* (Lallias et al., 2010). The main drawback of these markers is their development, since it is necessary to know the sequences that flank the microsatellite motifs for primers design. However, in recent years, microsatellite isolation has progressed thanks to the development of Next-generation DNA Sequencing (NGS) technologies, which have proved to be less time-consuming and more cost-effective than traditional methodologies. While NGS approach results in large amounts of sequencing data from which to isolate and develop numerous microsatellite loci, traditional methods provide typically a few hundred. In fact, improvements in NGS techniques have supplied new opportunities to discover high quality microsatellite markers, with a good representativeness of loci across the genome (Martin et al., 2010), especially in non-model organisms, for which genomic information is scarce (Shendure and Ji, 2008; Zhang et al., 2011).

To assist a sustainable exploitation of *D. vittatus*, our aims were to develop in this species a panel of microsatellite markers using NGS technology, employ them to assess the genetic diversity and population structure along the Atlantic cost of the Iberian Peninsula and contrast the results with those previously obtained based on the mitochondrial COI gene.

Materials and methods

Sample collection and DNA extraction

Fifty to sixty-one *D. vittatus* individuals were collected from seven localities along the Atlantic coast of the Iberian Peninsula (Figure 1, Table 1).

Total genomic DNA was extracted from muscle tissue preserved in ethanol following one of the two protocols described below, depending on the subsequent use. For library construction, DNA was extracted using DNAeasy Blood and Tissue Kit (Qiagen, Germany) according to manufacturer´s instructions with only a minor modification, namely EB (10mM Tris-Cl, pH 8.5) rather than AE (10mM Tris-Cl, 0.5 mM EDTA, pH 9.0) buffer was used to avoid possible

interference of EDTA with Nextera enzyme. For population genetic analysis, DNA extraction was performed using a Chelex-100 (Sigma-Aldrich, USA) protocol based on Walsh et al. (1991).

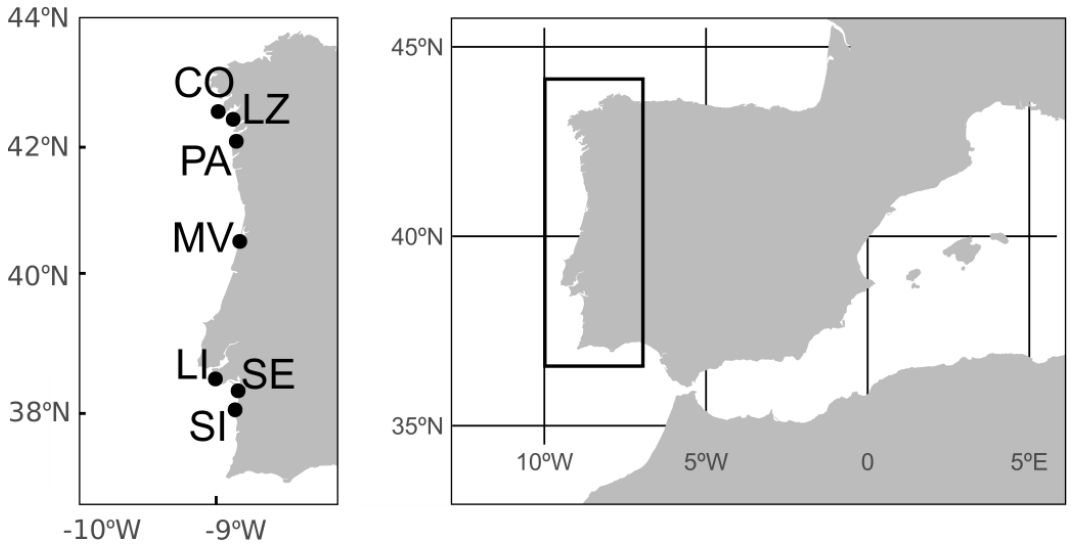


Figure 1. Geographical location of the sampling sites of *D. vittatus* along the Atlantic coast of the Iberian Peninsula (CO: Corrubedo, LZ: A Lanzada, PA: Praia América, MV: Mira-Vagueira, LI: Lisboa, SE: Setúbal, SI: Sines).

Table 1. Sampling details for *D. vittatus.*

Code	Locality	Region	Country	Sample size	Latitude	Longitude
CO	Corrubedo	A Coruña	Spain	50	42°34´N	09°03´W
LZ	A Lanzada	Pontevedra	Spain	61	42°27´N	08°52´W
PA	Praia América	Pontevedra	Spain	50	42°07´N	08°49´W
MV	Mira-Vagueira	Beira Litoral	Portugal	54	40°36´N	08°45´W
LI	Lisboa	Estremadura	Portugal	50	38°31´N	09°11´W
SE	Setúbal	Estremadura	Portugal	50	38°20´N	08°47´W
SI	Sines	Baixo Alentejo	Portugal	52	38°03´N	08°50´W

Molecular procedures, library construction and microsatellite sequencing

The quality and quantity of genomic DNA was assessed by spectrophotometry (NanoDrop ND-1000, Technologies, Inc.), fluorometry (Qubit HS, Invitrogen, USA) and 1% agarose gel electrophoresis. After quality control, DNA from eight individuals, representing different localities, was pooled for library construction at Genoscreen© (Lille, France). NGS libraries were prepared using the TruSeq Nano DNA Library Preparation kit and then sequenced on an Illumina MiSeq 2x300 bp platform.

Sequences were analysed using QDD v3 software (Meglécz et al., 2014) that treats all steps from raw sequences until obtaining PCR primers: removing adapters and vectors, detection of microsatellites, redundancy and possible mobile element association, selection of sequences with target microsatellites and primer design using BLAST (ftp://ftp.ncbi.nih.gov/blast/executables), Clustal W (Larkin et al., 2007), and Primer3 (Rozen and Skaletsky, 2000) programs.

Microsatellite validation, multiplex design and genotyping

Sequences suitable for primers design were ranked according to motif type (tetra-, tri-, and di-nucleotide repeats), number of repeats (the higher the better), and PCR product size (>100 bp) considering only sequences with perfect repeats. From this list, the 47 top ranking primers were selected for the identification of suitable microsatellites (i.e. consistent amplification and ease to score) from eight individuals of different population origin. PCR reactions were performed in a final volume of 10 µl containing 5 ng of DNA, 0.5 U of Taq polymerase, 6 pmol of dNTPs, 37.5 pmol of $MgCl_2$, and 10 pmol of each primer. PCR cycling profile consisted of a 10 min initial denaturation at 95 ^{o}C; followed by 40 cycles of denaturation for 30 s at 95 ^{o}C, annealing for 30 s at 55 ^{o}C, and extension for 1 min at 72 ^{o}C; and a final extension step for 10 min at 72 ^{o}C. One negative control was included in all PCR reactions. One µl from each PCR product was analysed on QIAxcel (Qiagen) to check amplification efficiency and specificity. Of the markers that amplified

successfully from eight individuals, 22 were selected for initial polymorphism tests in 15 individuals. In this step, PCRs were performed as described above, but using forward primers labelled with a fluorescent dye (6FAM, VIC, NED or PET; Applied Biosystems, Foster City, California, USA). One µl of PCR product was migrated on ABI 3730XL capillary sequencer (Applied Biosystems) using GeneScan™-500 LIZ® (Applied Biosystems) as size standard. Multiplex Manager v1.1 program (Holleley and Geerts, 2009) was employed to combine the polymorphic microsatellite loci into multiplex PCR reactions. This software uses prior marker information to minimize the total number of reactions, maximize the spacing between markers, minimize the variance of annealing temperature in each reaction, and mainly to minimise complementarity between markers in the same reaction. Each set of multiplexed loci was tested on the same DNA templates as used for the initial testing. PCR products were migrated on the capillary sequencer and genotyped using the program GeneMapper v5.0 software (Applied Biosystems).

Data analysis: genetic diversity and population genetic structure

Allele frequencies, observed number of alleles per locus (N_A), observed heterozygosity (H_o), and unbiased expected heterozygosity (H_e) of Nei (1978) were calculated with Genetix v.4.03 software (Belkhir et al., 2004). Allelic richness (*R*s) per locus and per locus-locality combination were computed with Fstat v.2.9.3.2 (Goudet, 2001). Friedman tests were carried out to compare heterozygosity values and allelic richness between localities with the statistical package SPSS 16.0 (SPSS Inc.).

Deviations from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium between pairs of loci within each locality were tested with Genepop v.4.0 (Rousset, 2008). This program was also used to calculate the inbreeding coefficient (F_{IS}) following Weir and Cockerham (1984). The significance was determined by a Markov chain method using 10,000 dememorizations, 5,000 batches and 5,000 iterations per batch. FreeNa software (Chapuis and Estoup, 2007) was used to estimate the null allele frequencies with a number of replicates fixed to 10,000. Whenever multiple

tests were performed, *P* values were adjusted using the sequential Bonferroni correction (Rice, 1989).

Both the Ewens-Watterson homozygosity test (Watterson, 1978, 1986) and the Ewens-Watterson-Slatkin exact test (Slatkin, 1994, 1996) were carried out to check for departures from selective neutrality using Arlequin v.3.11 (Excoffier and Lischer, 2010).

Population differentiation was assessed using Weir and Cockerham´s *F* statistics (1984). In order to estimate the extent of bias possibly introduced by the presence of null alleles, these statistics were also computed with FreeNa (Chapuis and Estoup, 2007) with and without null alleles correction (ENA). This correction method is used to efficiently correct for the positive bias induced by the presence of null alleles on F_{ST} estimation. Ninety-five per cent confidence intervals for the F_{ST} values were obtained using 10,000 iterations. F_{ST} matrices were visualised in a non-metric Multidimensional Scaling (nMDS) plot constructed in SPSS 16.0 (SPSS Incl.). Based on the F_{ST} values, the gene flow between localities was also evaluated estimating the number of migrants per generation (*Nm*) between pairs of localities with the software GenAlEx (Peakall and Smouse, 2006).

A Bayesian clustering analysis using Structure v.2.3.1 (Pritchard et al., 2000) was performed with and without recessive alleles option. In the latter case, the program assumes that the recessive allele is never observed in homozygous state, but it might be present (e. g. when there might be null alleles). The range of possible clusters tested (K) was set from 1 to the total number of sampling sites included in the analysis. Ten independent runs were carried out for each K, using prior location information and assuming a non-admixture model with correlated allele frequencies. The lengths of the Markov chain Monte Carlo and burn-in were set at 250,000 and 100,000, respectively. Results were processed using the program Structure Harvester (Earl and von Holdt, 2012).

Population structure was examined by locus-by-locus AMOVA using the program Arlequin v.3.5.1.2 (Excoffier and Lischer, 2010).

Results and discussion

Development of novel microsatellite loci in *D. vittatus* and multiplex PCRs

In this study, we use NGS data to provide the first panel of microsatellite loci for *D. vittatus*. Microsatellite isolation through NGS approach has great advantages over traditional methods based mainly on the construction of enriched genomic libraries (Zane et al., 2002) in terms of cost effectiveness, production of a large number of microsatellites and time efficiency (Abdelkrim et al., 2009; Castoe et al., 2010). This method has allowed isolating microsatellite loci in other species within the genus *Donax* (*Donax deltoides,* Miller et al., 2012; *D. trunculus,* Rico et al., 2017), and in other bivalve molluscs such as *Aequipecten tehuelchus* (Domínguez-Contreras et al., 2017), *Anodonta anatina* (Lopes-Lima et al., 2016), *Bathymodiolus manusensis* (Schultz et al., 2011), *Lutraria rhynchaena* (Thai et al., 2016), *Mytilus coruscus* (An and Lee, 2012; Kang et al., 2013), *Perna viridis* (Lukehurst et al., 2017) and *Potomida littoralis* (Froufe et al., 2013).

In the present work, a total of 985,748 sequences were examined from which 1647 primer pairs were design on microsatellite flanking regions, and a subset of 177 were selected as the most promising according to QDD parameters. Finally, 15 polymorphic markers with well-scorable peaks were selected. Thirteen markers were arranged into three multiplex PCRs (Multiplex-1, Multiplex-2 and Multiplex-3) with four, six and three microsatellites, respectively; and the two remaining loci were amplified by simplex PCRs (Table 2). The use of multiplex PCR reduces time and associated costs and offers a simple yet powerful technique that can increase the productivity of genetic studies (Neff et al., 2000). The sequences containing the 15 polymorphic microsatellite loci have been deposited in the GenBank database of the NCBI (Table 2).

Table 2. Characteristics of 15 newly polymorphic microsatellite loci identified in the wedge clam *D. vittatus*.

Locus name	Repeat motif	Forward primer sequence (5′-3′)	Dye	Reverse primer sequence (5′-3′)	Primer pair concentration (pmol)	Accession numbers
Multiplex-1						
Dvit13	(CT)16	ACACCAGATAAACAGAAGAGATTAAC	PET	AAGTATATTGATTACACTTTGGAGACC	8	MH215262
Dvit26	(TTG)11	ATGCTAAACATCGCTCACCC	6FAM	TCTGGAGCTTTCGACCTTGT	7	MH215264
Dvit31	(GTT)10	TGCCCAGCAGACAATTGCAA	VIC	CCAGCCATATGAGCACCTCT	10	MH215267
Dvit47	(AC)9	TGCCGCTCTATCATCACAGA	VIC	GGGAGGGTGTATGGATGAGG	2.5	MH215273
Multiplex-2						
Dvit12	(AT)16	GCGGATTAGCTCTCGTGGTT	6FAM	CTGTGTTACATTCAGTAGCTAGCA	4.5	MH215261
Dvit33	(TAA)10	GAGCCCTCTGATCTGGGTCT	PET	ACAGCCTCCTATGTGATTTCTCG	3	MH215268
Dvit35	(TGT)9	AGGAATGGTGGCGCTTGTTA	VIC	AGCCTCCGTCACTTCTTGAA	3	MH215269
Dvit37	(GAT)9	GAGGCATGTACAGCAAAGGC	PET	ACCGGATATTTATAATCTTTCCATGGC	2.2	MH215270
Dvit40	(CAT)9	GCTCAGTCAGTAACGCCCAG	VIC	TATTCAGCTGCGCATGGACA	2.5	MH215271
Dvit46	(TGA)9	CGGTGTTTATGTTAAACATGCAAT	6FAM	TCACCGAAAGCAAGTACACT	2.5	MH215272
Multiplex-3						
Dvit07	(ATCT)21	TGTACATGTTTGCTATTAATTCCATGT	NED	TGTGATGTCAGAAGCTTGTGTT	6	MH215259
Dvit24	(GAT)12	TGCTGATGATGTCGATGACG	VIC	AATCCACGCCTGGCTTCAAA	3.5	MH215263
Dvit27	(GATA)11	GCTGTACAGCGCATTTGAACA	6FAM	TCAGTTCCCTCGCAATCACC	6	MH215265
Simplex-1						
Dvit09	(AG)20	ATGTCAACCTAATGTTATTGTGGTT	NED	CTTCCACGCCTTTATCTCCCA	5	MH215260
Simplex-2						
Dvit29	(TGT)10	CCCAATTCAATCTTCCTTAACTCTG	6FAM	GAAATGACATACTGCAACACCGT	5	MH215266

Genetic diversity

A total of 367 individuals from seven different localities were genotyped for the 15 developed microsatellite loci. Genetic variation statistics by locus, locality and overall are available as Supplementary Material (see Supporting Information on CD-ROM: S1 Table).

All loci analysed were polymorphic in all localities with the number of alleles per locus ranging from 8 (Dvit33) to 65 (Dvit07). Allelic richness across localities per locus varied between 3.262 (Dvit33) and 21.181 (Dvit09) and was lower than the values reported for *D. trunculus* (4.714-35.905, Nantón et al., 2017). Expected and observed heterozygosity values per locus ranged from 0.446 (Dvit33) to 0.960 (Dvit09) and from 0.150 (Dvit24) to 0.759 (Dvit37), respectively. These values are in the range obtained for other *Donax* species, *D. trunculus* (H_e: 0.557-0.947 and H_o: 0.207-0.738, Marie et al., 2016; H_e: 0.082-0.903 and H_o: 0.033-0.758, Nantón et al., 2017) and *D. deltoides* (H_e: 0.189-0.860 and H_o: 0.207-0.857, Miller et al., 2012).

Regarding localities, all showed private alleles at the 15 loci examined, ranging from 1 to 25. They were always at frequencies below <6%. The mean number of alleles and the allelic richness across loci per locality varied between 13.867 and 15.733, and 10.750 and 15.201, respectively. In both cases, the lowest value was observed in MV and the highest one in SE. All localities showed similar mean values of expected and observed heterozygosity, ranging from 0.813 (MV) to 0.845 (LI) and from 0.389 (LI) to 0.460 (LZ), respectively. Significant differences in allelic richness (P = 0.070), expected heterozygosity (P = 0.070) and observed heterozygosity (P = 0.346) were not detected using the Friedman test. These results are in line with the previous study for *D. vittatus* (Fernández-Pérez et al., 2017a) where the localities showed similar diversity at several mtDNA markers (COI, Cytb and 16S) and other nuclear repetitive sequences (H3, 18S and 28S genes).

Globally, the mean number of microsatellite alleles in *D. vittatus* (30.130) was superior to that obtained for *D. trunculus* (22.750, Marie et al., 2016; 15.067, Nantón et al., 2017; 17.800, Rico et al., 2017) and other marine bivalves such as *C. edule* (27.250, Martínez et al., 2015), *O. edulis* (18.533, Launey et al., 2002), and *R. decussatus* (11.83, Borrell et al., 2014), but allelic richness

(12.250) was lower than the value reported for *D. trunculus* (14.664, Nantón et al., 2017) or *C. edule* (27.030, Martínez et al., 2015). While expected heterozygosity value obtained for *D. vittatus* (0.832) was similar or higher than those values observed for *D. trunculus* (0.806, Marie et al., 2016; 0.591, Nantón et al., 2017; 0.843, Rico et al., 2017) and other bivalve species from the same studied region (0.746 for *C. edule*, Martínez et al., 2015; 0.663 for *R. decussatus*, Borrell et al., 2014), observed heterozygosity value (0.422) was lower (0.493 for *D. trunculus*, Marie et al., 2016; 0.434 for *D. trunculus*, Nantón et al., 2017; 0.572 for *D. trunculus*, Rico et al., 2017; 0.665 for *C. edule*, Martínez et al., 2015; 0.638 for *R. decussatus*, Borrell et al., 2014). In line with mitochondrial markers results, microsatellites revealed moderate levels of genetic diversity in *D. vittatus*.

Linkage disequilibrium tests performed for all pair of loci across localities did not display significant values after sequential Bonferroni correction for the 735 comparisons analysed, except for comparison Dvit24-Dvit33 in LI. Seventy-nine out of the 105 locality-locus combinations showed no significant deviations from HWE after sequential Bonferroni correction. Three out of 15 loci conformed to HWE in almost all localities (Dvit13, Dvit37 and Dvit47), while ten (Dvit07, Dvit09, Dvit12, Dvit24, Dvit26, Dvit27, Dvit29, Dvit31, Dvit33 and Dvit40) showed significant departures from HWE in all localities. All combinations departing from HWE expectations showed positive F_{IS} values ($\geq$0.056), indicating the existence of a heterozygote deficit.

Several biological and technical factors could account for the observed heterozygote deficits, such as inbreeding, Wahlund effect, selection and the presence of null alleles. Although, all localities examined in this study showed heterozygote deficit in most of the analysed loci, pairwise F_{ST} values did not show evidences of genetic differentiation (see below). So, it is highly unlikely that there is genetic structure in each locality. Therefore, the Wahlund effect does not seem to be the cause of the heterozygote deficit. In species with separate sexes, partial inbreeding can occur through the mating of related individuals and produce a deficit of heterozygotes. Nevertheless, given that most bivalves present external fertilization and extended larval dispersal phase, it is unlikely that this factor is the main cause of the observed heterozygous deficit. Moreover, inbreeding should affect all loci equally, generating uniform heterozygote deficiencies across loci (Astanei et al., 2005).

However, in this study positive F_{IS} values (see Supporting Information on CD-ROM: S1 Table) were heterogeneous among loci. Microsatellites are typically considered as neutral markers but some can be directly or indirectly under the effect of selection (Li et al., 2002). In this study, the Ewens-Watterson homozygosity test and the Ewens-Watterson-Slatkin exact test reported deviations from neutrality at seven loci in some localities (see Supporting Information on CD-ROM: S1 Table). Although the values of these tests indicate the possibility that the seven loci are subject to balancing selection, this would only explain part of the observed deficit for these loci. Therefore, selection does not seem to be the main cause of the global observed deficit of heterozygotes in this work. The most likely cause for the heterozygote deficiencies detected here seems to be null allele occurrence whose allele frequencies estimates using FreeNa ranged between 0 and 0.440 (see Supporting Information on CD-ROM: S1 Table). Moreover, 69 out of the 79 locality-locus combinations departing HWE showed evidence of the presence of null alleles in ten loci (Dvit07, Dvit09, Dvit12, Dvit24, Dvit26, Dvit27, Dvit29, Dvit31, Dvit33 and Dvit40) with frequencies >0.2. Null alleles are frequent in many taxonomic groups (Dakin and Avise, 2004) but seem to be particularly common in populations with high effective population sizes (Chapuis and Estoup, 2007), as is the case of molluscs. They have been reported in up to 50% of the examined loci in several bivalve species (Launey et al., 2002; Li et al., 2003; Hedgecock et al., 2004; Carlsson and Reece, 2007; Galindo-Sánchez et al., 2008; Gardeström et al., 2008; Kim et al., 2014; Hargrove et al., 2015; Chiesa et al., 2016; Marie et al., 2016; Nantón et al., 2017; Rico et al., 2017) explaining the heterozygous deficit reported as it occurs in this work.

Population genetic structure

To estimate the impact of null alleles, F_{ST} statistic was estimated with and without ENA correction. In both cases, global multilocus F_{ST} had a value of 0.003 and it was significant, with confidence intervals (95%, 10,000 replicates) not including zero. Estimates of F_{ST} per locus ranged from -0.008 (Dvit33) to 0.024 (Dvit09) when were calculated without ENA correction and ranged from -0.003 (Dvit33) to 0.019 (Dvit09) when the correction was used (Table 3).

Table 3. F_{ST} values per microsatellite marker without and with applying the ENA correction for all individual from the 7 localities of *D. vittatus*.

Locus	F_{ST} not corrected for null alleles	F_{ST} corrected for null alleles
Dvit07	0.001	0.001
Dvit09	0.024	0.019
Dvit12	0.001	0.001
Dvit13	0.002	0.002
Dvit24	0.008	0.013
Dvit26	-0.001	-0.001
Dvit27	0.000	0.002
Dvit29	0.011	0.008
Dvit31	0.002	0.001
Dvit33	-0.008	-0.003
Dvit35	0.001	0.002
Dvit37	0.004	0.004
Dvit40	-0.003	0.001
Dvit46	-0.001	0.000
Dvit47	0.000	0.001

Pairwise F_{ST} values were also calculated with and without ENA correction to determinate whether their presence created a substantial bias (Table 4). Even though there are some differences between the corrected and uncorrected estimates of genetic differentiation, the significant F_{ST} estimates were concentrated in the pairwise combinations CO-SI, CO-SE and MV-PA. Nevertheless, only CO-SI, out of the 21 pairwise combinations, was significant using null alleles correction (Table 4). Overall, these estimates indicated that localities from the Atlantic coast are genetically homogeneous

although some samples showed significant pairwise F_{ST} values. Visualisation of pairwise F_{ST}, with and without ENA correction, using nMDS did not show a clear clustering (Figure 2). In addition, in the spatial Bayesian analysis all individuals of seven localities were grouped in one single clade, suggesting lack of genetic structure. LnPr(X|K) values increase slightly as the value of K (number of clusters) increases, without there being a clear jump between one and other values (Figure 3A). Moreover, although related to the above, ΔK value, which can be used as an indicator of the strength of the structure, is small (Figure 3B). This result suggests that there are no large differences between the values of K, contrary to what would be expected if there were a true K (Pritchard et al., 2010). The estimates of *Nm* (Table 5) were well above one migrant per generation, a value that is considered sufficient to cushion genetic differentiation among populations caused by drift effect (Slatkin, 1987). The seven localities showed similar values, suggesting that the larval exchange could be responsible for their genetic homogeneity. Although *Nm* is a useful relative measure of migration among populations it should be interpreted with caution, since it is calculated based on the F_{ST} value and this is usually estimated from a small number of loci in limited samples of individuals from a few populations. Furthermore, F_{ST} is a non-linear function of *Nm* and, therefore, small differences in F_{ST} can give rise to important variations in estimates of *Nm* (Whitlock and McCauley, 1999).

Table 4. Pairwise F_{ST} values for seven localities and 15 microsatellite markers in *D. vittatus* computed both using (below diagonal) and without using (above diagonal) correction for null alleles.

	CO	LZ	PA	MV	LI	SE	SI
CO		0.003	0.003	0.003	0.003	0.004	0.008*
LZ	0.003		-0.001	0.004	-0.003	-0.002	0.004
PA	0.005	0.000		0.005	0.001	-0.001	0.004
MV	0.002	0.002	0.006*		0.005	0.003	0.008
LI	0.002	0.002	0.003	0.003		0.001	0.006
SE	0.004*	-0.001	0.001	0.003	0.002		0.002
SI	0.008*	0.003	0.004	0.006	0.005	0.002	

* 95% confidence interval does not include zero.

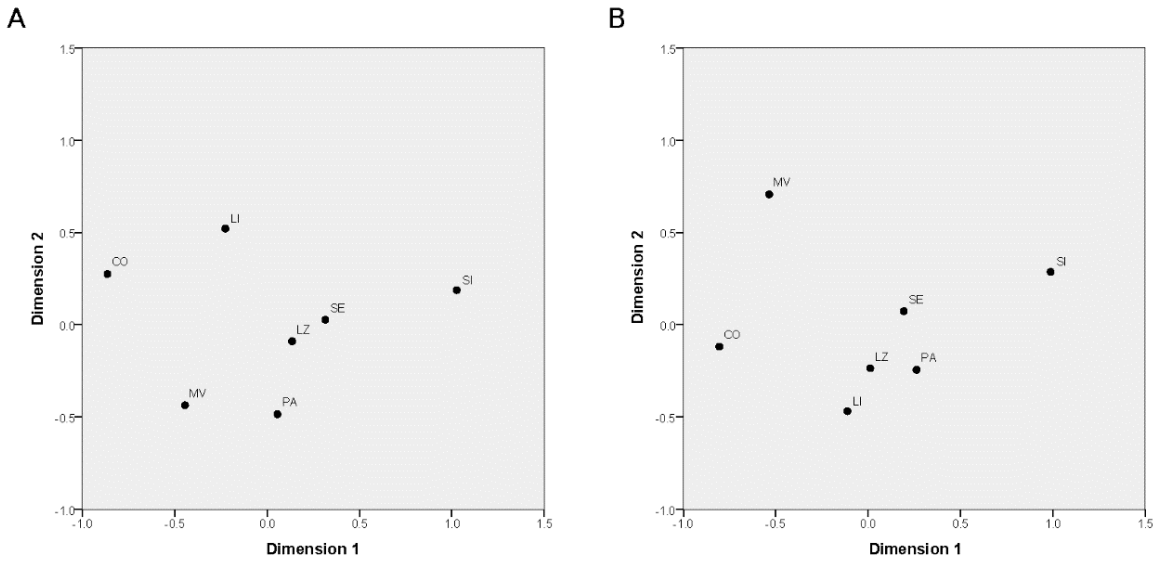


Figure 2. nMDS plots of pairwise F_{ST} values of *D. vittatus* in the Atlantic coast of the Iberian Peninsula without (A) and with (B) applying the ENA correction.

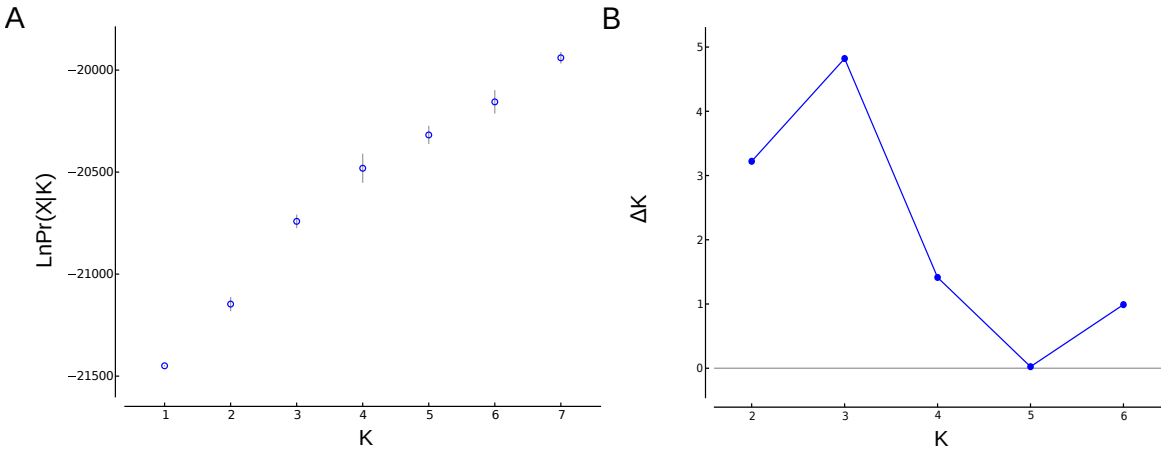


Figure 3. Distribution of the estimated likelihood of K (A) and ΔK (B). In the likelihood graph (A), the mean of the Ln of the estimated probability of the data is represented for each K value [LnPr (X|K)] using 20 runs. In the ΔK graph (B) the values of K = 1 and K = 7 are not represented as ΔK is computed as the difference between consecutive values of K.

Table 5. Estimates of *Nm* between pairs of localities.

	CO	LZ	PA	MV	LI	SE	SI
CO		24.465	20.592	22.610	21.986	20.600	18.094
LZ			30.246	24.046	29.021	32.503	23.140
PA				18.276	29.021	27.210	18.599
MV					19.553	21.295	18.229
LI						23.847	18.978
SE							23.692
SI							

Following the F_{ST} pairwise results (Table 4), nMDS plots (Figure 2), Bayesian clustering, and the estimates of *Nm* between pairs of localities (Table 5), samples were pooled into one group and the results of AMOVA revealed a small but significant among populations component, explaining 0.99% of the total of variance. However, most of the variation corresponds to differences within localities (99.01%).

The analyses performed with microsatellite markers did not show genetic structure in *D. vittatus* along the Atlantic coast of Iberian Peninsula. Microsatellites are expected to be able to reveal a finer resolution of genetic structuring than mtDNA, not only because of their sensitivity to gene flow (Ross et al., 1999) but also because their high variability gives them greater statistical power (Estoup et al., 1998). This lack of differentiation among studied populations is congruent with the previously published study based on the COI gene (Fernández-Pérez et al., 2017a). A similar pattern of homogeneity in the same geographic region has been observed in other marine invertebrates (Sanjuan et al., 1996; Duran et al., 2004; Triantafyllidis et al., 2005; Couceiro et al., 2007; Xavier et al., 2012), including the congeneric species *D. trunculus* (Marie et al., 2016; Nantón et al., 2017; Fernández-Pérez et al., 2018). In addition, other population genetics studies with a wider geographic range do not report genetic structure for samples in an area that goes from Africa or the Iberian Peninsula to the British Isles (Launey et al., 2002; Roman and Palumbi 2004; Darling et al., 2008; Krakau et al., 2012; Martínez et al., 2015; Morvezen et al., 2016).

The genetic homogeneity detected in this region may be the result of both ocean currents and demographic processes. The prevailing current patterns during the reproductive season could be responsible for the observed genetic homogeneity because it is during the larval planktonic phase when *D. vittatus* shows a greater dispersal ability. Similar explanations have been suggested for other species of marine organisms that inhabit this area (Ribeiro et al., 2010; Xavier et al., 2012; Fernández-Pérez et al., 2018).

Conclusion

In this study, NGS-based approach was used to provide the first microsatellite markers for *D. vittatus*. Fifteen loci were analysed to estimate genetic diversity and population differentiation along the Atlantic coast of the Iberian Peninsula. Results revealed moderate levels of genetic diversity without population differentiation along the studied area, supporting the previous results based on COI gene (Fernández-Pérez et al., 2017a). These data are useful to assist the fishery development and the conservation plans of this species. Furthermore, the panel of microsatellites will allow to undertake further genetic population studies throughout the geographic distribution of the species, as well as other genetic analysis needed for aquaculture activities.

Acknowledgements

We would like to thank Dr. M. B. Gaspar from Instituto Português do Mar e da Atmosfera – IPMA (Portugal) and Dras. D. Martínez-Patiño and S. Nóvoa from Centro de Cultivos Mariños de Ribadeo – CIMA (Xunta de Galicia) from Xefatura de Coordinación da Área do Mar de Vigo (Xunta de Galicia) for providing *D. vittatus* specimens. We also thank Dra. E. Froufe from Interdisciplinary Centre of Marine and Environmental Research – CIIMAR (Portugal) for her technical advice. We would also like acknowledge the assistance of staff at Genoscreen. This work was supported by the Ministerio de Economía y Competitividad (Spain) through project AGL2016-75288-R AEI/FEDER, UE.

Chapter 4

Sequence characterization of the 5S ribosomal DNA and the internal transcribed spacer (ITS) region in four European *Donax* species (Bivalvia: Donacidae)

Jenyfer Fernández-Pérez[1], Ana Nantón[1], Josefina Méndez[1]

[1] Grupo Xenomar, Departamento de Bioloxía, Facultade de Ciencias and Centro de Investigaciones Científicas Avanzadas (CICA), Universidade da Coruña, Campus de A Zapateira, 15071 A Coruña, Spain

(In Preparation)

Abstract

The whole repeat unit of 5S rDNA and the internal transcribed spacer (ITS) of the four European *Donax* species were analysed. After amplifying, cloning and sequencing several 5S and ITS units, their basic characteristics and their variation were reported. The length of the 5S repeat presented little variation among species, except *Donax trunculus* that differed from the rest of the *Donax* species in 170-210 bp. The deduced coding region covered 120 bp and showed recognizable internal control regions (ICRs) involved in the transcription. The length of non-transcribed spacer region (NTS) ranged from 157 bp to 165 bp in *D. trunculus* and from 335 bp to 367 bp in the other three species. The conservation degree of transcriptional regulatory regions was analysed revealing a conserved TATA-like box in the upstream region. Regarding ITS sequences, the four *Donax* species showed slight size differences among clones due to the variation occurring in the ITS1 and ITS2, except *Donax variegatus* did not display size differences in the ITS2. The total length of the ITS sequence ranged between 814 and 1014 bp. This is not only a basic research work, where new data and new knowledge is provided for the scientific community about *Donax* species, of which there is not much genetic information, but also offers a future and feasible application in aquaculture.

Introduction

In higher eukaryotes, rDNA comprises two different multigene families (Long and Dawid, 1980), including the major 45S rDNA family encoding 18S, 5.8S, and 28S rRNA, and the minor 5S rDNA family encoding 5S rRNA, each composed of hundreds to thousands of copies, organized in tandem repeats, and consisting of coding regions and transcribed and non-transcribed spacers. The different evolutionary rates among different regions, the secondary structural of these genes and their organization in tandem repeats, make rDNA attractive candidate for species identification, population characterization, phylogenetic studies and evolutionary relationships and genomic structure analysis (Jansen et al., 2006; Hurtado et al., 2011, Pinhal et al., 2011).

The 5S rDNA consists of one coding region of about 120 bp (5S rRNA gene), which is separated from the next unit by a non-transcribed spacer (NTS). The coding sequence and the NTS together form a repeat unit that can found in hundreds to thousands of copies tandemly repeated in the genomes. Even though the 5S rRNA gene is highly conserved, even among unrelated species, NTSs are variable, both in length and sequence (Rebordinos et al., 2013). These variations have been used for evolutionary studies and as species-specific or population-specific markers in several bivalve mollusc species (Freire et al., 2005; Fernández-Tajes and Méndez., 2007; López-Piñón et al., 2008; Pereira et al., 2012), so that the 5S rDNA is a good candidate to identify molecular markers suitable for distinguish related species. On the other hand, both the 5S rRNA gene and NTS have also been examined for phylogenetic analyses. The phylogenetic value of the 5S rRNA gene is limited, since it provides few characters owing to its small size and because the variable position shows high rate of change (Halanych, 1991; Steele et al., 1991). It has too few informative sites to allow for analysis of close relatives, and it evolves too rapidly for reliable comparisons among distantly related taxa. Nevertheless, spacer sequences have the potential to resolve phylogenetic relationships at low taxonomic levels (Suzuki et al., 1994; Ferreira et al., 2006; Infante et al., 2007).

By the same token, the internal transcribed spacer (ITS) region of rDNA consists of one coding region (5.8S rRNA gene) and two non-coding regions (ITS1 and ITS2) located in the rDNA between 18S and 5.8S rRNA genes and between 5.8S and 28S rRNA genes, respectively. Due to ITS sequences show more variation than their flanking coding region (Hillis and Dixon, 1991), they have regularly been used to infer phylogenetic relationships from populations to families, even higher taxonomic levels (Coleman and Vacquier, 2002; Insua et al., 2003; Vierna et al., 2010), and to differentiate related bivalve species (López-Piñón et al., 2002; Santaclara et al., 2006; Hurtado et al., 2011). ITS1 has been a widely chosen marker for assessing variation within species due to its high level of divergence, while ITS2 region has been proposed as an effective barcode similar to the cytochrome c oxidase subunit I (COI) for identifying species that are difficult to distinguish morphologically (Yao et al., 2010).

To date, several reports on the characterization of the 5S rDNA and the ITS region in various mollusc species, including bivalves, gastropods, and cephalopods have been published (Insua et al., 2001; Manchado et al., 2006a; Vizoso et al., 2011). Bivalve molluscs stand out for being one of the most widely studied group of organisms regarding 5S rDNA and the ITS region, showing high levels of gene organization as well as a vast diversity of gene arrangements. Molecular organization of the ITS region and 5S rDNA has been studied in cockles (Insua et al., 1999; Freire et al., 2005, 2010a,b), mussels (Insua et al., 2001), oysters (Cross and Rebordinos, 2005, 2006), scallops (Insua et al., 1998, 2003, 2006; Huang et al., 2007; López-Piñón et al., 2008), razor clams (Fernández-Tajes and Méndez, 2009; Vierna et al., 2009, 2010) and Veneroida clams (Cheng et al., 2006), but it has never been studied in *Donax* species, of which there is not much genetic information available.

Four *Donax* species, *Donax semistriatus*, *Donax trunculus*, *Donax variegatus* and *Donax vittatus*, are common along the European littoral and live sympatrically in some areas (Gaspar et al., 2002; Bejega et al., 2010; Rufino et al., 2010; Fernández-Pérez et al., 2017a). These marine bivalves play an important socioeconomic role in some European coastal regions. For instance, the wedge clam *D. trunculus* is one of the most exploited and economically important bivalve species in several European countries, including France (Thébaud et al., 2005), Italy (Zeichen et al., 2002), Portugal (Chícharo et al., 2002), Spain (Molares et al., 2008) and Turkey (Özden et al., 2009), and may account for most of the recorded catches of FAO in these countries for genus *Donax*. Nevertheless, FAO statistics (FAO-FIGIS, 2017) do not distinguish between species. Therefore, it is likely to find in the fish market other *Donax* species being sold as *D. trunculus*. However, despite the economic importance that genus *Donax* has for the European seafood sector and of being an overexploited species, genetic studies in this organism are very scarce.

In this work, the 5S rDNA and the ITS region of four wedge clams, *D. semistriatus*, *D. trunculus*, *D. variegatus* and *D. vittatus*, of the bivalve family Donacidae present in Europe were analysed. The aims of this analysis were to amplify, clone, and sequence the 5S rDNA and ITS repeats units to provide their basic characteristics, to assess their variability, and to estimate their divergence.

Material and methods

Sampling and DNA extraction

Donax trunculus specimens were collected from natural beds in Vilarrube (northwestern Spain) while *D. semistriatus*, *D. variegatus* and *D. vittatus* samples came from Portuguese coast (Table 1). All clams handling was conducted in accordance with the guidelines and regulations established by the University of A Coruña and the local governments. Specimens were taxonomically identified using a species-specific PCR-RFLP analysis of COI capable discriminating among the four *Donax* species (Nantón et al., 2015). Total genomic DNA was extracted from ethanol-preserved foot using a Chelex-100 (Sigma-Aldrich, USA) protocol based on Walsh et al. (1991).

PCR amplification, cloning and sequencing

For 5S rDNA, amplification reactions were carried out using a set of primers designed by Fernández-Tajes and Méndez (2007) (5SF: 5`-CGTCCGATCACCGAAGTTAA-3´ and 5SR: 5`-ACCGGTGTTTTCAACGTCAT-3´) annealing to the coding region in opposite orientations. They were performed in 25 µl containing 150 ng of genomic DNA, 0.25 µM of each dNTP, 2 mM of $MgCl_2$, 0.6 µM of each primer, 0.6 U of *Taq* polymerase (Roche Applied Science) and the buffer recommended by the polymerase suppliers. PCR cycling protocol consisted of an initial denaturation step of 2 min at 95 °C, 35 cycles with a denaturation step at 95 °C for 30 s, an annealing step at 55 °C for 30 s and an extension step for 1 min at 72 °C, and a final extension step at 72 °C for 5 min.

For ITS1 and ITS2, PCR reactions were performed with a pair of primers that anneal at the 3´end of the 18S ribosomal gene and the 5´end of the 28S ribosomal gene (Heath et al., 1995) (ITSF: 5´-GTTTCCGTAGGTGAACCTG-3´ and ITSR: 5´-CTCGTCTGATCTGAGGTCG-3´). They were carried out in 25 µl containing 100 ng of genomic DNA, 0.25 µM of each dNTP, 1.5 mM of $MgCl_2$, 1 µM of each primer, 0.625 U of *Taq* polymerase (Roche Applied Science) and the buffer recommended by the polymerase suppliers. PCR cycling protocol consisted of an initial

denaturation step of 3 min at 94 °C, 30 cycles with a denaturation step at 94 °C for 20 s, an annealing step at 55 °C for 20 s and an extension step for 45 s at 72 °C, and a final extension step at 72 °C for 5 min.

PCR products were migrated on 2% agarose gel electrophoresis. The gels were stained by immersion in 0.5 µg/ml ethidium bromide solution for 30 min, visualized and recorded on a transilluminator Gel Doc XR Systems (Bio-Rad, Barcelona, Spain).

For three or four individuals of each species, the product obtained was ligated into the T&A™ cloning vector and transformed into *Escherichia coli* ECOS™ JM109 strain competent cells using T&A™ Cloning Vector Kit (Yeastern Biotech Co., Ltd). Recombinant colonies were screened by PCR amplifying with M13 forward and reverse primers to assess the size of the insert. PCR reaction mixture contained 5 µl of recombinant cells, 1x PCR buffer, 1.5 mM $MgCl_2$, 0.2 µM of each dNTP, 0.6 µM of each primer, and 0.3 U of *Taq* polymerase (Roche Applied Science) in a final volume of 12.5 µl. The thermal cycle profile consisted of an initial denaturation of 10 min at 94 °C, 30 cycles of 1 min at 94 °C, 1 min at 55 °C and 1 min at 72 °C; and a final extension of 10 min at 72 °C. Several recombinant colonies (3–10 per individual) were selected at random and grown in LB medium. A QIAprep Spin Miniprep Kit (Qiagen) was used to purify the plasmids, which were sequenced using both M13 forward and M13 reverse primers on an ABI PRISM 3120xl (Applied Biosystems, Foster City, CA, USA) at the Molecular Biology Unit of the University of A Coruña (Spain). The corresponding nucleotide sequences have been deposited in the GenBank database under accession numbers MG041608 – MG041761 (Table 1).

Table 1. Sampling details of species and GB Accession numbers.

Species	Locality	Country	Coordinates		GB Accession numbers	
			Latitude	Longitude	5S rDNA	ITS
D. semistriatus	Monte Gordo	Portugal	37.167	-7.503	MG041608 - MG041634	MG041692 - MG041713
D. trunculus	Vilarrube	España	43.644	-8.077	MG041635 - MG041654	MG041714 - MG041736
D. variegatus	Monte Gordo	Portugal	37.100	-7.633	MG041655 - MG041676	MG041737 - MG041749
D. vittatus	Mira-Vagueira	Portugal	40.614	-8.769	MG041677 - MG041691	MG041750 - MG041761

Sequence analysis

The identity of sequences obtained was corroborated using BLASTn searches of the NCBI database (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Sequence data were aligned via MAFFT (Katoh and Toh, 2008) using the L-INS-i algorithm (recommended for <200 sequences with one conserved domain and long gaps) and manually checked using the BioEdit v.7.2.5 sequence editor (Hall, 1999). The number of variable sites, nucleotide diversity and sequence divergence were estimated using DnaSP v5.10.01 (Librado and Rozas, 2009). Differences between sequence pairs and *Donax* consensus sequences were calculated using Geneious Pro v.4.8.5 (Drummond et al., 2009).

Results

5S rDNA

The 5S rDNA repeat unit was PCR amplified in at least 20 individuals of each wedge clam species, except for *D. semistriatus* and *D. variegatus* of which the number of available individuals was low (12 and 8, respectively). The length of the 5S units was about 275–300 bp for *D. trunculus*, about 450 bp for *D. variegatus* and around 500 bp for fragments obtained from the other two *Donax* species with minimal variation (1-12 bp) among clones (Table 2). Taking as reference the 5S rRNA sequences available in other bivalve species (Fang et al., 1982; Insua et al., 2001; Freire et al., 2005; Cross and Rebordinos, 2005; López-Piñón et al., 2008; Fernández-Tajes and Méndez, 2009), the coding region was assigned to 120 bp in the four cases and the non-transcribed spacer (NTS) region to the remaining sequence (Table 2). BLAST analysis corroborated the identity of the limited region and indicated that no other coding sequence was included in the 5S rDNA repeat unit. The GC content of the repeat units ranged from 38.2% to 43.3% among the wedge clams (Table 2), with higher values in the coding region (53.4-55%) than in the spacer region (33.1-39.1%).

Table 2. Size (bp) and mean value of the GC content (%) of the 5S rDNA repeat unit.

Species	No. of clones	Repeat unit		Coding region		NTS	
		Length	GC	Length	GC	Length	GC
D. semistriatus	27	475–487	40.5	120	54.6	355–367	35.9
D. trunculus	20	277–285	41.5	120	54.3	157–165	32.3
D. variegatus	22	455–456	43.3	120	55.0	335–336	39.1
D. vittatus	15	480-481	38.2	120	53.4	360-361	33.1

The alignment of all the 5S rDNA wedge clam sequences consisted of 568 pb and showed 130 variable sites of which 108 were parsimony informative, and 342 indels (due mainly to the fact that the sequence of *D. trunculus* is smaller than for the rest of the species). Almost entirely the variation was located in the spacer region; the sequence corresponding to gene showed 38 variable sites (see Supporting Information on CD-ROM: S1 File). Intraindividual variation was minimal within control region and moderate within the spacer region in *D. trunculus*, *D. variegatus* and *D. vittatus.* But in *D. semistriatus* the sequences displayed considerable variation within the coding and spacer regions. For each species, in *D. semistriatus* the alignment of 27 clones showed 139 variable sites (17 nucleotide substitutions located in the coding region and 122 in the NTS) and 29 indels located in the NTS. The percentage of differences in pairwise comparisons ranged from 5.2% to 7.5% in intraindividual comparisons and from 6.3% to 7.3% in interindividual comparisons. Global nucleotide diversity was 0.05744 (0.01360 in 5S and 0.07342 in NTS). In *D. trunculus* the alignment of 20 clones showed 25 variable sites (6 nucleotide substitutions located in the coding region and 19 in the NTS; and 11 indels located in the NTS). The percentage of differences in pairwise comparisons ranged from 2.0% to 2.8% in intraindividual comparisons and from 2.5% to 3.3% in interindividual comparisons. Global nucleotide diversity was 0.02176 (0.01948 in 5S and 0.02359 in NTS). In the case of *D. variegatus* the alignment of 22 clones presented 13 variable sites (4 nucleotide substitutions located in the coding region and 9 in the NTS) and an indel located in the NTS. Only four clones (*Dvar*1/5, *Dvar*1/8, *Dvar*2/3 and *Dvar*4/3) displayed a nucleotide substitution in the coding region. The percentage of differences in pairwise comparisons ranged from 0.1% to 0.4% in intraindividual comparisons and from 0.2% to 0.6% in interindividual comparisons. Overall nucleotide diversity was 0.00393 (0.00305 in 5S and 0.00425 in NTS). In *D.*

vittatus the alignment of 15 clones displayed 69 variable sites (15 nucleotide substitutions located in the coding region and 54 in the NTS) and an indel located in the NTS. The percentage of differences in pairwise comparisons ranged from 0.0% to 2.4% in intraindividual comparisons and from 3.2% to 5.3% in interindividual comparisons. Overall nucleotide diversity was 0.05607 (0.04828 in 5S and 0.05881 in NTS). The values of nucleotide divergence (D_{xy}) and the net number of nucleotide substitutions between groups (D_a) according to the Jukes and Cantor (1969) method are shown in Table 3. The values obtained between species are similar, with the highest values being found between *D. variegatus* and *D. vittatus*, and lower between *D. semistriatus* and *D. vittatus*.

Table 3. D_a (above diagonal) and D_{xy} values (below diagonal) and their standard deviation in the four *Donax* species analysed. The values of the diagonal correspond to the values of nucleotide diversity (π) of the 5S rDNA repeat unit.

	D. semistriatus	***D. trunculus***	***D. variegatus***	***D. vittatus***
D. semistriatus	0.05744	0.31435±0.02371	0.53230±0.05356	0.13307±0.01849
D. trunculus	0.35038±0.02368	0.02176	0.28505±0.04513	0.36153±0.04897
D. variegatus	0.56132±0.05355	0.29773±0.04512	0.00393	0.65667±0.09996
D. vittatus	0.18950±0.01841	0.39845±0.04895	0.68411±0.09994	0.05607

D_{xy} denotes the average number of substitutions per site between species and D_a the number of net substitutions between species (Nei, 1987).

In the four *Donax* species, the internal control regions (ICRs) described in other organisms were identified. A graphical representation of the 5S internal promoters and their consensus sequences is shown in Figure 1. The stretches from 3 to 18, from 37 to 44, from 48 to 61, and from 78 to 98 in the alignment displayed high homology with their orthologues ICR I, II, III, and IV of *Drosophila melanogaster* (Sharp and Garcia, 1988) (12/16, 8/8, 12/14, and 18/21 matches, respectively) (see Figure 1), and the stretches from 50 to 64, from 67 to 72, and from 80 to 97 were also similar to box A, intermediate element, and box C of *Xenopus laevis* somatic 5S RNA gene (Pieler et al., 1987) (11/15, 5/6, and 17/18 matches, respectively) (see Figure 1). Moreover, the NTS region of *Donax* species contain TATA-like motif recognized at around -28 nucleotides, other potential

transcription control sequences that may be involved in 5S transcription initiation, such as in the silkworm *Bombyx mori* (Morton and Sprague, 1984), in *Neurospora crassa* (Tyler, 1987) and *D. melanogaster* (Sharp and Garcia, 1988). The NTS sequences of *D. semistriatus* and *D. vittatus* retained the complete block TATATA at the 3´end; but not the other species, *D. trunculus* and *D. variegatus*, because one insertion T(G)ATATA and a point mutation (TATTTA) occurred within, respectively. Finally, a T-rich stretch, potentially related to transcription termination (Bogenhagen and Brown, 1981; Geiduschek and Tocchini-Valentini, 1988; Huang and Maraia, 2001), is located a few residues downstream of the coding region in the four *Donax* species.

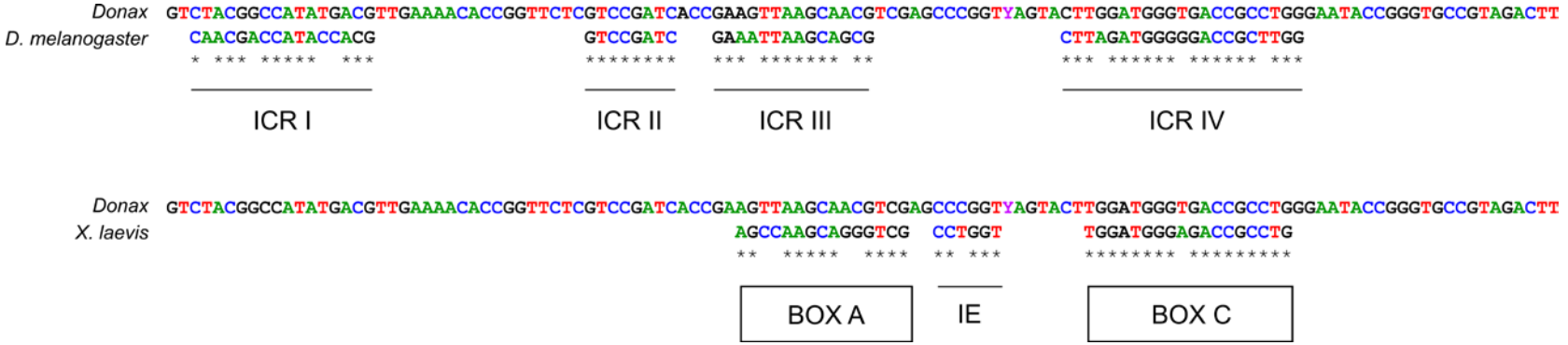


Figure 1. Graphical representation of the control elements involved in the transcription of 5S rDNA. The top sequences represent a schematic comparison of the ICRs between *Donax* consensus sequence and *D. melanogaster*. The bottom sequences represent a schematic comparison of the sequence elements (box A, intermediate element (IE), and box C) between *Donax* consensus sequence and *X. laevis*. The asterisks indicate similarities respect to consensus sequences described for *D. melanosgaster* and *X. laevis*.

ITS1 and ITS2

The ITS region was PCR amplified in at least 20 individuals of *D. trunculus* and *D. vittatus*, and 12 individuals of *D. semistriatus* and 8 individuals of *D. variegatus*, yielded a single product about 1000 bp for *D. vittatus*, about 800 bp for *D. trunculus* and about 900 bp for the rest of species. Table 4 shows the size and GC of the ITS region delimited according to the BLAST analysis result. The four *Donax* species showed slight size differences among clones due to the variation occurring in the ITS1 and ITS2, except *D. variegatus* did not display size differences in the ITS2. The total length of the ITS region was 814-1014 bp with 58-62.4% GC content; ITS1 was 400-542 bp and 59.7-62.6% GC; the 5.8S rRNA gene was 157 bp and 57.3% GC in all clones; and ITS2 was 254-316 bp and 55.7-64.8% GC. The GC content was similar in the four species, with the highest values in the ITS2 followed ITS1 and 5.8S gene, except for *D. trunculus* that displayed higher content in ITS1 than in ITS2 (Table 4).

The alignment of the different ITS sequences of *Donax* species consisted of 1048 pb and showed 258 variable sites of which 248 were parsimony informative sites, and 279 indels (see Supporting Information on CD-ROM: S2 File). The largest differences were found in ITS1 (162 variable sites), followed by ITS2 (90 variable sites) and 5.8 gene (6 variable sites). For each species, in *D. semistriatus* the alignment of 22 clones showed 33 variable sites (20 nucleotide substitutions located in the ITS1, one in the 5.8 gene and 12 in the ITS2) and 6 and 4 indels located in the ITS1 and ITS2, respectively. The percentage of differences in pairwise comparisons ranged from 0.2% to 2.5% in intraindividual comparisons and from 1.7% to 3.7% in interindividual comparisons. Overall nucleotide diversity was 0.01415 (0.01827 in ITS1, 0.00323 in ITS2 and 0.00058 in 5.8S). In *D. trunculus* the alignment of 23 clones presented 53 variable sites (21 nucleotide substitutions located in the ITS1, 2 in the 5.8 gene and 30 in the ITS2) and 7 and 17 indels located in the ITS1 and ITS2, respectively.The percentage of differences in pairwise comparisons ranged from 0.0% to 1.3% in intraindividual comparisons and from 0.03% to 2.1% in interindividual comparisons. Global nucleotide diversity was 0.01923 (0.01229 in ITS1, 0.04105 in ITS2 and 0.00314 in 5.8S). In *D. variegatus* the alignment of 13 clones displayed 28 variable sites (22 nucleotide substitutions

located in the ITS1, and 6 in the ITS2) and 4 indels located the ITS1. The percentage of differences in pairwise comparisons ranged from 0.00% to 0.01% in intraindividual comparisons and from 0.04% to 1.80% in interindividual comparisons. Overall nucleotide diversity was 0.01225 (0.01998 in ITS1, 0.00703 in ITS2 and 0.00000 in 5.8S). In *D. vittatus* the alignment of 12 clones showed 59 variable sites (17 nucleotide substitutions located in the ITS1, one in the 5.8 gene and 41 in the ITS2) and 11 and 38 indels located the ITS1 and ITS2, respectively. The percentage of differences in pairwise comparisons ranged from 0.2% to 2.20% in intraindividual comparisons and from 2.7% to 3.6% in interindividual comparisons. Global nucleotide diversity was 0.02166 (0.01272 in ITS1, 0.05163 in ITS2 and 0.00107 in 5.8S). In addition, the alignment of the four species revealed four stretches of 12, 23, 13, and 59 nucleotides in ITS1 (alignment positions 241-252, 316-338, 387-399 and 499-557, respectively) and four stretches of 17, 33, 22 and 17 nucleotides in ITS2 (alignment positions 715-731, 754-786, 790-811 and 862-878, respectively), all being highly conserved among *Donax* species (percentages of similarity higher than 96.7%). Sequence similarity of ITS2 (76.4%) was higher than that of ITS1 (70.6%) across species. The values of nucleotide divergence (D_{xy}) and the net number of nucleotide substitutions between groups (D_a) with Jukes and Cantor (1969) are shown in Table 5. The highest values being found between *D. trunculus* and *D. vittatus*, and lower between *D. semistriatus* and *D. variegatus*.

Table 4. Size (bp) and mean value of the GC content (%) of the ITS region.

Species	No. of clones	ITS1		5.8 gene		ITS2		ITS	
		Length	GC	Length	GC	Length	GC	Length	GC
D. semistriatus	22	452-457	62.6	157	57.3	283-287	64.0	892-900	62.1
D. trunculus	23	400-405	59.7	157	57.3	254-269	55.7	814-828	58.0
D. variegatus	13	452-453	62.6	157	57.3	283	64.8	892-893	62.4
D. vittatus	12	534-542	61.0	157	57.3	307-316	61.1	998-1014	60.5

Table 5. D_a (above diagonal) and D_{xy} values (below diagonal) and their standard deviation in the four *Donax* species analysed. The values of the diagonal correspond to the values of nucleotide diversity (π) of the ITS region.

	D. semistriatus	*D. trunculus*	*D. variegatus*	*D. vittatus*
D. semistriatus	0.01415	0.30096±0.03807	0.00386±0.00439	0.09768±0.01797
D. trunculus	0.31946±0.03804	0.01923	0.30675±0.04396	0.32262±0.04058
D. variegatus	0.01805±0.00459	0.32206±0.04394	0.01225	0.09384±0.02086
D. vittatus	0.11563±0.01780	0.33954±0.04054	0.11027±0.02065	0.02166

D_{xy} denotes the average number of substitutions per site between species and D_a the number of net substitutions between species (Nei, 1987).

Discussion

This work provides the nucleotide sequences of the 5S rDNA and the ITS region of four European *Donax* species, describes and characterizes for the first time in this group of organisms the general characteristics of these sequences and analyses their variation.

5S rDNA

Regarding the 5S rDNA of the four wedge clams studied show, at least in part, the conventional tandem arrangement, as deduced from successful amplification using contiguous primers with opposite orientation. Moreover, the BLAST analysis indicated that the 5S rRNA genes are separated from one another by a non-coding spacer region; the coding region was assigned to 120 bp in the four species and the NTS to the remaining sequence. The length of the characterized repeat units presented little variation among species (5-32 bp), except *D. trunculus* with a repeat unit of 277-285 bp and differing from the rest of the *Donax* species in 170-210 bp. Compared to other bivalve species, the *Donax* 5S rDNA units are among the shortest with the scallops *Aequipecten opercularis* (433-465 bp), *Mimachlamys varia* (453-455 bp), *Hinnites distortus* (451 bp) and *Pecten maximus* (463 bp) (Insua et al., 1998; López-Piñón et al., 2008) and the razor clams *Ensis arcuatus* (420 bp), *Ensis siliqua* (422 bp), *Ensis directus* (443 bp) and *Ensis macha* (434 bp) (Fernández-Tajes and Méndez, 2007); although the cockles *Cerastoderma edule* and *Cerastoderma glaucum* have a repeat unit of 544-546 and 539-568 bp (Insua et al., 1999; Freire et al., 2005), respectively; the oysters about 1100 bp in *Crassostrea* and 2000 bp in *Ostrea* (Cross and Rebordinos, 2005, 2006); and the mussels *Mytilus edulis*, *Mytilus galloprovincialis* and *Mytilus trossulus* have three types,~260 bp (α band), ~770 bp (β band) and ~1000 bp (γ band) in length, the mussel *Mytilus californianus* has other three types ~240 bp (small-β band), ~730 bp (β band) and ~980 bp (γ band), and the mussel *Mytilus coruscus* has a repeat unit of ~300 bp (Freire et al., 2010a). These differences in size are due to the NTS and can be very useful to differentiate among wedge clams and from other bivalves when morphological criteria are difficult, for instance processed samples, samples without shell or during the larval stage. Despite the fact that *Donax* species display length differences in the NTS

and a high sequence divergence deduced from the difficulty in obtaining unambiguous sequence alignments, all four species studied show similar GC content (38.2-43.35%), with higher values in the gene (53.4-55%) than in the spacer region (33.1-39.1%). This difference between coding and spacer region also occurs in the mussels, cockles and scallops previously studied (Insua et al., 2001; Freire et al., 2005, 2010a; López-Piñón et al., 2008), but also in species of crustaceans (Pelliccia et al., 1998) and fish (Sajdak et al., 1998; Sola et al., 2003), for example, which have AT rich spacers (73% and >57%, respectively). The mammalian 5S NTS has been shown to be GC rich (>60%) (Suzuki et al., 1994, 1996), and the spacer of the oocyte-type 5S rDNA of *Xenopus* is AT rich but that of the somatic type is GC rich (Peterson et al., 1980).

The four *Donax* 5S rDNA units consisted of a coding region linked to a spacer without any other coding sequence associated. This result agrees with that observed in cockles (Freire et al., 2005) and scallops (López-Piñón et al., 2008), but contrasts with alternative arrangements that have been described, such as the linkage between 5S rDNA and small nuclear RNA (snRNA) in the crustacean *Asellus aquaticus* (Pelliccia et al., 2001), the oysters *Crassostrea gigas* and *Crassostrea angulata* (Cross and Rebordinos, 2005) and the sole *Solea senegalensis* (Manchado et al., 2006a), or the linkage of 5S gene with histone genes in the mussel *M. galloprovincialis* (Eirín-López et al., 2004). Although the use of other PCR amplification conditions or the analysis of a genomic library could reveal 5S rDNA units linked to other multigene families, it is not surprising that the *Donax* 5S rDNA arrangement differ from other bivalve or species, as the 5S rRNA gene linkages seem to be repeatedly established and lost during the evolution of eukaryotic genomes (Drouin and Moniz de Sá, 1995).

The ICRs involved in the transcription of 5S rDNA and the sequence elements box A, IE, and box C were identified in the four *Donax* species (Figure 1). Moreover, all 5S rDNA sequences analysed here displayed TATA-like motifs upstream of the coding region very similar to that reported in *B. mori* (Morton and Sprague, 1984), in *N. crassa* (Tyler, 1987), *D. melanogaster* (Sharp and Garcia, 1988) and several fish species (Martins and Galetti, 2001). Although functional assays are necessary to know the role of these TATA-like motifs, according to the position, the one located nearest the

gene is good candidate for the interaction with TFIIIB (Freire et al., 2005). Recently, Raha et al. (2010) have proposed that this region could be involved in RNA pol III transcription together with RNA pol II-like transcriptional factors. Nevertheless, it was less conserved in *D. trunculus* and *D. variegatus* because one insertion (T(G)ATATA) and a point mutation (TATTTA) occurred within, respectively, as it happens in cockle, razor clam and scallop species (Vizoso et al. 2011). Therefore, as previous authors indicate (Vizoso et al., 2011), this could imply that in these molluscan groups i) the 5S rDNA transcription could not specifically be regulated by RNA pol II–like transcriptional factors, ii) they could present lower transcriptional activities, or iii) they do not require the same level of sequence specificity. Moreover, all 5S rDNA sequences showed a T-rich stretch potentially related to transcription termination (Bogenhagen and Brown, 1981; Geiduschek and Tocchini-Valentini, 1988; Huang and Maraia, 2001).

ITS1 and ITS2

Regarding the ITS region, the lengths determined for both ITS1 (400-542 bp) and ITS2 (254-316 bp) in the four *Donax* species are in line with those of other bivalve species. Average ITS1 length was 461.9 bp, and GC content was 61.5%, values very similar to those obtained by Chow et al. (2009), who studied the ITS1 of several marine animals and reported wide data regarding length and GC content for marine mollusc species (in Mollusca average ITS length was 492.5 bp, and GC content was 55.9%). The ITS1 and ITS2 lengths for the clams *Venerupis pullastra*, *Ruditapes decussatus* and *Ruditapes philippinarum* ranged between 600-715 and 316-396 pb, respectively (Fernández et al., 2001). The ITS1 and ITS2 of four scallops (*A. opercularis*, *M. varia*, *H. distortus*, and *P. maximus*) are 209-277 and 270-294 bp (Insua et al., 2003; Wang et al., 2007), respectively; and their GC content was 43-49% and 44-49%. They ranged between 367-514 and 317-446 bp in the Unionoidea species *Unio pictorum*, *Unio tumidus*, *Unio crassus*, *Anodonta anatina*, *A nodonta cygnea*, *Pseudanodonta complanata*, and *Margaritifera margaritifera* (Källersjö et al., 2005). In the Veneridae species *Meretrix meretrix*, *Cyclina sinensis*, *Mercenaria mercenaria*, *Protothaca jedoensis*, *Dosinia corrugata* and *R. philippinarum*, ITS1 and ITS2 length were 522-900 and 281-

412 bp, respectively; and their GC content were 57.66-65.62% and 65.21-67.87% (Cheng et al., 2006). In the two cockles *C. edule* and *C. glaucum*, ITS1 and ITS2 length ranged between 226-251 and 305-325 bp, and their GC content was 52-62% and 61-63%, respectively (Freire et al., 2010b). Data on ITS1 and ITS2 length in the razor shell *E. directus* ranged between 484-510 and 295-299 bp, and their average GC content were 58.9% for ITS1 and 63% for ITS2 (Vierna et al., 2010). Thus, *Donax* species ITS length and GC content were similar to those found in other bivalve species. Just as in the clams *V. pullastra*, *R. decussatus* and *R. philippinarum* (Fernández et al., 2001) and other Veneridae species (Cheng et al., 2006), the ITS1 *Donax* spacer is more longer than ITS2, while in other bivalves the two spacers differ by <100 bp (Kenchington et al., 2002; Källersjö et al., 2005; Freire et al., 2010b). As bivalve data accumulate, it seems that there are few constraints affecting the variation in spacer length because all kinds of situations may occur, e.g. ITS1>ITS2, ITS1<ITS2, and ITS1 and ITS2 of similar size (Freire et al., 2010b). ITS GC content in *Donax* species is similar to that in venerids, *E. directus* and *Cerastoderma* species, as would be expected considering that scallops are Pteriomorphia bivalves, and venerids, *Ensis* and *Cerastoderma* species are Heteroconchia. The high GC content of the ITS1 and ITS2 contrasted with the very low GC content of the NTS. This could be due to the fact the NTS region is not transcribed or folded into a secondary structure, whereas both ITS1 and ITS2 are transcribed and have known secondary structures. Maybe the high GC content is related to secondary structure stability. The length showed here for the 5.8S gene (157 bp) was previously described for the ocean quahog *Arctica islandica* (Dahlgren et al., 2000), the four scallops studied by Insua et al. (2003), and the six Veneridae species studied by Cheng et al. (2006); although sizes of 158 bp (Fernández et al., 2001), 156 bp (Kenchington et al., 2002), and 158-161 bp (Freire et al., 2010b) were reported in some species, but all of them agree with the average length of eukaryote 5.8S rRNA of about 160 bp deduced from direct sequencing (Nazar, 1984). As expected for a high conserved sequence, the GC content of the 5.8S gene did not show variation (57.3%) and the values correspond to those observed in bivalves (Freire et al., 2010b) and other animals (Tautz et al., 1988, Wesson et al., 1992; Odorico and Miller, 1997).

The four *Donax* species showed intraindividual variation mainly in the spacers, ITS1 being more variable than ITS2 in *D. semistriatus* and *D. variegatus*, as evidenced by the number of variable sites in the sequence alignments and the distance values in pairwise comparisons. This is in line with that observed in other organisms such as *D. melanogaster* (Schlötterer and Tautz, 1994), *Similium damnosum* (Tang et al., 1996), and *Cerastoderma* species (Freire et al., 2010b). By the contrast, ITS2 being more variable than ITS1 in *D. trunculus* and *D. vittatus*, as in scallop species (Insua et al., 2003) and with similar values to that described in ITS2 in the pearl oysters *Pinctada martensi*, *Pinctada maxima*, *Pinctada margaritifera*, *Pinctada chemnitzi*, *Pinctada nigra*, and *Pteria penguin* (He et al., 2005). Nevertheless, intraindividual variation of the ITS sequences for *D. variegatus* was minimal or almost nonexistent as in the *M. varia* and *P. maximus* scallop species (Insua et al., 2003). Therefore, it seems that intradindividual variation in *D. variagatus* is more moderate than that showed in the other three *Donax* species and that described in some other animal species (Wesson et al., 1992; Tang et al., 1996; Odorico and Miller, 1997; Gandolfi et al., 2001). Globally, both ITS1 and ITS2 show sequence variation among wedge clams, with sequence similarity of ITS2 higher than that of ITS1 across species. However, blocks highly conserved across the *Donax* species were distinguished both in ITS1 and ITS2, which may suggest that they play a role in rRNA processing.

Future implications

Despite the variation observed in the 5S rDNA and ITS region among *Donax* species, these sequences could allow the identification of reliable molecular markers than can be used to differentiate this wedge clams. DNA-based markers have been shown to be very useful for this purpose and, specifically, the use of the 5S rDNA and ITS have been reported to be valuable for discrimination of several bivalve species with commercial value, such as clams (Fernández et al., 2001; Hurtado et al., 2011), cockles (Freire et al., 2005), mussels (Dixon et al., 1995; Heath et al., 1995; Toro, 1998; Santaclara et al., 2006), oysters (Cross and Rebordinos, 2006), razor clams

(Fernández-Tajes and Méndez, 2007), scallops (López-Piñón et al., 2002) and wedge clams (Pereira et al., 2012).

Conclusions

In this work, we describe and characterize for the first time the 5S rDNA and the ITS region of four species of the genus *Donax* present in Europe. Not only has provided new interesting data for the scientific community about *Donax* species, of which there is not much genetic information, but also it can be feasible for application in aquaculture, for the development of genetic markers useful in wedge clams identification and authentication of commercial species, and aquaculture stock management.

Acknowledgements

We would like to thank Dra D. Martínez Patiño and S. Nóvoa from Centro de Cultivos Marinos de Ribadeo – CIMA (Xunta de Galicia) and Dr. M. Gaspar from Instituto Português do Mar e da Atmosfera – IPMA (Portugal) for supplying the samples. This work was supported by the Ministerio de Economía y Competitividad (Spain) through project AGL2016-75288-R AEI/FEDER, UE.

Supporting information (see on CD-ROM)

S1 File. Alignment of the 5S rDNA sequences of the four European *Donax* species.

S2 File. Alignment of the ITS sequences of the four European *Donax* species.

Chapter 5

An alternative method for rapid and specific authentication of four European *Donax* species, including *D. trunculus*, a commercially-important bivalve

Jenyfer Fernández-Pérez[1], Ana Nantón[1], Josefina Méndez[1]

[1] Grupo Xenomar, Departamento de Bioloxía, Facultade de Ciencias and Centro de Investigaciones Científicas Avanzadas (CICA), Universidade da Coruña, Campus de A Zapateira, 15071 A Coruña, Spain

Published in *European Food Research and Tecnology* (2018).

https://doi.org/10.1007/s00217-018-3093-5

Abstract

Seafood fraud or mislabelling has attracted much attention as it is widespread worldwide, and the scientific name of the species must be indicated on commercial products to satisfy labelling and traceability requirements. Species identification using morphological criteria is difficult to non-specialist, and shell-less specimens, processed products or larval and juvenile forms are morphologically unidentifiable. On the other hand, conventional methods, such as PCR-RFLPs, are tedious, expensive and time-consuming. As an alternative, a multiplex PCR of two rDNA regions, containing the 5S and the ITS, was developed for the identification of the four European *Donax* species: *Donax semistriatus, Donax trunculus, Donax variegatus* and *Donax vittatus*. The species-specific electrophoretic profiles allowed their identification in an efficient, fast, simple and low-cost method, offering a feasible application in aquaculture. These genetic markers can be very useful for traceability of the species, application in the management of wedge clam cultures, and conservation of these resources.

Introduction

Four *Donax* species, *Donax semistriatus* (Poli, 1775), *Donax trunculus* (Linaeus, 1758), *Donax variegatus* (Gmelin, 1791) and *Donax vittatus* (Da Costa, 1778), are common along the Atlantic and the Mediterranean coasts and live sympatrically in some areas of the European littoral, such as Portugal (Gaspar et al., 2002; Rufino et al., 2010) and Spain (Salas et al., 2001; Bejega García et al., 2010; Fernández-Pérez et al., 2017a). These marine bivalves play an important socioeconomic role in some European seaside regions. From 1999, when fishermen were obliged to declare their captures (Gaspar et al., 1999), to 2015 Europe produced nearly 14,000 tons of wedge clams (FAO-FIGIS, 2017). Despite the importance of *Donax* species for artisanal and small-scale fishing, FAO statistics do not differentiate among species. *D. trunculus* represents an important fishery resource in France (Thébaud et al., 2005), Italy (Zeichen et al., 2002), Portugal (Gaspar et al., 1999; Chícharo et al., 2002), Spain (Molares et al., 2008) and Turkey (Özden et al., 2009), and may account for

most of the catches recorded by FAO in these countries for the genus *Donax*. Even though *Donax* species can be differentiated by their morphological traits, due to their variability in shape, size and colour, and cohabitation of several species in one region, captures of *D. trunculus* could include other species of the genus with lesser economic value and they might all be sold as *D. trunculus*. In the case of Galicia (northwest of Spain) and northern Portugal, it cannot be excluded that captures contain other *Donax* species and that they are sold as *D. trunculus*, a species with great commercial interest. Moreover, processed products, shell-less specimens or larval and juvenile forms are morphologically unclassifiable. European Union labelling regulations specify that seafood products must be correctly labelled with their commercial and scientific names to guarantee their traceability and correct identification throughout the value chain (EC 104/2000; EC 2065/2001; EU 1379/2013). The recent Art. 35 of the 1379/2013 EU Regulation on Common Market Organization requires that consumers receive precise information about the seafood they purchase, such as scientific and commercial names, production method, catch area and fishing gears used (D´Amico et al., 2016). For this reason, it is necessary to develop analytical methods for the identification and authentication of commercial food species to protect consumers' rights, to detect and avoid commercial fraud, to avoid the unintentional substitution of different wedge clam species, to enforce labelling regulations and also to achieve other quality objectives, such as a certificate of origin. To achieve this goal, molecular techniques based on DNA have been used in different bivalve molluscs by several authors (e.g., Santaclara et al., 2006; Fernández-Tajes et al., 2010, 2011; Freire et al., 2011; Marín et al., 2013) for being a very stable and long-lived biological molecule present in all tissues of all organisms. Molecular methods based on polymerase chain reaction (PCR) are used to verify fish and seafood species due to their sensitivity, simplicity and specificity. Moreover, these techniques could allow the species authentication independently of their life stage (larva, juvenile or adult) or state (fresh or processed).

The genomic organization of rDNA is characterized by showing multiple clusters of tandem repeat units, each of them including highly conserved coding regions with slow evolution rates, and variable non-transcribed spacers which show fast evolution. While the slowly evolving conserved

regions (coding regions) are useful for comparing distantly related species, the faster evolving external transcribed spacer (ETS), non-transcribed spacer (NTS) regions and the internal transcribed spacer (ITS) are appropriate for differentiating more closely related groups, for instance species within a genus. The 5S rDNA, forms a multigene family of tandem arrays, whose unit of repetition is composed of a highly conserved coding region of 120 bp, that facilitates the primer design for its amplification, and a NTS variable in both sequence and length that evolves more rapidly. These features make it attractive as a tool in the search of species-specific DNA molecular markers. This marker has been reported to be useful for the identification of several bivalve mollusc species with commercial importance, such as cockles (Freire et al., 2005), oysters (Cross and Rebordinos, 2006), razor clams (Fernández-Tajes and Méndez, 2007), and wedge clams (Pereira et al., 2012). Moreover, the internal transcribed spacer regions (ITS1 and ITS2), flanked by the nuclear ribosomal RNA subunits 18S, 5.8S and 28S, evolve rapidly compared with other nuclear regions (Gerbi, 1985) and have also been commonly used to distinguish several bivalve species. ITS size and Restriction Fragment Length Polymorphisms (RFLPs) identify some *Mytilus* mussels (Heath et al., 1995; Toro, 1998), Veneridae clams (Fernández et al., 2001; Hurtado et al., 2011), and Pectinidae scallops (López-Piñón et al., 2002); ITS1-RFLPs discriminate *Perna* and *Choromytilus* mussels (Santaclara et al., 2006), and several freshwater mussels (White et al., 1994); and ITS2 size was used to differentiate *M. edulis* and two *Modiolus* species (Dixon et al., 1995).

In *Donax*, DNA based molecular analyses for species authentication are scarce. Pereira et al. (2012) developed one method based on the variable length of 5S rDNA PCR products, but it only allows for the discrimination of *D. variegatus* and *D. trunculus* from other *Donax* species. Another method is based on PCR-RFLP analysis of cytochrome c oxidase subunit I (COI) (Nantón et al., 2015) with which *D. semistriatus*, *D. trunculus*, *D. variegatus* and *D. vittatus* have been identified. Although this last method is capable of distinguishing among the four *Donax* species, RFLP techniques are time-intensive, laborious and expensive, especially when large numbers of samples need to be analysed. Moreover, given that Doubly Uniparental Inheritance (DUI, an unusual pattern of mitochondrial DNA inheritance that involves two separate lineages; for review see Zouros, 2013),

has been described in *D. trunculus* (Theologidis et al., 2008) and we have recently found evidence for it in *D. vittatus* and *D. semistriatus* (Fernández-Pérez et al., 2017a), a method based on nuclear DNA markers could avoid presence of biparental transmission affecting species identification.

The goal of the present study was to develop an alternative, simple, rapid and inexpensive authentication protocol, based on multiplex PCR, that provided a clear electrophoretic profile for the differentiation of the four *Donax* species present in Europe, including the wedge clam *D. trunculus* with high commercial importance.

Material and methods

Sample collection and DNA extraction

All individuals were collected from natural beds. In the case of *D. trunculus*, specimens were obtained from Vilarrube (nothwestern Spain), while *D. semistriatus*, *D. variegatus* and *D. vittatus* samples were collected from the Portuguese coast (Table 1). Specimens were taxonomically identified using a species-specific PCR-RFLP analysis of cytochrome c oxidase subunit I (COI) capable of discriminating among the four *Donax* species (Nantón et al., 2015). Total genomic DNA was extracted from ethanol-preserved foot using a Chelex-100 (Sigma-Aldrich, USA) protocol based on Walsh et al. (1991).

Table 1. Sampling details of species.

Species	Locality	Country	Latitude	Longitude
D. semistriatus	Monte Gordo	Portugal	37.167	-7.503
D. trunculus	Vilarrube	España	43.644	-8.077
D. variegatus	Monte Gordo	Portugal	37.100	-7.633
D. vittatus	Mira-Vagueira	Portugal	40.614	-8.769

PCR amplification of the 5S rDNA and the ITS region

The 5S rDNA repeat unit and the ITS region were PCR amplified in at least 20 individuals of each wedge clam species, except for *D. semistriatus* and *D. variegatus* for which the number of available individuals was low (12 and 8, respectively).

For the amplification of 5S rDNA, primers 5SF (5`-CGTCCGATCACCGAAGTTAA-3´) and 5SR (5`-ACCGGTGTTTTCAACGTCAT-3´) were used (Fernández-Tajes and Méndez, 2007). These primers have opposite orientation, are separated by 3 bp and anneal at positions 13-33 y 36-55 of the 5S transcribing region. They were designed for the amplification of one unit of any tandemly arranged 5S rDNA in razor clam species (Fernández-Tajes and Méndez, 2007). Amplification reactions were performed in 25 µl containing 150 ng of genomic DNA, 0.25 µM of each dNTP, 2 mM of $MgCl_2$, 0.6 µM of each primer, 0.6 U of *Taq* polymerase (Roche Applied Science) and the buffer recommended by the polymerase suppliers. PCR cycling protocol consisted of an initial denaturation step of 2 min at 95 °C, 35 cycles with a denaturation step at 95 °C for 30 s, an annealing step at 55 °C for 30 s and an extension step for 1 min at 72 ºC, and a final extension step at 72 °C for 5 min.

For the amplification of ITS1 and ITS2, PCR reactions were performed with a pair of primers that anneal at the 3´end of the 18S ribosomal gene and the 5´end of the 28S ribosomal gene (ITSF: 5´-GTTTCCGTAGGTGAACCTG-3´ and ITSR: 5´-CTCGTCTGATCTGAGGTCG-3´) (Heath et al., 1995). They were carried out in 25 µl containing 100 ng of genomic DNA, 0.25 µM of each dNTP, 1.5 mM of $MgCl_2$, 1 µM of each primer, 0.625 U of *Taq* polymerase (Roche Applied Science) and the buffer recommended by the polymerase suppliers. PCR cycling protocol consisted of an initial denaturation step of 3 min at 94 °C, 30 cycles with a denaturation step at 94 °C for 20 s, an annealing step at 55 °C for 20 s and an extension step for 45 s at 72 °C, and a final extension step at 72 °C for 5 min. PCR products were migrated in 2% agarose gel electrophoresis. Gels were stained by immersion in a 0.5 µg/ml ethidium bromide solution for 30 min, visualised and recorded on a transilluminator Gel Doc XR Systems (Bio-Rad, Barcelona, Spain).

Multiplex PCR design

Given that, taken together, the PCR-amplified products obtained for 5S rDNA and ITS allow for the discrimination of the four *Donax* species, the primers 5SF, 5SR, ITSF, and ITSR were combined into a multiplex PCR. For multiplex PCR amplifications, 200 ng of Chelex-extracted DNA were used in a reaction volume of 12.5 µl with 1x Qiagen Multiplex PCR Master Mix (containing HotStarTaq® polymerase, 3Mm $MgCl_2$ and dNTPs), 0.25x Q-solution (Qiagen), 5 µg bovine serum albumin (BSA), all primers at a final concentration of 0.2 µM and RNase-Free water (Qiagen) to adjust the volume. Cycling conditions were 15 min at 95 ºC, followed by 35 cycles of 30 s at 94 ºC, 90 s at 58 ºC, 90 s at 72 ºC, and final elongation at 72 ºC for 10 min. PCR products were migrated in 3% agarose gel electrophoresis and then visualised and recorded as described above.

Results and discussion

Developing proper tools that could be easily implemented by the government or private entities to deal with food safety and consumers' rights is essential to guarantee species identification and to ensure composition and safety of commercial products. In addition, these methods could be useful to achieve other quality goals, such as a certificate of origin. DNA based markers have been shown to be very useful for this purpose, mainly because the generalized application of PCR has simplified the technology. PCR has been used to differentiate bivalve species, such as mussels (Santaclara et al., 2006; Fernández-Tajes et al., 2011), scallops (López-Piñón et al., 2002; Marín et al., 2013), oysters (Cross and Rebordinos, 2006), razor clams (Fernández-Tajes and Méndez, 2007; Fernández-Tajes et al., 2010), cockles (Freire et al., 2005; Freire et al., 2011) or clams (Fernández et al., 2001; Hurtado et al., 2011). In this work, we developed a molecular protocol, based on multiplex PCR amplification of two rDNA regions containing the 5S and the ITS, to discriminate among four different wedge clam species belonging to the *Donax* genus.

When simplex PCR was used, a single product of ~275 bp was obtained in *D. trunculus*, ~450 bp in *D. variegatus* and ~500 bp in the other two *Donax* species for the 5S gene, while ITS yielded a

fragment of ~1000 bp in *D vittatus*, ~800 bp in *D. trunculus* and ~ 900 bp in the remaining species. Figures 1 and 2 show the usual amplification patterns.

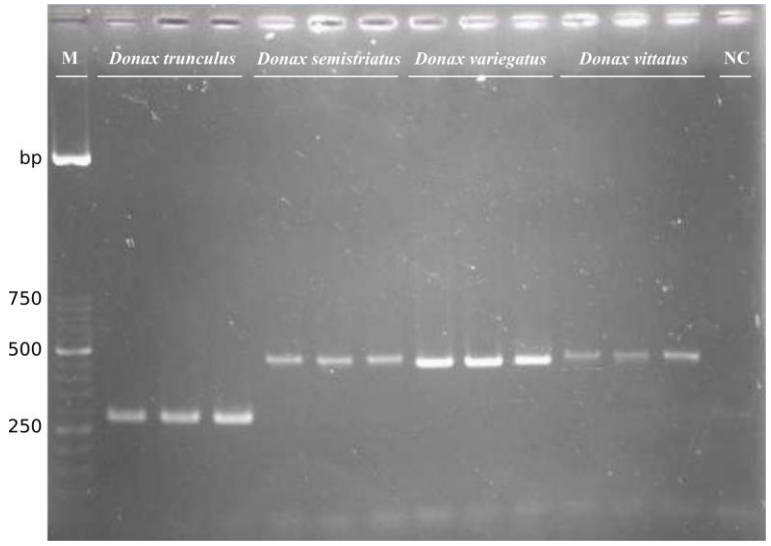


Figure 1. Image of a 2% agarose gel electrophoresis of 5S rDNA repeat unit PCR products of *D. trunculus* (lanes 2-4), *D. semistriatus* (lanes 5-7), *D. variegatus* (lanes 8-10) and *D. vittatus* (lanes 11-13). M: A 50-bp ladder was used as DNA marker (lane 1). NC: Negative control (lane 14).

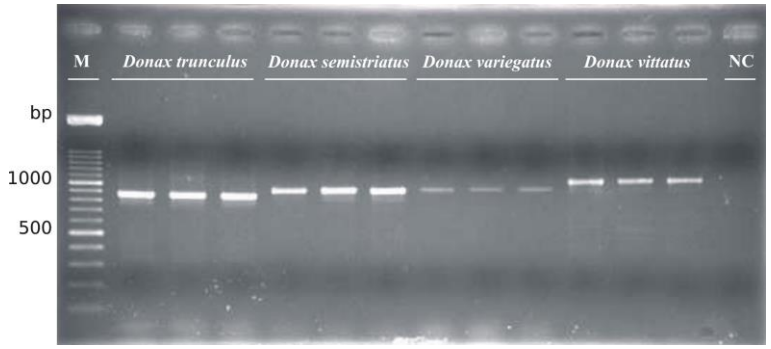


Figure 2. Image of a 2% agarose gel electrophoresis of ITS1 and ITS2 PCR products of *D. trunculus* (lanes 2-4), *D. semistriatus* (lanes 5-7), *D. variegatus* (lanes 8-10) and *D. vittatus* (lanes 11-13). M: A 100-bp ladder was used as the DNA marker (lane 1). NC: Negative control (lane 14).

In multiplex PCR, 5S and ITS primers have allowed us to differentiate among species. The results obtained (Figure 3) showed that primers did not interfere with each other and that they bound

specifically to the DNA target. Amplified fragments showed the same length as they did when primer pairs were used separately.

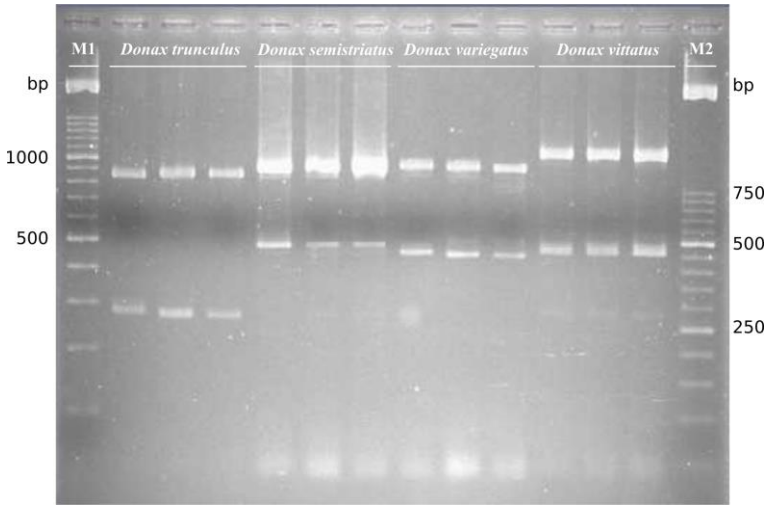


Figure 3. Image of a 3% agarose gel electrophoresis of the 5S rDNA and ITS products obtained from four *Donax* species. *D. trunculus* (lanes 2-4), *D. semistriatus* (lanes 5-7), *D. variegatus* (lanes 8-10) and *D. vittatus* (lanes 11-13). M1 and M2: Both the 100-bp and 50-bp ladders were used as DNA marker (lane 1 and lane 14, respectively).

Although two available and effective methods exist to identify *Donax* species (Pereira et al., 2012; Nantón et al., 2015), the former only allows for the differentiation of *D. trunculus* and *D. variegatus* from the remaining *Donax* species, while the latter is a species-specific PCR-RFLP analysis, tedious and expensive to carry out in a large number of samples. Therefore, our goal was to develop a faster, cheaper and more effective method that can be carried out by governmental institutions to detect and avoid commercial fraud. Furthermore, given that DUI has been described in some *Donax* species (Theologidis et al., 2008; Fernández-Pérez et al., 2017a), our method, based on nuclear DNA markers, avoids presence of biparental transmission affecting species determination. Relative to previously described molecular methods for *Donax* identification, the multiplex PCR described here reduces costs in terms of expendable materials as well as greatly reducing processing time. The cost of materials for one single multiplex reaction described in this study, including materials used in DNA extraction and electrophoresis, reduces the cost by half compared to RFLP techniques.

Therefore, our current method uses fewer reagents and is less expensive than any identification method that uses restriction endonuclease digestion.

The present study describes a technique that provides an easy, rapid, reliable and inexpensive protocol involving multiplex PCR amplification of the 5S rDNA and the ITS, which can be routinely applied to distinguish *Donax* species. Due to its speed, approximately one working day, and the fact that it does not require the digestion of PCR products with restriction endonucleases, this approach can be regularly carried out in food inspection laboratories, particularly when a high number of samples must be analysed. Moreover, it has been recently demonstrated that there is a real possibility of obtaining seed of *D. trunculus* in hatcheries and release it into natural beds to reach commercial size (Louzán et al., 2016). Therefore, the technique should be useful for not only the conservation of these marine resources, but also the characterization of stock breeders used in wedge clam culture in order to obtain their correct identity and choose the most suitable broodstock to obtain seed.

Conclusions

Genetic differentiation of *Donax* species *D. semistriatus*, *D. trunculus*, *D. variegatus* and *D. vittatus* was achieved by multiplex PCR amplification of two rDNA regions. This method offers sufficient accuracy to satisfy traceability requirements of *Donax* species in addition to applicability to the management of wedge clam culture and its ease of application even for non-specialised laboratories, since it is a fast, simple and efficient technique.

Acknowledgements

We would like to thank Dra. L. Martínez for her constructive and valuable comments. We also thank two anonymous referees for useful comments during manuscript review. This work was supported by the Ministerio de Economía y Competitividad (Spain) through project AGL2016-75288-R AEI/FEDER, UE.

Chapter 6

First complete female mitochondrial genome in four bivalve species genus *Donax* and their phylogenetic relationships within the Veneroida order

Jenyfer Fernández-Pérez[1], Ana Nantón[1], Francisco J. Ruiz-Ruano[2], Juan Pedro M. Camacho[2], Josefina Méndez[1]

[1] Grupo Xenomar, Departamento de Bioloxía, Facultade de Ciencias and CICA (Centro de Investigacións Científicas Avanzadas), Universidade da Coruña, Campus de A Zapateira, 15071 A Coruña, Spain

[2] Departamento de Genética, Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain

Published in *PLoS ONE*, 12(9): e0184464 (2017).

https://doi.org/10.1371/journal.pone.0184464

Abstract

Background: Four species of the genus *Donax* (*Donax semistriatus, Donax trunculus, Donax variegatus* and *Donax vittatus*) are common on Iberian Peninsula coasts. Nevertheless, despite their economic importance and overexploitation, scarce genetic resources are available. In this work, we newly determined the complete mitochondrial genomes of these four representatives of the family Donacidae, with the aim of contributing to unveil phylogenetic relationships within the Veneroida order, and of developing genetic markers being useful in wedge clam identification and authentication, and aquaculture stock management.

Principal Findings: The complete female mitochondrial genomes of the four species vary in size from 17,044 to 17,365 bp, and encode 13 protein-coding genes (including the *atp8* gene), 2 rRNAs and 22 tRNAs, all located on the same strand. A long non-coding region was identified in each of the four *Donax* species between *cob* and *cox2* genes, presumably corresponding to the Control Region. The Bayesian and Maximum Likelihood phylogenetic analysis of the Veneroida order indicate that all four species of *Donax* form a single clade as a sister group of other bivalves within the Tellinoidea superfamily. However, although Tellinoidea is actually monophyletic, none of its families are monophyletic.

Conclusions: Sequencing of complete mitochondrial genomes provides highly valuable information to establish the phylogenetic relationships within the Veneroida order. Furthermore, we provide here significant genetic resources for further research and conservation of this commercially important fishing resource.

Introduction

Bivalve molluscs of the genus *Donax* (Donacidae family) are an important constituent of the macrofauna of sandy beaches in temperate, tropical and subtropical zones, being the dominant organisms in this type of environment (Ansell, 1983). In the littoral of Iberian Peninsula, the five European species of *Donax* live sympatrically in the same beaches (Salas et al., 2001; Rufino et al., 2010): *Donax trunculus* (Linnaeus, 1758) (Atlantic and Mediterranean), *Donax vittatus* (Da Costa, 1778) (Atlantic), *Donax variegatus* (Gmelin, 1791) (Atlantic and Mediterranean), *Donax semistriatus* (Poli, 1775) (Atlantic and Mediterranean) and *Donax venustus* (Poli, 1775) (Atlantic and Mediterranean) (Ansell and Lagardère, 1980; Salas-Casanova, 1987; Salas-Casanova and Hergueta, 1990; Gaspar et al., 2002). Nevertheless, *D. venustus* is practically non-existent in the Iberian Peninsula as a single individual has been found between the years 2000 and 2006 along the south coast of Portugal (Rufino et al., 2010).

Few species of the genus *Donax* are commercially exploited, but some are consumed locally or used as fishing bait. *D. trunculus* is exploited in many countries bordering the Mediterranean Sea and Atlantic Ocean, including Portugal (Gaspar et al., 1999; Chícharo et al., 2002), Italy (Zeichen et al., 2002), France (Thébaud et al., 2005), and Spain (Ramón et al., 2005; Molares et al., 2008). Only in Iberian Peninsula, the recorded captures since 1999 to 2014 equal 10,156 tons, with a maximum production of 1,042 tons in 2005 followed by an incessant decline reaching only 250 tons in 2014 (FAO-FIGIS, 2017). Although this data only reflects production since fishermen were obliged to declare their captures (Gaspar et al., 1999), the species has been subjected to intense exploitation over the last decades and, currently, some *D. trunculus* populations seem to be at high long-term risk of extinction (Marie et al., 2016). Furthermore, this species constitutes an important shellfish resource due to its high economical value. For instance, in Galicia (northwest of Spain), *D. trunculus* is a species with a high contribution rate, being the bivalve with greater commercial value (38.52 €/kg in the year 2016) (data from http://www.pescadegalicia.gal/estadisticas/) in markets during last years. Due to the similarity in size, shape and colour of the *Donax* clams in different species, captures of *D. trunculus* in natural

beds may contain other species of the genus with lesser economical value and may be marketed together. However, despite their overexploitation and economic importance, relatively few genetic resources are available for this species (Nantón et al., 2014; Marie et al., 2016) and the whole genus (Pereira et al., 2012; Nantón et al., 2015).

In order to preserve this important fishing resource, genetic tools should be employed. Molecular genetics has proven highly informative to determine the level of genetic variability, which is an essential feature to consider when defining conservation priorities, as well as to better understand the (recent) evolutionary history of species groups. Within the molecular resources, mitochondrial (mt) genome stands out to be considered a useful tool for population genetic and phylogenetic studies, not only because complete mt genomes are often more informative than single genes, but also because they reveal some genome-level details, such as the rearrangement of genes, which are valuable information for studies of evolutionary relationships among species (Boore and Brown, 1998; Rokas and Holland, 2000; Gissi et al., 2008; Shen et al., 2009). Moreover, mitochondrial DNA (mtDNA) is particularly important in helping to differentiate species that are morphologically similar, contributing to the identification and authentication of commercial food species to detect and avoid fraud, to protect consumer rights and to achieve other quality objectives, such as certificate of origin.

Most metazoan mitochondrial genomes are typically closed circular molecules of ~16 kb, enconding 37 genes: 13 protein-coding genes (PCGs), 22 transfer RNA (tRNA) genes and two ribosomal RNA (rRNA) genes (Wolstenholme, 1992). In addition, at least one extensive non-coding sequence is present which contain elements that control the initiation of replication and transcription (Hoffman et al., 1992). Mitochondrial genome has several valuable features that make it exclusive, including its small size, high evolutionary rates, limited recombination, relatively conserved gene content and organization, and maternal inheritance (Gissi et al., 2008; Zhou et al., 2009). Though, an extreme exception to the paradigm of strict maternal inheritance of animal mtDNA (SMI) is found in some bivalve lineages, which possess an unusual system known as doubly uniparental inhertince (DUI) (Breton et al., 2007; Passamonti and Ghiselli, 2009;

Zouros, 2013 for reviews). Species showing DUI display two different kinds of mitochondrial genomes, i.e. male (M) and female (F) mitogenomes. While females have only the F genome, males are heteroplasmic and possess F and M genomes, which the F type predominating in somatic tissues and the M one in gonads (Stewart et al., 1995; Sutherland et al., 1998). To date, the vast majority of species with DUI which have been reported belong to the orders Mytiloida, Nuculanoida, Unionoida and Veneroida (Gusman et al., 2016), including the wedge clam *D. trunculus* (Theologidis et al., 2008).

In this study, we determine, for the first time, the complete female mitochondrial (mt) genome sequences in four species of *Donax* from the Iberian Peninsula, and compare them with those of other marine bivalves. In addition, the four newly sequenced mitogenomes, together with the veneroids mt genomes available in GenBank, were used to construct the phylogenetic relationships in the Veneroida order. This work should be of importance not only for better understanding the phylogenetic relationships within the Veneroida order, but also for the development of genetic markers useful in wedge clams aquaculture and restoration effects, as well as for the identification and authentication of commercial species.

Materials and methods

Ethics Statement

All clams handling was conducted in accordance with the guidelines and regulations established by the University of A Coruña and the local governments. Field sampling did not require specific permissions but was in accordance with general governmental regulations. No endangered or protected species were involved.

Samples collection and DNA extraction

Given that DUI has been described in *D. trunculus* (Theologidis et al., 2008) and we have found evidence for it in *D. vittatus* and *D. semistriatus* (Fernández-Pérez et al., 2017a), and since the goal of our work was on female mtDNA, we used somatic cells of female specimens as the only source for mtDNA sequencing. Therefore, each of the four *Donax* complete mt genomes sequenced here was obtained from a single female specimen in each species, sampled at natural beds. The *D. trunculus* sample was collected at Corrubedo (A Coruña, northwestern Spain) while the *D. semistriatus*, *D. variegatus* and *D. vittatus* samples came from the Portuguese coast (Table 1). Gender determination was performed on each individual by microscopic examination of gametogenic tissue from the visceral mass, and was based on the presence of eggs or sperm. Specimens were taxonomically identified using Pereira et al. (2012) and Nantón et al. (2015) molecular protocols developed in our laboratory. Voucher specimens and their shells were deposited at the malacology collections of the Museo Nacional de Ciencias Naturales (MNCN), Madrid (Spain) (Table 1).

Table 1. Sampling details.

Species	Sampling site	Country	Latitude	Longitude	Voucher no.
D. semistriatus	Monte Gordo	Portugal	37.167	-7.503	15.07/13263
D. trunculus	Corrubedo	Spain	42.566	-9.039	15.07/13264
D. variegatus	Monte Gordo	Portugal	37.100	-7.633	15.07/13265
D. vittatus	Mira-Vagueira	Portugal	40.614	-8.769	15.07/13266

Total genomic DNA was extracted from about 40 mg of ethanol-preserved foot muscle tissue of female specimens using DNAeasy Blood and Tissue Kit (Qiagen, Germany) following manufacturer´s instructions with only a minor modification, namely EB (10mM Tris-Cl, pH 8.5) rather than AE (10mM Tris-Cl, 0.5 mM EDTA, pH 9.0) buffer was used to avoid possible interference of EDTA with Nextera enzyme.

Molecular procedures and sequencing

The purified genomic DNA was assessed by spectrophotometry (NanoDrop ND-1000, Technologies, Inc.), fluorometry (Qubit HS, Invitrogen, USA) and 1% agarose gel electrophoresis. After quality controls, four libraries (one per species) were prepared using the NEBNext® Ultra™ DNA Library Prep Kit for Illumina® and sequenced in the Illumina HiSeq 4000 platform yielding about 20 Gb data for *D. vittatus* and 10 Gb for each of the three other species, subdivided into 2x150 nt paired-end reads.

Mitogenome assembly and annotation

The mt genomes were reconstructed using 2x1,000,000 reads per species with the MITObim assembler (Hahn et al., 2013). We performed a first assembly with the -quick option, which resulted in a partial mt genome sequence of about 10,000 bp. In order to get the complete sequence, we extracted the sequence of the COI gene from the previous assembly to be used as starting sequence in MITObim with the -seed option. This yielded sequence of about 17,000 bp whose quality and completeness were assessed on the basis of their average coverage along their whole length, by mapping, in each species, the same 2x1,000,000 reads used in the assembly against the inferred mitogenome sequence. For this purpose, we used the SSAHA2 software (Ning et al., 2001) with a minimum score of 100. Then we extracted coverage information from these mapping using pysamstats (available at: http://github.com/alimanfoo/pysamstats).

The mt genomes were annotated using the MITOS Web Server (Bernt et al., 2013) applying the invertebrate mitochondrial genetic code and followed by manual validation of the coding regions using the NCBI ORF Finder (https://www.ncbi.nlm.nih.gov/orffinder/). Based on ORF Finder result, the sqn files generated from MITOS were edited and submitted to NCBI. The annotations of PCGs were refined, while the annotations of tRNA genes were kept unchanged. tRNA genes were detected using MITOS, tRNAScan-SE v.2.0 (Lowe and Chan, 2016) and ARWEN v.1.2 (Laslett and Canbäck, 2008); and secondary structures of tRNAs were inferred using MITOS in default search mode. Mitogenome maps were drawn using GenomeVx online tool (Conant and

Wolfe, 2008) followed by manual modification. Repeat sequence patterns in the longest non-coding region (NCR) were checked using the web-based software server Tandem Repeats Finder (http://tandem.bu.edu/trf/trf.html) (Benson, 1999).

Phylogenetic analyses

To investigate the phylogenetic relationships between species of the Veneroida order, we used the 33 mitogenomes currently available in GenBank (last accessed 17 January 2017), in addition to the four newly determined in this work. *Lucinella divaricata* and *Loripes lacteus*, belonging to the order Lucinoida, were used as outgroups (Table 2). Owing to the fact that a lack of the *Atpase subunit 8* (*atp8*) gene has been reported in some bivalve species, we investigated the possibility that its presence might have gone unnoticed in these species by actively searching for *atp8* sequence in an annotation with MITOS and aligning with other mitogenomes using Geneious Pro v.4.8.5 (Drummond et al., 2009). We found the *atp8* gene in eight species where previous analyses had concluded the absence of this gene. The alignment of the amino acid sequences for each of the 13 mitochondrial PCGs was performed with the MUSCLE plug-in in Geneious Pro v.4.8.5 (Drummond et al., 2009) with default parameters. We removed poorly aligned regions with Gblocks v.0.91b (Castreana, 2000), with options allowing gaps for all positions and 85% of the number of sequences for flanking positions. The 13 separate amino acid sequence alignments were then concatenated into a single large dataset consisting of 2617 sites (see Supporting Information on CD-ROM: S1 File).

Table 2. List of the species whose mitogenome sequences were used in the phylogenetic analysis.

Species	Classification	GB Accession no.	Reference
Donax semistriatus	Veneroida; Tellinoidea; Donacidae	KY780363	This study
Donax trunculus	Veneroida; Tellinoidea; Donacidae	KY780364	This study
Donax variegatus	Veneroida; Tellinoidea; Donacidae	KY780365	This study
Donax vittatus	Veneroida; Tellinoidea; Donacidae	KY780366	This study
Macoma balthica	Veneroida; Tellinoidea; Tellinidae	KM373200	Saunier et al., 2014
Moerella iridescens	Veneroida; Tellinoidea; Tellinidae	JN398362	Yuan et al., 2012a
Nuttallia olivacea	Veneroida; Tellinoidea; Psammobiidae	JN398364	Yuan et al., 2012a
Semele scabra	Veneroida; Tellinoidea; Semelidae	JN398365	Yuan et al., 2012a
Solecurtus divaricatus	Veneroida; Tellinoidea; Solecurtidae	JN398367	Yuan et al., 2012a
Soletellina diphos	Veneroida; Tellinoidea; Psammobiidae	JN398363	Yuan et al., 2012a
Sinonovacula constricta	Veneroida; Solenoidea; Pharidae	JN398366	Yuan et al., 2012a
Solen grandis	Veneroida; Solenoidea; Solenidae	HQ703012	Yuan et al., 2012b
Solen strictus	Veneroida; Solenoidea; Solenidae	JN786377	Yuan et al., 2012c
Cyclina sinensis	Veneroida; Veneroidea; Veneridae	KU097333	Dong et al., 2016
Meretrix lamarckii	Veneroida; Veneroidea; Veneridae	GU071281	Wang et al., 2011
Meretrix lusoria	Veneroida; Veneroidea; Veneridae	GQ903339	Wang et al., 2010
Meretrix lyrata	Veneroida; Veneroidea; Veneridae	KC832317	Wu et al., 2014
Meretrix meretrix	Veneroida; Veneroidea; Veneridae	GQ463598	He et al., 2011
Meretrix petechialis	Veneroida; Veneroidea; Veneridae	EU145977	Ren et al., 2009
Paphia amabilis	Veneroida; Veneroidea; Veneridae	JF969276	Xu et al., 2012
Paphia euglypta	Veneroida; Veneroidea; Veneridae	GU269271	Xu et al., 2010
Paphia textile	Veneroida; Veneroidea; Veneridae	JF969277	Xu et al., 2012
Paphia undulata	Veneroida; Veneroidea; Veneridae	JF969278	Xu et al., 2012
Ruditapes philippinarum	Veneroida; Veneroidea; Veneridae	KT001084	Hwang et al., 2016
Saxidomus purpuratus	Veneroida; Veneroidea; Veneridae	KP419933	Bao et al., 2016
Acanthocardia tuberculata	Veneroida; Cardioidea; Cardiidae	DQ632743	Dreyer and Steiner, 2006
Fulvia mutica	Veneroida; Cardioidea; Cardiidae	NC_022194	Imanishi et al., 2013
Tridacna squamosa	Veneroida; Cardioidea; Cardiidae	KP205428	Gan et al., 2016a
Corbicula fluminea	Veneroida; Corbiculoidea; Corbiculidae	KX254564	Tao et al., unpublished
Geloina coaxans	Veneroida; Corbiculoidea; Corbiculidae	KP999913	Zhou, unpublished
Calyptogena magnifica	Veneroida; Glossoidea; Vesicomyidae	KR862368	Liu et al., 2016
Arctica islandica	Veneroida; Arcticoidea; Arcticidae	KF363951	Glöckner et al., 2013
Coelomactra antiquata	Veneroida; Mactroidea; Mactricidae	KC503290	Yuan et al., 2013
Lutraria rhynchaena	Veneroida; Mactroidea; Mactricidae	NC_023384	Gan et al., 2016b
Mactra chinensis	Veneroida; Mactroidea; Mactricidae	KJ754823	Shen et al., 2014
Lucinella divaricata	Lucinoida; Lucinoidea; Lucinidae	EF043342	Dreyer et al., unpublished
Loripes lacteus	Lucinoida; Lucinoidea; Lucinidae	EF043341	Dreyer et al., unpublished

Phylogenetic analyses were performed under Maximun Likelihood (ML) using RaxML (Stamatakis et al., 2008) in a web server (http://embnet.vital-it.ch/raxml-bb/) and Bayesian inference (BI) using MrBayes v3.2.6 (Ronquist et al., 2012) and PhyloBayes (Lartillot and Philippe, 2004). The best fit models of amino acid evolution were chosen by ProtTest v.3.4.2 (Darriba et al., 2011), with default settings, based on Akaike Information Criterion (AIC). The optimal chosen methods were: LG + I + G + F for *cox1*, *cox3* and *nad5* genes; LG + G + F for *cox2*, *nad6* and *atp8*; MtArt + I + G + F for *cob*, *atp6*, *nad2* and *nad4*; MtArt + I + G + F for *nad1*, *nad3* and *nad4l*. However, as the MtArt evolutionary model is not available in MrBayes, the LG model (the second best-fit model according to ProtTest) was used in Bayesian analysis, being therefore: LG + I + G + F for *cox1*, *cox3*, *cob*, *nad1*, *nad2*, *nad3*, *nad4* and *nad5* genes; LG + G + F for *cox2*, *atp6*, *nad6* and *atp8*; and LG + G for *nad4l*. The ML analyses consisted of 1000 bootstrap iterations using the CAT model for each partition. BI analysis consisted of two independent Markov chain Monte Carlo (MCMC) runs, each comprising four linked chains (one cold and three heated; as default settings). They were performed for 1,000,000 generations, sampling every 100 generations to allow adequate time for convergence. The convergence of the two runs was assessed by stopping the analysis when the average standard deviation was below 0.01 (stoprule= yes and stopval= 0.01 in the mcmc command). 1,000,000 generations were enough to reach adequate average standard deviation (<0.01). By default, the first 25% trees were discarded as burn-in. BI analyses were also conducted at the amino-acid level using the CAT + GTR model in PhyloBayes (Lartillot and Philippe, 2004). Two independent MCMC analyses were run in parallel for 4,000 generations. The first 1,000 samples were discarded as burn-in. From the remaining samples, we sampled a tree every 10 cycles to compute a consensus tree. The convergence between the two chains were considered acceptable when the maxdiff parameter was below 0.3 (maxdiff = 0.218586) and the minimum effective size (MES) was >50 (MES = 64).

Results and discussion

Sequencing and mitogenome assembly

A total of about 92,000,000 paired reads (2x150 nt) were obtained for *D. semistriatus*, about 85,000,000 for *D. trunculus*, about 82,000,000 for *D. variegatus* and about 185,000,000 for *D. vittatus*. We selected 2x1,000,000 reads that were used to assemble the mitogenome in each species, yielding average coverages of 45x in *D. semistriatus*, 31x in *D. trunculus*, 37x in *D. variegatus*, and 58x in *D. vittatus*. Coverage profiles were uniform along the mt genomes (see Supporting Information S2 Figure).

Genome composition

The mitogenomes of the four *Donax* species sequenced in this study were circular molecules, as revealed by the MITObim assembly. They are composed of 37 genes: 13 PCGs (including the *atp8* gene), two ribosomal RNA genes and 22 transfer RNA genes (Figure 1). Their main structural features are summarized in Table 3. The complete mt genomes of *D. semistriatus*, *D. trunculus*, *D. variegatus* and *D. vittatus* vary in size from 17,044 bp (*D. semistriatus*) to 17,365 bp (*D. trunculus*). Length differences are mostly due to the size variation of the non-coding region. The A+T content of the four mitogenomes ranges from 58.9% (*D. trunculus*) to 63.5% (*D. vittatus*). Although gene organization is known to vary extensively, even among species from the same genus (Gissi et al., 2008; Wu et al., 2012; Xu et al., 2012), all four complete *Donax* mt genomes showed the same gene order and they are located on the "+" strand, likewise in *Macoma balthica*, other member of the Tellinoidea superfamily for which the whole mt genome is available (Saunier et al., 2014). The only difference was noted in the location of the longest NCR which, in *M. balthica*, is situated between *rrnS* and *tRNA-Met*, whereas in *Donax* clams it is located between *cob* and *cox2* genes (Figure 1). Therefore, in consistency with the highly rearranged gene order in bivalves, the longest NCR is not conserved at the same position among bivalve mt genomes (Yuan et al., 2012a; Plazzi et al., 2016).

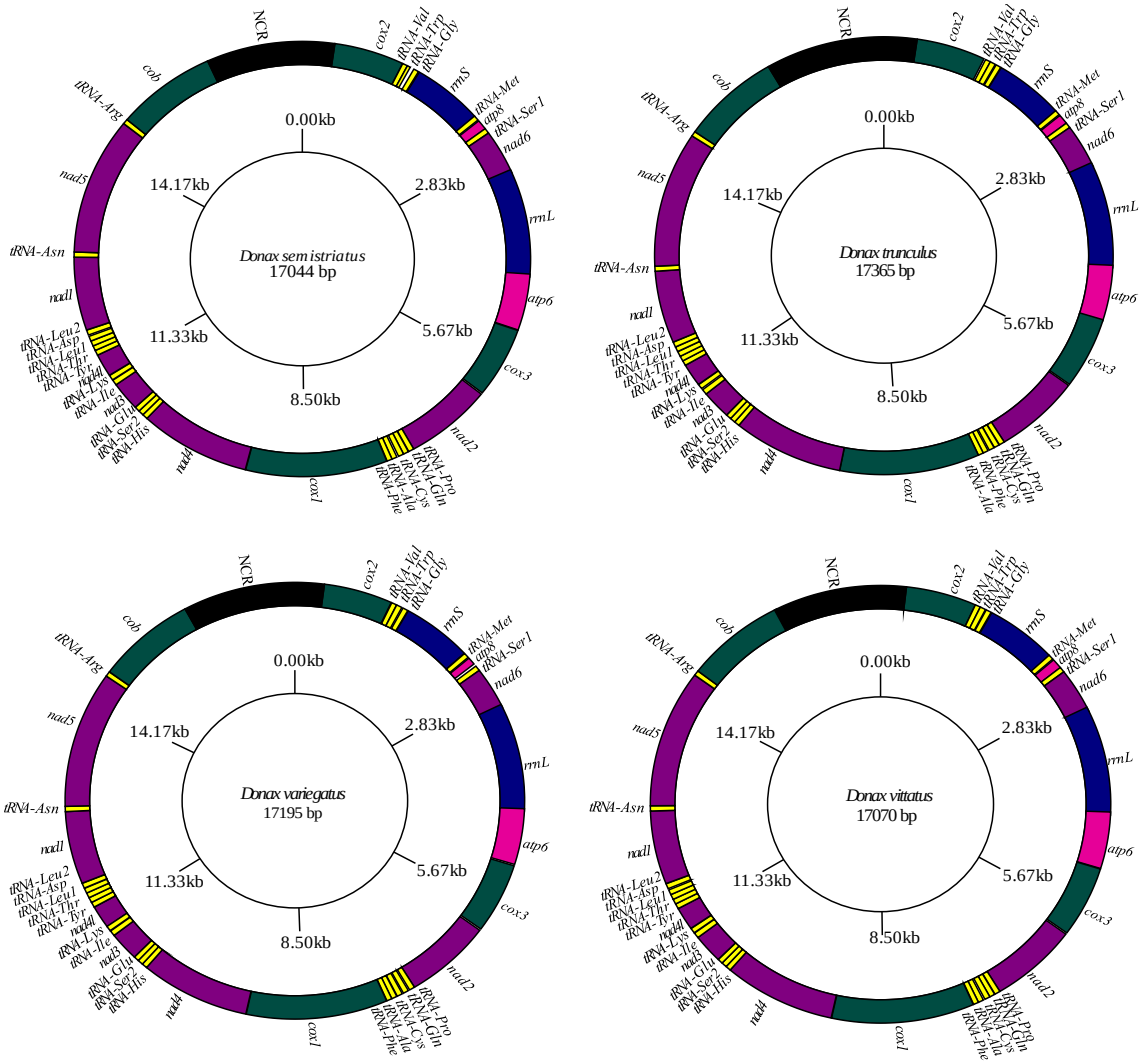


Figure 1. Maps of the mitochondrial genomes of *Donax* species. Genome lengths are shown in the middle of each map, genes are all on "+" strand and NCR indicates the longest non-coding region.

Table 3. Main structural features of the four sequenced mt genomes in this study.

	Donax semistriatus	***Donax trunculus***	***Donax variegatus***	***Donax vittatus***
Total length	17044	17365	17195	17070
A+T%	61.9	58.9	60.4	63.5
cox2	846 (ATG/TAA)	846 (ATG/TAA)	831 (ATG/TAG)	846 (ATG/TAA)
tRNA-Val	62	64	64	64
tRNA-Trp	69	68	69	69
tRNA-Gly	64	65	66	66
rrnS	863	860	859	865
tRNA-Met	65	65	65	65
atp8	126 (ATG/TAG)	126 (ATG/TAG)	126 (ATG/TAA)	126 (ATG/TAG)
tRNA-Ser1	68	69	69	68
nad6	576 (ATG/TAG)	573 (ATG/TAA)	540 (ATG/TAA)	576 (ATG/TAG)
rrnL	1373	1367	1383	1386
atp6	714 (ATG/TAA)	714 (ATG/TAA)	711 (ATG/TAG)	714 (ATG/TAG)
cox3	891 (ATG/TAG)	915 (ATA/TAA)	891 (ATG/TAG)	891 (ATG/TAG)
nad2	1062 (ATG/TAA)	1062 (TTG/TAG)	1062 (ATG/TAA)	1062 (ATG/TAA)
tRNA-Pro	67	68	67	67
tRNA-Gln	65	66	67	65
tRNA-Cys	66	66	68	66
tRNA-Ala	64	65	66	65
tRNA-Phe	63	64	64	63
cox1	1710 (ATG/TAA)	1710 (ATG/TAA)	1710 (ATG/TAA)	1710 (ATG/TAA)
nad4	1347 (TTG/TAA)	1356 (TTG/TAG)	1332 (TTG/TAA)	1347 (TTG/TAA)
tRNA-His	66	66	66	64
tRNA-Ser2	66	65	66	65
tRNA-Glu	63	64	63	63
nad3	363 (ATG/TAA)	363 (ATG/TAA)	363 (ATG/TAA)	363 (ATG/TAA)
tRNA-Ile	69	69	69	69
tRNA-Lys	65	63	64	64
nad4l	288 (TTG/TAG)	288 (TTG/TAG)	288 (ATG/TAA)	288 (TTG/TAG)
tRNA-Tyr	64	64	66	65
tRNA-Thr	63	65	66	64
tRNA-Leu1	65	66	65	65
tRNA-Asp	63	62	64	63
tRNA-Leu2	65	66	65	66
nad1	924 (ATG/TAG)	924 (ATG/TAA)	924 (ATG/TAG)	924 (ATG/TAG)
tRNA-Asn	65	64	66	65
nad5	1734 (ATG/TAA)	1734 (GTG/TAG)	1734 (ATG/TAA)	1734 (ATG/TAA)
tRNA-Arg	63	63	63	63
cob	1215 (ATG/TAA)	1218 (ATA/TAA)	1206 (ATG/TAA)	1215 (ATG/TAA)

For each mt genome, total length (in bp), the percent of overall A+T content, and size (bp) of the protein coding genes (start and stop codons in brackets), tRNAs, *rrnL* and *rrnS* are given.

Protein coding genes

The typical 13 PCGs were identified in the four new mitogenomes analyzed here, including the *atp8* gene, which had been reported as missing in several bivalve species (Smith and Snyder, 2007; Ren et al., 2010; Meng et al., 2012; Yuan et al., 2012a,b,c; Shen et al., 2016b), but subsequent analyses found its presence in several of them (Dreyer and Steiner, 2006; Breton et al., 2010; Śmietanka et al., 2010; Wang et al., 2010; Wu et al., 2012; Saunier et al., 2014; Plazzi et al., 2016). It was suggested that the short and variable length of this protein, along with its high variation in amino acid composition, might hinder the finding of this gene due to annotation difficulties (Gissi et al., 2008). However, using the same bioinformatic approach employed in *Donax* species, we found the *atp8* gene in publicly available mitogenome sequences of most Veneroida order species available in the databases (Table 4). Moreover, we found other discrepancies with GenBank annotations. The *tRNA-Lys* annotation for *Mactra chinensis* (KJ754823) was modified (from 9945-10028 to 13611-13677) and in the following cases, the previous *rrnS* annotations were also edited: *rrnS* for *M. meretrix* (GQ463598) and *M. petechialis* (EU145977) were edited from 7093-8673 to 7089-8569; for *C. antiquata* (KC503290) from 7898-9197 to 7898-9096; and for *L. rhynchaena* (NC_023384) from 6870-8244 to 6870-8161.

The location of the *atp8* gene within the mitogenome is the same in the eight species of the Tellinoidea superfamily (all four *Donax* species, *M. balthica*, *M. iridescens*, *S. divaricatus* and *S. diphos*), i.e. between *tRNA-Met* and *tRNA-Ser1*. In *Donax* species, this short gene encoded a 42 amino acids protein starting with methionine (ATG, in the four species) and ending with a stop codon (TAG in *D. semistriatus*, *D. trunculus* and *D. vittatus*; or TAA, in *D. variegatus*) (Table 4), so that ATP8 proteins show 83.7% amino acid identity among species. Finally, it has been suggested that the *atp6* and *atp8* genes are adjacent in most animal mitochondrial genomes, often with overlapping reading frames (Boore, 1999). However, in *Donax* species *atp6* and *atp8* genes are physically separated by 1,917 (*D. trunculus*) – 1,928 bp (*D. vittatus*). Likewise, these two genes also fail to be adjacent in the mitogenome of other heterodont bivalves, such as *Hiatella arctica* (Dreyer and Steiner, 2006), *M. balthica* (Saunier et al., 2014) and *Meretrix lamarckii*

(Bettinazzi et al., 2016). On the contrary, they are adjacent in the Unionidae (Breton et al., 2009) and Solemydae (Plazzi et al., 2013), as well as in basal molluscs like *Chaetoderma nitidilum* (EF211990) and *Katharina tunicata* (Boore and Brown, 1994). This suggests that the association of these genes might be an example of an ancestral state that has later been lost in derived bivalves.

Total length of the 13 PCGs ranged from 11,718 bp (*D. variegatus*) to 11,829 bp (*D. trunculus*), accounting for 68.1 – 69.2% of its total mt genome length. The longest PCG is *nad5,* with a size of 1,734 bp (577 aa), whereas *nad2*, *cox1, nad4* and *cob* exceed 1,000 bp. However, *nad3* and *nad4l* genes are shorter than 400 bp and *atp8* gene is the shortest PGC with 126 bp (41 aa). These features are similar to those previously reported in *M. balthica* (Saunier et al., 2014) and five other species of the Tellinoidea superfamily (*Moerella iridescens, Sanguilonaria diphos, Sanguinolaria olivacea, Semele scabra* and *Solecurtus divaricatus*) (Yuan et al., 2012a).

The ATN conventional start codon is used in most PCGs (ATG, N=41; ATA, N=2; the last codon being classically found in the invertebrate mitochondrial genetic code, particularly in bivalves (Saunier et al., 2014)). However, like most invertebrate mt genomes, *Donax* mtDNA shows alternative start codons, and some PCGs start with NTG codons (TTG, N=8; GTG, N=1). In contrast, the observed stop codons are TAA (N=32) and TAG (N=20), and all 13 PCGs of the four mt genomes end in a full termination codon.

Table 4. Presence of the *atp8* gene in the mitogenomes of the Veneroida order.

Species	*atp8*	Size	Position	Start/Stop codons	Reference
Donax semistriatus	Yes	126	2396-2521	ATG/TAG	This study
Donax trunculus	Yes	126	2419-2544	ATG/TAG	This study
Donax variegatus	Yes	126	2352-2477	ATG/TAA	This study
Donax vittatus	Yes	126	2310-2435	ATG/TAG	This study
Macoma balthica	Yes	129	75-203	ATT/TAA	Saunier et al., 2014
Moerella iridescens	Yes	132	11625-11756	ATA/TAG	Plazzi et al., 2016
Nuttallia olivacea	Yes	132	12930-13061	ATA/TAG	Plazzi et al., 2016
Semele scabra	Yes	129	11969-12100	ATT/TAA	Plazzi et al., 2016
Solecurtus divaricatus	Yes	135	11321-11455	GTG/TAG	Plazzi et al., 2016
Soletellina diphos	Yes	135	11214-11342	GTG/TAG	Plazzi et al., 2016
Sinonovacula constricta	Yes	114	14288-14401	ATG/TAA	This study
Solen grandis	Yes	114	13703-13816	GTG/TAG	This study
Solen strictus	Yes	114	13473-13586	ATG/TAG	This study
Cyclina sinensis	Yes	117	8568-8684	ATG/TAA	Dong et al., 2016
Meretrix lamarckii	Yes	120	8835-8954	ATG/TAA	Wang et al., 2011
Meretrix lusoria	Yes	120	8642-8761	ATG/TAG	Wang et al., 2010
Meretrix lyrata	Yes	120	8753-8872	ATG/TAG	Wu et al., 2014
Meretrix meretrix	Yes	141	8532-8672	ATA/TAG	Plazzi et al., 2016
Meretrix petechialis	Yes	141	8532-8672	ATA/TAG	Plazzi et al., 2016
Paphia amabilis	Yes	114	14035-14148	ATG/TAG	Xu et al., 2012
Paphia euglypta	Yes	117	12994-13110	ATA/TAA	Plazzi et al., 2016
Paphia textile	Yes	114	13019-13132	ATG/TAA	Xu et al., 2012
Paphia undulata	Yes	114	12642-12755	ATG/TAA	Xu et al., 2012
Ruditapes philippinarum	Yes	120	5968-6087	ATT/TAG	Plazzi et al., 2016
Saxidomus purpuratus	Yes	117	9557-9673	ATG/TAA	This study
Acanthocardia tuberculata	Yes	103	12546-12648	GTG/CCT	Plazzi et al., 2016
Fulvia mutica	Yes	114	11341-11454	TTG/TAA	Imanishi et al., 2013
Tridacna squamosa	Yes	117	8525-8641	ATG/TAG	This study
Corbicula fluminea	Yes	114	5480-5593	ATG/TAA	Tao et al., unpublished
Geloina coaxans	Yes	114	12249-12362	TTG/TAG	Zhou, unpublished
Calyptogena magnifica	Yes	114	5440-5553	ATG/TAA	Liu et al., 2016
Arctica islandica	Yes	151	10343-10493	TTG/AGT	Plazzi et al., 2016
Coelomactra antiquata	Yes	114	9097-9210	ATG/TAA	This study
Lutraria rhynchaena	Yes	118	8162-8275	ATG/TAA	This study
Mactra chinensis	Yes	114	10000-10113	ATG/TAG	This study
Lucinella divaricata	Yes	114	15861-15974	ATT/TAA	Dreyer et al., unpublished
Loripes lacteus	Yes	118	14442-14589	ATT/ACT	Dreyer et al., unpublished

For each *atp8* sequence, size (bp), position (from-to), and start and stop codons.

Transfer and ribosomal RNA genes

Standard rRNAs were found in the four mt genomes of *Donax* species analyzed here. The small-subunit ribosomal RNA (*rrnS*) was flanked by *tRNA-Gly* and *tRNA-Met* in all four mt genomes, and its size ranged from 859 bp (*D. variegatus*) to 865 bp (*D. vittatus*), with A+T content between 63.8 (*D. semistriatus*) and 68.5% (*D. vittatus*). On the other hand, the large-subunit ribosomal RNA (*rrnL*) was located between *nad6* and *atp6*, just like in *M. balthica* (Saunier et al., 2014), *M. iridescens, S. diphos, S. olivacea, S. scabra, S. constricta* and *S. divaricatus* (Yuan et al., 2012a). Its size varied from 1,367 bp (*D. semistriatus*) to 1,386 bp (*D. vittatus*), and its A+T content ranged between 63.5 (*D. variegatus*) and 67.2% (*D. semistriatus*).

Twenty-two discrete nucleotide sequences (ranging from 62 to 69 bp) were predicted to fold into the typical secondary structures of tRNAs (see Supporting Information S3-S6 Figures). The predicted structures of tRNA genes showed cloverleaf shape with four arms in the four species, although some of them exhibited folding differences. Sixteen *tRNAs* showed a small supplemental stem loop (four in *D. semistriatus*: *tRNA-Pro*, *tRNA-Phe*, *tRNA-Ile* and *tRNA-Leu2*; two in *D. trunculus*: *tRNA-Ile* and *tRNA-Thr*; six in *D. variegatus*: *tRNA-Val*, *tRNA-Pro*, *tRNA-Gln*, *tRNA-His*, *tRNA-Ile* and *tRNA-Arg*; and four in *D. vittatus*: *tRNA-Pro*, *tRNA-Phe*, *tRNA-Ile* and *tRNA-Leu2*). Seven *tRNAs* showed no terminal TΨC loop (three in *D. semistriatus*: *tRNA-His*, *tRNA-Thr* and *tRNA-Arg*; one in *D. trunculus*: *tRNA-Asn*; and three in *D. vittatus*: *tRNA-His*, *tRNA-Thr* and *tRNA-Asp*). In addition, *tRNA-Ser2* in *D. trunculus* showed the dihydrouracil (DHU) stem replaced by a big DHU loop. Finally, the single unpaired nucleotide, which is usually present at the 5´end in other tRNAs, appeared at the 3´end in *tRNA-Tyr*, with the only exception of *D. variegatus* where this tRNA lacks this unpaired nucleotide. These features have previously been found in mtDNAs of other bivalve species, such as *M. balthica* (Saunier et al., 2014) and *M. lamarckii* (Bettinazzi et al., 2016).

Non-coding regions

As in most bivalves, the four species of the genus *Donax* analyzed here contained a large number of NCRs. The number of intergenic sequences varied from 17 (*D. trunculus* and *D. vittatus*) to 22 (*D. variegatus*), with 1,679 bp (representing 9.9% of the whole mitogenome) in *D. semistriatus* to 1,985 bp (11.4% of the mt genome) in *D. trunculus* (Table 5). The longest NCR was located between *cob* and *cox2* genes in the four species, with length ranging from 1,549 bp (*D. semistriatus*) to 1,863 bp (*D. trunculus*). The other NCRs ranged fom 1 to 21 bp. The longest NCR is thought to contain the Control Region (CR) because it presents some peculiar patterns, such as AT-rich or tandem repeats, believed to play a role in initiating and/or regulating mitochondrial transcription and replication (Wolstenholme, 1992; Faber and Stepien, 1998; Saito et al., 2005). The A+T content of the longest NCR in each mt genome was higher (*D. semistriatus, D. variegatus and D. vittatus)* or slightly lower (*D. trunculus*) than that of the whole mt genome (Table 5).

Table 5. Comparison of non-coding regions (NCRs) within the four mt genomes.

Species	No. of NCR	Total length (bp)	Proportion of the mt genome (%)	Longest NCR	
				Length (bp)	A+T %
Donax semistriatus	18	1679	9.9	1549	66.6
Donax trunculus	17	1985	11.4	1863	51.8
Donax variegatus	22	1869	10.9	1718	62.6
Donax vittatus	17	1697	9.9	1580	67.5

Six tandem repeats were also found in the longest NCRs of the four mt genomes, four of which were distinct tandem repeat units. The first motif consisted of 2.7 nearly identical copies of a 122 bp unit located at positions 48-386 from the 5´-end of the longest NCR in *D. semistriatus*. The second was 2.1 copies of 126 bp located at positions 17042-17309 in *D. trunculus*. In addition, microsatellite-like repeats, $(TA)_{12}$ in *D. semistriatus* and $(TA)_{12}ACACTTGTGA(TA)_{10}$ in *D.*

trunculus, were detected near the 5´-end of the longest NCR. The third tandem repeat consisted in 2.1 copies of 137 bp located between positions 57 and 344 in *D. variegatus*, and the last one included 2 copies of 122 bp located at positions 47-304 in *D. vittatus*. Such long tandem repeats have also been reported in other bivalves of the Veneroida order (Dreyer and Steiner, 2006; Wang et al., 2010; Meng et al., 2012; Yuan et al., 2012a). The study of tandem repeats in the CR is important for the light it sheds on a variety of processes, including the molecular mechanisms arising them and their possible functional implications (Mundy et al., 1996).

Phylogenetic analysis in Veneroida

To further study the relationships among *Donax* species and its position within the Veneroida order, ML and BI trees based on amino acid sequences of 13 concatenated PCGs belonging to 37 species were performed (Figure 2). Tree topologies were congruent and received high support in most nodes, with the exception of *S. scabra*, which showed a less basal position in the PhyloBayes phylogeny ((*M. balthica* + *M. iridescens)* + *S. scabra*) with 0.57 posterior probability as branch support.

We perform here the first phylogeny including the species of the genus *Donax* from the Iberian Peninsula (*D. trunculus, D. semistriatus, D. variegatus* and *D.vittatus*). Our analysis has shown that the four species form a single clade as a sister group to other bivalves of the superfamily Tellinoidea. All ten species of this superfamily belong to five different families and form a strongly supported clade, thus corroborating the monophyly of this superfamily (Taylor et al., 2007; Combosch et al., 2017). Nevertheless, our phylogenetic tree indicated, with high support by BI and ML, that *S. diphos* (Psammobiidae) shows closer relationship with *S. divaricatus* (Solecurtidae), *M. balthica* and *M. iridescens* (Tellinidae), *S. scabra* (Semelidae) and *Donax* species (Donacidae) rather than with *N. olivacea* (Psammobiidae), which implies that these two species (*S. diphos* and *N. olivacea*) do not form monophyletic groups. This result is also reported by Yuan et al. (2012a) and Ozawa et al. (2017), and it is in agreement with the conclusion put forward by Taylor et al. (2007) when analysed familial relationships within Tellinoidea, as

Semelidae, Donacidae and Tellinidae do not form monophyletic groups. Tellinoidea is actually monophyletic, but none of its families are monophyletic (Combosch et al., 2017), suggesting the need for a more exhaustive study within this commercially important marine bivalve clade.

Gene arrangement within mitogenomes is highly conserved in many taxonomic groups. For instance, most vertebrates share the same gene order (Pereira, 2000). However, in other animal groups, like the class Bivalvia, the mitochondrial genome arrangement is more variable (Yuan et al., 2012a; Plazzi et al., 2016). We compare here the gene arrangements of four newly sequenced mitogenomes to other closed related species belonging to Tellinoidea superfamily. This comparison was previously done by Yuan et al. (2012a), without taking into account the *atp8* gene and without including *Donax* species and *M. balthica*, and their results supported the conclusion that comparisons of mitochondrial gene order rearrangements are, to some extent, a useful tool for phylogenetic studies. Seven out of the ten Tellinoidea mitogenomes hitherto analyzed (including the four *Donax* species analyzed by us, *M. balthica, M. iridescens* and *S. divaricatus*) show completely identical gene order, and *S. diphos* only differs in lacking a *tRNA-Phe*. Remarkably, the *atp8* gene shows the same location within the mitogenome of these eight species of the Tellinoidea superfamily, specifically between *tRNA-Met* and *tRNA*-Ser1. This result is consistent with the main phylogenetic conclusions from the 37 mitogenomes analyzed here (see above), and remarks the interest of performing additional full mitogenome sequencing, especially including more veneroid families and subfamilies, with gene order being a useful hallmark helping to clarify phylogenetic relationships within the order.

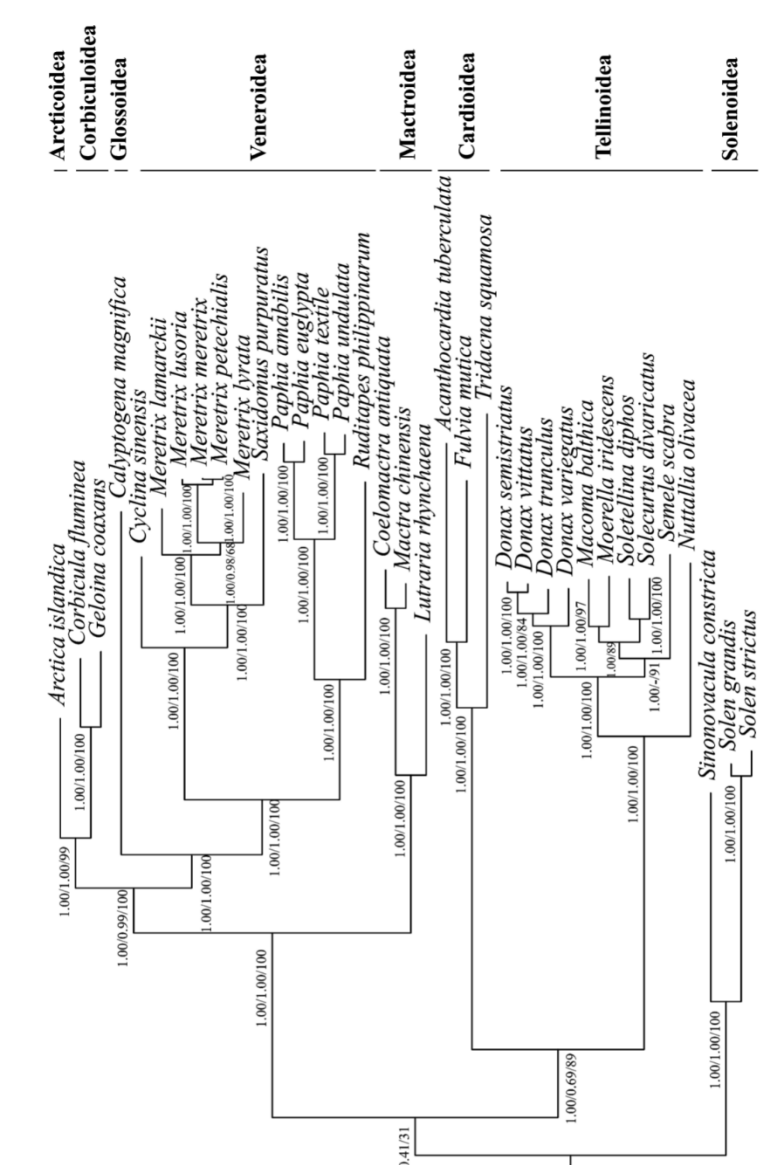


Figure 2. Phylogenetic tree of the Veneroida order based on concatenated amino acids of 13 protein-coding genes. Numbers at the nodes correspond to Bayesian posterior probabilities (left), PhyloBayes posterior probabilities (middle) and ML bootstrap proportions (right). Dash indicates the difference in the position for *S. scabra* in the PhyloBayes phylogeny.

Future implications

This is a basic research work where we describe and characterize, for the first time, the female mitochondrial genome in four bivalve molluscs belonging to the genus *Donax*. This has provided new interesting information for the scientific community which can be feasible for application in aquaculture. In fact, the mtDNA sequences contributed here add significantly useful genetic markers for i) helping to differentiate these commercial food species being morphologically similar, ii) detecting and avoiding fraud, iii) protecting consumer rights and achieving other quality objectives, such as certificate of origin, and iv) for using in population genetics studies and aquaculture stock management in *Donax* species. However, this possible applicability requires a broader work, where the different markers will be tested in a higher number of individuals, not only fresh individuals but also processed, packaged or frozen ones, as well as in a high number of females and males given that male genomes are still not available.

Conclusions

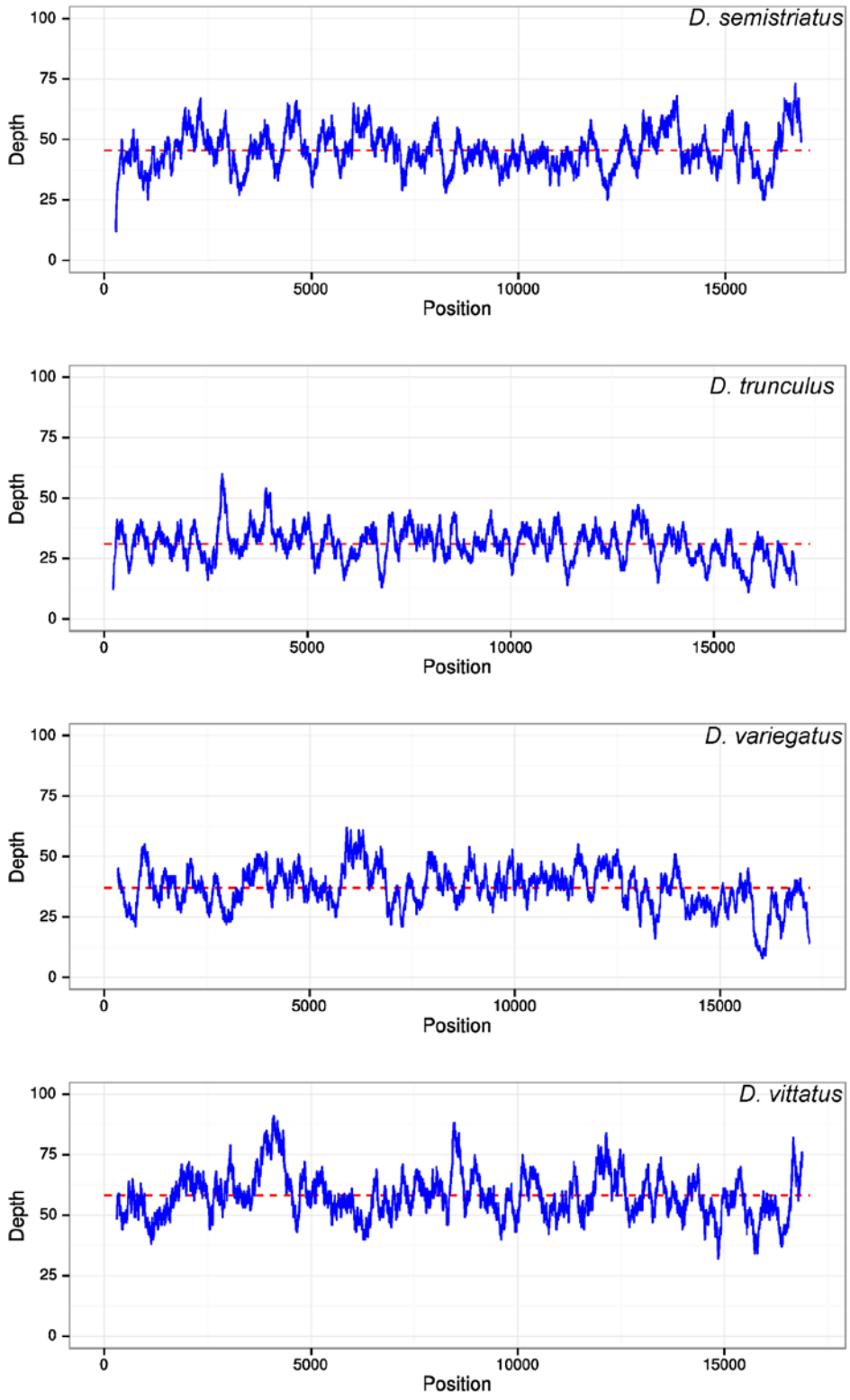
In this study, we determined the complete mt genomes of four bivalve species of the genus *Donax*, which are the first representatives from the family Donacidae being analyzed at this respect. Not only we have increased the number of complete mt genomes sequenced within Veneroida order, but also, we have illustrated the phylogenetic relationships among *Donax* species and their position within this order. Our results demonstrate that the sequencing of complete mitogenomes provides highly valuable information for phylogenetic analysis in bivalves. Furthermore, the mtDNA sequences contributed here add significantly useful genetic markers for use in species identification and authentication, phylogeny, population genetics, and aquaculture stock management in species of *Donax*.

Acknowledgements

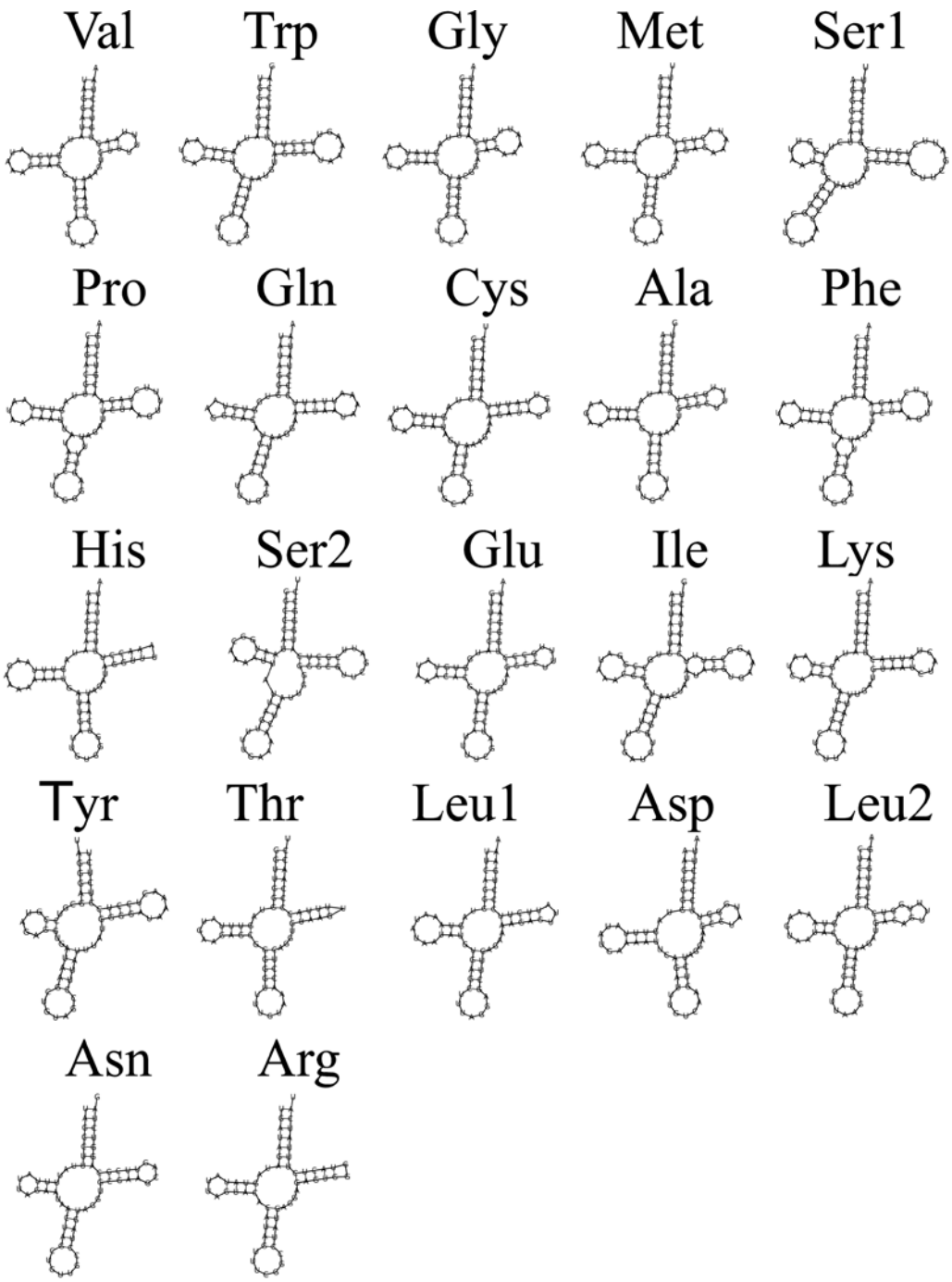
We would like to thank Dra D. Martínez Patiño and S. Nóvoa from Centro de Cultivos Marinos de Ribadeo – CIMA (Xunta de Galicia) and Dr. M.B. Gaspar from Instituto Português do Mar e da Atmosfera – IPMA (Portugal) for providing specimens. The authors wish to thank the three anonymous reviewers for helpful remarks and suggestions that improved the quality of the manuscript. We also thank Dr. D. Huchon for useful comments during manuscript review. This work was supported by the Ministerio de Economía y Competitividad (Spain) through project AGL2016-75288-R AEI/FEDER, UE.

Supporting information

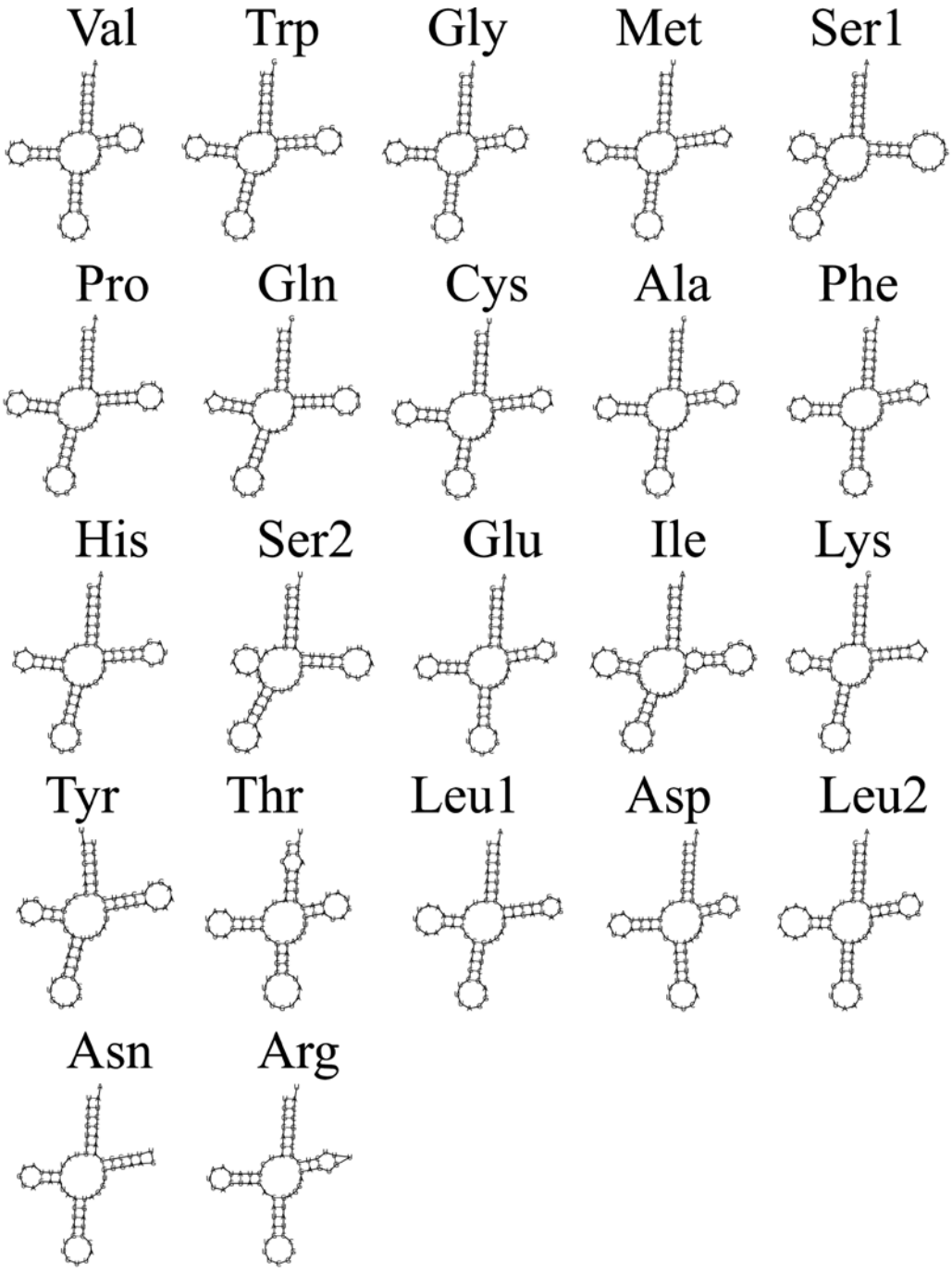
S1 File. The alignment of 37 mitogenomes sequences used for phylogenetic analyses. Sequences include concatenated thirteen mitochondrial protein-coding genes (see on CD-ROM).



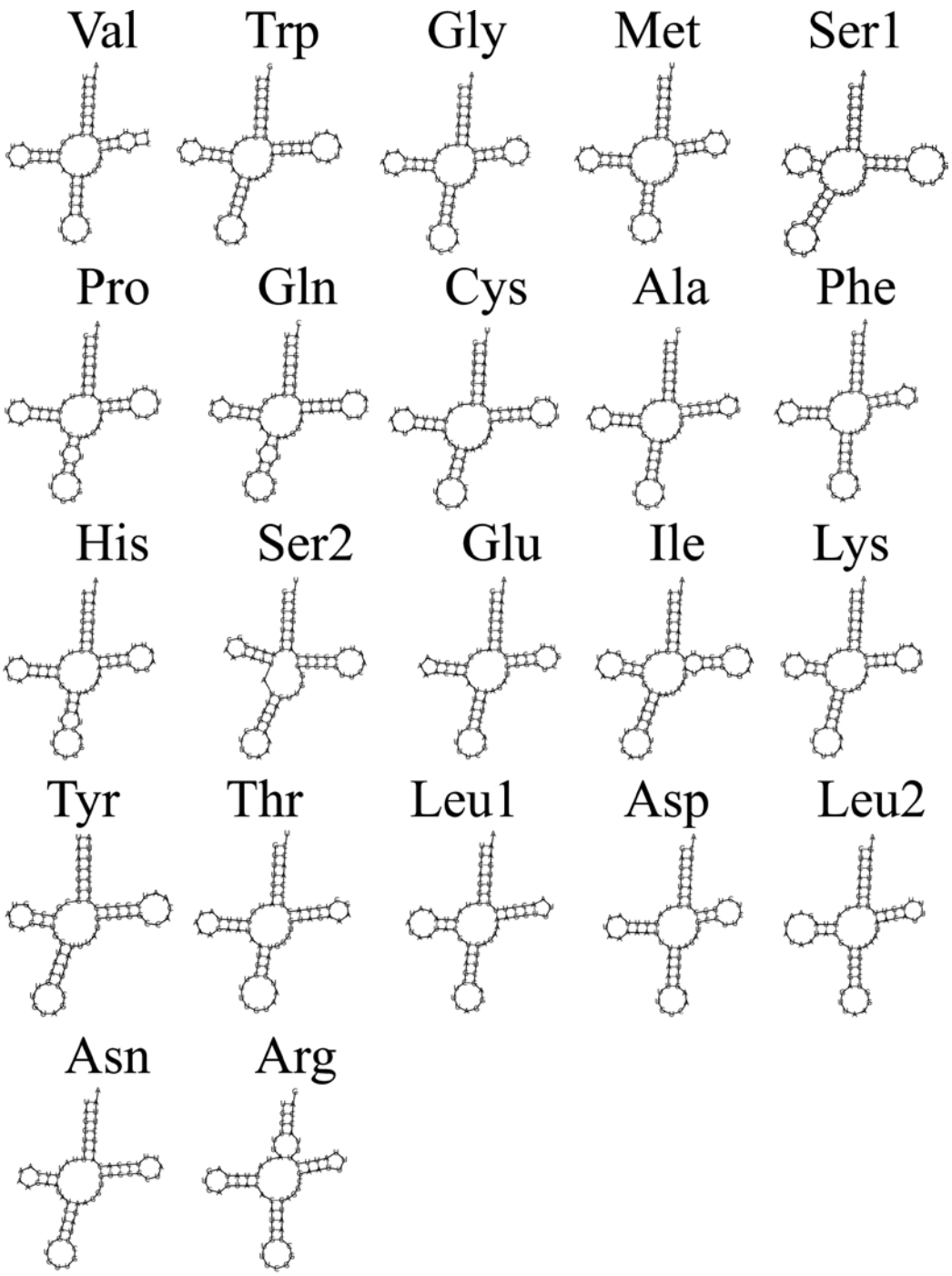
S2 Figure. Coverage profiles for the four newly sequenced mitochondrial genomes. Blue line represents coverage along the mitochondrial sequences for the four *Donax* species. Red dashed lines represent the average coverage values: 45.46x in *D. semistriatus*, 30.94x in *D. trunculus*, 37.12x in *D. variegatus*, and 58.10x in *D. vittatus*.



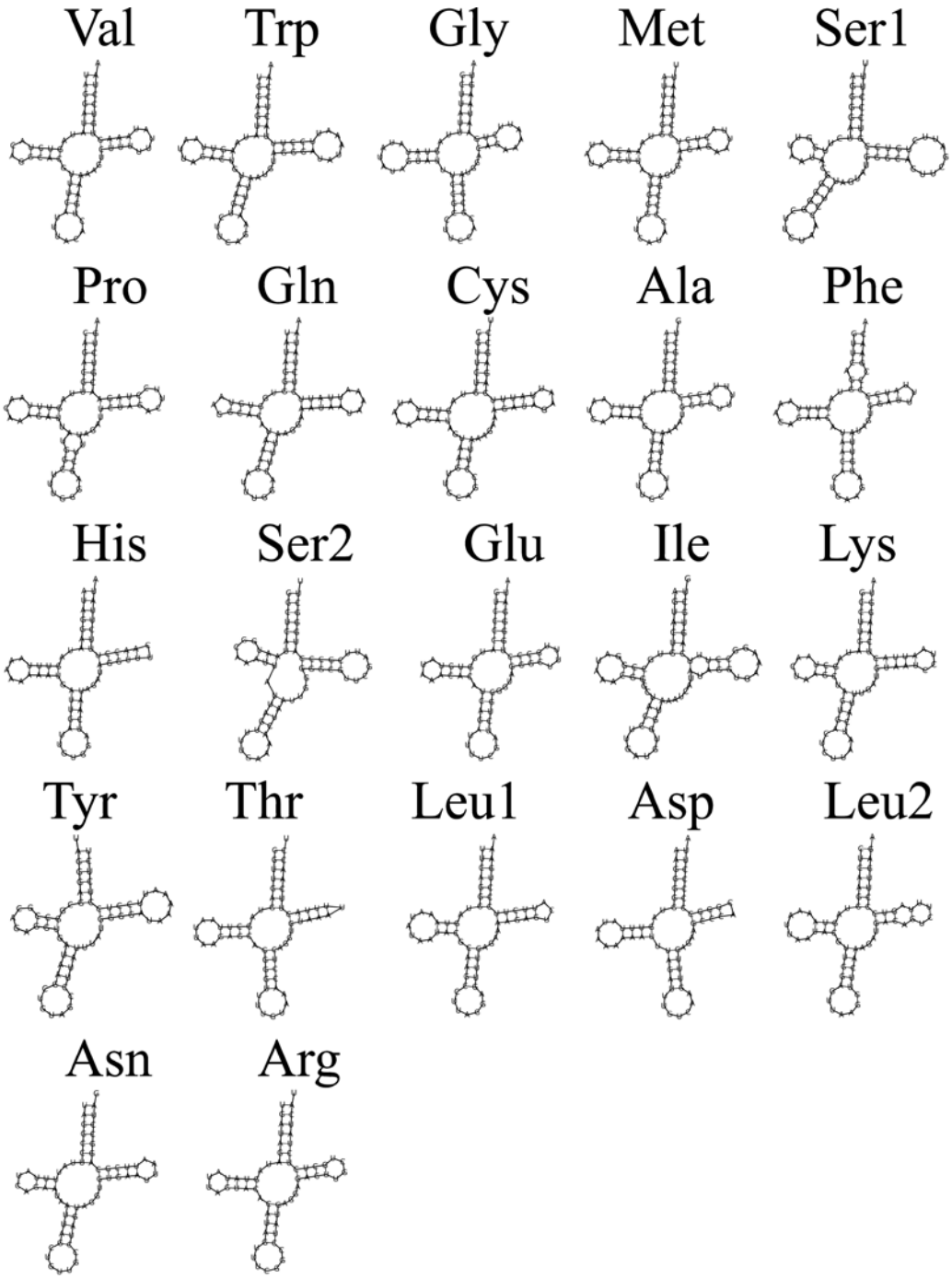
S3 Figure. Predicted tRNA structures in *D. semistriatus*. 22 tRNAs are identified in the mitogenome of *D. semistriatus* and their cloverleaf secondary structures are inferred with MITOS annotation pipeline.



S4 Figure. Predicted tRNA structures in *D. trunculus*. 22 tRNAs are identified in the mitogenome of *D. trunculus* and their cloverleaf secondary structures are inferred with MITOS annotation pipeline.



S5 Figure. Predicted tRNA structures in *D. variegatus*. 22 tRNAs are identified in the mitogenome of *D. variegatus* and their cloverleaf secondary structures are inferred with MITOS annotation pipeline.



S6 Figure. Predicted tRNA structures in *D. vittatus*. 22 tRNAs are identified in the mitogenome of *D. vittatus* and their cloverleaf secondary structures are inferred with MITOS annotation pipeline.

Concluding remarks

Concluding remarks

The main conclusions from the studies reported in this thesis can be summarized in the following sections:

Contributions of this thesis to the knowledge of the genomes of the genus *Donax*

1. Eighteen localities from the Iberian Peninsula analysed with F and M lineages for *D. trunculus* using mitochondrial 16S rDNA and Cytb sequences displayed similar levels of genetic diversity. F_{ST} values for the M type revealed significant differentiation between populations on both sides of the Almeria-Oran front, and between the locality from the Alboran Sea (FU) and localities from the Atlantic Ocean. However, the F type did not detect the differentiation of FU. This makes the M genome sequences more suitable for investigating genetic structure of *D. trunculus* populations.

2. The genetic diversity of *D. vittatus* across the Iberian Peninsula investigated using four mitochondrial (COI, Cytb, 16S F and M types) and three nuclear (H3, 18S and 28S) markers showed high haplotype diversity in combination with a low nucleotide diversity and star-shaped network with a predominant haplotype, indicating a recent population expansion for the examined sampling sites. Furthermore, this study presents for the first time the evidence of DUI in the clams *D. vittatus* and *D. semistriatus*, and its phylogenetic relationship with the other species of the genus *Donax* present in European coasts.

3. This study provides the first 15 microsatellite markers for *D. vittatus* and report estimates of genetic diversity and population differentiation in seven localities along the Atlantic coast of the Iberian Peninsula. The obtained results revealed moderate levels of genetic diversity without population structure along the studied area. Furthermore, these microsatellites will allow to undertake further genetic population studies throughout the geographic distribution of the species, as well as other genetic analysis needed for aquaculture activities.

4. The whole repeat unit of 5S rDNA and the internal transcribed spacer (ITS) of the four European *Donax* species were analysed for the first time.

5. In this work, we newly determined, for the first time, the complete mitochondrial genomes of the four representatives of the family Donacidae in the Iberian Peninsula. The phylogenetic relationships among *Donax* species and their position within Veneroida order were illustrated. Our results demonstrate that the sequencing of complete mitogenomes provides highly valuable information for phylogenetic analysis in bivalves.

Contributions of this thesis to the management of *Donax* populations

1. The mitochondrial genes analysed (F and M types in 16S and Cytb) for *D. trunculus* showed the existence of genetic differentiation between the Iberian Peninsula localities. The data obtained may serve as a reference in the monitoring of variation genetic of natural populations, necessary for the conservation and management of resource. Localities formed three groups that we suggest should be treated as separate management units, and besides transfers of seed or broodstock across these units should be restricted.

2. Analyses of population differentiation performed with COI mitochondrial marker and microsatellites indicated absence of significant genetic structure in *D. vittatus* of Atlantic Iberian populations. The obtained results are useful to assist the fishery development and the conservation plans of this resource, proposed as an alternative to the exploitation of *D. trunculus*.

3. Multiplex PCR of two rDNA regions (5S and ITS) developed in this work allowed to authenticate four *Donax* species (*D. semistriatus*, *D. trunculus*, *D. variegatus* and *D. vittatus*). This method should be useful for not only to the conservation of these marine resources, but also to the characterization of stock breeders used in wedge clam culture in order to obtain their correct identity, and to detect and avoid commercial fraud.

4. The mitochondrial sequences contributed here add significantly useful genetic markers for further research and conservation of this commercially important fishing resource.

References

References

Abdelkrim, J., Robertson, B.C., Stanton, J.A.L., Gemmell, N.J., 2009. Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. Biotechniques. 46,185-192.

Allcock, A.L., Strugnell, J.M., 2012. Southern Ocean diversity: new paradigms from molecular ecology. Trends Ecol. Evol. 27,520-528.

Allendorf, F.W., England, P.R., Luikart, G., Ritchie, P.A., Ryman, N., 2008. Genetic effects of harvest on wild animal populations. Trends Ecol. Evol. 23,327-337.

Allendorf, F.W., Luikart, G., Aitken, S.N., 2013. Conservation and the Genetics of Populations, second ed. Wiley-Blackwell Publising. West Sussex, UK. pp. 602.

An, H.S., Kim, E.M., Park, J.Y., 2009. Isolation and characterization of microsatellite markers for the clam *Ruditapes philippinarum* and cross-species amplification with the clam *Ruditapes variegate*. Conserv. Genet. 10,1821-1823.

An, H.S., Lee, J.W., 2012. Development of Microsatellite Markers for the Korean Mussel, *Mytilus coruscus* (Mytilidae) Using Next-Generation Sequencing. Int. J. Mol. Sci. 13,10583-10593.

Andolfatto, P., 2001. Adaptive hitchhiking on genome variability. Curr. Opin. Genet. Dev. 11,635-641.

Ansell, A.D., 1972. Distribution, growth and seasonal changes in biochemical composition for the bivalve *Donax vittatus* (Da Costa) from Kames Bay, Millport. J. Exp. Mar. Bio. Ecol. 108,137-150.

Ansell, A.D., 1983. The Biology of the Genus *Donax*. In: McLachlan A., Erasmus T. (Eds.), Sandy Beaches as Ecosystems. W. Junk, The Hague. pp. 607-636.

Ansell, A.D., Lagardère, F., 1980. Observations on the biology of *Donax trunculus* and *D. vittatus* at Ile d´Oléron (French Atlantic coast). Mar. Biol. 57,287-300.

Arias, A., Freire, R., Méndez, J., Insua, A., 2010. Isolation and characterization of microsatellite markers in the queen scallop *Aequipecten opercularis* and their application to a population genetic study. Aquat. Living Resour. 23,199-207.

Arias, A., Freire, R., Pablo De La Roche, J., Román, G., Méndez, J., Insua, A., 2011. Identification, inheritance, and variation of microsatellite markers in the black scallop *Mimachlamys varia*. Biochem. Genet. 49,139-152.

Arias-Pérez, A., Cordero, D., Borrell, Y., Sánchez, J.A., Blanco, G., Freire, R., Insua, A., Saavedra, C., 2016. Assessing the geographic scale of genetic population management with microsatellites and introns in the clam *Ruditapes decussatus*. Ecol. Evol. 6,3380-3404.

Arias-Pérez, A., Fernández-Tajes, J., Gaspar, M.B., Méndez, J., 2012. Isolation of microsatellite markers and analysis of genetic diversity among east Atlantic populations of the sword razor shell *Ensis siliqua*: a tool for population management. Biochem. Genet. 50,397-415.

Arnaud-Haond, S., Bonhomme, F., Blanc, F., 2003. Large discrepancies in differentiation of allozymes, nuclear and mitochondrial DNA loci in recently founded Pacific populations of the pearl oyster *Pinctada margaritifera*. J. Evol. Biol. 16,388-398.

Asensio, L., González, I., Fernández, A., Céspedes, A., Rodríguez, M.A., Hernández, P.E., García, T., 2001. Identification of Nile Perch (*Lates niloticus*) Grouper (*Epinephelus guaza*), and Wreck Fish (*Polyprion americanus*) Fillets by PCR Amplification of the 5S rDNA Gene. J. AOAC Int. 84,777-785.

Astanei, I., Gosling, E., Wilson, J., Powell, E., 2005. Genetic variability and phylogeography of the invasive zebra mussel, *Dreissena polymorpha* (Pallas). Mol. Ecol. 14,1655-1666.

Avise, J.C., 1998. Conservation genetics in the marine realm. J. Hered. 89,377-382.

Baharum, S.N., Nurdalila, A.A., 2012. Application of 16s rDNA and cytochrome *b* ribosomal markers in studies of lineage and fish populations structure of aquatic species. Mol. Biol. Rep. 39,5225-5232.

Baker, P., Austin, J.D., Bowen, B.W., Baker, S.M., 2008. Range-wide population structure and history of the northern quahog (*Merceneria merceneria*) inferred from mitochondrial DNA sequence data. J. Mar. Sci. 65,155-163.

Ballard, J.W.O., 2000a. Comparative genomics of mitochondrial DNA in *Drosophila simulans.* J. Mol. Evol. 51,64-75.

Ballard, J.W.O., 2000b. Comparative genomics of mitochondrial DNA in members of the *Drosophila melanogaster* subgroup. J. Mol. Evol 51,48-63.

Ballard, J.W.O., Whitlock, M.C., 2004. The incomplete natural history of mitochondria. Mol. Ecol. 13,729-744.

Bao, X., He, C., Gao, X., Li, Y., Gao, L., Jiang, B., Liu, W., 2016. The complete mitochondrial genome of *Saxidomus purpuratus* (Veneroida: Veneridae). Mitochondrial DNA A DNA Mapp. Seq. Anal. 27,3648-9.

Barbará, T., Palma-Silva, C., Paggi, G.M., Bered, F., Fay, M.F., Lexer, C., 2007. Cross-species transfer of nuclear microsatellite markers: potential and limitations. Mol. Ecol. 16(18),3759-3767.

Bayed, A., Guillou, J., 1985. Contribution à l'étude des populations du genre *Donax*: la population de *Donax trunculus* L. (Mollusca, Bivalvia) de Mehdia (Maroc). Ann. Inst. Oceanogr. 61,139-147.

Bejega García, V., González Gómez de Agüero, E., Fernández Rodríguez, C., Álvarez García, J.C., 2010. Los concheros de O Neixón (Boiro, A Coruña) y Punta Atalaia (San Cibrao, Lugo): Una propuesta de muestreo y excavación de depósitos de la Edad del Hierro y época romana en Galicia, en González Gómez de Agüero, E., Bejega García, V., Fernández Rodríguez, C. y Prieto Fuertes, N. (coord.), Actas de la I Reunión de Arqueomalacología de la Península Ibérica, Férvedes. 6,33-42.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., Bonhomme, F., 2004. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier.

Bell, J.D., Rothlisberg, P.C., Munro, J.L., Loneragan, N.R., Nash, W.J., Ward, R.D., Andrew, N.L., 2005. Restocking and stock enhancement of marine invertebrate fisheries. Adv. Mar. Biol. 49,1-392.

Benson, G., 1999. Tandem repeats finder: a program to analyse DNA sequences. Nucleic Acids Res. 27,573-580.

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M., Stadler, P.F., 2013. MITOS: Improved de novo Metazoan Mitochondrial Genome Annotation. Mol. Phylogenet. Evol. 69,313-319.

Bettinazzi, S., Plazzi, F., Passamonti, M., 2016. The Complete Female- and Male-Transmitted Mitochondrial Genome of *Meretrix lamarckii*. PLoS ONE. 11,e0153631.

Birky, C.W., 2001. The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. Annu. Revi. Genet. 35,125-148.

Bogenhagen, F.D., Sakonju, S., Brown, D.D., 1980. A control region in the center of the 5S RNA gene directs specific initiation of transcription: II. The 3′ border of the region. Cell. 19(1):27-35.

Bogenhagen, D.F., Brown, D.D., 1981. Nucleotide sequences in *Xenopus* 5S DNA required for transcription terminator. Cell. 24,261-270.

Boore, J.L., Brown, W.M., 1994. Complete DNA sequence of the mitochondrial genome of the black chiton, *Katharina tunicata*. Genetics. 138,423-443.

Boore, J.L., Brown, W.M., 1998. Big trees from little genomes: mitochondrial gene order as phylogenetic tool. Curr. Opin. Genet. Dev. 8,668-674.

Boore, JL., 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27,1767-1780.

Borrell, Y.J., Arias-Pérez, A., Freire, R., Valdés, A., Sánchez, J.A., Méndez, J., Martínez, D., López, J., Carleos, C., Blanco, G., Insua, A.M., 2014. Microsatellites and multiplex PCRs for assessing aquaculture practices of the grooved carpet shell *Ruditapes decussatus* in Spain. Aquaculture. 426, 49-59.

Bowden, K.F., 1975. Oceanic and estuarine mixing processes. In Press, J.P. Riley & Skirrow G. (Eds.), Chemical Oceanography. London: Academic Press. pp. 1-41.

Boyle, E.E., Etter, R.J., 2013. Heteroplasmy in a deep-sea protobranch bivalve suggests an ancient origin of doubly uniparental inheritance of mitochondria in Bivalvia. Mar. Biol. 160,413-422.

Breton, S., Doucet-Beaupré, H., Stewart, D.T., Hoeh, W.R., Blier, P.U., 2007. The unusual system of doubly uniparental inheritance of mtDNA: isn´t one enough? Trends Genet. 23,465-74.

Breton, S., Doucet-Beaupré, H., Stewart, D.T., Piontkivska, H., Karmakar, M., Bogan, A.E., Blier P.U., Hoeh, W.R., 2009. Comparative mitochondrial genomics of freshwater mussels

(Bivalvia:Unionoida) with doubly uniparental inheritance of mtDNA: gender-specific open reading frames and putative origins of replication. Genetics. 183,1575-1589.

Breton, S., Stewart, D.T., Hoeh, W.R., 2010. Characterization of a mitochondrial ORF from the gender-associated mtDNAs of *Mytilus* spp. (Bivalvia: Mytilidae): identification of the "missing" ATPase gene. Mar Genomics. 3,11-18.

Brito, P.H., 2007. Contrasting patterns of mitochondrial and microsatellite genetic structure among Western European populations of tawny owls (*Strix aluco*). Mol. Ecol. 16,3423-3437.

Brown, A.C., McLachlan, A., 1990. Ecology of sandy shores. Elsevier Science Publishers, Amsterdam.

Buroker, N.E., Hershberger, W.K., Chew, K.K., 1979a. Population genetics of the family ostreidae. I. Intraspecific studies of *Crassostrea gigas* and *Saccostrea commercialis.* Mar. Biol. 54,157-169.

Buroker, N.E., Hershberger, W.K., Chew, K.K., 1979b. Population genetics of the family ostreidae. II. Intraspecific studies of *Crassostrea* and *Saccostrea.* Mar. Biol. 54,171-184.

Burzyński, A., Zbawicka, M., Skibinski, D.O.F., Wenne, R., 2003. Evidence for Recombination of mtDNA in the Marine Mussel *Mytilus trossulus* from the Baltic. Mol. Biol. Evol. 20,388-392.

Camara, M.D., Vadopalas, B., 2009. Genetic aspects of restoring Olympia oyters and other native bivalves: balancing the need for action, good intentions, and the risks of making thing worse. J. Shellfish Res. 28,121-145.

Carlsson, J., Reece, K.S., 2007. Eight PCR primers to amplify EST-linked microsatellites in the eastern oyster, *Crassostrea virginica* genome. Mol. Ecol. Notes. 7,257-259.

Carrera, E., García, T., Céspedes, A., González, I., Fernández, A., Asensio, L.M., Hernández, P.E., 2000. Differentiation of smoked *Salmo salar, Oncorhynchus mykiss* and *Brama raii* using the nuclear marker 5S rDNA. Int. J. Food Sci. Technol. 35,401–406.

Castoe, T.A., Poole, A.W., Gu, W., Janson de Koning, A.P., Daza, J.M., Smith, E.N., Pollock, D.D., 2010. Rapid identification of thousands of copperhead snake (*Agkistrodon contortrix*) microsatellite loci from modest amounts of 454 shotgun genome sequence. Mol. Ecol.

Resour. 10,341-347.Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17,540-552.

Céspedes, A., García, T., Carrera, E., González, I., Fernández, A., Hernández, P.E., Martín, R., 1999. Identification of sole (*Solea solea*) and Greenland halibut (*Reinharius hippglossoides*) by PCR amplification of the 5S rDNA gene. J. Agric. Food Chem. 47,1046-1050.

Chabane, K., Ablett, G.A., Cordeiro, G.M., Valkoun, J., Henry, R.J., 2005. EST versus genomic derived microsatellite markers for genotyping wild and cultivated barley. Genet. Resour. Crop Evol. 52,903-909.

Chacón, G., Arias-Pérez, A., Méndez, J., Insua, A., Freire, R., 2013. Development and multiplex PCR amplification of microsatellite markers in the commercial clam *Venerupis rhomboides* (Mollusca:Bivalvia). Mol. Biol. Rep. 40,1625-1630.

Chambers, G.K., MacAvoy, E., 2000. Microsatellites: consensus and controversy. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 126,455-476.

Chapuis, M.P., Estoup, A., 2007. Microsatellite null alleles and estimation of population differentiation. Mol. Biol. Evol. 24,621-631.

Cheng, H.L., Meng, X.P., Ji, H.J., Dong, Z.G., Chen, S.Y., 2006. Sequence analysis of the ribosomal DNA internal transcribed and 5.8S ribosomal RNA gene in representatives of the clam family Veneridae (Mollusca: Bivalvia). J. Shellfish Res. 25,833-839.

Chairi, H., Rebordinos, L., 2014. A Rapid Method for Differentiation Four Species of the Engraulidae (Anchovy) Family. J. Agric. Food Chem. 62,2803-2808.

Chícharo, L., Chícharo, A., Gaspar, M., Alves, F., Regala, J., 2002. Ecological characterization of dredged and non-dredged bivalve fishing areas off south Portugal. J. Mar. Biol. Assoc. UK. 82,41-50.

Chiesa, S., Lucentini, L., Freitas, R., Nonnis Marzano, F., Ferrari, C., Filonzi, L., Breda, S., Minello, F., Figueira, E., Argese, E., 2016. Null alleles of microsatellites for Manila clam *Ruditapes philippinarum*. Anim. Genet. 47,135-136.

Chistiakov, D.A., Hellemans, B., Volckaert, F.A.M., 2006. Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics. Aquaculture. 255,1-29.

Chiu, Y.W., Bor, H., Tan, M.S., Lin, H.D., Jean, C.T., 2013. Phylogeography and Genetic Differentiation among Populations of the Moon Turban Snail *Lunella granulata* Gmelin, 1791 (Gastropoda: Turbinidae). Int. J. Mol. Sci.14,9062-9079.

Chow, S., Ueno, Y., Toyokawa, M., Oohara, I., Takeyama, H., 2009. Preliminary Analysis of Length and GC Content Variation in the Ribosomal First Internal Transcribed Spacer (ITS1) of Marine Animals. Mar. Biotechnol. 11,301-306.

Civera, T., 2003. Species identification and safety of fish products. Vet. Res. Commun. 27,481-489.

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9,1657-1659.

Coleman, A.W., Vacquier, V.D., 2002. Exploring the phylogenetic utility of ITS sequences for animals: a test case for abalone (*Haliotis*). J. Mol. Evol. 54,246-257.

Colgan, D. J., McLauchlan, A, Wilson, G.D.F., Livingston, S.P., Edgecombe, G.D., Macaranas, J., Cassis, G., Gray, M.R.., 1998. Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. Aust. J. Zool. 46,419-437.

Combosch, D.J., Collins, T.M., Glover, E.A., Graf, D.L., Harper, E.M., Healy, J.M., Kawauchi, G.Y., Lemer, S., McIntyre, E., Strong, E.E., Taylor, J.D., Zardus, J.D., Mikkelsen, P.M., Giribet, G., Bieler, R., 2017. A family-level Tree of Life for bivalves based on a Sanger-sequencing approach. Mol. Phylogenet. Evol. 107,191-208.

Conant, G.C., Wolfe, K.H., 2008. GenomeVx: Simple web-based creation of editable circular chromosome maps. Bioinformatics. 24,861-862.

Consellería do Mar, Xunta de Galicia. In *Pesca de Galicia - Plataforma tecnolóxica da pesca* http://www.pescadegalicia.gal/estadisticas/ (2018). Last access date 24 May 2018.

Cornet, M., Soulard, C., 1990. Chromosome number and karyotype of *Donax trunculus* L. (Mollusca, Bivalvia, Tellinacea). Genetica. 82,93-97.

Couceiro, L., Barreiro, R., Ruiz, J.M., Sotka, E.E., 2007. Genetic isolation by distance among populations of the netted dog whelk *Nassarius reticulatus* (L.) along the European Atlantic coastline. J. Hered. 98,603-610.

Coulibaly, I., Gharbi, K., Danzmann, R.G., Yao, J., Rexroad, C.E., 2005. Characterization and comparison of microsatellites derived from repeat-enriched libraries and expressed sequence tags. Anim Genet. 36, 309-315.

Cross, I., Rebordinos, L., 2005. 5S rDNA and U2 snRNA are linked in the genome of *Crassostrea angulata* and *Crassostrea gigas* oysters: does the (CT)n.(GA)n microsatellite stabilize this novel linkage of large tandem arrays? Genome. 48,1116-1119.

Cross, I., Rebordinos, L., 2006. Species identification of *Crassostrea* and *Ostrea* oysters by polymerase chain reaction amplification of the 5S rRNA gene. J. AOC Int. 89,144-148.

Cross, T.F., 2000. Genetic implications of translocation and stocking of fish species, with particular reference to Western Australia. Aquacult. Res. 31,83-94.

Curole, J.P., Kocher, T.D., 2002. Ancient sex-specific extension of the cytochrome c oxidase II gene in bivalves and the fidelity of doubly-uniparental inheritance. Mol. Biol. Evol. 19,1323-1328.

Curole, J.P., Kocher, T.D., 2005. Evolution of a unique mitotype-specific proteincoding extension of the cytochrome c oxidase II gene in freshwater mussels (Bivalvia: Unionoida). J. Mol. Evol. 61,381-389.

Da Costa, F., Aranda-Burgos, J.A., Cerviño-Otero, A., Fernández-Pardo, A., Louzán, A., Nóvoa, S., Ojea, J., Martínez-Patiño, D., 2012. Clam hatchery and nursery culture. In: Da Costa, F. (Ed.), Clams Fisheries and Aquaculture. Nova Pub., New York. pp. 217-254.

Da Costa, F., Martínez-Patiño, D., 2009. Culture potential of the razor clam *Solen marginatus* (Pennánt, 1777). Aquaculture. 288,57-64.

Dahlgren, T.G., Weinberg, J.R., Halanych, K.M., 2000. Phylogeography of the ocean quahog (*Arctica islandica*): influences of paleoclimate on genetic diversity and species range. Mar. Biol. 137,487-495.

Dakin, E.E., Avise, J.C., 2004. Microsatellite null alleles in parentage analysis. Heredity. 93,504-509.

Darling, J.A., Bagley, M.J., Roman, J., Telpot, C.K., Geller, J.B., 2008. Genetic patterns across multiple introductions of the globally invasive crab genus *Carcinus*. Mol. Ecol. 17,4992-5007.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2011. ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics. 27,1164-1165.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. "jModelTest 2: more models, new heuristics and parallel computing". Nature Methods. 9,772.

Dégletagne, C., Abele, D., Held, C., 2016. A distinct mitochondrial genome with DUI-like inheritance in the ocean quahog *Arctica islandica.* Mol. Biol. Evol. 33,375-383.

Deval, M.C., 2009. Growth and reproduction of the wedge clam (*Donax trunculus*) in the Sea of Marmara, Turkey. J. Appl. Ichthyol. 25,551-558.

Dixon, D.R., Sole-Cava, A.M., Pascoe, P.L., Holland, P.W.H., 1995. Periostracal adventitious hairs on spat of the mussel *Mytilus edulis*. J. Mar. Biol. Assoc. U.K. 75,363-372.

Diz, A.P., Presa, P., 2008. Regional patterns of microsatellite variation in *Mytilus galloprovincialis* from the Iberian Peninsula. Mar. Biol. 154,277-286.

Domínguez-Contreras, J.F., Munguía-Vega, A., Getino-Mamet, L.N., Soria, G., Parma, A., 2017. Characterization of 30 microsatellite loci for the Tehuelche scallop, *Aequipecten tehuelchus* (d´Orbigny, 1842) and their use for estimating demographic parameters relevant to fisheries management. Molluscan Res. 1-7.

Dong, P., Ma, G., Chang, L., Zhu, Y., Tian, X., 2016. The complete mitochondrial genome of *Cyclina sinensis* (Veneroida:Veneridae). Mitochondrial DNA Part B. 1,173-174.

Dreyer, H., Steiner, G., 2006. The complete sequences and gene organisation of the mitochondrial genomes of the heterodont bivalves *Acanthocardia tuberculata* and *Hiatella arctica* - and the first record for a putative Atp*ase subunit 8* gene in marine bivalves. Front. Zool. 3,13.

Drouin, G., Moniz de Sá, M., 1995. The concerted evolution of 5S ribosomal genes linked to the repeat unit of other multigene families. Mol. Biol. Evol. 12,481-493.

Drummond, A.J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Thierer, T., Wilson, A., 2009. Geneious v4.8.5. Available at: http://www.geneious.com/

Duran, S., Palacin, C., Becerro, M.A., Turon, X., Giribet, G., 2004. Genetic diversity and population structure of the commercially harvested sea urchin *Paracentrotus lividus* (Echinodermata, Echinoidea). Mol. Ecol. 13,3317-3328.

D´Amico, P., Armani, A., Gianfaldoni, D., Guidi, A., 2016. New provisions for the labelling of fishery and aquaculture products: Difficulties in the implementation of Regulation (EU) n. 1379/2013. Mar. Policy. 71,147-156.

Earl, D.A., von Holdt, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4,359-361.

EC (European Commission), 2000. European Council Regulation No 104/2000 of 17 December 1999 on the common organization of the markets in fishery and aquaculture products. Official Journal of the European Communities. L17,22-52.

EC (European Commission), 2001. Commission Regulation (EC) No 2065/2001 of 22 October 2001 laying down detailed rules for the application of Council Regulation (EC) No 104/2000 as regards informing consumers about fishery and aquaculture products. Official Journal of the European Communities. L278,6-8.

EC (European Commission), 2013. Commission Regulation (EC) No 1379/2013 of 11 December 2013 on the common organization of the markets in fishery and aquaculture products, amending Council Regulations (EC) No 1184/2006 and (EC) No 1124/2009 and repealing Council Regulation (EC) No 104/2000. Official Journal of the European Communities. L354,1-21.

Eirín-López, J.M., Ruiz, F., González-Tizón, A.M., Martínez, A., Sánchez, L., Méndez, J., 2004. Molecular evolutionary characterization of the mussel *Mytilus* histone multigene family: first record of a tandemly repeated unit of five histone genes containing an H1 subtype with "orphon" features. J. Mol. Evol. 58,131-144.

Ellegren, H., 2000. Microsatellite mutations in the germline: implications for evolutionary inference. Trends Genet. 16,551-558.

Ellers, O., 1995. Behavioural control of swash-riding in the clam *Donax variabilis*. Biol. Bull. 189,120-127.

Estoup, A., Rousset, F., Michalakis, Y., Cornuet, J.M., Adriamanga, m., Guyomard, R., 1998. Comparative analysis of microsatellite and allozyme markers a case study investigating microgeographic differentiation in brow trout (*Salmo trutta*). Mol. Ecol. 7,339-353.

Excoffier, L., Lischer, H.E.L., 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analysis under Linux and Windows. Mol. Ecol. Res. 10,564-567.

Faber, J.E., Stepien, C.A., 1998. Tandemly repeated sequences in the mitochondrial DNA control region and phylogeography of the Pike-perches *Stizostedion*. Mol. Phylogenet. Evol. 10,310-322.

Fang, B.L., De Baere, R., Vandenberghe, A., De Wachter, R., 1982. Sequences of the three molluscan 5S ribosomal RNAs confirm the validity of a dynamic secondary structure model. Nucleic Acids Res. 10,4679-4685.

FAO, 1997. Aquaculture Development. FAO Technical Guidelines for Responsible Fisheries. No. 5. Rome, FAO. (Available at ftp://ftp.fao.org/docrep/fao/003/W4493e/W4493e00.pdf).

FAO, 2002. Estado mundial de la pesca y la acuicultura 2002, Roma. (Available at http://www.fao.org/docrep/005/y7300s/y7300s00.htm).

FAO, 2008. Aquaculture Sevelopment. 3. Genetic resource management. FAO Technical Guidelines for Responsible Fisheries. No. 5, Suppl. 3. Rome, FAO. 125 p. (Available at www.fao.org/docrep/011/i0283e/i0283e00.htm).

FAO-FIGIS: In *Fisheries Global Information System* http://www.fao.org/fishery/statistics/global-capture-production/query/en (2017). Last access date 19 March 2017.

FAO-FIGIS: In *Fisheries Global Information System* http://www.fao.org/fishery/statistics/global-capture-production/query/en (2018). Last access date 6 May 2018.

Ferchaud, A.L., Eudeline, R., Arnal, V., Cheylan, M., Leblois, R., Crochet, P.A., 2015. Congruent signals of population history but radically different patterns of genetic diversity between mitochondrial and nuclear markers in a mountain lizard. Mol. Ecol. 24,192-207.

Ferguson, A., Taggart, J.B., Prodohl, P.A., McMeel, O., Thompson, C., Stone, C., McGinnity, P., Hynes, R.A., 1995. The application of molecular markers to the study and conservation of fish populations, with special reference to Salmo. J. Fish Biol. 47,103-126.

Fernández, A., García, T., Asensio, L., Rodríguez, M.A., González, I., Hernández, P.E., Martín, R., 2001. PCR-RFLP analysis of the internal transcribed spacer (ITS) region for identification of 3 clams species. J. Food Sci. 66,657-661.

Fernández-Pérez, J., Froufe, E., Nantón, A., Gaspar, M.B., Méndez, J., 2017a. Genetic diversity and population genetic analysis of *Donax vittatus* (Mollusca: Bivalvia) and phylogeny of the genus with mitochondrial and nuclear markers. Est. Coast. Shelf Sci. 197,126-135.

Fernández-Pérez, J., Nantón, A., Arias-Pérez, A., Martínez-Patiño, D., Méndez, J., 2018. Mitochondrial DNA analyses of *Donax trunculus* (Mollusca: Bivalvia) population structure in the Iberian Peninsula, a bivalve with high commercial importance. Aquatic Conserv.: Mar. Freshw. Ecosyst. DOI: 10.1002/aqc.2929.

Fernández-Pérez, J., Nantón, A., Ruiz-Ruano, F.J., Camacho, J.P.M., Méndez, J., 2017b. First complete female mitochondrial genome in four bivalve species genus *Donax* and their phylogenetic relationship within the Veneroida order. PloS ONE. 12, e0184464.

Fernández-Tajes, J., Freire, R., Méndez, J., 2010. A simple one-step PCR method for the identification between European and American razor clams species. Food Chem. 118,995-998.

Fernández-Tajes, J., Longa, A., García-Gil, J., Chiu, Y.W., Huang, Y.S., Méndez, J., Lee, R.S., 2011. Alternative PCR-RFLP methods for mussel *Mytilus* species identification. Eur. Food Res. Technol. 233,791-796.

Fernández-Tajes, J., Méndez, J., 2007. Identification of the Razor Clam Species *Ensis arcuatus*, *E. siliqua*, *E. directus*, *E. macha*, and *Solen marginatus* Using PCR-RFLP Analysis of the 5S rDNA Region. J. Agric. Food Chem. 55,7278-7282.

Fernández-Tajes, J., Méndez J., 2009. Two different size classes of 5S rDNA units coexisting in the same tandem array in the razor clam *Ensis macha*: is this region suitable for phylogeographic studies? Biochem. Genet. 47,775-788.

Ferreira, I.A., Oliveira, C., Venere, P.C., Galetti, P.M., Martins, C., 2006. 5S rDNA variation and its phylogenetic inference in the genus *Leporinus* (Characiformes: Anostomidae). Genetica. 129,253-257.

Fisher, C., Skibinski, D.O.F., 1990. Sex-biased mitochondrial DNA heteroplasmy in the marine mussel *Mytilus*. Proc R Soc Lond B Biol Sci.242, 149-156.

Frankham, R., Ballou, J.D., Briscoe, D.A., 2002. Introduction to Conservation Genetics. Cambridge University Press. Cambridge. pp. 617.

Freire, R., Arias, A., Insua, A.M., Méndez, J., Eirín-López, J.M., 2010a. Evolutionary dynamics of the 5S rDNA gene family in the mussel *Mytilus*: mixed effects of the birth-and-death and concerted evolution. J. Mol. Evol. 70,413-426.

Freire, R., Arias, A., Méndez, J., Insua, A., 2010b. Sequence variation of the internal transcribed spacer (ITS) region of ribosomal DNA in *Cerastoderma* species (Bivalvia: Cardiidae). J. Molluscan Stud. 76,77-86.

Freire, R., Arias, A., Méndez, J., Insua, A., 2011. Identification of European commercial cockles (*Cerastoderma edule* and *C. glaucum*) by species-specific PCR amplification of the ribosomal DNA ITS region. Eur. Food Res. Technol. 232,83-86.

Freire, R., Insua, A., Méndez, J., 2005. *Cerastoderma glaucum* 5S ribosomal DNA: characterization of the repeat unit, divergence with respect to *Cerastoderma edule*, and PCR-RFLPs for the identification of both cockles. Genome. 48,427-442.

Froufe, E., Sobral, C., Teixeira, A., Lopes, A., Sousa, R., Varandas, S., Lopez-Lima, M., 2013. Development and multiplexing of microsatellite loci for the near threatened freshwater mussel *Potomida littoralis* (Cuvier, 1789) using 454 sequencing. Aquatic Conserv.: Mar. Freshw. Ecosyst. 23,619-623.

Frouin, R., Fuiza, A.F., Ambar, I., Boyd, T.J., 1990. Observations of a poleward surface current off the coasts of Portugal and Spain during the winter. J. Geophys. Res. 95,679-691.

Fu, Y.X., 1997. Statistical Tests of Neutrality of Mutations Against Population Growth, Hitchhiking and Background Selection. Genetics. 147,915-925.

Funk, W. C., McKay, J. K., Hohenlohe, P. A., Allendorf, F. W., 2012. Harnessing genomics for delineating conservation units. Trends Ecol. Evol. 27,489-496.

Gaffney, PM., 2006. The role of genetics in shellfish restoration. Aquat. Living Resour. 19,277-282.

Galindo-Sánchez, C.E., Gaffney, P.M., Pérez-Rostro, C.I., De la Rosa-Vélez, J., Candela, J., Cruz, P., 2008. Assessment of genetic diversity of the eastern oyster *Crassostrea virginica* in Veracruz, Mexico using microsatellite markers. J. Shellfish Res. 27,721-727.

Galtier, N., Nabholz, B., Glemin, S., Hurst, G.D.D., 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Mol. Ecol. 18,4541-4550.

Gan, H.M., Gan, H.Y., Tan, M.H., Penny, S.S., Willan, R.C., Austin, C.M., 2016a. The complete mitogenome of the giant clam *Tridacna squamosa* (Heterodonta: Bivalvia: Tridacnidae). Mitochondrial DNA A DNA Mapp Seq. Anal. 27,3220-1.

Gan, H.M., Tan, M.H., Thai, B.T., Austin, C.M., 2016b. The complete mitogenome of the marine bivalve *Lutraria rhynchaena* Jonas 1844 (Hteredonta: Bivalvia: Mactridae). Mitochondrial DNA A DNA Mapp Seq. Anal. 27,335-6.

Gandolfi, A., Bonilauri, P., Rossi, V., Menozzi, P., 2001. Intraindividual and intraspecies variability of ITS 1 sequences in the ancient asexual *Darwinula stevensoni* (Crustacea: Ostracoda). Heredity. 87,449-455.

García-Merchán, V.H., Robainas-Barcia, A., Abelló, P., Macpherson, E., Palero, F., García-Rodríguez, M., Gil de Sola, L., Pascual, M., 2012. Phylogeographic patterns of decapod crustaceans at the Atlantic Mediterranean transition. Mol. Phylogenet. Evol. 62,664-672.

García-Souto, D., Pérez-García, C., Pasantes, J.J., 2017. Are Pericentric Inversions Reorganizing Wedge Shell Genomes? Genes. 8,370.

Gardeström, J., Pereyra, R.T., André, C., 2008. Characterization of six microsatellite loci in the Baltic blue mussel *Mytilus trossulus* and cross-species amplification in North Sea *Mytilus edulis*. Conserv. Genet. 9,1003-1005.

Gaspar, M.B., Ferreira, R., Monteiro, C.C., 1999. Growth and reproductive cycle of *Donax trunculus* L., (Mollusca: Bivalvia) off Faro, southern Portugal. Fish. Res. 41,309-316.

Gaspar, M.B., Santos, M.N., Vasconcelos, P., Monteiro, C.C., 2002. Shell morphometric relationships of the most common bivalve species (Mollusca: Bivalvia) of the Algarve coast (southern Portugal). Hydrobiologia. 477,73-80.

Gaspar, M.B., Sobral, M., Maia, F., Sobral, M.P., Viegas, M.C., Monteiro, C.C., 2004. A pescaria de moluscos bivalves na Zona Ocidental Norte e Zona Sul. Ponto de situação dos principais bancos (Setembro/Outubro 2002). Relat. Cient. Téc. IPIMAR, série digital, n.º 12, pp. 1-33.

Geiduschek, E.P., Tocchini-Valentini, G.P., 1988. Transcription by RNA polymerase III. Annu. Rev. Biochem. 57,873-914.

Geist, J., 2010. Strategies for the conservation of endangered freshwater pearl mussels (*Margaritifera margaritifera* L.): a synthesis of Conservation Genetics and Ecology. Hydrobiologia. 644,69-88.

Gerbi, S.A., 1985. Evolution of ribosomal DNA. In: McIntyre, R.J. (Eds.), Molecular Evolutionary Genetic. New York, N.Y.: Plenum Press. pp. 419-517.

Gissi, C., Iannelli, F., Pesole, G., 2008. Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. Heredity (Edinb). 101,301-320.

Glöckner, G., Heinze, I., Platzer, M., Held, C., Abele, D., 2013. The mitochondrial genome of *Artica islandica*; Phylogeny and variation. PLoS ONE. 8,e82857.

González, I.L., Sylvester, J.E., Smith. T.F., Stambolian, D., Schmickel, R.D., 1990. Ribosomal RNA gene sequences and hominoid phylogeny. Mol. Biol. Evol. 7,203-219.

González-Tizón, A., Martínez-Lage, A., Mariñas, L., Freire, R., Cornudella, L., Méndez, J., 1998. Cytogenetic characterization of *Donax trunculus* (Mollusca, Bivalvia). Cytogenet. Cell Genet. 81,109.

Goudet, J., 2001. FSTAT, a program to estimate and test gene diversities and fixation indices. Institut d'Ecologie, Université de Lausanne, Dorigny.

Guerra, D., Ghiselli, F., Passamonti, M., 2014. The largest unassigned regions of the male- and female- transmitted mitochondrial DNAs in *Musculista senhousia* (Bivalvia Mytilidae). Gene. 536,316-325.

Gupta, P.K., Rustgi, S., Sharma, S., Singh, R., Kumar, N., Balyan, H.S., 2003. Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. Mol. Genet. Genomics. 270,315-323.

Gusman, A., Lecomte, S., Stewart, D.T., Passamonti, M., Breton, S., 2016. Pursuing the quest for better understanding the taxonomic distribution of the system of doubly uniparental inheritance of mtDNA. PeerJ. 4,e2760.

Hahn, C., Bachmann, L., Chevreux, B., 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads-a baiting and iterative mapping approach. Nucleic Acids Res. 41,e129.

Halanych, K.M., 1991. 5S ribosomal RNA sequences inappropriate for phylogenetic reconstruction. Mol. Biol. Evol. 8,249-253.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41,95-98.

Hargrove, J.S., Sturmer, L., Scarpa, J., Austin, J.D., 2015. Assessment of genetic diversity in wild and aquaculture stocks of *Mercenaria mercenaria* in Florida. J. Shellfish Res. 34,355-365.

Harrison, R.G., 1989. Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. Trends. Ecol. Evol. 4,6-12.

Hasegawa, M., Cao, Y., Yang, Z., 1998. Preponderance of slightly deleterious polymorphism in mitochondrial DNA: Nonsynonymous/synonymous rate ratio is much higher within species than between species. Mol. Biol. Evol. 15,1499-1505.

Haynes, R., Barton, E.D., 1990. A poleward flow along the Atlantic coast of the Iberian Peninsula. J Geophys. Res. 95,11425-11441.

He, C.B., Wang, J., Gao, X.G., Song, W.T., Li, H.J., Li Y.F., Liu, W.D., Su, H., 2011. The complete mitochondrial genome of the hard clam *Meretrix meretrix.* Mol. Biol. Rep. 38,3401-3409.

He, M., Huang, L., Shi, J., Jiang, Y., 2005. Variability of ribosomal DNA ITS-2 and its utility in detecting genetic relatedness of Pearl Oyster. Mar. Biotechnol. 7,40-45.

Heath, D.D., Rawson, P.D., Hilbish, T.J., 1995. PCR-based nuclear markers identify alien blue mussel (*Mytilus spp.*) genotypes on the west coast of Canada. Can. J. Fish. Aquat. Sci. 52,2621-2627.

Hedgecock, D., Li, G., Hubert, S., Bucklin, K., Ribes, V., 2004. Widespread null alleles and poor cross-species amplification of microsatellite DNA loci cloned from the Pacific oyster, *Crassostrea gigas*. J. Shellfish Res. 23,379-385.

Hillis, D.M., Dixon, M.T., 1991. Ribosomal DNA: Molecular Evolution and Phylogenetic Inference. Q. Rev. Biol. 66,411-453.

Hixson, J.E., Brown, W.M., 1986. A comparison of the small ribosomal RNA genes from the mitochondrial DNA of the great apes and humans: sequence, structure, evolution, and phylogenetic implications. Mol. Biol. Evol. 3,1-18.

Hoeh, W.R., Blakley, K.H., Brown, W.M., 1991. Heteroplasmy suggests limited biparental inheritance of *Mytilus* mitochondrial DNA. Science. 251,1488-1490.

Hoeh, W.R., Stewart, D.T., Guttman, S.I., 2002. High fidelity of mitochondrial genome transmission under the doubly uniparental mode of inheritance in freshwater mussels (Bivalvia: Unionoidea). Evolution. 56,2252-2261.

Hoeh, W.R., Stewart, D.T., Sutherland, B.W., Zouros E., 1996. Cytochrome c oxidase sequence comparisons suggest an unusually high rate of mitochondrial DNA evolution in *Mytilus* (Mollusca: Bivalvia). Mol. Biol. Evol. 13,418-421.

Hoffman, R.J., Boore, J.L., Brown, W.M., 1992. A novel mitochondrial genome organization for the blue mussel, Mytilus edulis. Genetics. 131,397-412.

Hoffman, J.I., Dasmahapatra, K.K., Amos, W., Phillips, C.D., Gelatt, T.S., Bickham, J.W., 2009. Contrasting patterns of genetic diversity at three different genetic markers in a marine mammal metapopulation. Mol. Ecol. 18,2961-2978.

Holleley, C.E., Geerts, P.G., 2009. Multiplex Manager 1.0: a crossplatform computer program that plans and optimizes multiplex PCR. Biotechniques. 46,511-517.

Hu, Y.P., Lutz, R.A., Vrijenhoek, R.C., 1992. Electrophoretic identification and genetic analysis of bivalve larvae. Mar. Biol. 113,227-230.

Hu, J.J., Nakatani, M., Mizuno, K., Fujimura, T., 2004. Development and characterization of microsatellite markers in sweetpotato. Breed. Sci. 54,177-188.

Huang, X., Hu, J., Hu, X., Zhang, C., Zhang, L., Wang, S., Lu, W., Bao, Z., 2007. Cytogenetic characterization of the bay scallop, *Argopecten irradians irradians*, by multiple staining techniques and fluorescence *in situ* hybridization. Genes Genet. Syst. 82,257-263.

Huang, X-C., Rong, J., Liu, Y., Zhang, M-H., Wan, Y., Ouyang, S., Zhou, C-H., 2013. The Complete Maternally and Paternally Inherited Mitochondrial Genomes of the Endangered Freshwater Mussel *Solenaia carinatus* (Bivalvia: Unionidae) and Implications for Unionidae Taxonomy. PLoS ONE. 8,e84352.

Huang, Y., Maraia, R.J., 2001. Comparison of the RNA polymerase III transcription machinery in *Schizosaccharomyces pombe, Saccharomyces cerevisiae* and human. Nucleic Acids Res. 29,2675-2690.

Hurtado, S.N., Pérez-García, C., Morán, P., Pasantes, J.J., 2011. Genetic and cytological evidence of hybridization between native *Ruditapes decussatus* and introduced *Ruditapes philippinarum* (Mollusca, Bivalvia, Veneridae) in NW Spain. Aquaculture. 311,123-128.

Hwang, J.Y., Han, G.G., Park, J.Y., Kim, E.M., An, C.M., Kang, J.H., Choi, Y.J., Kim, E.B., 2016. Complete sequence and polymorphisms of female *Ruditapes philippinarum* (Mollusca: Bivalvia) mitochondria genome. Mitochondrial DNA A DNA Mapp. Seq. Anal. 27,3462-3.

Imanishi, Y., Tanaka, M., Fujiwara, M., 2013. Complete mitochondrial genome sequence of Japanese cockle *Fulvia mutica* (Cardiidae). Fish. Sci. 79,949-57.

Infante, C., Blanco, E., Zuasti, E., Crespo, A., Manchado, M., 2007. Phylogenetic differentiation between Atlantic *Scomber colias* and Pacific *Scomber japonicus* based on nuclear DNA sequences. Genetica. 130,1-8.

Infante, C., Manchado, M., 2007. Recursos genéticos de peces marinos de interés en acuicultura. En: Martínez P y Figueras A (eds) Genética y Genómica en Acuicultura. Publicaciones científicas y tecnológicas del Observatorio Español de Acuicultura, Madrid, pp. 23-64.

Insua, A., Freire, R., Méndez, J., 1999. The 5S rDNA of the bivalve *Cerastoderma edule*: nucleotide sequence of the repeat unit and chromosomal location relative to 18S-28S rDNA. Genet Select Evol 31,509-518.

Insua, A., Freire, A., Ríos, J., Méndez, J., 2001. The 5S rDNA of mussels *Mytilus galloprovincialis* and *M. edulis*: sequence variation and chromosomal location. Chromosome Res. 9,495-505.

Insua, A., López-Piñón, M.J., Freire, R., Méndez, J., 2003. Sequence analysis of the ribosomal DNA internal transcribed spacer in some scallop species (Mollusca: Bivalvia: Pectinidae). Genome. 46,595-604.

Insua, A., López-Piñón, M.J., Freire, R., Méndez, J., 2006. Karyotype and chromosomal location of 18S-28S and 5S ribosomal DNA in the scallops *Pecten maximus* and *Mimachlamys varia* (Bivalvia: Pectinidae). Genetica. 126,291-301.

Insua, A., López-Piñón, M.J., Méndez, J., 1998. Characterization of *Aequipecten opercularis* (Bivalvia: Pectinidae) chromosomes by different staining techniques and fluorescent *in situ* hybridization. Genes Genet. Syst. 73,193-200.

Jansen, G., Devaere, S., Weekers, P.H.H., Adriaens, D., 2006. Phylogenetic relationships and divergence time estimate of African anguilliform catfish (Siluriformes: Clariidae) inferred from ribosomal gene and spacer sequences. Mol. Phylogenet. Evol. 38,65-78.

Johnson, J.A., Toepfer, J.E., Dunn, P.O., 2003. Contrasting patterns of mitochondrial and microsatellite population structure in fragmented populations of greater prairie-chickens. Mol. Ecol. 12,3335-3347.

Johnson, M.S., 2000. Measuring and interpreting genetic structure to minimize the genetic risks of translocations. Aquacult. Res. 31,133-143.

Jukes, T.H., Cantor, C.R., 1969. Evolution of protein molecules. In Press: H.N., Munro (Ed.) Mammalian protein metabolism. Nueva York: Academic press. pp. 21-123.

Källersjö, M., Von Proschwitz, T., Lundberg, S., Eldenäs, P., Erséus, C., 2005. Evaluation of ITS rDNA as a complement to mitochondrial gene sequences for phylogenetic studies in freshwater mussels: an example using Unionidae from north-western Europe. Zoologica Scripta. 34,415-424.

Kang, J.H., Kim, Y.K., Park, J.Y., Noeh, E.S., Jeong, J.E., Lee, Y.S., Choi, T.J., 2013. Development of microsatellite markers for a hard-shelled mussel, *Mytilus coruscus*, and cross-species transfer. Genet. Mol. Res. 12, 4009-4017.

Katoh, K., Toh, H., 2008. Recent developments in the MAFFT multiple sequence alignment program. Brief. Bioinform. 9,286-298.

Katsares, V., Tsiora, A., Galinou-Mitsoudi, S., Imsiridou, A., 2008. Genetic structure of the endangered species *Pinna nobilis* (Mollusca: Bivalvia) inferred from mtDNA sequences. Biologia. 63,412-417.

Kenchington, E., Bird, C.J., Osborne, J., Reith, M., 2002. Novel repeat elements in the nuclear ribosomal RNA operon of the flat oysters *Ostrea edulis* C. Linnaeus, 1758 and *O. angasi* Sowerby, 1978. J. Shellfish Res. 21,679-705.

Kim, E.M., An, H.S., Kang, J.H., An, C.M., Dong, C.M., Hong, Y.K., Park, J.Y., 2014. New polymorphic microsatellite markers for the Korean manila clam (*Ruditapes philippinarum*) and their application to wild populations. Genet. Mol. Res. 13,8163-8173.

Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Pääbo, S. F, Villablanca, X., Wilson, A.C., 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proc. Natl. Acad. Sci. USA. 86,6196-6200.

Krakau, M., Jacobsen, S., Jensen, K.T., Reise, K., 2012. The cockle *Cerastoderma edule* at Northeast Atlantic shores: genetic signatures of glacial refugia. Mar, Biol. 159,221-230.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33,1870-1874.

Ladoukakis, E.D., Saavedra, C., Magoulas, A., Zouros, E., 2002. Mitochondrial DNA variation in a species with two mitochondrial genomes: the case of *Mytilus galloprovincialis* from the Atlantic, the Mediterranean and the Black Sea. Mol. Ecol. 11,755-769.

Ladoukakis, E.D., Zouros, E., 2001. Direct evidence for homologous recombination in mussel (*Mytilus galloprovincialis*) mitochondrial DNA. Mol. Biol. Evol. 18,1168-75.

Lagardère, F., 1972. Les fonds de pêche de la côte ouest de l'ile d'Oléron. Cartographie bionomique. Les peuplements benthiques. Thétys. 3,507-538.

Lallias, D., Taris, N., Boudry, P., Bonhomme, F., Lapègue, S., 2010. Variance in the reproductive success of flat oyster *Ostrea edulis* L. assessed by parentage analyses in natural and experimental conditions. Genet. Res. 92,175-187.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23,2947-2948.

Lartillot, N., Philippe, H. A., 2004. Bayesian Mixture Model for Across-Site Heterogeneities in the Amino-Acid Replacement Process. Mol. Biol. Evol. 21,1095-1109.

Laslett, D., Canbäck, B., 2008. ARWEN, a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics. 24,172-175.

Launey, S., Ledu, C., Boundry, P., Bonhomme, F., Naciri-Graven, Y., 2002. Geographic structure in the European flat oyster (*Ostrea edulis* L.) as revealed by microsatellite polymorphism. J. Hered. 93,331-338.

Leber, K. M., 1982. Seasonality of macroinvertebrates on a temperature, high wave, energy sandy beach. Bull. Mar. Sci. 32,86-98.

Li, G., Hubert, S., Bucklin, K., Ribes, V., Hedgecock, D., 2003. Characterization of 79 microsatellite DNA markers in the Pacific oyster *Crassostrea gigas*. Mol. Ecol. Notes. 3,228-232.

Li, H., Liang, Y., Sui, L., Gao, X., He, C., 2011a. Characterization of 10 polymorphic microsatellite markers for Mediterranean blue mussel *Mytilus galloprovincialis* by EST database mining and cross-species amplification. J. Genet. 90,e30-e33.

Li, H.J., Liu, W.D., Gao, X.G., Zhu, D., 2011b. Identification of host-defense genes and development of microsatellite markers from ETs of hard clam *Meretrix meretrix*. Mol. Biol. Rep. 38,769-775.

Li, Y.C., Korol, A.B., Fahima, T., Beiles, A., Nevo, E., 2002. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol. Ecol. 11,2453-2465.

Li, Y.C., Korol, A.B., Fahima, T., Nevo, E., 2004. Microsatellites within genes: Structure, function, and evolution. Mol. Biol. Ecol. 21,991-1007.

Lian, C., Zhou, Z., Hogetsu, T., 2001. A simple method for developing microsatellite markers using amplified fragments of inter-simple sequence repeat (ISSR). J. Plant. Res. 114,381-385.

Librado, P., Rozas, J., 2009. DnaSP v5: software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 25,1451-2.

Lim, H.C., Ahmad, A.T., Nuruddin, A.A., Mohd-Nor, S.A., 2016. *Cytochrome b* gene reveals panmixia among Japanese Threadfin Bream, *Nemipterus Japonicus* (Bloch, 1791) populations along the coasts of Peninsular Malaysia and provides evidence of a cryptic species. Mitochondrial DNA A DNA Mapp. Seq. Anal. 27,575-584.

Little, R.D., Braaten, D.C., 1989. Genomic organization of human 5S rDNA and sequence of one tandem repeat. Genomics. 4,376-383.

Liu, H., Cai, S., Zhang, H., Vrijenhoek, R.C., 2016. Complete mitochondrial genome of hydrothermal vent clam *Calyptogena magnifica*. Mitochondrial DNA A DNA Mapp. Seq. Anal. 27,4333-4335.

Liu, H.P., Mitton, J.B., Wu, S.K., 1996. Paternal mitochondrial DNA differentiation far exceeds maternal mitochondrial DNA and allozyme differentiation in the freshwater mussel, *Anodonta grandis grandis*. Evolution. 50, 952-957.

Liu, T., Li, Q., Song, J., Yu, H., 2017. Development of genomic microsatellite multiplex PCR using dye-labeled universal primer and its validation in pedigree analysis of Pacific oyster (*Crassostrea gigas*). J. Ocen Univ. China. 16,151-160.

Liu, Z.J., Cordes, J.F., 2004. DNA marker technologies and their applications in aquaculture genetics. Aquaculture. 238,1-37.

Long, E.O., Dawid, I.B., 1980. Repeated genes in eukaryotes. Annu. Rev. Biochem. 49,727-764.

Lopes-Lima, M., Sousa, R., Teixeira, A., Varandas, S., Riccardi, N., Aldridge, D.C., Froufe, E., 2016. Newly developed microsatellite markers for the pan-european duck mussel, *Anodonta anatina*: revisiting the main mitochondrial lineages. Aquatic Conserv.: Mar. Freshw. Ecosyst. 26,307-318.

López-Piñón, M.J., Freire, R., Insua, A., Méndez, J., 2008. Sequence characterization and phylogenetic analysis of the 5S ribosomal in some scallop (Bivalvia: Pectinidae). Hereditas. 145,9-19.

López-Piñón, M.J., Insua, A., Méndez, J., 2002. Identification of four scallop species using PCR and restriction analysis of the ribosomal DNA internal transcribed spacer region. Mar. Biotechnol. 4,495-502.

Louzán, A., Nóvoa, S., Ojea, J., Da Costa, F., Martínez-Patiño, D., 2016. Culture of the wedge-shell clam Donax trunculus: new developments. In J.L., Romalde, (Eds.), Oysters and Clams: Cultivation, Habitat Threats and Ecological Impact. New York: Nova Science Publishers. pp. 15-18.

Lowe, T.M., Chan, P.P., 2016. tRNAscan-SE On-line: integratinf search and context for analysis of transfer RNA genes. Nucleic Acids Res. 44(Web Server issue):W54-W57.

Luis, J.R., Comesaña, A.S., Sanjuan, A., 2011. mtDNA differentiation in the mussel *Mytilus galloprovincialis* Lmk. on the Iberian Peninsula coast: first results. Mar. Ecol-Evol Persp. 32,102-106.

Lukehurst, S.S., Dias, P.J., Huhn, M., Madduppa, H.H., Lee, S.S.C., Teo, S., Gardner, M.G., McDonald, J., 2017. Isolation and characterization of 16 polymorphic microsatellite loci for the Asian green mussel *Perna viridis* (Mollusca, Mytilidae). Manag. Biol. Invasion. 8,85-88.

Lunt, D.H., Hutchinson, W.F., Carvalho, G.R., 1999. An efficient method for PCR-based isolation of microsatellite arrays (PIMA). Mol. Ecol. 8,891-894.

Maggi, C., González-Wangüemert, M., 2015. Genetic differentiation among *Parastichopus regalis* populations from Western Mediterranean Sea: potential effects of its fishery and current connectivity. Mediterr. Mar. Sci. 16,489-501.

Maia, F., Pimenta, J., 2007. Estudo do crescimento de *Donax vittatus* na costa ocidental norte de Portugal. Dados científicos para a gestão de uma nova pescaria. Relat. Cient. Téc. IPIMAR. Lisboa, Portugal. pp. 1-27.

Manchado, M., Rebordinos, L., Infante, C., 2006b. U1 and U2 Small Nuclear RNA Genetic Linkage: A Novel Molecular Tool for Identification of Six Sole Species (Soleidae, Pleuronectiformes). J. Agric. Food Chem. 54,3765-3767.

Manchado, M., Zuasti, E., Cross, I., Merlo, A., Infante, C., Rebordinos, L., 2006a. Molecular characterization and chromosomal mapping of the 5S rRNA gene in *Solea senegalensis*: a new linkage to the U1, U2 and U5 small nuclear RNA genes. Genome. 49,79-86.

Marie, A.D., Lejeusne, C., Karapatsiou, E., Cuesta, J.A., Drake, P., Macpherson, E., Bernatchez, L., Rico, C., 2016. Implications for management and conservation of the population genetic structure of the wedge clam *Donax trunculus* across two biogeographic boundaries. Sci. Rep. 6,39152.

Marín, A., Fujimoto, T., Arai, K., 2013. Rapid species identification of fresh and processed scallops by multiplex PCR. Food Control. 32,472-476.

Martin, J.F., Pech, N., Meglécz, E. Ferreira, S., Costedoat, C., Dubut, V., Malausa, T., Gilles, A., 2010. Representativeness of microsatellite distribution in genomes, as revealed by 454 GS-FLX titanium pyrosequencing. BMC Genomics. 11,560.

Martínez, A., Mariñas, L., González-Tizón, A., Méndez, J., 2002. Cytogenetic characterization of *Donax trunculus* (Bivalvia: Dionacidae) by means of karyotyping, fluorochrome banding and fluorescent in situ hybridization. J. Molluscan Stud. 68,393-396.

Martínez, L., Arias, A., Méndez, J., Insua, A., Freire, R., 2009. Development of twelve polymorphic microsatellite markers in the edible cockle *Cerastoderma edule* (Bivalvia: Cardiidae). Conserv. Genet. Resour. 1,107-109.

Martínez, L., Freire, R., Arias-Pérez, A., Méndez, J., Insua, A., 2015. Patterns of genetic variation across the distribution range of the cockle *Cerastoderma edule* inferred from microsatellites and mitochondrial DNA. Mar. Biol. 162,1393-1406.

Martínez-Patiño, D., Nóvoa, S., Ojea, J., Rodríguez, E., 2003. Determinación de la talla de primera maduración sexual de la coquina, *Donax trunculus* (Linné, 1758) (Mollusca:Bivalvia), en un banco natural de la Ría del Barquero (NO Galicia). In: Libro de Actas IX Congreso Nacional de Acuicultura. Consejería de Agricultura y Pesca. Junta de Andalucía. pp. 215-217.

Martins, C., Galetti, P.M., 2001. Two 5S rDNA arrays in Neotropical fish species: is it a general rule for fishes? Genetica. 111,439-446.

Martins, C.S., Hamann, M., Fuiza, A.F.G., 2002. Surface circulation in the eastern North Atlantic from drifters and altimetry. J. Geophys. Res. 107,3217.

Masaoka, T., Kobayashi, T., 2005. Natural hybridization between *Pinctada fucata* and *Pinctada maculate* inferred from internal transcribed spacer regions of nuclear ribosomal RNA genes. Fish. Sci. 71,829-836.

May, B., Stoneking, M., Wright, J.E., 1980. Joint segregation of biochemical loci in Salmonidae: II. Linkage associations from a hybridized *Salvelinus* genome (*S. namaycush* X *S. fontinalis*). Genetics. 95,707-726.

McLachlan, A., Wooldridge, T., Van Der Horst, G., 1979. Tidal movements of the macrofaunal on an exposed sandy beach in South Africa. J. Zool. 188,433-442.

Meglécz, E., Pech, N., Gilles, A., Dubut, V., Trilles, A., Grenier, R., Martin, J.F., 2014. QDD version 3.1: a user-friendly computer program for microsatellite selection and primer design revisited: experimental validation of variables determining genotyping success rate. Mol. Ecol. Resour. 14,1302-1313.

Meng, X., Zhao, N., Shen, X., Hao, J., Liang, M., Zhu, X., Cheng, H., Yan, B., Liu, Z., 2012. Complete mitochondrial genome of *Coelomactra antiquata* (Mollusca: Bivalvia): The first representative from the family Mactridae with novel gene order and unusual tandem repeats. Comp. Biochem. Physiol. Part D Genomics Proteomics. 7,175-179.

Merritt, T.J., Shi, L., Chase, M.C., Rex, M.A., Etter, R.J., Quattro, J.M., 1998. Universal cytochrome b primers facilitate intraspecific studies in molluscan taxa. Mol. Mar. Biol. Biotechnol. 7,7-11.

Miller, A.D., Versace, V.L., Matthews, T., Bowie, K.C., 2012. The development of 20 microsatellite loci for the Australian marine mollusk, *Donax deltoides*, through next generation DNA sequencing. Conserv. Genet. Res. 4,257-260.

Miller, L.M., Kapuscinski, A.R., Hallerman, E., 2003. Genetic guidelines for hatchery supplementation programmes. In Press: E. Hallerman (Ed.), Population genetics: principles and

applications for fisheries scientists. Bethesda, Maryland: American Fisheries Society. pp. 329-355.

Molares, J., Parada, J.M., Navarro-Pérez, E., Fernández, A., 2008. Variabilidad interanual de las ventas de los principales recursos marisqueros de Galicia y su relación con las condiciones ambientales. Rev. Gal. Rec. Mar. 2,1-42.

Moritz, C., 1994. Applications of mitochondrial DNA analysis in conservation: a critical review. Mol. Ecol. 3,401-411.

Moritz, C., Brown, W.M., 1986. Tandem duplication of D-loop and ribosomal RNA sequences in lizard mitochondrial DNA. Science. 233,1425-1427.

Moritz, C., Brown, W.M., 1987. Tandem duplications in animal mitochondrial DNAs: variation in incidence and gene content among lizards. Proc. Natl. Acad. Sci. USA. 84,7183-7187.

Moritz, C., Dowling, T.E., Brown, W.M., 1987. Evolution of animal mitochondrial DNA: relevance for biology and systematic. Annu. Rev. Ecol. Evol. Syst. 18,269-292.

Morton, D.G., Sprague, K.U., 1984. In vitro transcription of a silkworm 5S RNA gene requires an upstream signal. Proc. Natl. Acad. Sci. U.S.A. 81,5519-5522.

Morvezen, R., Charrier, G., Boundry, P., Chauyaud, L., Breton, F., Strand, Ø., Laroche, J., 2016. Genetic structure of a commercially exploited bivalve, the great scallop *Pecten maximus*, along the European coasts. Conserv Genet. 17,57-67.

Mouëza, M., Chessel, D., 1976. Contribution à l'étude de la biologie de *Donax trunculus* L. (Mollusque: Lamellibranche) dans l'algerois: Analyse statistique de la dispersion le long d'un plage en baie de bou ismaïl. J. Exp. Mar. Biol. Ecol. 21,211-221.

Mundy, N.I., Winchell, C.S., Woodruf, D.S., 1996. Tandem Repeats and Heteroplasmy in the Mitochondrial DNA Control Region of the Loggerhead Shrike (*Lanius ludovicianus*). J. Hered. 87,21-26.

Naciri, M., Lemaire, C., Borsa, P., Bonhomme, F., 1999. Genetic study of the Atlantic/Mediterranean transition in sea bass (*Dicentrarchus labrax*). J. Hered. 90,591-596.

Nantón, A., Arias-Pérez, A., Freire, R., Fernández-Pérez, J., Nóvoa, S., Méndez, J. 2017. Microsatellite variation in *Donax trunculus* from the Iberian Peninsula, with particular attention to Galician estuaries (NW Spain). Est. Coast. Shelf Sci. 197,27-34.

Nantón, A., Arias-Pérez, A., Méndez, J., Freire, R., 2014. Characterization of nineteen microsatellite markers and development of multiplex PCRs for the wedge clam *Donax trunculus* (Mollusca: Bivalvia). Mol. Biol. Rep. 41,5351-5357.

Nantón, A., Freire, R., Arias-Pérez, A., Gaspar, M.B., Méndez, J., 2015. Identification of four *Donax* species by PCR-RFLP analysis of cytochrome c oxidase subunit I (COI). Eur. Food. Res. Technol. 240,1129-1133.

Nazar, R.N., 1984. The ribosomal 5.8S RNA: eukaryotic adaptation or processing variant? Can. J. Biochem. Cell Biol. 62,311-320.

Neff, B.D., Fu, P., Gross, M.R., 2000. Microsatellite multiplexing in fish. Trans. Am. Fish. Soc. 129,584-93.

Nei, M., 1987. Molecular Evolutionary Genetics. Columbia Univ. Press, New York.

Nelson, D.W., Linning, R.M., Davison, P.J., Honda, B.M.**,** 1998. 5′-flanking sequences required for efficient transcription in vitro of 5S RNA genes, in the related nematodes *Caenorhabditis elegans* and *Caenorhabditis briggsae*. Gene. 218,9-16.

Nie, H., Jiang, L., Gao, Y., Xu, H., Du, S., Yan, X., 2017. Polymorphic microsatellite markers for *Solen grandis* and their cross-species amplification in three other species. Anim. Biotechnol. 1-5.

Nielsen, J.N., Hallenberg, C., Frederiksen, S., Sørensen, P.D, Lomholt, B.**,** 1993. Transcription of human 5S rRNA genes is influenced by an upstream DNA sequence. Nucleic Acids Res**.** 21,3631-3636.

Nilsen, T.W., Shambaugh, J., Denker, J., Chubb, G., Faser, C., Putnam, L., Bennet, K., 1989. Characterization and expression of a spliced leader RNA in the parasitic nematode *Ascaris lumbricoides* var. *suum*. Mol. Cell Biol. 9,3543-3547.

Ning, Z., Cox, A.J., Mullikin, J.C., 2001. SSAHA: A Fast Method for Large DNA Databases. Genome Res. 11,1725-1729.

Odorico, D.M., Miller, D.J., 1997. Variation in the ribosomal internal transcribed spacers and 5.8S rDNA among five species of *Acropora* (Cnidaria: Scleractinia): patterns of variation consistent with reticulate evolution. Mol. Biol. Evol. 14,465-473.

Olsen, E.M., Heino, M., Lilly, G.R., Morgan, M.J., Brattey, J., Ernande, B., Dieckmann, U., 2004. Maturation trends indicative of rapid evolution preceded the collapse of northern cod. Nature 428,932-935.

Ort, B.S., Pogson, G.H. 2007. Molecular population genetics of the male and female mitochondrial DNA molecules of the california sea mussel, *Mytilus californianus.* Genetics. 177,1087-1099.

Ozawa, G., Shimamura, S., Takaki, Y., Yokobori, S.I., Ohara, Y., Takishita, K., Maruyama, T., Fujikura, K., Yoshida, T., 2017. Updated mitochondrial phylogeny of Pteriomorph and Heterodont Bivalvia, including deep-sea chemosymbiotic *Bathymodiolus* mussels, vesicomyd clams and the thyasirid clam *Conchocele* cf. *bisecta*. Mar. Genomics. 31,43-52.

Özden, Ö., Erkan, N., Deval, M.C., 2009. Trace mineral profiles of the bivalve species *Chamelea gallina* and *Donax trunculus*. Food Chem. 113,222-226.

Passamonti, M., 2007. An unusual case of gender-associated mitochondrial DNA heteroplasmy: The mytild *Musculista senhousia* (Mollusca Bivalvia). BMC Evol. Biol. 7(Suppl 2),S7.

Passamonti, M., Boore, J.L., Scali, V., 2003. Molecular evolution and recombination in gender-associated mitochondrial DNAs of the manila clam *Tapes philippinarum*. Genetics. 164,603-611.

Passamonti, M., Ghiselli, F., 2009. Doubly uniparental inheritance: two mitochondrial genomes, one precious model for organelle DNA inheritance and evolution. DNA Cell Biol. 28,79-89.

Passamonti, M., Scali, V. 2001. Gender-associated mitochondrial DNA heteroplasmy in the venerid clam *Tapes philippinarum* (Mollusca Bivalvia). Curr. Genet. 39,117-124.

Paule, M.R., White, R.J., 2000. Transcription by RNA polymerases I and III. Nucleic Acids Res. 28,1283-1298.

Peakall, R., Smouse, P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes. 6,288-295.

Peay, K.G.,Kennedy, P.G., Bruns, T.D., 2008. Fungal community ecology: a hybrid beast with a molecular master. Bioscience. 58,799-810.

Pelliccia, F., Barzotti, R., Bucciarelli, E., Rocchi, A., 2001. 5S ribosomal and U1 small nuclear RNA genes: a new linkage type in the genome of a crustacean that has 3 different tandemly repeated units containing 5S ribosomal sequences. Genome. 44,331-335.

Pelliccia, F., Barzotti, R., Volpi, E.V., Bucciarelli, E., Rocchi, A., 1998. Nucleotide sequence and chromosomal mapping of the 5S rDNA repeat of the crustacean *Proasellus coxalis*. Genome. 41,129-133.

Pendás, A.M., Morán, P., Martínez, J.L., García-Vázquez, E. 1995. Applications of 5S rDNA in Atlantic salmon, brown trout, and in Atlantic salmon brown trout hybrid identification. Mol. Ecol. 4,275-276.

Pereira, A.M., Fernández-Tajes, J., Gaspar, M.B., Méndez, J., 2012. Identification of the wedge clam *Donax trunculus* by a simple PCR technique. Food Control. 23, 268-270.

Pereira, S.L., 2000. Mitochondrial genome organization and vertebrate phylogenetics. Genet Mol Biol. 23,754-752.

Pereira, S.M., 2013. Estudio genético poblacional en la almeja babosa *Venerupis pullastra* (PhD thesis). Universidade da Coruña, Spain.

Pereira, S.M., Arias, A., Méndez, J., Insua, A., Freire, R., 2010. Isolation of twelve microsatellite markers in the pullet carpet shell *Venerupis pullastra* (Bivalvia: Veneridae). Conserv. Genet. Resour. 2, 201-203.

Perina, A., Seoane, D., González-Tizón, A.M., Rodríguez-Fariña, F., Martínez-Lage, A., 2011. Molecular organization and phylogenetic analysis of 5S rDNA in crustaceans of the genus *Pollicipes* reveal birth-and-death evolution and strong purifying selection. BMC Evol. Biol. 11,304-314.

Peterson, R.C., Doering, J.L., Brown, D.D., 1980. Characterization of two *Xenopus* somatic 5S DNAs and one minor oocyte-specific 5S DNA. Cell. 20,131-141.

Petrović, V., Pérez-García, C., Pasantes, J.J., Satović, E., Prats, E., Plohl, M., 2009. A GC-rich satellite DNA and karyology of the bivalve mollusk *Donax trunculus*: a dominance of GC-rich heterochromatin. Cytogenet. Genome Res. 124,63-71.

Petrović, V., Plohl, M., 2005. Sequence divergence and conservation in organizationally distinct subfamilies of *Donax trunculus* satellite DNA. Gene. 362,37-43.

Pieler, T., Hamm, J., Roeder, R.G., 1987. The 5S gene internal control region is composed of three distinct sequence elements, organized as two functional domains with variable spacing. Cell. 48,91-100.

Pinhal, D., Yoshimura, T.S., Araki, C.S., Martins, C., 2011. The 5S rDNA family evolves through concerted and birth-and-death evolution in fish genomes: an example from freshwater stingrays. BMC Evol. Biol. 11,151.

Plazzi, F., 2015. The detection of sex-linked heteroplasmy in *Pseudocardium sachalinense* (Bivalvia: Mactridae) and its implications for the distribution of doubly uniparental inheritance of mitochondrial DNA. J. Zool. Syst. Evol. Res. 53,205-210.

Plazzi, F., Cassano, A., Passamonti, M., 2015. The quest for Doubly Uniparental Inheritance in heterodont bivalves and its detection in *Meretrix lamarckii* (Veneridae: Meretricinae). J. Zool. Syst. Evol. Res. 53,87-94.

Plazzi, F., Puccio, G., Passamonti, M., 2016. Comparative Large-Scale Mitogenomics Evidences Clade-Specific Evolutionary Trends in Mitochondrial DNAs of Bivalvia. Genome Biol. Evol. 8,2544-2564.

Plazzi, F., Ribani, A., Passamonti, M., 2013. The complete mitochondrial genome of *Solemya velum* (Mollusca: Bivalvia) and its relationships with Conchifera. BMC Genomics. 14,409.

Plohl, M., Cornudella, L., 1996. Characterization of a complex satellite DNA in the mollusc *Donax trunculus*: Analysis of sequence variations and divergence. Gene. 169,157-164.

Plohl, M., Cornudella, L., 1997. Characterization of interrelated sequence motifs in four satellite DNAs and their distribution in the genome of the mollusc *Donax trunculus*. J. Mol. Evol. 44,189-198.

Plohl, M., Prats, E., Martínez-Lage, A., González-Tizón, A., Méndez, J., Cornudella, L., 2002. Telomeric localization of the vertebrate-type hexamer repeat, (TTAGGG)(n), in the wedgeshell clam *Donax trunculus* and other marine invertebrate genomes. J. Biol. Chem. 277,19839-19846.

Plough, L.V., Marko, P.B., 2014. Characterization of microsatellite loci and repeat density in the gooseneck barnacle, *Pollicipes elegans*, using Next Generation Sequencing. J. Hered. 105, 136-142.

Presa, P., Pérez, M., Diz, A.P., 2002. Polymorphic microsatellite markers for blue mussels (*Mytilus* spp.). Conserv. Genet. 3,441-443.

Primmer, C.R., Ellegren, H., 1998. Patterns of molecular evolution in avian microsatellites. Mol. Biol. Evol. 15,997-1008.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics. 155,945-959.

Pritchard, J.K., Wen, X., Falush, D., 2010. Documentation for *structure* software: Version 2.3. Available from http://pritch.bsd.uchicago.edu/structure.html

Quesada, H., Beynon, C.M., Skibinski, D.O., 1995. A mitochondrial DNA discontinuity in the mussel *Mytilus galloprovincialis* Lmk: pleistocene vicariance biogeography and secondary intergradation. Mol. Biol. Evol. 12,521-524.

R Core Team, 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Raha, D., Wan, Z., Moqtaderi, Z., Wu, L., Zhong, G., Gerstein, M., Struhl, K., Snyder, M., 2010. Close association of RNA polymerase II and many transcription factors with Pol III genes. Proc. Natl. Acad. Sci. U.S.A. 107,3639-3644.

Ramón, M., Cano, J., Peña, J.B., Campos, M.J., 2005. Current status and perspectives of mollusc (bivalves and gastropods) culture in the Spanish Mediterranean. Bol. Inst. Oceanogr. 21,361-373.

Ramos-Onsis, S.E., Rozas, J., 2002. Statistical properties of new neutrality tests against population growth. Mol. Biol. Evol. 19,2092-2100.

Rasmussen, R.S., Morrissey, M.T., 2008. DNA-Based methods for the identification of commercial fish and seafood species. Compr. Rev. Food Sci. Food Saf. 7,280-295.

Rassmann, K., Schlötterer, C., Tautz, D., 1991. Isolation of simple-sequence loci for use in polymerase chain reaction-based DNA fingerprinting. Electrophoresis. 12,113-118.

Rawson, P.D., Hilbish, T.J., 1995. Evolutionary relationships among the male and female mitochondrial DNA lineages in the *Mytilus edulis* species complex. Mol. Biol. Evol. 12,893-901.

Rebordinos, L., Cross, I., Merlo, A., 2013. High evolutionary dynamism in 5S rDNA of fish: state of the art. Cytogenet. Genome Res. 141,103-113.

Ren, J., Shen, X., Jiang, F., Liu, B., 2010. The mitochondrial genomes of two scallops, *Argopecten irradians* and *Chlamys farreri* (Mollusca: Bivalvia): The most highly rearranged gene order in the family Pectinidae. J. Mol. Evol. 70,57-68.

Ren, J., Shen, X., Sun, M., Jiang, F., Yu, Y., Chi, Z., Liu, B., 2009. The complete mitochondrial genome of the clam *Meretrix petechialis* (Mollusca: Bivalvia: Veneridae). Mitochondrial DNA. 20,78-87.

Ribeiro, P.A., Branco, M., Hawkins, S.J., Santos, A.M., 2010. Recent changes in the distribution of a marine gastropod, *Patella rustica*, across the Iberian Atlantic coast did not result in diminished genetic diversity or increased connectivity. J. Biogeogr.37,1782-1796.

Rice, W.R., 1989. Analyzing tables of statistical tests. Evolution. 43,223-225.

Rico, C., Cuesta, J.A., Drake, P., Macpherson, E., Bernatchez, L., Marie, A.D., 2017. Null alleles are ubiquitous at microsatellite loci in the wedge clam (*Donax trunculus*). PeerJ. 5,e3188.

Riginos, C., Nachman, M.W., 2001.Population subdivisión in marine environments: the contributions of biogeography, geographical distance and discontinuous habitat to genetic differentiation in a blennioid fish, *Axoclinus nigricaudus*. Mol Ecol. 10, 1439-1453.

Ríos, C., Sanz, S., Saavedra, C., Pena, J.B., 2002. Allozyme variation in populations of scallops, *Pecten jacobaeus* (L.) and *P. maximus* (L.) (Bivalvia: Pectinidae), across the Almeria-Oran front. J. Exp. Mar. Biol. Ecol. 267,223-244.

Rogers, A.R., Harpending, H., 1992. Population growth makes waves in the distribution of pairwise genetic differences. Mol. Biol. Evol. 9,552-569.

Rokas, A., Holland, P.W.H., 2000. Rare genomic changes as a tool for phylogenetics. Trends Ecol. Evol. 15,454-459.

Roman, J., Palumbi, S.R., 2004. A global invader at home: population structure of the green crab, *Carcinus maenas*, in Europe. Mol. Ecol. 13,2891-2898.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Hohna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: effective Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61,539-542.

Ross, C.L., Dyer, K.A., Erez, T., Miller, S.J., Jaenike, J., Markow, T.A., 2003. Rapid divergence of microsatellite abundance among species of Drosophila. Molecular Biology and Evolution 20,1143-1157.

Ross, K.G., Shoemaker, D.D., Krieger, M.J.B., DeHeer, C.J., Keller, L., 1999. Assessing genetic structure with multiple classes of molecular markers: a case study involving the introducer fire ant *Solenopsis invicta*. Mol. Biol. Evol. 16,525-543.

Rousset, F., 2008. GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol. Ecol. Res. 8,103-106.

Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. 132,365-386.

Rufino, M.M., Gaspar, M.B., Pereira, A.M., Maynou, F., Monteiro, C.C., 2010. Ecology of megabenthic bivalve communities from Sandy beaches on the south coast of Portugal. Sci. Mar. 74,163-178.

Saito, S., Tamura, K., Aotsuka, T., 2005. Replication origin of mitochondrial DNA in insects. Genetics. 171,1695-16705.

Sajdak, S.L., Reed, K.M., Phillips, R.B., 1998. Intraindividual and interspecies variation in the 5S rDNA of coregonid fish. J. Mol. Evol. 46,680-688.

Sakonju, S., Bogenhagen, F.D., Brown, D.D. 1980. A control region in the center of the 5S RNA gene directs specific initiation of transcription: I. The 5′ border of the region. Cell. 19,13-25.

Salas, C., Tirado, C., Manjón-Cabeza, M.E., 2001. Sublethal foot-predation on Donacidae (Mollusca: Bivalvia). J. Sea Res. 46,43-56.

Salas-Casanova, C., 1987. The Donacidae of the Bay of Malaga (Spain). Taxonomy. Basteria. 51,33-50.

Salas-Casanova, C., Hergueta, E., 1990. The functional morphology of the alimentary canal in *Donax venustus* Poli and *Donax semistriatus* Poli. In: B. Morton (Eds.). The bivalvia. Hong Kong University Press, Hong Kong. pp. 213-222.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular cloning: a laboratory manual. Cold Harbor Spring Press. New York.

Sanjuan, A., Comesaña, A.S., De Carlos, A., 1996. Macrogeographic differentiation by mtDNA restriction site analysis in SW European *Mytilus galloprovincialis* Lmk. J. Exp. Mar. Biol. Ecol. 198,89-100.

Sanjuan, A., Zapata, C., Álvarez, G., 1994. *Mytilus galloprovincialis* and *M. edulis* on the coasts of the Iberian Peninsula. Mar. Ecol. Prog. Ser. 113,131-146.

Sanjuan, A., Zapata, C., Álvarez, G., 1997. Genetic differentiation in *Mytilus galloprovincialis* LMK. throughout the world. Ophelia. 47,13-31.

Santaclara, F.J., Espiñeira, M., Cabado, A.G., Aldaroso, A., Gonzalez-Lavín, N., Vieites, J.M., 2006. Development of a method for the genetic identification of mussel species belonging to *Mytilus*, *Perna*, *Aulacomya*, and other genera. J. Agric. Food Chem. 54,8461-8470.

Santos, A.M., Cabezas, M.P., Travares, A.I., Xavier, R., Branco, M., 2016. tcsBU: a tool to extend TCS network layout and visualization. Bioinformatics. 32,627-628.

Satovic, E., Plohl, M., 2013. Tandem repeat-containing MITEs in the clam *Donax trunculus*. Genome Biol. Evol. 5,2549-2559.

Saunier, A., Garcia, P., Becquet, V., Marsaud, N., Escudié, F., Pante, E., 2014. Mitochondrial genomes of the Baltic clam *Macoma balthica* (Bivalvia: Tellinidae): settings the stage for studying mito-nuclear incompatibilities. BMC Evol. Biol. 14,259.

Saura, M., Faria, R., 2011. Genetic tools for restoration of fish populations. J. Appl. Inchthyol. 27,5-15.

Schlötterer, C., 2000. Evolutionary dynamics of microsatellite DNA. Chromosoma. 109,365-371.

Schlötterer, C., Tautz, D., 1994. Chromosomal homogeneity of *Drosophila* ribosomal DNA arrays suggests intrachromosomal exchanges drive concerted evolution. Curr. Biol. 4,777-783.

Schultz, T.F., Hsing, P.Y., Eng, A., Zelnio, K.A., Thaler, A.D., Carlsson, J., Van Dover, C.L., 2011. Characterization of 18 polymorphic microsatellite loci from *Bathymodiolus manusensis* (Bivalvia, Mytilidae) from deep-sea hrydrothermal vents. Conserv. Genet. Resour. 3,25-27.

Schunter, C., Carreras-Carbonell, J., Macpherson, E., Tintoré, J., Vidal-Vijande, E., Pascual, A., Guidetti, P., Pascual, M., 2011. Matching genetics with oceanography: directional gene flow in a Mediterranean fish species. Mol. Ecol. 20,5167-5181.

Seeb, J.E., Seeb, L.W., 1986. Gene mapping of isozyme loci in chum salmon. J. Hered. 77,399-402.

Sharp, S.J., Garcia, A.D., 1988. Transcription of the *Drosophila melanogaster* 5S RNA gene requires an upstream promoter and four intragenic sequence elements. Mol. Cell Biol. 8,1266-1274.

Shen, Y., Kou, Q., Chen, W., He, S., Yang, M., Li, X., Gan, X., 2016a. Comparative population structure of two dominant species, *Shinkaia crosnieri* (Munidopsidae: *Shinkaia*) and *Bathymodiolus platifrons* (Mytilidae: *Bathymodiolus*), inhabiting both deep-sea vent and cold seep inferred from mitochondrial multi-genes. Ecol. Evol. 6,3571-3582.

Shen, X., Ma, X., Ren, J., Zhao, F., 2009. A close phylogenetic relationship Sipuncula and Annelida evidenced from the complete mitochondrial genome sequence of *Phascolosoma esculenta*. BMC Genomics. 10,136.

Shen, X., Meng, X.P., Chu, K.H., Zhao, N.N., Tian, M., Liang, M., Hao, J., 2014. Comparative mitogenomic analysis reveals cryptic species: A case study in Mactridae (Mollusca: Bivalvia). Comp. Biochem. Physiol. Part D Genomics Proteomics. 12,1-9.

Shen, X., Song, J., Meng, X., Tian, M., Yan, B., Cheng, H., Cai, Y., Zhao, J., 2016b. The first representative of *Coelomactra antiquata* mitochondrial genome from Liaoning (China) and phylogenetic consideration. Mitochondrial DNA Part B. 1,525-527.

Shendure, J., Ji, H., 2008. Next-generation DNA sequencing. Nat. Biotechnol. 26,1135-1145.

Shippen-Lentza, D.E., Vezza, A.C. 1988. The three r5S RNA genes from the human malaria parasite *Plasmodium falciparum* are linked. Mol. Bichem. Parasit. 27,263-273.

Skibinski, D.O.F, Gallagher, C., Beynon, C.M., 1994a. Mitochondrial DNA inheritance. Nature. 368,817-818.

Skibinski, D.O.F, Gallagher, C., Beynon, C.M., 1994b. Sex-limited mitochondrial DNA transmission in the marine mussel *Mytilus edulis*. Genetics. 138,801-809.

Slatkin, M., 1987. Gene flow and the geographic structure of natural populations. Science. 236,787-792.

Slatkin, M., 1994. An exact test for neutrality based on the Ewens sampling distribution. Genet. Res. 64,71-74.

Slatkin, M., 1996. A correction to the exact test based on the Ewens sampling distribution. Genet. Res. 68,259-260.

Slatkin, M., Hudson, R.R., 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics. 129,555-562.

Śmietanka, B., Burzyński, A., Wenne, A., 2009. Molecular population genetics of male and female mitochondrial genomes in European mussels *Mytilus*. Mar. Biol. 156,913-925.

Śmietanka, B., Burzyński, A., Wenne, R., 2010. Comparative genomics of marine mussels (*Mytilus* spp.) gender associated mtDNA: rapidly evolving *atp8*. J. Mol. Evol. 71,385-400.

Śmietanka, B., Zbawicka, M., Sańko, T., Wenne, R., Burzyński, A., 2013. Molecular population genetics of male and female mitochondrial genomes in subartic *Mytilus trossulus*. Mar. Biol. 160,1709-1721.

Smith, D.R., Snyder, M., 2007. Complete Mitochondrial DNA Sequence of the Scallop *Placopecten magellanicus*: Evidence of Transposition Leading to an Uncharacteristically Large Mitochondrial Genome. J. Mol. Evol. 65,380-391.

Sogin, M.L., 1990. In Innes. M.A. et al. PCR Protocols. A Guide to Methods and Applications. Academic Press, San Diego, pp. 307-314.

Sola, L., Rossi, A.R., Annesi, F., Gornung, E., 2003. Cytogenetic studies in *Sparus auratus* (Pisces, Perciformes): molecular organization of 5S rDNA and chromosomal mapping of 5S and 45S ribosomal genes and of telomeric repeats. Hereditas. 139,232-236.

Sotelo, G., Posada, D., Moran, P., 2009. Low-mitochondrial diversity and lack of structure in the velvet swimming crab *Necora puber* along the Galician coast. Mar. Biol. 159,1039-1048.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30,1312-1313.

Stamatakis, A., Hoover, P., Rougemont, J.A., 2008. Rapid Bootstrap Algorithm for the RAxML Web-Servers, Syst. Biol. 57,758-771.

Steele, K.P., Holsinger, K.E., Jansen, R.K., Taylor, D.W., 1991. Assessing the reliability of 5S rRNA sequence data for phylogenetic analysis in green plants. Mol. Biol. Evol. 8,240-248.

Stewart, D.T., Kenchington, E.R., Singh, R.K., Zouros, E., 1996. Degree of selective constraint as an explanation of the different rates of evolution of gender-specific mitochondrial DNA lineages in the mussel *Mytilus*. Genetics. 143,1349-1357.

Stewart, D.T., Saavedra, C., Stanwood, R.R., Ball, A.O., Zouros, E., 1995. Male and female mitochondrial DNA lineages in the blue mussel (*Mytilus edulis*) species group. Mol. Biol. Evol. 12,735-747.

Strasser, C.A., Barber, P.H., 2009. Limited genetic variation and structure in softshell clams (*Mya arenaria*) across their native and introduced range. Conserv. Genet. 10,803-814.

Sutherland, B.W., Stewart, D.T., Kenchington, E., Zouros, E., 1998. The fate of paternal mitochondrial DNA in developing female mussels, *Mytilus edulis*: implications for the mechanism of doubly uniparental inheritance of mitochondrial DNA. Genetics. 148,341-347.

Suzuki, H., Moriwaki, K., Sakurai, S., 1994. Sequences and evolutionary analysis of mouse 5S rDNAs. Mol. Biol. Evol. 11,704-710.

Suzuki, H., Sakurai, S., Matsuda, Y., 1996. Rat 5S rDNA spacer sequences and chromosomal assignment of the genes to the extreme terminal region of chromosome 19. Cytogenet. Cell Genet. 72,1-4.

Sweijd, N.A., Bowie, R.C.K., Evans, B.S., Lopata, A.L., 2000. Molecular genetics and the management and conservation of marine organisms. Hydrobiologia. 420,153-164.

Sørensen, M.V., Dal Zotto, M., Rho, H.S., Herranz, M., Sánchez, N., Pardos, F., Yamasaki, H., 2015. Phylogeny of Kinorhyncha Based on Morphology and Two Molecular loci. PLoS ONE. 10,e0133440.

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

Tang, J., Toè, L., Back, C., Unnasch, T.R., 1996. Intra-specific heterogeneity of the rDNA internal transcribed spacer in the *Simulium damnosum* (Diptera: Simuliidae) complex. Mol. Biol. Evol. 13,244-252.

Tautz, D., Hancock, J.M., Webb, D.A., Tauzt, C., Dover, G., 1988. Complete sequences of the rRNA genes of *Drosophila melanogaster*. Mol. Biol. Evol. 5,366-376.

Taylor, J.D., Williams, S.T., Glover, E.A., Dyal, P., 2007. A molecular phylogeny of heterodonta bivalves (Molusca: Bivalvia: Heterodonta): new analyses of 18S and 28S rRNAgenes. Zool. Scr. 36,587-606.

Tebble, N., 1966. British bivalve sheashells. A Handbook for Identification. London: The British Museum (Natural History).

Thai, B.T., Tan, M.H., Lee, Y.P., Gan, H.M., Tran, T.T., Austin, C.M., 2016. Characterization of 12 microsatellite loci in the Vietnamese commercial clam *Lutraria rhynchaena* Jonas 1844 (Heterodonta: Bivalvia: Mactridae) through next-generation sequencing. Mol. Biol. Rep. 43,391-396.

Thébaud, O., Verón, G., Fifas, S., 2005. Incidences des épisodes d´efflorescences de micro algues toxiques sur les écosystèmes et sur les pêcheries de coquillages en baie de Douarnenez. In: Rapport Ifremer R.INT.DCB/DEM - DCB/STH/UDPP 05-010- Brest, France. pp 88.

Theologidis, I., Fodelianakis, S., Gaspar, M.B., Zouros, E., 2008. Doubly uniparental inheritance (DUI) of mitochondrial DNA in *Donax trunculus* (Bivalvia: Donacidae) and the problem of its sporadic detection in Bivalvia. Evolution. 62,959-70.

Tintore, J., Laviolette, P.E., Blade, I., Cruzado, A., 1988. A study of an intense density front in the eastern Alboran Sea: The Almeria-Oran front. J. Phys. Oceanogr. 18,1384-1397.

Toews, D.P.L., Brelsford, A., 2012. The biogeography of mitochondrial and nuclear discordance in animals. Mol. Ecol. 21,3907-3930.

Toro, J.E., 1998. Molecular identification of four species of mussels from southern Chile by PCR-based nuclear markers: the potential use in studies involving planktonic surveys. J. Shellfish Res. 17,1203-1205.

Tóth, G., Gaspari, Z., Jurka, J., 2000. Microsatellites in different eukaryotic genomes: Survey and analysis. Genome Res. 10,967-981.

Triantafyllidis, A., Apostolidis, A.P., Katsares, V., Kelly, E., Mercer, J., Hughes, M., Jorstad, K., Tsolou, A., Hynes, R., Triantafyllidis, C., 2005. Mitochondrial DNA variation in the European lobster (*Homarus gammarus*) throughout the range. Mar. Biol. 146,223-235.

Trueman, R.D., 1971. The control of burrowing and the migratory behaviour of *Donax denticulatus* (Bivalvia: Tellinacea). J. Zool. 165,453-469.

Tyler, B.M., 1987. Transcription of *Neurospora crassa* 5S rRNA genes requires a TATA Box and three internal elements. J. Mol. Evol. 196,801-811.

Vargas, J., Pérez, M., Toro, J., Astorga, M.P., 2015. Presence of two mitochondrial genomes in the mytild *Perumytilus purpuratus*: Phylogenetic evidence for doubly uniparental inheritance. Genet. Mol. Biol. 38,173-181.

Vierna, J., González-Tizón, A., Martínez-Lage, A., 2009. Long-term evolution of 5S ribosomal DNA seems to be driven by birth-and-death processes and selection in *Ensis* razor shells (Mollusca: Bivalvia). Biochem. Genet. 47,635-644.

Vierna, J., Jensen, K.T., González-Tizón, A.M., Martínez-Lage, A., 2012. Population genetic analysis of *Ensis directus* unveils high genetic variation in the introduced range and reveals a new species from the NW Atlantic. Mar. Biol. 159,2209-2227.

Vierna, J., Jensen, K.T., Martínez-Lage, A., González-Tizón, A.M., 2011. The linked units of 5S rDNA and U1 snDNA of razor shells (Mollusca: Bivalvia: Pharidae). Heredity. 107,127-142.

Vierna, J., Martínez-Lage, A., González-Tizón, A.M., 2010. Analysis of ITS1 and ITS2 sequences in *Ensis* razor shells: suitability as molecular marker at the population and species levels, and evolution of these ribosomal DNA spacers. Genome. 53,23-34.

Vizoso, M., Vierna, J., Gónzález-Tizón, A.M., Martínez-Lage, A., 2011. The 5S rDNA gene family in mollusks: characterization of transcriptional regulatory regions, prediction of secondary structures, and long-term evolution, with special attention to Mytilidae mussels. J. Heredity. 102,433-447.

Wade, B.A., 1964. Notes on the ecology of *Donax denticulatus* (Linné). Proc. Gulf. Carib. Fish. Inst. (17 Ann. Sess.),36-41.

Walsh, M.R., Munch, S.B., Chiba, S., Conover, D.O., 2006. Maladaptive changes in multiple traits caused by fishing: impediments to population recovery. Ecol. Lett. 9,142-148.

Walsh, P.S., Metzger, D.A., Higuchi, R., 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques. 10,506-513.

Wang, H., Zhang, S., Li, Y., Liu, B., 2010. Complete mtDNA of *Meretrix lusoria* (Bivalvia: Veneridae) reveals the presence of an atp8 gene, length variation and heteroplasmy in the control region. Comp Biochem Physiol Part D: Genomics and Proteomics. 5,256-264.

Wang, H., Zhang, S., Xiao, G., Liu, B., 2011. Complete mtDNA of the *Meretrix lamarckii* (Bivalvia: Veneridae) and molecular identification of suspected *M. lamarckii* based on the whole mitochondrial genome. Mar. Genomics. 4,263-71.

Wang, S., Zhang, L., Zhan, A., Wang, X., Liu, Z., Hu, J., Bao, Z., 2007. Patterns of concerted evolution of the rDNA family in a natural population of zhikong scallop, *Chlamys farreri.* J. Mol. Evol. 65,660-667.

Waples, R. S., Gaggiotti, O., 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Mol. Ecol. 15,1419-1439.

Ward, R.D., 2006. The importance of identifying spatial population structure in restocking and stock enhancement programmes. Fish. Res. 80,9-18.

Watterson, G.A., 1978. Homozygosity test of neutrality. Genetics. 88,405-417.

Watterson, G.A., 1986. The homozygosity test after a change in population size. Genetics. 112,899-907.

Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. Evolution. 38, 1358-1370.

Wenne, R., Boudry, P., Hemmer-Hansen, J., Lubieniecki, K.P., Was, A., Kause, A., 2007. What role for genomics in fisheries management and aquaculture? Aquat. Living Resour. 20,241-255.

Wesson, D.M., Porter, C.H., Collins, F.H., 1992. Sequence and secondary structure comparisons of ITS rDNA in mosquitoes (Diptera: Culicidae). Mol. Phylogenet. Evol. 1,253-269.

White, L.R., McPheron, B.A., Stauffer, J.R., 1994 Identification of freshwater mussel glochidia on host fishes using restriction fragment length polymorphisms. Mol. Ecol. 3,183-185.

Whitlock, M.C., McCauley, D.E., 1999. Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm+1)$. Heredity. 82,117-125.

Williams, S.T., Reid, D.G., Littlewood, D.T., 2003. A molecular phylogeny of the Littorininae (Gastropoda: Littorinidae): unequal evolutionary rates, morphological parallelism, and biogeography of the Southern Ocean. Mol. Phylogenet. Evol. 28,60-86.

Wolstenholme, D.R., 1992. Animal Mitochondrial DNA: Structure and Evolution. Int. Rev. Cytol. 141,173-216.

Wood, A.R., Apte, S., MacAvoy, E.S., Gardner, J.P.A., 2007. A molecular phylogeny of the marine mussel genus *Perna* (Bivalvia: Mytilidae) based on nuclear (ITS1&2) and mitochondrial (COI) DNA sequences. Mol. Phylogenet. Evol. 44,685-698.

Wood, V., Gwilliam, R., Rajandream, M.A., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat, N., Hayles, J., Baker, S., et al., 2002. The genome sequence of *Schizosaccharomyces pombe*. Nature. 415,871-880.

Wu, X., Li, X., Li, L., Xu, X., Xia, J., Yu, Z., 2012. New features of Asian *Crassostrea* oyster mitochondrial genomes: a novel alloacceptor tRNA gene recruitment and two novel ORFs. Gene. 507,112-118.

Wu, X., Xiao, S., Li, X., Li, L., Shi, W., Yu, Z., 2014. Evolution of the tRNA gene family in mitochondrial genomes of five *Meretrix* clams (Bivalvia, Veneridae). Gene. 533(1),439-46.

Xavier, R., Santos, A.M., Branco, M., 2012. MtDNA and nuclear data reveal patterns of low genetic differentiation for the isopods *Stenosoma lancifer* and *Stenosoma acuminatum*, with low dispersal ability along the northeast Atlantic coast. Sci. Mar. 76,133-140.

Xu, X., Wu, X., Yu, Z., 2010. The mitogenome of *Paphia euglypta* (Bivalvia: Veneridae) and comparative mitogenomic analyses of three venerids. Genome. 53,1041-52.

Xu, X., Wu, X., Yu, Z., 2012. Comparative studies of the complete mitochondrial genomes of four *Paphia* clams and reconsideration of subgenus Neotapes (Bivalvia: Veneridae). Gene. 494,17-23.

Yao, H., Song, J., Liu, C., Luo, K., Han, J., Li, Y., Pang, X., Xu, H., Zhu, Y., Xiao, P., Chen, S., 2010. Use of ITS2 Region as the Universal DNA Barcode for Plants and Animals. PLoS ONE 5,e13102.

Yasuda, N., Nagai, S., Yamaguchi, S., Lian, C.L., Hamaguchi, M., 2007. Development of microsatellite markers for the Manila clam *Ruditapes philippinarum*. Mol. Ecol. Notes. 7,43-45.

Yu, J.N., Kim, B.J., Yun, Y.E., Kim, H.S., Bae, C.H., Kim, S., Oh, K., Lim, C.E., 2014. Characterization of 11 novel polymorphic microsatellite loci in the threatened Korean loach, *Iksookimia koreensis*, isolated using a next-generation sequencing method. Biochem. Syst. Ecol. 56, 132-137.

Yuan, T., He, M., Huang, L., 2009. Intraespecific genetic variation in mitochondrial 16S rRNA and COI genes in domestic and wild populations of Huaguizhikong scallop *Chlamys nobilis* Reeve. Aquaculture. 289,19-25.

Yuan, Y., Kong, L., Li, Q., 2013. Mitogenome evidence for the existence of cryptic species in *Coelomactra antiquata*. Genes Genom. 35,693-701.

Yuan, Y., Li, Q., Kong, L., Yu, H., 2012b. The complete mitochondrial genome of the grand jackknife clam, *Solen grandis* (Bivalvia: Solenidae): a novel gene order and unusual non-coding región. Mol. Biol. Rep. 39,1287-1292.

Yuan, Y., Li, Q., Kong, L., Yu, H., 2012c. The complete mitochondrial genome of *Solen strictus* (Bivalvia: Solenidae). Mitochondrial DNA A DNA Mapp. Seq. Anal. 23,112-114.

Yuan, Y., Li, Q., Yu, H., Kong, L., 2012a. The Complete Mitochondrial Genomes of Six Heterodont Bivalves (Tellinoidea and Solenoidea): Variable Gene Arrangements and Phylogenetic Implications. PLoS ONE. 7,e32353.

Zane, L., Bargelloni, L., Patarnello, T., 2002. Strategies for microsatellite isolation: a review. Mol. Ecol. 11,1-16.

Zane, L., Ostellari, L., Maccatrozzo, L., Bargelloni, L., Cuzin-Roudy, J., Buchholz, F., Patarnello, T., 2000. Genetic differentiation in a pelagic crustacean (*Meganyctiphanes norvegica*: Euphausiacea) from the North East Atlantic and the Mediterranean Sea. Mar. Biol. 136,191-199.

Zeichen, M.M., Agnesi, S., Mariani, A., Maccaroni, A., Ardizzone, G.D., 2002. Biology and population dynamics of *Donax trunculus* L. (Bivalvia: Donacidae) in the South Adriatic Coast (Italy). Estuar. Coast. Shelf Sci. 54,971-982.

Zevering, C.E., Moritz, C., Heideman, A., Sturm, R.A., 1991. Parallel origins of duplications and the formation of pseudogenes in mitochondrial DNA from parthenogenetic lizards (*Heteronotia binoei*; Gekkonidae). J. Mol. Evol. 33,431-441.

Zhang, G.F., Yan, X.W., 2006. A new three-phase culture method for Manila clam, *Ruditapes philippinarum*, farming in northern China. Aquaculture. 258,452-461.

Zhang, J., Chiodini, R., Badr, A., Zhang, G., 2011. The impact of next-generation sequencing on genomics. J. Genet. Genomics. 38,95-109.

Zhang, J., Li, Q., Wang, Q., Cong, R., Ge, J., Kong, L., 2018. The impact of successive mass selection on population genetic structure in the Pacific oyster (*Crassostrea gigas*) revealed by microsatellite markers. Aquacult. Int. 26,113-125.

Zhou, Y., Zhang, J.Y., Zheng, R.Q., Yu, B.G., Yang, G., 2009. Complete nucleotide sequence and gene organization of the mitochondrial genome of *Paa spinosa* (Anura: Ranoidae). Gene. 447,86-96.

Zink, R.M., Barrowclough, G.F., 2008. Mitochondrial DNA under siege in avian phylogeography. Mol. Ecol. 17,2107-2121.

Zitari-Chatti, R., Chatti, N., Fulgione, D., Caiazza, I., Aprea, G., Elouaer, A., Said, K., Capriglione, T., 2009. Mitochondrial DNA variation in the caramote prawn *Penaeus* (*Melicertus*) *kerathurus* across a transition zone in the Mediterranean Sea. Genetica. 136,439-447.

Zouros, E., 2013. Biparental Inheritance Through Uniparental Transmission: The Doubly Uniparental Inheritance (DUI) of Mitochondrial DNA. Evol. Biol. 40,1-31.

Zouros, E., Ball, A.O., Saavedra, K., Freeman, R., 1994a. An unusual type of mitochondrial DNA inheritance in the blue mussel *Mytilus*. Proc Natl Acad Sci U S A. 91,7463-7467.

Zouros, E., Ball, A.O., Saavedra, K., Freeman, R., 1994b. Mitochondrial DNA inheritance. Nature. 368,818.