

# UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Agronomia, Animali, Alimenti, Risorse Naturali e Ambiente

Dipartimento di Biologia

Corso di laurea magistrale in Scienze e Tecnologie per

l'Ambiente e il Territorio

## **Caratterizzazione e variabilità genetica in specie di pesci antartici appartenenti ai generi *Lepidonotothen* e *Nototheniops* (Perciformes, Notothenioidei).**

**Relatore:**

Prof.ssa Chiara Papetti

**Correlatore:**

Dott. Luca Schiavon

**Laureanda:**

Cecilia Melarancio

ANNO ACCADEMICO 2021/2022



## INDICE

<b>1. INTRODUZIONE</b> .....	<b>2</b>
1.1 Ambiente Antartico .....	3
1.2 Fauna dell'Oceano Meridionale.....	8
1.3 Famiglia Nototheniidae .....	11
1.3.1 Genere <i>Lepidonotothen</i> .....	12
1.3.2 Genere <i>Nototheniops</i> .....	14
1.4 Radiazione adattativa.....	19
1.4.1 <i>La radiazione adattativa nei Notothenioidei</i> .....	22
1.5 Struttura di popolazione nei Notothenioidei.....	28
<b>2. SCOPO DELLA TESI</b> .....	<b>33</b>
<b>3. MATERIALI E METODI</b> .....	<b>35</b>
3.1 Campionamento ed estrazione del DNA .....	35
3.2 Marcatori microsatellite.....	37
3.2.1 <i>Amplificazione dei marcatori microsatellite tramite PCR</i> .....	42
3.2.2 <i>Genotipizzazione o scoring degli alleli</i> .....	43
3.2.3 <i>Binning o allele calling</i> .....	45
3.2.4 <i>Correzione dataset</i> .....	46
3.3 DNA mitocondriale (mtDNA).....	46
3.3.1 <i>Amplificazione del DNA mitocondriale tramite PCR</i> .....	47
3.4 Analisi dei dati .....	48
3.4.1 <i>Marcatori microsatellite</i> .....	50
3.4.2 <i>DNA mitocondriale</i> .....	61
<b>4. RISULTATI</b> .....	<b>63</b>
4.1 Marcatori microsatellite.....	63
4.1.1 <i>Correzione dataset</i> .....	63
4.1.2 <i>Equilibrio di Hardy-Weinberg e linkage disequilibrium</i> .....	67
4.1.3 <i>Numero di alleli e frequenze alleliche</i> .....	69
4.1.4 <i>Allelic richness</i> .....	71

4.1.5	<i>Eterozigosità osservata ed attesa</i> .....	72
4.1.6	<i>Differenziamento genetico inter ed intraspecifico</i> .....	76
4.1.7	<i>Identificazione di gruppi geneticamente distinti</i> .....	78
4.2	DNA mitocondriale .....	90
4.2.1	<i>Statistiche di base e variabilità genetica</i> .....	90
4.2.2	<i>Differenziamento genetico inter ed intraspecifico</i> .....	91
<b>5.</b>	<b>DISCUSSIONE</b> .....	<b>100</b>
<b>6.</b>	<b>CONCLUSIONI</b> .....	<b>106</b>
<b>7.</b>	<b>BIBLIOGRAFIA</b> .....	<b>108</b>
<b>8.</b>	<b>APPENDICE</b> .....	<b>121</b>

## 1. INTRODUZIONE

La conoscenza della composizione in specie nella regione antartica è un prerequisito fondamentale per gli studi di biologia evolutiva, ecologia, biogeografia e genetica della conservazione.

L'identificazione delle specie, fino agli anni '60, avveniva mediante l'analisi della morfologia, delle caratteristiche comportamentali e delle caratteristiche fisiologiche. Questi metodi portarono ad evidenti incertezze tassonomiche dovute alla presenza di specie criptiche, all'esistenza di ibridi interspecifici e a fenomeni di evoluzione convergente e plasticità morfologica. Le specie criptiche sono specie diverse, isolate dal punto di vista riproduttivo e diverse geneticamente, ma così simili tra loro da essere indistinguibili. Specie ad apparente ampia distribuzione possono essere in realtà costituite da numerose specie criptiche con distribuzione e numerosità molto ridotta, il cui prelievo in natura può portare ad estinzione.

Dalla seconda metà del secolo scorso l'identificazione delle specie iniziò ad avvalersi di informazioni sulle componenti molecolari delle cellule (macromolecole). L'analisi di proteine fu il primo metodo molecolare impiegato per l'identificazione di specie, tramite l'individuazione di allozimi (varianti di proteine codificate da alleli diversi di uno stesso gene). Le differenti varianti della proteina presentano differente carica elettrica netta evidenziabile mediante un'elettroforesi proteica.

In seguito alla scoperta, nel 1953 da parte di Watson e Crick, della struttura molecolare del DNA, il suo utilizzo negli studi molecolari apportò numerosi vantaggi. Il DNA è stabile anche se sottoposto a condizioni estreme, è universale poiché presente in tutti i tessuti e fluidi biologici con cellule nucleate ed è più informativo rispetto alle proteine. Da quel momento è stato possibile condurre ricerche il cui fine è l'identificazione, la descrizione e la quantificazione della variabilità genetica inter e intraspecifica, tra popolazioni e tra individui. All'interno delle sequenze di DNA è possibile individuare delle regioni, dette *loci*, che presentano un livello adeguato di polimorfismo. Questi *loci* particolari

sono definiti marcatori molecolari e sfruttano il fenomeno biologico della diversità genetica, necessaria per la caratterizzazione di un dato sistema genetico mediante l'individuazione delle differenti varianti alleliche a livello di *loci* diversi. La scelta del *locus*/marcatore è di estrema importanza e si compie considerando l'informazione ottenuta in relazione alle risorse investite e allo scopo perseguito. Vale la regola generale che minore è la distanza genetica tra gli individui (o i gruppi di individui) che si vuole confrontare, maggiore deve essere il grado di variabilità della regione che si vuole analizzare.

In questa tesi sono stati analizzati i dati genetici di *Lepidonotothen kempfi*, *Lepidonotothen squamifrons*, *Nototheniops larseni* e *Nototheniops nudifrons*, quattro specie di pesci antartici. L'obiettivo principale è di stabilire se gli individui assegnati alla specie *Lepidonotothen kempfi* appartengono effettivamente ad una specie distinta oppure se costituiscono una popolazione della specie *Lepidonotothen squamifrons*. Le analisi sono state condotte su due tipi di marcatori molecolari: 21 *loci* microsatellite e il gene mitocondriale ND2.

## **1.1 Ambiente Antartico**

Il continente antartico si estende per circa 13.8 milioni di km<sup>2</sup> ed è un ambiente particolare, con caratteristiche uniche che lo rendono diverso da tutti gli altri continenti della Terra. Caratterizzato da una forma vagamente circolare, si estende intorno al Polo Sud ed è completamente confinato oltre i 60 °S. Il Canale di Drake, largo circa 645 km, separa la Penisola Antartica, la porzione più settentrionale del continente, dall'America Meridionale. I monti Transantartici e due ampie insenature, occupate dal Mare di Ross e dal Mare di Weddell, suddividono il continente in due porzioni principali: l'Antartide Orientale e l'Antartide Occidentale. La maggior parte delle isole e degli arcipelaghi si trovano a ridosso delle coste, soprattutto in corrispondenza della Penisola Antartica. Una calotta glaciale ricopre quasi totalmente (~98%) la

superficie emersa del continente e costituisce la più importante riserva d'acqua dolce del nostro pianeta.

L'Oceano Meridionale (chiamato anche Oceano Antartico o Oceano Australe) circonda il continente antartico e presenta una profondità media compresa tra i 4000 e i 5000 m. L'Oceano Meridionale è un sistema di bacini profondi separati da tre ampie creste oceaniche sottomarine: la Cresta di Macquarie a sud della Nuova Zelanda e della Tasmania, la Cresta di Kerguelen-Gaussberg a circa 80°E e l'Arco di Scozia che si estende ad arco verso est, dalla piattaforma meridionale della Patagonia alle isole South Shetland e alla Penisola Antartica (Kock, 1992).

La caratteristica peculiare di questo oceano è di non essere contenuto in un bacino distinto delimitato da terre emerse. Esso è delimitato e definito a sud dal continente antartico e a nord da un fattore oceanografico, il Fronte Polare Antartico (solitamente indicato come APF, Antarctic Polar Front). L'APF è situato circa tra 47 °S e 63 °S (Shaw *et al.*, 2004), dove l'acqua superficiale antartica che fluisce verso nord affonda al di sotto dell'acqua sub-antartica più calda e meno densa. In prossimità del Fronte Polare Antartico, o Convergenza Antartica, la temperatura superficiale dell'acqua subisce un calo di temperatura di 2-3 °C, oltre al cambiamento di numerosi altri parametri oceanografici, come la salinità (Kock, 1992). L'area compresa tra il Fronte Polare Antartico e il margine continentale comprende oltre 35 milioni di km<sup>2</sup>, circa il 10% degli oceani mondiali (Eastman, 2005).

La piattaforma continentale antartica è stretta, ad eccezione della porzione nel Mare di Weddell e nel Mare di Ross (Gon e Heemstra, 1990), e rappresenta solo il 3-5% dell'area totale dell'Oceano Meridionale (Kock, 1992). L'isostasi e l'erosione glaciale, insieme all'assenza di sedimentazione fluviale, fanno sì che la piattaforma continentale sia caratterizzata da una profondità media di circa 500 m, circa otto volte la media mondiale (Eastman, 2005). In alcuni punti sono presenti solchi profondi fino a 900-1000 m e che corrono paralleli alla

costa (Gon e Heemstra, 1990), determinando una superficie topografica molto irregolare.

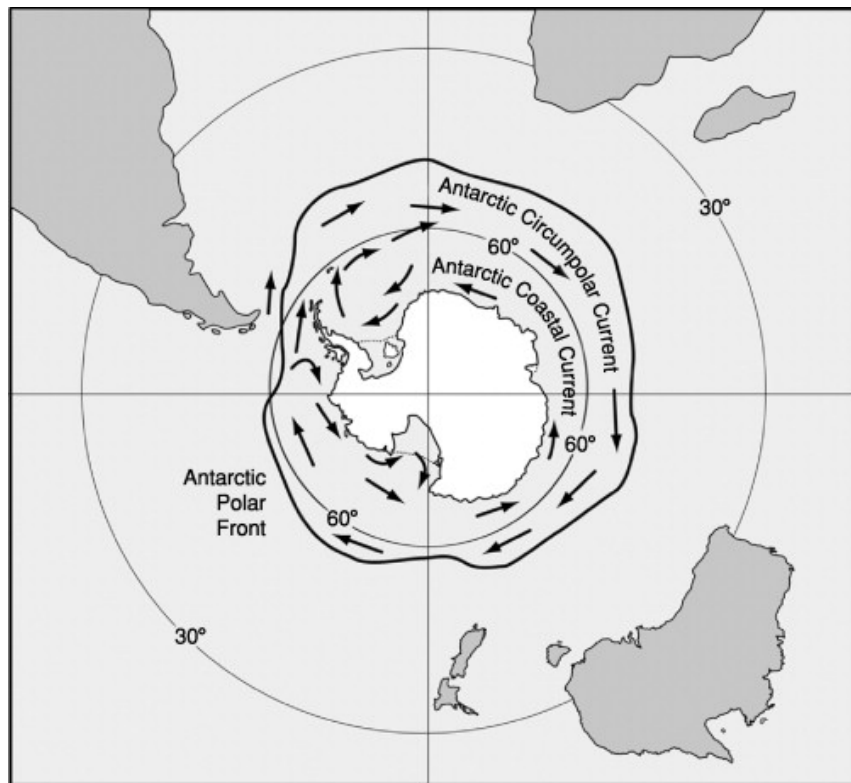
Le acque dell'Oceano Meridionale sono ambienti marini unici, caratterizzati da temperature sotto lo zero, estese fluttuazioni stagionali della produzione primaria e la presenza diffusa di ghiaccio marino (banchisa). La banchisa si forma attorno al continente antartico e presenta uno spessore di 2-3 m. Essa si forma per il congelamento dell'acqua marina e presenta un'estensione che va da un minimo di 3-4 milioni di km<sup>2</sup> in estate fino ad un massimo di 18-20 milioni di km<sup>2</sup> in inverno, estendendosi ampiamente al di sopra della piattaforma continentale (Griffiths, 2010).

All'interno della banchisa restano localmente isolate aree di mare completamente libere dai ghiacci, note come *polynya* (Eastman, 1993). La temperatura di porzioni estese dell'Oceano Meridionale sono al di sotto di 0°C e c'è una differenza di meno di 2 °C nelle temperature medie dell'acqua superficiale tra i mesi più caldi e più freddi, mentre la differenza di temperatura tra la superficie e il fondo è di soli 4-5 °C. In prossimità della piattaforma continentale la temperatura dell'acqua è sempre prossima al punto di congelamento dell'acqua marina (-1.86 °C) ed è termicamente stabile lungo tutto l'anno (Eastman, 1993).

L'Oceano Meridionale svolge un ruolo di primo piano nella regolazione del clima a scala globale e funge da motore alla circolazione oceanica globale (Griffiths, 2010).

Le correnti dell'Oceano Meridionale si possono suddividere in due categorie (Gon e Heemstra, 1990) e si estendono fino a grandi profondità (Tomczak e Godfrey, 2003). La prima categoria riguarda l'avvezione orizzontale degli strati superiori che genera la Corrente Circumpolare Antartica (chiamata anche *West Wind Drift*) e la Corrente Costiera Antartica (detta anche *East Wind Drift*) (figura 1.1). La seconda categoria consiste invece in un regime di flusso in profondità caratterizzato da un movimento verso sud di masse d'acqua profonde (Kock, 1992).



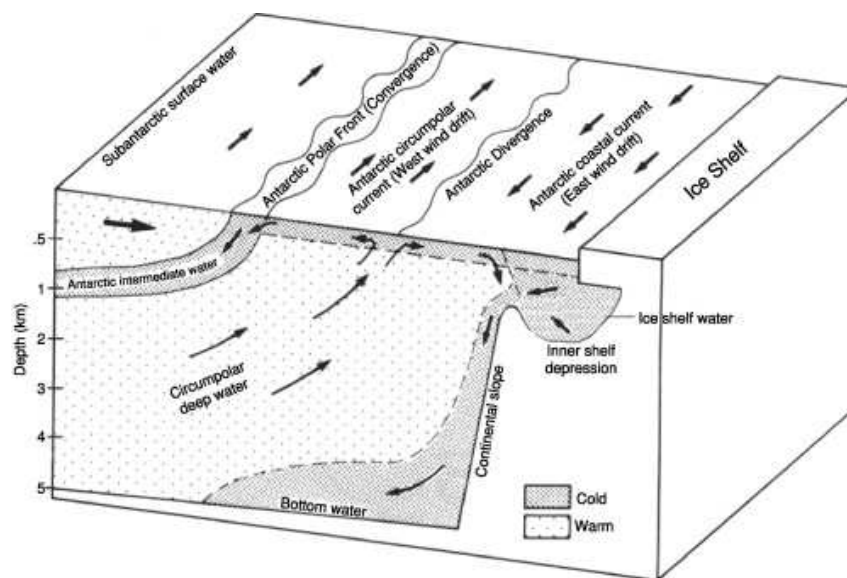


**Figura 1.1:** Rappresentazione schematica della Corrente Circumpolare Antartica (*West Wind Drift*) e della Corrente Costiera Antartica (*East Wind Drift*) intorno al continente antartico. La linea continua indica la posizione approssimativa del Fronte Polare Antartico, o Convergenza Antartica (DeVries e Steffensen, 2005).

La Corrente Circumpolare Antartica (indicata come ACC, *Antarctic Circumpolar Current*) costituisce la maggior parte del flusso oceanico superficiale ed è alimentata da fortissimi venti provenienti da occidente. Fluisce da ovest verso est, si estende dalla superficie del mare fino a circa 4000 m di profondità (Lebedev, 2006) e può essere larga fino a 2000 km e lunga 23000 km. La ACC è inoltre la corrente che trasporta la più grande massa d'acqua di tutti gli oceani, circa 100-150 milioni di metri cubi d'acqua al secondo (Klinck e Nowlin, 2001). L'ACC collega l'Oceano Atlantico, Pacifico e Indiano fungendo da principale percorso di scambio di masse d'acqua, calore ed altre proprietà. Scorrendo in senso orario attorno al continente la ACC è inoltre il principale meccanismo di trasporto circumantartico, in particolare di organismi marini.

La Corrente Costiera Antartica (nota come AACC, *Antarctic Coastal Current*) fluisce in direzione opposta, scorrendo verso ovest in prossimità della costa.

Queste due correnti opposte sono separate dalla Divergenza Antartica a circa 65 °S, una zona dove avviene una consistente risalita di acque profonde circumpolari (figura 1.2). La loro interazione determina la formazione di due importanti vortici oceanici che si muovono in senso orario, presenti nel Mare di Weddell (*Weddell Gyre*) e nel Mare di Ross (*Ross Gyre*).



**Figura 1.2:** Sezione trasversale dell'Oceano Meridionale che mostra l'andamento delle principali correnti superficiali e il movimento delle masse d'acqua calde e fredde in profondità (DeVries e Steffensen, 2005).

Nutrienti come i nitrati (0.16 mmol/kg a 50 m di profondità) e i fosfati (1.55 mmol/kg a 50 m di profondità) si trovano ad alte concentrazioni, così come il silicato (0.10 mmol/kg a 50 m di profondità). I livelli di nutrienti sono spesso più alti nelle aree vicine al continente e nel Mare di Weddell, dove i valori dei silicati possono superare i 60 mmol/kg. In generale, l'Oceano Australe è considerato ricco di sostanze nutritive ma povero di clorofilla (Griffiths, 2010).

La piattaforma continentale intorno al continente antartico include alcune delle aree meno conosciute al mondo, ma nonostante questo le ricerche compiute fino ad oggi indicano che vi è una fauna molto più ricca e varia

di quello che ci aspetterebbe in una zona caratterizzata da condizioni climatiche estreme. Molte specie sono endemiche dell'Oceano Meridionale, o addirittura di aree ancora più ristrette, e questa zona ha subito una limitata pressione antropica (Barnes e Peck, 2008).

## **1.2 Fauna dell'Oceano Meridionale**

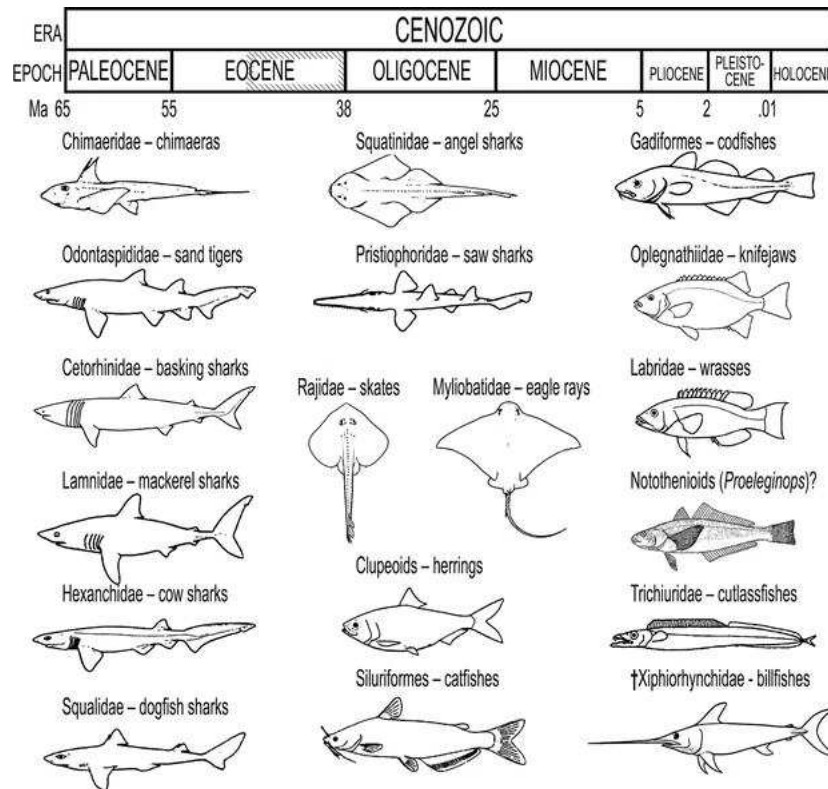
L'Oceano Meridionale ospita una ricca e variegata comunità marina che prospera in un ambiente dominato da ghiaccio e forti correnti. I primi studi di biologia marina in questa regione risalgono al XIX secolo, ma nonostante questo si sa ancora relativamente poco circa le complesse interazioni tra l'ambiente fisico e le specie che vi abitano (Griffiths, 2010). Secondo il Registro delle Specie Marine Antartiche (noto come RAMS, Register of Antarctic Marine Species) le specie presenti nell'ambiente marino antartico sono circa 8400 (De Broyer *et al.*, 2021), distribuite nei tre regni dell'Oceano Meridionale: il fondale, la colonna d'acqua e il ghiaccio marino.

Anche se l'Oceano Meridionale rappresenta il 10% degli Oceani della Terra, la sua fauna attuale consiste di sole 322 specie di pesci antartici raggruppati in 50 famiglie, poche considerando la diversità globale di circa 25000-28000 specie (Eastman, 2005). Tuttavia, il numero di specie dovrebbe essere considerato ragionevolmente alto vista l'assenza di alcuni habitat bentonici chiave sulla piattaforma continentale causata dalla presenza di ghiaccio ancorato che ricopre il fondale fino a profondità di 30 m e di iceberg che erodono il substrato fino a profondità di 550 m. La calotta glaciale antartica e il ghiaccio marino occupano la maggior parte delle aree costiere causando la scomparsa di spiagge, zone intertidali e subtidali, scogliere ed estuari. In questo modo gli habitat costieri poco profondi, aree di primaria importanza per la diversità dei pesci nelle altre parti del mondo, sono ristretti (Eastman, 2005).

La composizione della ittiofauna antartica è cambiata nel corso degli ultimi 40 milioni di anni, dall'Eocene fino ad oggi, a causa di eventi

tettonici, climatici e oceanografici che si sono susseguiti nel tempo. Attualmente nell'Oceano Meridionale è presente una fauna endemica e tassonomicamente ristretta che è succeduta alla fauna dell'Eocene cosmopolita e diversificata tassonomicamente (Eastman, 2005).

La fauna fossile dell'Eocene medio-superiore (~ 40 milioni di anni fa) è rappresentata da *La Meseta Formation* (figura 1.3) a Seymour Island, vicino alla sommità della Penisola Antartica. Essa fornisce un'idea di come fosse composta in quel periodo una fauna di piattaforma continentale a latitudini di 60 °S: una fauna tipica di acque con temperature che andavano dai 10 ai 15 °C (Eastman, 1993). Questa era cosmopolita e costituita da almeno 29 taxa, inclusi 21 pesci cartilaginei e 8 attinopterigi. Razze e Gadiformi sono gli unici taxa presenti ancora oggi (Eastman, 2005). La presenza di Notothenioidei (l'attuale taxon dominante nell'Oceano Meridionale) nella fauna fossile dell'Eocene è ancora aperta a dibattito, anche se Balushkin (1994) attribuì ad un esemplare basale, appartenente alla famiglia Eleginopsidae, i resti fossili di un neurocranio che inizialmente era stato attribuito ad un gadiforme. La formazione dei Notothenioidei come un ramo separato dei pesci perciformi avvenne nelle acque tiepide e poco profonde che connettevano l'Australia, l'Antartide e il Sud America. Alla fine dell'Eocene, ci fu la separazione di almeno quattro famiglie di Notothenioidei tra cui i Nototheniidae (Balushkin, 2000). I cambiamenti climatici e geologici che avvennero in questa regione definirono largamente la successiva storia evolutiva di queste famiglie.



**Figura 1.3:** Rappresentazione della fauna fossile della *La Meseta Formation* a Seymour Island tra la metà e la fine dell'Eocene (~40 milioni di anni fa) (Eastman, 2005).

Nella piattaforma continentale e nella parte alta della scarpata continentale la fauna attuale include 222 specie (19 famiglie) di pesci bentonici. I taxa più numerosi sono i Notothenioidei, i Liparidae e i Zoarcidae, che da soli costituiscono circa l'88% della ricchezza di specie (Eastman, 2005). I due gruppi più grandi (Notothenioidei e Liparidae) occupano principalmente habitat non sovrapposti in quanto i Liparidae si trovano quasi esclusivamente nel mare profondo sotto gli 800 m di profondità e sono poco abbondanti, mentre i Notothenioidei antartici dominano la piattaforma continentale e la scarpata continentale alta (Matschiner *et al.*, 2015).

Il tasso di endemismo delle specie di pesci antartici è dell'88%, circa tre volte tanto il valore riscontrato per altre specie presenti in località marine isolate, ed è dovuto all'isolamento termico e geografico. La fauna è ragionevolmente ben conosciuta da una prospettiva tassonomica ma nuove specie vengono regolarmente descritte. Una stima conservativa

indica che devono essere ancora scoperte circa 30-60 specie, molte delle quali Liparidae (Eastman, 2005).

Il sottordine dei Notothenioidei è costituito da 140 specie (8 famiglie e 45 generi), di cui 30 non sono antartiche, 33 sono sub-antartiche e 77 sono antartiche (Eastman e Eakin, 2021). Rispetto ad una prima lista di controllo del 2000 il numero di specie è aumentato di 18 (15%). Costituiscono il 77% della ricchezza di specie di pesci, il 92% dell'abbondanza e il 91% della biomassa nella piattaforma e scarpata continentale dell'Antartide e delle isole sub-antartiche (Eastman, 2005). La famiglia con distribuzione geografica più estesa è quella dei Bovichtidae che comprende 9 specie in 3 generi (Matschiner *et al.*, 2015). I Notothenioidei mostrano una varietà di strategie di vita adulta che vanno da stili di vita bentonici a stili di vita pelagici. Comprendono specie importanti non solo dal punto di vista economico (Kock, 1992), ma anche da punto di vista trofico, fungendo da collegamento essenziale tra i livelli più bassi (krill e altri invertebrati) e i livelli più alti (balene, foche, pinguini) della catena alimentare antartica (La Mesa *et al.*, 2004). Questo clade rappresenta uno dei migliori esempi di radiazione adattativa in pesci marini (Eastman e Eakin, 2000).

### **1.3 Famiglia Nototheniidae**

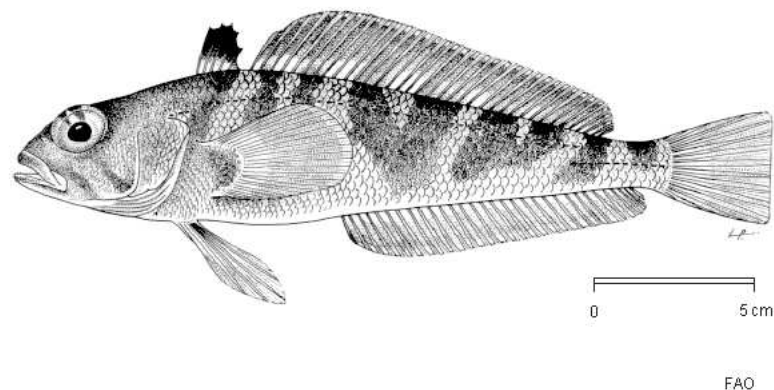
La famiglia dei Nototheniidae è costituita da 57 specie suddivise in 16 generi, tutti presenti nella regione dell'Arco di Scozia (Balushkin, 2000). Questa famiglia è andata incontro alla più importante diversificazione ecologica e morfologica in relazione alla dimensione, alla forma del corpo e alla distribuzione (Gon e Heemstra, 1990). Le diversificazioni morfologiche ed ecologiche riguardano principalmente la galleggibilità (Eastman, 2005), permettendo loro di riempire nicchie ecologiche che, in acque temperate, vengono occupate da pesci tassonomicamente differenti (Gon e Heemstra, 1990). Circa metà delle specie antartiche occupano l'habitat ancestrale bentonico, ma altre sono semipelagiche,

epibentoniche, criopelagiche e pelagiche (Eastman, 1993). Questo tipo di radiazione, chiamata *pelagization*, è avvenuta indipendentemente in numerosi cladi, portando raramente a galleggiabilità neutra. Alcune specie mostrano un cambiamento di habitat passando dalla fase giovanile a quella adulta.

Essi vanno da specie bentoniche di 15-65 cm che si cibano di anfipodi e krill antartico a specie pelagiche lunghe fino a 2 metri che si cibano principalmente di molluschi come i calamari. In base alla specie, la maturità sessuale avviene tipicamente tra i 15 e i 45 cm (misurata come lunghezza totale LT) e tra i 5 e i 9 anni di età. La fecondità è bassa o modesta, generalmente con una densità totale di 5-80 uova per grammo. Le uova possono essere sia demersali che pelagiche e misurano circa 1.5-6 mm di diametro. La stagione riproduttiva avviene comunemente tra il tardo inverno e la prima estate mentre le larve presentano spesso una fase pelagica della durata di diversi mesi (North, 2001).

### 1.3.1 Genere *Lepidonotothen*

Attualmente, secondo Eastman e Eakin (2021), il genere *Lepidonotothen* comprende solo la specie *Lepidonotothen squamifrons* (Günther 1880), con distribuzione biogeografica antartica.



**Figura 1.4:** Illustrazione di un esemplare di *Lepidonotothen squamifrons* (Fischer e Hureau, 1985).

Nel corso degli anni c'è stata poca chiarezza in merito al numero effettivo di specie che fanno parte di questo genere. DeWitt (1966) descrisse il gruppo 'squamifrons' formato dalle specie *L. squamifrons* (Günther 1880), *L. kempfi* (Norman 1937) e *L. macrophthalma* (Norman 1937) con differenti distribuzioni geografiche mentre Permitin e Sazonov (1974) divisero *L. squamifrons* in due sub-specie: *L. s. squamifrons* nel settore dell'Oceano Indiano e *L. s. atlantica* nel settore dell'Oceano Atlantico. Successivamente, basandosi su dati derivati da elettroforesi enzimatica, Schneppenheim *et al.* (1994) stabilirono che le specie che si sovrapponevano dal punto di vista morfologico e meristico (*L. squamifrons*, *L. kempfi* e *L. macrophthalma*) fossero in realtà una sola specie: *L. squamifrons* (Günther 1880) e le altre due specie furono poste in sinonimia. Balushkin (2000), al contrario, stabilì che il genere comprendeva due specie: *L. kempfi* (Norman, 1937) e *L. squamifrons* (Günther 1880). Uno studio recente di Miya *et al.* (2016), basato sull'analisi di relazioni filogenetiche e di divergenza di sequenze di DNA mitocondriale e nucleare, indica che *L. kempfi* è la stessa specie di *L. squamifrons*. Le sequenze di DNA analizzate (geni mitocondriali ND2 e COI; primo introne del gene nucleare S7) mostrano tassi di divergenza bassi (da 0% a 0.4%), suggerendo l'esistenza di due popolazioni della stessa specie.

*Lepidonotothen squamifrons* (*sensu* Günther 1880) (figura 1.4) è una specie bentonica che si trova a profondità comprese tra i 10-900 m, con una predilezione per habitat a profondità comprese tra i 195-312 m con temperature intorno a 0 °C (Gon e Heemstra, 1990; Damerou *et al.*, 2012). Questa specie è presente nell'Oceano Meridionale estendendosi a nord fino all'isola di South Georgia e all'isola di Bouvet. È stato trovato anche nelle isole sub-antartiche e nelle porzioni più meridionali degli Oceani Indiano e Pacifico.

*Lepidonotothen squamifrons* presenta un corpo oblungo e compresso verso la coda. Mediamente raggiunge una lunghezza totale di 35 cm, ma può raggiungere una lunghezza totale massima di 55 cm. Gli occhi sono ravvicinati, posizionati dorsalmente e possono sporgere sopra il profilo



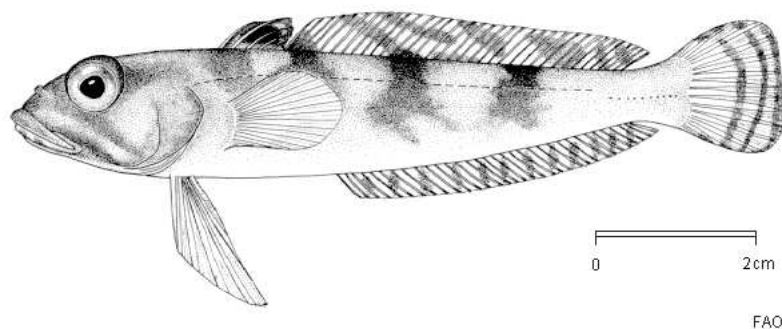
dorsale della testa. Il corpo è ricoperto da scaglie ctenoidi, ad eccezione della parte inferiore della bocca. Le pinne pettorali sono più piccole delle pinne pelviche. La pinna caudale è arrotondata o tronca e presenta 15-17 raggi ramificati. La prima pinna dorsale ha 4-5 spine, la seconda pinna dorsale ha 36-37 raggi molli mentre la pinna anale contiene 29-33 raggi molli. Il colore del corpo è grigiastro e presenta 9 bande scure, ampie e irregolari sulla parte superiore, che si uniscono nella parte inferiore. La parte posteriore della prima pinna dorsale è nerastra (Gon e Heemstra, 1990). La maturità sessuale viene raggiunta alla lunghezza di 34-36 cm e all'età di 8 o 9 anni. La deposizione di uova demersali, del diametro di circa 1.4 mm, avviene a febbraio-marzo (Kock e Jones, 2005). Si stima che lo sviluppo embrionale duri dai 2 ai 3 mesi (Damerou *et al.*, 2012). Informazioni sulla fecondità assoluta non sono ancora disponibili per questa specie (Kock e Jones, 2005). Questa specie si nutre principalmente di krill e macrozooplancton (principalmente crostacei, cnidari e salpe), nonché di pesci che si trovano nella colonna d'acqua (Gon e Heemstra, 1990).

### **1.3.2 Genere *Nototheniops***

Attualmente il genere *Nototheniops* comprende tre specie: *Nototheniops larseni* (Lönnberg 1905), *Nototheniops mizops* (Günther 1880) e *Nototheniops nudifrons* (Lönnberg 1905), tutte con distribuzione biogeografica sub-antartica (Eastman e Eakin, 2021). Le specie *N. nudifrons* e *N. larseni* erano in precedenza attribuite al genere *Lepidonotothen*.

*Nototheniops nudifrons* e *Nototheniops larseni* sono due dei Notothenioidei più numerosi e ampiamente distribuiti nella zona dell'Arco di Scozia (Llompert *et al.*, 2015). Rivestono un ruolo importante nella catena trofica antartica all'interno dell'ecosistema marino costiero, a causa della loro elevata abbondanza locale (Barrera-Oro, 2002). Larve di *N. larseni* sono state trovate disperse su ampie aree, mentre larve di

*N. nudifrons* sono state trovate principalmente vicino alla costa (Deli Antoni *et al.*, 2019).



**Figura 1.5:** Illustrazione di un esemplare di *Nototheniops nudifrons* (Fischer e Hureau, 1985).

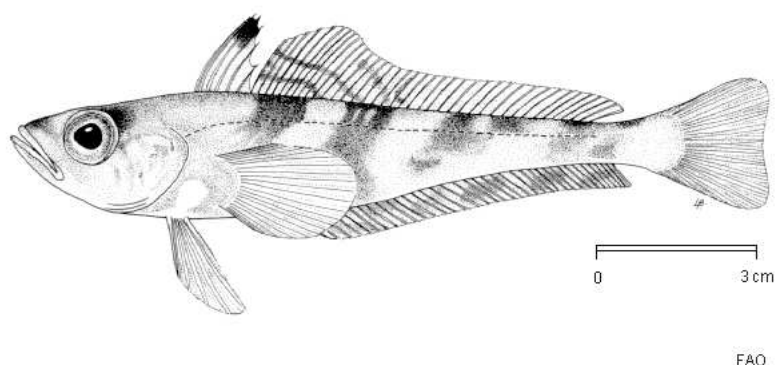
*Nototheniops nudifrons* (Lönnberg 1905) (figura 1.5) è una specie bentonica che occupa habitat a profondità comprese tra i 5 e i 350 m, in acque con temperature intorno a 0 °C. La sua distribuzione è limitata all'Arco di Scozia: dalla Penisola Antartica all'isola di South Georgia, comprese le isole South Shetland, South Orkneys e South Sandwich (Gon e Heemstra, 1990). Il corpo è oblungo, compresso verso la coda e può raggiungere una lunghezza totale di 19.5 cm. Ha occhi ravvicinati posizionati dorsalmente che possono sporgere sopra il profilo dorsale della testa. Il corpo è in gran parte ricoperto da scaglie ctenoidi, ad eccezione della parte superiore e ventrale della testa. La prima pinna dorsale contiene 4-5 spine ed è unita alla seconda pinna dorsale da una membrana compresa tra la sua ultima spina e i due terzi inferiori del primo raggio della seconda pinna dorsale. La seconda pinna dorsale ha 34-37 raggi molli mentre la pinna pelvica ha 33-35 raggi. Le pinne pettorali sono più lunghe delle pinne pelviche in individui di piccole dimensioni, ma diventano simili in individui di lunghezza superiore ai 12 cm. I maschi di questa specie sono uno dei pesci più colorati dell'Oceano Meridionale. Nei maschi la parte superiore e laterale del corpo presenta un caldo colore arancio-marrone, con bande marroni ai lati. La zona ventrale è pallida, con un'area argentea alla base delle pinne pelviche e

pettorali. La prima pinna dorsale è arancione con una grande macchia nera, mentre la seconda pinna dorsale è arancione con corte bande gialle e poche macchie brunastre distribuite su alcuni raggi. La pinna anale e la pinna caudale presentano un'alternanza di bande gialle e arancio-marroni. Le pinne pettorali sono arancioni, con colorazione più scura nella porzione distale. Le pinne pelviche sono di colore giallo pallido con deboli bande marroni. Le femmine adulte e gli esemplari immaturi sono quasi uniformemente di colore giallo pallido o giallo-brunastro con bande marroni sulle pinne (Gon e Heemstra, 1990). È evidente un dicromismo sessuale tra gli individui adulti (Kock e Jones, 2005). La deposizione di uova demersali, del diametro di circa 2.5 mm, avviene da aprile a ottobre, con la maturità sessuale raggiunta tra 9-14 cm. Le uova vengono deposte in nidi in prossimità o sotto la copertura di rocce e custodite dal maschio per circa 4 mesi. La presenza di cure parentali può ridurre la predazione, la mortalità e il flusso genetico tra popolazioni mediante il trattenimento delle uova vicino alla costa (Papetti *et al.*, 2012). La schiusa inizia a settembre e le larve misurano circa 7 mm. La fecondità assoluta decresce gradualmente con la latitudine, passando da 1646-6886 uova intorno all'isola di South Georgia a 110-3460 uova intorno alle isole South Shetland. Da adulti si nutrono principalmente di policheti, anfipodi e isopodi; tuttavia, individui più grandi mangiano anche uova di pesce e gamberetti. I giovani e gli adulti più piccoli si nutrono principalmente di copepodi (Gon e Heemstra, 1990).

Usando una combinazione di dati morfologici e genetici, Dornburg *et al.* (2016) determinarono che *N. nudifrons* (indicata come *Lepidonotothen nudifrons* nello studio di Dornburg *et al.* 2016) è suddivisa in due specie criptiche: *L. nudifrons* intorno alle isole di South Georgia e di South Sandwich e *L. cf. nudifrons* nell'area della Penisola Antartica. Secondo questo studio la presenza di specie criptiche è un fenomeno ampiamente diffuso tra gli invertebrati antartici. Se ciò fosse riscontrato anche tra i pesci, allora la diversità di specie nei Notothenioidei potrebbe al momento essere sottostimata.

Differenze morfologiche tra individui della Penisola Antartica e individui delle isole di South Sandwich sono state trovate anche in uno studio di

Deli Antoni (2019). Egli attribuisce queste differenze a differenti tipi di alimentazione. Pesci che si trovano ad avere a che fare con risorse eterogenee in termini di cibo, spazio e habitat tendono ad evolvere morfologie diverse che riflettono adattamenti a differenze ecologiche locali. Secondo questo studio, adattamenti locali, in assenza di un flusso genetico significativo, possono portare al trend osservato di differenziamento genetico.



**Figura 1.6:** Illustrazione di un esemplare di *Nototheniops larseni* (Fischer e Hureau, 1985).

*Nototheniops larseni* (Lönnberg 1905) (figura 1.6) è una specie bentonica che occupa habitat a profondità comprese tra i 30 e i 750 m, in acque con temperature intorno a 0 °C. Questa specie abita le piattaforme continentali della Penisola Antartica così come la maggior parte delle isole sub-antartiche. Presenta una distribuzione geografica molto ampia e nel settore Atlantico si trova lungo la catena di isole dalle South Shetland a ovest fino a Bouvet a est, incluse le isole South Orkney, Shag Rocks, South Georgia e South Sandwich (Gon e Heemstra, 1990). Presenta un corpo compresso e fusiforme, ma nonostante possa sembrare un pesce pelagico o semipelagico, è una tra le specie più pesanti. Può raggiungere una lunghezza totale massima di 24 cm. Ha occhi grandi con un diametro nettamente maggiore della lunghezza del muso. Il corpo e la testa sono in gran parte ricoperti da scaglie ctenoidi. Le pinne pettorali sono approssimativamente della stessa lunghezza o leggermente più lunghe delle pinne pelviche. La prima pinna dorsale presenta 6-7 spine, la prima e la seconda spina sono all'incirca uguali in

altezza ma sono più alte della terza spina. L'ultima spina della prima pinna dorsale è collegata al primo raggio della seconda pinna dorsale da una membrana. La seconda pinna dorsale contiene 37-39 raggi molli mentre la pinna anale ha 37-38 raggi molli. La pinna caudale è ramificata e presenta 14-15 raggi (Gon e Heemstra, 1990). Nella parte dorsale e laterale del corpo sono presenti 3-4 bande scure, oblique e irregolari che si estendono fino alla zona prossimale della pinna dorsale come macchie bruno-nerastre. La punta della prima pinna dorsale è nera mentre la seconda pinna dorsale presenta bande scure oblique. L'area tra le pinne pettorali e pelviche è pallida. Si nota un certo dicromismo sessuale in individui in età riproduttiva. La maturità sessuale è raggiunta alla lunghezza totale di 11-13 cm e all'età di 4 o 5 anni (Gon e Heemstra, 1990). La deposizione delle uova, del diametro di circa 2 mm, avviene in giugno-luglio nell'isola di South Georgia e circa un mese dopo nelle isole South Shetland. Il luogo preciso dove vengono deposte è ancora sconosciuto (vicino alla riva o più in profondità). Le uova si schiudono a settembre al largo dell'isola di South Georgia e in ottobre al largo della Penisola Antartica e misurano circa 8 mm. *Nototheniops larseni* presenta una fase pelagica prolungata, tra le più lunghe dei Notothenioidei, che continua nei mesi invernali e dura circa 1 anno, incrementando il potenziale di dispersione e il flusso genico (North, 2001). La fecondità assoluta varia da 2000-12000 uova nell'isola di South Georgia a 1851-7127 uova nelle isole South Orkney e Elephant. *Nototheniops larseni* è molto abbondante ma a causa della sua dimensione ridotta la biomassa totale di questa specie è abbastanza bassa se comparata ad altre specie (Jones *et al.*, 2008). Da adulti si nutrono principalmente di krill, anfipodi, iperidi e mysidi, mentre gli individui giovani si nutrono anche di copepodi (Gon e Heemstra, 1990). È un predatore opportunistico, capace di migrare dal fondo fino alla superficie dell'acqua per nutrirsi, sviluppando un'ampia serie di strategie che gli permettono di utilizzare risorse nutritive in una varietà di habitat differenti (Curcio *et al.*, 2013).

Secondo Miya (2016) gli individui di *N. larseni* (indicati come *Lepidonotothen larseni*) rappresentano una singola unità genetica, con sequenze di DNA mitocondriale (geni ND2 e COI) e nucleare (gene S7)

che presentano tassi di divergenza bassi (da 0.3% a 0.6%) tra la sezione Atlantica e Indiana dell'Oceano Meridionale.

#### **1.4 Radiazione adattativa**

La radiazione adattativa è l'evoluzione rapida di una moltitudine di specie distinte morfologicamente ed ecologicamente a partire da un antenato comune, come conseguenza dell'adattamento a nuove nicchie ecologiche. La radiazione adattativa è un processo estremamente complesso e influenzato da una varietà di fattori ecologici, genetici ed evolutivi. È considerata di primaria importanza e potenzialmente responsabile della maggior parte della diversità della vita presente sulla Terra (Schluter, 2000).

Molto spesso la diversificazione all'interno di un clade avviene durante un periodo di tempo lungo, innescata da processi di adattamento. È quindi di fondamentale importanza individuare dei criteri oggettivi che permettano di discriminare tra un processo di radiazione adattativa e un normale processo evolutivo. La radiazione adattativa è caratterizzata da un tasso eccezionale di diversificazione all'interno di una varietà di nicchie ecologiche in tempi estremamente rapidi.

Una radiazione adattativa può essere individuata attraverso quattro criteri principali. Il primo criterio consiste nella discendenza comune di tutte le specie indagate, il secondo criterio nella rapida diversificazione. Il terzo e il quarto criterio riguardano la correlazione fenotipo-ambiente e la *trait utility*, ovvero la presenza di un'associazione significativa tra l'ambiente e i tratti morfologici e fisiologici impiegati per sfruttarlo. Le differenze fenotipiche, associate a differenti usi delle risorse o ad altre caratteristiche dell'ambiente, apportano vantaggi in termini di prestazioni o *fitness* (Schluter, 2000).

Losos e Miles (2002) propongono che il termine radiazione adattativa venga ristretto a quei cladi che mostrano un'elevata ed inusuale divergenza in una varietà di forme adattative. La formulazione di una metodologia adeguata potrebbe permettere di individuare casi di

radiazione adattativa. Un primo fattore da prendere in considerazione è il differenziamento dal punto di vista adattativo che deve portare alla formazione di specie che hanno sviluppato adattamenti utili ad interagire con l'ambiente in modi diversi o ad utilizzare differenti habitat o risorse. Questa forma di adattamento non deve necessariamente accompagnarsi alla formazione di un numero elevato di specie. Un altro fattore importante consiste nel determinare se la quantità di diversificazione è inusualmente elevata. Questo è possibile mediante analisi comparative tra un numero sufficientemente grande di gruppi che hanno attributi fenotipici paragonabili e che interagiscono con l'ambiente in modi simili.

Quello che viene considerato come il più eccezionale esempio di radiazione adattativa è rappresentato dai pesci ciclidi africani che sono andati incontro a radiazione adattativa in più di 30 laghi diversi (Seehausen, 2006). Nel lago Tanganica sono presenti circa 240 specie (figura 8.1 in appendice) il cui antenato comune risale a circa 9.7 milioni di anni fa, momento di formazione del lago. I pesci ciclidi del lago Tanganica si sono quindi evoluti e diversificati *in situ* e presentano un elevato grado di diversità morfologica, ecologica e comportamentale. In accordo con la teoria della radiazione adattativa, secondo cui la diversificazione avviene grazie all'insediamento di nicchie ecologiche differenti, le specie attuali presentano associazioni forti tra l'ambiente occupato e le caratteristiche morfologiche che hanno sviluppato per sfruttarlo (Ronco *et al.*, 2021). Un altro classico esempio di radiazione adattativa è rappresentato dai fringuelli delle Isole Galápagos. Di questi ultimi ne esistono circa 15 specie, tutte derivate da antenati comuni che colonizzarono le Isole Galápagos dopo aver attraversato in volo l'Oceano per circa 1000 km. In seguito alla colonizzazione di un nuovo habitat iniziò la rapida diversificazione, non prima di 3 milioni di anni fa, di nuove specie. Attualmente ciascuna specie occupa una nicchia ecologica esclusiva alla quale si è adattata. L'adattamento è avvenuto principalmente attraverso modificazioni nelle dimensioni e nella forma del becco, in relazione al cibo consumato (Grant e Grant, 2002).

Alcune condizioni ambientali possono determinare un incremento nel differenziamento e i primi modelli di radiazione adattativa erano meramente teorici e non descrivevano i meccanismi attraverso i quali l'ambiente poteva influenzare i tassi di diversificazione delle specie e dei fenotipi. In generale si pensava che la radiazione adattativa derivasse da opportunità ecologiche sotto forma di nicchie ecologiche vacanti che potevano diventare disponibili grazie alla colonizzazione di nuovi ambienti, all'estinzione di antagonisti o all'emergere di innovazioni chiave evolutive che permettevano di interagire con l'ambiente in un modo nuovo. Si conoscono bene gli effetti a breve termine di queste opportunità ecologiche, ma gli effetti a lungo termine sulla speciazione e sulla radiazione adattativa sono ancora poco conosciuti. Gli effetti a breve termine portano ad un incremento delle popolazioni, un aumento dell'utilizzo delle risorse e degli habitat e un incremento nella variazione dei tratti (Yoder *et al.*, 2010).

La radiazione adattativa consiste di due componenti: la produzione di nuove specie (speciazione) e l'adattamento delle specie a una varietà di nicchie ecologiche disponibili.

Durante una radiazione adattativa, ci si aspetta che speciazione e differenziamento morfologico rallentino dopo una fase iniziale di adattamento rapido (*early bursts*) a nicchie ecologiche vacanti. Questo rallentamento si verifica poiché, con il passare del tempo, lo spazio disponibile nelle nicchie inizia ad esaurirsi (Colombo *et al.*, 2015). Si nota inoltre che non solo i tassi di speciazione, ma anche il numero totale di specie, può diminuire in seguito ad un evento di *early bursts*, un fenomeno chiamato *overshooting*.

Gavrilets e Losos (2009) proposero un meccanismo secondo cui la radiazione adattativa procede seguendo *step* successivi, con la diversificazione che avviene lungo differenti assi e a differenti intervalli di tempo. La divergenza avviene inizialmente in relazione ai macrohabitat, seguita da un'evoluzione legata ai microhabitat. Le prime due fasi sono seguite da una divergenza legata ai tratti che controllano sia l'adattamento locale che l'accoppiamento non casuale e infine una divergenza rispetto ai tratti che controllano la sopravvivenza e la



riproduzione. Eventi di radiazione adattativa sono facilitati inoltre dalla presenza di aree geografiche estese e monodimensionali e dall'assenza di speciazione allopatrica. L'ordine con cui avvengono questi *step* sembra dipendere da diversi fattori, variando da una radiazione ad un'altra (Colombo *et al.*, 2005).

#### **1.4.1 La radiazione adattativa nei Notothenioidei**

La notevole importanza e l'elevato interesse dimostrato negli ultimi anni dalla comunità scientifica per la fauna antartica non dipendono dal numero assoluto di specie, ma dalla natura della diversità che si ritrova in quest'area. Le acque antartiche della piattaforma continentale vengono considerate come un *hot spot* evolutivo, inteso come una concentrazione di diversità che non si trova da nessun'altra parte. (Eastman, 2005).

I Notothenioidei rappresentano uno dei migliori esempi di radiazione adattativa in ambiente marino (Matschiner *et al.*, 2015). È sorprendente pensare che si conoscano poche radiazioni adattative in ambiente marino, considerando il grande numero di esempi di radiazione adattativa in acqua dolce. Si potrebbe trovare una spiegazione nel fatto che le radiazioni adattative sono più evidenti in aree ben definite dal punto di vista geografico, come isole e laghi mentre sono meno identificabili in sistemi aperti come gli oceani. La piattaforma continentale antartica risulta però isolata dalle altre piattaforme continentali dalla distanza, dalle correnti, dalle acque profonde e dalle temperature sotto lo zero (Eastman e McCune, 2000).

I Notothenioidei si sono evoluti *in situ* nelle acque sottozero della piattaforma continentale (Eastman, 2000) e formano un tipo speciale di radiazione adattativa conosciuta come *species flock* (Eastman e McCune, 2000). Dal punto di vista teorico uno *species flock* è un numero sproporzionatamente alto di specie strettamente imparentate che si sono

evolute rapidamente all'interno di un'area circoscritta, dove la maggior parte delle specie sono endemiche (Ribbink, 1984). Eastman e McCune (2000) identificarono cinque criteri per riconoscere uno species flock: monofilia, elevata diversità di specie, alto livello di endemismo, dominanza dell'habitat (in termini di biomassa) e diversità morfologica ed ecologica. Come già visto, i Notothenioidei costituiscono il 77% della ricchezza di specie e il 91% della biomassa nella piattaforma continentale antartica e il tasso di endemismo è del 97%. Inoltre, un aspetto distintivo della biodiversità dei Notothenioidei è che mostrano diversità morfologiche ed ecologiche considerevoli (Eastman, 2005).

Si ritiene che la radiazione adattativa dei Notothenioidei sia seguita a opportunità ecologiche causate da cambiamenti climatici, oceanografici e tettonici intercorsi durante gli ultimi 50 milioni di anni. Questi cambiamenti isolarono la piattaforma continentale antartica, abbassarono la temperatura dell'acqua e introdussero ripetute alterazioni degli habitat in seguito alle continue espansioni e ritirate della calotta glaciale antartica.

Durante i periodi di maggiore glaciazione, la calotta glaciale si estendeva fino al bordo esterno della piattaforma continentale, causando la distruzione delle comunità marine di fondo su larga scala. Evidenze geologiche marine però hanno mostrato che i ghiacciai non sempre avanzarono fino all'estremità della piattaforma continentale, lasciando libere aree che possono aver agito da rifugi per gli organismi bentonici. Sembra però più probabile che gli organismi marini si siano rifugiati nella parte alta della scarpata continentale o nelle acque profonde (Barnes e Peck, 2006). Un'altra alternativa probabile è che la copertura di ghiaccio sia avvenuta con tempistiche diverse, lasciando rifugi isolati sulla piattaforma continentale che cambiavano posizione durante i periodi di massima glaciazione, permettendo alla fauna bentonica di migrare da un rifugio al successivo (Thatje *et al.*, 2005). Come risultato, le popolazioni furono periodicamente isolate in rifugi rimasti liberi dal ghiaccio. La frammentazione delle popolazioni e l'isolamento in rifugi durante i massimi glaciali probabilmente furono meccanismi chiave per la

speciazione allopatrica (Rogers, 2007), seguita poi dalla ricolonizzazione degli habitat distrutti in seguito al ritiro dei ghiacci. L'uso di dati molecolari a livello di popolazione ha permesso a molti studiosi di investigare i cambiamenti passati nella dimensione delle popolazioni. Zane *et al.* (2006) suggeriscono che vi sia un collegamento tra eventi demografici in *Pleuragramma antarctica* (famiglia Nototheniidae) ed eventi paleoclimatici passati.

Circa 34 milioni di anni fa, al confine Eocene-Oligocene, il continente antartico divenne isolato termicamente ed oceanograficamente in seguito all'instaurarsi della Corrente Circumpolare Antartica. La regione antartica subì una drastica diminuzione delle temperature, culminando in un'estesa glaciazione dell'intero continente. Questi eventi forzarono alcune specie antartiche a spostare la loro distribuzione, come visto in precedenza, ma portarono molte altre specie all'estinzione. Una delle conseguenze di questo raffreddamento, per l'ecosistema marino dell'Oceano Meridionale, fu un drastico cambiamento della biodiversità. Si assistette alla perdita dei principali gruppi tassonomici, come i predatori apicali bentonici e pelagici (squali e granchi), e una riduzione della biodiversità in gruppi come i molluschi bivalvi, i pesci teleostei e i decapodi (Thatje *et al.*, 2005). La fauna cambiò drasticamente, ma probabilmente i Notothenioidei persisterono dopo l'estinzione della maggior parte della fauna grazie all'acquisizione di numerosi adattamenti, morfologici e fisiologici, alle mutate condizioni dell'Oceano Meridionale. Questi adattamenti permisero loro di diversificarsi e di occupare con successo una grande varietà di nicchie ecologiche lasciate libere, diventando il gruppo di pesci dominante nella piattaforma antartica.

In presenza di competizione ridotta con altri taxa essi andarono incontro ad una diversificazione per riempire una varietà di nicchie disponibili a varie profondità nella colonna d'acqua. La loro diversificazione fu centrata sull'alterazione della galleggiabilità (Eastman, 2005). Partendo da un antenato bentonico privo di vescica natatoria, diversificazioni morfologiche sostanziali portarono a fenotipi adatti ad alimentazioni pelagiche o parzialmente pelagiche, zooplanctivore o piscivore

(Eastman e Eakin, 2000). Questo fenomeno, detto *pelagization*, è considerato come il segno distintivo della radiazione dei Notothenioidei (Eastman, 2000).

I Notothenioidei mostrano inoltre adattamenti, a livello molecolare, alle basse temperature (Detrich *et al.*, 2000). Un esempio sono i pesci antartici della famiglia Channichthyidae il cui sangue presenta una ridotta viscosità a causa dell'assenza di globuli rossi ed emoglobina, condizione possibile in acque a basse temperature nelle quali l'ossigeno è molto più solubile e tende ad essere assorbito dal sangue con più facilità.

Matschiner *et al.* (2011) e Near *et al.* (2012) stimano l'inizio della radiazione adattativa dei Notothenioidei antartici vicino al limite tra Oligocene e Miocene (23.9 Ma e 22.4 Ma rispettivamente), coincidente con un periodo di raffreddamento globale e di espansione della calotta glaciale antartica.

Colombo *et al.* (2015) invece stimano il momento della radiazione adattativa a circa 13.4 Ma, data che è molto più recente rispetto ad altri studi precedenti. Se questa stima fosse corretta, l'inizio della radiazione coincise con la transizione climatica del Medio Miocene MMCT (14.1-13.9 Ma), durante la quale la temperatura superficiale dell'Oceano Meridionale diminuì di 6-7 °C e un clima polare si stabilì nel continente antartico.

Un tema molto dibattuto consiste nell'individuazione del fattore, probabilmente un'innovazione chiave, che innescò il processo di radiazione adattativa, permettendo la colonizzazione delle acque ghiacciate. Molti studi individuano questo fattore nella comparsa di glicoproteine antigelo (AFGPs), che si stima sia avvenuta tra 42 e 22 milioni di anni fa (Near *et al.*, 2012).

Studi filogenetici recenti (Near *et al.*, 2012) scoprirono che un impulso maggiore alla diversificazione avvenne molto dopo l'evoluzione delle AFGPs, il che implica che altri fattori furono più importanti nel guidare la diversificazione e agirono in tempi successivi.

La radiazione dei Notothenioidei antartici è una radiazione adattativa in quanto soddisfa tutti i criteri evidenziati da Schluter (2000): antenato comune, rapida speciazione, correlazione fenotipo ambiente e utilità dei tratti.

Numerosi studi sulla filogenesi e i tassi di diversificazione dei Notothenioidei (Eastman, 2005; Matschiner *et al.*, 2011; Near *et al.*, 2012) mostrano tuttavia tassi di diversificazione che sembrano essere minori rispetto a radiazioni adattative come quella dei pesci ciclidi dell'Africa orientale (Rutschmann *et al.*, 2011). Questi tassi inferiori sembrano essere dovuti esclusivamente a fattori come la mancanza di eterogeneità di habitat, l'assenza di alcuni habitat primari costieri, l'abilità migratoria nella lunga distanza delle fasi larvali pelagiche (Damerau *et al.*, 2014), l'assenza di strutturazione genetica di popolazione su grandi distanze e l'ipotesi che la radiazione adattativa sia ancora in corso.

La correlazione fenotipo-ambiente e l'utilità dei tratti sono perfettamente soddisfatte dal fenomeno della *pelagization* (il passaggio da uno stile di vita bentonico a uno pelagico) che è sorto più volte e indipendentemente in numerosi cladi (Bargelloni *et al.*, 2000). L'esempio più completo è rappresentato dalla famiglia dei Nototheniidae. Circa la metà delle specie occupano l'habitat bentonico ancestrale, mentre l'altra metà presentano uno stile di vita semipelagico, epibentonico, criopelagico o pelagico (Eastman, 2005). Furono necessari molti aggiustamenti morfologici e fisiologici per permettere alle specie di colonizzare l'intera colonna d'acqua, principalmente per raggiungere effettive capacità di nuoto e compensare la mancanza della vescica natatoria che la maggior parte degli altri pesci teleostei usano per regolare il loro galleggiamento. Alcune specie raggiunsero la galleggiabilità neutra attraverso una riduzione della mineralizzazione dello scheletro (Balushkin, 2000; Eastman, 2005) e un accumulo di depositi di lipidi in cellule adipose per diminuire la densità generale e produrre una spinta statica verso l'alto (Eastman, 1993). Adattamenti per la vita nella colonna d'acqua inclusero anche cambiamenti morfologici per alterare le capacità di nutrirsi e di

nuotare, come diversificazioni nella forma del corpo e della testa che permisero ai Notothenioidei di specializzarsi in diete differenti (Colombo *et al.*, 2015).

Un altro tratto, oltre alla galleggiabilità, che serve a mostrare sia la correlazione fenotipo-ambiente sia l'utilità dei tratti, è la capacità dei Notothenioidei di sintetizzare glicoproteine antigelo (AFGPs). In un ambiente ghiacciato, la maggiore sfida fisiologica per organismi ectotermi è quella di prevenire il congelamento del sangue e dei tessuti. Dal momento che i pesci teleostei marini hanno un punto di congelamento più alto rispetto ai pesci di acqua dolce, il contatto con l'acqua ghiacciata porterebbe al rapido congelamento dei fluidi corporei, che è letale per quasi tutti i vertebrati. L'insorgere della capacità di produrre glicoproteine antigelo ha permesso ai pesci antartici di abbassare efficientemente il loro punto di congelamento, prevenendo il congelamento in seguito al contatto con ghiaccio marino. Le AFGPs sono sintetizzate principalmente nel pancreas come lunghi precursori poliproteici che vengono poi spezzati per produrre gli otto differenti tipi di AFGPs (Evans *et al.*, 2012). Dal pancreas esocrino, le AFGPs vengono dislocate nel tratto gastrointestinale dove si legano a cristalli di ghiaccio, ingeriti con il cibo o l'acqua, e impediscono la loro crescita fino a quando non vengono escreti con le feci. Le AFGPs libere vengono poi riassorbite attraverso l'epitelio rettale (Evans *et al.*, 2012) ed entrano nel sangue e nei fluidi interstiziali. Queste AFGPs raggiungono la bile attraverso il fegato e sono riportate nuovamente nel tratto gastrointestinale. Il muco che ricopre la superficie del corpo degli adulti contiene AFGPs, impedendo la crescita di ghiaccio sulla superficie del corpo e impedendo ai cristalli di ghiaccio di entrare nel sistema circolatorio attraverso l'epitelio (Matschiner *et al.*, 2015). Uno studio di Bilyk e DeVries (2010) su 11 specie di Notothenioidei, appartenenti alla famiglia Channichthyidae, mostra una maggiore espressione di AFGPs in specie che si trovano alle più alte latitudini e quindi a temperature più basse.

In merito alla radiazione adattativa dei Notothenioidei rimane ancora poco chiaro come le specie si siano potute evolvere rapidamente dal momento che il differenziamento è in contrasto con un elevato flusso genico favorito dalla presenza di una fase larvale pelagica. Sembra probabile che gli eventi di speciazione dei Notothenioidei siano stati ristretti a periodi in cui la dispersione larvale era ostacolata come, per esempio, durante l'estesa copertura di ghiaccio nel corso del massimo glaciale (Damerou *et al.*, 2012). Il flusso genico, attraverso la dispersione larvale, può d'altra parte aver facilitato la sopravvivenza ai periodi glaciali, permettendo la ricolonizzazione degli habitat divenuti nuovamente disponibili e contrastando la divergenza genetica tra popolazioni.

### **1.5 Struttura di popolazione nei Notothenioidei**

Nel corso degli anni sono stati condotti numerosi studi sulla struttura di popolazione in specie diverse di Notothenioidei, sia pelagiche che bentoniche, basati sia su caratteristiche morfologiche che genetiche.

La dispersione nelle specie marine modella la biodiversità ed è influenzata da processi oceanografici. Alle alte latitudini, in particolare nell'emisfero meridionale, gli organismi marini presentano fasi larvali planctoniche lunghe e le forti correnti marine determinano una dispersione fino a centinaia di chilometri di distanza (Álvarez-Noriega *et al.*, 2020). Le caratteristiche oceanografiche sono importanti fattori che regolano la connettività tra popolazioni sia aumentando che attenuando la dispersione larvale. Come è già stato visto, le correnti dell'Oceano Meridionale ricadono in due categorie, una circolazione superficiale e una profonda. Le correnti superficiali sono le più importanti per la maggior parte della catena trofica, dal fitoplancton ai pesci, essendo implicate nella dispersione degli organismi e nella creazione di specifici habitat marini.

L'ipotesi generale è quella di una mancanza di struttura genetica di popolazione nei Notothenioidei antartici, dovuta alla durata della fase

larvale pelagica e alla forza delle correnti. Il trasporto di uova o larve tramite le correnti oceaniche è stato considerato un fattore chiave in grado di modulare la connettività tra le popolazioni. Ci si potrebbe aspettare che specie con una fase larvale pelagica più lunga abbiano una dispersione maggiore, e quindi, una struttura di popolazione più debole, rispetto alle specie con una fase larvale più breve.

I primi studi di struttura di popolazione condotti negli anni '80 riguardarono il settore Atlantico. Non erano studi genetici e si basavano su caratteristiche morfologiche e meristiche, tassi di infestazione di parassiti e chimica degli otoliti. Il trend individuato era un differenziamento tra le popolazioni a nord (isole di Shag Rock e South Georgia) e a sud (isole di South Shetland, di Elephant e di South Orkney) del Mare di Scozia, indicando bassa connettività tra popolazioni separate da distanze superiori a 900 km, anche se connesse dalla Corrente Circumpolare Antartica. Queste scoperte furono messe in dubbio con l'avvento degli studi genetici i cui risultati furono molto variabili così come i marker genetici utilizzati, il disegno sperimentale (dimensione dei campioni) e la scelta delle specie target (Damerau *et al.*, 2014).

Numerosi studi genetici sulla struttura di popolazione in specie di Notothenioidei antartici hanno evidenziato una certa omogeneità genetica tra popolazioni separate geograficamente (Matschiner *et al.*, 2009, Papetti *et al.*, 2012, Damerau *et al.*, 2012).

Uno studio di genetica di popolazione di Matschiner *et al.* (2009) fu condotto su popolazioni di specie strettamente bentoniche, come *Gobinotothen gibberifrons*, confinate quindi alle aree della piattaforma continentale da adulti. I risultati indicarono che le popolazioni distribuite nelle isole del Mare di Scozia e nella porzione nordoccidentale della Penisola Antartica non erano significativamente differenziate, anche se separate geograficamente da circa 2000 km e da bacini profondi. Per capire se questo potesse essere dovuto a dispersione larvale passiva venne utilizzato un modello per approssimare il movimento di particelle passive superficiali trasportate da correnti oceaniche. I risultati del



modello indicarono che la ACC poteva trasportare in direzione ovest-est le larve, ostacolando il differenziamento genetico tra popolazioni distanti.

Un successivo studio genetico di Papetti *et al.* (2012), condotto sulla specie *Chionodraco rastrospinosus*, mostrò una mancanza di struttura genetica, senza differenze a livello geografico o temporale, e un flusso genetico unidirezionale verso est lungo i siti di campionamento nell'area dell'Arco di Scozia. Questo risultato suggerì una mancanza di barriere genetiche per questa specie, attribuibile a una dispersione larvale che seguiva l'ACC. Questo studio riportò un risultato opposto rispetto ad uno studio precedente del 2009, condotto su una specie (*Chaenocephalus aceratus*) strettamente imparentata, presente nella stessa zona e che mostra tratti ecologici e di stile di vita simili, ma che tuttavia presentava moderati livelli di differenziamento genetico tra la Penisola Antartica e le Isole dell'Arco di Scozia.

Per ultimo si può citare lo studio di Damerau *et al.* (2012) condotto su popolazioni di 7 specie di Notothenioidei (tra cui *L. squamifrons*) nella zona dell'Arco di Scozia meridionale (dalla punta della Penisola Antartica alle South Orkney). I risultati mostrarono un basso o inesistente differenziamento genetico tra le popolazioni di tutte le specie analizzate, indipendentemente dallo stile di vita adulta o dalla durata della fase larvale pelagica. L'unica caratteristica comune delle sette specie è la presenza di una fase larvale pelagica che si ritenne essere la causa del flusso genico osservato. Lo studio dimostrò inoltre che particelle passive potevano essere trasportate tramite le correnti dalla Penisola Antartica alle isole South Shetland, supponendo che le larve potessero essere trasportate nello stesso modo.

Altri studi hanno osservato invece una certa eterogeneità genetica in alcune specie di Notothenioidei antartici (Patanello *et al.*, 2003; Zane *et al.*, 2006; Kuhn e Gaffney, 2006; Papetti *et al.*, 2009). Patanello *et al.* (2003) mostrò un differenziamento genetico significativo, in *Chionodraco hamatus*, tra le popolazioni del Mare di Weddell e del Mare di Ross. Zane

*et al.* (2006) trovò, a scala geografica ampia, un differenziamento geografico e temporale significativo tra popolazioni di *Pleuragramma antarctica* campionati ai lati opposti del continente antartico. Kuhn e Gaffney (2006) trovarono una struttura genetica significativa in *Champscephalus gunnari* tra le popolazioni del settore Atlantico e del settore Indiano dell'Oceano Meridionale. Infine, uno studio genetico di Papetti *et al.* (2009), condotto sulla specie *Chaenocephalus aceratus* (simile a *G. gibberifrons* in termini di distribuzione e periodo di deposizione delle uova), trovò un differenziamento genetico non solo nel tempo, ma anche nello spazio tra le popolazioni delle isole South Shetland, Elephant Island e South Orkney. Nonostante un alto potenziale di dispersione, la specie mostrava una struttura genetica significativa. Questo studio dimostrò che elevate dimensioni effettive delle popolazioni e assenza di barriere al flusso genico non possono evitare la strutturazione di popolazione in alcuni pesci marini, nonostante quanto si potrebbe supporre da una fase larvale planctonica prolungata.

Le tre specie di Notothenioidei considerati in questo studio (*N. larseni*, *N. nudifrons* e *L. squamifrons*) mostrano strutture genetiche di popolazione differenti.

Lo studio di Jones *et al.* (2008) sulla strutturazione di popolazione in *L. squamifrons* e *N. larseni* nel settore Atlantico dell'Oceano Meridionale (Shag Rocks, South Georgia, South Sandwich e Bouvet Island) mostrò risultati contrastanti per le due specie. Intorno all'isola di Bouvet *N. larseni* non mostrava differenze genetiche significanti rispetto a campioni raccolti in altre aree dell'Arco di Scozia e della Penisola Antartica. *L. squamifrons* invece mostrava differenze genetiche significative tra popolazioni intorno alle isole South Shetland e popolazioni intorno all'isola di Bouvet ( $F_{ST} = 0.189$ ,  $p = 0.015$ ). Non furono evidenziate invece differenze significative, a scala ridotta, tra campioni provenienti dall'Arco di Scozia e dalla Penisola Antartica.

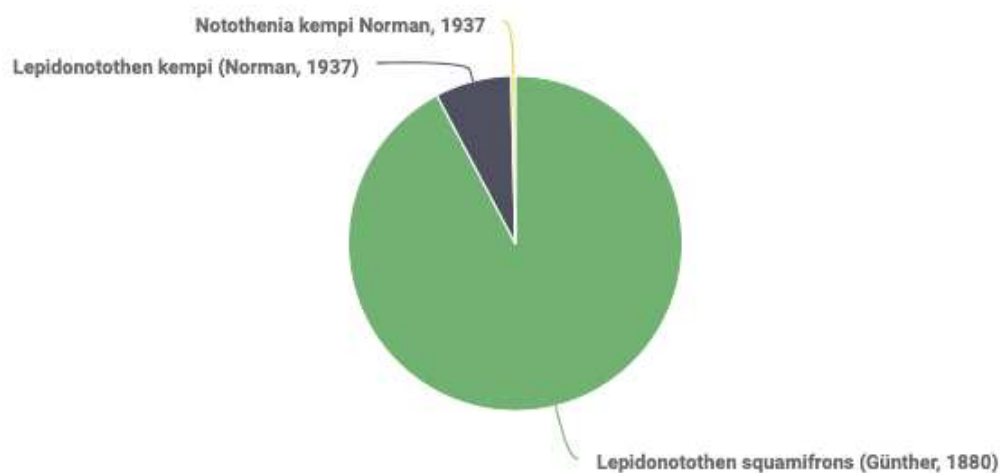
Lo studio di Damerau *et al.* (2014) esaminò invece la struttura genetica di popolazione in *N. larseni* nel settore Atlantico dell'Oceano Meridionale, mostrando differenze genetiche significative tra popolazioni e presumendo l'esistenza di tre popolazioni distinte geneticamente nell'area di studio. La scoperta di un differenziamento significativo tra popolazioni di *N. larseni*, la specie di Notothenioidei con lo sviluppo larvale conosciuto più lungo, mise in dubbio il concetto popolare che una lunga durata della fase larvale pelagica potesse portare ad alti livelli di flusso genico.

Uno studio recente di Deli Antoni *et al.* (2019) confrontò la struttura di popolazione, sia dal punto di vista morfologico che genetico, delle specie *N. larseni* e *N. nudifrons*, due tra le specie più numerose e con una maggiore distribuzione. Il confronto venne fatto tra le popolazioni della Penisola Antartica e quelle delle isole South Shetland, separate dallo Stretto di Bransfield, largo 100 km e con una profondità massima di 1000 m. I risultati relativi a *N. nudifrons* indicarono che le popolazioni erano significativamente differenti sia dal punto di vista genetico ( $F_{ST} = 0.0897$ ,  $p < 0.01$ ) che morfologico. I risultati relativi a *N. larseni* indicarono invece un'assenza di differenziamento significativo tra le popolazioni, sia dal punto di vista genetico ( $F_{ST} = -0.0154$ ,  $p > 0.01$ ) che morfologico.

La presenza di risultati contrastanti per i nototenioidei è evidenziata da studi recenti (Papetti *et al.*, 2012; Damerau *et al.*, 2014; Deli Antoni *et al.*, 2019) che mostrano infatti una relazione più complessa di quanto ci si potrebbe aspettare tra il potenziale di dispersione e il flusso genico in ambiente marino. Oltre alla circolazione oceanica e alla dispersione larvale, una serie di differenti fattori fisici e biologici è in grado di determinare la connettività tra popolazioni. I fattori fisici possono essere quelli relativi alla ritenzione fisica attorno alle isole, in grado di trattenere le larve vicino al luogo natale. I fattori biologici che possono influenzare la dispersione larvale sono molti e includono caratteristiche legate allo stile di vita dell'organismo: il tempo e il luogo dello sviluppo e della deposizione delle uova, la tipologia di uova, il tempo di incubazione, il periodo di dispersione delle larve e il loro comportamento.

## 2. SCOPO DELLA TESI

La tassonomia è un campo della biologia molto complesso e in qualche modo controverso. L'applicazione di diversi approcci nell'identificazione spesso porta ad identificare un numero diverso di specie. Come anticipato, al momento non vi è ancora un'idea condivisa in merito alla separazione tra *L. squamifrons* e *L. kempfi*. I nomi di entrambe le specie sono utilizzati anche in studi recenti (figura 2.1).



**Figura 2.1:** Rappresentazione schematica dell'utilizzo dei termini *Lepidonotothen squamifrons* (92,2%), *Lepidonotothen kempfi* (7,4%) e *Notothenia kempfi* (0,4%) in letteratura. (Fonte: FishBase DOI 10.15468/dggejk).

Lo scopo di questa tesi consiste nel determinare se gli individui assegnati alla specie *L. kempfi*, sulla base di tratti morfologici, appartengano effettivamente ad una specie distinta da *L. squamifrons*, informazione necessaria per una corretta gestione delle specie in ambiente antartico. L'ambiente antartico e la sua fauna stanno ricevendo un'attenzione crescente data dalla consapevolezza del cambiamento climatico globale, della distruzione degli habitat naturali e della continua perdita di biodiversità. Dati sulla diversità di specie, la distribuzione spaziale, l'abbondanza e la struttura genetica delle popolazioni sono informazioni molto importanti per determinare la struttura delle comunità e valutare le risorse disponibili in vista di una gestione sostenibile dello sfruttamento

da parte dell'uomo e della necessità di proteggere biotipi unici applicando principi di conservazione agli organismi antartici marini.

### 3. MATERIALI E METODI

In questa tesi, mi sono occupata dell'analisi dei dati con strumenti bioinformatici. Per completezza riporto una sintesi delle fasi precedenti che hanno permesso di ottenere i dati dei genotipi dei *loci* microsatellite e delle sequenze mitocondriali.

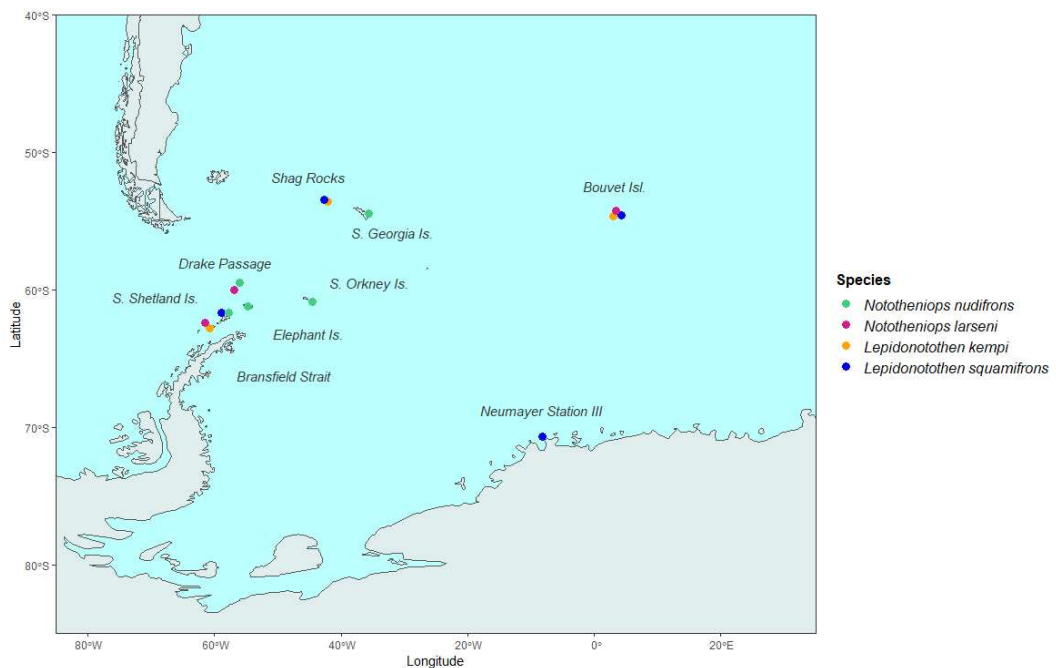
#### 3.1 Campionamento ed estrazione del DNA

I campioni di tessuto analizzati (tabella 3.1) provengono da individui appartenenti alle specie *Lepidonotothen kempfi*, *Lepidonotothen squamifrons*, *Nototheniops larseni* e *Nototheniops nudifrons* che sono stati raccolti durante cinque spedizioni scientifiche antartiche tra il 2005 e il 2012. I campionamenti sono stati effettuati a Bouvet Island, Shag Rocks, Bransfield Strait, Neumayer Station, South Shetland, Drake Passage, Elephant Island, South Georgia e South Orkney (figura 3.1). Tutti i campioni provengono dalla regione dell'Arco di Scozia, dove sono presenti tutte e quattro le specie analizzate.

Gli individui, pescati con rete a strascico, sono stati pesati, misurati e assegnati *in loco* ad una delle quattro specie sulla base delle caratteristiche morfologiche descritte da Gon e Heemstra (1990). Da ogni individuo è stato prelevato un campione di tessuto (pinna o muscolo) che è stato conservato in etanolo 95% per le successive analisi molecolari. Il DNA è stato estratto da tutti i campioni di tessuto seguendo il protocollo *Salting Out* (Patwary *et al.*, 1994) che prevede l'isolamento del DNA dopo l'eliminazione delle proteine e di eventuali contaminanti per precipitazione.

Specie	Siti di campionamento	N. individui (mtDNA)	N. individui (microsatelliti)
<i>Nototheniops nudifrons</i>	Elephant Island Drake Passage South Shetland South Georgia South Orkney	39	40
<i>Nototheniops larseni</i>	Bouvet Island South Shetland Drake Passage	97	94
<i>Lepidonotothen kempfi</i>	Bouvet Island Shag Rocks Bransfield Strait	23	52
<i>Lepidonotothen squamifrons</i>	Bouvet Island Neumayer Station Shag Rocks South Shetland	73	101
<b>Dataset completo</b>		232	287

**Tabella 3.1:** Lista dei campioni analizzati. Sono riportati il nome della specie, il sito di campionamento e il numero di individui rappresentanti il campione nei *dataset* utilizzati per le analisi molecolari mediante mtDNA e microsatelliti.



**Figura 3.1:** Mappa dei siti di campionamento. La figura rappresenta la regione della Penisola Antartica e il Mare di Weddell. I punti colorati rappresentano i siti di campionamento degli individui analizzati in questa tesi (verde per *Nototheniops nudifrons*, rosa per *Nototheniops larseni*, giallo per *Lepidonotothen kempfi* e blu per *Lepidonotothen squamifrons*). La mappa è stata generata con il pacchetto ggOceanMaps del software R.

La buona riuscita di ogni estrazione è stata verificata tramite la stima della concentrazione e della qualità del DNA estratto utilizzando il *NanoDrop Spectrophotometer*. L'integrità del DNA estratto è stata poi verificata mediante corsa elettroforetica su gel d'agarosio all'1% in tampone TBE 1X.

### **3.2 Marcatori microsatellite**

I microsatelliti, detti anche *Simple Sequence Repeats* (SSR) o *Short Tandem Repeats* (STR), sono i marcatori molecolari più utilizzati negli studi di genetica della conservazione, di genetica di popolazione e di genetica forense.

Sono presenti nel genoma di procarioti ed eucarioti e sono distribuiti in modo non casuale in regioni sia codificanti che non codificanti (Li *et al.*, 2002). Sono sequenze ripetute di DNA costituite da unità di ripetizione molto corte (1-6 bp) disposte secondo una ripetizione in tandem. Le sequenze sono ripetute circa 10-40 volte e in media si trovano ogni 10.000 basi nel genoma eucariote. La loro distribuzione non è casuale in quanto svolgono indispensabili funzioni biologiche coinvolte nella regolazione dell'attività genica, nella replicazione e ricombinazione del DNA, nell'organizzazione della cromatina e nel ciclo cellulare (Li *et al.*, 2002).

Sono localizzati su cromosomi autosomici e presentano un alto livello di polimorfismo dato dal differente numero di unità di ripetizione. Gli alleli differiscono per la diversa lunghezza delle sequenze ripetute. Le analisi molecolari si effettuano distinguendo gli alleli di diversa lunghezza tramite elettroforesi su gel. L'elevato polimorfismo è dovuto al fatto che presentano un tasso di mutazione maggiore rispetto ad altre regioni del genoma, il che li rende marcatori molecolari molto efficaci. Le mutazioni avvengono in seguito a inserzione e/o delezione dell'unità di ripetizione, causando variazioni nella lunghezza complessiva della sequenza. Le cause delle mutazioni possono essere attribuite prevalentemente ad uno slittamento della DNA-polimerasi, detto *slippage*, durante la replicazione



o riparazione del DNA, oppure ad un evento di ricombinazione tra molecole di DNA (Li *et al.*, 2002). Gli elevati tassi di mutazione sono compresi tra  $10^{-2}$  e  $10^{-6}$  mutazioni per *locus* per generazione (Selkoe e Toonen, 2006), circa 10 ordini di grandezza in più rispetto alle mutazioni puntiformi (Vieira *et al.*, 2016).

In base alla struttura (tabella 3.2) possono essere distinti in:

- perfetti se composti interamente da ripetizioni della stessa unità di ripetizione;
- imperfetti se una coppia di basi non appartenente al motivo appare tra le ripetizioni;
- interrotti se una sequenza di poche paia di basi è inserita nel motivo;
- composti se formati da motivi di ripetizione diversi in posizione adiacente (Senan *et al.*, 2014).

TIPO	SEQUENZA TIPO
Perfetti	CACACACACACACACACA
Imperfetti	ACACACATACACACACACAC
Interrotti	GATGATGATGATTGAATGAT
Composti	CACACACAGAGAGAGAGA

**Tabella 3.2:** Esempi di tipi di microsatellite conosciuti.

Sulla base del numero di basi per unità di ripetizione i microsatelliti possono essere suddivisi in mono-, di-, tri-, tetra- e esa-nucleotidici. Le ripetizioni con motivo di-nucleotidico rappresentano la maggior parte dei microsatelliti in molte specie e sono più frequenti nelle regioni non codificanti (Li *et al.*, 2002). Sono inoltre usati con maggiore frequenza poiché risultano di più facile isolamento. I microsatelliti di-, tri- e tetra-nucleotidici sono i più utilizzati in studi di genetica molecolare (Selkoe e Toonen, 2006). I microsatelliti tri- ed esa-nucleotidici sono più frequenti nelle regioni codificanti perché non causano mutazioni *frameshift*. Infine, i microsatelliti mono-nucleotidici sono i meno facilmente rilevabili a causa

di problemi legati all'amplificazione, mentre ripetizioni più lunghe sono meno comuni (Li *et al.*, 2002).

A differenza dell'elevato polimorfismo dei microsatelliti, il DNA che li circonda, detto regione fiancheggiante (*flanking region*), è maggiormente conservato sia tra individui della stessa specie che di specie filogeneticamente vicine. Un particolare *locus* microsatellite può essere identificato dalla sua regione fiancheggiante. È stato possibile disegnare e sintetizzare artificialmente piccole catene di DNA, chiamate *primer*, che si appaiano alla regione fiancheggiante e guidano l'amplificazione selettiva del *locus* microsatellite mediante *Polymerase Chain Reaction* (PCR). Nel corso degli anni sono stati isolati numerosi *primer* specifici per specie target. Se una specie viene analizzata per la prima volta è necessario isolare *de novo* specifiche coppie di *primer* (Zane *et al.*, 2002). Risulta evidente, di conseguenza, la necessità di avere informazioni a priori sul DNA da analizzare, ovvero sapere quali *primer* utilizzare. Questo problema viene attenuato dall'alto grado di conservazione di queste regioni che permette la *cross-species amplification*. Gli stessi *primer* possono essere utilizzati in specie diverse ma filogeneticamente vicine a quella per cui erano stati disegnati in origine grazie alla natura omologa delle sequenze di DNA delle regioni fiancheggianti i microsatelliti. La presenza di sequenze fiancheggianti altamente conservate permette la *cross-species amplification* tra specie che si sono separate fino a 470 milioni di anni fa (Zane *et al.*, 2002).

L'utilizzo dei marcatori microsatellite presenta numerosi vantaggi:

- facilità di isolamento a partire da uno scarso materiale di partenza;
- presenza di *flanking region* altamente conservate che permettono l'amplificazione del *locus* anche tra specie diverse ma filogeneticamente vicine;
- polimorfismo di lunghezza con un numero di alleli per *locus* che può superare i 14-15 alleli;
- potenza statistica e accuratezza nella stima delle differenze genetiche tra le specie ed all'interno delle stesse;
- elevata riproducibilità;
- alta risoluzione;

- co-dominanza data dalla rilevabilità di entrambi gli alleli;
- ereditarietà mendeliana;
- isolamento di *primer* specifici per specie target.

Per il loro elevato grado di polimorfismo e per il fatto di essere marcatori co-dominanti che vengono trasmessi alla progenie con le modalità della eredità mendeliana semplice, rappresentano un validissimo strumento molecolare per l'identificazione individuale (identità molecolare), consentono di effettuare indagini per il controllo della parentela (*parentage identification* e *breeding programs*) e sono molto utilizzati in genetica di popolazione per la stima del grado di differenziamento tra popolazioni (*analysis of genetic diversity* e *population genetic structure*). Il microsatellite ideale presenta elevati livelli di eterozigotità (media superiore al 50%, con *loci* che raggiungono picchi del 100%), un range allelico contenuto, una possibile amplificazione in corti prodotti di PCR, alleli ben separabili (accuratezza), un ridotto fenomeno di *stuttering* (scivolamento della DNA polimerasi durante l'amplificazione di ripetizioni in tandem) e un tasso di mutazione contenuto.

In questa tesi sono stati analizzati 21 *loci* microsatellite, 12 dei quali sono stati isolati dal trascrittoma della milza di *Lepidonotothen nudifrons* (Papetti *et al.*, 2016), 2 sono stati isolati da *Chaenocephalus aceratus* (Susana *et al.*, 2007) e i restanti 7 sono stati isolati da *Chionodraco hamatus* (Marino *et al.*, 2011) (tabella 3.3).

<b>Locus</b>	<b>Repeat</b>	<b>Sequenza primer</b>	<b>Fluoroforo</b>
Ln 41281	(GT) <sub>9</sub>	F: TGTGTTTTTGGACGTATGGCA R: GCTTCCATTTTGGGAGTACCT	HEX
Ln 40551	(TA) <sub>9</sub>	F: TAACTGATGCATGCCAGGAACTT R: GCTTCCATAAAGCTCCACAGAT	FAM
Ln 42701	(AT) <sub>9</sub>	F: TGGACTTCACAAGCGTGGCT R: ATTGAAAGTGTGCATAGTCCGT	FAM

**Tabella 3.3:** Elenco dei 21 *loci* microsatellite analizzati in questa tesi. La tabella riporta il nome del *locus*, il pattern di *repeat*, la sequenza dei *primer* (F *forward* e R *reverse*) e il fluoroforo associato al *primer*.

<i>Locus</i>	<i>Repeat</i>	<i>Sequenza primer</i>	<i>Fluoroforo</i>
Ln 42233	(AT) <sub>9</sub>	F: ATTCCTGACTGCCAAAGACG R: TGTTGACTTGCAGACTGAACG	FAM
Ln 36100	(AG) <sub>9</sub>	F: TAAAGCTCCAGCCGATACTGG R: TGGGAATTGAACAGATGTCACC	TAMRA
Ln 35217	(CT) <sub>9</sub>	F: CCACTACACAGCTGAAATGGT R: GGCATGATTGGAGCACTTCC	HEX
Ln 23194	(TA) <sub>9</sub>	F: TCCGGAGGCTGCGTTCAGAG R: CTGCTTGAGGACCGGCTCAG	HEX
Ln 22268	(GA) <sub>9</sub>	F: TGGATGGTTTTTCGTTTGACCAA R: TGAAGAGCTTTGGAGCAACAA	TAMRA
Ln 36156	(GT) <sub>9</sub>	F: GTCTGCTCATGTCCAACACAT R: AGGTGTTTCTGTAAACCCTCG	FAM
Ln 45589	(AG) <sub>9</sub>	F: CGTGTCTAGGAGCTACAGCA R: CCAGGTTGGTGGCATCTCTT	TAMRA
Ln 45257	(TG) <sub>9</sub>	F: AAGACCAGGGCGAGTCTGAC R: CTCTAACCAGAACACCGCCT	HEX
Ln 42016	(GT) <sub>9</sub>	F: ACATCACTTGCAAATAGGCCGGTG R: CCGTCATGCACCTGGGAGTT	HEX
Ch 126	(AAT) <sub>7</sub>	F: CGGTTTTTATGCATGTTGCCA R: ACTGCTCATTACACTGGTTC	NED
Ch 3866	(CA) <sub>9</sub>	F: AGCGTTACACACTCCCATCCGT R: CTGCACCACTCTACCAGGGACG	NED
Ca 35	(AGAC) <sub>5</sub>	F: CCCAACTCCTCTCAATGTCTG R: TCGGCTGTGTAACCTCCTCCA	HEX
Ch 5817	(GT) <sub>9</sub>	F: TTTAAAGCTGGGAAACACAGG R: GAAACTGTGACAAACACAGGCT	NED
Ch 623	(AAC) <sub>7</sub>	F: GCTGTTTGATTCCCTCGTGAGG R: AAAAGTGGTCCCTCCGCTGCAGT	PET
Ch 1968	(ACA) <sub>7</sub>	F: AGCTCAAGGTGTTGCGAGACG R: TAGCCAGCAGCGCTAATCCTG	PET
Ch 3603	(TTTAA) <sub>5</sub>	F: TGAGACAAGTCAAATTCCAACC R: GGCATAAAGCTATTTGAGGCT	VIC
Ch 2788	(CT) <sub>9</sub>	F: TGGTTTCGATGAAGAATGCTC R: TGATGATATTGGTCGTGGTCCG	VIC
Ca 48	(AC) <sub>15</sub>	F: CATTCTCCCTGAGTTTACCAC R: TCTTCTCTTTACCTCTAC(AG)G	TAMRA

**Tabella 3.3 (continuazione):** Elenco dei 21 *loci* microsatellite analizzati in questa tesi. La tabella riporta il nome del *locus*, il pattern di *repeat*, la sequenza dei *primer* (F *forward* e R *reverse*) e il fluoroforo associato al *primer*.

### 3.2.1 Amplificazione dei marcatori microsatellite tramite PCR

La *Polymerase Chain Reaction* (PCR) è una tecnica scoperta e pubblicata nel 1985 che consiste nella replicazione del DNA *in vitro*, durante la quale una sequenza di DNA bersaglio si duplica ad ogni ciclo. La PCR permette di selezionare ed amplificare esponenzialmente una particolare regione di DNA sfruttando la capacità di sintesi dell'enzima DNA polimerasi, ottenendo come risultato numerose copie partendo da quantità estremamente ridotte di DNA. È una tecnica veloce, sensibile e richiede una conoscenza parziale della sequenza di DNA bersaglio in quanto è sufficiente conoscere le sequenze, ai due estremi della sequenza bersaglio, a cui far associare i *primer*. La DNA polimerasi utilizzata (Taq polimerasi) deve essere termostabile in modo da poter lavorare a temperature elevate senza denaturarsi.

Il materiale di partenza per la PCR è un campione di DNA a doppio filamento, che contiene la sequenza nucleotidica "selezionata" per la copiatura. A questo campione vengono aggiunti la DNA polimerasi, i 4 nucleotidi (dNTP's) e gli inneschi specifici (*primer*) per il particolare segmento di DNA che deve essere amplificato.

Un ciclo di PCR consiste in tre fasi:

- **denaturazione**: breve riscaldamento a 94°C per 30/60" per separare i due singoli filamenti di DNA;
- **annealing**: raffreddamento a 50-65°C per 30/60" per permettere ai *primer* di associarsi ai filamenti di DNA denaturati mediante la formazione di legami a idrogeno;
- **elongazione** per reazione enzimatica: sintesi di un nuovo filamento complementare al DNA stampo da parte della DNA polimerasi mediante l'aggiunta di nucleotidi all'estremità 3' di ogni *primer* (72°C);

Alla fine di ciascun ciclo la sequenza di DNA selezionata (lunga anche centinaia di basi) è stata duplicata. Si può procedere a numerosi cicli di amplificazione semplicemente variando la temperatura della miscela di reazione. In media, la reazione viene ripetuta per 30 cicli generando anche 1 milione di copie della sequenza bersaglio.

L'esito di una PCR viene generalmente visualizzata mediante elettroforesi su gel che consiste nella migrazione su matrice porosa di molecole di DNA (cariche negativamente) sfruttando un campo elettrico. Il DNA viene separato in base alle dimensioni poiché generalmente molecole più piccole corrono più rapidamente.

L'analisi dei *loci* microsatellite è stata svolta in un'unica reazione di amplificazione detta PCR *multiplex*, utilizzando più coppie di *primer* insieme per produrre più frammenti corrispondenti a più marcatori.

Il successo delle reazioni di amplificazione in PCR è stato poi verificato mediante una corsa elettroforetica su gel d'agarosio all'1.8 % in tampone TBE 1X. I prodotti di amplificazione ottenuti sono stati utilizzati per procedere alla genotipizzazione di tutti i campioni.

### **3.2.2 Genotipizzazione o *scoring* degli alleli**

Il servizio di genotipizzazione mediante separazione dei prodotti di PCR tramite elettroforesi capillare con sequenziatore automatico Applied Biosystems 3130 XL è stato svolto presso l'azienda Life Technologies (USA).

Per l'elaborazione dei risultati è stato utilizzato il software GENEMARKER VER. 2.6.3 (SoftGenetics LLC) che permette la caratterizzazione del genotipo in ogni *locus* microsatellite per ogni individuo, confrontando le sequenze di DNA. Il programma permette di visualizzare ogni frammento amplificato corrispondente ad un *locus* microsatellite. Ogni *locus* viene visualizzato in un grafico (elettroferogramma) avente in ascissa la dimensione del frammento in paia di basi (bp) derivata dal tempo di migrazione su gel e in ordinata l'intensità del prodotto amplificato. Il dimensionamento dei singoli alleli avviene per confronto con lo *standard* interno. In un genoma diploide, un omozigote presenta un singolo picco e i due alleli hanno la stessa lunghezza mentre un eterozigote presenta due picchi e i due alleli sono diversi, permettendo di discriminare tra uno e l'altro. Esaminando il profilo

di ciascun *locus* microsatellite è possibile individuare tutti gli alleli presenti in un *locus* e quindi identificare il genotipo di tutti gli individui analizzati. Questo processo è noto come *scoring* e può essere compromesso da alcuni problemi biologici o tecnici che possono verificarsi durante l'amplificazione dei microsatelliti, portando a falsificare le reali frequenze alleliche e genotipiche.

Alcuni problemi biologici riguardano:

- *Stuttering*: scivolamento della DNA polimerasi durante l'amplificazione di ripetizioni in tandem (fenomeno analogo a quello che porta alla mutazione in natura dei microsatelliti). Alcuni prodotti di PCR possono quindi essere più corti di una o più unità di ripetizione rispetto al numero di ripetizioni dell'allele iniziale. Nell'elettroferogramma si producono picchi in una posizione diversa da quella dell'allele parentale e in genere i prodotti dello *stuttering* hanno un'altezza inferiore al 15% del picco reale. È più frequente con strutture a ripetizione semplice e con alleli lunghi. L'interpretazione di questi *loci* può risultare difficoltosa quando si è in presenza di un genotipo eterozigote per alleli adiacenti. In questo caso, infatti, lo *stuttering* può portare a considerare questo genotipo eterozigote come un omozigote per l'allele più grande;
- *Large allele drop-out*: si osserva principalmente nei *loci* che presentano un'elevata differenza di dimensioni tra i due alleli ed è causato dall'amplificazione preferenziale dell'allele più piccolo in un genotipo eterozigote. Se non rilevato, il *large allele drop-out* può causare una sovrastima delle frequenze di alleli più corti e risultare nella totale omissione dal *dataset* di alleli rari e grandi e in un deficit dell'eterozigosità osservata;
- Pattern multiallelici: possono essere causati da un errato appaiamento dei *primer* al filamento stampo;
- *Split peaks*: adenilazione accidentale al 3' dei prodotti di PCR. Per ciascun allele si ottengono due picchi distanziati di una sola base nell'elettroferogramma. Di conseguenza diventa difficoltoso il riconoscimento dei picchi, soprattutto in presenza di genotipi eterozigoti per alleli molto vicini tra loro;

- *Allele drop-out* (alleli nulli): mutazione nel sito di *annealing* di un *primer* che determina il suo mancato attacco e la non amplificazione di uno dei due alleli in un eterozigote. Gli alleli nulli sono i più difficili da individuare in quanto si tratta di alleli presenti nella popolazione che però non vengono amplificati. La presenza di alleli nulli in un *dataset* fa sì che gli eterozigoti per l'allele nullo vengano classificati come omozigoti. Gli omozigoti per l'allele nullo non sono individuati. In studi di genetica di popolazione, un'elevata percentuale di omozigoti, associata a disequilibrio di Hardy-Weinberg (si veda la sezione 3.4.1.3), è spesso indicativa della presenza di alleli nulli ad un certo *locus*;
- Omoplasia: fenomeno che non permette di distinguere, mediante elettroforesi, due alleli non omologhi, perché di discendenza diversa, ma di uguale lunghezza.

Lo *scoring* deve sempre essere eseguito da almeno due operatori in modo indipendente in modo da ridurre al minimo ulteriori errori dovuti alla distrazione o all'interpretazione dei pattern allelici più complessi.

### 3.2.3 Binning o allele calling

Il passo successivo allo *scoring* consiste nell'assegnare un numero intero, o *bin*, agli alleli osservati sulla base della loro lunghezza in paia di basi (bp). Il software GENEMARKER fornisce i risultati di lunghezza dei frammenti nella forma di un numero intero con due numeri decimali. Risulta necessario arrotondare i dati di *scoring* e creare categorie alleliche, un processo noto come *binning* o *allele calling*, in quanto tutti i programmi di analisi di microsatelliti lavorano con dimensioni alleliche intere.

In questa tesi è stato utilizzato il metodo di *binning* automatico del software FLEXIBIN (Amos *et al.*, 2007) che ha permesso di generare un profilo completo per ogni campione composto da una coppia di numeri per ciascun *locus*, rappresentativi del numero di ripetizioni presenti in ciascun allele.



### 3.2.4 Correzione *dataset*

Prima di procedere all'analisi dei dati molecolari è stato necessario eliminare i campioni poco informativi dal *dataset*. Nel caso specifico di questa tesi, il primo passo è consistito nell'identificazione ed eliminazione di eventuali *loci* monomorfici per tutte e quattro le specie. Successivamente è stato necessario scartare tutti gli individui o *loci* con meno dell'80% dell'informazione. Infine è stato necessario eliminare tutti i *loci* che non rispettano l'equilibrio di Hardy-Weinberg, i *loci* in *linkage disequilibrium* e i *loci* che presentavano alleli nulli. Le prime due correzioni vengono fatte mediante l'utilizzo del programma GENEPOP ON THE WEB ver. *online* 4.7 (Rousset, 2008) mentre l'ultima correzione viene fatta utilizzando il *software* MICROCHECKER ver. 2.2.3 (Van Oosterhout *et al.*, 2004).

A partire del *dataset* corretto, con il *software* CREATE ver. 1.38 (Coombs *et al.*, 2008) sono stati creati tutti i file di *input* idonei per i programmi utilizzati successivamente. CREATE formatta i valori interi, risultanti dal *binning*, in codici allelici necessari per le analisi con altri programmi.

I problemi sopra individuati potrebbero ridurre sia la precisione che l'accuratezza delle successive analisi genetiche, rendendo perciò necessario l'utilizzo del *dataset* risultante da questo processo di correzione. È necessario però tenere a mente che maggiore è il numero di *loci* analizzati più affidabile sarà il *dataset* risultante e di conseguenza riducendo il numero di *loci* si potrebbe anche ridurre il potere statistico delle successive analisi. Chiaramente è necessario trovare un compromesso tra le due differenti necessità.

### 3.3 DNA mitocondriale (mtDNA)

Il DNA mitocondriale è localizzato nei mitocondri, piccoli organelli presenti nel citoplasma delle cellule, dove si trova in un numero elevato di copie (1000-10000 copie per cellula). È piccolo, circolare e compatto.

Il 93-95% del mtDNA è codificante, superavvolto e non organizzato in nucleosomi. Costituisce un patrimonio genetico non correlato al genoma nucleare. Si tratta di un genoma aploide ad eredità uniparentale. Nella maggior parte dei vertebrati ma anche in molti invertebrati il 99.99% dei genomi mitocondriali provengono dalla madre (eredità matrilineare). Presenta un tasso di mutazione 6-7 volte più elevato rispetto al genoma nucleare. Questa differenza dipende dall'accumulo di mutazioni e non da ricombinazione in quanto viene ereditato in modo diretto. Alcune regioni sono molto conservate, mentre altre sono ipervariabili (es. *D-loop Control Region*). Molti marcatori molecolari frequentemente utilizzati in analisi molecolari sono localizzati in particolari porzioni del genoma mitocondriale. Sono marcatori aploidi, meno "potenti" dei marcatori autosomici, ma molto utili grazie ad alcune caratteristiche chiave:

- Elevato numero di copie che permette l'analisi di campioni con materiale cellulare ridotto o degradato;
- Eredità matrilineare che permette la ricostruzione della linea parentale materna;
- Possibilità di analisi di cellule anucleate;
- Tasso di accumulo di polimorfismi più elevato che nel genoma nucleare.

In questa tesi è stata amplificata la regione di DNA mitocondriale codificante per il gene ND2. Questa è una delle regioni maggiormente ipervariabile, permettendo il confronto tra specie filogeneticamente vicine e tra individui appartenenti alla stessa specie.

### **3.3.1 Amplificazione del DNA mitocondriale tramite PCR**

Per l'amplificazione in PCR del gene mitocondriale ND2 sono stati utilizzati i *primers* GLN 5'-CTACCTGAAGAGATCAAAAC-3' e ASN 5'-CGCGTTTAGCTGTAACTAA-3', seguendo la procedura descritta in Kocher, Conroy, McKaye, Stauffer e Lockwood (1995).

La buona riuscita delle reazioni di amplificazione in PCR è stata poi verificata mediante una corsa elettroforetica su gel d'agarosio all'1.8 % in tampone TBE 1X.

La qualità del sequenziamento, effettuato dalla ditta Eurofins Genomics (Germania), è stata verificata tramite il software MACVECTOR VER. 12.6 (Rastogi, 2000). Le estremità sono state tagliate fino a comprendere unicamente il gene ND2, ottenendo una sequenza di 1047 paia di basi.

### 3.4 Analisi dei dati

A partire del *dataset* corretto sono stati creati tutti i file di *input* compatibili per le analisi di differenziamento genetico svolte successivamente. Le analisi dei dati sono state condotte utilizzando programmi diversi, ciascuno di essi idoneo a studiare e descrivere particolari parametri genetici.

I programmi utilizzati in questa tesi sono:

- CREATE *ver. 1.38* (Coombs *et al.*, 2008), utilizzato per formattare gli *input file* relativi ai *loci* microsatellite, espressi in valori interi risultanti dal *binning*, in codici allelici necessari per le analisi con altri programmi;
- Funzione *divBasic* del pacchetto *diveRsity* in RStudio *ver. 4.0.5*, utilizzato per calcolare le statistiche di base dei *loci* microsatellite. Sono stati calcolati: il numero di individui genotipizzati per ciascun *locus* (N), il numero di alleli osservati per ciascun *locus* (A), la percentuale di alleli totali osservati per campione per locus (%), l'*allelic richness* ( $A_r$ ), l'eterozigosità osservata ( $H_o$ ) e l'eterozigosità attesa ( $H_e$ );
- MICRO-CHECKER *ver. 2.2.3* (Van Oosterhout *et al.*, 2004), utilizzato per individuare la presenza di alleli nulli nel *dataset* relativo ai *loci* microsatellite;

- GENEPOP ON THE WEB ver. *online* 4.7 (Rousset, 2008), utilizzato per individuare i *loci* microsatellite che non rispettano l'Equilibrio di Hardy-Weinberg e/o sono in *linkage disequilibrium*;
- MYRIADS ver. 1.2 (Carvajal-Rodriguez, 2018), utilizzato per correggere la significatività statistica (*p-value*) ottenuta mediante test multipli di verifica dell'Equilibrio di Hardy-Weinberg e del *linkage disequilibrium*, utilizzando il metodo di correzione Benjamini-Hochberg (Benjamini e Hochberg, 1995);
- ARLEQUIN ver. 3.5.2.2 (Excoffier e Lischer, 2010), utilizzato per indagare la struttura di popolazione e il differenziamento tra specie attraverso la stima dell'indice di fissazione ( $F_{ST}$ ) e del numero medio di differenze nucleotidiche tra coppie di popolazioni e/o specie ( $P_i$ );
- STRUCTURE ver. 2.3.4 (Pritchard *et al.*, 2000), utilizzato per suddividere il *dataset* relativo ai *loci* microsatellite in gruppi (K) di individui geneticamente differenziati, mostrando la struttura di popolazione e permettendo l'identificazione di specie distinte;
- Pacchetto *pophelper* ver. 2.3.1. in RStudio ver. 4.0.5, utilizzato per rappresentare graficamente i risultati ottenuti con *Structure*;
- Pacchetti *adegenet* ver. 2.1.4 per il calcolo e *ggplot2* ver. 4.0.5 per la visualizzazione grafica dell'analisi delle componenti principali (PCA) svolta in RStudio ver. 4.0.5 a partire dal *dataset* relativo ai 16 *loci* microsatellite;
- DNA Sequence Polymorphism ver. 6.12.03 (Rozas *et al.*, 2017), utilizzato per calcolare le statistiche descrittive partendo dagli *input file* relativi alle sequenze del gene mitocondriale ND2. Sono stati calcolati, per ogni specie e popolazione, il numero di siti segreganti (S), il numero di aplotipi ( $H_{ap}$ ), la *haplotype diversity* ( $H_d$ ) e la *nucleotide diversity* ( $\pi$ );
- POPART ver. 1.7.2 (Leigh e Bryant, 2015), che permette la visualizzazione grafica delle relazioni genetiche (*haplotype network*) tra tutti gli individui analizzati.

### 3.4.1 Marcatori microsatellite

#### 3.4.1.1 *Correzione dataset*

Prima di procedere con le analisi di differenziamento genetico è stato necessario eliminare dal *dataset* di partenza i *loci* o gli individui poco informativi e di conseguenza poco utili alle successive analisi, a causa dei seguenti motivi:

- Presenza di *loci* monomorfici caratterizzati da una diversità genetica pari a zero a causa della presenza di un solo allele all'interno delle specie in esame;
- Mancanza di più del 20% dei dati;
- Presenza di alleli nulli (si veda la sezione 3.2.2).

#### 3.4.1.2 *Frequenze alleliche*

Due concetti essenziali da comprendere per poter descrivere la variabilità e la struttura genetica di una popolazione o specie sono quelli di frequenza allelica e frequenza genotipica. La frequenza genotipica corrisponde alla somma degli individui che possiedono un certo genotipo divisa per il numero totale di individui. In una popolazione che si riproduce sessualmente i genotipi sono insiemi temporanei di alleli, si suddividono a ogni generazione e i singoli alleli vengono trasmessi alla generazione successiva. Il tipo e il numero di alleli, al contrario, presentano una continuità da una generazione all'altra e rappresentano il *pool* genico di una popolazione o specie. La frequenza allelica corrisponde alla frequenza relativa di un allele ad un *locus* rispetto alla totalità degli alleli presenti nella popolazione, espressa come frazione o percentuale. Per calcolare la frequenza allelica ( $p_i$ ) di un allele *i*-esimo di un determinato *locus*, a partire dal numero di genotipi osservati, si può utilizzare la seguente formula:

$$p_i = \frac{(2n_1 + n_2)}{2N}$$

dove  $n_1$  rappresenta il numero di omozigoti per l'allele i-esimo,  $n_2$  il numero di eterozigoti per l'allele i-esimo ed N il numero di individui. La somma di tutte le frequenze alleliche in una popolazione è sempre uguale a 1.

In genetica di popolazione, i valori delle frequenze alleliche sono parametri importanti per stimare la diversità genetica e la ricchezza del *pool* genico di una popolazione o specie.

### **3.4.1.3 Equilibrio di Hardy-Weinberg**

La legge di Hardy-Weinberg è uno dei più importanti principi della genetica di popolazione. Questa legge è in realtà un modello matematico che valuta gli effetti della riproduzione sulle frequenze alleliche e genotipiche di una popolazione. La legge di Hardy-Weinberg stabilisce che in una popolazione ideale di dimensioni infinitamente grandi, caratterizzata da accoppiamenti casuali e non affetta da mutazioni, migrazione o selezione naturale, le frequenze alleliche non cambiano e le frequenze genotipiche si stabilizzano dopo una generazione.

Nel caso di un *locus* autosomico con n alleli ( $A_1, A_2, \dots, A_n$ ), le frequenze genotipiche attese secondo la legge di Hardy-Weinberg sono:

$p_i^2$  (frequenza omozigoti  $A_iA_i$  per l'allele i-esimo)

$2p_i p_j$  (frequenza eterozigoti  $A_iA_j$ )

dove  $p_i$  è la frequenza allelica dell'allele i-esimo (si veda sezione 3.4.1.2). In una popolazione in equilibrio di Hardy-Weinberg le frequenze genotipiche sono legate dalla seguente equazione:

$$p_i^2 + 2p_i p_j + p_j^2 = 1$$

La legge di Hardy-Weinberg indica che, quando le assunzioni sono soddisfatte, la sola riproduzione non può dare luogo a evoluzione. Perché una popolazione si evolva sono necessari altri processi come la selezione naturale, la mutazione, la migrazione o il caso. Di conseguenza una popolazione può trovarsi in disequilibrio a causa dei seguenti fattori:

- Accoppiamento non casuale rispetto al genotipo: un'unione non casuale influenza il modo in cui gli alleli si combinano per formare i genotipi e altera le frequenze genotipiche della popolazione. Tipi di accoppiamenti non casuali sono: l'unione assortativa positiva (preferenza ad accoppiarsi tra individui somiglianti), l'unione assortativa negativa (preferenza ad accoppiarsi tra individui non somiglianti) e l'*inbreeding* (unione preferenziale tra parenti biologici);
- Mutazione: mutazioni ricorrenti determinano una variazione nelle frequenze degli alleli. All'equilibrio le frequenze alleliche sono funzione dei tassi di mutazione in avanti e inversa. Dato che i tassi di mutazione sono molto bassi, l'effetto delle mutazioni sulle frequenze alleliche per generazione è assai ridotto;
- Migrazione: la presenza di migrazioni determina variazioni nelle frequenze alleliche delle popolazioni introducendovi alleli provenienti da altre popolazioni. L'entità della variazione legata alla migrazione dipende sia dall'estensione della migrazione sia dalla differenza nelle frequenze alleliche fra popolazione sorgente e ricevente. La migrazione fa diminuire le differenze genetiche fra popolazioni e aumenta la variabilità genetica al loro interno;
- Deriva genetica: origina dal fatto che le dimensioni di una popolazione rimangono basse a lungo, dall'effetto del fondatore (la formazione di una popolazione a partire da pochi individui) e dall'effetto collo di bottiglia (riduzione della popolazione). La deriva genetica determina variazione nelle frequenze alleliche all'interno delle popolazioni, perdita di variabilità genetica per fissazione degli alleli e divergenza genetica fra popolazioni;
- Selezione naturale: riproduzione differenziale di genotipi che si misura in termini di *fitness*, ovvero il successo riproduttivo di un

genotipo rispetto a quello degli altri genotipi presenti nella popolazione.

Per determinare se i genotipi di una popolazione sono in equilibrio di Hardy-Weinberg, occorre confrontare le frequenze genotipiche attese per la legge di Hardy-Weinberg con le frequenze genotipiche osservate. Per farlo si può effettuare un test del chi-quadro ( $\chi^2$ ), valutando se il discostamento tra valori attesi e osservati sia imputabile esclusivamente al caso.

Prima di procedere alle analisi di differenziamento genetico è stato necessario eliminare dal *dataset* tutti i *loci* che non rispettavano l'equilibrio di Hardy-Weinberg utilizzando il programma GENEPOP ON THE WEB (Rousset, 2008).

#### **3.4.1.4 Linkage disequilibrium**

Prima di introdurre il concetto di *linkage disequilibrium* è utile riportare brevemente i principi di Mendel della segregazione e dell'assortimento indipendente. Il principio della segregazione stabilisce che ogni organismo diploide possiede due alleli ad ogni *locus* che si separano durante la meiosi e ciascuno finisce in un gamete. Secondo il principio dell'assortimento indipendente, al momento della separazione i due alleli di un *locus* si separano indipendentemente dagli alleli di altri *loci*. Questa separazione indipendente degli alleli dà luogo alla ricombinazione, cioè la formazione di nuove combinazioni di alleli. Alcuni *loci* sono però collocati sullo stesso cromosoma e potrebbero non assortire in modo indipendente, segregando insieme durante la meiosi e venendo ereditati congiuntamente. Il *linkage* fra *loci* è contrastato da un processo detto *crossing-over* che determina la ricombinazione: in altre parole, rompe l'associazione fra *loci* che si trovano vicini sullo stesso cromosoma. Il livello di ricombinazione fra i *loci* è però direttamente collegato alla distanza tra loro. Di conseguenza, *loci* che si trovano così vicini tra loro



sul medesimo cromosoma da non essere soggetti a *crossing-over* si trovano in *linkage disequilibrium*.

Il *linkage disequilibrium* consiste quindi nell'associazione non casuale tra alleli presenti in due o più *loci* (Slatkin, 2008) all'interno di una popolazione, che porta a specifiche combinazioni alleliche che presentano una frequenza diversa da quella attesa in caso di *loci* indipendenti. Avviene quindi uno scostamento dalle assunzioni relative all'indipendenza tra *loci* e all'accoppiamento casuale all'interno delle popolazioni.

Oltre alla presenza di *loci* vicini sullo stesso cromosoma, la presenza di *linkage disequilibrium* può essere determinata anche dal mescolamento di popolazioni con differenti frequenze alleliche, dalla selezione naturale positiva per particolari combinazioni di *loci* che aumentano la *fitness*, dalla deriva genica casuale, dall'*inbreeding* o da forti riduzioni della dimensione della popolazione (Slatkin, 2008).

Il *linkage disequilibrium* viene calcolato a partire dalla differenza tra la frequenza osservata di una particolare combinazione di alleli in due *loci* e la frequenza attesa secondo un'associazione casuale.

È stato molto importante verificare che i *loci* analizzati in questa tesi non fossero statisticamente associati, ovvero che il genotipo di ciascun *locus* fosse indipendente da quello di tutti gli altri. Per farlo è stato utilizzato il programma GENEPOP ON THE WEB (Rousset, 2008). La necessità di questa verifica dipende dal fatto che le analisi statistiche di differenziamento genetico prevedono come assunzione di base un'associazione casuale tra gli alleli dei diversi *loci*.

#### **3.4.1.5 Allelic richness**

L'*allelic richness*, detta anche ricchezza allelica, consiste nel numero medio di alleli per *locus*, indipendentemente dalla dimensione del campione (Foulley e Ollivier, 2006). Il numero di alleli osservati per ciascun *locus* dipende dalla dimensione del campione in quanto ci si

aspetta che campioni di grandi dimensioni contengano molti più alleli rispetto a campioni più piccoli. L'*allelic richness* fornisce quindi una misura del numero di alleli standardizzata rispetto al campione più piccolo, in modo da poter confrontare in modo coerente campioni con numerosità differente (Leberg, 2002).

L'*allelic richness* è un parametro molto importante per misurare la diversità genetica all'interno di una popolazione o di una specie ed è molto utilizzato in studi di genetica di conservazione, in quanto evidenzia eventi di *bottleneck* recenti e i potenziali evolutivi in modo più efficiente rispetto alla stima dell'eterozigosità osservata.

#### **3.4.1.6 Eterozigosità osservata e attesa**

Un altro indice molto importante della variabilità genetica è l'eterozigosità (H), indice molto versatile che consente di avere stime a livello di popolazione e di specie, confrontabili tra loro. Valori elevati di eterozigosità indicano una buona diversità genetica all'interno delle popolazioni o specie esaminate. Con eterozigosità si intende la frequenza di eterozigoti a un dato *locus*. Si può distinguere tra eterozigosità osservata ( $H_o$ ), che indica la percentuale di genotipi eterozigoti effettivamente osservati per ciascun *locus*, e eterozigosità attesa ( $H_e$ ). La stima più importante della variabilità genetica è l'eterozigosità attesa la cui importanza deriva dal fatto che nella sua stima sono comprese le frequenze alleliche del *locus* studiato:

$$H_e = 1 - \sum x_i^2$$

dove con  $x_i$  si intende ciascuna delle frequenze degli alleli al *locus*. Si parla di eterozigosità attesa in quanto è il valore atteso per la frequenza degli eterozigoti se la popolazione fosse all'Equilibrio di Hardy-Weinberg e, nel caso di un locus a due alleli, corrisponde a:

$$H_e = 2pq$$

dove  $p$  e  $q$  indicano le frequenze dei due alleli.

L'entità dello scostamento tra  $H_e$  e  $H_o$  si utilizza spesso per verificare se un *locus* rispetta l'equilibrio di Hardy-Weinberg.

### **3.4.1.7 Differenziamento genetico inter ed intraspecifico**

Per studiare il livello di differenziamento delle popolazioni o delle specie, si utilizza principalmente il concetto di *inbreeding*, che permette di stabilire l'esistenza di un'origine comune, sulla base del numero di alleli uguali per discesa (copie dello stesso allele di un antenato) che condividono. Nelle popolazioni reali gli incroci non sono mai completamente casuali, ma comprendono un certo grado di assortatività (si veda la sezione 3.4.1.3). In genetica di popolazione, il caso più rilevante di unione assortativa è dato dalle unioni tra consanguinei, detta inincrocio o *inbreeding*. Questo fenomeno non determina la variazione delle frequenze alleliche, ma solo di quelle genotipiche, in quanto ha l'effetto di aumentare gli omozigoti e ridurre gli eterozigoti.

Il coefficiente di *inbreeding* ( $F$ ) è la probabilità che un dato individuo possieda ad un dato *locus* due alleli uguali per discendenza. Il valore di  $F$  può essere stimato secondo la seguente formula:

$$F = 1 - \left( \frac{H_o}{H_e} \right)$$

dove  $H_o$  è l'eterozigosità osservata e  $H_e$  è l'eterozigosità attesa all'equilibrio di Hardy-Weinberg.

In una popolazione caratterizzata da *inbreeding*, con un valore di  $F > 0$ , le frequenze genotipiche all'equilibrio cambiano e si dice che la popolazione è all'equilibrio di Wright. Si può notare come l'equilibrio di Hardy-Weinberg sia un caso particolare di quello di Wright quando  $F = 0$ . Invece nel caso in cui  $F = 1$ , la popolazione sarà composta solo da omozigoti, con frequenze genotipiche pari a quelle alleliche, cioè i due alleli si saranno fissati. Da qui deriva la denominazione dell'indice  $F$  come indice di fissazione. L'indice  $F$  è molto versatile in quanto mette in

relazione i livelli di variabilità genetica con la discendenza diretta degli alleli stessi.

Un eccesso di omozigosità può essere causato, oltre che dal fenomeno di *inbreeding*, anche da una suddivisione interna delle popolazioni in più sottopopolazioni (effetto di Wahlund). Wright derivò tre tipi diversi di indice F per lo studio della struttura di una popolazione e la sua suddivisione in sottopopolazioni, che vanno sotto il nome di “*F statistics*” e sono  $F_{IS}$ ,  $F_{IT}$  e  $F_{ST}$ . I tre indici possono essere visti come la suddivisione in tre componenti della varianza allelica di un dato gene nell'intera popolazione. Nel modello di Wright, una popolazione può essere considerata come l'insieme di più sottopopolazioni. Facilmente si può ampliare il concetto a una specie divisa nelle popolazioni che la compongono. Gli F-statistici rappresentano un set di strumenti individuati da Wright per ripartire il deficit di eterozigosi all'interno degli individui appartenenti alla stessa popolazione e tra individui appartenenti a popolazioni diverse:

- $F_{IT}$  misura il deficit globale di eterozigosi;
- $F_{IS}$  misura il deficit di eterozigosi all'interno di ogni popolazione;
- $F_{ST}$  (indice di fissazione) misura il deficit di eterozigosi tra le popolazioni.

$F_{ST}$  è uno degli indici più usati in genetica di popolazione e corrisponde alla probabilità che due alleli estratti a caso da sottopopolazioni (o popolazioni) diverse siano identici per discendenza. Dato che la presenza di alleli identici per discendenza implica un antenato comune, è un ottimo indice del grado di differenziamento tra popolazioni dovuto alla struttura genetica. Valori bassi di  $F_{ST}$  indicano bassi livelli di differenziamento, mentre livelli più alti sono tipici di popolazioni che non sono più in contatto genetico da diverse generazioni. Un valore pari a zero implica una completa panmissia, ovvero il caso in cui due popolazioni si incrociano liberamente. Un valore pari a uno implica invece che tutta la variazione genetica è dovuta alla struttura della popolazione, e che le due popolazioni non condividono alcuna diversità genetica. La

presenza di valori anche molto bassi di  $F_{ST}$  rende necessario l'uso di metodologie statistiche per stabilire se i valori ottenuti siano significativamente diversi da zero. Anche se risulta molto difficile costruire una scala di valori universale, si possono definire valori bassi quelli minori di 0.1, valori medi quelli compresi tra 0.1 e 0.15-0.2, valori alti quelli superiori a 0.2.

Con il programma ARLEQUIN *ver.* 3.5.2.2 (Excoffier e Lischer, 2010) si è indagato il differenziamento genetico tra specie o popolazioni, mediante il calcolo dell'indice  $F_{ST}$ .

#### **3.4.1.8 Identificazione di gruppi geneticamente distinti**

Per poter rappresentare graficamente i dati contenuti nel *dataset* è stato utilizzato il metodo dell'analisi delle componenti principali (detta anche PCA) che è una tecnica utilizzata nell'ambito della statistica multivariata per la semplificazione dei dati d'origine. Lo scopo primario di questa tecnica è la riduzione del numero di variabili che influiscono sui dati. Ciò avviene tramite una trasformazione lineare delle variabili che determina la proiezione di quelle originarie all'interno di un sistema cartesiano nel quale le nuove variabili vengono ordinate in ordine decrescente di varianza. La variabile con maggiore varianza viene proiettata sul primo asse, la seconda sul secondo asse e così via. La riduzione della complessità consiste nell'analizzare esclusivamente le due variabili (due assi cartesiani X e Y) che descrivono la varianza maggiore.

In questa tesi la PCA è stata calcolata mediante il pacchetto *ade4* *ver.* 2.1.4 e il suo risultato rappresentato tramite il pacchetto *ggplot2* *ver.* 4.0.5 in RStudio, permettendo di condensare le informazioni contenute nel *dataset* in un grafico.

In qualsiasi specie, la somiglianza genetica tra gli individui è strutturata tramite la presenza di sottogruppi, spesso isolati geograficamente. Capire la struttura di popolazione di una specie è di fondamentale

importanza e per fare questo è stato utilizzato il programma STRUCTURE ver. 2.3.4 (Pritchard *et al.*, 2000) che permette di identificare possibili gruppi di popolazioni o di specie a partire dai dati genotipici *multilocus*. Il programma viene generalmente utilizzato per determinare la presenza di popolazioni distinte, assegnare individui di origine sconosciuta alle diverse popolazioni o specie, identificare migranti e individui ibridi e stimare le frequenze alleliche delle popolazioni in situazioni in cui molti individui sono migranti o ibridi. Può essere applicato alla maggior parte dei marcatori genetici comunemente usati, inclusi i marcatori microsatellite (Porras-Hurtado *et al.*, 2013). Questo programma utilizza una metodologia di *clustering* per dedurre la struttura di popolazione usando i dati relativi a marcatori molecolari *multilocus*. L'assunzione di base è un modello con K gruppi, ciascuno dei quali caratterizzato da uno specifico set di frequenze alleliche per ciascuno dei *loci* utilizzati, assumendo che i *loci* siano in equilibrio di Hardy-Weinberg ed in *linkage equilibrium* all'interno delle diverse popolazioni. Utilizzando un algoritmo iterativo bayesiano e un metodo di simulazione basato sulle catene di Markov MonteCarlo (MCMC), il programma calcola la probabilità che individui di origine sconosciuta appartengano a ciascuno dei K gruppi in base al loro genotipo, in modo da minimizzare il *linkage disequilibrium* e il disequilibrio di Hardy-Weinberg. Il processo inizia assegnando casualmente gli individui a un numero predeterminato di gruppi (K), quindi le frequenze alleliche vengono stimate in ciascun gruppo e gli individui riassegnati sulla base di tali stime. La simulazione comprende un periodo di *burnin*, in cui non avviene il campionamento dei dati, ed una fase di corsa vera e propria in cui i valori dei parametri vengono campionati. Poiché ogni passaggio dipende dai valori del parametro del passaggio precedente, e si possono formare delle correlazioni tra i diversi *step*, è opportuno che vengano impostati periodi sufficientemente lunghi sia per il *burnin* sia per la corsa stessa, per limitare tali correlazioni. Questo processo viene ripetuto molte volte, traducendosi in una progressiva convergenza verso stime affidabili delle frequenze alleliche e il calcolo delle probabilità di appartenenza dei diversi individui a ciascun gruppo (Q).

Informazioni preliminari sui campioni di studio (come area geografica d'origine o presenza di un certo fenotipo) possono essere fornite per modellare l'analisi. Ad esempio, è possibile inserire informazioni sul sito di campionamento, una caratteristica che può essere associata alla loro vicinanza genetica, applicando il parametro *locprior* al modello di popolazione. Il parametro *locprior* è particolarmente informativo quando sono presenti deboli segnali di struttura di popolazione, una situazione che può derivare dall'utilizzo di un numero ridotto di marcatori, da piccole dimensioni del campione o da strette relazioni tra le popolazioni.

Per determinare l'origine degli individui, STRUCTURE permette di scegliere tra quattro modelli alternativi: il *no admixture model*, l'*admixture model*, il *linkage model* ed il *model with informative priors*. Se si possiedono conoscenze preliminari sull'origine delle popolazioni studiate o se vi è motivo di considerare ciascuna popolazione come completamente distinta, è appropriato utilizzare il *no admixture model*. Tuttavia, la mescolanza tra le popolazioni è un fenomeno molto comune e fa sì che un'ampia percentuale di individui campionati possa avere antenati recenti provenienti da più popolazioni. In questi casi è più appropriato utilizzare l'*admixture model*, secondo cui ciascun individuo può aver ereditato una frazione del proprio genoma da ognuna delle possibili popolazioni o specie, ammettendo la presenza di rimescolamento genico tra i diversi gruppi.

Sono disponibili inoltre due modelli di frequenza allelica. Il modello delle frequenze alleliche correlate presuppone un livello di non indipendenza, quindi è più conservativo. Il modello delle frequenze alleliche indipendenti richiede la conoscenza dei livelli di correlazione tra i gruppi e viene utilizzato quando si analizzano specie distinte.

Dal momento che l'obiettivo del programma è stabilire il numero di gruppi (K) più probabile in cui possono essere suddivisi gli individui, prima di procedere all'analisi è necessario fornire un intervallo entro il quale individuare il numero di gruppi (K). Generalmente l'intervallo parte da 2 e arriva ad un valore che corrisponde al numero di gruppi che ci si aspetterebbe di trovare a cui viene aggiunto 3. Per ogni valore di K

vengono fatte corse multiple indipendenti e ogni corsa produce un risultato diverso.

L'*output* generato da STRUCTURE consiste in una tabella in cui viene stimato il valore del logaritmo naturale della *likelihood* ( $\ln P(D)$ ) per ciascuna corsa. Questo valore è una misura della bontà del risultato e rappresenta la probabilità che gli individui vengano ripartiti in un certo numero  $K$  di gruppi caratteristico di ciascuna corsa. La corsa migliore, ovvero la suddivisione più probabile in  $K$  gruppi, si individua trovando il valore di  $\ln P(D)$  massimo. Il programma calcola inoltre le probabilità che ogni raggruppamento e ogni singolo individuo hanno di appartenere a ciascuno dei  $K$  gruppi individuati. Questo coefficiente di appartenenza ( $Q$ ) viene stimato per ciascun individuo e può essere visualizzato graficamente per ogni corsa. All'interno dei grafici ogni individuo è rappresentato da una singola linea verticale, che si divide in  $K$  segmenti colorati di lunghezza proporzionale al coefficiente di appartenenza stimato per ciascun gruppo  $K$ .

### **3.4.2 DNA mitocondriale**

#### **3.4.2.1 Statistiche di base e variabilità genetica**

In questa tesi è stata analizzata una sequenza costituita da 1047 paia di basi del gene mitocondriale ND2 in 231 individui (22 *L. kempfi*, 73 *L. squamifrons*, 97 *N. larseni* e 39 *N. nudifrons*).

Per descrivere la variabilità genetica tra popolazioni o specie è stato utilizzando il programma DNA *Sequence Polymorphism* (Rozas *et al.*, 2017) che ha permesso di calcolare i seguenti parametri:

- Numero di aplotipi ( $H_{ap}$ );
- Numero di siti polimorfici ( $S$ );
- Diversità aplotipica ( $H_d$ ), ovvero la probabilità che due aplotipi presi a caso nella popolazione siano diversi;



- Diversità nucleotidica ( $p$ ), ovvero la probabilità che due nucleotidi omologhi presi a caso siano differenti.

### **3.4.2.2 Differenziamento genetico inter ed intraspecifico**

Con il programma ARLEQUIN *ver.* 3.5.2.2 (Excoffier e Lischer, 2010) si è indagato il differenziamento genetico tra specie o popolazioni. Si sono calcolati gli indici  $F_{ST}$  e tramite l'analisi della varianza molecolare (AMOVA) si è verificato quanta della varianza totale delle frequenze alleliche calcolate è spiegata dalle differenze tra gruppi di popolazioni e quanta dalle differenze tra individui entro i gruppi.

Il programma POPART *ver.* 1.7.2 (Leigh e Bryant, 2015) è stato utilizzato per l'analisi genetica delle popolazioni o delle specie mediante un *network* di aplotipi. La sua funzione primaria è l'inferenza e la visualizzazione delle relazioni genetiche tra dati di sequenze genetiche. In questa tesi, le relazioni tra le diverse sequenze genetiche sono state dedotte mediante il metodo implementato nel programma TCS (Clement et al., 2002).

## 4. RISULTATI

### 4.1 Marcatori microsatellite

Il *dataset* di partenza era costituito da un numero di 287 individui (52 *L. kemp*; 101 *L. squamifrons*; 94 *N. larseni*; 40 *N. nudifrons*) e 21 *loci* microsatellite.

#### 4.1.1 Correzione dataset

Per l'identificazione dei *loci* monomorfici è stato utilizzato il pacchetto *divRsity* in RStudio ver. 4.0.5, utilizzando la funzione *divBasic*. Sono state calcolate le statistiche di base di tutti i 21 *loci* microsatellite per ciascuna specie (tabella 8.2 in appendice), al fine di eliminare dal *dataset* i *loci* con un numero di alleli osservati pari a 1. Il *locus* Ln42701 (tabella 4.1) presenta un numero di alleli osservati (A) pari a 1 in tutte e quattro le specie e di conseguenza è stato eliminato dal *dataset*.

Specie	Statistiche	Ln42701
<i>L. kemp</i>	N	52
	A	1
	%	100
	Ar	1
	Ho	0
	He	0
<i>L. squamifrons</i>	N	101
	A	1
	%	100
	Ar	1
	Ho	0
	He	0

**Tabella 4.1:** Statistiche di base per il *locus* microsatellite Ln42701 nelle quattro specie analizzate in questa tesi, prima della correzione del *dataset*. In giallo è evidenziato il numero effettivo di alleli (A) per il *locus* microsatellite Ln42701 in ognuna delle quattro specie.

Specie	Statistiche	Ln42701
<i>N. larseni</i>	N	94
	A	1
	%	100
	Ar	1
	Ho	0
	He	0
<i>N. nudifrons</i>	N	40
	A	1
	%	100
	Ar	1
	Ho	0
	He	0

**Tabella 4.1 (continuazione):** Statistiche di base per il *locus* microsatellite Ln42701 nelle quattro specie analizzate in questa tesi, prima della correzione del *dataset*. In giallo è evidenziato il numero effettivo di alleli (A) per il *locus* microsatellite Ln42701 in ognuna delle quattro specie.

All'interno del *dataset* un valore pari a 0, associato ad un allele, indica che non è stato possibile ottenere informazioni. Se tra i dati associati ad un *locus* o ad un individuo, all'interno di ciascuna specie, i valori pari a 0 sono più del 20%, quel *locus* o individuo va eliminato dall'analisi. Il *locus* Ln40551 nella specie *N. larseni* mostra il 21% di dati mancanti e il *locus* Ln22268 nella specie *N. nudifrons* mostra il 23% di dati mancanti (tabella 4.2). Di conseguenza entrambi i *loci* sono stati eliminati dal *dataset*.

Specie	Locus	Percentuale dati mancanti
<i>N. larseni</i>	Ln41281	0%
	Ln40551	21%
	Ln42701	0%
	Ln42233	0%
	Ln36100	12%
	Ln35217	0%
	Ln23194	0%
	Ln22268	0%

**Tabella 4.2:** Percentuale di dati mancanti per ciascun *locus* nelle specie *N. larseni* e *N. nudifrons*. I *loci* che presentano una percentuale di dati mancanti superiore al 20% sono evidenziati in giallo.

Specie	Locus	Percentuale dati mancanti
<i>N. larseni</i>	Ln36156	0%
	Ln45589	0%
	Ln45257	0%
	Ln42016	0%
	Ch126c	0%
	Ch3866	0%
	Ca35	0%
	Ch5817	0%
	Ch623	0%
	Ch1968	0%
	Ch3603	1%
	Ch2788	0%
	Ca48	1%
	<i>N. nudifrons</i>	Ln41281
Ln40551		0%
Ln42701		0%
Ln42233		0%
Ln36100		0%
Ln35217		0%
Ln23194		0%
Ln22268		23%
Ln36156		0%
Ln45589		0%
Ln45257		0%
Ln42016		0%
Ch126c		0%
Ch3866		0%
Ca35		3%
Ch5817		0%
Ch623		0%
Ch1968		0%
Ch3603		5%
Ch2788		3%
Ca48	0%	

**Tabella 4.2 (continuazione):** Percentuale di dati mancanti per ciascun *locus* nelle specie *N. larseni* e *N. nudifrons*. I loci che presentano una percentuale di dati mancanti superiore al 20% sono evidenziati in giallo.

La presenza di alleli nulli è stata verificata utilizzando il programma MICRO-CHECKER ver. 2.2.3 (Van Oosterhout *et al.*, 2004) per ogni *locus*, in ciascuna popolazione delle quattro specie analizzate in questa tesi (tabella 8.3 in appendice). Il *locus* Ln42016 mostra la presenza di alleli nulli in tutte e quattro le specie e il *locus* Ch5817 mostra la presenza di alleli nulli in tre specie su quattro (tabella 4.3). Di conseguenza entrambi i *loci* sono stati eliminati dal *dataset*.

Locus	<i>L. kempfi</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
Ln 41281				
Ln 40551			AN	
Ln 42701				
Ln 42233				
Ln 36100				
Ln 35217				
Ln 23194				
Ln 22268				
Ln 36156				
Ln 45589				
Ln 45257		AN		
Ln 42016	AN	AN	AN	AN
Ch 126				
Ch 3866				
Ca 35				
Ch 5817	AN	AN	AN	
Ch 623				
Ch 1968				
Ch 3603				
Ch 2788				
Ca 48				

**Tabella 4.3:** Presenza di alleli nulli per ciascun *locus* nelle quattro specie analizzate in questa tesi. Per ciascuna specie sono evidenziati in giallo i *loci* caratterizzati dalla presenza di alleli nulli (AN).

#### 4.1.2 Equilibrio di Hardy-Weinberg e *linkage disequilibrium*

Il programma GENEPOP ON THE WEB *ver. online 4.7* (Rousset, 2008) è stato utilizzato per verificare la presenza di equilibrio di Hardy-Weinberg per ogni *locus* in ciascuna specie (tabella 4.4). Il programma svolge un test statistico di verifica d'ipotesi (*probability test*) che testa l'ipotesi nulla  $H_0$ : "l'unione tra i gameti avviene in maniera casuale". Il calcolo della probabilità che le frequenze alleliche osservate siano in equilibrio di Hardy-Weinberg si basa sul metodo delle catene di Markov (Raymond e Rousset, 1995a; parametri: 10000 *dememorization steps*, 500 *batches* e 10000 iterazioni per ogni *batch*). Trattandosi di test multipli, i valori dei *p-value* ottenuti sono stati corretti con il programma MYRIADS *ver. 1.2* (Carvajal-Rodriguez, 2018) applicando la correzione di Benjamini-Hochberg (Benjamini e Hochberg, 1995). Al test è stato assegnato un livello di significatività ( $\alpha = 0.05$ ), che consiste nella probabilità massima di errore che si è disposti ad accettare nel respingere l'ipotesi nulla. Se il valore del *p-value* è inferiore a 0.05 significa che il *locus* è in disequilibrio. Se un *locus* è in disequilibrio in almeno tre specie su quattro deve essere eliminato dal *dataset*. I loci che risultano in disequilibrio di Hardy-Weinberg in almeno tre specie su quattro sono Ln42016 e Ch5817. Questi stessi *loci* erano già stati eliminati in precedenza dal *dataset* in quanto mostravano la presenza di alleli nulli. Questi risultati confermano quindi che la presenza di disequilibrio potrebbe essere dovuto ad errori commessi durante lo *scoring* (si veda la sezione 3.2.2).

Locus	<i>L. kempii</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
Ln41281	1	0.558	1	
Ln40551	0.43	0.62	< 0.001	0.293

**Tabella 4.4:** *Loci* che non rispettano l'equilibrio di Hardy-Weinberg secondo i test svolti con GENEPOP ON THE WEB per ogni locus nelle quattro specie analizzate in questa tesi. I valori riportati in tabella si riferiscono ai *p-value* ottenuti dopo la correzione di Benjamini-Hochberg per test multipli. I valori evidenziati in giallo risultano statisticamente significativi (*p-value* < 0.05) ed indicano i *loci* in disequilibrio di Hardy-Weinberg. Le celle vuote indicano l'assenza di dati, la presenza di un solo allele o la presenza di due alleli di cui uno presente in una sola copia.

Locus	<i>L. kempii</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
Ln42701				
Ln42233			1	
Ln36100			< 0.001	
Ln35217	0.346	0.818	0.871	0.55
Ln23194	1	1		0.124
Ln22268			0.112	
Ln36156	0.333	0.03	0.876	0.854
Ln45589	1	0.914	0.871	0.15
Ln45257	< 0.001	< 0.001	1	0.124
Ln42016	< 0.001	< 0.001	< 0.001	< 0.001
Ch126		1	1	
Ch3866	1	1	1	0.55
Ca35	1	0.012	0.736	0.55
Ch5817	< 0.001	< 0.001	< 0.001	0.549
Ch623	0.03		0.167	
Ch1968	0.03	1	< 0.001	0.082
Ch3603	0.03	1	0.005	
Ch2788	0.333	0.914	0.29	1
Ca48			< 0.001	0.549

**Tabella 4.4 (continuazione):** *Loci* che non rispettano l'equilibrio di Hardy-Weinberg secondo i test svolti con GENEPOP ON THE WEB per ogni locus nelle quattro specie analizzate in questa tesi. I valori riportati in tabella si riferiscono ai *p-value* ottenuti dopo la correzione di Benjamini-Hochberg per test multipli. I valori evidenziati in giallo risultano statisticamente significativi (*p-value* < 0.05) ed indicano i *loci* in disequilibrio di Hardy-Weinberg. Le celle vuote indicano l'assenza di dati, la presenza di un solo allele o la presenza di due alleli di cui uno presente in una sola copia.

Il programma GENEPOP ON THE WEB *ver. online 4.7* (Rousset, 2008) è stato utilizzato anche per calcolare la probabilità di *linkage disequilibrium* in ciascuna coppia di *loci* in ciascuna specie. Il programma svolge un test statistico di verifica d'ipotesi (*likelihood-ratio statistical significance test*) che testa l'ipotesi nulla  $H_0$ : "i genotipi di ciascun *locus* sono indipendenti dai genotipi degli altri *loci*". Il programma si avvale di tabelle di contingenza corrispondenti a tutte le possibili coppie di *loci* in ciascuna specie ed esegue il test di probabilità utilizzando le catene di Markov (Raymond e Rousset, 1995a; parametri: 10000 *dememorization*

*steps*, 500 *batches* e 10000 iterazioni per ogni *batch*) per stimare il valore esatto del *p-value*. Il valore soglia di significatività ( $\alpha$ ) del test è pari a 0.05, con correzione di Benjamini-Hochberg (Benjamini e Hochberg, 1995) per test multipli effettuata con il programma MYRIADS *ver.* 1.2 (Carvajal-Rodriguez, 2018). Nessuna coppia di *loci* risulta significativamente in disequilibrio in almeno tre specie su quattro (dati non riportati), di conseguenza non è necessario apportare ulteriori modifiche al *dataset*.

A questo punto, eliminando tutti i *loci* individuati nelle sezioni 4.1.1 e 4.1.2, si è ottenuto il *dataset* corretto. Gli *input file* sono stati creati utilizzando il software CREATE *ver.* 1.38 (Coombs *et al.*, 2008). Il *dataset* corretto finale è costituito da 16 *loci* microsatellite al posto dei 21 *loci* di partenza.

#### 4.1.3 Numero di alleli e frequenze alleliche

Tutte le statistiche di base sono state calcolate a partire dal *dataset* corretto (tabella 8.4 in appendice) mediante il pacchetto *diveRsity* in RStudio *ver.* 4.0.5, utilizzando la funzione *divBasic*.

Il numero di alleli osservati per ciascun *locus* in ciascuna specie, il numero medio di alleli per specie e il numero totale di alleli per *locus* sono riportati nella tabella 4.5.

Locus	<i>L. kempi</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>	Totale
Ln41281	3	2	2	1	4
Ln42233	2	1	2	1	3
Ln36100	1	2	3	1	4
Ln35217	6	7	9	3	10
Ln23194	2	2	1	2	3

**Tabella 4.5:** Numero effettivo di alleli (A) calcolati mediante il pacchetto *diveRsity* in RStudio *ver.* 4.0.5. In tabella sono riportati il numero di alleli osservati per i 16 *loci* microsatellite in ognuna delle quattro specie, il numero medio di alleli per specie e il totale di alleli per *locus*.



Locus	<i>L. kempfi</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>	Totale
Ln36156	19	27	25	23	32
Ln45589	13	17	5	5	20
Ln45257	2	4	2	2	6
Ch126	1	3	3	1	5
Ch3866	2	6	8	4	8
Ca35	16	19	24	9	26
Ch623	2	1	2	1	4
Ch1968	2	2	2	3	4
Ch3603	2	2	2	1	3
Ch2788	5	5	6	3	12
Ca48	1	1	41	14	44
Media	4.9	6.3	8.6	4.6	11.8

**Tabella 4.5 (continuazione):** Numero effettivo di alleli (A) calcolati mediante il pacchetto *diveRsity* in RStudio ver. 4.0.5. In tabella sono riportati il numero di alleli osservati per i 16 *loci* microsatellite in ognuna delle quattro specie, il numero medio di alleli per specie e il totale di alleli per *locus*.

Considerando tutte le specie insieme, il numero totale di alleli per *locus* varia da un minimo di 3 (*loci* Ln42233, Ln23194 e Ch3603) ad un massimo di 44 (*locus* Ca48).

Senza considerare i *loci* monomorfici, il numero effettivo di alleli per *locus*, entro le specie, varia da un minimo di 2 ad un massimo di 41 (*locus* Ca48 nella specie *N. larseni*).

Il numero medio di alleli per ciascuna specie, considerando tutti i 16 *loci* microsatellite, può essere considerato come un'indicazione della variabilità genetica e corrisponde a 4.9 in *L. kempfi*, 6.3 in *L. squamifrons*, 8.6 in *N. larseni* e 4.6 in *N. nudifrons*. Il numero medio totale di alleli, per tutte le specie e tutti i *loci*, corrisponde a 11.8.

Tutte le frequenze alleliche per i 16 *loci* microsatellite (tabelle da 8.5 a 8.20 in appendice) sono state calcolate tramite il programma GENEPOP ON THE WEB ver. *online* 4.7 (Rousset, 2008) in tutte e quattro le specie analizzate in questa tesi.

#### 4.1.4 Allelic richness

L'*allelic richness* (tabella 4.6) è stata calcolata, a partire dal *dataset* corretto, mediante la funzione *divBasic* del pacchetto *diveRsity* in RStudio ver. 4.0.5, utilizzando 1000 ricampionamenti.

Locus	<i>L. kempfi</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
Ln41281	2.52	2	1.99	1
Ln42233	1.54	1	1.89	1
Ln36100	1	1.33	2.58	1
Ln35217	4.85	5.27	6.44	3
Ln23194	1.92	1.57	1	1.88
Ln36156	16.52	20.85	18.33	19.67
Ln45589	10.64	10.2	4.36	4.53
Ln45257	1.78	2.57	2	1.88
Ch126	1	2.13	2.15	1
Ch3866	1.81	3.7	5.48	3.64
Ca35	13.79	14.59	19.11	8.12
Ch623	1.54	1	2	1
Ch1968	1.54	1.86	2	2.3
Ch3603	1.53	1.69	2	1
Ch2788	4.01	3.94	5.16	2.95
Ca48	1	1	26.58	12.14
Media	4.19	4.67	6.44	4.13

**Tabella 4.6:** *Allelic richness* ( $A_r$ ) calcolata mediante il pacchetto *diveRsity* in RStudio ver. 4.0.5 per ogni *locus* e valori medi per ciascuna specie analizzata in questa tesi. In giallo sono evidenziati i *loci* monomorfici.

Per le analisi dei valori di *allelic richness* di ciascuna specie non sono stati considerati i *loci* monomorfici.

In *L. kempfi* i valori di *allelic richness* per *locus* variano da un minimo di 1.53 (*locus* Ch3603) a un massimo di 16.52 (*locus* Ln36156), mentre il valore medio è di 4.19.

In *L. squamifrons* i valori di *allelic richness* per *locus* variano da un minimo di 1.33 (*locus* Ln36100) a un massimo di 20.85 (*locus* Ln36156), mentre il valore medio è di 4.67.

In *N. larseni* i valori di *allelic richness* per *locus* variano da un minimo di 1.89 (*locus* Ln42233) a un massimo di 26.58 (*locus* Ca48), mentre il valore medio è di 6.44.

In *N. nudifrons* i valori di *allelic richness* per *locus* variano da un minimo di 1.88 (*loci* Ln23194 e Ln45257) a un massimo di 19.67 (*locus* Ln36156), mentre il valore medio è di 4.13.

#### 4.1.5 Eterozigosità osservata ed attesa

I valori di eterozigosità osservata ( $H_o$ ) ed attesa ( $H_e$ ) sono stati calcolati, a partire dal *dataset* corretto, mediante la funzione *divBasic* del pacchetto *diveRsity* in RStudio ver. 4.0.5. Per ciascuna specie non sono stati considerati i *loci* monomorfici.

<i>Locus</i>	$H_o$ ; $H_e$	<i>L. kempfi</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
Ln41281	$H_o$	0.38	0.22	0.12	0
	$H_e$	0.38	0.25	0.11	0
Ln42233	$H_o$	0.02	0	0.05	0
	$H_e$	0.02	0	0.05	0
Ln36100	$H_o$	0	0.01	0.06	0
	$H_e$	0	0.01	0.26	0
Ln35217	$H_o$	0.6	0.59	0.66	0.4
	$H_e$	0.63	0.53	0.66	0.45
Ln23194	$H_o$	0.06	0.02	0	0.03
	$H_e$	0.06	0.02	0	0.07
Ln36156	$H_o$	0.94	0.86	0.87	0.95
	$H_e$	0.91	0.93	0.9	0.93
Ln45589	$H_o$	0.85	0.65	0.54	0.65
	$H_e$	0.79	0.7	0.59	0.62
Ln45257	$H_o$	0.03	0.01	0.19	0.03
	$H_e$	0.08	0.15	0.19	0.07

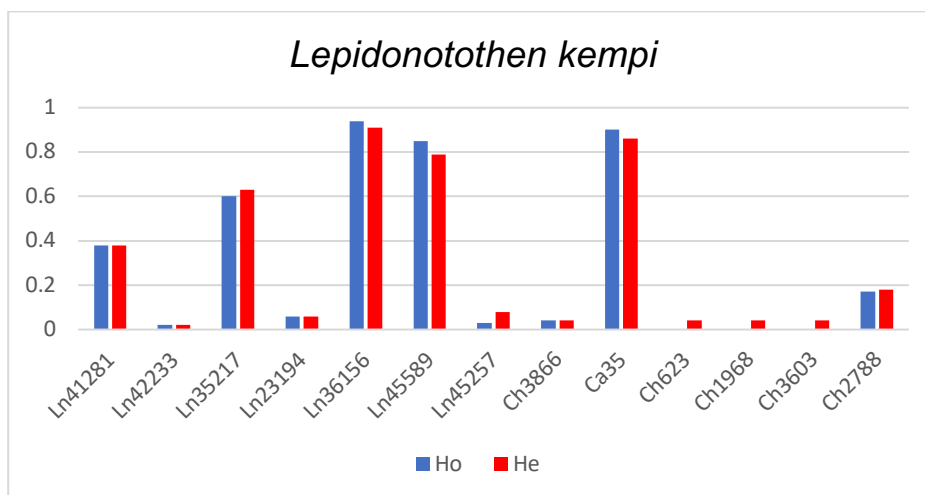
**Tabella 4.7:** Eterozigosità osservata ( $H_o$ ) ed attesa ( $H_e$ ) calcolate mediante il pacchetto *diveRsity* in RStudio ver. 4.0.5 per ogni *locus* e la media per ciascuna specie. In giallo sono evidenziati i *loci* monomorfici.

<i>Locus</i>	$H_o; H_e$	<i>L. kempfi</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
Ch126	$H_o$	0	0.04	0.05	0
	$H_e$	0	0.04	0.05	0
Ch3866	$H_o$	0.04	0.12	0.65	0.72
	$H_e$	0.04	0.12	0.63	0.66
Ca35	$H_o$	0.9	0.89	0.87	0.74
	$H_e$	0.86	0.89	0.92	0.8
Ch623	$H_o$	0	0	0.38	0
	$H_e$	0.04	0	0.47	0
Ch1968	$H_o$	0	0.05	1	0.03
	$H_e$	0.04	0.05	0.5	0.07
Ch3603	$H_o$	0	0.03	0.23	0
	$H_e$	0.04	0.03	0.34	0
Ch2788	$H_o$	0.17	0.23	0.61	0.44
	$H_e$	0.18	0.23	0.68	0.41
Ca48	$H_o$	0	0	0.8	0.75
	$H_e$	0	0	0.94	0.82
Media	$H_o$	0.25	0.23	0.44	0.3
	$H_e$	0.25	0.25	0.46	0.31

**Tabella 4.7 (continuazione):** Eterozigosità osservata ( $H_o$ ) ed attesa ( $H_e$ ) calcolate mediante il pacchetto *diveRsity* in RStudio ver. 4.0.5 per ogni *locus* e la media per ciascuna specie. In giallo sono evidenziati i *loci* monomorfici.

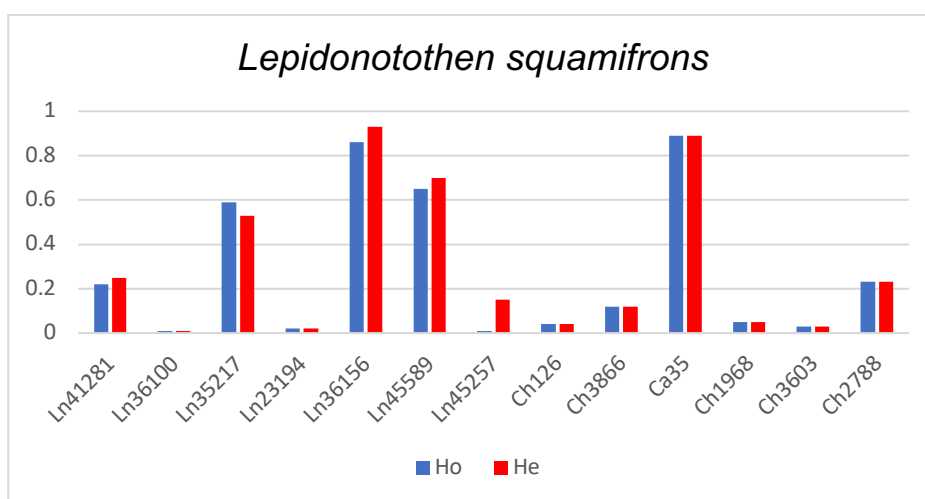
I valori di eterozigosità osservata ( $H_o$ ) ed attesa ( $H_e$ ), in ciascuna specie analizzata, sono simili in quasi tutti i *loci* (tabella 4.7).

In *L. kempfi* (grafico 4.1) il valore minimo di eterozigosità osservata ( $H_o$ ) è stato trovato per i *loci* Ch623, Ch1968 e Ch3603 ( $H_o=0$ ), mentre il valore massimo è stato trovato per il *locus* Ln36156 ( $H_o=0.94$ ). Lo scostamento massimo tra eterozigosità osservata ( $H_o$ ) ed attesa ( $H_e$ ) è stato trovato per il *locus* Ln45589 ( $H_o=0.85$  e  $H_e=0.79$ ; scostamento=0.06).



**Grafico 4.1:** Eterozigosità osservata ( $H_o$ ) ed attesa ( $H_e$ ) per la specie *L. kempii*.

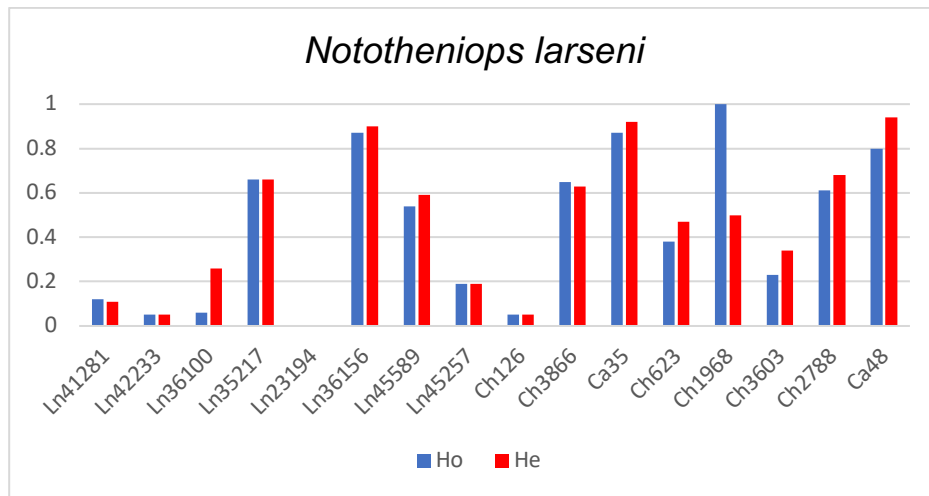
In *L. squamifrons* (grafico 4.2) il valore minimo di eterozigosità osservata ( $H_o$ ) è stato trovato per i *loci* Ln36100 e Ln45257 ( $H_o=0.01$ ), mentre il valore massimo è stato trovato per il *locus* Ca35 ( $H_o=0.89$ ). Lo scostamento massimo tra eterozigosità osservata ( $H_o$ ) ed attesa ( $H_e$ ) è stato trovato per il *locus* Ln45257 ( $H_o=0.01$  e  $H_e=0.15$ ; scostamento=0.14).



**Grafico 4.2:** Eterozigosità osservata ( $H_o$ ) ed attesa ( $H_e$ ) per la specie *L. squamifrons*.

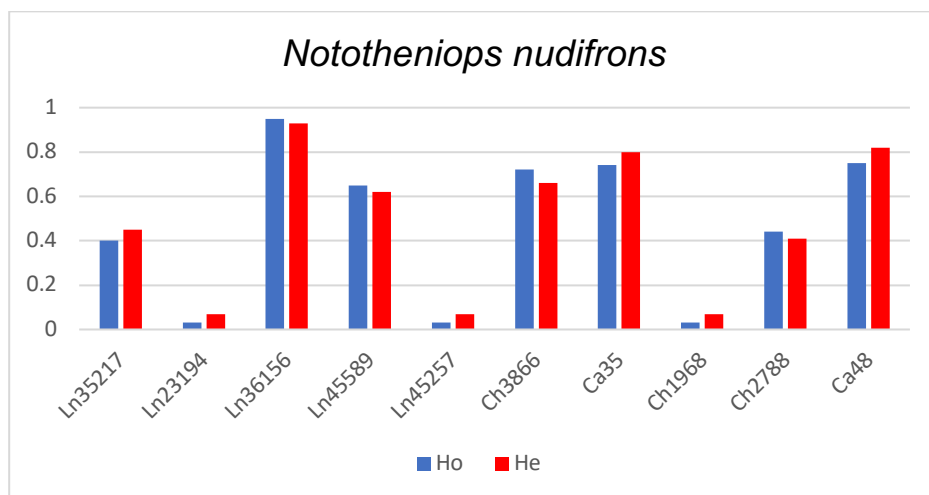
In *N. larseni* (grafico 4.3) il valore minimo di eterozigosità osservata ( $H_o$ ) è stato trovato per i *loci* Ln42233 e Ch126 ( $H_o=0.05$ ), mentre il valore massimo è stato trovato per il *locus* Ch1968 ( $H_o=1$ ). Lo scostamento

massimo tra eterozigotità osservata ( $H_o$ ) ed attesa ( $H_e$ ) è stato trovato per il *locus* Ch1968 ( $H_o=1$  e  $H_e=0.50$ ; scostamento=0.50).



**Grafico 4.3:** Eterozigotità osservata ( $H_o$ ) ed attesa ( $H_e$ ) per la specie *N. larseni*.

In *N. nudifrons* (grafico 4.4) il valore minimo di eterozigotità osservata ( $H_o$ ) è stato trovato per i *loci* Ln23194, Ln45257 e Ch1968 ( $H_o=0.03$ ), mentre il valore massimo è stato trovato per il *locus* Ln36156 ( $H_o=0.95$ ). Lo scostamento massimo tra eterozigotità osservata ( $H_o$ ) ed attesa ( $H_e$ ) è stato trovato per il *locus* Ca48 ( $H_o=0.75$  e  $H_e=0.82$ ; scostamento=0.07).



**Grafico 4.4:** Eterozigotità osservata ( $H_o$ ) ed attesa ( $H_e$ ) per la specie *N. nudifrons*.

Per quanto riguarda i valori medi, l'eterozigotità osservata minima è stata trovata in *L. squamifrons* (0.23), mentre quella massima è stata trovata

in *N. larseni* (0.44). L'eterozigotità attesa minima, invece, è stata trovata in *L. kempfi* e in *L. squamifrons* (0.25), mentre quella massima è stata trovata in *N. larseni* (0.46). I valori medi di eterozigotità osservata ( $H_o$ ) ed attesa ( $H_e$ ), nelle quattro specie analizzate, sono molto simili e lo scostamento massimo è di 0.02 in *L. squamifrons* e *N. larseni*.

#### 4.1.6 Differenziamento genetico inter ed intraspecifico

La struttura di popolazione e di specie è stata analizzata con il programma ARLEQUIN ver. 3.5.2.2 (Excoffier e Lischer, 2010). Il programma permette di quantificare l'indice  $F_{ST}$  a coppie. La significatività dell'indice  $F_{ST}$  è stata testata attraverso un approccio di permutazioni non parametriche, con un numero di permutazioni pari a 1000 e una soglia di significatività pari a 0.05.

Tutti i valori di differenziamento genetico tra le quattro specie analizzate in questa tesi, stimati mediante l'indice  $F_{ST}$ , sono statisticamente significativi (tabella 4.8). È presente un basso livello di differenziamento tra le specie *L. kempfi* e *L. squamifrons* ( $F_{ST}=0.013$ ,  $p\text{-value} < 0.001$ ). I livelli di differenziamento tra tutte le altre coppie di specie sono più elevati, nonostante si possa vedere che *N. larseni* e *N. nudifrons* siano più simili tra loro ( $F_{ST}=0.333$ ,  $p\text{-value} < 0.001$ ) di quanto non siano con *L. kempfi* e *L. squamifrons* ( $F_{ST}$  compreso tra 0.611 e 0.499). Le specie che risultano geneticamente più distanti sono *N. nudifrons* e *L. squamifrons* ( $F_{ST} = 0.611$ ,  $p\text{-value} < 0.001$ ).

	<i>Lepidonotothen kempfi</i>	<i>Lepidonotothen squamifrons</i>	<i>Nototheniops larseni</i>	<i>Nototheniops nudifrons</i>
<i>Lepidonotothen kempfi</i>	*	< 0.001	< 0.001	< 0.001
<i>Lepidonotothen squamifrons</i>	0.013	*	< 0.001	< 0.001
<i>Nototheniops larseni</i>	0.499	0.533	*	< 0.001
<i>Nototheniops nudifrons</i>	0.591	0.611	0.333	*

**Tabella 4.8:**  $F_{ST}$  a coppie tra le quattro specie ipotetiche *L. kempfi*, *L. squamifrons*, *N. larseni* e *N. nudifrons* basati sulla variabilità dei 16 loci microsatellite. I valori di  $F_{ST}$  sono riportati sotto la

diagonale, mentre sopra la diagonale sono riportati i valori del *p-value* ( $\alpha = 0.05$ ). I valori significativi sono evidenziati in grassetto.

Sono stati poi condotti confronti tra le diverse popolazioni all'interno di ciascuna specie, con l'intento di individuare eventuali differenze genetiche legate alla distanza geografica.

In *N. larseni* si possono distinguere tre popolazioni in base al sito di campionamento (South Shetland, Drake Passage e Bouvet Island). Il differenziamento genetico tra le tre popolazioni, stimato mediante l'indice  $F_{ST}$  (tabella 4.9), indica che la popolazione delle isole South Shetland non è differenziata significativamente dalla popolazione del Passaggio di Drake. Le popolazioni dell'isola di Bouvet sono differenziate significativamente dalle popolazioni del Passaggio di Drake e delle isole South Shetland, mostrando un livello di differenziamento molto basso ( $F_{ST}=0.016$ ,  $p\text{-value}<0.05$ ).

	South Shetland	Drake Passage	Bouvet Island
South Shetland	*	0.968	0.002
Drake Passage	0	*	0.007
Bouvet Island	<b>0.016</b>	<b>0.016</b>	*

**Tabella 4.9:**  $F_{ST}$  a coppie tra le tre popolazioni di *N. larseni*, basati sulla variabilità dei loci microsatellite. I valori di  $F_{ST}$  sono riportati sotto la diagonale, mentre sopra la diagonale sono riportati i valori del *p-value* ( $\alpha = 0.05$ ). I risultati significativi sono evidenziati in grassetto. I valori di  $F_{ST}$  negativi sono stati cambiati con 0.

In *N. nudifrons* si possono distinguere cinque popolazioni in base al sito di campionamento (Drake Passage, South Shetland, Elephant Island, South Orkney e South Georgia). Il differenziamento genetico tra le tre popolazioni, stimato mediante l'indice  $F_{ST}$ , non indica differenziamento genetico significativo tra le popolazioni in esame (valori non riportati).

Unendo i *dataset* di *L. kemp* e *L. squamifrons*, si possono distinguere cinque popolazioni in base al sito di campionamento (Shag Rocks, Bransfield Strait, Bouvet Island, South Shetland e Neumayer Station). Il



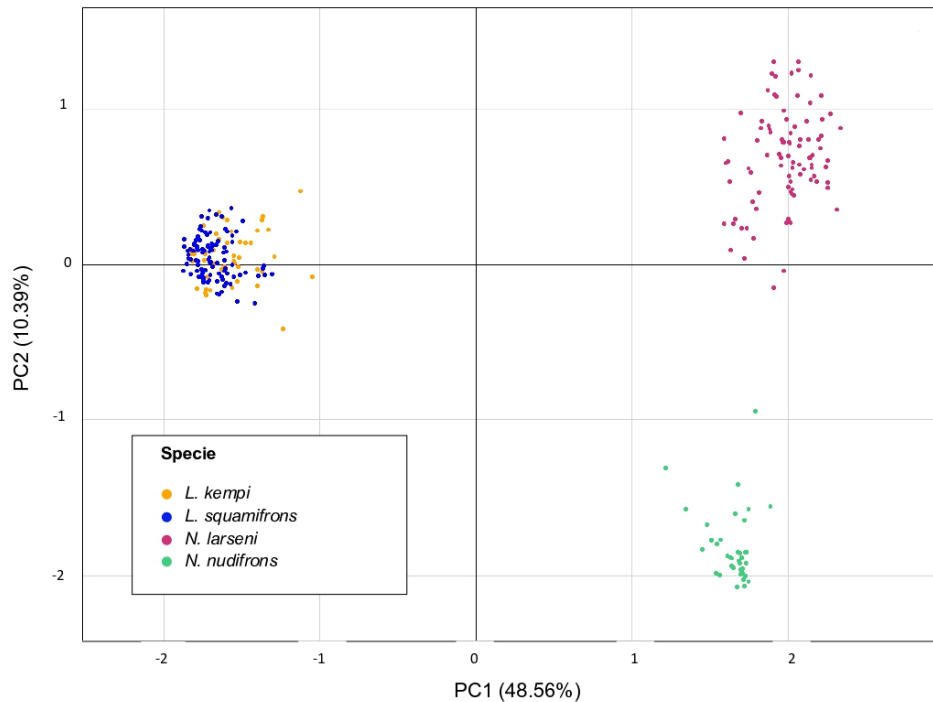
differenziamento genetico tra le cinque popolazioni, stimato mediante l'indice  $F_{ST}$  (tabella 4.10), indica che solo la popolazione delle isole di Shag Rocks è differenziata significativamente dalla popolazione dello Stretto di Bransfield ( $F_{ST}=0.048$ ,  $p\text{-value}<0.001$ ), dalla popolazione dell'isola di Bouvet ( $F_{ST}=0.044$ ,  $p\text{-value}<0.001$ ) e dalla popolazione delle isole South Shetland ( $F_{ST}=0.057$ ,  $p\text{-value}<0.001$ ). Tutte le altre popolazioni non sono differenziate significativamente.

	Shag Rocks	Bransfield Strait	Bouvet Island	South Shetland	Neumayer Station
Shag Rocks	*	< 0.001	< 0.001	< 0.001	0.999
Bransfield Strait	<b>0.048</b>	*	0.651	0.628	0.999
Bouvet Island	<b>0.044</b>	0.753	*	0.104	0.999
South Shetland	<b>0.057</b>	0.269	0.538	*	0.999
Neumayer Station	0	0	0	0	*

**Tabella 4.10:**  $F_{ST}$  a coppie tra le cinque popolazioni di *L. kempfi* e *L. squamifrons*, basati sulla variabilità dei *loci* microsatellite. I valori di  $F_{ST}$  sono riportati sotto la diagonale, mentre sopra la diagonale sono riportati i valori del  $p\text{-value}$  ( $\alpha = 0.05$ ). I risultati significativi sono evidenziati in grassetto. I valori di  $F_{ST}$  negativi sono stati cambiati con 0.

#### 4.1.7 Identificazione di gruppi geneticamente distinti

L'analisi delle componenti principali (PCA) è stata svolta mediante l'utilizzo del pacchetto *adegenet* in RStudio ver. 2.1.4, che permette di individuare i due assi che catturano maggiormente la variabilità dei dati. In base ai valori delle frequenze alleliche, è stato possibile fornire una rappresentazione grafica a due dimensioni in cui si possono visualizzare le componenti principali della variabilità genetica tra tutti gli individui analizzati (figura 4.1).



**Figura 4.1:** Analisi delle componenti principali (PCA) svolta a partire dal *dataset* relativo ai 16 *loci* microsatellite. L'analisi è stata svolta in R tramite i pacchetti *adegenet* ver. 2.1.4 per il calcolo e *ggplot2* ver. 4.0.5 per la visualizzazione grafica. Ogni punto rappresenta un individuo appartenente ad una delle quattro specie analizzate in questa tesi (giallo: *L. kempfi*, rosa: *N. larseni*, verde: *N. nudifrons* e blu: *L. squamifrons*). L'asse X (PC1) rappresenta il 47.72% della variabilità totale. L'asse Y (PC2) rappresenta il 9.92% della variabilità totale.

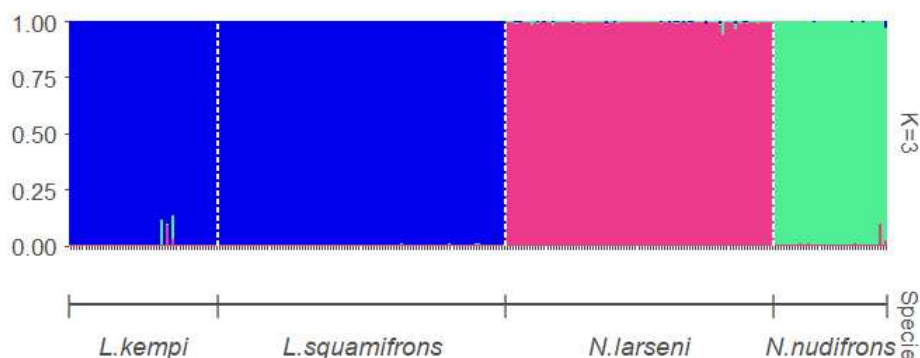
L'analisi della prima componente principale spiega il 48.56% della varianza totale, mentre la seconda spiega il 10.39% della varianza totale. Dalla figura 4.1 è possibile notare una chiara suddivisione in tre gruppi distinti, corrispondenti alle specie *N. larseni*, *N. nudifrons* e *L. kempfi* insieme a *L. squamifrons*.

La stessa analisi è stata svolta suddividendo gli individui in base al sito di campionamento (figura 8.1 in appendice). È stato possibile identificare tre raggruppamenti principali, all'interno dei quali non vi è una netta distinzione e differenziamento tra individui provenienti da aree geografiche differenti.

L'analisi della struttura di popolazione o di specie, effettuata secondo il metodo bayesiano con il programma STRUCTURE ver. 2.3.4 (Pritchard *et al.*, 2000), ha permesso di testare quale fosse il numero più probabile di raggruppamenti in cui è possibile suddividere il *dataset*.

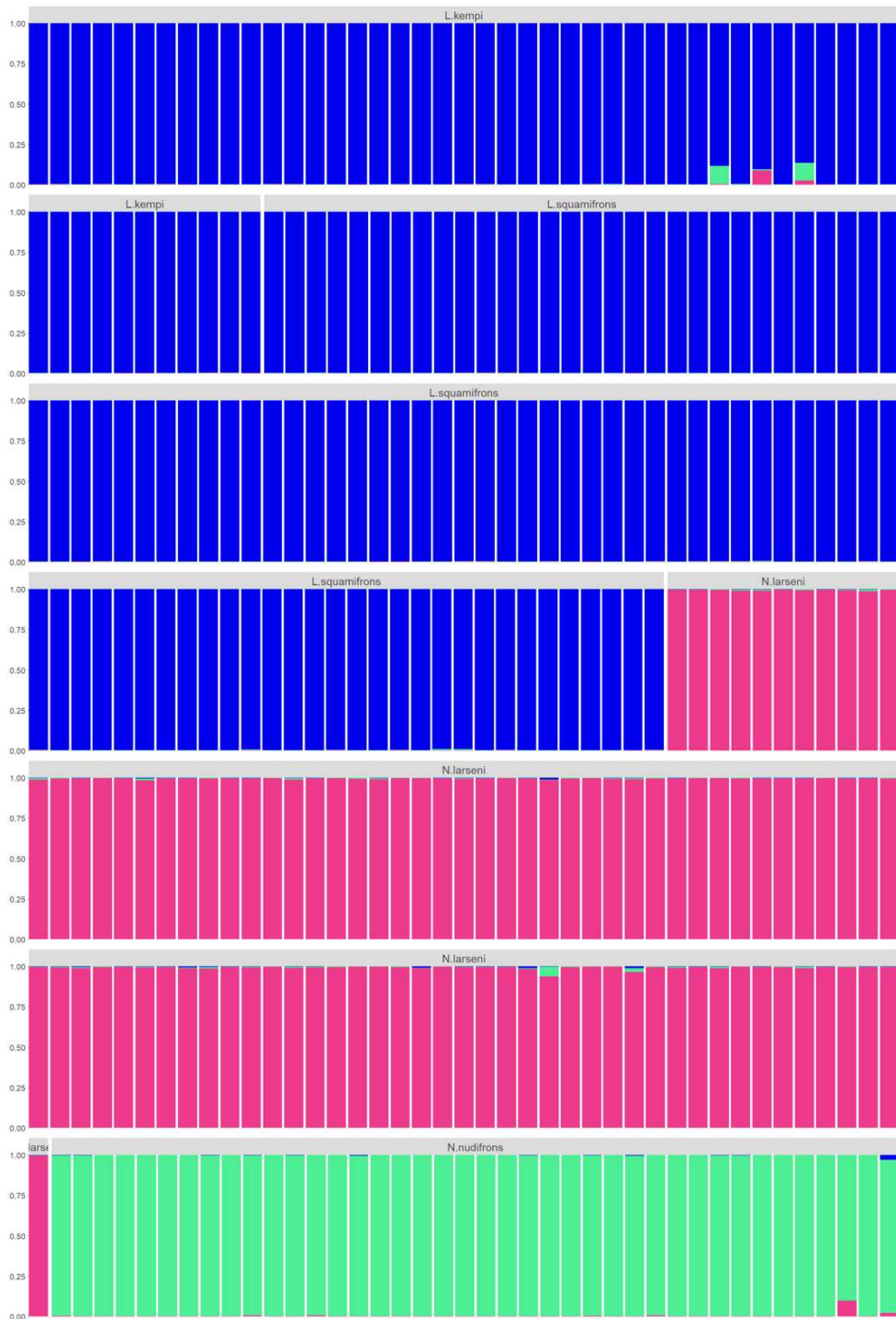
A partire dal *dataset* completo, si è voluto capire se si era in presenza di quattro specie distinte (*L. kempfi*, *L. squamifrons*, *N. larseni* e *N. nudifrons*) oppure se le specie *L. kempfi* e *L. squamifrons* potessero essere considerate un'unica specie. Il raggruppamento in quattro specie distinte era stato definito a priori basandosi sull'assegnazione ad una delle quattro specie, avvenuta al momento del campionamento, sulla base di caratteristiche morfologiche.

La prima analisi è stata eseguita utilizzando i 16 *loci* microsatelliti presenti nel *dataset* corretto, composto da 287 campioni, analizzando congiuntamente tutte e quattro le specie. Per ciascun valore di K, compreso fra 2 e 7, sono state impostate 10 corse. Per ciascuna corsa è stato fissato un valore di 10000 per il *burnin* e di 100000 per le corse vere e proprie. È stato utilizzato l'*admixture model* e frequenze alleliche indipendenti, in quanto si tratta di specie differenti. Non è stato impostato il parametro *locprior*. Per ogni corsa è stato stimato il valore del logaritmo naturale della *likelihood* ( $\ln P(D)$ ). I risultati di tale analisi indicano come situazione più probabile quella che prevede la suddivisione del *dataset* in 3 gruppi; infatti, il valore massimo di  $\ln P(D)$  corrisponde ad una corsa con K pari a 3 (figura 4.2).



(a)

**Figura 4.2:** Ripartizione delle probabilità (Q) che ogni individuo ha di appartenere ad uno dei tre gruppi individuati da STRUCTURE, senza *locprior*. Raffigurazione ottenuta in R tramite il pacchetto *pophelper* ver. 2.3.1. a) in ascissa è indicata la suddivisione a priori nelle quattro specie (*L. kempfi*, *L. squamifrons*, *N. larseni* e *N. nudifrons*), in ordinata la probabilità di appartenenza al singolo gruppo (K). b) struttura mostrata nel dettaglio. Ogni individuo è rappresentato da un segmento verticale, partizionato in diversi colori. L'altezza di ciascun segmento colorato indica la probabilità di appartenenza dell'individuo ai diversi gruppi (con K=3).

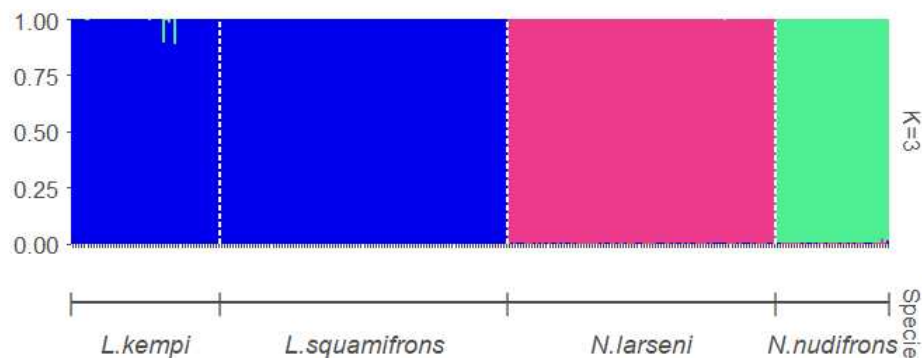


(b)

**Figura 4.2 (continuazione):** Ripartizione delle probabilità (Q) che ogni individuo ha di appartenere ad uno dei tre gruppi individuati da STRUCTURE, senza *locprior*. Raffigurazione ottenuta in R tramite il pacchetto *pophelper* ver. 2.3.1. a) in ascissa è indicata la suddivisione a priori nelle quattro specie (*L. kempi*, *L. squamifrons*, *N. larseni* e *N. nudifrons*), in ordinata la probabilità di appartenenza al singolo gruppo (K). b) struttura mostrata nel dettaglio. Ogni individuo è rappresentato da un segmento verticale, partizionato in diversi colori. L'altezza di ciascun segmento colorato indica la probabilità di appartenenza dell'individuo ai diversi gruppi (con K=3).

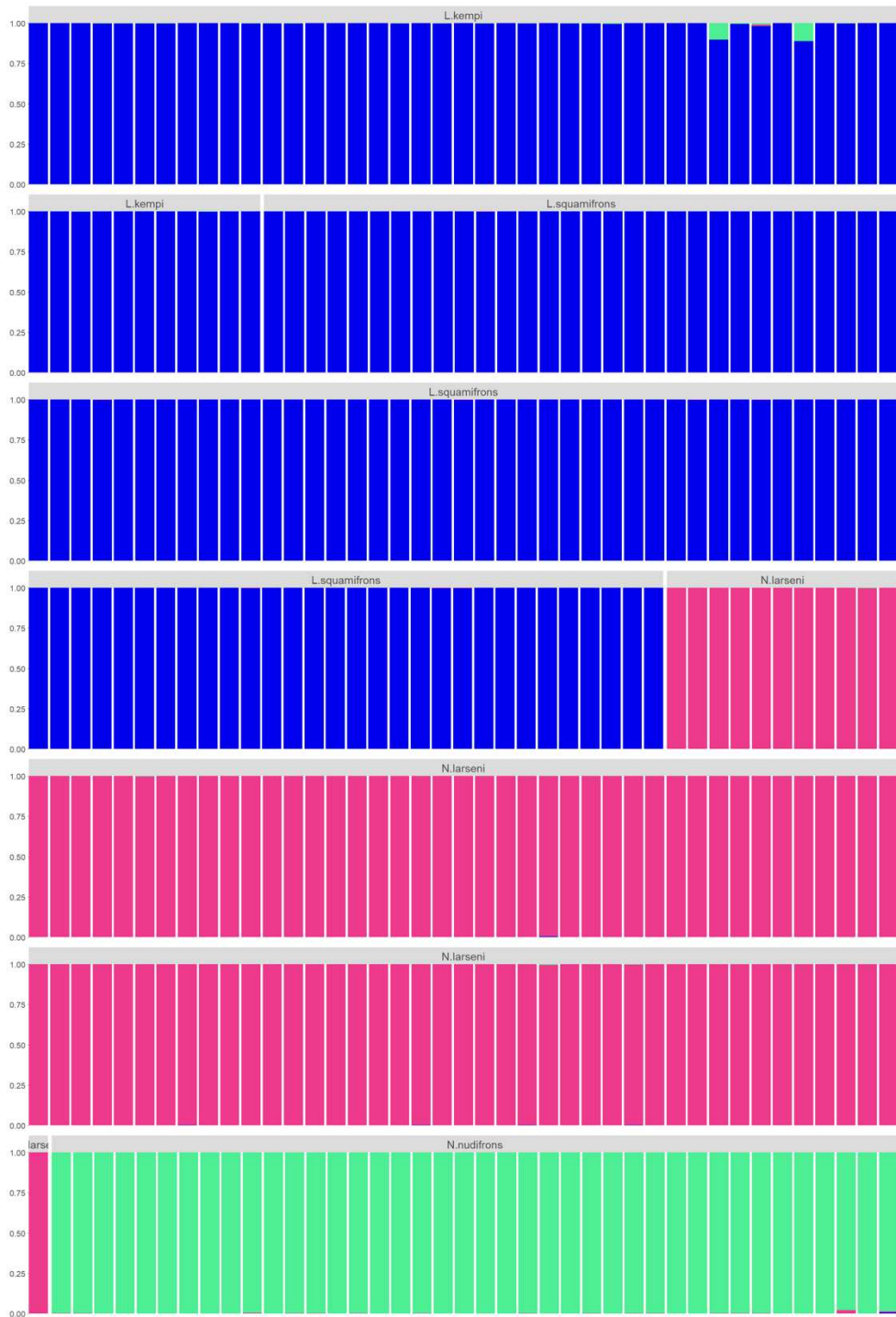
Gli individui appartenenti a *L. kempfi* e *L. squamifrons* risultano tutti appartenenti ad uno stesso gruppo. Si nota inoltre una netta separazione tra *N. larseni* e *N. nudifrons*.

La seconda analisi è stata svolta con le stesse impostazioni della precedente, con l'unica differenza che è stato impostato il parametro *locprior*, utilizzando come *prior* l'assegnazione di ciascun individuo ad una specie sulla base della morfologia. Come prima, i risultati indicano come situazione più probabile quella che prevede la suddivisione del *dataset* in 3 gruppi; infatti, il valore massimo di  $\ln P(D)$  corrisponde ad una corsa con  $K$  pari a 3 (figura 4.3).



(a)

**Figura 4.3:** Ripartizione delle probabilità ( $Q$ ) che ogni individuo ha di appartenere ad uno dei tre gruppi individuati da STRUCTURE, con *locprior*. Raffigurazione ottenuta in R tramite il pacchetto *pophelper* ver. 2.3.1. a) in ascissa è indicata la suddivisione a priori nelle quattro specie (*L. kempfi*, *L. squamifrons*, *N. larseni* e *N. nudifrons*), in ordinata la probabilità di appartenenza al singolo gruppo ( $K$ ). b) struttura mostrata nel dettaglio. Ogni individuo è rappresentato da un segmento verticale, partizionato in diversi colori. L'altezza di ciascun segmento colorato indica la probabilità di appartenenza dell'individuo ai diversi gruppi (con  $K=3$ ).



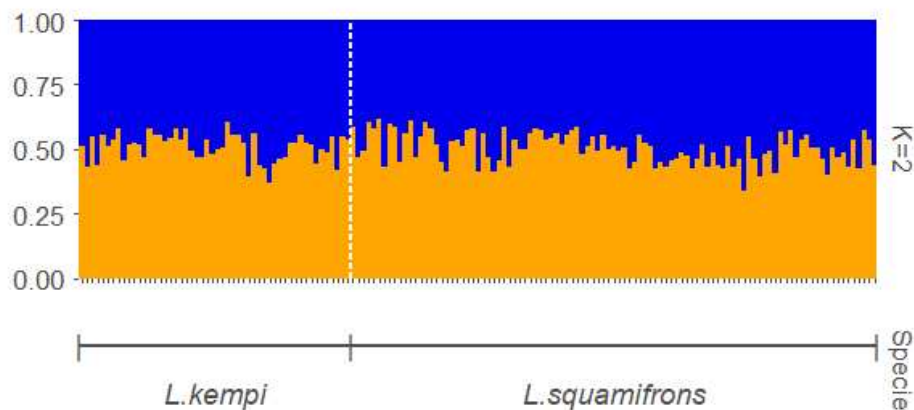
(b)

**Figura 4.3 (continuazione):** Ripartizione delle probabilità ( $Q$ ) che ogni individuo ha di appartenere ad uno dei tre gruppi individuati da STRUCTURE, con *locprior*. Raffigurazione ottenuta in R tramite il pacchetto *pophelper* ver. 2.3.1. a) in ascissa è indicata la suddivisione a priori nelle quattro specie (*L. kempi*, *L. squamifrons*, *N. larseni* e *N. nudifrons*), in ordinata la probabilità di appartenenza al singolo gruppo ( $K$ ). b) struttura mostrata nel dettaglio. Ogni individuo è rappresentato da un segmento verticale, partizionato in diversi colori. L'altezza di ciascun segmento colorato indica la probabilità di appartenenza dell'individuo ai diversi gruppi (con  $K=3$ ).

Un'ulteriore verifica è stata fatta a partire dal *dataset* relativo alle sole specie *L. kempi* e *L. squamifrons*, basandosi sul raggruppamento a priori in due specie distinte.

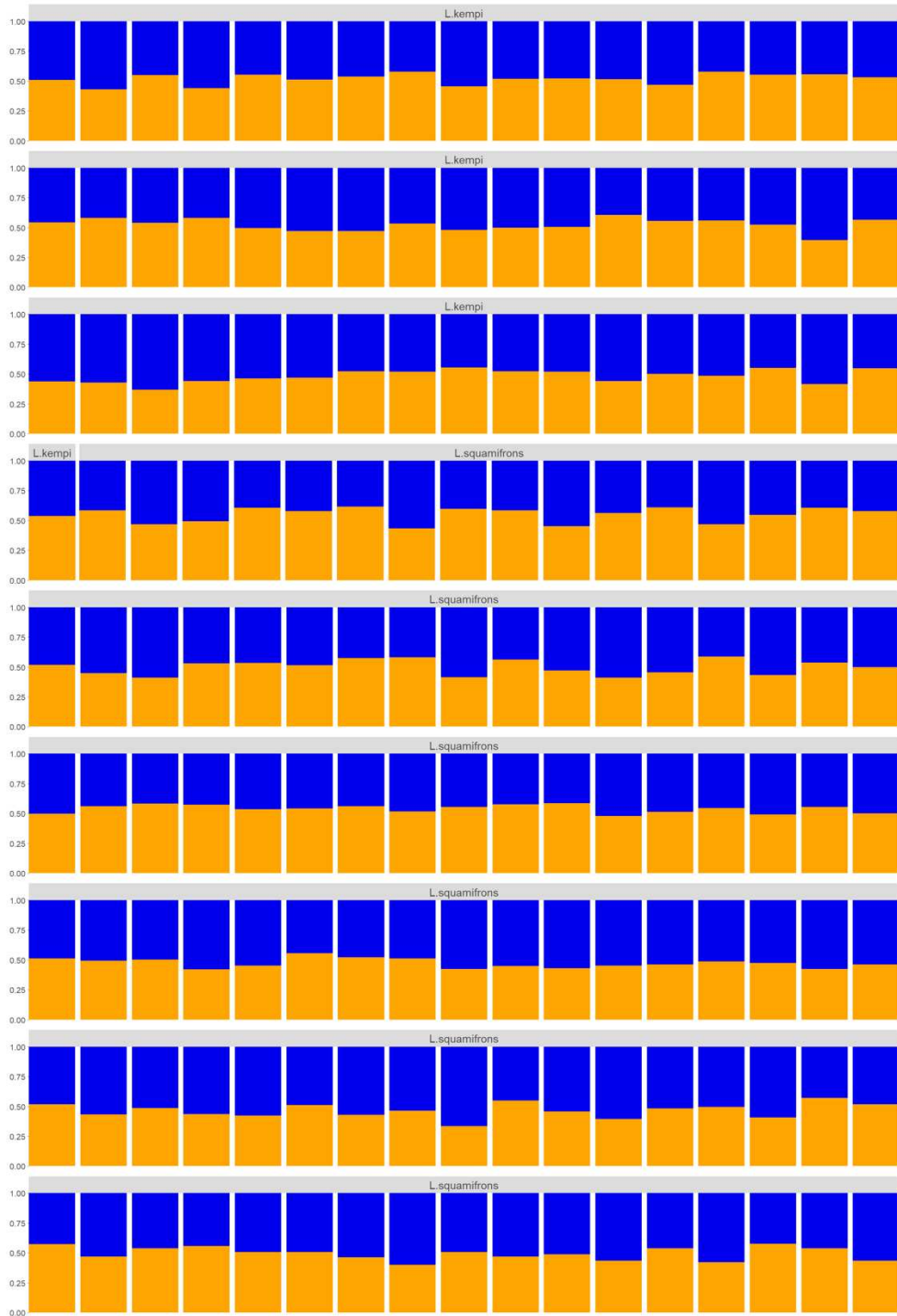
La terza analisi è stata eseguita utilizzando i 16 *loci* microsatellite presenti nel *dataset* corretto di *L. kempi* e *L. squamifrons*, composto da 153 campioni. Per ciascun valore di K, compreso fra 2 e 5, sono state impostate 10 corse. Per ciascuna corsa è stato fissato un valore di 10000 per il burnin e di 100000 per le corse vere e proprie. È stato utilizzato l'*admixture model* e frequenze alleliche correlate. Non è stato impostato il parametro *locprior*. Per ogni corsa è stato stimato il valore del logaritmo naturale della *likelihood* ( $\ln P(D)$ ). I risultati di tale analisi indicano come situazione più probabile quella che prevede la suddivisione del *dataset* in 2 gruppi; infatti, il valore massimo di  $\ln P(D)$  corrisponde ad una corsa con K pari a 2 (figura 4.4).

Nessun individuo mostra una probabilità di appartenenza maggiore ad uno dei due gruppi.



(a)

**Figura 4.4:** Ripartizione delle probabilità (Q) che ogni individuo ha di appartenere ad uno dei due gruppi individuati da STRUCTURE, senza *locprior*. Rappresentazione ottenuta in R tramite il pacchetto *pophelper* ver. 2.3.1. a) in ascissa è indicata la suddivisione a priori nelle due specie (*L. kempi* e *L. squamifrons*), in ordinata la probabilità di appartenenza al singolo gruppo (K). b) struttura mostrata nel dettaglio. Ogni individuo è rappresentato da un segmento verticale, partizionato in due colori. L'altezza di ciascun segmento colorato indica la probabilità di appartenenza dell'individuo ai due gruppi.



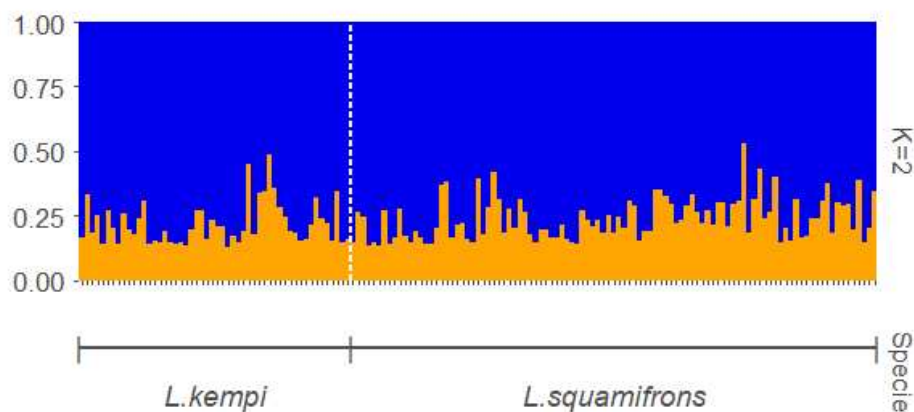
(b)

**Figura 4.4 (continuazione):** Ripartizione delle probabilità (Q) che ogni individuo ha di appartenere ad uno dei due gruppi individuati da STRUCTURE, senza *locprior*. Rappresentazione ottenuta in R tramite il pacchetto *pophelper* ver. 2.3.1. a) in ascissa è indicata la suddivisione a priori nelle due specie (*L. kempi* e *L. squamifrons*), in ordinata la probabilità di appartenenza al singolo gruppo (K). b) struttura mostrata nel dettaglio. Ogni individuo è rappresentato da un segmento verticale, partizionato in due colori. L'altezza di ciascun segmento colorato indica la probabilità di appartenenza dell'individuo ai due gruppi.



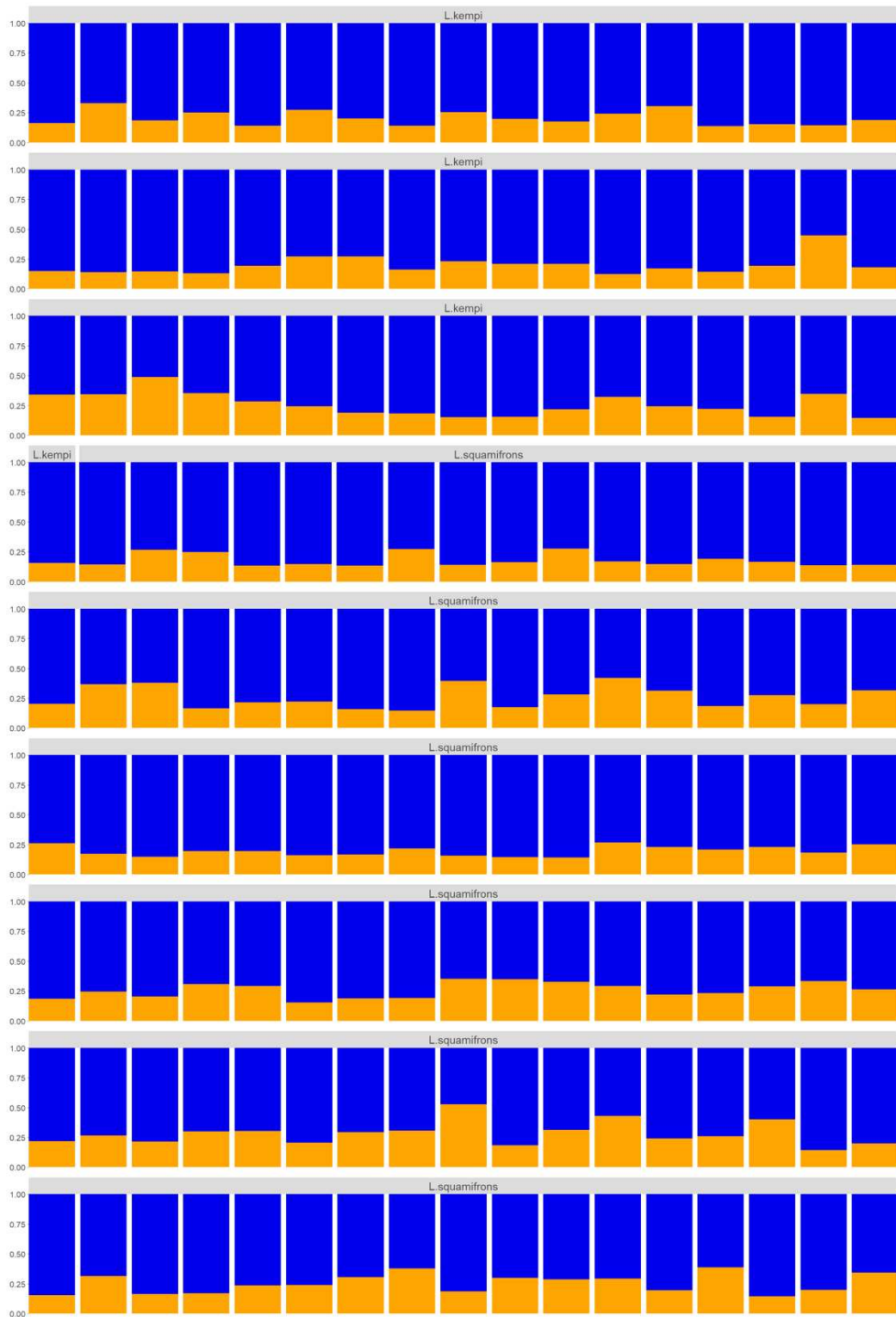
La quarta analisi è stata svolta con le stesse impostazioni della precedente, con l'unica differenza che è stato impostato il parametro *locprior*, utilizzando come *prior* l'assegnazione di ciascun individuo ad una specie sulla base della morfologia. Come prima, i risultati indicano come situazione più probabile quella che prevede la suddivisione del *dataset* in 2 gruppi; infatti, il valore massimo di  $\ln P(D)$  corrisponde ad una corsa con  $K$  pari a 2 (figura 4.5).

Tutti gli individui mostrano una probabilità di appartenenza maggiore allo stesso gruppo (in blu nella figura 4.5).



(a)

**Figura 4.5:** Ripartizione delle probabilità ( $Q$ ) che ogni individuo ha di appartenere ad uno dei due gruppi individuati da STRUCTURE, con *locprior*. Rappresentazione ottenuta in R tramite il pacchetto *pophelper* ver. 2.3.1. a) in ascissa è indicata la suddivisione a priori nelle due specie (*L. kempi* e *L. squamifrons*), in ordinata la probabilità di appartenenza al singolo gruppo ( $K$ ). b) struttura mostrata nel dettaglio. Ogni individuo è rappresentato da un segmento verticale, partizionato in due colori. L'altezza di ciascun segmento colorato indica la probabilità di appartenenza dell'individuo ai due gruppi.

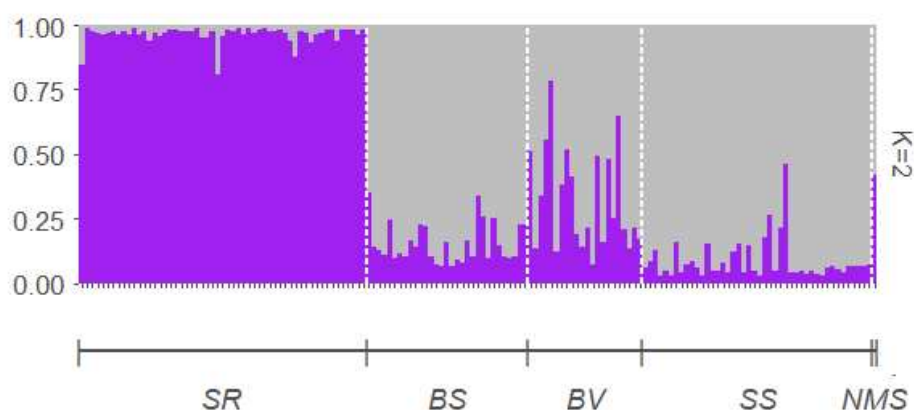


(b)

**Figura 4.5 (continuazione):** Ripartizione delle probabilità (Q) che ogni individuo ha di appartenere ad uno dei due gruppi individuati da STRUCTURE, con *locprior*. Raffigurazione ottenuta in R tramite il pacchetto *pophelper* ver. 2.3.1. a) in ascissa è indicata la suddivisione a priori nelle due specie (*L. kempi* e *L. squamifrons*), in ordinata la probabilità di appartenenza al singolo gruppo (K). b) struttura mostrata nel dettaglio. Ogni individuo è rappresentato da un segmento verticale, partizionato in due colori. L'altezza di ciascun segmento colorato indica la probabilità di appartenenza dell'individuo ai due gruppi.

Infine, è stato analizzato il *dataset* relativo alle sole specie *L. kemp* e *L. squamifrons* effettuando un raggruppamento a priori sulla base della provenienza geografica degli individui. Lo scopo era quello di verificare che la suddivisione in popolazioni, fatta in base all'area geografica di provenienza, fosse coerente con il differenziamento genetico degli individui.

L'ultima analisi è stata eseguita utilizzando i 16 *loci* microsatellite presenti nel *dataset* corretto di *L. kemp* e *L. squamifrons*, composto da 153 campioni. Per ciascun valore di K, compreso fra 2 e 8, sono state impostate 10 corse. Per ciascuna corsa è stato fissato un valore di 10000 per il *burnin* e di 100000 per le corse vere e proprie. È stato utilizzato l'*admixture model* con frequenze alleliche correlate ed è stato impostato il parametro *locprior*, utilizzando come prior il luogo in cui ciascun individuo è stato campionato. I risultati dell'analisi indicano come situazione più probabile quella che prevede la suddivisione del *dataset* in 2 gruppi, infatti, il valore massimo di  $\ln P(D)$  corrisponde ad una corsa con K pari a 2 (figura 4.6).



(a)

**Figura 4.6:** Ripartizione delle probabilità (Q) che ogni individuo ha di appartenere ad uno dei due gruppi individuati da STRUCTURE. Raffigurazione ottenuta in R tramite il pacchetto *pophelper* ver. 2.3.1. a) in ascissa è indicata la suddivisione a priori nelle cinque aree geografiche (SR: Shag Rock, BS: Bransfield Strait, BV: Bouvet Island, SS: South Shetland e NMS: Neumayer Station), in ordinata la probabilità di appartenenza al singolo gruppo (K). b) struttura mostrata nel dettaglio. Ogni individuo è rappresentato da un segmento verticale, partizionato in due colori. L'altezza di ciascun segmento colorato indica la probabilità di appartenenza dell'individuo ai due gruppi.



(b)

**Figura 4.6 (continuazione):** Ripartizione delle probabilità (Q) che ogni individuo ha di appartenere ad uno dei due gruppi individuati da STRUCTURE. Raffigurazione ottenuta in R tramite il pacchetto *pophelper* ver. 2.3.1. a) in ascissa è indicata la suddivisione a priori nelle cinque aree geografiche (SR: Shag Rock, BS: Bransfield Strait, BV: Bouvet Island, SS: South Shetland e NMS: Neumayer Station), in ordinata la probabilità di appartenenza al singolo gruppo (K). b) struttura mostrata nel dettaglio. Ogni individuo è rappresentato da un segmento verticale, partizionato in due colori. L'altezza di ciascun segmento colorato indica la probabilità di appartenenza dell'individuo ai due gruppi.

Gli individui provenienti dalle isole di Shag Rocks mostrano una maggiore probabilità di appartenere al primo gruppo (viola in figura 4.6), a differenza di tutti gli altri individui che mostrano una maggiore probabilità di appartenere al secondo gruppo (grigio in figura 4.6).

## 4.2 DNA mitocondriale

### 4.2.1 Statistiche di base e variabilità genetica

A partire dalla sequenza costituita da 1047 paia di basi del gene mitocondriale ND2 sono state calcolate le statistiche descrittive (tabella 4.11) mediante il programma *DNA Sequence Polymorphism* (Rozas *et al.*, 2017). La sequenza contiene 16 siti segreganti (siti polimorfici) in *L. kemp*, 42 in *L. squamifrons*, 99 in *N. larseni* e 39 in *N. nudifrons*. Considerando l'intero *dataset*, i siti polimorfici (196 in totale) determinano 150 aplotipi. I due parametri  $H_d$  e  $\pi$  sono misure di diversità genetica e sono stati calcolati sia entro le specie che entro le popolazioni di ciascuna specie.

A	Specie	S	$H_{ap}$	$H_d$	$\pi$
	<i>Lepidonotothen kemp</i>	16	14	0.922	0.003
	<i>Lepidonotothen squamifrons</i>	42	38	0.896	0.004
	<i>Nototheniops larseni</i>	99	79	0.992	0.005
	<i>Nototheniops nudifrons</i>	39	19	0.823	0.004

B	<i>N. larseni</i>	S	$H_{ap}$	$H_d$	$\pi$
	South Shetland	43	24	1	0.005
	Drake Passage	37	20	0.995	0.005
	Bouvet Island	55	40	0.983	0.005

**Tabella 4.11:** Statistiche descrittive basate sulla variabilità genetica del gene mitocondriale ND2 nelle quattro specie analizzate in questa tesi (A), nelle popolazioni di *N. larseni* (B), nelle popolazioni di *N. nudifrons* (C) e nelle popolazioni di *L. kemp* (ke) e *L. squamifrons* (sq) (D). Vengono riportati in tabella il numero di siti segreganti (S), il numero di aplotipi ( $H_{ap}$ ), la haplotype diversity ( $H_d$ ) e la nucleotide diversity ( $\pi$ ).

C	<i>N. nudifrons</i>	S	H <sub>ap</sub>	H <sub>d</sub>	$\pi$
	Elephant Island	14	8	0.859	0.004
	South Shetland	7	5	0.667	0.002
	South Orkney	15	10	0.923	0.004

D	<i>L. kempfi</i> e <i>L. squamifrons</i>	S	H <sub>ap</sub>	H <sub>d</sub>	$\pi$
	Bransfield Strait (ke)	3	4	1	0.001
	Bouvet Island (ke)	13	10	0.875	0.003
	South Shetland (sq)	16	16	0.824	0.002
	Shag Rocks (sq)	23	20	0.755	0.001

**Tabella 4.11 (continuazione):** Statistiche descrittive basate sulla variabilità genetica del gene mitocondriale ND2 nelle quattro specie analizzate in questa tesi (A), nelle popolazioni di *N. larseni* (B), nelle popolazioni di *N. nudifrons* (C) e nelle popolazioni di *L. kempfi* (ke) e *L. squamifrons* (sq) (D). Vengono riportati in tabella il numero di siti segreganti (S), il numero di aplotipi (H<sub>ap</sub>), la haplotype diversity (H<sub>d</sub>) e la nucleotide diversity ( $\pi$ ).

#### 4.2.2 Differenziamento genetico inter ed intraspecifico

La struttura di popolazione e di specie è stata analizzata con il programma ARLEQUIN ver. 3.5.2.2 (Excoffier e Lischer, 2010). Il programma permette di quantificare l'indice  $F_{ST}$  e il numero medio di differenze nucleotidiche tra coppie di popolazioni ( $P_i$ ) mediante l'analisi della varianza molecolare (AMOVA), utilizzata per valutare l'esistenza di una struttura genetica all'interno dei vari campioni mediante il confronto della variabilità genetica all'interno e tra le popolazioni o specie. La significatività dell'indice  $F_{ST}$  è stata testata attraverso un approccio di permutazioni non parametriche, con un numero di permutazioni pari a 1000 e una soglia di significatività pari a 0.05. Sono stati inoltre calcolati i confronti a coppie degli indici  $F_{ST}$  per tutte le popolazioni o specie, in quanto gli  $F_{ST}$  a coppie sono importanti per valutare le distanze genetiche tra popolazioni o specie.

Tutti i valori di differenziamento genetico tra le quattro specie analizzate in questa tesi, stimati mediante l'indice  $F_{ST}$ , sono statisticamente

significativi (tabella 4.12). È presente un basso livello di differenziamento tra le specie *L. kemp* e *L. squamifrons* ( $F_{ST}=0.118$ ,  $p\text{-value} < 0.05$ ). I livelli di differenziamento tra tutte le altre coppie di specie sono più elevati, nonostante si possa vedere che *N. larseni* e *N. nudifrons* siano più simili tra loro ( $F_{ST}=0.954$ ,  $p\text{-value} < 0.001$ ) di quanto non siano con *L. kemp* e *L. squamifrons* ( $F_{ST}$  compreso tra 0.973 e 0.968). Le specie che risultano geneticamente più distanti sono *L. nudifrons* e *L. kemp* ( $F_{ST}=0.973$ ,  $p\text{-value} < 0.001$ ).

	<i>Lepidonotothen kemp</i>	<i>Lepidonotothen squamifrons</i>	<i>Nototheniops larseni</i>	<i>Nototheniops nudifrons</i>
<i>Lepidonotothen kemp</i>	*	0.003	< 0.001	< 0.001
<i>Lepidonotothen squamifrons</i>	<b>0.118</b>	*	< 0.001	< 0.001
<i>Nototheniops larseni</i>	<b>0.968</b>	<b>0.969</b>	*	< 0.001
<i>Nototheniops nudifrons</i>	<b>0.973</b>	<b>0.972</b>	<b>0.954</b>	*

**Tabella 4.12:**  $F_{ST}$  a coppie tra le quattro specie ipotetiche *L. kemp*, *L. squamifrons*, *N. larseni* e *N. nudifrons* basati sulla variabilità del gene mitocondriale ND2. I valori di  $F_{ST}$  sono riportati sotto la diagonale, mentre sopra la diagonale sono riportati i valori del  $p\text{-value}$  ( $\alpha = 0.05$ ). I valori significativi sono evidenziati in grassetto.

È stato calcolato inoltre il numero medio di differenze nucleotidiche tra coppie di specie ( $P_i$ ) e i risultati confermano quanto detto in precedenza (tabella 4.13). Tra le specie *L. kemp* e *L. squamifrons* si può notare un differenziamento genetico minore ( $P_i=4.215$ ) rispetto a tutte le altre coppie di specie. *N. larseni* e *N. nudifrons* sono più simili tra loro ( $P_i=105.115$ ) di quanto non lo siano con *L. kemp* e *L. squamifrons* ( $P_i$  compreso tra 146.013 e 147.917). Le specie che risultano più distanti sono *L. squamifrons* e *N. nudifrons* ( $P_i=147.917$ ).

L'indice  $F_{ST}$  totale è pari a 0.965 ( $p\text{-value} < 0.001$ ) indicando un differenziamento altamente significativo tra le specie. Secondo l'analisi della varianza molecolare (AMOVA) la varianza totale osservata nel campione è dovuta per il 96.47% a variazioni tra specie e per il 3.53% a variazioni all'interno delle singole specie, a conferma della presenza di specie distinte geneticamente.

	<i>Lepidonotothen kempfi</i>	<i>Lepidonotothen squamifrons</i>	<i>Notototheniops larseni</i>	<i>Notototheniops nudifrons</i>
<i>Lepidonotothen kempfi</i>	3.281	<b>4.215</b>	<b>146.014</b>	<b>146.541</b>
<i>Lepidonotothen squamifrons</i>	0.003	4.082	<b>147.349</b>	<b>147.917</b>
<i>Notototheniops larseni</i>	< 0.001	< 0.001	5.025	<b>105.115</b>
<i>Notototheniops nudifrons</i>	< 0.001	< 0.001	< 0.001	4.321

**Tabella 4.13:** Differenze nucleotidiche (Pi) a coppie tra le quattro specie ipotetiche *L. kempfi*, *L. squamifrons*, *N. larseni* e *N. nudifrons* basate sulla variabilità del gene mitocondriale ND2. Sopra la diagonale è riportato il numero medio di differenze a coppie tra le specie ( $Pi_{XY}$ ), sotto la diagonale sono riportati i corrispondenti *p-value* e lungo la diagonale è riportato il numero medio delle differenze a coppie all'interno delle specie ( $Pi_X$ ). I valori significativi sono evidenziati in grassetto.

Sono stati poi condotti confronti tra le diverse popolazioni all'interno di ciascuna specie, con l'intento di individuare eventuali differenze genetiche legate alla distanza geografica.

In *N. larseni* si possono distinguere tre popolazioni in base al sito di campionamento (South Shetland, Drake Passage e Bouvet Island). Il differenziamento genetico tra le tre popolazioni, stimato mediante l'indice  $F_{ST}$  (tabella 4.14), indica che la popolazione delle isole South Shetland non è differenziata significativamente dalla popolazione del Passaggio di Drake ( $F_{ST}=0.006$ , *p-value*=0.234) e dalla popolazione dell'isola di Bouvet ( $F_{ST}=0.014$ , *p-value*=0.091). Solamente le popolazioni dell'isola di Bouvet e del Passaggio di Drake sono differenziate significativamente, mostrando un livello di differenziamento molto basso ( $F_{ST}=0.022$ , *p-value*<0.05).

	South Shetland	Drake Passage	Bouvet Island
South Shetland	*	0.234	0.091
Drake Passage	0.006	*	0.027
Bouvet Island	0.014	<b>0.022</b>	*

**Tabella 4.14:**  $F_{ST}$  a coppie tra le tre popolazioni di *N. larseni*, basati sulla variabilità del gene mitocondriale ND2. I valori di  $F_{ST}$  sono riportati sotto la diagonale, mentre sopra la diagonale sono riportati i valori del *p-value* ( $\alpha = 0.05$ ). I risultati significativi sono evidenziati in grassetto.



Il numero medio di differenze nucleotidiche tra coppie di popolazioni ( $P_i$ ) (tabella 4.15) conferma il risultato precedente in quanto solo le popolazioni dell'isola di Bouvet e del Passaggio di Drake sono differenziate in maniera significativa ( $P_i=5.037$ ,  $p\text{-value}<0.05$ ).

L'indice  $F_{ST}$  totale è pari a 0.016 ( $p\text{-value} < 0.05$ ) indicando un basso livello di differenziamento tra le popolazioni. Secondo l'analisi della varianza molecolare (AMOVA) la varianza totale osservata nel campione è dovuta per l'1.59% a variazioni tra popolazioni e per il 98.41% a variazioni all'interno delle singole popolazioni, a conferma l'elevata somiglianza genetica tra le popolazioni in esame.

	South Shetland	Drake Passage	Bouvet Island
South Shetland	5.429	5.319	5.136
Drake Passage	0.234	5.149	<b>5.037</b>
Bouvet Island	0.110	0.033	4.721

**Tabella 4.15:** Differenze nucleotidiche ( $P_i$ ) a coppie tra le tre popolazioni di *N. larseni*, basate sulla variabilità del gene mitocondriale ND2. Sopra la diagonale è riportato il numero medio di differenze a coppie tra le popolazioni ( $P_{i_{xy}}$ ), sotto la diagonale sono riportati i corrispondenti  $p\text{-value}$  e lungo la diagonale è riportato il numero medio delle differenze a coppie all'interno delle popolazioni ( $P_{i_x}$ ). I risultati significativi sono evidenziati in grassetto.

In *N. nudifrons* si possono distinguere tre popolazioni in base al sito di campionamento (South Shetland, Elephant Island e South Orkney). Il differenziamento genetico tra le tre popolazioni, stimato mediante gli indici  $F_{ST}$  e  $P_i$ , non indica differenziamento genetico significativo tra le popolazioni in esame (valori non riportati). Secondo l'analisi della varianza molecolare (AMOVA) la varianza totale osservata nel campione è dovuta per lo 0.91% a variazioni tra popolazioni e per il 99.09% a variazioni all'interno delle singole popolazioni, a conferma dell'assenza di differenziamento genetico tra le popolazioni in esame.

Unendo i *dataset* di *L. kempi* e *L. squamifrons*, si possono distinguere quattro popolazioni in base al sito di campionamento (South Shetland, Bransfield Strait, Shag Rocks e Bouvet Island). Il differenziamento

genetico tra le quattro popolazioni, stimato mediante l'indice  $F_{ST}$  (tabella 4.16), indica che la popolazione delle isole South Shetland non è differenziata significativamente dalla popolazione dello Stretto di Bransfield. Tutte le altre popolazioni sono differenziate significativamente. La popolazione delle isole di Shag Rocks mostra i livelli più alti di differenziamento rispetto alle popolazioni delle isole South Shetland ( $F_{ST}=0.733$ ,  $p\text{-value}<0.001$ ), dello Stretto di Bransfield ( $F_{ST}=0.753$ ,  $p\text{-value}<0.001$ ) e dell'isola di Bouvet ( $F_{ST}=0.538$ ,  $p\text{-value}<0.001$ ). Le popolazioni più simili tra loro sono le popolazioni dell'isola di Bouvet con le popolazioni dello Stretto di Bransfield ( $F_{ST}=0.269$ ,  $p\text{-value}<0.05$ ) e con le popolazioni delle isole South Shetland ( $F_{ST}=0.275$ ,  $p\text{-value}<0.001$ ).

	South Shetland	Bransfield Strait	Shag Rocks	Bouvet Island
South Shetland	*	0.706	< 0.001	< 0.001
Bransfield Strait	0	*	< 0.001	0.019
Shag Rocks	<b>0.733</b>	<b>0.753</b>	*	< 0.001
Bouvet Island	<b>0.275</b>	<b>0.269</b>	<b>0.538</b>	*

**Tabella 4.16:**  $F_{ST}$  a coppie tra le tre popolazioni di *L. kempfi* e *L. squamifrons*, basati sulla variabilità del gene mitocondriale ND2. I valori di  $F_{ST}$  sono riportati sotto la diagonale, mentre sopra la diagonale sono riportati i valori del  $p\text{-value}$  ( $\alpha = 0.05$ ). I risultati significativi sono evidenziati in grassetto. I valori di  $F_{ST}$  negativi sono stati cambiati con 0.

Il numero medio di differenze nucleotidiche tra coppie di popolazioni ( $P_i$ ) conferma il risultato precedente (tabella 4.17).

L'indice  $F_{ST}$  totale è pari a 0.613 ( $p\text{-value} < 0.001$ ) indicando un differenziamento altamente significativo tra le popolazioni. Secondo l'analisi della varianza molecolare (AMOVA) la varianza totale osservata nel campione è dovuta per l'61.25% a variazioni tra popolazioni e per il 38.75% a variazioni all'interno delle singole popolazioni.

	South Shetland	Bransfield Strait	Shag Rocks	Bouvet Island
South Shetland	1.798	1.605	<b>6.533</b>	<b>3.244</b>
Bransfield Strait	0.660	1.500	<b>6.750</b>	<b>3.434</b>
Shag Rocks	< 0.001	< 0.001	1.695	<b>4.783</b>
Bouvet Island	< 0.001	0.015	< 0.001	3.006

**Tabella 4.17:** Differenze nucleotidiche ( $P_i$ ) a coppie tra le tre popolazioni di *L. kempfi* e *L. squamifrons*, basate della variabilità del gene mitocondriale ND2. Sopra la diagonale è riportato il numero medio di differenze a coppie tra le popolazioni ( $P_{i_{XY}}$ ), sotto la diagonale sono riportati i corrispondenti  $p$ -value e lungo la diagonale è riportato il numero medio delle differenze a coppie all'interno delle popolazioni ( $P_{i_X}$ ). I risultati significativi sono evidenziati in grassetto.

L'ultima analisi è stata svolta considerando *L. kempfi* e *L. squamifrons* come due gruppi distinti, ciascuno dei quali suddiviso in popolazioni sulla base del sito di campionamento. *L. kempfi* è stata suddivisa in due popolazioni (Bransfield Strait e Bouvet Island) e *L. squamifrons* è stata suddivisa in due popolazioni (South Shetland e Shag Rocks). Come nel caso precedente, il differenziamento genetico tra le tre popolazioni, stimato mediante l'indice  $F_{ST}$  (tabella 4.18), indica che la popolazione di *L. squamifrons* nelle isole South Shetland non è differenziata significativamente dalla popolazione di *L. kempfi* dello Stretto di Bransfield. Tutte le altre popolazioni sono differenziate significativamente. La popolazione di *L. squamifrons* nelle isole di Shag Rocks mostra i livelli più alti di differenziamento rispetto alla popolazione di *L. squamifrons* nelle isole South Shetland ( $F_{ST}=0.767$ ,  $p$ -value<0.001), alla popolazione di *L. kempfi* dello Stretto di Bransfield Strait ( $F_{ST}=0.801$ ,  $p$ -value<0.001) e alla popolazione di *L. kempfi* dell'isola di Bouvet ( $F_{ST}=0.576$ ,  $p$ -value<0.001). Le popolazioni più simili tra loro sono la popolazione di *L. kempfi* dell'isola di Bouvet con la popolazione di *L. kempfi* dello Stretto di Bransfield ( $F_{ST}=0.264$ ,  $p$ -value<0.05) e con la popolazione di *L. squamifrons* delle isole South Shetland ( $F_{ST}=0.277$ ,  $p$ -value<0.001).

	<b>Bransfield Strait (ke)</b>	<b>Bouvet Island (ke)</b>	<b>South Shetland (sq)</b>	<b>Shag Rocks (sq)</b>
<b>Bransfield Strait (ke)</b>	*	0.030	0.689	< 0.001
<b>Bouvet Island (ke)</b>	<b>0.264</b>	*	< 0.001	< 0.001
<b>South Shetland (sq)</b>	0	<b>0.277</b>	*	< 0.001
<b>Shag Rocks (sq)</b>	<b>0.801</b>	<b>0.576</b>	<b>0.767</b>	*

**Tabella 4.18:**  $F_{ST}$  a coppie tra le tre popolazioni di *L. kempfi* e *L. squamifrons*, basati sulla variabilità del gene mitocondriale ND2. I valori di  $F_{ST}$  sono riportati sotto la diagonale, mentre sopra la diagonale sono riportati i valori del  $p$ -value ( $\alpha = 0.05$ ). I risultati significativi sono evidenziati in grassetto. I valori di  $F_{ST}$  negativi sono stati cambiati con 0.

Il numero medio di differenze nucleotidiche tra coppie di popolazioni ( $P_i$ ) conferma il risultato precedente (tabella 4.19).

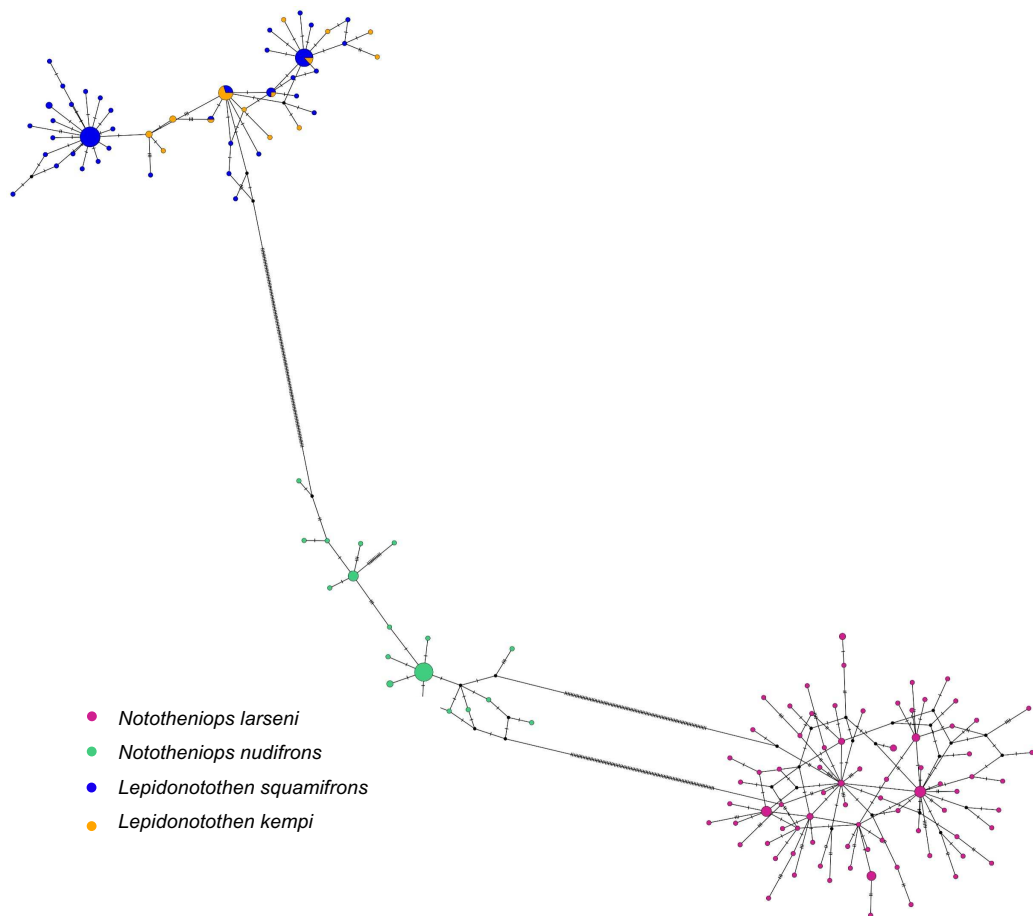
Gli andamenti del differenziamento tra popolazioni sembrano legati solo alla posizione geografica e non all'appartenenza al gruppo di *L. kempfi* o *L. squamifrons*.

	<b>Bransfield Strait (ke)</b>	<b>Bouvet Island (ke)</b>	<b>South Shetland (sq)</b>	<b>Shag Rocks (sq)</b>
<b>Bransfield Strait (ke)</b>	1.500	<b>3.456</b>	1.605	<b>6.850</b>
<b>Bouvet Island (ke)</b>	0.021	3.074	<b>3.277</b>	<b>4.681</b>
<b>South Shetland (sq)</b>	0.633	< 0.001	1.798	<b>6.623</b>
<b>Shag Rocks (sq)</b>	< 0.001	< 0.001	< 0.001	1.340

**Tabella 4.19:** Differenze nucleotidiche ( $P_i$ ) a coppie tra le tre popolazioni di *L. kempfi* e *L. squamifrons*, basate sulla variabilità gene mitocondriale ND2. Sopra la diagonale è riportato il numero medio di differenze a coppie tra le popolazioni ( $P_{iXY}$ ), sotto la diagonale sono riportati i corrispondenti  $p$ -value e lungo la diagonale è riportato il numero medio delle differenze a coppie all'interno delle popolazioni ( $P_{iX}$ ). I risultati significativi sono evidenziati in grassetto.

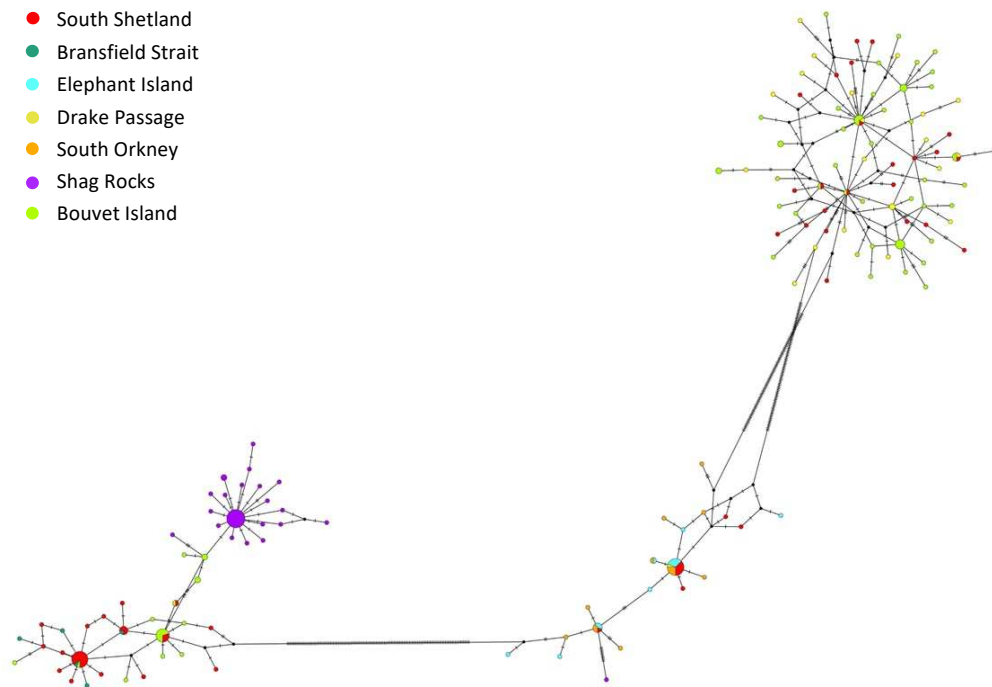
L'indice  $F_{ST}$  totale è pari a 0.567 ( $p$ -value < 0.001) indicando un differenziamento altamente significativo tra le popolazioni.

La rappresentazione grafica del *network* di aplotipi, utilizzando il programma POPART ver. 1.7.2 (Leigh e Bryant, 2015), è stata svolta mostrando le relazioni tra le quattro specie (figura 4.7). Il *network* di aplotipi suggerisce la presenza di tre raggruppamenti principali che corrispondono alle specie *N. larseni*, *N. nudifrons* e *L. kempi* insieme a *L. squamifrons*.



**Figura 4.7:** *Network* aplotipico costruito dalle sequenze mitocondriali del gene ND2 nelle specie analizzate in questa tesi. Il *network* è stato sviluppato mediante il software POPART usando l'algoritmo descritto da Clement et al. (2002) sulla base di una sequenza di 1047 bp del gene mitocondriale ND2 in 232 individui. I cerchi corrispondono agli aplotipi di *L. kempi* (giallo), *L. squamifrons* (blu), *N. larseni* (rosa) e *N. nudifrons* (verde) e la dimensione è proporzionale al numero di individui che possiedono quello specifico aplotipo. I segni di tratteggio sui collegamenti tra cerchi rappresentano il numero di mutazioni intervenute tra gli aplotipi.

La stessa analisi è stata poi riportata colorando i campioni sulla base dell'area geografica di origine (figura 4.8). Gli individui provenienti dalle isole di Shag Rocks appaiono molto simili tra loro e geneticamente distinti dagli altri individui.



**Figura 4.8:** *Network* aplotipico costruito dalle sequenze mitocondriali del gene ND2 nei campioni analizzati in questa tesi e suddivisi in base al sito di campionamento. Il network è stato sviluppato mediante il software POPART usando l'algoritmo descritto da Clement et al. (2002) sulla base di una sequenza di 1047 bp del gene mitocondriale ND2 in 232 individui. I cerchi corrispondono agli aplotipi provenienti da South Shetland (rosso), Bransfield Strait (verde scuro), Elephant Island (azzurro), Drake Passage (giallo), South Orkney (arancione), Shag Rocks (viola) e Bouvet Island (verde chiaro). La dimensione dei cerchi è proporzionale al numero di individui che possiedono quello specifico aplotipo. I segni di tratteggio sui collegamenti tra cerchi rappresentano il numero di mutazioni intervenute tra gli aplotipi.

## 5. DISCUSSIONE

In questo lavoro di tesi è stata analizzata la variabilità genetica in *Lepidonotothen kemp*, *Lepidonotothen squamifrons*, *Nototheniops larseni* e *Nototheniops nudifrons*, quattro specie di pesci antartici provenienti dall'area dell'Arco di Scozia, utilizzando come marcatori molecolari 21 *loci* microsatellite e il gene mitocondriale ND2.

L'obiettivo principale consisteva nello stabilire se gli individui assegnati alle specie *L. kemp* e *L. squamifrons*, sulla base di caratteristiche morfologiche, presentassero una variabilità genetica tale da considerarli come una sola specie o come due specie distinte. Questa suddivisione è stata a lungo dibattuta ed è tuttora fonte di dubbi.

L'analisi svolta tramite il programma STRUCTURE, utilizzando i dati relativi ai *loci* microsatellite, ha individuato tre gruppi geneticamente distinti, corrispondenti alle specie *N. larseni*, *N. nudifrons* e *L. kemp* insieme a *L. squamifrons* (figure 4.2 e 4.3 in risultati). Questo conferma la recente lista di controllo compilata da Eastman e Eakin (2021) che, all'interno del genere *Lepidonotothen*, riconosce solo la specie *L. squamifrons*.

I valori di  $F_{ST}$  a coppie, sia per i *loci* microsatellite (tabella 4.8 in risultati) che per il gene mitocondriale ND2 (tabella 4.12 in risultati), indicano un differenziamento altamente significativo tra le quattro specie analizzate ( $F_{ST}$  globale=0.965,  $p$ -value<0.001 per il gene mitocondriale ND2). Come atteso è stato trovato un basso livello di differenziamento tra *L. kemp* e *L. squamifrons* ( $F_{ST}$ =0.118,  $p$ -value<0.05 per il gene mitocondriale ND2;  $F_{ST}$ =0.013,  $p$ -value<0.001 per i *loci* microsatellite) e un differenziamento molto elevato tra le altre coppie di specie ( $F_{ST}$  compreso tra 0.973 e 0.954 per il gene mitocondriale ND2;  $F_{ST}$  compreso tra 0.611 e 0.333 per i *loci* microsatellite). Questo conferma il recente studio di Miya *et al.* (2016), basato anch'esso sul gene mitocondriale ND2, che ha riconosciuto *L. kemp* come appartenente alla stessa specie di *L. squamifrons*.

Una successiva analisi svolta tramite il programma STRUCTURE, condotta considerando solo le specie *L. kemp*i e *L. squamifrons* e utilizzando i dati relativi ai *loci* microsatellite, ha confermato l'assenza di differenziamento genetico (figure 4.4 e 4.5 in risultati). La suddivisione del campione in due gruppi riguarda una mera impostazione del programma in quanto non vi è la possibilità di svolgere un'analisi considerando la presenza di un unico gruppo. Nonostante questo, si nota come tutti gli individui abbiano una probabilità simile di appartenere ai due gruppi, senza differenze tra individui inizialmente attribuiti a specie diverse.

In questo lavoro di tesi sono state inoltre analizzate le strutture di popolazione in ciascuna specie, con l'obiettivo di identificare eventuali differenze genetiche legate alla distanza geografica. Dal momento che la maggior parte dell'Arco di Scozia è situato a sud del Fronte Polare Antartico e all'interno della Corrente Circumpolare Antartica, è considerato un centro di dispersione per gli organismi marini. Nonostante le isole siano molto distanti tra loro, la distribuzione di molte specie di Nototenoidei comprende tutto l'Arco di Scozia, fino all'isola di Bouvet (Jones *et al.*, 2008).

All'interno del campione di *L. kemp*i e *L. squamifrons*, l'unico differenziamento che si può notare è basato sulla provenienza geografica dei diversi individui. L'analisi svolta tramite il programma STRUCTURE mostra come i campioni provenienti dalle isole di Shag Rocks presentino un netto differenziamento rispetto a tutti gli altri individui (figura 4.6 in risultati).

Questo *pattern* è evidente anche dai valori di  $F_{ST}$  a coppie, calcolati sulla base dei *loci* microsatellite (tabella 4.10 in risultati) e del gene mitocondriale ND2 (tabella 4.16 in risultati). Per il gene mitocondriale ND2 la popolazione delle isole di Shag Rocks presenta livelli alti e significativi di differenziamento genetico rispetto alle popolazioni delle isole South Shetland ( $F_{ST}=0.733$ ,  $p\text{-value}<0.001$ ), dello Stretto di Bransfield ( $F_{ST}=0.753$ ,  $p\text{-value}<0.001$ ) e dell'isola di Bouvet ( $F_{ST}=0.538$ ,



$p\text{-value}<0.001$ ). Per i *loci* microsatellite, i livelli di differenziamento genetico tra la popolazione delle isole di Shag Rocks e le altre popolazioni sono molto bassi ( $F_{ST}$  compreso tra 0.057 e 0.044,  $p\text{-value}<0.001$ ), ma sono gli unici che risultano altamente significativi. Le popolazioni delle isole South Shetland e le popolazioni dello Stretto di Bransfield, le più vicine geograficamente, non mostrano un differenziamento significativo. Per il gene mitocondriale ND2 è stato trovato inoltre un differenziamento significativo, anche se basso, tra le popolazioni dell'isola di Bouvet e le popolazioni delle isole South Shetland ( $F_{ST}=0.275$ ,  $p\text{-value}<0.001$ ) e dello Stretto di Bransfield ( $F_{ST}=0.269$ ,  $p\text{-value}<0.05$ ). Questo conferma i risultati di Jones *et al.* (2008), basati sul gene mitocondriale ND2, che mostravano un differenziamento genetico significativo tra popolazioni di *L. squamifrons* delle isole South Shetlands e popolazioni di *L. squamifrons* dell'isola di Bouvet.

La popolazione dell'isola di Bouvet Island, che è più distante geograficamente dalle popolazioni dell'isola di South Shetland e dello Stretto di Bransfield di quanto non lo siano le popolazioni delle isole di Shag Rocks, mostra un differenziamento non significativo per i *loci* microsatellite e molto basso per il gene mitocondriale ND2. Una possibile spiegazione potrebbe essere legata al fatto che il Mare di Weddell è limitato a nord dalle piattaforme continentali delle isole di South Georgia e di South Orkney, che determinano la formazione di un fronte (La Mesa *et al.*, 2004) che esclude le isole di Shag Rock dalla circolazione oceanica, portando ad una distanza genetica dovuta ad una mancanza di connettività fisica. Inoltre, l'isola di Bouvet è situata al confine nord della Corrente Circumpolare Antartica ma risulta ancora a sud del Fronte Polare Antartico, e può giocare un ruolo importante nel trasporto di organismi marini all'interno del Mare di Weddell grazie al *Weddell Gyre* (Arntz, 2006). Quest'isola si trova all'estremità orientale dell'Arco di Scozia, a circa 2500 km dall'isola di South Georgia, senza isole o montagne sottomarine a dividerle ed è un esempio tipico di isola oceanica remota.

La separazione genetica netta degli individui provenienti dalle isole di Shag Rocks è evidente anche dal *network* aplo-tipico (figura 4.8 in risultati).

Gli unici individui che sono stati campionati presso le isole di Shag Rocks appartengono a *L. kempi* e *L. squamifrons*, di conseguenza non è stato possibile verificare la presenza dello stesso *pattern* anche nelle specie *N. nudifrons* e *N. larseni*. In un'analisi futura sarebbe interessante raccogliere individui appartenenti a tutte le specie in ciascun sito di campionamento, in modo da rendere possibili confronti più accurati.

In *N. larseni*, analizzando i valori di  $F_{ST}$  a coppie (tabelle 4.9 e 4.14 in risultati), si può notare un differenziamento significativo, ma estremamente ridotto, tra la popolazione del Passaggio di Drake e la popolazione dell'isola di Bouvet ( $F_{ST}=0.022$ ,  $p\text{-value}<0.05$  per il gene mitocondriale ND2;  $F_{ST}=0.016$ ,  $p\text{-value}<0.05$  per i *loci* microsatellite) a cui si aggiunge, solo per i *loci* microsatellite, un differenziamento significativo molto basso tra la popolazione dell'isola di Bouvet e la popolazione delle isole South Shetland ( $F_{ST}=0.016$ ,  $p\text{-value}<0.05$ ). Anche se significativo, il differenziamento è estremamente ridotto, mostrando un'omogeneità genetica tra le tre popolazioni. Questo conferma sia i risultati di Jones *et al.* (2008) secondo cui *N. larseni* non mostrava alcun differenziamento genetico significativo nell'area dell'Arco di Scozia, sia i risultati di Deli Antoni *et al.* (2019) secondo cui non vi è alcun differenziamento tra la popolazione della Penisola Antartica e la popolazione delle isole South Shetland. *Nototherniops larseni* è una specie caratterizzata dallo sviluppo larvale più lungo di cui si è a conoscenza, con una durata di circa un anno, il che potrebbe determinare un incremento del suo potenziale di dispersione e di flusso genico (North, 2001). Larve di *N. larseni* sono state inoltre trovate disperse su ampie aree (Deli Antoni *et al.*, 2019). Questo sembra confermare l'aspettativa generale secondo cui specie con una fase larvale pelagica più lunga abbiano una dispersione maggiore, e quindi, una struttura di popolazione più debole, rispetto alle specie con una fase larvale più breve.

In *N. nudifrons*, analizzando i valori di  $F_{ST}$  a coppie, non è stato riscontrato alcun differenziamento significativo tra le popolazioni delle isole South Shetland, dell'isola di Elephant e delle isole di South Orkney, mostrando un'omogeneità genetica tra le popolazioni. Questo andamento potrebbe essere facilmente spiegato per le popolazioni delle isole South Shetland e dell'isola di Elephant in quanto si trovano a breve distanza geografica. Lo stesso non si può dire per la popolazione delle isole di South Orkney che si trovano a circa 250 km di distanza, separate da acque che hanno una profondità compresa tra 2000 e 3000 (La Mesa *et al.*, 2004). Una spiegazione potrebbe essere attribuita ad una dispersione larvale legata alla circolazione oceanica, in quanto tutte le isole si trovano lungo aree interessate dalla Corrente Circumpolare Antartica. Questo risultato è in contrasto sia con l'evidenza secondo cui questa è una specie caratterizzata dalla presenza di cure parentali, che potrebbero causare una limitazione al flusso genico tra popolazioni, sia con il fatto che le larve di *N. nudifrons* sono state trovate principalmente vicino alla costa (Deli Antoni *et al.*, 2019). Uno studio di Deli Antoni *et al.* (2019), in linea con quanto appena detto, trovò un differenziamento significativo tra le popolazioni della Penisola Antartica e delle isole di South Shetland, separate solo dallo Stretto di Bransfield, largo 100 km e con una profondità massima di 1000 m.

Le statistiche descrittive relative al gene mitocondriale ND2 hanno mostrato la presenza di 150 aplotipi differenti. Sulla base delle misure di  $H_d$  e  $\pi$  (tabella 4.11 in risultati) sono state fatte alcune deduzioni riguardo la variabilità genetica delle specie analizzate in questa tesi. Secondo Grant e Bowen (1998) le popolazioni di specie marine possono essere distinte in quattro categorie sulla base di diverse combinazioni di  $H_d$  e  $\pi$ , permettendo di fare ipotesi riguardo alla storia passata della popolazione in esame. Le specie e le popolazioni analizzate in questa tesi presentano tutti alti valori di  $H_d$  ( $H_d > 0.5$ ) e bassi valori di  $\pi$  ( $\pi < 0.05$ ) e questo potrebbe indicare che è avvenuta un'espansione rapida partendo da una popolazione di piccole dimensioni. Questi risultati sembrerebbero

confermare la teoria secondo cui durante i periodi di massima glaciazione la fauna bentonica subì delle frammentazioni, trovando rifugio in zone isolate e libere dai ghiacci (Thatje *et al.*, 2005). In seguito al ritiro dei ghiacci queste popolazioni di piccole dimensioni riconquistarono gli habitat, prima occupati dai ghiacci, e furono caratterizzate da una rapida espansione. Le isole dell'Antartide, grazie alla loro esposizione alla Corrente Circumpolare Antartica, sono state importanti rifugi per specie bentoniche di acqua bassa. In qualche modo le isole intorno al continente antartico possono aver contribuito alla sopravvivenza della fauna della piattaforma e della scarpata continentale durante i periodi di massima glaciazione, a condizione che le specie fossero in grado di distribuirsi grazie alla presenza di fasi larvali planctoniche (Arntz, 2006).

## 6. CONCLUSIONI

La conoscenza della reale composizione in specie dell'Oceano Meridionale è un'informazione necessaria per una corretta gestione della biodiversità in ambiente antartico. La comprensione della variabilità genetica all'interno e tra le popolazioni è, allo stesso tempo, di fondamentale importanza in quanto, come esplicitato dal teorema di Fisher, "la capacità adattativa di una popolazione è direttamente proporzionale alla sua variabilità genetica in un dato momento" (Fisher, 1930). Quindi, a livelli più elevati di variabilità corrisponde una maggiore capacità della popolazione di adattarsi a mutate condizioni ambientali. Questo è un tema che oggi è al centro dell'attenzione a causa dei sempre più pressanti cambiamenti ambientali dovuti al riscaldamento globale.

C'è una grande necessità di studiare aree che possono essere in grado di aumentare la nostra conoscenza in merito alla capacità di resistenza degli organismi al riscaldamento globale e rafforzare la nostra capacità di prevedere la propensione o l'incapacità delle specie di sopravvivere. La Penisola Antartica è uno dei posti più adatti per queste ricerche in quanto è uno dei tre luoghi al mondo che ha registrato il più intenso riscaldamento negli ultimi 50 anni, circa cinque volte la media mondiale. A differenza dell'Artico, le condizioni climatiche in Antartide sono state storicamente molto più stabili e la fauna che si è evoluta *in situ* è probabilmente molto più sensibile ai cambiamenti ambientali (Barnes e Peck, 2008).

La risposta degli ecosistemi ai cambiamenti climatici è uno dei temi più urgenti nelle future ricerche in ambiente antartico. Studi sperimentali mostrano come le specie bentoniche sono molto sensibili ai cambiamenti di temperatura. Anche incrementi ridotti della temperatura, dell'ordine di 2 o 3 °C, possono impedire la loro capacità di eseguire funzioni critiche, come la capacità di evitare i predatori (Barnes e Peck, 2008). Gli ecosistemi bentonici possono essere considerati come sentinelle per monitorare gli effetti dei cambiamenti climatici (Sahade *et al.*, 2015).

Il continente antartico è uno dei posti che si sta riscaldando di più e capire gli effetti che questo può avere in quella regione può permettere anche di capire quello che potrebbe succedere globalmente.

## 7. BIBLIOGRAFIA

Álvarez-Noriega, M., Burgess, S., Byers, J.E., Pringle, J.M., Wares, J.P. & Marshall, D.J. (2020). Global biogeography of marine dispersal potential. *Nature Ecology & Evolution* 4: 1196-1203. <https://doi.org/10.1038/s41559-020-1238-y>.

Amos, W., Hoffman, J.I., Frodsham, A., Zhang, L., Best, S. & Hill, A.V.S. (2007). Automated binning of microsatellite alleles: problems and solutions. *Molecular Ecology* 7: 10-14. <https://doi.org/10.1111/j.1471-8286.2006.01560.x>.

Arntz, W.E. (2006). Bouvet Island: a stepping stone in the Southern Ocean. *Polar Biology* 29: 81-82. <https://doi.org/10.1007/s00300-005-0092-3>.

Balushkin, A. V. (1994). *Proeleginops grandeastmanorum* gen. et sp. nov. (Perciformes, Notothenioidei, Eleginopsidae) from the Late Eocene of Seymour Island (Antarctica) is a fossil Notothenioid, not a gadiform. *Journal of Ichthyology* 34: 10-23.

Balushkin, A. V. (2000). Morphology, Classification, and Evolution of Notothenioid Fishes of the Southern Ocean (Notothenioidei, Perciformes). *Journal of Ichthyology* 40: S74-S109.

Bargelloni, L., Marcato, S., Zane, L. & Patarnello, T. (2000). Mitochondrial Phylogeny of Notothenioids: A Molecular Approach to Antarctic Fish Evolution and Biogeography. *Systematic Biology* 49: 114-129. <https://doi.org/10.1080/10635150050207429>.

Barnes, D.K.A. & Peck, L.S. (2008). Vulnerability of Antarctic shelf biodiversity to predicted regional warming. *Climate Research* 37: 149-163. <https://doi.org/10.3354/cr00760>.

Barrera-Oro, E. (2002). The role of fish in the Antarctic marine food web: differences between inshore and offshore waters in the southern Scotia Arc and west Antarctic Peninsula. *Antarctic Science* 14: 293-309. <https://doi.org/10.1017/S0954102002000111>.

Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57: 289-300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.

Bilyk, K.T. & DeVries, A.L. (2010). Freezing avoidance of the Antarctic icefishes (Channichthyidae) across thermal gradients in the Southern Ocean. *Polar Biology* 33: 203-213. <https://doi.org/10.1007/s00300-009-0697-z>.

Carvajal-Rodriguez, A. (2018). Myriads: p-value-based multiple testing correction. *Bioinformatics* 34: 1043-1044. <https://doi.org/10.1093/bioinformatics/btx746>.

Clarke, A. & Johnston, N. M. (2003). Antarctic marine benthic diversity. *Oceanography and Marine Biology Annual Review* 41: 47-114.

Clement, M., Snell, Q., Walker, P., Posada, D. & Crandall, K. (2002). TCS: estimating gene genealogies. In *Parallel and Distributed Processing Symposium, International* (Vol. 3, pp. 0184-0184). IEEE Computer Society.

Colombo, M., Damerau, M., Hanel, R., Salzburger, W. & Matschiner, M. (2015). Diversity and disparity through time in the adaptive radiation of Antarctic notothenioid fishes. *Journal of Evolutionary Biology* 28: 376-394. <https://doi.org/10.1111/jeb.12570>.



Coombs, J.A., Letcher, B. H. & Nislow, K. H. (2008). CREATE: a software to create input files from diploid genotypic data for 52 genetic software programs. *Molecular Ecology Resources* 8: 578-580. <https://doi.org/10.1111/j.1471-8286.2007.02036.x>.

Curcio, N., Tombari, A. & Capitano, F. (2013). Otolith morphology and feeding ecology of an Antarctic nototheniid, *Lepidonotothen larseni*. *Antarctic Science* 26: 124-132. <https://doi.org/10.1017/S0954102013000394>.

Damerau, M., Matschiner, M., Salzburger, W. & Hanel, R. (2012). Comparative population genetics of seven notothenioid fish species reveals high levels of gene flow along ocean currents in the southern Scotia Arc, Antarctica. *Polar Biology* 35: 1073-1086. <https://doi.org/10.1007/s00300-012-1155-x>.

Damerau, M., Salzburger, W. & Hanel, R. (2014). Population genetic structure of *Lepidonotothen larseni* revisited: cyb and microsatellites suggest limited connectivity in the Southern Ocean. *Marine Ecology Progress Series* 51: 251-263. <https://doi.org/10.3354/meps11061>.

De Broyer, C., Clarke, A., Koubbi, P., Pakhomov, E., Scott, F., Vanden Berghe, E. & Danis, B. (2021). Register of Antarctic Marine Species. <http://www.marinespecies.org/rams> on 2021-12-22.

Deli Antoni, M.Y., Delpiani, S.M., González-Castro, M., Blasina, G.E., Spath, M.C., Delpiani, G.E., Ashikaga, F.Y., Cruz, V.P., Oliveira, C. & Díaz de Astarloa, J.M. (2019). Comparative populational study of *Lepidonotothen larseni* and *L. nudifrons* (Teleostei: Nototheniidae) from the Antarctic Peninsula and the South Shetland Islands, Antarctica. *Polar Biology* 42: 1537-1547. <https://doi.org/10.1007/s00300-019-02540-1>.

Detrich, H.W., Parker, S.K., Williams, R.C., Nogales, E. & Downing, K.H. (2000). Cold Adaptation of Microtubule Assembly and Dynamics. *The*

*Journal of Biological Chemistry* 275: 37038-37047.  
<https://doi.org/10.1074/jbc.M005699200>.

DeVries, A.L. & Steffensen, J.F. (2005). The Arctic and Antarctic Polar Marine Environments. *Fish Physiology* 22: 1-24.  
[https://doi.org/10.1016/S1546-5098\(04\)22001-5](https://doi.org/10.1016/S1546-5098(04)22001-5).

Dornburg, A., Federman, S., Eytan, R. & Near, T.J. (2016). Cryptic species diversity in sub-Antarctic islands: A case study of *Lepidonotothen*. *Molecular Phylogenetics and Evolution* 104: 32-43.  
<https://doi.org/10.1016/j.ympev.2016.07.013>.

Eastman, J.T. (1993). *Antarctic Fish Biology: evolution in a unique environment*. Academic Press, San Diego, CA.

Eastman, J.T. (2005). The nature of the diversity of Antarctic fishes. *Polar Biology* 28: 93-107. <https://doi.org/10.1007/s00300-004-0667-4>.

Eastman, J.T. & Eakin R.R. (2000). An updated species list for notothenioid fish (Perciformes; Notothenioidei), with comments on Antarctic specie. *Archive of Fishery and Marine Research* 48: 11-20.

Eastman, J.T. & Eakin, R.R. (2021). Checklist of the species of notothenioid fishes. *Antarctic Science* 33: 273-280.  
<https://doi.org/10.1017/S0954102020000632>.

Eastman, J. & McCune, A. (2000). Fishes on the Antarctic continental shelf: evolution of a marine species flock? *Journal of Fish Biology* 57: 84-102. <https://doi.org/10.1111/j.1095-8649.2000.tb02246.x>.

Evanno, G., Regnaut, S. & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular ecology* 14: 2611-2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>.

Evans, C.W., Hellman, L., Middleditch, M., Wojnar, JM., Brimble, M.A. & Devries, A. (2012). Synthesis and recycling of antifreeze glycoproteins in polar fishes. *Antarctic Science* 24: 259-268. <https://doi.org/10.1017/S0954102012000119>.

Excoffier, L. & Lischer, H.E.L. (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564-567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>.

Fisher, R.A. (1930). *The Genetical Theory of Natural Selection*. Oxford University Press.

Fischer, W. & Hureau, J.C. (1985). *FAO species identification sheets for fishery purposes Southern Ocean: Fishing Areas 48, 58 and 88 (CCAMLR Convention Area)*. Prepared and published with the support of the Commission for the Conservation of Antarctic Marine Living Resources. Rome, FAO 1985. Vol. 2: 233-470.

Foulley, J.L. & Ollivier, L. (2006). Estimating allelic richness and its diversity. *Livestock Science* 101: 150-158. <https://doi.org/10.1016/j.livprodsci.2005.10.021>.

Gavrilets, S. & Losos, J.B. (2009). Adaptive Radiation: Contrasting Theory with Data. *Science* 323: 732-737. DOI: 10.1126/science.1157966.

Gon, O. & Heemstra, P.C. (1990). *Fishes of the Southern Ocean*. J.L.B. smith Institute of Ichthyology, Grahamstown.

Grant, W.A.S. & Bowen, B.W. (1998). Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and

anchovies and lessons for conservation. *Journal of heredity* 89: 415-426.  
<https://doi.org/10.1093/jhered/89.5.415>.

Grant, P.R. & Grant, R. (2002). Adaptive Radiation of Darwin's Finches: Recent data help explain how this famous group of Galápagos birds evolved, although gaps in our understanding remain. *American Scientist* 90: 130-139.

Griffiths, H. J. (2010). Antarctic Marine Biodiversity – What Do We Know About the Distribution of Life in the Southern Ocean? *PloS one* 5: e11683. <https://doi.org/10.1371/journal.pone.0011683>.

Jones, C.D., Anderson, M.E., Balushkin, A.V., Duhamel, G., Eakin, R.R., Eastman, J.T., Kuhn, K., Lecointre, G., Near, T.J., North, A.W., Stein, D.L., Vacchi, M. & Detrich III, H.W. (2008). Diversity, relative abundance, new locality records and population structure of Antarctic demersal fishes from the northern Scotia Arc islands and Bouvetøya. *Polar Biology* 31: 1481-1497. <https://doi.org/10.1007/s00300-008-0489-x>.

Klinck, J. & Nowlin, W.D. (2001). Southern Ocean. Antarctic Circumpolar Current. *Encyclopedia of Oceanography* 151-159.  
doi:10.1006/rwos.2001.0370.

Kocher, T.D., Conroy, J.A., McKaye, K.R., Stauffer, J.R. & Lockwood, S.F. (1995). Evolution of NADH dehydrogenase subunit 2 in east African cichlid fish. *Molecular Phylogenetics and Evolution* 4: 420-432.  
<https://doi.org/10.1006/mpev.1995.1039>.

Kock, K.H. (1992). *Antarctic fish and fisheries*. Cambridge University Press, Cambridge, UK.

Kock, K.H. & Jones, C.D. (2005). Fish stocks of the southern Scotia Arc region – a review and prospects for future research. *Reviews in Fisheries Science* 13: 75-108. <https://doi.org/10.1080/10641260590953900>.

Kuhn, K.L. & Gaffney, P.M. (2006). Preliminary assessment of population structure in the mackerel icefish (*Champsocephalus gunnari*). *Polar Biology* 29: 927-935. <https://doi.org/10.1007/s00300-006-0134-5>.

La Mesa, M., Eastman, J.T. & Vacchi, M. (2004). The role of notothenioid fish in the food web. *Polar Biology* 27: 321-338. <https://doi.org/10.1007/s00300-004-0599-z>.

Lebedev, S.A. (2006). Interannual and Seasonal Variation of Axis Position and Intensity of the Antarctic Circumpolar Current by Satellite Altimetry. Conference: Proceedings of the Symposium on 15 Years of Progress in Radar Altimetry. Venice, Italy. 13-18 March 2006.

Leberg, P.L. (2002). Estimating allelic richness: effects of sample size and bottlenecks. *Molecular Ecology* 11: 2445-2449. <https://doi.org/10.1046/j.1365-294X.2002.01612.x>.

Leigh, J.W. & Bryant, D. (2015). PopART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution* 6: 1110–1116. doi: 10.1111/2041-210X.12410.

Li, Y., Korol, A.B., Fahima, T., Beiles, A. & Nevo, E. (2002). Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology* 11: 2453-2465. <https://doi.org/10.1046/j.1365-294X.2002.01643.x>.

Llompert, F., Delpiani, M., Lattuca, E., Delpiani, G., Cruz-Jiménez, A., Orlando, P., Ceballos, S., De Astarloa, J.M.D., Vanella, F. & Fernández, D. (2015). Spatial patterns of summer demersal fish assemblages around the Antarctic Peninsula and South Shetland Island. *Antarctic Science* 27: 109-117. <https://doi.org/10.1017/S0954102014000352>.

Losos, J.B. & Miles, D.B. (2002). Testing the Hypothesis That a Clade Has Adaptively Radiated: Iguanid Lizard Clades as a Case Study. *The American Naturalist* 160: 147-157. <https://doi.org/10.1086/341557>.

Marino, I.A.M., Agostini, C., Papetti, C., Bisol, P.M., Zane, L. & Patarnello, T. (2011). Isolation, characterization, and multiplexing of novel microsatellite markers for the tropical scalloped spiny lobster (*Panulirus homarus*). *Molecular Ecology Resources* 11: 418-421. <http://dx.doi.org/10.4238/2015.December.29.14>.

Matschiner, M., Hanel, R. & Salzburger, W. (2009). Gene flow by larval dispersal in the Antarctic notothenioid fish *Gobionotothen gibberifrons*. *Molecular Ecology* 18: 2574-2587. <https://doi.org/10.1111/j.1365-294X.2009.04220.x>.

Matschiner, M., Hanel, R. & Salzburger, W. (2011). On the Origin and Trigger of the Notothenioid Adaptive Radiation. *PloS one* 6: e18911. <https://doi.org/10.1371/journal.pone.0018911>.

Matschiner, M., Colombo, M., Damerau, M., Ceballos, S., Hanel, R. & Salzburger, W. (2015). The Adaptive Radiation of Notothenioid Fishes in the Waters of Antarctica. *Extremophile Fishes* (pp. 35-57). Springer International Publishing.

Miya, T., Gon, O., Mwale, M. & Poulin, E. (2016). Molecular systematics and taxonomic status of three latitudinally widespread nototheniid (Perciformes: Notothenioidei) fishes from the Southern Ocean. *Zootaxa* 4061: 381-396. <https://scite.ai/reports/10.11646/zootaxa.4061.4.4>.

Near, T.J., Dornburg, A., Kuhn, K.L., Eastman, J.T., Pennington, J.N., Patarnello, T., Zane, L., Fernández, D.A. & Jones, C.D. (2012). Ancient climate change, antifreeze, and the evolutionary diversification of Antarctic fishes. *Proceedings of the National Academy of Sciences* 109: 3434-3439. <https://doi.org/10.1073/pnas.1115169109>.

North, A. W. (2001). Early life history strategies of notothenioids at South Georgia. *Journal of Fish Biology* 58: 496-505. <https://doi.org/10.1111/j.1095-8649.2001.tb02268.x>.

Papetti, C., Susana, E., Patarnello, T. & Zane, L. (2009). Spatial and temporal boundaries to gene flow between *Chaenocephalus aceratus* populations at South Orkney and South Shetlands. *Marine Ecology Progress Series* 376: 269-281. <https://doi.org/10.3354/meps07831>.

Papetti, C., Pujolar, J. M., Mezzavilla, M., La Mesa, M., Rock, J., Zane, L. & Patarnello, T. (2012). Population genetic structure and gene flow patterns between populations of the Antarctic icefish *Chionodraco rastrospinosus*. *Journal of Biogeography* 39: 1361-1372. <https://doi.org/10.1111/j.1365-2699.2011.02682.x>.

Papetti, C., Harms, L., Jürgens, J., Sandersfeld, T., Koschnick, N., Windisch, H.S., Knust, R., Pörtner, H.O. & Lucassen, M. (2016). Microsatellite markers for the notothenioid fish *Lepidonotothen nudifrons* and two congeneric species. *BMC Research Notes* 9: 238. <https://doi.org/10.1186/s13104-016-2039-x>.

Patarnello, T., Marcato, S., Zane, L., Varotto, V. & Bargelloni, L. (2003). Phylogeography of the *Chionodraco* genus (Perciformes, Channichthyidae) in the Southern Ocean. *Molecular Phylogenetics and Evolution* 28: 420-429. [https://doi.org/10.1016/S1055-7903\(03\)00124-6](https://doi.org/10.1016/S1055-7903(03)00124-6).

Patwary, M.U., Kenchington, E.L., Bird, C.J. & Zouros, E. (1994). The use of random amplified polymorphic DNA markers in genetic studies of the sea scallop *Placopecten magellanicus*. *Journal of Shellfish Research* 13: 547-553.

Porrás-Hurtado, L., Ruiz, Y., Santos, C., Philips, C., Carracedo, A. & Lareu, M.V. (2013). An overview of STRUCTURE: applications,

parameter setting, and supporting software. *Frontiers in genetics* 4 :1-13. <https://doi.org/10.3389/fgene.2013.00098>.

Pritchard, J., Stephens, M. & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics* 155: 945-959. <https://doi.org/10.1093/genetics/155.2.945>.

Rastogi, P.A. (2000). MacVector. In *Bioinformatics methods and protocols* (pp. 47-69). Humana Press, Totowa, NJ.

Raymond, M. & Rousset, F. (1995). Genepop (Version 1.2): Population genetics software for exact tests and ecumenicism. *J. Heredity* 86: 248-249.

Ribbink, A.J. (1984). *Is the species flock concept tenable?* In: Evolution of fish species flocks. University of Maine at Orono Press, Orono, pp 21–25.

Rogers, A.D. (2007). Evolution and biodiversity of Antarctic organisms: a molecular perspective. *Philosophical Transactions of the Royal Society B* 362: 2191-2214. <https://doi.org/10.1098/rstb.2006.1948>.

Ronco, F., Matschiner, M., Böhna, A., Boila, A., Büscher, H.H., El Taher, A., Indermaur, A., Malinsky, M., Ricci, V., Kahmen, A., Jentoft, S. & Salzburger, W. (2021). Drivers and dynamics of a massive adaptive radiation in cichlid fishes. *Nature* 589: 76-81. <https://doi.org/10.1038/s41586-020-2930-4>.

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

Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E. & Sanchez-Gracia, A. (2017). DnaSP



6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution* 34: 3299-3302. DOI: 10.1093/molbev/msx248.

RStudio Team (2021). RStudio: Integrated Development Environment for R. RStudio, PBC, Boston. <http://www.rstudio.com/>.

Rutschmann, S., Matschiner, M., Damerau, M., Muschick, M., Lehmann, M.F., Hanel, R. & Salzburger, W. (2011). Parallel ecological diversification in Antarctic notothenioid fishes as evidence for adaptive radiation. *Molecular Ecology* 20: 4707-4721. <https://doi.org/10.1111/j.1365-294X.2011.05279.x>.

Sahade, R., Lager, C., Torre, L., Momo, F., Monien, P., Schloss, I., Barnes, D.K.A., Servetto, N., Tarantelli, S., Tatián, M., Zamboni, N. & Abele, D. (2015). Climate change and glacier retreat drive shifts in an Antarctic benthic ecosystem. *Science Advances* 1: e1500050. DOI: 10.1126/sciadv.1500050.

Selkoe, K.A. & Toonen R.J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* 9: 615-629. <https://doi.org/10.1111/j.1461-0248.2006.00889.x>.

Senan, S., Kizhakayil, D., Sasikumar, B. & Sheeja, T.E. (2014). Methods for Development of Microsatellite Markers: An Overview. *Notulae Scientia Biologicae* 6: 1-13. <https://doi.org/10.15835/nsb619199>.

Schluter, D. (2000). The Ecology of Adaptive Radiation. *Oxford University Press, Oxford*.

Schneppenheim, R., Kock, K.H., Duhamel, G. & Janssen, G. (1994). On the taxonomy of the *Lepidonotothen squamifrons* group (Pisces, Perciformes, Notothenioidei). *Archive of Fishery and Marine Research* 422: 137-148.

Seehausen, O. (2006). African cichlid fish: a model system in adaptive radiation research. *Proceedings of the Royal Society B* 273: 1987-1998. <https://doi.org/10.1098/rspb.2006.3539>.

Shaw, P.W., Arkhipkin, A.I & Al-Khairulla, H. (2004). Genetic structuring of Patagonian toothfish populations in the Southwest Atlantic Ocean: the effect of the Antarctic Polar Front and deep-water troughs as barriers to genetic exchange. *Molecular Ecology* 13: 3293-3303. <https://doi.org/10.1111/j.1365-294X.2004.02327.x>.

Slatkin, M. (2008). Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics*, 9: 477-485. <https://doi.org/10.1038/nrg2361>.

Susana, E., Papetti, C., Barbisan, F., Bortolotto, E., Buccoli, S., Patarnello, T. & Zane, L. (2007). Isolation and characterization of eight microsatellite loci in the icefish *Chaenocephalus aceratus* (Perciformes, Notothenioidei, Channichthyidae). *Molecular Ecology Notes* 7: 791-793. <https://doi.org/10.1111/j.1471-8286.2007.01703.x>.

Thatje, S., Hillenbrand, C.-H. & Larter, R. (2005). On the origin of Antarctic marine benthic community structure. *Trends in Ecology and Evolution* 20: 534-540. <https://doi.org/10.1016/j.tree.2005.07.010>.

Tomczak, M. & Godfrey, J.S. (2003). *Regional oceanography: an introduction*. 2nd Edition, Daya Publishing House.

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P. & Shipley, P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535-538. <https://doi.org/10.1111/j.1471-8286.2004.00684.x>.

Vieira, M.L.C., Santini, L., Lima Diniz, A. & Munhoz, C.F. (2016). Microsatellite markers: what they mean and why they are so useful.

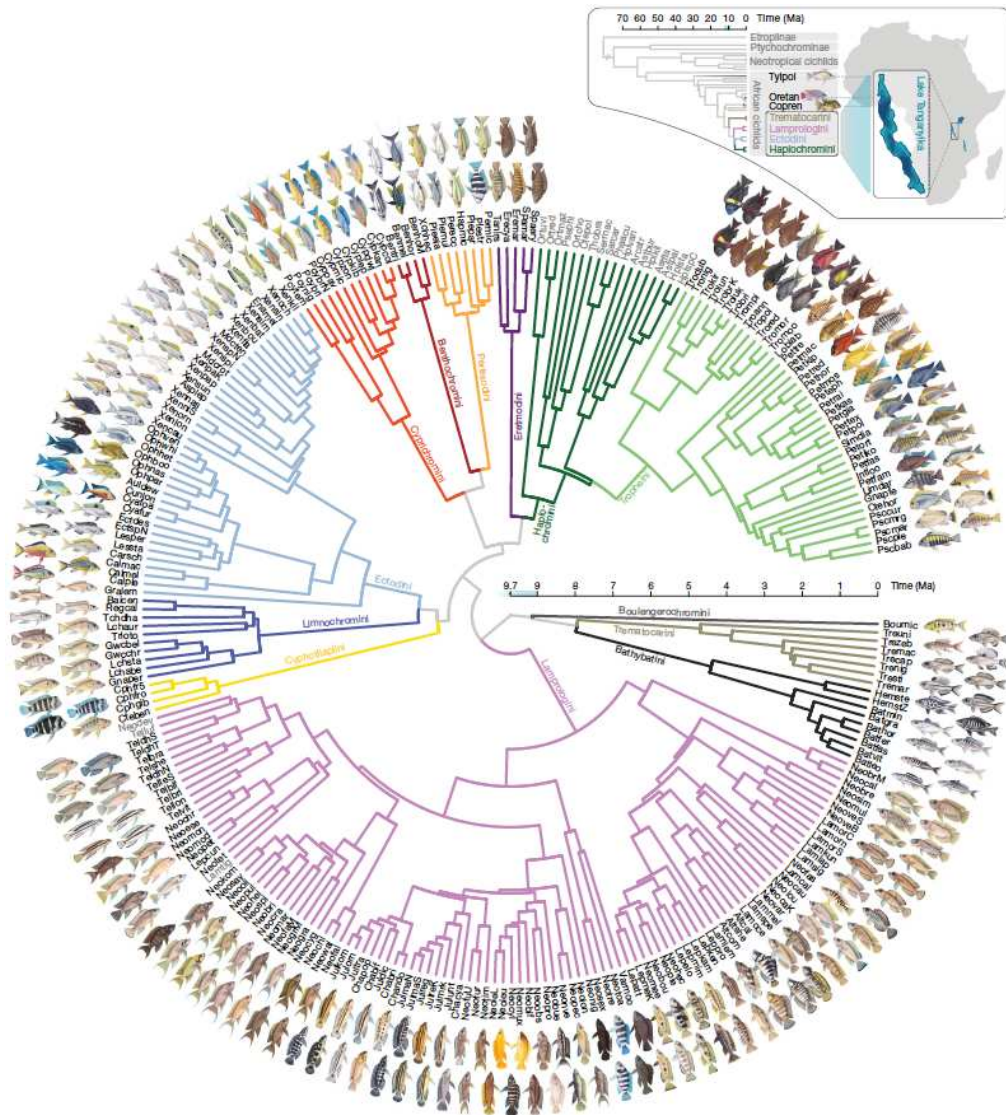
*Genetics and Molecular Biology* 39: 312-328.  
<https://doi.org/10.1590/1678-4685-GMB-2016-0027>.

Yoder, J.B., Clancey, E., Des Roches, S., Eastman, J.M., Gentry, L., Godsoe, W., Hagey, T.J., Jochimsen, D., Oswald, B.O., Robertson, J., Sarver, B.A.J., Schenk, J.J., Spear, S.F. & Harmon, L.J. (2010). Ecological opportunity and the origin of adaptive radiations. *Journal of Evolutionary Biology* 23: 1581-1596. <https://doi.org/10.1111/j.1420-9101.2010.02029.x>.

Zane, L., Bargelloni, L. & Patarnello, T. (2002). Strategies for microsatellite isolation: a review. *Molecular Ecology* 11 :1-16.  
<https://doi.org/10.1046/j.0962-1083.2001.01418.x>.

Zane, L., Marcato, S., Bargelloni, L., Bortolotto, E., Papetti, C., Simionato, M., Varotto, V. & Patarnello, T. (2006). Demographic history and population structure of the Antarctic silverfish *Pleuragramma antarcticum*. *Molecular Ecology* 15: 4499-4511.  
<https://doi.org/10.1111/j.1365-294X.2006.03105.x>.

## 8. APPENDICE



**Figura 8.1:** Albero filogenetico calibrato nel tempo dei pesci ciclidi del Lago Tanganica. I colori dei rami dell'albero rappresentano le diverse tribù. I nomi delle specie sono abbreviati utilizzando un codice di sei lettere; le prime tre lettere indicano il genere mentre le ultime tre lettere indicano la specie (Ronco *et al.*, 2021).

Specie	Statistiche	Ln41281	Ln40551	Ln42701	Ln42233	Ln36100	Ln35217	Ln23194	Ln22268	Ln36156	Ln45589	Ln45257
<i>L. kempi</i>	N	52	52	52	52	52	52	49	52	52	52	35
	A	3	6	1	2	1	6	2	1	19	13	2
	%	75	18.75	100	66.67	25	60	66.67	25	59.38	65	33.33
	A <sub>r</sub>	2.54	4.92	1	1.55	1	4.88	1.9	1	16.54	10.68	1.8
	H <sub>o</sub>	0.38	0.37	0	0.02	0	0.6	0.06	0	0.94	0.85	0.03
	H <sub>e</sub>	0.38	0.35	0	0.02	0	0.63	0.06	0	0.91	0.79	0.08
<i>L. squamifrons</i>	N	101	101	101	101	101	101	99	101	101	101	100
	A	2	5	1	1	2	7	2	1	27	17	4
	%	50	15.62	100	33.33	50	70	66.67	25	84.38	85	66.67
	A <sub>r</sub>	2	3.56	1	1	1.31	5.3	1.56	1	20.75	10.04	2.59
	H <sub>o</sub>	0.22	0.28	0	0	0.01	0.59	0.02	0	0.86	0.65	0.01
	H <sub>e</sub>	0.25	0.25	0	0	0.01	0.53	0.02	0	0.93	0.7	0.15
<i>N. larseni</i>	N	94	74	94	94	83	94	94	94	94	94	94
	A	2	29	1	2	3	9	1	4	25	5	2
	%	50	90.62	100	66.67	75	90	33.33	100	78.12	25	33.33
	A <sub>r</sub>	1.99	17.84	1	1.88	2.58	6.45	1	2.49	18.36	4.34	2
	H <sub>o</sub>	0.12	0.27	0	0.05	0.06	0.66	0	0.04	0.87	0.54	0.19
	H <sub>e</sub>	0.11	0.94	0	0.05	0.26	0.66	0	0.06	0.9	0.59	0.19
<i>N. nudifrons</i>	N	40	40	40	40	40	40	40	31	40	40	40
	A	1	11	1	1	1	3	2	1	23	5	2
	%	25	34.38	100	33.33	25	30	66.67	25	71.88	25	33.33
	A <sub>r</sub>	1	8.7	1	1	1	3	1.88	1	19.67	4.52	1.88
	H <sub>o</sub>	0	0.58	0	0	0	0.4	0.03	0	0.95	0.65	0.03
	H <sub>e</sub>	0	0.65	0	0	0	0.45	0.07	0	0.93	0.62	0.07

**Tabella 8.2:** statistiche di base per i 21 *loci* microsatellite delle quattro specie analizzate in questa tesi prima della correzione del *dataset*, ottenute mediante la funzione *divBasic* del pacchetto *diveRsity* in RStudio ver. 4.0.5. Sono riportati il numero di individui genotipizzati per ciascun *locus* (N), il numero di alleli osservati per ciascun *locus* (A), la percentuale di alleli totali osservati per campione per locus (%), l'*allelic richness* (A<sub>r</sub>), l'eterozigotità osservata (H<sub>o</sub>) e l'eterozigotità attesa (H<sub>e</sub>).

Specie	Statistiche	Ln42016	Ch126c	Ch3866	Ca35	Ch5817	Ch623	Ch1968	Ch3603	Ch2788	Ca48	Media
<i>L. kempi</i>	N	52	52	52	52	52	52	52	50	52	51	50.9
	A	2	1	2	16	7	2	2	2	5	1	96
	%	66.67	20	25	61.54	36.84	50	50	66.67	41.67	2.27	48.36
	A <sub>r</sub>	1.98	1	1.77	13.83	6.76	1.57	1.57	1.53	4	1	3.94
	H <sub>o</sub>	0	0	0.04	0.9	0.48	0	0	0	0.17	0	0.23
	H <sub>e</sub>	0.17	0	0.04	0.86	0.71	0.04	0.04	0.04	0.18	0	0.25
<i>L. squamifrons</i>	N	101	101	100	99	99	101	101	101	101	101	100.62
	A	2	3	6	19	10	1	2	2	5	1	120
	%	66.67	60	75	73.08	52.63	25	50	66.67	41.67	2.27	55.22
	A <sub>r</sub>	2	2.11	3.72	14.6	7.56	1	1.89	1.68	3.95	1	4.27
	H <sub>o</sub>	0	0.04	0.12	0.89	0.6	0	0.05	0.03	0.23	0	0.22
	H <sub>e</sub>	0.27	0.04	0.12	0.89	0.77	0	0.05	0.03	0.23	0	0.25
<i>N. larseni</i>	N	94	94	94	94	94	94	94	93	94	93	92.43
	A	3	3	8	24	11	2	2	2	6	41	185
	%	100	60	100	92.31	57.89	50	50	66.67	50	93.18	69.62
	A <sub>r</sub>	2.57	2.15	5.39	19.03	8.55	2	2	2	5.18	26.42	6.44
	H <sub>o</sub>	0.02	0.05	0.65	0.87	0.48	0.38	1	0.23	0.61	0.8	0.38
	H <sub>e</sub>	0.47	0.05	0.63	0.92	0.65	0.47	0.5	0.34	0.68	0.94	0.45
<i>N. nudifrons</i>	N	40	40	40	39	40	40	40	38	39	40	39.38
	A	3	1	4	9	2	1	3	1	3	14	92
	%	100	20	50	34.62	10.53	25	75	33.33	25	31.82	41.66
	A <sub>r</sub>	3	1	3.65	8.16	2	1	2.29	1	2.96	12.09	3.9
	H <sub>o</sub>	0.32	0	0.72	0.74	0.15	0	0.03	0	0.44	0.75	0.28
	H <sub>e</sub>	0.62	0	0.66	0.8	0.18	0	0.07	0	0.41	0.82	0.3

**Tabella 8.2 (continuazione):** statistiche di base per i 21 loci microsatellite delle quattro specie analizzate in questa tesi prima della correzione del *dataset*, ottenute mediante la funzione *divBasic* del pacchetto *diveRsity* in RStudio ver. 4.0.5. Sono riportati il numero di individui genotipizzati per ciascun locus (N), il numero di alleli osservati per ciascun locus (A), la percentuale di alleli totali osservati per campione per locus (%), l'*allelic richness* (A<sub>r</sub>), l'eterozigotità osservata (H<sub>o</sub>) e l'eterozigotità attesa (H<sub>e</sub>).

Specie	Pop.	Ln 41281	Ln 40551	Ln 42701	Ln 42233	Ln 36100	Ln 35217	Ln 23194	Ln 22268	Ln 36156	Ln 45589	Ln 45257	Ln 42016	Ch 126	Ch 3866	Ca 35	Ch 5817	Ch 623	Ch 1968	Ch 3603	Ch 2788	Ca 48	
<i>N. larseni</i>	SS		AN										AN			AN							
	DP		AN			AN							AN				AN						
	BV		AN														AN			AN		AN	
<i>N. nudifrons</i>	SS												AN										
	EI												AN										
	SO																						
	DP																						
	SG																						
<i>L. kempii</i>	BS												AN				AN						
	BV																AN						
	SR																						
<i>L. squamifrons</i>	SS											AN	AN				AN						
	SR											AN											
	BV																						
	NMS																						

**Tabella 8.3:** Presenza di alleli nulli nei 21 *loci* microsatellite in ciascuna popolazione delle quattro specie analizzate in questa tesi. Le specie sono suddivise in popolazioni (SS=South Shetland; DP= Drake Passage; BV= Bouvet Island; EI= Elephant Island; SO= South Orkney; BS= Bransfield Strait; SR= Shag Rocks; SG= South Georgia; NMS= Neumayer Station). Per le popolazioni evidenziate in rosso non è stato possibile ottenere risultati in quanto composte da un numero troppo ristretto di campioni (1/2 individui). I *loci* caratterizzati dalla presenza di alleli nulli sono indicati con AN.

Specie	Statistiche	Ln41281	Ln42233	Ln36100	Ln35217	Ln23194	Ln36156	Ln45589	Ln45257	Ch126	Ch3866	Ca35	Ch623	Ch1968	Ch3603	Ch2788	Ca48	Media
<i>L. kempii</i>	N	52	52	52	52	49	52	52	35	52	52	52	52	52	50	52	51	50.56
	A	3	2	1	6	2	19	13	2	1	2	16	2	2	2	5	1	79
	%	75	66.67	25	60	66.67	59.38	65	33.33	20	25	61.54	50	50	66.67	41.67	2.27	48.01
	Ar	2.52	1.54	1	4.85	1.92	16.52	10.64	1.78	1	1.81	13.79	1.54	1.54	1.53	4.01	1	4.19
	Ho	0.38	0.02	0	0.6	0.06	0.94	0.85	0.03	0	0.04	0.9	0	0	0	0.17	0	0.25
	He	0.38	0.02	0	0.63	0.06	0.91	0.79	0.08	0	0.04	0.86	0.04	0.04	0.04	0.18	0	0.25
<i>L. squamifrons</i>	N	101	101	101	101	99	101	101	100	101	100	99	101	101	101	101	101	100.62
	A	2	1	2	7	2	27	17	4	3	6	19	1	2	2	5	1	101
	%	50	33.33	50	70	66.67	84.38	85	66.67	60	75	73.08	25	50	66.67	41.67	2.27	56.23
	Ar	2	1	1.33	5.27	1.57	20.85	10.2	2.57	2.13	3.7	14.59	1	1.86	1.69	3.94	1	4.67
	Ho	0.22	0	0.01	0.59	0.02	0.86	0.65	0.01	0.04	0.12	0.89	0	0.05	0.03	0.23	0	0.23
	He	0.25	0	0.01	0.53	0.02	0.93	0.7	0.15	0.04	0.12	0.89	0	0.05	0.03	0.23	0	0.25
<i>N. larseni</i>	N	94	94	83	94	94	94	94	94	94	94	94	94	94	93	94	93	93.19
	A	2	2	3	9	1	25	5	2	3	8	24	2	2	2	6	41	137
	%	50	66.67	75	90	33.33	78.12	25	33.33	60	100	92.31	50	50	66.67	50	93.18	63.35
	Ar	1.99	1.89	2.58	6.44	1	18.33	4.36	2	2.15	5.48	19.11	2	2	2	5.16	26.58	6.44
	Ho	0.12	0.05	0.06	0.66	0	0.87	0.54	0.19	0.05	0.65	0.87	0.38	1	0.23	0.61	0.8	0.44
	He	0.11	0.05	0.26	0.66	0	0.9	0.59	0.19	0.05	0.63	0.92	0.47	0.5	0.34	0.68	0.94	0.46
<i>N. nudifrons</i>	N	40	40	40	40	40	40	40	40	40	40	39	40	40	38	39	40	39.75
	A	1	1	1	3	2	23	5	2	1	4	9	1	3	1	3	14	74
	%	25	33.33	25	30	66.67	71.88	25	33.33	20	50	34.62	25	75	33.33	25	31.82	37.81
	Ar	1	1	1	3	1.88	19.67	4.53	1.88	1	3.64	8.12	1	2.3	1	2.95	12.14	4.13
	Ho	0	0	0	0.4	0.03	0.95	0.65	0.03	0	0.72	0.74	0	0.03	0	0.44	0.75	0.3
	He	0	0	0	0.45	0.07	0.93	0.62	0.07	0	0.66	0.8	0	0.07	0	0.41	0.82	0.31

**Tabella 8.4:** Statistiche di base per i 16 *loci* microsatellite delle quattro specie analizzate in questa tesi dopo la correzione del *dataset*, ottenute mediante la funzione *divBasic* del pacchetto *diversity* in RStudio ver. 4.0.5. Sono riportati il numero di individui genotipizzati per ciascun *locus* (N), il numero di alleli osservati per ciascun *locus* (A), la percentuale di alleli totali osservati per campione per locus (%), l'*allelic richness* (Ar), l'eterozigotà osservata (H<sub>o</sub>) e l'eterozigotà attesa (H<sub>e</sub>).



Allele	<i>L. kempi</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
4	0.010	0.000	0.000	0.000
5	0.750	0.851	0.941	1.000
7	0.000	0.000	0.059	0.000
8	0.240	0.149	0.000	0.000

**Tabella 8.5:** Frequenze alleliche del *locus* Ln41287.

Allele	<i>L. kempi</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
1	0.010	0.000	0.973	1.000
2	0.990	1.000	0.000	0.000
15	0.000	0.000	0.027	0.000

**Tabella 8.6:** Frequenze alleliche del *locus* Ln42233.

Allele	<i>L. kempi</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
2	0.000	0.000	0.012	0.000
9	0.000	0.005	0.000	0.000
11	0.000	0.000	0.139	0.000
14	1.000	0.995	0.849	1.000

**Tabella 8.7:** Frequenze alleliche del *locus* Ln36100.

Allele	<i>L. kempi</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
2	0.010	0.005	0.011	0.000
3	0.000	0.000	0.005	0.000
4	0.000	0.000	0.005	0.113
5	0.250	0.193	0.261	0.713
6	0.519	0.649	0.074	0.175
7	0.192	0.104	0.500	0.000
8	0.019	0.015	0.117	0.000
9	0.010	0.030	0.021	0.000
10	0.000	0.000	0.005	0.000
11	0.000	0.005	0.000	0.000

**Tabella 8.8:** Frequenze alleliche del *locus* Ln35217.

Allele	<i>L. kempii</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
4	0.031	0.010	0.000	0.000
7	0.969	0.990	0.000	0.963
8	0.000	0.000	1.000	0.037

**Tabella 8.9:** Frequenze alleliche del *locus* Ln23194.

Allele	<i>L. kempii</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
2	0.000	0.000	0.005	0.013
3	0.000	0.000	0.000	0.025
4	0.000	0.000	0.000	0.025
5	0.000	0.005	0.011	0.013
6	0.038	0.050	0.021	0.000
7	0.019	0.015	0.000	0.000
8	0.010	0.015	0.027	0.050
9	0.183	0.158	0.229	0.013
10	0.048	0.025	0.069	0.113
11	0.067	0.079	0.122	0.062
12	0.067	0.040	0.037	0.075
13	0.000	0.010	0.106	0.025
14	0.087	0.045	0.074	0.013
15	0.096	0.069	0.053	0.000
16	0.077	0.064	0.032	0.000
17	0.038	0.040	0.053	0.000
18	0.038	0.064	0.011	0.000
19	0.106	0.074	0.005	0.163
20	0.029	0.074	0.021	0.037
21	0.029	0.054	0.043	0.013
22	0.010	0.015	0.016	0.013
23	0.000	0.010	0.021	0.025
24	0.000	0.010	0.000	0.050
25	0.010	0.015	0.005	0.062
26	0.000	0.005	0.011	0.062
27	0.000	0.010	0.011	0.062
28	0.000	0.020	0.005	0.025
29	0.038	0.020	0.005	0.000
30	0.000	0.000	0.000	0.050
31	0.010	0.010	0.000	0.013
32	0.000	0.005	0.000	0.000
33	0.000	0.000	0.005	0.000

**Tabella 8.10:** Frequenze alleliche del *locus* Ln36156.

Allele	<i>L. kempfi</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
1	0.000	0.005	0.000	0.000
2	0.010	0.020	0.000	0.000
3	0.173	0.376	0.202	0.000
4	0.375	0.386	0.590	0.125
5	0.115	0.069	0.128	0.338
6	0.019	0.005	0.074	0.000
8	0.000	0.000	0.005	0.000
9	0.000	0.005	0.000	0.025
10	0.000	0.000	0.000	0.500
13	0.029	0.005	0.000	0.000
14	0.010	0.005	0.000	0.013
17	0.000	0.005	0.000	0.000
18	0.029	0.010	0.000	0.000
19	0.029	0.005	0.000	0.000
20	0.048	0.045	0.000	0.000
21	0.144	0.020	0.000	0.000
22	0.000	0.015	0.000	0.000
23	0.010	0.020	0.000	0.000
24	0.000	0.005	0.000	0.000
27	0.010	0.000	0.000	0.000

Tabella 8.11: Frequenze alleliche del *locus* Ln45589.

Allele	<i>L. kempfi</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
1	0.000	0.010	0.000	0.000
3	0.957	0.920	0.000	0.000
8	0.000	0.005	0.000	0.000
10	0.043	0.065	0.000	0.000
13	0.000	0.000	0.894	0.963
14	0.000	0.000	0.106	0.037

Tabella 8.12: Frequenze alleliche del *locus* Ln45257.

Allele	<i>L. kempfi</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
1	0.000	0.000	0.005	0.000
2	0.000	0.010	0.000	0.000
3	0.000	0.000	0.973	1.000
4	1.000	0.980	0.021	0.000
5	0.000	0.010	0.000	0.000

Tabella 8.13: Frequenze alleliche del *locus* Ch126.

Allele	<i>L. kemp</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
1	0.000	0.005	0.005	0.000
2	0.000	0.010	0.011	0.000
3	0.000	0.005	0.011	0.013
4	0.981	0.9	0.431	0.000
5	0.019	0.030	0.415	0.375
6	0.000	0.010	0.112	0.225
7	0.000	0.000	0.011	0.388
12	0.000	0.000	0.005	0.000

Tabella 8.14: Frequenze alleliche del *locus* Ch3866.

Allele	<i>L. kemp</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
4	0.000	0.000	0.048	0.000
6	0.000	0.000	0.005	0.026
7	0.000	0.005	0.021	0.000
8	0.000	0.000	0.011	0.244
9	0.067	0.000	0.149	0.000
10	0.000	0.010	0.080	0.295
11	0.029	0.056	0.122	0.000
12	0.010	0.000	0.080	0.154
13	0.279	0.177	0.032	0.000
14	0.019	0.015	0.101	0.128
15	0.087	0.056	0.037	0.000
16	0.019	0.010	0.021	0.077
17	0.087	0.131	0.059	0.000
18	0.000	0.010	0.011	0.051
19	0.144	0.162	0.032	0.000
20	0.000	0.005	0.005	0.013
21	0.115	0.121	0.037	0.000
22	0.000	0.020	0.032	0.013
23	0.048	0.091	0.048	0.000
24	0.000	0.005	0.005	0.000
25	0.019	0.035	0.027	0.000
27	0.019	0.051	0.016	0.000
29	0.029	0.015	0.016	0.000
31	0.019	0.025	0.000	0.000
33	0.000	0.000	0.005	0.000
37	0.010	0.000	0.000	0.000

Tabella 8.15: Frequenze alleliche del *locus* Ca35.

Alelle	<i>L. kempfi</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
1	0.000	0.000	0.383	0.000
2	0.000	0.000	0.617	1.000
7	0.019	0.000	0.000	0.000
9	0.981	1.000	0.000	0.000

**Tabella 8.16:** Frequenze alleliche del *locus* Ch623.

Alelle	<i>L. kempfi</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
1	0.000	0.000	0.000	0.013
2	0.019	0.000	0.500	0.963
4	0.981	0.975	0.500	0.025
5	0.000	0.025	0.000	0.000

**Tabella 8.17:** Frequenze alleliche del *locus* Ch1968.

Alelle	<i>L. kempfi</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
1	0.020	0.000	0.780	0.000
3	0.980	0.985	0.220	1.000
4	0.000	0.015	0.000	0.000

**Tabella 8.18:** Frequenze alleliche del *locus* Ch3603.

Alelle	<i>L. kempfi</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
1	0.000	0.010	0.000	0.000
2	0.010	0.025	0.000	0.000
3	0.0385	0.000	0.000	0.731
4	0.904	0.871	0.000	0.038
5	0.010	0.010	0.000	0.231
6	0.038	0.084	0.000	0.000
7	0.000	0.000	0.250	0.000
8	0.000	0.000	0.021	0.000
9	0.000	0.000	0.181	0.000
10	0.000	0.000	0.473	0.000
11	0.000	0.000	0.069	0.000
12	0.000	0.000	0.005	0.000

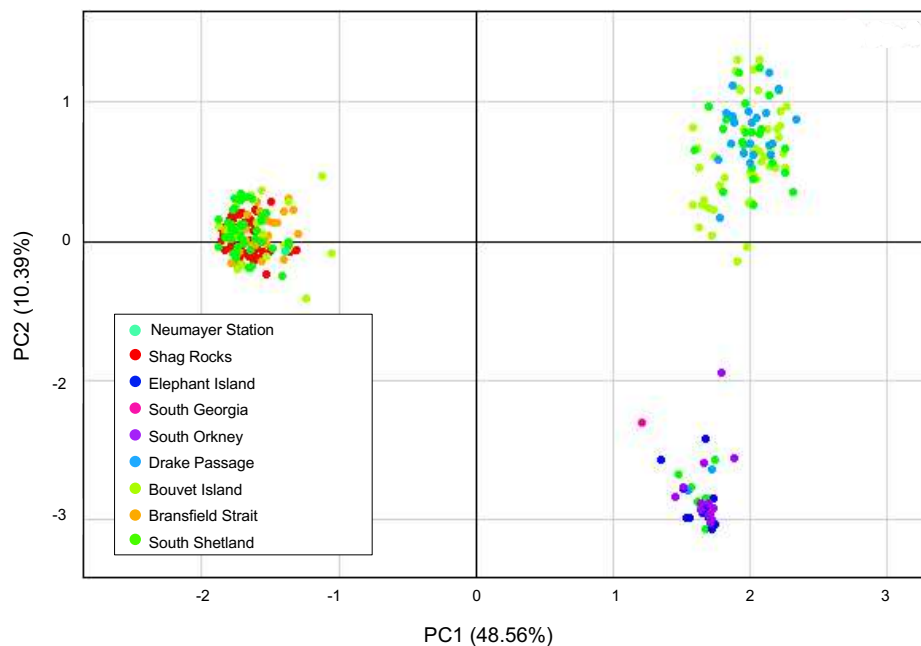
**Tabella 8.19:** Frequenze alleliche del *locus* Ch2788.

Allele	<i>L. kempii</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
12	0.000	0.000	0.005	0.000
13	0.000	0.000	0.000	0.287
14	0.000	0.000	0.000	0.050
15	1.000	1.000	0.005	0.000
16	0.000	0.000	0.016	0.000
17	0.000	0.000	0.005	0.025
18	0.000	0.000	0.043	0.025
19	0.000	0.000	0.054	0.237
20	0.000	0.000	0.027	0.025
21	0.000	0.000	0.145	0.000
22	0.000	0.000	0.011	0.000
23	0.000	0.000	0.081	0.037
24	0.000	0.000	0.070	0.163
25	0.000	0.000	0.048	0.050
26	0.000	0.000	0.043	0.013
27	0.000	0.000	0.027	0.025
28	0.000	0.000	0.059	0.013
29	0.000	0.000	0.059	0.000
30	0.000	0.000	0.038	0.000
31	0.000	0.000	0.011	0.037
32	0.000	0.000	0.016	0.000
33	0.000	0.000	0.016	0.000
34	0.000	0.000	0.000	0.013
35	0.000	0.000	0.022	0.000
36	0.000	0.000	0.038	0.000
37	0.000	0.000	0.005	0.000
38	0.000	0.000	0.011	0.000
39	0.000	0.000	0.011	0.000
40	0.000	0.000	0.011	0.000
41	0.000	0.000	0.027	0.000
42	0.000	0.000	0.005	0.000
43	0.000	0.000	0.005	0.000
44	0.000	0.000	0.005	0.000
46	0.000	0.000	0.011	0.000
47	0.000	0.000	0.011	0.000
49	0.000	0.000	0.005	0.000
53	0.000	0.000	0.005	0.000
54	0.000	0.000	0.005	0.000
57	0.000	0.000	0.011	0.000
59	0.000	0.000	0.011	0.000

**Tabella 8.20:** Frequenze alleliche del *locus* Ca48.

Allele	<i>L. kemp</i>	<i>L. squamifrons</i>	<i>N. larseni</i>	<i>N. nudifrons</i>
62	0	0	0.005	0
63	0	0	0.005	0
68	0	0	0.005	0
88	0	0	0.005	0

Tabella 8.20 (continuazione): Frequenze alleliche del locus Ca48.



**Figura 8.1:** Analisi delle componenti principali (PCA) svolta a partire dal *dataset* relativo ai 16 *loci* microsatellite. L'analisi è stata svolta in R tramite i pacchetti *adegenet ver. 2.1.4* per il calcolo e *ggplot2 ver. 4.0.5* per la visualizzazione grafica. Ogni punto rappresenta un individuo appartenente ad una delle quattro specie analizzate in questa tesi. I diversi colori indicano il sito di campionamento (Shag Rocks: rosso, Bransfield Strait: arancione, Bouvet Island: verde chiaro, South Shetland: verde, Neumayer Station: verde acqua, Drake Passage: azzurro, Elephant Island: blu, South Orkney: viola e South Georgia: rosa).