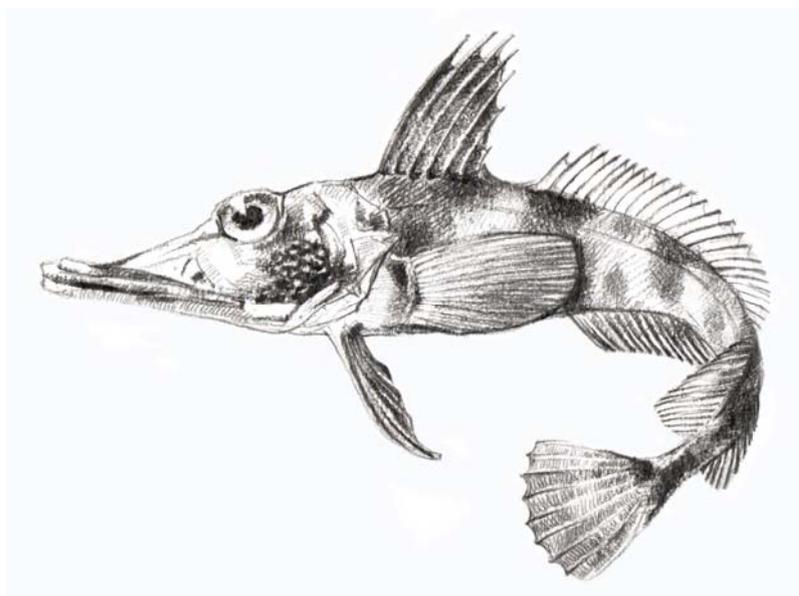




UNIVERSITÀ DEGLI STUDI DI PADOVA
Facoltà di SCIENZE MM. FF. NN.
Corso di Laurea in BIOLOGIA MARINA

DIFFERENZIAMENTO GENETICO IN
Chaenocephalus aceratus
NELL'INTERO RANGE DI DISTRIBUZIONE



Relatore: **LORENZO ZANE**, dipartimento di Biologia
Correlatore: **CHIARA PAPETTI**, dipartimento di Biologia

Laureanda: **MARYNA BARODZICH**

Anno Accademico: **2008/2009**

INDICE

Introduzione	pag. 1
L'ambiente antartico	3
Connettività nello <i>Scotia Sea</i>	5
Struttura di popolazione in <i>Scotia Sea icefish</i>	7
<i>Notothenioidei</i>	8
Channichthyidae, o <i>icefish</i>	9
<i>Chaenocephalus aceratus</i>	10
Scopo della tesi	12
Materiali e Metodi	13
Campionamento	13
Estrazione del DNA	13
Marcatori microsatellite: vantaggi e svantaggi	16
Marcatori microsatellite utilizzati per questo studio	19
Condizioni di amplificazione in multiplex	21
Analisi dei frammenti	23
<i>Allele calling</i> o <i>binning</i>	23
Analisi dei dati	24
Variabilità genetica, equilibrio di Hardy-Weinberg e <i>linkage disequilibrium</i>	25
Differenziamento genetico	27
Risultati	31
Estrazione del DNA	31
Amplificazione in multiplex	31
<i>Allele calling</i> o <i>binning</i>	32
Analisi dei dati	33
Variabilità genetica, equilibrio di Hardy-Weinberg e <i>linkage disequilibrium</i>	33
Alleli nulli ed errori di <i>scoring</i>	38
Differenziamento genetico	39
Tassi di <i>inbreeding</i> F_{IS}	39
Test di differenziamento fra popolazioni	40
Indici di differenziamento F_{ST} e F'_{ST}	41
Discussione	43
Ringraziamenti	48
Bibliografia	49

INTRODUZIONE

Lo studio della connettività tra popolazioni di organismi marini e del differenziamento delle popolazioni sulla scala spazio-temporale è di vitale importanza per la comprensione dei meccanismi che stanno alla base della loro distribuzione e abbondanza. L'utilizzo di marcatori molecolari ha rivoluzionato l'idea classica che gli organismi marini, altamente mobili e caratterizzati da un'elevata capacità di dispersione durante la fase pelagica, siano caratterizzati da popolazioni essenzialmente omogenee (Ward et al., 1994), contribuendo alla scoperta dell'esistenza di strutturazione genetica delle popolazioni ittiche e di una dimensione effettiva degli stock nettamente inferiore alle dimensioni fornite da censimento.

Recenti studi genetici hanno, infatti, dimostrato la suddivisione delle popolazioni nei pesci marini su una ridotta scala geografica che varia da decine a poche centinaia di chilometri (Nielsen et. al., 2004; Olsen et. al., 2008). Inoltre, dati recenti indicano che le dimensioni effettive di queste popolazioni (N_e), soprattutto per le specie caratterizzate da un'elevata fecondità ed elevata mortalità larvale, sono tipicamente di 2-6 volte più piccole rispetto alle dimensioni delle popolazioni stimate tramite il censimento, suggerendo l'esistenza di dinamiche di reclutamento più complesse di quanto fosse assunto inizialmente (Hauser e Carvalho, 2008).

Il tasso di differenziamento (F_{ST}) rilevato tra popolazioni di specie marine è tuttavia relativamente basso, ad esempio rispetto a quello di specie d'acqua dolce, a causa di una più elevata dimensione effettiva delle popolazioni marine e di una loro maggiore connettività (cioè maggiore migrazione di individui) (Ward et. al., 1994). Infatti, il differenziamento delle popolazioni ittiche marine rappresenta spesso soltanto l'1 % della variazione totale riscontrata (cioè l'indice $F_{ST} < 0.01$) (Hauser e Carvalho, 2008). Il differenziamento, inoltre, non avviene solo su scala spaziale, ma anche su quella temporale, che può essere più o meno ampia. La variazione della struttura di popolazione lungo un continuum temporale può dipendere da un contatto secondario tra popolazioni isolate, da una recente divisione di una popolazione in differenti sottopopolazioni, oppure da un evento di ricolonizzazione (effetto fondatore) (Hauser e Carvalho, 2008). La variazione su scala temporale può inoltre essere dovuta a fluttuazioni annuali delle condizioni oceanografiche che risultano in considerevoli variazioni nella dispersione larvale (Mitarai et al., 2008).

È noto che il differenziamento genetico in popolazioni marine caratterizzate da una vita sedentaria è guidato essenzialmente dalla dispersione passiva durante la fase pelagica. Le caratteristiche ecologiche della specie e l'oceanografia su scala locale e globale risultano quindi fondamentali nel determinare il loro *pattern* di dispersione. Tuttavia, studi sempre più numerosi evidenziano che la dispersione

reale su scala geografica di una specie ittica marina è quasi sempre inferiore a quella attesa in base alle caratteristiche intrinseche come durata dello stadio pelagico (larvale o giovanile) e circolazione regionale.

La distanza di dispersione durante la fase pelagica può essere stimata mediante l'utilizzo di marcatori molecolari, calcolando l'indice F_{ST} e applicando il modello *stepping stone*. L'assunzione di questo modello consiste nel fatto che popolazioni scambino immigranti solo con popolazioni immediatamente adiacenti. In questo modo, le distanze genetiche aumentano linearmente con le distanze geografiche e diventa possibile spiegare l'esistenza di differenziamento genetico tra popolazioni distanti, anche quando quelle adiacenti rimangono indistinguibili (a causa di un elevato reciproco flusso genico) (Palumbi, 2003). Questo scenario si osserva, infatti, spesso nelle popolazioni naturali. Il modello *stepping stone* sembra riflettere al meglio l'organizzazione di molte specie marine costiere, per le quali è ragionevole assumere che lo scambio di individui avvenga principalmente tra popolazioni confinanti (Hellberg, 1994). Studi analitici e di simulazione hanno mostrato che, nel caso valga un modello *stepping stone*, le distanze di dispersione larvale si possono stimare attraverso l'utilizzo di distanze genetiche rappresentate dall'indice F_{ST} . Questo approccio ha evidenziato come la distanza di dispersione larvale effettivamente realizzata in natura (misurata in termini di individui reclutati nella popolazione adulta) sia molto inferiore rispetto a quella potenzialmente possibile in base alle condizioni oceanografiche, e per molti pesci risulta compresa fra 20 e 200 km (Palumbi, 2004).

Il modello *stepping stone*, che si basa su un'assunzione relativamente semplice, non può essere applicato a situazioni più complesse, in cui la dispersione larvale non avvenga lungo la linea di costa, ma sia determinata da un pattern oceanografico più complicato. In questo caso la distanza genetica tra popolazioni non aumenta necessariamente in modo lineare all'aumento della distanza geografica, ma può variare in relazione alle condizioni oceanografiche locali.

In questa tesi si analizzerà il differenziamento genetico in *Chaenocephalus aceratus*, una specie antartica caratterizzata da una fase di dispersione larvale e giovanile estremamente lunga (3 anni). L'oceanografia dell'area in cui è distribuita la specie è caratterizzata da una corrente unidirezionale in grado potenzialmente di determinare una relazione *source-sink* tra le diverse popolazioni. L'utilizzo di marcatori molecolari permetterà di verificare se la dispersione potenziale della specie corrisponda a quella effettiva riscontrata dai dati genetici.

L'ambiente Antartico

L'Antartide è un'isola di dimensioni continentali, isolata geograficamente per distanza da altri continenti, e isolata oceanograficamente e termicamente dalle correnti e temperature sotto lo zero (Eastman, 2000). Essa è circondata dall'oceano Australe (Fig. 1) che rappresenta un sistema chiuso, caratterizzato da un *up-welling* di masse d'acqua fredde che si sollevano dal fondo oceanico verso il continente Antartico. L'influenza delle acque di latitudini più basse è esclusa dalla potente Corrente Circumpolare Antartica (CCA), in particolare da uno dei suoi fronti, il Fronte Polare (o *Polar Front*), che forma la cosiddetta Convergenza Antartica. La CCA, oltre a rivestire un ruolo fondamentale nella circolazione globale e avere profonde ricadute sul clima del pianeta, ha agito storicamente come una barriera biologica, favorendo l'evoluzione di forme di vita uniche dal punto di vista degli adattamenti. Questo ambiente freddo e costante, in cui le fluttuazioni della temperatura raramente eccedono 2°C, ha favorito l'evoluzione di molti *taxa* stenotermi, caratterizzati da un bilancio energetico estremamente efficiente entro questo *range* ristretto di temperatura (Portner, 2006).

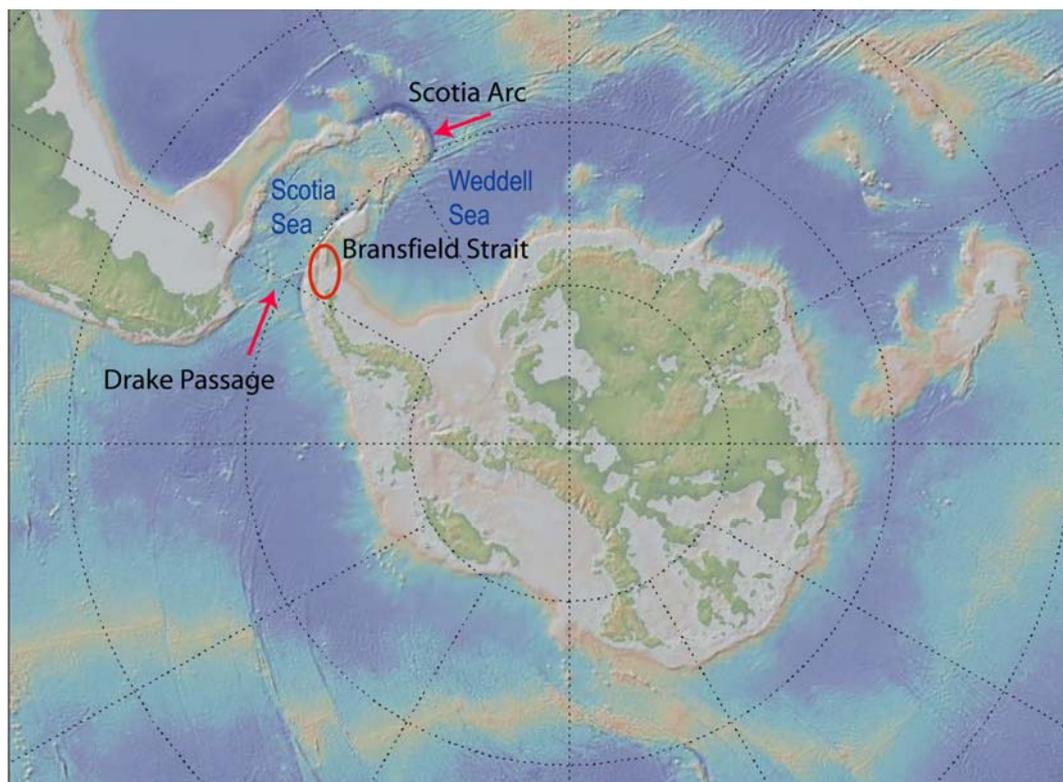


Figura 1. Mappa dei continenti sudamericano e antartico e delle relative piattaforme continentali. Sono mostrate le localizzazioni dello stretto di Bransfield, *Drake Passage* e *Scotia Arc*. Immagine tratta da *Sounds of the Southern Ocean* 2006.

Sicuramente, uno dei *taxon* dell'ittiofauna che è riuscito ad adattarsi meglio all'ambiente antartico è rappresentato dal subordine dei Nototenioidi, che ha subito un'impressionante diversificazione negli ultimi milioni di anni seguendo le dinamiche geologiche e la storia climatica dell'oceano Australe (Eastman, 1993; Chen et al., 1998; Bargelloni et al., 2000a). Attualmente, i Nototenioidi occupano numerose nicchie ecologiche che erano state liberate in seguito all'instaurarsi di temperature glaciali e del definitivo isolamento dell'oceano Australe dalla CCA (Eastman, 1993; Near et al., 2004).

La storia evolutiva dei Nototenioidi risulta, in effetti, strettamente legata agli eventi geologici che si sono susseguiti dopo la rottura del Gondwana (iniziata 167 Ma fa) nei moderni continenti che attualmente occupano l'emisfero Australe. Questi eventi hanno portato, alla fine, all'isolamento del continente Antartico, che ha assunto la sua attuale configurazione e ha raggiunto l'attuale posizione geografica circa 65 Ma fa. In questo stadio, le masse d'acqua tra l'Antartide e la maggior parte degli altri continenti erano ancora poco profonde e le terre emerse erano probabilmente unite da contatti intermittenti.

Negli ultimi 40 Ma, la piattaforma Antartica ha subito una serie di eventi di natura tettonica, climatica e oceanografica, che l'hanno portata all'isolamento dagli altri oceani e hanno contribuito all'instaurarsi di un clima più freddo, provocando una completa sostituzione e impoverimento dell'ittiofauna. L'approfondirsi delle separazioni dalle altre masse continentali e la rimozione delle ultime barriere che impedivano un flusso circumpolare, ha permesso l'instaurarsi della CCA e, sul suo confine settentrionale, del Fronte Polare. Attualmente, la CCA rappresenta un sistema approssimativamente circolare che corre fra 50°S e 60°S e che si estende fino a 2000 m in profondità. Appena a nord del Fronte Polare, la temperatura superficiale dell'acqua sale improvvisamente di circa 3°C, agendo come barriera al flusso generico in entrambe le direzioni, e isolando i processi evolutivi dell'oceano Australe.

Due eventi chiave hanno permesso l'instaurarsi della CCA: l'apertura del passaggio della Tasmania (*Tasman Seaway*), avvenuto approssimativamente 35.5-32 Ma fa (Kenneth, 1977; Lawver e Gahagan, 2003) e l'apertura del *Drake Passage* tra il Sud America meridionale e la Penisola Antartica (PA), la cui datazione è incerta e varia tra 40 e 17 Ma fa. Recentemente, secondo un nuovo sistema di datazione, l'apertura del *Drake Passage* è stata datata a 41 Ma fa (Scher e Martin, 2006) e sarebbe in questo modo antecedente allo stabilirsi del *Tasmania Seaway*.

La datazione dell'instaurazione della CCA è incerta (Barker et al., 2007) e va dal tardo Eocene (40 Ma fa), al tardo Miocene (8-10 Ma fa) (Scher e Martin, 2006). Nonostante le incertezze sull'influenza della CCA sul cambiamento climatico, le ricostruzioni paleoclimatiche lasciano pochi dubbi sul fatto che processi

oceanografici e climatici che hanno portato all'attuale pattern glaciale dell'Antartico, abbiano avuto inizio nello stesso periodo. Infatti, il raffreddamento del clima, iniziato nel tardo Eocene, 42 Ma fa, con una fase iniziale di forti ma brevi glaciazioni, coincide con le più recenti stime della formazione del *Drake Passage* (Scher e Martin, 2006). Un ulteriore drastico cambiamento è avvenuto al confine tra l'Eocene e l'Oligocene, 33-34 Ma fa (Tripathi et al., 2005; Barker et al., 2007), conducendo alla permanente copertura glaciale simile a quella attuale. Quest'ultima datazione "fitta" perfettamente con l'apertura del *Tasmania Seaway*.

Connettività nello Scotia Sea

Il mare di Scozia si estende dal 20° al 60° O e confina con il mare di Weddell a sud, l'Oceano Atlantico Meridionale a nord e a est, e l'oceano Pacifico a ovest del *Drake Passage* (Fig. 2). Il bacino è caratterizzato da una complessa topografia del fondale ed è separato dal mare di Weddell a sud dal *South Scotia Ridge* (Meredith et al., 2008). La circolazione superficiale su grande scala nello *Scotia Sea* è dominata dal flusso in direzione nord-est (orario) della Corrente Circumpolare Antartica (CCA), nella quale sono incorporati getti frontali ad alta velocità (Orsi et al., 1995). I getti frontali costituiscono una parte importante della CCA, dal momento che forniscono velocità al flusso fino a 0.5 m/s, che supera di 2 - 4 volte la velocità tra i getti (Nowlin e Clifford, 1982). Questi fronti, che raggiungono il fondo sono, da nord a sud, il *Subantarctic Front (SAF)*, il *Polar Front (PF)*, il *Southern Antarctic Circumpolar Current Front (SACCF)*, e il confine meridionale della CCA (*Bndry*), a sud del quale si trovano le acque della *Weddell-Scotia Confluence*.

I fronti della CCA formano dei meandri e si possono spostare lateralmente anche di 100 km in 10 giorni (Nowlin e Klinck, 1986). Il loro trasporto complessivo attraverso il *Drake Passage* è di circa 140 Sv ($1 \text{ Sv} = 10^6 \text{ m}^3 \text{ s}^{-1}$), contro i 10 Sv trasportati dal *SACCF*, il più piccolo dei fronti.

Il confine meridionale della CCA, o *southern boundary (Bndry)* e il *SACCF*, il più meridionale dei fronti, fluiscono lungo la scarpata continentale della Penisola Antartica occidentale. Dopo la PA, il *Bndry* si trova a est nella parte meridionale dello *Scotia Sea* e passa fuori nell'Atlantico meridionale attraverso una depressione tra le isole *South Sandwich* e *South Georgia*. Dall'altra parte, il *SACCF* gira attorno alla parte orientale della piattaforma di *South Georgia* e si dirige verso il mare aperto attraverso il *North-West Georgia Rise*. All'ovest, il fronte polare attraversa il *North Scotia Ridge*.

Le correnti associate con il *SACCF* e *Bndry* sembrano costituire le principali vie di trasporto dalla PA (Hoffman et al., 1998) per organismi e stadi vitali planctonici. Considerando che la CCA fluisce in modo consistente in senso orario, tale trasporto è essenzialmente unidirezionale nell'area di studio e determina potenzialmente una relazione *source-sink* tra popolazioni della PA e quelle poste più a Nord Est sullo *Scotia Arc*.

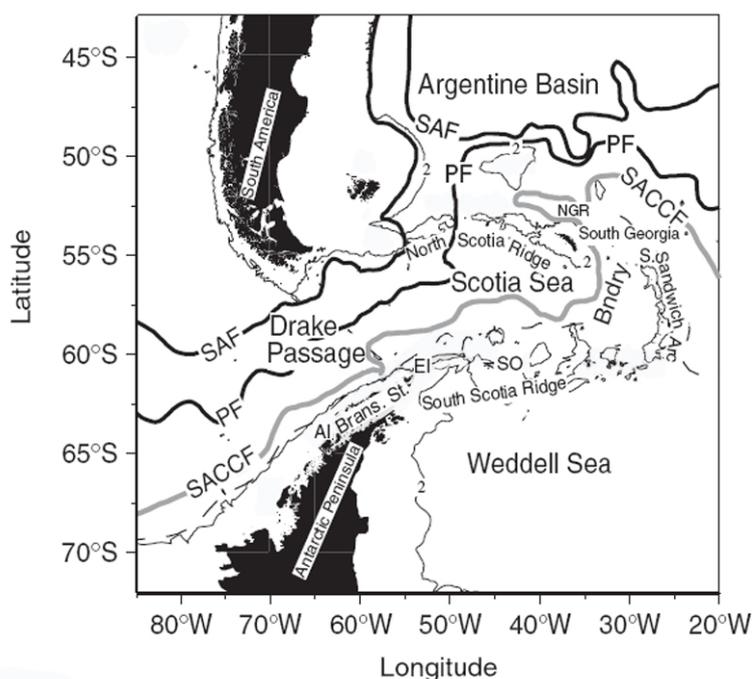


Figura 2. Mappa rappresentante la localizzazione del *Subantarctic Front* (SF), il *Polar Front* (PF), il *Southern Antarctic Circumpolar Current Front* (SACCF), e il confine meridionale della Corrente Circumpolare Antartica (linea tratteggiata; *Bndry*). La linea sottile rappresenta l'isobata dei 2000 m. I nomi delle località geografiche sono abbreviati come North Georgia Rise (NGR), *South Orkney Islands* (SO), *Elephant Island* (EI), *Anvers Island* (AI) e *Bransfield Strait* (Brans. St). Modificato da Fach e Klinck, 2006.

Per quanto riguarda le acque profonde, lo *Scotia Sea* riceve le acque principalmente dal mare di Weddell, via il limbo settentrionale del *Weddell Gyre*. Le acque che originano dal mare di Weddell comprendono le *Warm Deep Water* (WDW) e le più profonde *Weddell Sea Deep Water* (WSDW), le masse d'acqua più dense che si trovano nello *Scotia Sea* (Meredith et al., 2008). Queste acque rappresentano una componente essenziale delle *Antarctic Bottom Water* (ABW), contribuendo con il 60-70% alla loro produzione (Orsi et al., 1999). Le ABW formano il livello abissale della maggior parte degli oceani del mondo (Meredith et al., 2008).

L'ingresso nello *Scotia Sea* delle WSDW avviene principalmente attraverso i passaggi profondi e stretti nelle dorsali di *South Scotia* (*South Scotia Ridge*)

(Naveira Garabato et al., 2002). Le WDSW, dopo essersi unite alle ABW, abbandonano lo *Scotia Sea* attraverso il *Georgia Passage* profondo 3200 m, a sud-est di *South Georgia* (Meredith et al., 2008; Locarnini et al., 1993). Tuttavia, la componente più densa di queste acque, nota come *Lower WSDW (LWDSW)* è troppo profonda per abbandonare lo *Scotia Sea* attraverso il *Georgia Passage*. L'unica via di sfuggire per queste acque dense e fredde resta il rimescolamento con le acque soprastanti ed è ciò che si osserva ad alte intensità in tutto il bacino del mare (Heywood et al., 2002). L'*upwelling* delle *LWDSW* avviene fino alla superficie con importanti effetti a livello di produttività dell'area, mentre una parte considerevole diventa *ABW* che raffredda l'oceano Mondiale.

Struttura di popolazione in *Scotia Sea icefish*

In questa tesi viene investigata la struttura di popolazione dello *Scotia Sea icefish*, *Chaenocephalus aceratus* (Perciformes, Notothenioidei, Channichthyidae). Gli *icefish*, o pesci a sangue bianco, rappresentano una famiglia unica di pesci che vivono quasi esclusivamente nelle acque stabilmente fredde e ricche di ossigeno dell'oceano Australe. Tra questi, *C. aceratus* è distribuito esclusivamente nel settore Atlantico dell'oceano, in particolare nello *Scotia Sea* lungo le isole dello *Scotia Arc* e l'estremità della Penisola Antartica (PA), oltre che attorno all'isola di *Bouvet* (Iwami e Kock, 1990). Nella parte settentrionale di PA, in vicinanza dello stretto di Bransfield, e nelle isole di *Scotia Arc*, *C. aceratus* costituisce la specie ittica più abbondante (Kock et al., 2000). Gli adulti di *C. aceratus* si trovano tipicamente alle profondità comprese fra 200 e 400 m in corrispondenza della piattaforma continentale, anche se alcuni individui sono stati trovati a più di 700 m dalla superficie (Kock e Jones, 2002; Kock, 2005I); sono negativamente galleggianti (Eastman e Sidell, 2002) e conducono una vita bentonica sedentaria. La fase larvale è invece pelagica, di durata che può arrivare anche a 400-500 giorni (La Mesa e Ashford, 2008).

Dati parassitologici e sui tempi di riproduzione/maturità indicano la presenza di popolazioni differenziate nelle isole di *South Shetland* e *South Orkney* (Siegel, 1980; Sosinski e Janusz, 2000). Questi risultati sono supportati da Papetti et al. (2009), dove viene rilevato un differenziamento genetico su scala geografica fra le popolazioni provenienti da *southern South Shetland - Elephant Island* e le isole *South Orkney*, mentre le isole *South Shetland* (distinte in *southern South Shetland* e *Elephant Island* in Papetti et al. (2007) supportano un popolazione panmittica di *C. aceratus* (Papetti et al., 2007). Studi sulla chimica degli otoliti (Ashford et al.,

in stampa) rivelano invece l'esistenza di più popolazioni lungo la PA e di una popolazione differenziata a *South Georgia*.

Notothenoidei

La diversificazione dei nototenioidi, stimata da Near (2004) mediante la correlazione fra la divergenza dei lineamenti antartici del subordine, caratterizzati dalla presenza delle proteine anticongelamento (AFGP), e l'instaurazione delle temperature gelide dell'oceano, viene fatta risalire a circa 41 Ma fa. Questa stima si basa sulla datazione del fossile *Proeleginops grandeastmanorum*, un nototenioido non antartico, diversificatosi sicuramente antecedentemente a quella datazione. Il fossile viene anche utilizzato da Near per calibrare la stima dei nodi più recenti. La stima della divergenza di Near risulta più abbondante rispetto a quella riportata in altri lavori (Bargelloni e Lecointre, 1998; Chen et al., 1998). Tuttavia, la maggior parte degli autori che basano le inferenze filogenetiche su dati morfologici e molecolari, concordano sul fatto che il *clade* dei nototenioidi si sia diversificato in seguito all'isolamento e il raffreddamento dell'oceano Australe (Near, 2004). La datazione di quest'ultimo evento resta però tuttora poco chiara. A sostegno di tale teoria esistono numerose evidenze sul *pattern* di distribuzione attuale dei nototenioidi. Per esempio, le specie non antartiche del subordine, classificati come Bovichtidae, Pseudaphritidae e Eleginopidae, sono distribuiti nelle acque costiere del Sud America meridionale, dell'Australia meridionale e della Tasmania. Queste linee risultano filogeneticamente alla base del *clade*, distribuito quasi esclusivamente nell'Antartico, fornendo l'evidenza che la vicarianza, associata con la rottura del Gondwana, abbia costituito un fattore importante nella diversificazione dei nototenioidi (Eastman, 1993; Bargelloni et al., 2000).

La maggior parte delle specie del subordine è distribuita sulla Piattaforma Continentale Antartica. Inoltre, i nototenioidi rappresentano la componente dominante dei pesci teleostei antartici, sia in termini di diversità di specie, che di biomassa (Eastman, 1993). I nototenioidi antartici sono caratterizzati dalla presenza di glicoproteine anticongelamento (AFGP), citate sopra, che aiutano a prevenire il congelamento dei liquidi corporei a temperature sotto zero delle acque antartiche (DeVries e Lin, 1977). Questa caratteristica viene considerata come un'innovazione chiave, che ha permesso ai nototenioidi di diversificarsi nelle acque gelide e ricoperte dal ghiaccio dell'Antartico (Eastman, 1993, Chen et al., 1997). I nototenioidi antartici che presentano AFGP sono classificati nei Nototheniidae, Harpagiferidae, Artedidraconidae, Bathydraconidae e

Channichthyidae. Le analisi filogenetiche basate sulle inferenze morfologiche e molecolari forniscono sempre più prove a sostegno della monofilia all'interno del gruppo dei nototenioidi caratterizzati dalla presenza di AFGP (Iwami, 1985; Bargelloni et al., 1994; 2000).

I nototenioidi, come il loro comune antenato ancestrale, sono per lo più pesci di profondità o prettamente bentonici, come i sedentari artedidraconidi. Essi presentano però una fase larvale pelagica che ne incrementa il potere dispersivo e l'elasticità ecologica (Eastman 1993, Bargelloni et al., 2000). Solo poche specie hanno evoluto un *habitus* pelagico o mesopelagico (ad esempio il più grande nototenioido, *Dissostichus mawsoni*, e la sua preda principale, *Pleuragramma antarcticum*), mentre alcuni, come *Pagothenia borchgrevinki*, vivono proprio sotto il ghiaccio (specie criopelagica, Eastman, 1993). Si ritiene che una delle caratteristiche peculiari dei Nototenioidi sia quella di aver perso la vescica natatoria. Infatti, quasi tutte le specie sono bentoniche o bento-pelagiche. Le rare specie pelagiche hanno acquisito una serie di adattamenti compensatori che hanno portato all'alleggerimento del loro corpo e nel contempo all'abbattimento della richiesta di energia per mantenere la posizione nella colonna d'acqua. Il galleggiamento passivo è possibile grazie ad una consistente riduzione della mineralizzazione delle ossa, il mantenimento della notocorda (conseguentemente le vertebre sono cave) e lo stoccaggio di riserve lipidiche sottocutanee, intramuscolari o in vescicole (Eastman, 1993).

Tra gli adattamenti fisio-morfologici messi in atto dai Nototenioidi, i più peculiari sono costituiti dallo sviluppo delle già citate glicoproteine anticongelamento (AFGP) e dalla riduzione dell'emoglobina. La completa assenza di emoglobina si riscontra, tuttavia, solo nella famiglia dei cannittidi. Tutti i nototenioidi, infatti, hanno raggiunto una condizione di ridotta viscosità ed abbassamento dell'ematocrito, ma solo i membri di questa famiglia hanno portato il carattere all'estremo, perdendo l'emoglobina. Per questo motivo i Channichthyidae vengono comunemente chiamati *icefishes* (Bargelloni et al., 1994).

Channichthyidae, o *icefish*

La comparsa e la diversificazione dei Channichthyidae resta ancora poco chiara e in genere viene fatta risalire in seguito alla comparsa della Convergenza Antartica. Il progenitore comune più recente dei cannittidi risale, secondo Near (2004) a 8.5 Ma fa, ad indicare che probabilmente, gli *icefish* si siano evoluti quando le temperature dell'oceano Australe erano già vicine ai livelli attuali. La famiglia è

attualmente considerata un gruppo monofiletico (Chen et. al., 1998; Near et al., 2003), costituito da 15 specie.

Gli *icefish*, o pesci dal sangue bianco, rappresentano una famiglia unica di specie tra i vertebrati in quanto non possiedono emoglobina. Tutte le specie dei cannittidi, tranne una (*Champscephalus esox*, distribuito nella piattaforma patagonica meridionale), vivono esclusivamente nelle acque sempre gelide e ricche di ossigeno dell'oceano Australe, entro la Convergenza Antartica.

La temperatura delle acque antartiche varia da meno di -1.8°C a 3.0°C (Clarke, 1990; Kock, 1992). Questo range di $4-5^{\circ}\text{C}$ può essere fisiologicamente troppo impegnativo per molte delle specie antartiche, che per questo motivo vengono distinte in due ittiofaune, dell'alto Antartico e del basso Antartico. Della prima fanno parte due terzi delle specie *icefish* che possiedono il proprio *optimum* compreso fra -1.86°C e 0°C . La specie considerata in questa tesi, *C. aceratus*, appartiene alla categoria della fauna ittica del basso Antartico, avendo un *optimum* compreso fra -0.5 e 1.5°C .

Chaenocephalus aceratus

C. aceratus (*Black-fin icefish* o *Scotia icefish*) (Figura 3), un tipico cannittide bentonico, presenta un corpo robusto, una forma meno affusolata rispetto alle altre specie della stessa famiglia e un caratteristico esteso ispessimento della pelle nella parte distale delle pinne pelviche che vengono utilizzate per appoggiarsi al substrato (Eastman, 1993). La specie viene comunemente trovata fino a 450 m di profondità con i singoli individui che si trovano a 500 - 770 m.

C. aceratus è una specie a sessi separati a fecondazione esterna. I maschi raggiungono la maturità intorno a 35 – 45 cm di lunghezza, mentre le femmine a 45 - 55 cm. La riproduzione avviene tra Aprile e Luglio in acque costiere, dove la profondità supera i 100 m. Le femmine producono 3000 - 22000 uova dal diametro di 4.4 - 4.7 mm che vengono deposte sul fondo in buche poco profonde (Kock, 2005II). I maschi svolgono le cure parentali, difendendo i nidi (Detrich et al., 2005).

Le fasi larvale, postlarvale e giovanile, come negli altri *icefish*, prediligono le acque superficiali e i primi 100 m della colonna d'acqua (Loeb et al., 1993).

Le larve nascono tra Agosto e Ottobre (Kock, 2000). La fase giovanile pelagica, che può durare 2 - 4 anni, è ancora poco compresa. Le classi di età 0+ sono spesso associate con le aggregazioni di krill e diventano un regolare *by-catch* nella sua pesca presso l'isola di *South Georgia* e lo *Scotia Arc* meridionale (Iwami et al., 1996). Durante la fase adulta bentonica le femmine possono raggiungere 70 – 75

cm di lunghezza superando i 3800 g di peso. I maschi, più piccoli, arrivano invece a 55 – 58 cm e non superano i 1300 g (Kock, 2000).

Il cambiamento di habitat nel corso del ciclo vitale di *C. aceratus* è associato a una modificazione delle sue abitudini alimentari. La fase larvale della specie, infatti, si nutre principalmente di larve furcilia e individui giovanili di krill, mentre i giovanili prediligono misidi e krill adulto. Da adulti (all'età di 2-3 anni), gli individui di *C. aceratus* diventano sedentari e cambiano la propria dieta, adottando una strategia *sit-and-wait* e specializzandosi su alcune specie *target* di pesci (Kock et al., 2000). Solo occasionalmente gli adulti ascendono ai livelli d'acqua vicini al fondo o ai livelli intermedi per nutrirsi di krill (Siegel, 1980; Takahasi e Iwami, 1997)

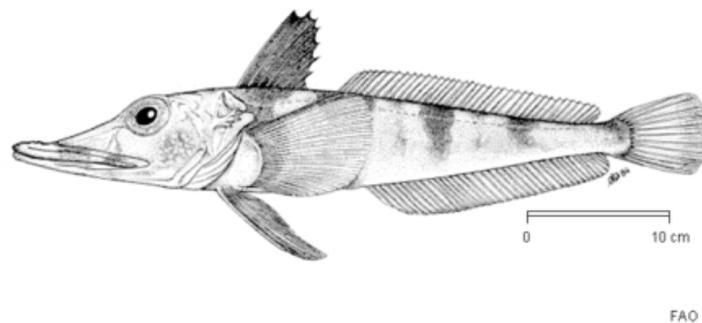


Figura 3. *Chaenocephalus aceratus* nella fase adulta. Fonte immagine: FAO

Le caratteristiche ecologiche come la sedentarietà degli individui adulti, la deposizione delle uova in profondità delle buche, la presenza di cure parentali da parte del maschio fino alla schiusa delle uova, fanno pensare che la dispersione in questa specie avvenga pressoché esclusivamente attraverso la fase larvale. In questo quadro, l'investimento parentale considerevole (basti pensare al diametro delle uova) da parte di entrambi i sessi potrebbe significare che la quantità di uova disperse dalle correnti sia limitata, e che ci sia una misura di *self retention* delle larve nella fase dispersiva.

Come tutti i Nototenioidi, questa specie è molto longeva. Nelle vicinanze dell'isola *South Georgia* sono stati trovati individui di 13-15 anni.

Chaenocephalus aceratus ha costituito una delle principali specie *by-catch* nella pesca di fondo che aveva come *target* il *mackerel icefish* *Champosocephalus gunnari* (Kock, 2005I). Il suo sfruttamento è iniziato negli anni 1970 presso *South Georgia*, ma solo occasionalmente, come nella stagione 1977/1978 (in seguito al declino dello *stock* di *C. gunnari*), *C. aceratus* ha rappresentato il *target* diretto.

La pesca di *C. aceratus* si è poi estesa alle altre aree del settore Atlantico, come le isole *South Orkney*, dove la specie è stata sfruttata dagli anni 1977/1978 agli 1989/1990 (Kock 1991; Sosinski e Paciorkowski, 1993). Lo sfruttamento presso le isole *South Shetland* e la Penisola Antartica è stato considerevole nel periodo che va dagli anni 1978/1979 agli 1981/1982, mentre in seguito, e fino agli anni 1989/1990, queste zone sono state visitate solo occasionalmente (Kock 1991; Sosinski e Paciorkowski, 1993; Kock et al., 2004). A partire dagli anni 1990, molti degli *stock* dei Channichthyidae hanno cominciato a subire un rapido declino, sia in seguito a un eccessivo sfruttamento, sia per cause naturali (Kock, 2005I). Questo ha portato alla moratoria della pesca commerciale dopo la stagione 1989/1990 da parte della *Commission for the Conservation of Antarctic Marine Living Resources* (CCAMLR) lungo lo *Scotia Arc* meridionale, comprendendo le isole di *South Georgia*, *South Orkneys*, *South Shetland* e la PA. Nel 1995 la pesca commerciale è stata riaperta nell'area di *South Georgia* (CCAMLR *Conservation Measures* 32-03 e 32-04), consentendo soltanto l'utilizzo degli attrezzi da traino pelagici, in grado di pescare selettivamente *C. gunnari* e limitando la cattura di *by-catch*, costituito prevalentemente dalle specie bentoniche. La moratoria, tuttora in vigore, ha permesso a *C. aceratus* di ristabilire le dimensioni originali degli *stock* (Kock, 1998; Jones et al., 2000; Kock e Jones, 2005).

Scopo della tesi

In questa tesi ci si propone di studiare il differenziamento genetico della specie *Chaenocephalus aceratus* che, come descritto in precedenza, è endemica della Penisola Antartica, delle isole dello *Scotia Arc* e di *Bouvet Island*. Per questo sono state analizzate alcune popolazioni di *C. aceratus*, provenienti da aree finora non considerate in lavori precedenti. I dati ottenuti da questo studio sono stati inoltre confrontati con i risultati degli studi precedenti sulla specie. Sono state formulate e verificate alcune ipotesi di differenziamento: I) geografico, II) tra repliche temporali dello stesso campione di popolazione.

MATERIALI E METODI

Questo studio comprende due fasi principali:

- Genotipizzazione di **176** individui suddivisi in **5** campioni di popolazione di *Chaenocephalus aceratus*.
- Confronto dei dati ottenuti con informazioni precedenti: completamento dell'areale di distribuzione della specie, considerazioni finali.

Campionamento

I campioni di *C. aceratus* analizzati per questo studio sono stati raccolti durante diverse crociere. In maggior parte sono stati ottenuti grazie a scambi con gruppi internazionali con lo scopo di ottenere un campionamento più ampio possibile, in grado di coprire l'intero areale di distribuzione di questa specie. In particolare, i campioni analizzati in questo studio provengono da *Bouvet Island*, *Crystal Sound*, *Dallmann Bay*, *Elephant Island* e *South Georgia*, come indicato nella tabella 1 e nella figura 4 di riferimento.

La tabella 2 descrive invece i campioni analizzati in due studi precedenti (Papetti et al., 2007; Papetti et al., 2009), confrontati con le popolazioni genotipizzate in questa tesi. Per la popolazione di *Elephant Island* sono disponibili tre repliche temporali. La figura 4 indica riassuntivamente la localizzazione di tutti i campioni a disposizione per la specie *C. aceratus*.

Estrazione del DNA

Il DNA è stato estratto da **122** individui, ad eccezione dei campioni di *Bouvet Island* e *Crystal Sound* per i quali era già a disposizione il DNA genomico. L'estrazione è stata effettuata seguendo il protocollo *Salting out* (Patwary et al., 1994) che prevede l'isolamento del DNA dopo l'eliminazione delle proteine per precipitazione.

Sito di campionamento	Anno	Crociera	Acronimo del campione	Numero degli individui
Bouvet Island	2004	ICEFISH 2004 ¹	BI04	4
Crystal Sound	2008	LMG08-07 ²	CS08	50
Dallmann Bay	2005	LMG05-05/06 ³	DB05	35
Elephant Island	2007	ANT-XXIII/8 ⁴	EI07	54
South Georgia	2004	DOSG0104 ¹	SG04	33
Totale				176

¹ Raccolto da Mark Belchier, British Antarctic Survey (BAS), a bordo del RV James Clark Ross (Regno Unito)

² Raccolti da William Detrich III a bordo del RV Laurence M. Gould (USA)

³ Raccolto da Bruce Sidell a bordo del RV Laurence M. Gould (USA)

⁴ Alfred Wegener Institute (AWI), a bordo del RV Polarstern (Germania)

Tabella 1. Campioni di *Chaenocephalus aceratus* analizzati per la prima volta in questo studio.

Sito di campionamento	Anno	Campagna	Acronimo del campione	Numero degli individui
Elephant Island	2002	ANT-XIX/3 ¹	EI02	202
Southern South Shetlands	2002	ANT-XIX/3 ¹	SS02	45
Elephant Island	2006	JR145 ²	EI06	47
South Orkney	2006	JR145 ²	SO06	48
Totale				342

¹ Alfred Wegener Institute (AWI), a bordo del RV Polarstern (Germania)

² British Antarctic Survey (BAS), a bordo del RV James Clark Ross (Regno Unito)

Tabella 2. Campioni di *Chaenocephalus aceratus* utilizzati negli studi precedenti.

Tutte le soluzioni utilizzate per l'estrazione sono state preventivamente sterilizzate in autoclave, e, qualora possibile, si è utilizzato materiale plastico monouso autoclavato. La strumentazione da dissezione è stata accuratamente lavata con acqua deionizzata (dH₂O) nel passaggio da un individuo all'altro, per evitare di contaminare tra loro i campioni.

In una provetta *Eppendorf* da 2 ml, pochi milligrammi (5-10mg) di muscolo sono stati trattati con 300 µl di tampone (0.1 M EDTA, 0.05 M TRIS pH 8), 37.5 µl di SDS 10 % e 12 µl di proteinasi K (20 ng/ml) successivamente riscaldato a 55 °C

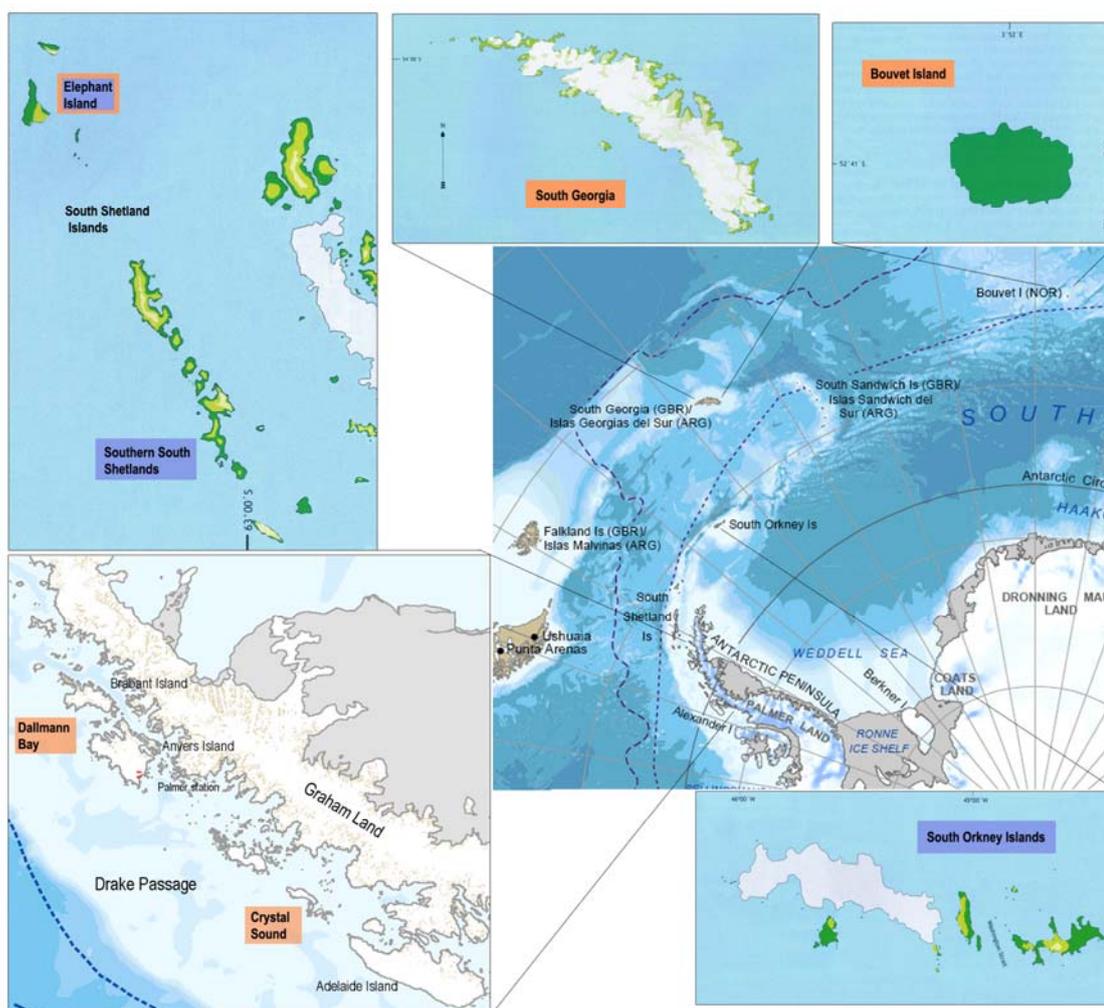


Figura 4. Mappa della Penisola Antartica comprendente l'intero areale di distribuzione di *Chaenocephalus aceratus*. Vengono riportati gli ingrandimenti delle aree corrispondenti ai siti di raccolta dei 9 campioni totali di *C. aceratus*. Il campione di *Elephant Island* è costituito da tre repliche temporali di individui raccolti negli anni 2002, 2006 e 2007. I siti di raccolta in rosso indicano i campioni analizzati nel presente lavoro, mentre quelli in blu indicano i campioni analizzati negli studi precedenti (Papetti et al., 2007) (Papetti et al., 2009).

per un'ora per la completa dissoluzione dei tessuti. Sono stati aggiunti 2 μ l di RNasi A (10 mg/ml) a temperatura ambiente e 90 μ l di una soluzione satura (>6 M) di NaCl che facilita la precipitazione delle proteine. Dopo aver agitato la soluzione per 20 minuti con un *vortex* ed averla centrifugata a freddo a 14000 RPM per 30 minuti, le proteine precipitate si sono concentrate sul fondo. Il sovranatante (contenente il DNA) è stato prelevato e trasferito in un'altra provetta *Eppendorf* da 1.5 ml. Il DNA è stato fatto precipitare con un volume di isopropanolo freddo e centrifugato per 20 minuti a 14000 RPM. L'isopropanolo, ed eventuali tracce di altri coprecipitati, sono stati allontanati con etanolo 70 %.

Un'ultima centrifugazione è servita ad eliminare l'etanolo e a raccogliere il DNA in un *pellet*. Il DNA è stato infine disciolto in dH₂O sterile ed autoclavata. Le soluzioni di DNA sono state poi conservate a -20°C fino al momento dell'utilizzo.

Marcatori microsatellite: vantaggi e svantaggi

Per gli scopi di questa tesi sono stati utilizzati dei marcatori nucleari: i *loci* microsatellite. Microsatelliti sono corti frammenti di DNA contenenti un modulo di 1-6 basi ripetuto fino a 60 volte. La lunghezza massima di questi frammenti è limitata, probabilmente da un meccanismo selettivo. Grazie alla loro eccezionale variabilità e ad una relativamente facile genotipizzazione, vengono ora generalmente considerati come tipo di marcatore genetico più potente (Schlötterer, 2004). Oltre ad essere altamente variabili, sono anche ampiamente distribuiti in tutti i genomi eucariotici, principalmente nelle regioni non codificanti, anche se a volte vengono ritrovati nelle regioni codificanti. Il meccanismo di mutazione principale porta all'estensione di unità di *repeat* a causa dello slittamento della DNA-polimerasi durante la fase di replicazione alterando il numero di moduli ripetuti (*repeat*) del microsatellite (Tautz e Schlötterer, 1994).

Tasso di mutazione elevato, di circa 10^{-3} - 10^{-4} eventi di mutazione per generazione (Goldstein e Schlötterer, 1998), alto numero di alleli (fino a 50) ed elevati livelli di eterozigotità (compresa fra 50 e 100%) costituiscono alcune delle caratteristiche peculiari dei microsatelliti, grazie alle quali essi vengono attualmente considerati come strumenti ideali negli studi di genetica di popolazione e per l'identificazione di popolazioni differenziate (Schlötterer, 2004).

Alcuni aspetti costituiscono però una limitazione, conducendo a errori nella genotipizzazione (o *scoring*). Lo *scoring* rappresenta il primo passo di analisi di dati grezzi di lunghezza dei frammenti e consiste nell'individuare tutti gli alleli presenti ad un dato *locus*. Esistono numerose fonti d'errore legate allo *scoring* dei genotipi microsatellite, tra cui *stuttering*, *large allele drop-out* e presenza di alleli nulli, che sono di particolare importanza in quanto contribuiscono ad una errata valutazione delle frequenze alleliche e genotipiche che può, in linea definitiva, causare l'errore nell'interpretazione dei dati (DeWoody et al., 2006). In particolare, nella maggior parte dei casi, questi errori portano ad una sottostima

degli eterozigoti che genera una deviazione dall'equilibrio di Hardy-Weinberg (HWE) non dovuta a fenomeni evolutivi, ma ad artefatti della fase sperimentale.

Il termine *stuttering* si riferisce alla produzione di picchi in una posizione diversa da quella dell'allele parentale, causata dallo slittamento lungo un intero *repeat* della polimerasi durante la PCR, che provoca la perdita di quest'unità. L'allele genotipizzato risulta essere quindi più piccolo rispetto all'allele parentale. L'errore di *stuttering* è tipico dei *loci* di nucleotidici (Fig. 5). 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. Una serie di errori commessi in questa maniera potrà di conseguenza portare a sovrastimare le frequenze degli alleli più grandi, diminuendo l'eterozigosità osservata e aumentando apparentemente il livello di *inbreeding* nei loci affetti (DeWoody et al., 2006).

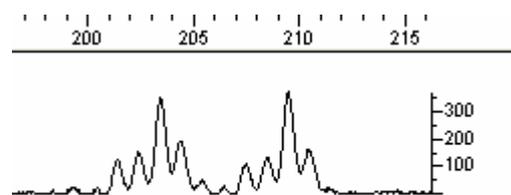


Figura 5. Esempio di stuttering. In ascissa la dimensione in paia di basi (pb) e in ordinata l'intensità di emissione del campione.

Il *Large allele drop-out* costituisce un altro potenziale errore di scoring in grado di falsificare le reali frequenze alleliche e genotipiche. Il *Large allele drop-out* (o perdita di alleli grandi in bp), è osservata principalmente nei *loci* che presentano grandi differenze nelle dimensioni alleliche (Björklund, 2005) ed è causata dall'amplificazione preferenziale dell'allele più piccolo in un genotipo eterozigote (Gagneux et al., 1997). Se non rilevato, il *large allele drop-out* può causare una sovrastima delle frequenze di alleli più corti e risultare nella totale omissione dal *data set* di alleli rari e grandi. Come nel caso dello *stuttering*, il *large allele drop-out* potrà causare in questo modo un deficit dell'eterozigosità osservata aumentando apparentemente il tasso di *inbreeding* nei *loci* affetti (DeWoody et al., 2006).

Infine, gli alleli nulli sono i più difficili da individuare in quanto sono alleli presenti nella popolazione che però non vengono amplificati. Ciò avviene a causa di una mutazione nel sito di *annealing* del primer che non permette il suo aggancio, impedendo l'amplificazione del frammento. La presenza di alleli nulli in un data set fa sì che gli eterozigoti per l'allele nullo possano essere classificati

come omozigoti, mentre individui la cui amplificazione fallisce completamente potrebbero essere omozigoti per tale allele (DeWoody et al., 2006).

A questi errori, si affianca la soggettività dell'interpretazione dell'operatore. Questo possibile problema viene solitamente risolto confrontando risultati dello scoring effettuato da più operatori.

La strategia ideale per minimizzare gli errori di scoring prevede un'accurata serie di controlli. La natura di questa tesi di laurea, intesa sia come ampliamento di un *data set* precedente (Papetti et al., 2007; 2009), sia come revisione completa di tale *data set*, ha permesso di applicare una strategia di minimizzazione degli errori di scoring a diversi stadi della genotipizzazione, dall'amplificazione alla correzione del *data set* finale.

In particolare sono stati applicati i seguenti accorgimenti:

- Amplificazione in multiplex (kit QUIAGEN). Grazie a una *Taq* polimerasi altamente affidabile e il kit di amplificazione PCR ideato appositamente per l'analisi dei microsatteliti da QUIAGEN, si è in grado di minimizzare lo *stuttering* e prevenire il *large allele drop-out*. Inoltre, riunendo più loci in un'unica reazione e riducendo la necessità di riamplicazione di loci soggetti ad errore, si è ridotto il numero totale di amplificazioni necessario per analizzare un *data set* (DeWoody et al., 2006):

- Rianalisi di un subset di individui: sono stati indipendentemente riamplicati e rigenotipizzati individui che mostravano picchi sospetti, indicativi di alleli adiacenti eterozigoti a un locus soggetto a *stuttering*; individui la cui amplificazione è fallita; campioni di individui con una sospetta contaminazione reciproca durante la procedura di amplificazione.

- Binning (o *allele calling*): questa fase può introdurre errori, legati principalmente al fatto che dati grezzi continui delle lunghezze dei frammenti forniti in decimali devono essere trasformati in numeri interi che rappresentano classi alleliche distinte (binning è il processo di assegnazione di un codice allelico intero a un valore continuo di lunghezza del frammento ottenuto dal sequenziatore). L'errore può anche essere dovuto al fatto che la lunghezza degli alleli misurata è influenzata dalla velocità di migrazione del frammento, che a sua volta è legata al contenuto in GC del frammento (Rosenblum et al., 1997; Wenz et al., 1998). Di conseguenza, in base al contenuto in GC, gli alleli possono mostrare una periodicità che si avvicina, ma che non è uguale all'unità di *repeat* (Amos et al., 2007). La lunghezza stimata del frammento può, inoltre, variare da esperimento a esperimento a causa della variazione stocastica della temperatura ambientale e della differente qualità della matrice di separazione (Ghosh et al., 1997; Rosenblum et al., 1997). Errori introdotti durante il *binning* si ripercuotono in generale sulla stima e il calcolo delle frequenze alleliche distorcendo parametri chiave come l'eterozigosità (Amos et al., 2007).

Esistono diverse strategie che permettono di ridurre il tasso d'errore di *binning*. Tali strategie si basano essenzialmente sul confronto con i genotipi di cui si conoscono già le distribuzioni delle classi alleliche (ad esempio nel caso delle popolazioni umane europee) o sulla genotipizzazione d'interesse famiglie di individui, in cui gli alleli presenti ad ogni singolo locus vengono confrontati fra individui all'interno della stessa famiglia e tra famiglie diverse. Per *C. aceratus* non sono a disposizione genotipi standard di riferimento o *pedigree* di famiglie genotipizzate, per cui il *binning* è stato eseguito basandosi esclusivamente sulla lunghezza dei frammenti. Gli stessi dati sono stati trattati con due metodi di *binning*, implementati in due *software* differenti, FLEXIBIN e TANDEM. Gli *output* dei due programmi sono stati in seguito trattati come dati indipendenti e sono serviti per confrontare la validità dei due metodi di *binning*.

Per individuare errori di *binning*, sono stati anche usati i seguenti accorgimenti: controllo manuale dello *scoring*, test per gli errori di *scoring* con un programma statistico. Di seguito vengono riportati alcuni dettagli dei due livelli di controllo eseguiti.

- il controllo manuale dello *scoring* è stato eseguito per correzione di eventuali errori dei programmi di *binning*, soprattutto in loci con numero elevato di alleli;
- il controllo dell'intero *data set* è stato svolto con il *software* statistico MICROCHECKER che corregge per la presenza di errori di *scoring* sulla base delle frequenze alleliche, delle dimensioni degli alleli e dello scostamento dall'equilibrio di HW.

Va però considerato, in conclusione, che queste strategie non sono infallibili, così come i programmi per identificare gli errori di *scoring* possono identificare falsi positivi, forzando i dati verso l'equilibrio di Hardy-Weinberg; in questo caso, tuttavia, anche se viene persa parte dell'informazione, i successivi test di differenziamento risultano essere più conservativi.

Marcatori microsatellite utilizzati per questo studio

Per questo studio sono stati utilizzati undici *loci* microsatellite. Sette *loci* sono specie-specifici, dal momento che sono stati isolati da *Chaenocephalus aceratus* (Susana et al., 2007), mentre quattro *loci*, isolati da *Chionodraco rastrispinosus* (Notothenoidei, Channichthyidae) (Papetti et al., 2006), sono stati cross-amplificati in *C. aceratus*. Tutti gli undici *loci* sono elencati nella tabella 3.

Come in parte descritto nella sezione sugli errori di *scoring*, diversi loci possono essere combinati in un'unica reazione di amplificazione, qui di seguito indicata

<i>Locus</i>	<i>Repeat</i>	Sequenza <i>primer</i> (5'-3')	Marcatore fluorescente
Ca21*	(CA) ₁₄	F: GGACGAAGAAAGCAAGAAGAACGA R: CTATGTGGGC(AG)TGTTTGTGATTAT	6-FAM
Ca26	(TGCGTG) ₇	F: AAGGTGGGCAACAGGTTAGAGT R: ATGAACACATACAAGTGGTCACAT	HEX
Ca40	(AC) ₂₁	F: GCTCGCTGTGGTATGTGCTAAT R: AAAGTGAGATATGTA(ACT)CGTAACTG	TAMRA
Ca48	(GT) ₁₅	F: CATTCTCCCTGAGTTTACCAC R: TCTTCTCTTTCACCTCTAC(AG)G	HEX
Ca55	(ACGC) ₆ (AC) ₂₂	F: GAAATAAAGGATGGCAAGGGAG R: AACAAACAGGCAAAAAGACAGGC	HEX
Ca86	(AC) ₃₈	F: AGTTATTATGGTGGTAAATCGCT R: GCTGCTGGT(CT)ACAGGAGTGT	6-FAM
Ca88	(AC) ₂₄	F: GATTGAGGTCCCCACAAGTAA R: AAGA(GT)ACAAGAATAATAAAAT(AG)ACA	6-FAM
Cr15	(AC) ₈	F: TAACATGCTCCCGTCTCC R: ATATCATGTGGGCAGAATGGT	HEX
Cr127	(GT) ₄ ATAA	F: CGTATAGGGCCGTACCTCA R: GCTCCATCATAGATCCAGTCA	TAMRA
Cr171	(AC) ₂₁	F: CTCGTGGTGGATGATATAAAA R: AAGTCATATGGTAGCACTTTTAGC	TAMRA
Cr259	(AG) ₁₇ GG(AG) ₅	F: TGATTACTTCCATCTTCACACATA R: CACAAAGAATTCTGGGAACAG	6-FAM

Tabella 3: Elenco dei *loci* microsatellite genotipizzati in questa tesi. La tabella riporta in particolare il nome del *locus*, la sequenza dei *primers* (F *Forward*, R, *Reverse*) utilizzati per amplificare ogni *locus* e il fluoroforo con cui sono stati marcati tutti i *primers forward* (F).

come PCR in *multiplex*. Per questo studio, sono state ottimizzate le condizioni di amplificazione in *multiplex* con l'uso del kit QIAGEN.

Per consentire la corsa elettroforetica simultanea di più frammenti nello stesso capillare del sequenziatore automatico, sono stati utilizzati i *primers forward*, specifici per ogni *locus*, legati ad un marcatore fluorescente (6-FAM, HEX o TAMRA). La scelta del marcatore specifico per ogni *primer* è basata sulla lunghezza dei frammenti microsatellite che si vogliono amplificare nella *mix* PCR, in modo tale da permettere la distinzione di frammenti di lunghezza molto diversa, marcati con lo stesso fluoroforo. Sia i *primers* utilizzati che i fluorofori corrispondenti sono riportati nella tabella 3.

Condizioni di amplificazione in multiplex

Per raggiungere delle condizioni di amplificazione PCR in *multiplex* ottimali e per verificare la compatibilità in termini sia di dimensione dei frammenti che di *performance* di ogni *locus*, sono state eseguite varie prove di amplificazione per diverse combinazioni degli undici *loci* utilizzando sette individui di una delle popolazioni. Per ogni reazione sono stati impiegati 2 µl di DNA, amplificato in 10 µl di volume totale partendo dalle seguenti concentrazioni standard:

- *Master mix* Quiagen 1X,
- *primers Forward* e *Reverse* 0.20 µM ciascuno.

Le amplificazioni sono state svolte al *Thermal Cycler GeneAmp PCR system 9700 Applied Biosystem*

Per tutti i *loci* è stato usato un profilo termico comune:

- (i) predenaturazione 95°C per 15 min
- (ii) 30 cicli:
 - denaturazione 94°C per 30 s
 - *annealing* 57°C per 1 min 30 s
 - estensione 72°C per 1min
- (iii) estensione finale 60°C per 30 min

Nella sezione dei risultati si riportano le concentrazioni e le combinazioni di *loci* finali ottenuti con questa strategia di ottimizzazione.

Individui la cui amplificazione o corsa al sequenziatore erano fallite, sono stati amplificati singolarmente, seguendo le concentrazioni standard messe a punto in Papetti *et al.* 2006 e Susana *et al.* 2007. Il volume totale PCR previsto da tali protocolli era di 20 µL contenenti: *Taq* buffer 1× (Promega, 50 mM KCl, 10 mM Tris-HCl pH9 a 25 ° C, 0.1% TritonX-100), MgCl₂ 1 mM, 150 nM di ciascun *primer* (tranne che per il *locus* Ca88 per il quale sono stati utilizzati 300 nM di *primer forward*), 70 µM di dNTP, 0.8 U di *Taq* polimerasi (Promega) e 50 ng di DNA genomico (tranne che per il *locus* Ca88 per il quale sono stati utilizzati 100 ng di DNA genomico). I *loci* erano amplificati con un profilo PCR *touch down* (tranne che per i *loci* Cr15 e Cr171).

Il profilo base di amplificazione per tutti i *loci* era:

- (i) predenaturazione 94°C per 3 min (non ripetuto)
- (ii) da 12 a 16 cicli di *touch down*:
 - denaturazione 94°C per 40 s
 - *annealing* 65÷58 ° C per 30 s con un decremento di 0.7 ° C per ogni ciclo
 - estensione 72°C per 1 min

- (iii) 30 cicli con *annealing* a temperatura fissa:
 - denaturazione 95°C per 40 s
 - *annealing* 59÷50 ° C per 30 s
 - estensione 72°C per 1 min
- (iv) estensione finale 72°C per 10 min

I protocolli e i profili di amplificazione standard per tutti i *loci* considerati sono riportati nella tabella 4 (Papetti et al., 2006; Susana et al., 2007) .

<i>Locus</i>	<i>Annealing</i> dei cicli TDN	Numero cicli TDN	<i>Annealing</i> cicli a T° fissa	Numero cicli a T° fissa
Ca21*	Da 60 a 52 °C, -0.5 °C ad ogni ciclo	12	50 °C	30
Ca26	Da 60 a 52 °C, -0.5 °C ad ogni ciclo	12	50 °C	30
Ca40	Da 60 a 52 °C, -0.5 °C ad ogni ciclo	12	50 °C	30
Ca48	Da 58 a 50 °C, -0.5 °C ad ogni ciclo	16	50 °C	30
Ca55	Da 58 a 50 °C, -0.5 °C ad ogni ciclo	16	50 °C	30
Ca86	Da 58 a 50 °C, -0.5 °C ad ogni ciclo	16	50 °C	30
Ca88	Da 58 a 50 °C, -0.5 °C ad ogni ciclo	16	50 °C	30
Cr15	-	-	59°C	35
Cr127	Da 65 a 55°C, -0.7 °C ad ogni ciclo	12	55°C	30
Cr171	-	-	57°C	35
Cr259	Da 65 a °C, -0.7 °C ad ogni ciclo	12	55°C	30

Tabella 4. Sono riportati i profili PCR utilizzati per l'amplificazione dei loci singolarmente. Vengono riportati tempi e temperature di *annealing*, numero di cicli *touch down* (TDN), numero e temperatura di cicli a temperatura fissa. *Ogni acronimo del singolo *locus* indica la specie (Ca: *Chaenocephalus aceratus*, Cr: *Chionodraco rastrispinosus*) e il numero della colonia batterica da cui è stato isolato il *locus*.

Il successo delle reazioni di PCR è stato verificato mediante corsa elettroforetica in gel d'agarosio all'1.8%, TAE 0.5X (Tris 20 mM, acido acetico 10 mM, EDTA 0.5 mM, pH 8), a 3 V/cm per 30 minuti. Una volta ottenuta una quantità di template sufficiente, si è proceduto, attraverso diluizioni e corse di prova, alla messa a punto della quantità di prodotto di PCR da sottoporre a *genotyping* per i diversi microsatelliti individuati.

Analisi dei frammenti

L'elettroforesi su capillare dei frammenti (sequenziatore automatico *ABI PRISM 3100 DNA Analyzer*) è stata svolta presso il laboratorio del prof. Patarnello Tomaso del dipartimento di Sanità Pubblica, Patologia Comparata e Igiene Veterinaria dell'Università degli Studi di Padova. Per l'elaborazione dei risultati è stato utilizzato il programma GENOTYPER 3.7 (*ABI Prism Applied Biosystems*) che permette di effettuare la genotipizzazione degli individui. Il programma permette, infatti, un'ispezione visiva di ogni frammento amplificato di un *locus* microsatellite nella ricerca di alleli. Ogni frammento è visualizzato in un grafico avente per ascissa il tempo di migrazione su gel e per ordinata l'intensità del prodotto amplificato. Il dimensionamento dei singoli alleli avviene per confronto con lo *standard* interno (GS 400 Hd Rox) marcato con fluoroforo Rox.

I *file* pervenuti dal servizio di *genotyping* sono stati importati nel programma, quindi i profili microsatellite sono stati separati per colore del fluoroforo, esaminati e gli alleli di ogni *locus* microsatellite dimensionati. L'analisi di ogni picco allelico ed il suo dimensionamento in paia di basi ha permesso di ricavare il genotipo per ogni individuo.

Allele calling o binning

Il passo successivo allo *scoring* consiste nell'assegnare un numero intero, o *bin*, agli alleli osservati in base alla loro lunghezza in paia di basi (bp). GENOTYPER fornisce, infatti, i risultati di lunghezza dei frammenti relativa allo standard interno, che consiste in un numero intero con due decimali. Tuttavia, tutti i programmi di analisi di microsatelliti lavorano con dimensioni alleliche intere. Da qui nasce quindi l'esigenza di arrotondare i dati di *scoring*, un processo noto come *binning*.

In questa tesi sono stati utilizzati due metodi di *binning*, implementati nei programmi FLEXIBIN (Amos et al., 2007), disponibile sul sito web www.zoo.cam.ac.uk/zoostaff/amos/, e TANDEM (Matschiner e Salzburger, 2009), disponibile sul sito www.evolution.unibas.ch/salzburger/software.htm/.

FLEXIBIN, utilizzato per il binning automatico, opera mediante un algoritmo che fornisce alleli sotto forma di numero di *repeat* e permette il controllo di altri eventuali problemi, come alleli di lunghezza intermedia tra due classi (tipo mono-, di-, trinucleotidi). Il numero di *repeat* deve essere impostato manualmente per ogni *locus*.

TANDEM opera attraverso un algoritmo che fornisce alleli per *bin* sotto forma di lunghezza in bp (un numero intero). Il modulo di *repeat* può essere sia stabilito sulla base dei dati osservati dal programma stesso, che impostato manualmente.

In questo lavoro, le prove di *binning* con FLEXIBIN sono state effettuate usando sia 2 che 1 *repeat*. Anche con TANDEM sono state effettuate le due prove di impostazione sia automatica che manuale dell'unità di *repeat*.

Mediante FLEXIBIN e TANDEM sono stati generati due *data set*, in seguito confrontati in termini di numero di alleli (N_a) per *locus*, frequenza dell'allele più comune (MCAF) per ogni campione, eterozigosità osservata (H_o) e probabilità di equilibrio di HW (pHWE). Per le successive analisi di differenziamento genetico è stato usato uno solo dei due *data set*.

ANALISI DEI DATI

Tutti i *file* di *input* utilizzati per l'analisi di differenziamento genetico sono stati preparati con il programma CREATE ver. 1.1 (Coombs et al., 2008). Le statistiche descrittive, come il numero di alleli (N_a), l'eterozigosità osservata (H_o) e la diversità genica *unbiased* (H_E), corrispondente all'eterozigosità attesa pesata sulla dimensione della popolazione, sono stati calcolati utilizzando il programma GENETIX ver. 4.03 (Belkhir et al., 2001). POWERMAKER ver. 3.25 (Liu e Muse, 2005) è stato impiegato invece per calcolare le frequenze dell'allele più comune (MCAF) per ogni *locus*.

Le probabilità di equilibrio di Hardy-Weinberg per ogni *locus* per ogni popolazione sono state stimate mediante i test di probabilità esatti, implementati nel programma GENEPOP ver. *online* della ver. 3.4 (Raymond e Rousset, 1995a) e mediante i test del χ^2 implementati in GENALEX ver. 6 (Peakall e Smouse, 2006). Le probabilità di equilibrio globale sono state combinate attraverso il metodo esatto di Fisher (Manly, 1984).

I test di *linkage disequilibrium* tra coppie di *loci* per ogni popolazione, i test di differenziamento esatti per tutte le popolazioni e per tutte le coppie di popolazioni sono stati implementati in GENEPOP.

MICROCHECKER ver. 2.2.3 (van Oosterhout et al., 2004) è stato utilizzato per verificare la presenza di errori di *scoring* e alleli nulli.

FSTAT ver. 2.9.3.2 (Goudet, 2002) è stato utilizzato per calcolare l'*allelic richness* (A_R), gli indici F_{ST} complessivi e per coppie di popolazioni e gli indici F_{IS} per ogni locus per ogni popolazione.

Gli indici F_{ST} standardizzati (F'_{ST}) sono stati calcolati utilizzando RECODEDATA ver. 0.1 (Meirmans, 2006) e FSTAT.

L'uso di più programmi è dovuto al fatto che non sempre un *software* è adatto a calcolare tutte le statistiche per la genetica di popolazione e al fatto che, per le elaborazioni comuni, questo permette una verifica incrociata dei dati.

Variabilità genetica, equilibrio di Hardy-Weinberg e *linkage disequilibrium*

Il principio d'equilibrio di Hardy-Weinberg rappresenta la legge base della genetica di popolazioni. Esso afferma che in una popolazione infinitamente grande, in cui si verifica un incrocio casuale (popolazione panmittica), sulla quale non agiscono mutazione, migrazione e selezione naturale, le frequenze alleliche non variano di generazione in generazione, mentre le frequenze genotipiche sono determinate dal prodotto delle frequenze alleliche corrispondenti (Hartl e Clark, 1997). Per un *locus* autosomico con m alleli (A_1, A_2, \dots, A_m), le frequenze genotipiche date dalla legge di Hardy-Weinberg sono semplicemente:

per gli omozigoti per l'allele i -esimo (A_iA_i) $\rightarrow p_i^2$,
per gli eterozigoti (A_iA_j) $\rightarrow 2p_i p_j$,

dove p_i = frequenza allelica di $A_i = (2n_1+n_2)/2N$, n_1 è il numero di omozigoti per l'allele i -esimo, n_2 il numero di eterozigoti per l'allele i -esimo e N è il numero di individui.

Il test di verifica dell'equilibrio di HW (Raymond e Rousset, 1995a) normalmente utilizzato calcola le frequenze genotipiche attese in base a quelle alleliche osservate, e le confronta con le frequenze genotipiche osservate. L'ipotesi di equilibrio di HW viene rigettata valutando il grado di deviazione tra frequenze genotipiche attese ed osservate mediante il test statistico del χ^2 . Per i microsatelliti, la difficoltà di applicare questo tipo di test dipende dal fatto che alcuni loci possono avere un elevato numero di alleli e quindi molti possibili genotipi a bassa frequenza.

Eterozigosità osservata (H_O) e eterozigosità *unbiased* (H_E = eterozigosità attesa, pesata sulla dimensione della popolazione) sono state stimate dal *software* GENETIX.

I test di probabilità di equilibrio di HW sono stati implementati per ogni *locus* per ogni popolazione nel programma GENEPOP. Nel caso di presenza di quattro o meno alleli, il programma utilizza il test esatto di HW. Se il numero di alleli è superiore a quattro, viene effettuata una stima *unbiased* della probabilità esatta di equilibrio di HW che si basa sul metodo delle catene di Markov (Guo e Thompson, 1992). In entrambi i casi, GENEPOP restituisce la probabilità di errore di I tipo, ovvero di rifiutare H_0 quando questa è vera (cioè nel caso di equilibrio entro una popolazione). La soglia α utilizzata era pari a 0.05.

Nei casi in cui non è stato possibile utilizzare GENEPOP (per il *locus Ca21*) per il calcolo di equilibrio di HW, è stato utilizzato il programma GENEALLEX. In questo caso, la stima dell'equilibrio è basata sulla la statistica del chi quadro (χ^2). La soglia α era pari anche in questo caso a 0.05.

Il disequilibrio di HW può verificarsi a causa degli errori commessi durante la fase di *scoring* dei dati delle sequenze microsatellite (DeWoody et al., 2006) o per effetto dell'esistenza di fenomeni biologici e/o evolutivi. Gli errori di *scoring*, già descritti in precedenza, possono avere un effetto sull'interpretazione dei dati in chiave ecologica o evolutiva.

I fenomeni biologici e/o evolutivi che sono in grado di provocare il disequilibrio di Hardy-Weinberg sono costituiti da *inbreeding*, cioè riproduzione non casuale, presenza di una struttura di popolazione, deriva genetica e selezione naturale.

Il *linkage disequilibrium* (associazione gametica preferenziale) è definito come il fenomeno per cui, a livello di popolazione, specifiche combinazioni di alleli a due *loci* abbiano una frequenza diversa da quella attesa nel caso in cui i due *loci* fossero indipendenti (Hartl e Clark, 1997). È importante verificare questa associazione in quanto le analisi di differenziamento presuppongono associazione casuale tra alleli di diversi *loci* (*linkage equilibrium*).

La condizione di disequilibrio si verifica quando i *loci* sono strettamente concatenati per cui gli eventi di ricombinazione tra essi sono rari. Essa però può anche essere determinata dal mescolamento di popolazioni con differenti frequenze gametiche, dalla selezione naturale in favore di alcuni genotipi eterozigoti o dalla deriva genetica casuale.

Per verificare la condizione di *linkage disequilibrium* è stato utilizzato il programma GENEPOP che crea automaticamente le tabelle di contingenza

corrispondenti a tutte le possibili coppie di *loci* in ogni popolazione. Le tabelle vengono analizzate utilizzando il metodo delle catene di Markov (privo di *bias*) per stimare il *P-value* esatto partendo dall'ipotesi H_0 che i genotipi di un *locus* siano indipendenti dai genotipi di un altro *locus* (Raymond e Rousset, 1995b). Il valore dei *P-value* ottenuti mediante i confronti multipli è stato aggiustato con la correzione di Bonferroni (α globale = 0,05)

Allelic richness (A_R), o numero medio degli alleli per *locus*, è una misura di diversità genetica legata al numero di alleli diversi presenti in un campione. Poiché il numero degli alleli dipende dalla dimensione del campione, in quanto ci si attende che grandi campioni contengano molti più alleli rispetto a quelli piccoli, l'*allelic richness* fornisce una misura del numero di alleli standardizzata rispetto al campione più piccolo, in modo da poter confrontare tra loro, in maniera coerente, tutti i campioni di popolazioni (Leberg, 2002). La stima di A_R è stata eseguita attraverso il programma FSTAT .

Differenziamento genetico

Per analizzare la distribuzione delle frequenze alleliche è stato utilizzato il test esatto per il differenziamento genico (Raymond e Rousset, 1995b), implementato nel programma GENEPOP (Raymond e Rousset, 1995a) che si basa sull'analisi di tabelle di contingenza in cui le righe corrispondono ad ogni popolazione e le colonne ai conteggi di ogni allele. Questo test applica un metodo di simulazione che utilizza le catene di Markov per ottenere delle stime non distorte della probabilità globale di omogeneità genetica. In dettaglio, il test esplora in modo efficiente l'insieme di tutte le possibili tabelle di contingenza caratterizzate dalle stesse somme marginali osservate nei dati reali. Viene testata l'ipotesi nulla dell'assenza di differenziamento tra popolazioni. In altri termini, il test verifica l'ipotesi di distribuzione allelica omogenea per tutti i campioni.

Questo approccio è stato poi esteso a più *loci* combinando i risultati mediante il test esatto di Fisher.

Il fenomeno dell'eccesso di omozigotità, indice di suddivisione delle popolazioni e dell'*inbreeding*, viene comunemente studiato applicando le statistiche F, definite da Wright (1978) e riderivate da Weir e Cockerham in termini di analisi della varianza (1984).

Secondo Wright, le popolazioni possiedono tre livelli di complessità, che possono essere stimati mediante tre indici:

- F_{IS} , che misura la consanguineità (*inbreeding*) relativa degli individui (I) rispetto alla popolazione (S) a cui appartengono.
- F_{ST} , che misura la fissazione di un allele, dovuta alla deriva genetica casuale, in termini di riduzione dell'eterozigosità della popolazione (S) rispetto alla popolazione complessiva (T). Questo indice viene utilizzato per stimare la suddivisione delle popolazioni (S).
- F_{IT} , che misura la consanguineità relativa degli individui (I), rispetto alla popolazione totale (T).

In questo modo, la variabilità genetica può essere suddivisa in una componente intra- (F_{IS}) e una tra-popolazioni (F_{ST}). Sotto l'ipotesi nulla (cioè di panmixia per F_{IS} e omogeneità delle frequenze alleliche per F_{ST}), questi indici sono uguali a zero.

Un valore positivo di F_{IS} corrisponde a un deficit di eterozigoti, mentre un valore negativo indica un loro eccesso. Il deficit può indicare una strutturazione delle popolazioni (S), mentre un eccesso di eterozigoti può essere un indice di accoppiamento non casuale (*inbreeding*) o maggiore *fitness* degli eterozigoti.

L'indice F_{ST} è sempre positivo e varia fra 0 e 1 per un *locus* biallelico, per cui, se tutte le popolazioni (S) sono in equilibrio di HW e hanno tutte le stesse frequenze alleliche, $F_{ST} = 0$, ad indicare che le popolazioni (S) fanno parte della stessa popolazione complessiva (T); viceversa, $F_{ST} = 1$ indica una separazione totale fra le popolazioni (S).

Weir e Cockerham (1984) forniscono una stima di questi indici basata sul modello delle isole, dove la migrazione e il *drift* sono i fenomeni principali. Questi indici vengono chiamati θ (theta, che corrisponde all' F_{ST} di Wright), f (f piccolo, che corrisponde all' F_{IS} di Wright) e F (che corrisponde all' F_{IT} di Wright).

Attualmente, molti studi di genetica di popolazione usano gli indici F di Weir e Cockerham, il cui calcolo è incluso in molti pacchetti statistici. Gli indici theta (θ) e f piccolo (f) sono stati utilizzati anche per questa tesi e vengono riportati rispettivamente come $F_{ST}(\theta)$ e $F_{IS}(f)$, dove F_{ST} è definito come la frazione della variazione genetica totale attribuibile alle differenze tra popolazioni, mentre F_{IS} è definito come la frazione attribuibile alle differenze entro gli individui.

La suddivisione delle popolazioni per deriva genetica casuale (*drift*) e *inbreeding* sono fenomeni in parte sovrapponibili che provocano entrambi una deviazione dalle frequenze genotipiche attese previste dall'equilibrio di HW verso un eccesso di omozigoti. La loro sovrapposizione deriva dal fatto che il *drift* produce una perdita della variabilità genetica, aumentando così la probabilità di unire insieme geni derivanti da un antenato comune.

Un fenomeno opposto al *drift*, che tende a riportare le frequenze genotipiche e geniche delle popolazioni all'equilibrio, è rappresentato dalla migrazione. L'indice F_{ST} rappresenta una stima sia del flusso genico che della fissazione (che avviene per effetto della deriva genetica casuale nelle popolazioni di dimensioni finite). In questo quadro, la condizione d'equilibrio è determinata dall'equilibrio drift-migrazione, mentre la mutazione può essere trascurata in quanto molto inferiore alle prime due. Questa situazione è rappresentata nel modello infinito delle isole (Maynard Smith, 1970), collegate da un flusso migratorio consistente in m migranti per generazione che provengono da un'intera metapopolazione. Secondo tale modello le due quantità hanno un identico effetto atteso sull' F_{ST} .

In questo lavoro, il differenziamento genico è stato quindi prima calcolato in termini di probabilità complessiva e per coppie di popolazioni attraverso il programma GENEPOP ver. *online* della ver. 3.4 (Raymond e Rousset, 1995a), e poi espresso in termini di distanze genetiche utilizzando il parametro F_{ST} (θ) di Weir e Cockerham (1984).

Il calcolo degli indici F_{ST} , implementato in FSTAT, è stato eseguito prima a livello globale, su tutte le popolazioni, e, successivamente, su coppie di popolazioni (pairwise F_{ST}). La stima delle probabilità di F_{ST} è stata determinata con permutazioni 1000 del *data set* originale. Gli intervalli di confidenza al 95 % per ogni singolo valore sono stati stimati mediante il metodo di ricampionamento numerico *bootstrap*.

L'interpretazione del differenziamento genetico basata sull'indice F_{ST} può essere resa difficoltosa dal fatto che alcuni *loci* possono contribuire in modo diverso o disuguale alla variabilità genetica totale. Una misura standardizzata del differenziamento tra popolazioni (F'_{ST}), indipendente dai livelli di eterozigotità e dalla dimensione del campione, è stata quindi utilizzata per permettere confronti tra le popolazioni (Hedrick, 2005). Il programma RECODEDATA è stato utilizzato per ricodificare il *data set*, in modo che ogni campione di popolazione avesse alleli unici, permettendo così la stima dei livelli massimi di differenziamento tra popolazioni. Per stimare il livello standardizzato della suddivisione tra popolazioni (F'_{ST}), ogni attuale stima del differenziamento tra popolazioni (F_{ST}) è stata divisa per il valore massimo di F_{ST} , ottenuto dal *data set* ricodificato. Gli intervalli di confidenza al 95% sono stati stimati con FSTAT mediante 1000 ricampionamenti *bootstrap* su tutti i *loci*, mentre le probabilità sono state determinate attraverso 1000 permutazioni.

RISULTATI

ESTRAZIONE DEL DNA

L'estrazione del DNA con il protocollo *Salting Out* (Patwary et al., 1994) si è rivelato un metodo affidabile ed efficace per il DNA genomico di questa specie. L'amplificazione dell'estratto è riuscita per tutti i campioni, fornendo, per ognuno, una quantità di DNA stampo sufficiente per l'attività della *Taq* polimerasi. Non è stato quindi necessario riestrarre campioni applicando altri protocolli. È probabile che ulteriori fattori coadiuvanti l'efficacia dell'estrazione siano il tipo e i tempi di conservazione degli organismi. Infatti i campioni di *C. aceratus* sono stati raccolti specificamente allo scopo di condurre analisi genetiche, manipolati rapidamente al momento della raccolta e immediatamente posti in etanolo al 90%.

AMPLIFICAZIONE IN MULTIPLEX

Inizialmente, pochi individui sono stati genotipizzati (7) per ottimizzare le condizioni di corsa al sequenziatore e verificare la qualità dei singoli *loci* microsatellite.

In tutto, sono state messe a punto 3 combinazioni di *loci* da amplificare in *multiplex*.

L'ottimizzazione delle quantità dei *primers* è stata resa necessaria a causa della diversa intensità dei marcatori nella genotipizzazione. I *primers Forward* marcati che davano un segnale troppo intenso, sono stati diluiti con *Forward* non marcato. Le concentrazioni per la *mix* di amplificazione e le combinazioni dei *loci* ottenuti con la strategia di ottimizzazione in *multiplex* sono riportate nella tabella 5.

MULTIPLEX 1			MULTIPLEX 2			MULTIPLEX 3		
Conc. primers			Conc. primers			Conc. primers		
	F	R		F	R		F	R
<i>Ca86</i>	0.1 µM	0.1 µM	<i>Cr259</i>	0.2 µM	0.2 µM	<i>Ca88</i>	0.2 µM	0.2 µM
<i>Ca21</i>	0.2 µM	0.2 µM	<i>Ca26</i>	0.3 µM	0.3 µM	<i>Ca48*</i>	0.1µM	0.1µM
<i>Cr121</i>	0.4 µM	0.4 µM	<i>Ca40</i>	0.3 µM	0.3 µM	<i>Cr15*</i>	0.1µM	0.1µM
<i>Cr171</i>	0.2 µM	0.2 µM				<i>Ca55*</i>	0.1µM	0.1µM

Tabella 5. Combinazioni dei *loci* e concentrazioni ottimizzate dei *primers*. Sono riportati i *loci* amplificati in ognuna delle tre *mixmultiplex* e le relative concentrazioni dei *primers Forward* (F) e *Reverse* (R). Nei *loci* segnati con l'asterisco (*) la concentrazione riportata per i *primers Forward* comprende metà volume del *primer* marcato col fluoroforo e metà non marcato. Il profilo di amplificazione è identico per le tre *mix* ed è riportato nei Materiali e Metodi.

ALLELE CALLING O BINNING

Sono stati generati due *data set* con i programmi FLEXIBIN e TANDEM. In FLEXIBIN, il *binning* è stato effettuato impostando il modulo dinucleotidico per tutti i *loci* tranne che per i *loci Ca21, Ca40 e Cr171*, per i quali è stato ritenuto più adatto scegliere un modulo mononucleotidico. Infatti, le classi alleliche generate per questi tre *loci* con modulo dinucleotidico, presentavano deviazione standard troppo grande ($DS \geq 0.4$) e comprendevano più di 1 bp. In TANDEM, il processo di assegnazione automatica delle classi alleliche ha assegnato un *repeat* dinucleotidico a tutti i *loci*.

ANALISI DEI DATI

Variabilità genetica, equilibrio di Hardy-Weinberg e *linkage disequilibrium*

Le analisi statistiche descrittive di variabilità genetica sono state condotte utilizzando gli input dei programmi di *scoring* FLEXIBIN e TANDEM. Le analisi descrittive sono state eseguite quasi esclusivamente su 518 individui che costituivano l'intero *set* di campioni (5 analizzati per la prima volta in questo studio e 4 analizzati negli studi precedenti). I dati utilizzati e i *loci* considerati vengono specificati caso per caso nella descrizione dei risultati di questa sezione.

Per la stima del numero di alleli (N_a) per popolazione e per *locus*, implementata con GENETIX (Belkhir *et al.* 2001), è stato utilizzato l'intero *data set* di individui e tutti gli undici *loci* (Tab. 6).

Il numero di alleli per *locus* (N_a) vanno da 121 (*Ca21*) a 6 (*Cr15*) per il *data set* di FLEXIBIN e da 72 (*Ca21*) a 6 (*Cr15*) per il *data set* di TANDEM. All'interno delle popolazioni, N_a varia da un massimo di 43 (*Ca21*) per il campione EI07, a un minimo di 2 (*Cr15*) per i campioni SS02, SO06 e EI07 secondo il metodo di *binning* di FLEXIBIN. Secondo il *data set* di TANDEM, N_a varia da un massimo di 39 (*Ca21*) per la popolazione EI07 ad un minimo di 2 (*Cr15*) per le popolazioni SS02, SO06 e EI07. Il numero d'alleli della popolazione BI04, composta da soli 4 individui, non viene considerato.

	CS08		DB05		SS02		EI02		EI06		EI07		SO06		SG04		BI04		TOT	
<i>Locus</i>	F	T	F	T	F	T	F	T	F	T	F	T	F	T	F	T	F	T	F	T
<i>Cr15</i>	3	3	3	3	2	2	5	5	3	3	2	2	2	2	4	4	1	1	6	6
<i>Ca21</i>	41	34	39	35	27	24	56	43	42	31	43	39	36	28	30	29	7	7	121	72
<i>Ca26</i>	10	10	10	10	19	19	25	25	16	16	9	9	14	14	10	10	3	3	31	31
<i>Ca40</i>	35	26	22	20	30	22	44	28	28	21	31	25	31	24	32	23	6	6	50	30
<i>Ca48</i>	4	4	4	4	4	4	5	5	5	5	5	5	5	5	4	4	2	2	7	7
<i>Ca55</i>	18	18	24	24	21	21	26	26	22	22	24	24	19	19	20	20	5	5	31	31
<i>Ca86</i>	29	29	27	27	22	22	42	42	19	19	28	28	22	22	28	28	3	3	52	52
<i>Ca88</i>	14	14	20	20	14	14	15	15	16	16	15	14	16	16	17	17	5	5	25	25
<i>Cr127</i>	6	6	8	8	7	7	9	9	7	7	6	6	7	7	8	8	3	3	10	10
<i>Cr171</i>	29	24	22	19	25	20	40	26	30	21	23	20	38	28	19	15	3	3	67	39
<i>Cr259</i>	18	18	24	24	16	16	21	21	17	17	17	17	19	19	25	24	7	7	32	32

Tabella 6. Numero di alleli (N_a) stimato per tutti gli undici *loci* per ognuna delle nove popolazioni considerate. Legenda: F = N_a stimato con l'input di FLEXIBIN; T = N_a stimato con l'input di TANDEM. TOT = N_a totale per ogni *locus* stimati con gli *input* di FLEXIBIN e di TANDEM.

Le frequenze dell'allele più comune (MCAF) sono state calcolate attraverso POWERMAKER ver. 3.25 (Liu e Muse, 2005) usando tutti gli individui e tutti gli 11 *loci* (Tab. 7) MCAF varia da 0.888 (*Ca48*) a 0.091 (*Ca 21*) per il metodo di

binning di FLEXIBIN e da 0.888 (*Ca48*) a 0.116 (*Ca40*) per quello di TANDEM. I risultati dei due *data set* risultano molto simili, eccetto che per i *loci Ca21* e *Cr171*. In particolare, la frequenza di MCAF è pari a 0.091 (*Ca21*) e a 0.099 (*Cr171*) per il *data set* ottenuto mediante FLEXIBIN, mentre è uguale a 0.135 (*Ca21*) e a 0.157 (*Cr171*) per il *data set* ottenuto con TANDEM.

Loc	F	T
<i>Cr15</i>	0.830	0.830
<i>Ca21</i>	0.091	0.135
<i>Ca26</i>	0.292	0.292
<i>Ca40</i>	0.116	0.116
<i>Ca48</i>	0.888	0.888
<i>Ca55</i>	0.126	0.126
<i>Ca86</i>	0.139	0.139
<i>Ca88</i>	0.161	0.161
<i>Cr127</i>	0.401	0.401
<i>Cr171</i>	0.099	0.157
<i>Cr259</i>	0.320	0.318
MEDIA	0.315	0.324

Tabella 7. Frequenze dell'allele più comune (MCAF) per ogni *locus* ottenute con POWERMAKER per i risultati di *binning* di FLEXIBIN (F) e di TANDEM (T) per l'intero *data set* di individui e per tutti gli undici *loci*.

Anche l'eterozigosità, stimata mediante GENETIX, è stata calcolata sull'intero *data set* di individui usando tutti gli undici *loci* per metodi di *binning* di TANDEM e di FLEXIBIN (Tab. 8 a) e b)). I risultati sono complessivamente molto simili, tranne che per i *loci Ca21* e *Cr171*, per i quali si osserva una discrepanza nell'eterozigosità attesa (H_E) a seconda del metodo utilizzato. Per il *locus Cr171*, in relazione al metodo di *binning*, si osservano anche differenze nella stima dell'eterozigosità osservata (H_O).

Da entrambi i metodi di *binning* risulta un eccesso di omozigoti, dal momento che l'eterozigosità osservata (H_O) è quasi sempre inferiore a quella attesa (H_E). Anche nei casi, in cui H_O è maggiore di H_E , i risultati dei due programmi sono simili, ad eccezione della stima di H_E per il *locus Cr171* per la popolazione EI06.

L'eterozigosità complessiva osservata per il *data set* di TANDEM varia da un massimo di 0.7233 per la popolazione SG04, ad un minimo di 0.6863 e di 0.6867 per le popolazioni, rispettivamente, di SO06 e di DB05, escludendo la popolazione BI04 in quanto limitante per questa analisi. I due metodi di *binning*, anche in questo caso, forniscono la stessa

a)

<i>Locus T</i>	CS08	DB05	SS02	EI02	EI06	EI07	SO06	SG04	BI04
<i>Cr15</i>	0.1400/ 0.2883	0.2353/ 0.4372	0.1364/ 0.1285	0.2284/ 0.2794	0.2766/ 0.3301	0.1667/ 0.2137	0.1458/ 0.2050	0.5152/ 0.4811	0.0000/ 0.0000
<i>Ca21</i>	0.9388/ 0.9495	0.8000/ 0.9549	0.7556/ 0.9401	0.8400/ 0.9368	0.8913/ 0.9274	0.9630/ 0.9524	0.7708/ 0.9388	0.8182/ 0.9552	0.7500/ 0.9643
<i>Ca26</i>	0.5417/ 0.7711	0.5172/ 0.8403	0.8000/ 0.8594	0.8144/ 0.8437	0.7727/ 0.8529	0.5435/ 0.8223	0.6889/ 0.7576	0.3548/ 0.7070	0.5000/ 0.7500
<i>Ca40</i>	0.8200/ 0.9414	0.9310/ 0.9359	0.8889/ 0.9413	0.8995/ 0.9353	0.6818/ 0.9441	0.8704/ 0.9408	0.7826/ 0.9398	0.9091/ 0.9580	0.7500/ 0.8929
<i>Ca48</i>	0.1200/ 0.1164	0.0571/ 0.3573	0.2000/ 0.1880	0.1782/ 0.1807	0.2222/ 0.2454	0.1852/ 0.2075	0.2500/ 0.2292	0.2424/ 0.2256	0.2500/ 0.5357
<i>Ca55</i>	0.9200/ 0.9222	0.8529/ 0.9368	0.8222/ 0.9378	0.9000/ 0.9355	0.8936/ 0.9387	0.9057/ 0.9441	0.8958/ 0.9202	0.9375/ 0.9479	1.0000/ 0.8571
<i>Ca86</i>	0.8367/ 0.9356	0.8286/ 0.9462	0.8864/ 0.9269	0.8769/ 0.9288	0.8478/ 0.9152	0.8519/ 0.9413	0.8085/ 0.9250	0.9697/ 0.9534	0.2500/ 0.4643
<i>Ca88</i>	0.9184/ 0.9049	0.7941/ 0.9241	0.8837/ 0.8908	0.8750/ 0.9017	0.9787/ 0.9092	0.9074/ 0.8846	0.8125/ 0.9145	0.8485/ 0.9212	0.7500/ 0.7857
<i>Cr127</i>	0.8600/ 0.7903	0.7714/ 0.8265	0.7333/ 0.7146	0.7015/ 0.7418	0.7234/ 0.7680	0.6981/ 0.7332	0.7708/ 0.7175	0.6364/ 0.7566	0.5000/ 0.4643
<i>Cr171</i>	0.8163/ 0.9131	0.8519/ 0.8784	0.8095/ 0.9028	0.6648/ 0.8986	0.9302/ 0.9239	0.7879/ 0.9282	0.9149/ 0.9494	0.7857/ 0.9127	0.5000/ 0.8333
<i>Cr259</i>	0.8600/ 0.8390	0.9143/ 0.8961	0.8140/ 0.8219	0.8200/ 0.8410	0.6596/ 0.8273	0.8148/ 0.8584	0.7083/ 0.8996	0.9394/ 0.9478	1.0000/ 0.9643
MEDIA	0.7065/ 0.7611	0.6867/ 0.8121	0.7027/ 0.7502	0.7090/ 0.7658	0.7162/ 0.7802	0.6995/ 0.7660	0.6863/ 0.7633	0.7233/ 0.7970	0.5682/ 0.6829

b)

<i>Locus F</i>	CS08	DB05	SS02	EI02	EI06	EI07	SO06	SG04	BI04
<i>Cr15</i>	0.1400/ 0.2883	0.2353/ 0.4372	0.1364/ 0.1285	0.2284/ 0.2794	0.2766/ 0.3301	0.1667/ 0.2137	0.1458/ 0.2050	0.5152/ 0.4811	0.0000/ 0.0000
<i>Ca21</i>	0.9388/ 0.9647	0.8000/ 0.9665	0.7556/ 0.9423	0.8400/ 0.9418	0.8913/ 0.9446	0.9630/ 0.9595	0.7708/ 0.9500	0.8182/ 0.9557	0.7500/ 0.9643
<i>Ca26</i>	0.5417/ 0.7711	0.5172/ 0.8403	0.8000/ 0.8594	0.8144/ 0.8437	0.7727/ 0.8529	0.5435/ 0.8223	0.6889/ 0.7576	0.3548/ 0.7070	0.5000/ 0.7500
<i>Ca40</i>	0.8200/ 0.9602	0.9310/ 0.9425	0.8889/ 0.9608	0.8995/ 0.9465	0.6818/ 0.9577	0.8704/ 0.9476	0.7826/ 0.9472	0.9091/ 0.9692	0.7500/ 0.8929
<i>Ca48</i>	0.1200/ 0.1164	0.0571/ 0.3573	0.2000/ 0.1880	0.1782/ 0.1807	0.2222/ 0.2454	0.1852/ 0.2075	0.2500/ 0.2292	0.2424/ 0.2256	0.2500/ 0.5357
<i>Ca55</i>	0.9200/ 0.9222	0.8529/ 0.9368	0.8222/ 0.9378	0.9000/ 0.9355	0.8936/ 0.9387	0.9057/ 0.9441	0.8958/ 0.9202	0.9375/ 0.9479	1.0000/ 0.8571
<i>Ca86</i>	0.8367/ 0.9356	0.8286/ 0.9462	0.8864/ 0.9269	0.8769/ 0.9282	0.8478/ 0.9152	0.8519/ 0.9413	0.8085/ 0.9250	0.9697/ 0.9534	0.2500/ 0.4643
<i>Ca88</i>	0.9184/ 0.9057	0.7941/ 0.9263	0.8837/ 0.8908	0.8750/ 0.9017	0.9787/ 0.9092	0.9074/ 0.8856	0.8125/ 0.9145	0.8485/ 0.9212	0.7500/ 0.7857
<i>Cr127</i>	0.8600/ 0.7903	0.7714/ 0.8265	0.7333/ 0.7146	0.7015/ 0.7418	0.7234/ 0.7714	0.6981/ 0.7332	0.7708/ 0.7175	0.6364/ 0.7566	0.5000/ 0.4643
<i>Cr171</i>	0.8163/ 0.9352	0.8519/ 0.9043	0.8333/ 0.9263	0.7374/ 0.9392	0.9302/ 0.9513	0.7879/ 0.9455	0.9362/ 0.9693	0.8571/ 0.9524	0.5000/ 0.8333
<i>Cr259</i>	0.8600/ 0.8390	0.9143/ 0.8957	0.8140/ 0.8219	0.8200/ 0.8410	0.6596/ 0.8273	0.8148/ 0.8584	0.7083/ 0.8996	0.9394/ 0.9483	1.0000/ 0.9643
MEDIA	0.7084/ 0.7662	0.6867/ 0.8163	0.7049/ 0.7543	0.7156/ 0.7709	0.7162/ 0.7858	0.6995/ 0.7690	0.6882/ 0.7668	0.7298/ 0.8017	0.5682/ 0.6829

Tabella 8: Eterozigosità osservata e attesa per ogni popolazione per ogni singolo *locus* realizzate con l'output di TANDEM (a) e di FLEXIBIN (b). Sono riportate l'eterozigosità osservata (H_O) / eterozigosità *unbiased* (H_E) e la media (H_O) / (H_E) per ogni popolazione.

Le probabilità di equilibrio di Hardy-Weinberg (Tab. 9 a) e b)) sono state calcolate per tutti gli undici *loci* e per tutte le popolazioni utilizzando i *data set* di FLEXIBIN e di TANDEM. L'analisi è stata effettuata mediante GENEPOP, ver. on-line della ver.3.4 (Raymond e Rousset, 2005a) per tutti i *loci* eccetto che per il *locus Ca21* dove la presenza di un numero di alleli superiore a 100 nel data set di FLEXIBIN ha richiesto l'utilizzo di GENEALEX ver. 6 (Peakall e Smouse, 2006). Nove *loci* su undici (*Cr15, Ca21, Ca26, Ca40, Ca48, Ca55, Ca86, Cr171 e Cr259*) si trovano in disequilibrio di HW combinando le popolazioni con il metodo esatto di Fisher (soglia di significatività $\alpha = 0.05$) ed il risultato non dipende dal metodo di *binning*.

a)

<i>Locus</i>	CS08	DB05	SS02	EI02	EI06	EI07	SO06	SG04	BI04	TOT
<i>Cr15</i>	0.001	0.001	1	0.018	0.465	0.15	0.095	1	-	<0.001
<i>Ca21</i>	0.645	0.164	<0.001	<0.001	0.004	0.995	<0.001	0.101	0.293	<0.001
<i>Ca26</i>	<0.001	<0.001	0.35	0.044	0.591	0.001	0.738	<0.001	0.659	<0.001
<i>Ca40</i>	<0.001	0.911	0.431	0.248	<0.001	0.092	0.002	0.619	0.43	<0.001
<i>Ca48</i>	1	<0.001	1	0.287	0.084	0.392	1	1	0.428	<0.001
<i>Ca55</i>	0.402	0.074	0.003	0.262	0.295	0.043	0.639	0.623	0.656	0.017
<i>Ca86</i>	0.014	0.036	0.942	0.264	0.033	0.071	0.083	0.927	0.143	0.003
<i>Ca88</i>	0.527	0.022	0.149	0.446	0.751	0.837	0.057	0.075	0.773	0.084
<i>Cr127</i>	0.625	0.131	0.709	0.033	0.214	0.795	0.805	0.099	1	0.273
<i>Cr171</i>	0.006	0.22	<0.001	<0.001	0.574	0.018	0.375	0.138	0.333	<0.001
<i>Cr259</i>	0.449	0.123	0.664	0.581	0.001	0.397	0.002	0.614	1	0.007
TOT T	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.001	<0.001	0.861	<0.001

b)

<i>Locus</i>	CS08	DB05	SS02	EI02	EI06	EI07	SO06	SG04	BI04	TOT
<i>Cr15</i>	0.001	0.001	1	0.018	0.464	0.15	0.096	1	-	<0.001
<i>Ca21</i>	0.357	0.009	<0.001	<0.001	<0.001	0.928	<0.001	0.068	0.293	<0.001
<i>Ca26</i>	<0.001	<0.001	0.384	0.044	0.586	<0.001	0.736	<0.001	0.657	<0.001
<i>Ca40</i>	0.001	0.805	0.065	0.009	<0.001	0.015	<0.001	0.286	0.432	<0.001
<i>Ca48</i>	1	<0.001	1	0.292	0.086	0.393	1	1	0.428	<0.001
<i>Ca55</i>	0.403	0.052	0.004	0.253	0.301	0.041	0.621	0.629	0.659	0.016
<i>Ca86</i>	0.01	0.019	0.939	0.266	0.028	0.068	0.094	0.896	0.143	0.001
<i>Ca88</i>	0.481	0.009	0.149	0.436	0.747	0.894	0.054	0.082	0.772	0.053
<i>Cr127</i>	0.629	0.129	0.714	0.028	0.21	0.795	0.812	0.107	1	0.265
<i>Cr171</i>	0.001	0.076	<0.001	<0.001	0.486	0.007	0.425	0.23	0.333	<0.001
<i>Cr259</i>	0.365	0.083	0.665	0.559	0.005	0.409	0.002	0.263	0.421	0.004
TOT F	<0.001	0.009	0.768	<0.001						

Tabella 9: Probabilità di equilibrio di Hardy-Weinberg (pHWE) per ogni *locus* per ciascuna popolazione e la probabilità totale per ogni popolazione. a) probabilità per il metodo di *binning* di TANDEM. b) Probabilità per il metodo di *binning* di FLEXIBIN; Le probabilità complessive di HWE sono state ottenute applicando il metodo esatto di Fisher. In *grassetto corsivo* sono evidenziate le probabilità significative ($P\text{-value} < 0.05$).

La condizione di *linkage disequilibrium*, stimato con GENEPOP, è stata trovata solo fra la coppia di *loci* *Cr15* e *Ca48* della popolazione DB05 cioè, in solo 1 su 360 test effettuati. Dal momento che in altre popolazioni questi stessi *loci* si trovavano in equilibrio si può ritenere che tale disequilibrio sia dovuto ad effetti di campionamento casuale. A livello globale, infatti, tutti i *loci* sono risultati in *linkage equilibrium*

Sulla base dei risultati ottenuti, ed in particolare per le differenze in MCAF, H_E e H_O , rilevate tra metodi di *binning*, i *loci* *Ca21* e *Cr171* sono stati ritenuti inaffidabili ed esclusi dalle analisi successive in quanto potenzialmente in grado di interferire con i risultati. L'esclusione è anche giustificata dal fatto che il *locus* *Ca21* presentava un numero eccessivo di alleli ($N_a = 121$ stimati con FLEXIBIN) rispetto alla soglia considerata come accettabile in letteratura ed il *locus* *Cr171* presentava un numero elevato di individui la cui genotipizzazione era fallita, soprattutto appartenenti ai campioni analizzati in questo studio (22 su 54 per la popolazione EI07 e 19 su 33 per la popolazione SG04).

Dopo l'eliminazione dei due *loci* *Ca21* e *Cr171*, le analisi successive hanno fornito risultati molto simili, indipendentemente dal metodo di *binning*. Per questo motivo si riportano nel seguito le sole analisi effettuate col *data set* ottenuto mediante TANDEM.

Per quanto riguarda la stima dell'*allelic richness* (A_R), implementata con FSTAT ver. 2.9.3.2 (Goudet, 2002), oltre all'esclusione dei *loci* *Ca21* e *Cr171*, è stata eliminata la popolazione di *Bouvet Island* (BI04) e tutti gli individui la cui genotipizzazione era fallita per almeno uno dei nove *loci* considerati (Tab. 10)

La popolazione BI04, costituita da un numero troppo piccolo di individui (4), è stata esclusa per rendere la stima di A_R più precisa. Infine, l'esclusione di individui non genotipizzati è stata richiesta dal supporto di analisi statistica utilizzato (FSTAT).

Locus	CS08	DB05	SS02	EI02	EI06	EI07	SO06	SG04	TOT
<i>Cr15</i>	2.522	2	1.993	2.276	2.615	2	1.999	3.6	2.322
<i>Ca26</i>	7.84	8	14.163	13.397	13.354	8.374	10.898	9.087	12.208
<i>Ca40</i>	20.149	17	19.132	18.415	16.164	19.979	19.374	21.591	19.317
<i>Ca48</i>	3.322	3	3.704	3.29	4.41	3.439	3.922	3.919	3.497
<i>Ca55</i>	15.869	17	17.824	17.971	18.691	18.159	16.144	18.191	17.945
<i>Ca86</i>	20.402	18	16.865	19.76	15.777	20.774	17.645	24.599	19.805
<i>Ca88</i>	12.407	14	11.581	12.663	13.729	12.105	14.034	15.68	13.057
<i>Cr127</i>	5.774	6	6.605	6.669	6.593	5.979	6.598	6.792	6.514
<i>Cr259</i>	14.492	16	13.289	14.934	13.46	14.468	14.411	20.835	14.981

Tabella 10. *Allelic richness* (A_R) stimata escludendo dagli undici *loci* iniziali i *loci* *Ca21* e *Cr171*, la popolazione di *Bouvet Island* e gli individui, la cui genotipizzazione era fallita per almeno uno dei *loci* considerati utilizzando il metodo di *binning* di TANDEM. TOT = A_R totale per *locus* ottenuti con FSTAT.

Il numero medio di alleli per *locus* (A_R) varia complessivamente da 19.805 (*Ca86*) e 19.317 (*Ca40*) a 2.322 (*Cr15*). Questi risultati sono confermati a livello intrapopolazione, con il valore di A_R che varia da 24.599 (*Ca86*) e 21.591 (*Ca40*) nella popolazione SG04 a 1.993 (*Cr15*) nella popolazione SS02.

Alleli nulli ed errori di *scoring*

La presenza di alleli nulli e di altri errori di *scoring* è stata verificata mediante MICROCHECKER ver. 2.2.3 (van Oosterhout et al., 2004) su tutte e nove le popolazioni eliminando gli individui la cui genotipizzazione non era riuscita per almeno uno dei *loci* considerati. I *loci* *Ca21* e *Cr171* sono stati esclusi definitivamente dall'analisi.

L'analisi statistica ha rivelato un generale eccesso di genotipi omozigoti. Il tipo principale d'errore di *scoring* evidenziato è costituito dalla presenza di alleli nulli, suggerita dal generale eccesso di omozigoti per la maggior parte dei *loci*.

Gli errori di *stuttering*, evidenziati da una mancanza significativa di genotipi eterozigoti nel caso di alleli che differiscono di un *repeat*, sono stati rilevati solo in alcune popolazioni per i *loci* *Cr15* e *Cr259*.

Non è stata evidenziata alcuna presenza di *large allele dropout*.

Locus	CS08	DB05	SS02	EI02	EI06	EI07	SO06	SG04	BI04
<i>Cr15</i>	AN/ST	AN		AN/ST					NT
<i>Ca26</i>	AN	AN				AN		AN	NT
<i>Ca40</i>	AN				AN		AN		NT
<i>Ca48</i>		AN							NT
<i>Ca55</i>			AN						NT
<i>Ca86</i>	AN	AN		AN		AN	AN		NT
<i>Ca88</i>		AN					AN		NT
<i>Cr127</i>									NT
<i>Cr259</i>					AN/ST		AN/ST		NT

Tabella 11. Risultati dei test per gli errori di *scoring* per *locus* per ogni popolazione messi a confronto con risultati del test di equilibrio di HW. Risultati dei test su MICROCHECKER per il *data set* di TANDEM. Legenda: AN = presenza di alleli nulli; AD = *large allele drop-out*; ST = *stuttering*; NT = *not possibile* (campione troppo piccolo); le caselle vuote corrispondono all'assenza degli errori di *scoring*. In verde vengono evidenziati i *loci* per le popolazioni in disequilibrio di Hardy-Weinberg.

I risultati dei test per gli errori di *scoring*, implementati con il *data set* di TANDEM, sono stati messi a confronto con i risultati del test di equilibrio di HW (vedi Tab. 9a sulle p_{HWE}) nella tabella 11. La sovrapposizione dei due risultati mostra come la situazione di disequilibrio di HW si osserva anche in *loci* privi di

alleli nulli o problemi legati allo *stuttering*, come nel caso di *Ca26* e *Cr127* in EI02, di *Ca86* in EI06 e di *Ca55* in EI07; tale condizione suggerisce che il disequilibrio complessivo di HW possa avere una reale base biologica ed essere dovuto alla presenza di *inbreeding*, piuttosto che di artefatti sperimentali.

DIFFERENZIAMENTO GENETICO

Tassi di *inbreeding* F_{IS}

Il tasso di *inbreeding* F_{IS} è stato stimato per tutte le popolazioni per tutti i *loci* attraverso il programma statistico FSTAT ver. 2.9.3.2 (Goudet, 2002) (Tab. 12). Quasi tutti gli indici F_{IS} risultano positivi, ad indicare un deficit di eterozigoti.

F_{IS}	CS08	DB05	SS02	EI02	EI06	EI07	SO06	SG04	BI04	TOT
<i>Cr15</i>	0.517	0.466	-0.062	0.183	0.164	0.222	0.291	-0.072	NA	0.218
<i>Ca26</i>	0.300	0.389	0.070	0.035	0.095	0.342	0.092	0.502	0.368	0.152
<i>Ca40</i>	0.130	0.005	0.056	0.038	0.280	0.076	0.169	0.052	0.182	0.086
<i>Ca48</i>	-0.032	0.842	-0.065	0.014	0.096	0.108	-0.092	-0.076	0.571	0.113
<i>Ca55</i>	0.002	0.091	0.124	0.038	0.049	0.041	0.027	0.011	-0.200	0.043
<i>Ca86</i>	0.107	0.126	0.044	0.056	0.074	0.096	0.127	-0.017	0.500	0.075
<i>Ca88</i>	-0.015	0.142	0.008	0.030	-0.077	-0.026	0.113	0.080	0.053	0.027
<i>Cr127</i>	-0.089	0.068	-0.027	0.055	0.059	0.048	-0.075	0.161	-0.091	0.029
<i>Cr259</i>	-0.025	-0.021	0.010	0.025	0.204	0.051	0.214	0.009	-0.043	0.051
TOT	0.076	0.171	0.039	0.045	0.101	0.093	0.100	0.080	0.143	0.074
IC	(0.029 – 0.117)	(0.079 – 0.245)	(-0.014 – 0.066)	(0.021 – 0.061)	(0.038 – 0.169)	(0.037 – 0.121)	(0.046 – 0.133)	(0.013 – 0.111)	(-1.000 – 0.156)	

Tabella 12. Tassi di *inbreeding* F_{IS} per ogni *locus* per ogni popolazione. Risultati test su FSTAT per il *data set* di TANDEM. Vengono riportati i valori di F_{IS} complessivo (TOT) per ogni *locus* e per ogni popolazione e gli intervalli di confidenza (IC) al 95 % per tutte le popolazioni. In arancione sono evidenziate le popolazioni, in cui i *loci* si trovano in disequilibrio di HW.

Test di differenziamento fra popolazioni

Per tutte le successive analisi di differenziamento sono stati esclusi i *loci* *Ca21* e *Cr171*. Queste analisi sono quindi state eseguite con i rimanenti nove *loci*. Si è inoltre deciso di tenere solo l'input di TANDEM.

I Test di differenziamento genico sono stati eseguiti con il programma GENEPOP, ver. *online* della ver. 3.4 (Raymond e Rousset, 1995a). L'ipotesi nulla di identica distribuzione allelica lungo tutte le popolazioni è stata testata sia a livello complessivo per le nove popolazioni considerate (Tab.13) che a livello più specifico, facendo i test di differenziamento fra coppie di popolazioni (Tab.14). Per entrambi tipi di test sono stati usati 1000 *batches* e 10000 iterazioni per *batch*.

Il test di differenziamento complessivo tra popolazioni è risultato altamente significativo ($P < 0.000001$) permettendo di rigettare l'ipotesi di omogeneità.

Locus	P-Value
<i>Cr15</i>	0.00120
<i>Ca26</i>	<0.00001
<i>Ca40</i>	0.00029
<i>Ca48</i>	<0.00001
<i>Ca55</i>	0.00246
<i>Ca86</i>	0.00435
<i>Ca88</i>	<0.00001
<i>Cr127</i>	0.00434
<i>Cr259</i>	<0.00001

Tabella 13. Test di differenziamento genico per tutte le popolazioni. Vengono riportate le probabilità di differenziamento (*P-value*) per ogni *locus* e totali. I valori significativi ($P < 0.05$) sono evidenziati in grassetto.

Per verificare fra quali coppie di popolazioni esisteva un livello significativo di differenziamento, è stato eseguito il test di differenziamento per coppie di popolazioni. I *P-values* risultanti sono riportati nella tabella 14.

P-val.	CS08	DB05	SS02	EI02	EI06	EI07	SO06	SG04	BI04
CS08	-	0.00001	0.00095	0.12923	0.06555	0.05382	0.00021	<0.00001	<0.00001
DB05		-	<0.00001	<0.00001	0.00005	0.00001	<0.00001	0.00001	0.00118
SS02			-	0.41264	0.00045	0.22674	0.00031	<0.00001	0.00050
EI02				-	0.02484	0.12963	<0.00001	<0.00001	<0.00001
EI06					-	0.08710	0.00006	<0.00001	0.00002
EI07						-	0.05262	<0.00001	0.00003
SO06							-	0.00020	0.00031
SG04								-	0.00709
BI04									-

Tabella 14. Test di differenziamento genico per confronto a coppie tra popolazioni. Vengono riportate le probabilità di differenziamento per ogni coppia di popolazioni per tutti i *loci* stimate con il metodo esatto di Fisher. I valori significativi, ricavati in seguito alla correzione con il metodo di Bonferroni ($P < 0.05/36$, soglia = 0.001389), sono evidenziati in grassetto.

Indici di differenziamento F_{ST} e F'_{ST}

F_{ST} globale è stato misurato su 9 popolazioni usando 9 *loci*. F_{ST} globale, calcolato con le statistiche F di Weir e Cockerham, implementate in FSTAT ver. 2.9.3 2 (Goudet, 1995), è pari a 0.008 (SE = \pm 0.002, CI 0.005 ÷ 0.012). L'intervallo di confidenza risultante conferma la presenza di differenziamento tra le popolazioni. Si è quindi deciso di verificare fra quali coppie di popolazioni esistevano distanze genetiche significative. A tal fine si è effettuato un confronto a coppie tra le singole popolazioni.

I risultati, riportati nella tabella 15, mostrano un chiaro differenziamento delle popolazioni SG04 e SO06 dalle altre popolazioni. Anche la popolazione di BI04 risulta diversa da tutte le altre popolazioni. Questo risultato va però considerato con cautela a causa del numero troppo basso degli individui per campione. Le repliche temporali provenienti da *Elephant Island* risultano complessivamente omogenee, anche se si nota comunque un leggero differenziamento temporale fra i campioni. La popolazione DB04 risulta essere geneticamente differenziata da SS02.

F_{ST}	CS08	DB05	SS02	EI02	EI06	EI07	SO06	SG04	BI04
CS08	-	0.005	0.009	0.002	0.001	0.003	0.008	0.018	0.103
DB05	-0.000 – 0.015	-	0.014	0.005	0.001	0.004	0.011	0.010	0.071
SS02	0.002 – 0.020	0.005 – 0.033	-	0.003	0.007	0.000	0.010	0.026	0.073
EI02	-0.001 – 0.005	0.001 – 0.016	-0.001 – 0.008	-	0.001	0.000	0.007	0.021	0.083
EI06	-0.002 – 0.004	-0.003 – 0.008	0.001 – 0.014	-0.001 – 0.003	-	0.001	0.008	0.014	0.075
EI07	-0.001 – 0.008	-0.003 – 0.016	-0.003 – 0.003	-0.003 – 0.002	-0.001 – 0.003	-	0.002	0.017	0.072
SO06	0.002 – 0.016	0.004 – 0.022	0.002 – 0.020	0.002 – 0.012	0.004 – 0.014	-0.002 – 0.006	-	0.012	0.067
SG04	0.007 – 0.033	0.002 – 0.019	0.010 – 0.054	0.010 – 0.034	0.003 – 0.026	0.007 – 0.035	0.001 – 0.031	-	0.085
BI04	0.066 – 0.149	0.051 – 0.096	0.039 – 0.114	0.053 – 0.120	0.054 – 0.095	0.043 – 0.105	0.035 – 0.100	0.051 – 0.128	-

Tabella 15. Stima del differenziamento genico basata sulle frequenze alleliche. Sopra la diagonale della matrice vengono riportati i valori di $F_{ST}(\theta)$, ottenuti tramite il confronto a coppie tra le singole popolazioni. Nella parte inferiore sono riportati gli intervalli di confidenza al 95 % ($F_{ST} 95\% CI$) ottenuti con il metodo di *bootstrap*.

I tassi di differenziamento standardizzati F'_{ST} sono stati ottenuti mediante RECODEDATA ver. 0.1 (Meirmans, 2006) e FSTAT (Tab. 16)

F'_{ST}	CS08	DB05	SS02	EI02	EI06	EI07	SO06	SG04	BI04
CS08	-	0.020	0.032	0.007	0.004	0.011	0.029	0.070	0.340
DB05		-	0.056	0.020	0.004	0.017	0.045	0.045	0.276
SS02			-	0.011	0.026	0.000	0.036	0.100	0.234
EI02				-	0.004	0.000	0.026	0.083	0.280
EI06					-	0.004	0.030	0.058	0.263
EI07						-	0.007	0.067	0.241
SO06							-	0.047	0.222
SG04								-	0.309
BI04									-

Tabella 16. Divergenza standardizzata fra le popolazioni F'_{ST} (F_{ST} standardizzati) per TANDEM.

DISCUSSIONE

In questa tesi si riportano i risultati del più completo studio di genetica di popolazione di una specie *icefish*, *Chaenocephalus aceratus*. La difficoltà di campionare in un ambiente come l'Antartide evidenzia ulteriormente l'importanza di un campionamento così esteso, rappresentativo di tutto l'areale di distribuzione di *C. aceratus*. Questa tesi amplia due lavori di Papetti et al. (2007; 2009), e permette, attraverso l'applicazione di marcatori ad elevata risoluzione come i microsatelliti, di delineare un quadro di differenziamento completo per la specie *C. aceratus*, rilevando anche differenze molto piccole della struttura genetica.

In particolare, in questa tesi sono stati ottenuti *ex novo* dati genetici per i campioni di *Crystal Sound* (CS08), *Dallmann Bay* (DB05), *South Georgia* (SG04) e *Bouvet Island* (BI04). Questi dati sono stati confrontati con quelli ottenuti in precedenti lavori (Papetti et al., 2007; 2009), ricavando un'immagine esaustiva del livello di differenziamento di popolazioni campionate in tutto il *range* di distribuzione della specie. L'inclusione del campione proveniente da *Elephant Island* (EI07), ha inoltre permesso di estendere l'analisi di differenziamento su scala temporale, già considerata da Papetti et. al. (2009) per le repliche temporali del 2002 e del 2006.

I risultati ottenuti mostrano differenze significative tra le popolazioni di *South Georgia*, *Dallmann Bay*, *Bouvet Island* e le popolazioni provenienti dalla Penisola Antartica (PA). Questo dato è in accordo con i risultati ottenuti da studi di infestazione parassitica che mostrano la presenza di più popolazioni differenziate nell'areale di distribuzione della specie (Siegel, 1980; Sosinski e Janusz, 2000).

Il differenziamento significativo, trovato per la popolazione di *South Georgia*, è in accordo con i risultati dello studio di Ashford et al. (in stampa). Gli autori di questo lavoro hanno evidenziato che la struttura degli otoliti e la loro chimica negli individui di *C. aceratus*, campionati a *South Georgia*, presentavano una consistente omogeneità, giustificabile dalla presenza di una singola popolazione, ed avevano una composizione nettamente diversa da quella trovata in individui della PA. Nello stesso lavoro si è usato un modello oceanografico per simulare la dispersione passiva di larve rilasciate presso la PA verso gli altri siti. Si è rilevato che, in accordo con la chimica degli otoliti, le larve dalla PA vengono trasportati vicino al confine meridionale della CCA (*Bndry*), ma non arrivano però a *South Georgia*.

La popolazione di *Bouvet Island*, una piccola isola subantartica investita dalla CCA, è risultata, nonostante il *sample size* ridotto, geneticamente differenziata

dalle altre popolazioni campionate. I risultati ottenuti vanno considerati con cautela, a causa del numero ridotto di individui campionati ($n = 4$), sicuramente non rappresentativi dell'intera popolazione. Va tuttavia considerato che la piccola dimensione del campione riduce la potenza del test e quindi la possibilità di rigettare l'ipotesi zero di omogeneità; da questo punto di vista, l'aver rilevato differenze significative nonostante il *sample size* ridotto suggerisce che queste siano in realtà molto forti. Il differenziamento rilevato per *Bouvet Island* può essere facilmente spiegato dall'isolamento su scala geografica, dal momento che l'isola dista ben 2500 km da *South Georgia*. Il *pattern* osservato può però essere più complesso e in disaccordo con altri lavori.

Secondo Loeb et al. (1993) l'estensione delle fasi pelagiche di molte specie dei nototenioidi è indicativa del ruolo fondamentale della CCA nella dispersione larvale di queste specie. Jones et al. (2008), durante il *survey* sull'ittiofauna di *Bouvet Island*, hanno ritrovato solo nototenioidi e cannittidi caratterizzati da uno stadio larvale di durata superiore a 1-2 anni, senza incontrare alcuna specie endemica dell'isola. Gli autori hanno concluso quindi che l'ittiofauna di *Bouvet Island* deriva primariamente da *South Georgia*, attraverso la dispersione delle larve pelagiche con la CCA. Queste osservazioni, però, non concordano pienamente con i nostri risultati, in quanto la popolazione di *South Georgia* risulta geneticamente fra le più distanziate da quella di *Bouvet Island*, ad indicare che altri fattori (come la ritenzione larvale e il *self-recruitment*) possano essere responsabili del differenziamento osservato. Non è da escludere però, che il differenziamento riscontrato per la popolazione di *Bouvet Island* sia dovuto al campionamento.

La popolazione di *Dallmann Bay* risulta significativamente differenziata da tutte le popolazioni ad esclusione di *Crystal Sound* e due repliche di *Elephant Island*. Nell'ambito del quadro di differenziamento trovato per *C. aceratus*, la popolazione di *Dallmann Bay* sembra essere isolata dalla maggior parte delle altre aree di campionamento. Il *pattern* osservato in questo caso potrebbe essere causato da un isolamento su scala locale dovuto a ritenzione larvale, favorita dalla conformazione della baia. Il differenziamento tra il campione di *Dallmann Bay* con quello di *Crystal Sound* è tuttavia, non significativo, ad indicare un possibile scambio tra le due popolazioni. Questo scambio, infatti, può essere garantito da una corrente costiera, l'*Antarctic Peninsula Coastal Current (APCC)*, che fluisce lungo la costa occidentale della Penisola Antartica in direzione opposta del fronte più interno della CCA, il *Southern Antarctic Circumpolar Current Front (SACCF)*, da *Adelaide Island* ad *Alexander Island* (Moffat et al., 2007). La *APCC* è fortemente stagionale e si forma in primavera/estate, quando la costa interessata è libera dal ghiaccio.

La popolazione di *Crystal Sound* potrebbe costituire la popolazione *source* per il sistema *Elephant Island*, le cui repliche temporali non risultano significativamente differenziate dalla popolazione di *Crystal Sound*. Questa relazione *source-sink* era già stata predetta da Fack e Klinck (2006) in base alle simulazioni del *drift* delle particelle che mostravano come le popolazioni *source* per il krill, presente a *South Georgia*, potessero essere costituite dalle popolazioni distribuite lungo la piattaforma continentale della PA; le larve e i giovanili potevano infatti essere trasportati a *South Georgia* dopo essere stati incorporati nel SACCF. Per la popolazione di *Crystal Sound*, il flusso genico potrebbe essere infatti garantito dal trasporto passivo larvale unidirezionale attraverso il SACCF, che rappresenta la corrente dominante del sistema.

Le repliche temporali di *Elephant Island* non presentano differenze significative. Questi risultati appaiono in parziale contraddizione con i risultati ottenuti da Papetti et al. (2009) per le repliche temporali di *EI02* e *EI06*, fra le quali era stato trovato un differenziamento significativo ($F_{ST} = 0.009$, $p < 0.0001$). Queste differenze potrebbero essere dovute alla eliminazione nel presente lavoro del locus *Ca21* (che era risultato altamente polimorfico) e del locus *Cr171* (la cui genotipizzazione era fallita per molti campioni). Le differenze riscontrate in Papetti et al. (2009) erano infatti ridotte e possono essere state appiattite con l'esclusione dei due loci. La variabilità nel reclutamento e nel successo riproduttivo, utilizzati per spiegare le differenze temporali riscontrate in Papetti et al. (2009), possono quindi essere ancora una causa attendibile delle piccole, seppur non significative, differenze su scala temporale per le repliche di *Elephant Island*, riscontrate anche in questa tesi. L'ipotizzata advezione delle larve e degli individui giovanili provenienti da sud potrebbe, infatti, variare in seguito alle fluttuazioni oceanografiche che influiscono sulla sopravvivenza e sulla dispersione larvale (Mitarai et al., 2008).

Il differenziamento tra la popolazione di *Southern South Shetland - Elephant Island* (panmittica secondo Papetti et al., 2007) e la popolazione di *South Orkney*, riscontrato in Papetti et al. (2009), è stato confermato in questa tesi, anche se la replica temporale *EI07* non ha mostrato differenze significative dal campione di *South Orkney*. I valori di differenziamento ottenuti per queste popolazioni sembrano seguire le distanze oceanografiche, aumentando man mano che ci si sposta lungo la costa dalla popolazione di riferimento. L'integrazione di questi risultati con i dati oceanografici, che indicano la presenza della corrente prevalente SACCF lungo la piattaforma continentale della PA in direzione nord-est, potrebbe confermare la nostra ipotesi sulle popolazioni *source* lungo la

penisola che contribuiscono a breve raggio alla variabilità genetica delle popolazioni che si trovano più a monte.

Questi risultati di differenziamento lungo la PA richiamano il modello sulla dispersione larvale *stepping stone*, descritto brevemente nell'introduzione. Le distanze genetiche sembrano, infatti, correlare con le distanze geografiche, mentre l'assenza di differenziamento significativo fra popolazioni confinanti e la sua presenza, invece, fra popolazioni più distanziate, indica che la dispersione larvale avviene entro un raggio relativamente ridotto. Dall'altra parte, il modello potrebbe semplificare eccessivamente la situazione reale, dal momento che non è in grado di prendere in considerazione il flusso unidirezionale della corrente dominante, nonché l'esistenza di correnti locali e della conformazione specifica della costa, probabilmente in grado sia di favorire la ritenzione dello stadio pelagico (costituito da larve e individui giovanili), che di esporlo alle correnti.

Studi simili sulla genetica di popolazione dei nototenioidei mostrano risultati contrastanti; il differenziamento sembra infatti dipendere in queste specie non solo dalle caratteristiche oceanografiche e quelle di *life history* (come la durata della fase pelagica, notevole in molte specie dei nototenioidei), ma anche dalle caratteristiche ecologiche della specie, come migrazione attiva durante la fase pelagica e maggiore o minore sedentarietà durante la fase adulta. Il movimento attivo degli individui potrebbe, infatti, essere cruciale nel determinare il *pattern* di differenziamento genetico osservato. Ad esempio, un recente studio di Matschiner et al. (2009) rivela l'assenza di struttura genetica su un'ampia scala geografica in *Gobionotothen gibberifrons*, un nototenioide bentonico caratterizzato da distribuzione sovrapponibile a quella di *C. aceratus* e caratteristiche simili dello stadio larvale. L'autore giustifica l'assenza di differenziamento genetico mediante un trasporto unidirezionale (ovest-est) con la CCA. Matschiner conclude che la corrente risulta fondamentale nel prevenire il differenziamento genetico in *G. gibberifrons*, favorendo il flusso genico unidirezionale dalla PA verso le isole dello *Scotia Arc*. Tuttavia, alla luce delle simulazioni di flusso di particelle descritte da Ashford et al. (in stampa), la CCA manca *South Georgia*. Quindi un'ipotesi alternativa a quella suggerita da Matschiner per spiegare l'assenza di differenziamento genetico in *G. gibberifrons* prevedrebbe un movimento controcorrente attivo anche da parte di giovanili o adulti di questa specie. Questa ipotesi indica che gli aspetti ecologici delle specie come *C. aceratus* e *G. gibberifrons*, a parità di condizioni oceanografiche, possono svolgere un ruolo importante nel determinare la struttura genetica di popolazione.

Va infine considerato che il trasporto passivo larvale, che sembra rivestire il principale ruolo nel differenziamento genetico osservato per le popolazioni di *C. aceratus*, avviene probabilmente su scale più ridotte rispetto alle distanze

potenzialmente percorribili dalla specie, data l'estensione della sua fase larvale (400-500 giorni secondo La Mesa e Ashford, 2008) e l'esistenza di una potente corrente unidirezionale che gira attorno al continente Antartico, la CCA. Questo *pattern* potrebbe essere primariamente dovuto alla ritenzione larvale, la cui importanza è stata recentemente riconsiderata (Swearer et al., 2002). Molte larve sono, infatti, capaci della migrazione verticale attiva nella colonna d'acqua, caratteristica che, in combinazione con la stratificazione verticale dei flussi, può permettere loro di evitare il trasporto passivo. Va inoltre aggiunto che, soprattutto verso la fine della fase pelagica, gli individui giovanili sono in molti casi capaci di nuotare più velocemente delle correnti ambientali (Leis, 2006). È noto che le larve di molte specie dei nototenioidi sono in grado di effettuare la migrazione verticale (North e Murray, 1992). In questo modo, i meccanismi attivi di ritenzione sarebbero potenzialmente in grado di contrastare il trasporto con la CCA. Da questo punto di vista è importante notare come la persistenza delle popolazioni nella PA, per le quali proprio per l'unidirezionalità delle correnti non esiste la possibilità di dispersione passiva da altre aree geografiche, sia di per sé una chiara evidenza dell'importanza del *self-recruitment*, confermata anche dall'esistenza di popolazioni geneticamente distinte trovate in questa tesi.

Una completa comprensione della connettività e dei meccanismi che permettono la persistenza delle popolazioni marine nello *Scotia Sea* potrebbe quindi richiedere di tener conto, oltre che del potenziale legato alla dispersione passiva, anche del ruolo di fattori ecologici ed etologici nel determinare il reclutamento effettivo a livello locale. Tale comprensione è di fondamentale importanza per una corretta politica di gestione della pesca, tenuto conto della presenza di aree soggette a diversi livelli di protezione. Da questo punto di vista, i dati ottenuti in questa tesi suggeriscono che l'attuale regolamentazione non colga la complessità dei livelli di connessione tra popolazioni di *C. aceratus*, poiché impedisce *tout court* la pesca nell'area delle *South Shetlands*, *Elephant Island*, *South Orkneys* e *PA*, caratterizzata probabilmente da popolazioni parzialmente aperte e legate da relazioni *source-sink*, mentre consente la pesca commerciale, pur con attrezzi pelagici che limitano il *by-catch* di specie demersali, a *South Georgia*, dove la popolazione è, in base ai nostri dati, isolata e dipendente dal *self-recruitment*, quindi potenzialmente più vulnerabile.

RINGRAZIAMENTI

Un sentito grazie va innanzitutto al prof. Tomaso Patarnello, dott. Massimiliano Babbucci e dott.ssa Rafaella Franch per il servizio di genotyping. Ringrazio anche con tutto il cuore Stefania e Ilaria per avermi insegnato e aiutato nella mia esperienza di laboratorio. Un infinito grazie va sicuramente al dott. Lorenzo Zane e a Chiara per avermi insegnato, aiutato e sopportato, rendendo possibile lo svolgimento di questa tesi.

Ringrazio infinitamente mia mamma e mio papà, per aver sempre creduto in me e sostenuto nei momenti più difficili, mia nonna, che ha dedicato la propria vita ai suoi cari, e Carrie, per avermi fatto compagnia in tutti questi lunghi anni di studio. Non posso sicuramente scordarmi degli amici di università e dei mitici coinquilini, che mi hanno accompagnato in questi lunghi anni, vivendo con me momenti belli e un po' meno belli, e per avermi dato un'enorme mano in tutto.

Ringrazio infine tutte quelle persone che mi accompagnano nella vita di tutti i giorni e che mi vogliono bene; Ginino e Berta con i quali resterò eternamente in debito. Il ringraziamento più grande va però a Ivan per avermi aspettato, sopportato, capito e coccolato, anche quando non me lo meritavo.

BIBLIOGRAFIA

- Amos W, et al. (2007). Automated binning of microsatellite alleles: problems and solutions. *Mol. Ecol. Notes* 7:10–14.
- Ashford J, La Mesa M, Fach BA, and Jones C. (2009). Testing early life connectivity in Scotia Sea icefish using otolith chemistry and particle tracking simulations. *Canadian Journal of Fisheries and Aquatic Sciences*, in stampa.
- Bargelloni L, Ritchie PA, Patarnello T, Battaglia B, Lambert DM, Meyer A (1994). Molecular evolution at subzero temperatures mitochondrial and nuclear phylogenies of fishes from Antarctica (suborder Notothenioidei) and the evolution of antifreeze glycopeptides. *Molecular Biology Evolution*, 11:854-863.
- Bargelloni L, Lecointre G (1998). Four years in notothenioid systematics: a phylogenetic perspective. In: di Prisco G, Pisano E, Clarke A (eds) *Fishes of Antarctica*. Springer, Milano, pp 259–273
- 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.
- Barker, PF, Filippini GM, Florindo F, Martin EE & Scher HD (2007). Onset and role of the Antarctic Circumpolar Current. *Deep Sea Research II* 54:2388-2398.
- Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme F (2001). GENETIX, Logiciel sous Windows pour la Genetique des Populations. *Laboratoire Genome et Populations, Montpellier, France*.
- Björklund M (2005). A method for adjusting allele frequencies in the case of microsatellite allele drop-out. *Molecular Ecology Notes*, 5:676-679.
- Carvalho GR, Hauser L (1998). Advances in the molecular analysis of fish population structure. *Italian Journal of Zoology*, 65:21-33.
- Chen L, DeVries AL, & Cheng CHC (1997). Evolution of antifreeze glycoprotein gene from a trypsinogen gene in Antarctic notothenioid fish. *Proceedings of the National Academy of Sciences of the United States of America*, 94: 3811-3816.
- Chen WJ, Bonillo C, Lecointre G (1998). Phylogeny of the Channichthyidae (Notothenioidei, Teleostei) based on two mitochondrial genes. In: di Prisco G, Pisano E, Clarke A (eds) *Fishes of Antarctica. A biological overview*. Springer, Milano, pp 287–298.
- Clarke A (1990). Temperature and evolution: southern Ocean cooling and the Antarctic marine fauna. In: Kerry KR, Hempel G. (eds) *Antarctic*

- ecosystems. *Ecological change and conservation*. Springer, Berlin Heidelberg New York, pp 9–22
- 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(3), pp 578 - 580
- Detrich HW, Jones CD, Kim S, North AW, Thurber A, Vacchi M (2005). Nesting behavior of the icefish *Chaenocephalus aceratus* at Bouvetøya Island, Southern Ocean. *Polar Biol* 28:828–832.
- DeVries AL & Lin Y (1977). The role of glycoprotein antifreezes in the survival of Antarctic fishes. In Llano, G. A., ed. *Adaptations within Antarctic Ecosystems*, Washington, DC: Smithsonian Institution, pp 439-458.
- DeWoody J, Nason JD and Hipkins VD (2006). Mitigating scoring errors in microsatellite data from wild populations. *Molecular Ecology Notes*, 6:95-57
- Eastman JT (1993). *Antarctic fish biology: Evolution in a unique environment*. San Diego, Academic Press, 322 pp.
- Eastman JT (2000). Antarctic notothenioid fishes as subjects for research in evolutionary biology. *Antarctic Science* 12, 276-287.
- Eastman JT, Sidell BD (2002). Measurements of buoyancy for some Antarctic notothenioid fishes from the South Shetland Islands. *Polar Biol* 25:753–760
- Fach BA, Klinck JM (2006). Transport of Antarctic krill (*Euphausia superba*) across the Scotia Sea. I. Circulation and particle tracking simulations. *Deep-Sea Res I* 53:987-1010.
- Gagneux P, Boesch C, Woodruff DS (1997). Microsatellite scoring errors associated with non-invasive genotyping based on nuclear DNA amplified from shed hair. *Molecular Ecology*, 6, 861–868.
- Ghosh S, Karanjawala ZE, Hauser ER et al. (1997). Methods for precise sizing, automated binning of alleles, and reduction of error rates in large-scale genotyping using fluorescently labelled dinucleotide markers. *FUSION (Finland–US Investigation of NIDDM Genetics) Study Group. Genome Research*, 7, 165–178.
- Goldstein DB, Schlötterer C (1998). Microsatellite evolution and applications. Goldstein D.B., Schlötterer eds. *Oxford University Press*, 350pp.
- Goudet J (2002). FSTAT: a program to estimate and test gene diversities and fixation indices. Version 2.9.3.2.
- Guo S, Thompson E (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, 48:361-372.
- Hartl DL, Clark AG (1997). Principle of population genetics. *Sinauer Associates. Third edition*.
- Hauser L, Carvalho GR (2008). Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish and Fish*. 9:333-362.

- Hedrick PW (2005). A standardized genetic differentiation measure. *Evolution* 59:1633-1638.
- Hellberg ME (1994). Relationships between inferred levels of gene flow and geographic distance in a philopatric coral *Balanophylla elegans*. *Evolution* 48:1829-1854.
- Heywood KJ, Naveira Garabato AC and Stevens DP (2002). High mixing rates in the abyssal Southern Ocean. *Nature*, 415:1011-1014.
- Hofmann EE, Klinck JM (1998). Thermohaline variability of the waters overlying the west Antarctic Peninsula continental shelf. In: Jacobs SS, Weiss RF (Eds.), *Ocean, Ice, and Atmosphere: Interactions at the Antarctic Continental Margin. Antarctic Research Series*, vol 75, pp. 67–81.
- Iwami T (1985). Osteology and relationships of the family Channichthyidae. *Mem Natl Inst Polar Res Tokyo*, Ser E 36:1–69
- Iwami T, Kock KH (1990). Channichthyidae. In: Gon O, Heemstra PC (eds) *Fishes of the Southern Ocean. Smith Institute of Ichthyology, Grahamstown JLB*, pp 381–399.
- Iwami T, Cielniaszek J, Pakhomov EA (1996). Results on by-catch of fish during Ukrainian, Polish and Japanese krill fishery in the South Orkney Islands, South Georgia and Shetland Islands areas. WG-FSA-96/19, 20 pp, *CCAMLR, Hobart, Australia* (mimeogr)
- Jones CD, Kock KH, Balguerias E (2000). Changes in biomass of eight species of finfish around the South Orkney Islands (Subarea 48.2) from three bottom trawl surveys. *CCAMLR Sci* 7:53–74
- Jones CD, Anderson ME, Balushkin AV et al. (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.
- Kennett, JP (1977). Cenozoic evolution of Antarctic Glaciation, the Circum-Antarctic Ocean, and Their Impact on Global Paleooceanography. *Journal of Geophysical Research* 82:3843-3860.
- Kock KH (1991). The state of exploited fish stocks in the Southern Ocean - a review. *Arch FischWiss*, 41:1–66.
- Kock KH (1992). Antarctic fish and fisheries. Bundesforschung für Fischerei, Hamburg, Germany. *Cambridge University Press*, Cambridge, 359 pp.
- Kock KH, Döllefeld L, Hulley PA, Jährig H, Petzel W, Pusch C, Siegel V, White MG (1998). The composition of the fish fauna of the northern slope of King George Island. *Ber Polarforsch* 274:14–17
- Kock KH (2000). Understanding CCAMLR's Approach to Management. *Edited by Kock K.H.*
- Kock KH, Jones CD (2002). The biology of the icefish *Cryodraco antarcticus* Dollo, 1900 (Pisces, Channichthyidae) in the southern Scotia Arc

- (Antarctica). *Polar Biol* 25:416–424
- Kock KH, Pshenichnov L, Jones CD, Shust KV, Skora KE, Frolkina ZhA (2004). The area north of Joinville-D’Urville Islands (Subarea 48.1)—a former fishing ground at the tip of the Antarctic Peninsula—revisited. *CCAMLR Sci* 11:1–20
- Kock KH (2005). Antarctic icefishes (Channichthyidae): a unique family of fishes. A review, Part I. *Polar Biol*, 28(11):862-895.
- Kock KH (2005). Antarctic icefishes (Channichthyidae): a unique family of fishes. A review, Part II. *Polar Biol*, 28(12):897-909.
- Kock KH, Jones CD (2005). Fish stocks in the southern Scotia Arc region—a review and prospects for future research. *Rev Fish Res* (accepted)
- Lawver LA & Gahagan LM (2003). Evolution of Cenozoic seaways in the Circum-Antarctic region. *Palaeogeography, Palaeoclimatology, Palaeoecology* 198, 11-38.
- La Mesa M, Ashford J (2008). Age and early life history of juvenile Scotia Sea icefish, *Chaenocephalus aceratus*, from Elephant and the South Shetland islands. *Polar Biology*, 31(2):221-228
- Leberg PL (2002). Estimating allelic richness: Effects of sample size and bottlenecks. *Molecular Ecology*, 11:2445-2449.
- Leis JM (2006). Are larvae of demersal fishes plankton or nekton? Advances in *Marine Biology*, 51: 57–141.
- Liu K and Muse SV (2005). PowerMarker: an integrated analysis environment for genetic marker analysis
- Locarnini, RA, Whithworth T, and Nowlin WD (1993). The importance of the Scotia Sea on the outflow of Weddell Sea Deep Water. *J. Mar. Res.*, 51:135-153.
- Loeb VJ, Kellermann AK, Koubbi P, North AW, White MG (1993). Antarctic larval fish assemblages: a review. *Bull Mar Sci* 53:416–449
- Manly B. F. (1984). The statistics of natural selection. *Chapman & Hall, New York*.
- Matschiner M, Hanel R and Salzburger W (2009). Gene flow by larval dispersal in the Antarctic notothenioid fish *Gobionotothen gibberifrons*. *Mol. Ecol.* 18, 2574-2587.
- Matschiner M and Salzburger W (2009). TANDEM: integrating automated allele binning into genetics and genomics workflows. *Bioinformatics* 25(15):1982-1983.
- Maynard Smith J 1970). Population size, polymorphism and the rate of non-Darwinian evolution. *Amer. Nat.*, 104:231-236.
- Meirmans, PG 2006. Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution*, 60: 2399–2402.
- Meredith, MP, Naveira Garabato AC, Gordon AL & Johnson GC (2008).

- Evolution of the deep and bottom waters of the Scotia Sea, Southern Ocean, during 1995 - 2005. *J. Clim.* 21(7): 327-3343.
- Mitarai S, Siegel DA and Winters KB (2008). A numerical study of stochastic larval settlement in the California Current system. *Journal of Marine Systems* 69:295–309.
- Moffat C, Beardsley RC, Owens B, Nicole van Lipzig (2008). A first description of the Antarctic Peninsula Coastal Current. *Deep-Sea Research II* 55, 277-293.
- Naveira Garabato AC, McDonagh EL, Stevens DP, Heywood KJ & Sanders RJ (2002). On the export of Antarctic Bottom Water from the Weddell Sea. *Deep-Sea Res. II*.
- Near TJ, Pesavento JJ, Cheng CHC (2003). Mitochondrial DNA, morphology and the phylogenetic relationships of Antarctic icefish (Notothenioidei: Channichthyidae). *Mol Phyl Evol* 28:87–98
- Near TJ (2004). Estimating divergence times of notothenioid fishes using a fossil-calibrated molecular clock. *Antarct Sci* 16(1):37–44
- Near TJ, Chen CHC (2004). Phylogenetic investigations of Antarctic notothenioid fishes (Perciformes: Notothenioidei) using complete gene sequences of the mitochondrial encoded 16S rRNA. *Molecular Phylogenetic and Evolution*, 32:881-891.
- Nielsen EE, Nielsen PH, Meldrup D and Hansen MM (2004). Genetic population structure of turbot (*Scophthalmus maximus* L.) supports the presence of multiple hybrid zones for marine fishes in the transition zone between the Baltic Sea and the North Sea. *Molecular Ecology* 13, 585–595.
- North AW, Murray AWA (1992). Abundance and diurnal vertical distribution of fish larvae in early spring and summer in a fjord at South Georgia. *Antarctic Science*, 4, 405–412.
- Nowlin WD Jr. and Clifford M (1982). The kinematic and thermohaline zonation of the Antarctic Circumpolar Current at Drake Passage, *Journal of Marine Science* 40, pp. 481–507.
- Nowlin WD Jr. and Klinck JM (1986). The physics of the Antarctic Circumpolar Current, *Reviews of Geophysics* 24 (3), pp. 469–491.
- Olsen EM, Knutsen H, Gjøsæter J, Jorde PE, Knutsen JA and Stenseth NC (2008). Small-scale biocomplexity in coastal Atlantic cod supporting a Darwinian perspective on fisheries management. *Evolutionary Applications* 1, 524–533.
- Orsi AH, Whitworth T III and Nowlin WD Jr. (1995). On the meridional extent and fronts of the Antarctic Circumpolar Current, *Deep Sea Research I* 42, pp. 641–673.
- Orsi AH, Johnson GC & Bullister JL (1999). Circulation, mixing and production of Antarctic Bottom Water. *Prog. Oceanogr.* 43, 55-109.

- Palumbi SR (2003). Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Applications*: Vol. 13, Supplement: Marine Reserves, pp. 146-158.
- Palumbi SR (2004). Marine reserves and ocean neighborhoods: the spatial scale of marine populations and their management. *Annu Rev Environ Resour* 29:31–68
- Papetti C, Zane L, Patarnello T (2006). Isolation and characterization of microsatellite loci in the icefish *Chionodraco rastrispinosus* (Perciformes, Notothenioidea, Channichthyidae). *Molecular Ecology Notes*, 6: 207-209.
- Papetti C, Susana E, La Mesa M, Patarnello T, Zane L (2007). Microsatellite analysis reveals genetic differentiation between year-classes in the icefish *Chaenocephalus aceratus* at South Shetlands and Elephant Island. *Polar Biology*, 30(12): 1605-1613
- 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.
- Patwary MU, Kenchington EL, Birol CJ, Zauros E (1994). The use of random amplified polymorphic DNA markers in genetic studies of the scallop *Plactopecten magellanicus* (GMELLIN, 1791). *Journal of Shellfish Research*, 13, 547-553.
- Peakall R, Smouse PE (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6, 288-295.
- Rosenblum BB, Oaks F, Menchen S, Johnson B (1997). Improved single-stranded DNA sizing accuracy in capillary electrophoresis. *Nucleic Acids Research*, 25, 3925–3929.
- Raymond M, Rousset F (1995a). GENEPOP, version 1.2: population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86, 248–249.
- Raymond M and Rousset F (1995b). An exact test for population differentiation. *Evolution*, 49(6), 1280-1283.
- Scher HD & Martin EE (2006). Timing and climatic consequences of the opening of the Drake Passage. *Science* 312, 428-430.
- Schlotterer C (2004). The evolution of molecular markers-just a matter of fashion? *Nature Reviews Genetics*, 5, 63–69.
- Siegel V (1980). Quantitative investigations on parasites of Antarctic channichthyid and notothenioid fishes. *Meeresforsch* 28:146–156.
- Sosinski J, Paciorkowski A (1993). State of mackerel icefish (*Champsocephalus gunnari* Lönnberg, 1905) stock from South Georgia area based on Polish biological investigations in 1975–1992. *Pol Polar Res* 14:407–431
- Sosinski J., Janusz J. (2000). Infection variability of the parasitic copepod *Eubrachiella antarctica* (Quidor, 1906) on fishes in the Atlantic sector of

- the Antarctic. *Bulletin of the Sea Fisheries Institute*, 2(150):25-41.
- Susana E, Papetti C, Barbisan F, Bortolotto E, Buccoli S, Patarnello T and Zane L (2007). Isolation and characterization of eight microsatellite loci in the icefish *Chaenocephalus aceratus* (Perciformes, Notothenioidei, Channichthyidae), *Molecular Ecology Notes* 7, 791-793.
- Swearer SE, Shima JS, Hellberg ME et al. (2002). Evidence of self-recruitment in demersal marine populations. *Bulletin of Marine Science*, 70, 251–271.
- Takahasi M, Iwami T (1997). Summer diet of demersal fish at the South Shetland Islands. *Antarct Sci* 9:407–413
- Tautz D and C Schlötterer (1994). Simple sequences. *Curr. Opin. Genet. Dev.* 4:832–837.
- Tripathi, A., Backman, J., Elderfield, H., Ferretti, P. (2005). Eocene bipolar glaciation associated with global carbon cycle changes. *Nature* 436, 341-346.
- Van Oosterhout C, Hutchinson WF, Wills DP, Shipley P (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, 4(3): 535-538.
- Ward RD, Woodwark M, and Skibinski, DOF (1994). A comparison of genetic diversity levels in marine, fresh-water, and anadromous fishes. *Journal of Fish Biology* 44:213-224.
- Weir BS and Cockerham CC (1984). Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358-1370.
- Wenz HM, Robertson JM, Menchen S et al. (1998) High-precision genotyping by denaturing capillary electrophoresis. *Genome Research*, 8, 69–80.
- Wright, S 1978. Evolution and the genetics of population, variability within and among natural populations. *Univ. of Chicago Press, Chicago*.