



# UNIVERSITA' DEGLI STUDI DI PADOVA

Università degli Studi di Padova – Dipartimento di Biologia

DOTTORATO DI RICERCA IN BIOLOGIA EVOLUZIONISTICA  
CICLO XXI

## APPLICAZIONI DI MARCATORI MICROSATELLITE PER LO STUDIO DELLA FILOGEOGRAFIA DI ORGANISMI LAGUNARI DELL'ADRIATICO

**Coordinatore:** Ch.mo Prof. Giorgio Casadoro

**Supervisore:** Dott. Lorenzo Zane

**Dottorando:** Ilaria Anna Maria Marino

2 FEBBRAIO 2009



# INDICE

|  |           |
|--|-----------|
| <b>RIASSUNTO</b> .....   | <b>1</b>  |
| Abstract .....   | 4         |
| <b>1. INTRODUZIONE</b> .....   | <b>7</b>  |
| 1.1 Organismo di studio: <i>Carcinus aestuarii</i> .....                   | 12        |
| 1.2 Lagune e filogeografia .....   | 15        |
| 1.3 Mar Adriatico, cenni storici .....                                     | 17        |
| 1.4 Marcatori molecolari .....   | 19        |
| 1.4.1 DNA mitocondriale (mtDNA) .....                                      | 19        |
| 1.4.2 Marcatori microsatellite .....                                       | 20        |
| 1.5 Scopo della tesi .....   | 23        |
| <b>2. MATERIALI E METODI</b> .....   | <b>25</b> |
| 2.1 Estrazione del DNA genomico .....                                      | 27        |
| 2.2 Amplificazione di marcatori mitocondriali .....                        | 28        |
| 2.2.1 Condizioni di amplificazione .....                                   | 31        |
| 2.2.2 Purificazione e sequenziamento .....                                 | 31        |
| 2.3 Isolamento e messa a punto di marcatori microsatellite .....           | 32        |
| 2.3.1 Costruzione di una <i>library</i> genomica parziale arricchita ..... | 32        |
| 2.3.2 Purificazione e sequenziamento del DNA .....                         | 36        |
| 2.3.3 Disegno dei <i>primer</i> .....                                      | 36        |
| 2.3.4 Condizioni di amplificazione .....                                   | 37        |
| 2.4 Applicazione dei marcatori .....                                       | 39        |
| 2.4.1 Campionamento .....  | 39        |
| 2.4.2 Analisi dati .....   | 41        |
| 2.4.2.1 DNA mitocondriale .....  | 41        |
| 2.4.2.2 Marcatori microsatellite .....                                     | 47        |
| <b>3. RISULTATI</b> .....  | <b>51</b> |
| 3.1 Estrazione del DNA genomico .....                                      | 53        |
| 3.2 Amplificazione del DNA mitocondriale .....                             | 54        |
| 3.3 Isolamento e messa a punto di marcatori microsatellite .....           | 54        |
| 3.4 Analisi dati .....   | 55        |
| 3.4.1 Analisi del DNA mitocondriale .....                                  | 55        |
| 3.4.2 Analisi dei marcatori microsatellite .....                           | 64        |
| <b>4. DISCUSSIONE</b> .....  | <b>81</b> |
| 4.1 Analisi comparata .....  | 93        |
| 4.2 Conclusioni .....  | 96        |
| <b>5. BIBLIOGRAFIA</b> .....   | <b>97</b> |



## ***RIASSUNTO***



In questo lavoro di tesi sono stati applicati due tipi di marcatori molecolari, il DNA mitocondriale e i microsatelliti, per analizzare la struttura genetica di campioni di popolazione adriatici di *Carcinus aestuarii* (Decapoda: Portunidae, Nardo, 1847). Questo ha implicato l'amplificazione di un frammento di 482 paia di basi del gene mitocondriale codificante per la subunità I della citocromo c ossidasi (COI) al fine di indagare alcuni aspetti di filogeografia e di demografia storica della specie. Inoltre, è stato effettuato l'isolamento *ex novo* di 8 marcatori microsatellite specie-specifici per *C. aestuarii*, a cui sono stati affiancati 3 *loci* specifici per la specie atlantica *C. maenas*, per ricavare informazioni sulla genetica di popolazione del granchio verde.

*C. aestuarii* può essere considerato a tutti gli effetti un valido modello di studio delle lagune, essendo un tipico rappresentante della fauna di questi ambienti in tutto il Mediterraneo. Per l'elevato potere dispersivo larvale e per la facilità di campionamento, *C. aestuarii* può ricoprire un ruolo determinante nella comprensione delle possibili connessioni tra popolazioni di lagune differenti.

A livello di analisi di DNA mitocondriale, sono stati sequenziati complessivamente 255 individui (suddivisi per 8 campioni di popolazione provenienti da altrettante lagune adriatiche, ioniche e tirreniche). Sono state trovate 164 diverse varianti di sequenza (aplotipi). Il calcolo degli indici  $F_{st}$ , come pure l'analisi molecolare della varianza (*AMOVA*) hanno permesso di evidenziare un basso ma significativo livello di differenziamento genico tra i campioni di popolazione analizzati, confermando la presenza di una lieve strutturazione genetica e permettendo di rigettare l'ipotesi di panmixia. In particolare, è stato visto che una quota significativa della variabilità (3.90%) è dovuta alla suddivisione in gruppi, riconducibili rispettivamente al bacino tirrenico e a quelli adriatico-ionico.

Inoltre, è stato possibile evidenziare come tutte le popolazioni di *C. aestuarii* analizzate abbiano subito, in passato, fenomeni di espansione. Questo è stato possibile attraverso l'utilizzo dei test di neutralità-equilibrio, delle *mismatch distribution* e dei *bayesian skyline plot*, che hanno permesso di trarre indicazioni riguardo ai fenomeni demografici avvenuti in passato. In particolare, pare che per i campioni adriatico-ionici tali espansioni si siano realizzate in un intervallo di tempo antecedente le ultime glaciazioni pleistoceniche; mentre, per il campione tirrenico una variazione nelle dimensioni di popolazione sembra collocarsi in un periodo che coincide con gli ultimi cambiamenti climatici avvenuti in Mediterraneo.

Anche l'analisi attraverso i microsatelliti ha evidenziato, confermando i risultati mitocondriali, un debole ma significativo differenziamento tra i campioni di popolazione analizzati. L'uso dei marcatori microsatellite si è dimostrato di fondamentale importanza per rilevare, inoltre, piccole differenze presenti tra i campioni di popolazione dell'Adriatico e dello Ionio, dato non riscontrabile con il solo impiego del DNA mitocondriale.

Attraverso i microsatelliti, infine, è stato possibile verificare la presenza di isolamento per distanza e stimare i tassi di migrazione tra i campioni analizzati. Ne è emersa una situazione di non facile interpretazione: i flussi migratori nella maggior parte dei casi presentano direzione nord-sud con, tuttavia, due rilevanti eccezioni. Sia nel caso dei campioni di Venezia e Marano (alto Adriatico), che in quello dei campioni di Aquatina (bacino Adriatico meridionale) e Marano, la

migrazione si inverte, andando da sud a nord. Il motivo di un tale andamento potrebbe essere attribuito alle correnti oceanografiche presenti nel Mar Adriatico. Nella parte finale della tesi, vi è poi una sezione dedicata all'analisi di due specie lagunari (*Zosterisessor ophiocephalus* e *Atherina boyeri*) con lo scopo di condurre un'indagine comparata sulla genetica di popolazione di organismi che occupano abitualmente le lagune costiere del Mediterraneo.

## Abstract

Molecular markers, at the mitochondrial and nuclear level, were applied to the study of the population structure of shore crab *Carcinus aestuarii* (Decapoda: Portunidae, Nardo, 1847).

A 482-base-pair fragment of the mitochondrial cytochrome c oxidase I (COI) gene was analysed to examine the phylogeography and demographic history of *C. aestuarii*. Moreover, 8 microsatellites markers specific for shore crab were isolated *ex novo*; additional three microsatellites *loci*, specific for the sibling species *C. maenas*, already reported to cross-amplify in *C. aestuarii*, were also amplified to study population genetic of shore crab.

Due to its high dispersive planktonic larval stage and ease of sampling sample, *C. aestuarii* is a good model for studying lagoon ecosystems and understanding connectivity patterns of populations from different lagoons.

COI was analysed from 255 crabs collected in 8 different lagoons of the Mediterranean Sea. 164 sequence variants were found among the 255 individuals studied. Fixation indices ( $F_{st}$ ) and Analysis of Molecular Variance (*AMOVA*) showed a significant, though weak, genetic difference between samples, confirming the existence of a slight population structure and rejecting the panmixia hypothesis. *AMOVA* also showed that the 3.90% of the total genetic variability was explained by differences between groups of population that resemble the Tyrrhenian and Adriatic-Ionian Sea.

We also investigated the demographic history of *C. aestuarii* populations. We found that all the samples showed departures from neutrality that are consistent with massive population expansions. Neutrality tests, *mismatch distribution* and *bayesian skyline plot* confirmed the exponential growth in effective population size in all eight population samples. Estimated times for these expansions for Adriatic and Ionian population samples fall before the Last Glacial Maximum. Instead, population expansion for Tyrrhenian sample falls well within the last pleistocenic glaciation.

Microsatellites markers confirmed the results obtained with mitochondrial markers: a significant, though weak, genetic differentiation was found between samples. In particular, the application of microsatellite *loci* was very important for detecting a slight differentiation between population samples of Adriatic and Ionian Sea, that DNA mitochondrial marker did not find.

Microsatellites also revealed the presence of isolation by distance and were useful in estimating migration rates between samples. The data are not simply explainable: the migration flows have most often a north to south direction with two considerable exceptions. Both for Venezia-Marano and for Aquatina-Marano the migration has a different direction: from south to north. The cause of this pattern is probably due to the circulations of the Adriatic Sea.

In the last section of this study, two typical lagoon species (*Zosterisessor ophiocephalus* and *Atherina boyeri*) were used for a comparative analysis of the population genetic of Mediterranean lagoons organisms.



## ***1. INTRODUZIONE***



Negli ultimi decenni, gli studi di genetica molecolare degli organismi marini hanno modificato e ampliato sensibilmente la percezione della biodiversità nel mare, contribuendo in maniera significativa alla comprensione di quali siano i meccanismi evolutivi che agiscono sulle popolazioni naturali.

La diversità genetica, è una delle componenti fondamentali della biodiversità e può essere definita in modi diversi: a livello intraspecifico rappresenta l'intervallo delle differenze ereditabili tra gli individui di una stessa specie ed include le diversità esistenti tra individui all'interno di una popolazione e tra popolazioni diverse.

La diversità genetica di una specie e delle popolazioni ad essa appartenenti è condizionata dalla loro storia e influisce sulla capacità di adattamento delle specie e delle popolazioni future (Bagley *et al.*, 2002).

Può essere utilizzata come indicatore delle condizioni ambientali, infatti fattori di stress ambientale portano tipicamente ad una riduzione della diversità genetica, principalmente attraverso le forze della selezione e della deriva genetica. Quindi, una riduzione nella diversità genetica è indicativa del deterioramento dell'ambiente (Bagley *et al.*, 2002). Inoltre, la capacità di una specie di adattarsi ai cambiamenti ambientali è direttamente legata all'ammontare della variabilità genetica presente: più una specie presenta elevata variabilità genetica più è in grado di adattarsi facilmente alla variazione delle condizioni ambientali (Soule, 1987). Non sempre però questa regola è confermata: se individui appartenenti a popolazioni adattate alle condizioni locali si accoppiano con organismi adattati ad altre condizioni, allora l'aumento della variabilità genetica può portare ad una riduzione dell'adattamento con effetti avversi a livello di popolazione (depressione da esoincrocio). In linea generale, popolazioni piccole perdono la diversità genetica per effetto della deriva e vanno incontro ad una riduzione della fitness con un aumento del rischio di estinzione (Moritz, 1995).

Una bassa variabilità genetica all'interno di una popolazione, inoltre, è spesso indicativa di consanguineità (inbreeding), evento che porta effetti negativi a breve termine soprattutto sulla fitness degli individui.

La variabilità genetica rappresenta la base necessaria per i futuri cambiamenti evolutivi delle specie ed i livelli di tale diversità nelle popolazioni possono fornire le prove di eventi evolutivi avvenuti in tempi passati più o meno recenti (Weir, 1990).

Le modalità d'introduzione di nuova variabilità genetica in una popolazione sono ritenute classicamente:

- la mutazione;
- il flusso genico.

La mutazione è la fonte di tutta la diversità genetica. Essa è rappresentata da cambiamenti nel DNA all'interno di un dato *locus* e porta alla conversione della forma allelica di un gene in un'altra. Il tasso con cui avvengono le mutazioni è generalmente basso. La maggior parte di tali mutazioni o è neutrale o è dannosa rispetto alla fitness e verrà eliminata; tuttavia alcune risulteranno vantaggiose, quindi apporteranno vantaggi agli individui che le possiedono e si diffonderanno nella popolazione (Russel, 2003).

Il flusso genico, o migrazione, rappresenta il movimento degli individui riproduttivi tra le popolazioni. Generalmente è una forza che agisce molto pesantemente sulla diversità genetica aumentandola all'interno delle popolazioni rendendole così omogenee tra loro. Tale effetto di omogeneizzazione può

comportare una riduzione della fitness in popolazioni adattate localmente (Bagley *et al.*, 2002). Il flusso genico si realizza in proporzione alle possibilità d'incontro e quindi di scambio genico tra individui o gruppi di individui di una determinata specie. Le capacità di movimento caratteristiche dei vari stadi della vita di un organismo, gameti, embrioni, larve ed adulti, condizionano questo fattore microevolutivo.

Le popolazioni naturali di una determinata specie, che appartengono a regioni geograficamente distinte, presentano quasi sempre differenze nella distribuzione dei geni al loro interno: ciò indica la presenza di strutturazione genetica in quella specie. Tuttavia in ambiente marino differenze genetiche significative possono essere trovate su scale spaziali diverse che vanno dai pochi metri alle migliaia di chilometri (Procaccini and Maltagliati, 2003). Generalmente, specie con alte capacità di dispersione tendono ad essere meno geneticamente strutturate grazie all'azione omogeneizzante del flusso genico (Procaccini and Maltagliati, 2003).

Una tale situazione sembra si sia realizzata in molte specie di invertebrati e vertebrati marini, come evidenziato dalla scarsità di differenziazione di allozimi e del DNA mitocondriale all'interno di vaste aree geografiche (Avise, 2004). Questo è vero, per esempio, per il gambero tigre (*Penaeus monodon*) del sud-ovest dell'Oceano Indiano (Forbes *et al.*, 1999); per popolazioni di aragosta (*Jasus edwardsii*) in cui non è stata riscontrata alcuna suddivisione genetica per oltre 4600 km di habitat australasiano (Ovenden *et al.*, 1992); per i molluschi abissali (*Bathymodiolus thermophilus*) delle sorgenti idrotermali disseminate per tutto il Pacifico orientale (Craddock *et al.*, 1995); per popolazioni di numerose specie di ricci di mare appartenenti ai generi *Echinotrix* e *Strongylocentrotus* dell'Oceano Pacifico in cui si sono riscontrati elevati livelli di flusso genico (Palumbi and Wilson, 1990; Lessios *et al.*, 1998). Anche per quanto riguarda i vertebrati marini, e più in particolare i pesci, non è stata riscontrata alcuna struttura genetica: ad esempio, nel gadide *Theragra chalcogramma* campionato all'interno del Mare di Bering (Shields and Gust, 1995); nelle popolazioni di donzella (*Stegastes fasciolatus*) all'interno dei 2500 km dell'arcipelago hawaiano (Shaklee, 1984) e in ognuna delle numerose specie di pesci dei reef caraibici (Lacson, 1992; Shulman and Bermingham, 1995). Inoltre, anche molte specie di pesci pelagici tendono ad essere geneticamente più omogenee rispetto a specie anadrome (Ward *et al.*, 1994).

Per decenni, infatti, si è ritenuto che la struttura genetica degli organismi marini fosse piuttosto omogenea, questo per effetto dell'elevata capacità dispersiva degli stadi larvali, per gli eventuali spostamenti attivi degli adulti e per l'assenza, in ambiente marino, di ovvie barriere al flusso genico (Pampoulie *et al.*, 2008). Tuttavia, quest'idea ora è cambiata e complesse strutture genetiche sono state descritte per diverse specie marine. Infatti, molte specie marine, che presentano stadi larvali pelagici, hanno mostrato in alcuni casi un elevato differenziamento di popolazione (Burton, 1983, 1986; Hedgecock, 1986; Avise 1987, 2000; Palumbi, 1996). Nel *Limulus polyphemus* sono state riportate distinzioni nel mtDNA tra due popolazioni distribuite in maniera contigua nel Golfo del Messico e nella costa Atlantica del sud degli USA (Saunders *et al.*, 1986), nonostante la presenza in questa specie di larve trilobite specializzate per la dispersione (Rudloe, 1979). Ci sono anche fattori addizionali che possono spiegare perché un elevato potenziale di dispersione di gameti e larve non sempre è traducibile con una omogeneità spaziale nella struttura genetica di una popolazione (Hedgecock, 1986). Il flusso

genico può essere, infatti, condizionato da impedimenti fisici al movimento larvale e inoltre le larve non sempre si disperdono in maniera passiva ma possono adottare meccanismi di migrazione attiva e scegliere il luogo dove insediarsi. Fattori come la migrazione verticale, il nuoto attivo e una relativa tendenza a mantenere il contatto con il fondo possono far sì che le larve si allontanino meno del previsto dal loro luogo d'origine (Raimondi and Keough, 1990). Se le larve sono trattenute vicino alle loro popolazioni natali da meccanismi comportamentali o da meccanismi fisici e oceanografici, le popolazioni hanno l'opportunità di differenziarsi geneticamente e di adattarsi alle condizioni locali (Taylor and Hellberg, 2003). La ritenzione larvale locale potrebbe persistere per alcune generazioni e popolazioni marine che sono rimaste omogenee nonostante la presenza di barriere fisiche, potrebbero formare nuove specie o quantomeno differenziarsi (Taylor and Hellberg, 2003). Studi in cui sono stati utilizzati tags fluorescenti per marcare gli otoliti di individui giovanili di alcune specie di pesci di barriera (*Pomacentrus amboinensis* e *Thalassoma bifasciatum*) hanno mostrato che dal 15 al 60% dei giovani può reclutare nel sito di origine della popolazione natale, nonostante una durata dello stadio larvale da tre a sette settimane (Jones *et al.*, 1999; Swearer *et al.*, 1999; Taylor and Hellberg, 2003). Nel gamberetto *Alpheus immaculatus*, una specie in cui gli adulti vivono in simbiosi con anemoni ma producono larve liberamente natanti che possono rimanere pelagiche per lungo tempo, sono state trovate larve che si sono insediate in anemoni distanti pochi metri dai gamberetti parentali (Knowlton and Keller, 1986).

Le caratteristiche oceanografiche, inoltre, hanno promosso una differenziazione genetica tra popolazioni su piccola e larga scala geografica, come nel caso della stella di mare *Coscinasterias mircata* in cui si è riscontrato un ristretto flusso genico su scala macrogeografica (>1000 km) e una significativa struttura di popolazione su scala mesogeografica (10-200 km, Pérrin *et al.*, 2004) o come nel *Dissostichus eleginoides* della Patagonia (Shaw *et al.*, 2004). La dispersione larvale può essere limitata da barriere fisiche come sistemi frontali (ad esempio, il Fronte Polare Antartico, Shaw *et al.*, 2004), correnti oceaniche (Ruzzante *et al.*, 1998) e circolazioni estuarine (Pérrin *et al.*, 2004). Inoltre, una differenziazione genetica tra popolazioni potrebbe essere legata ad un limitato flusso genico degli adulti in seguito a isolamento per distanza, soprattutto su vasta scala geografica. Per altro, ci sono evidenze sempre più forti che eventi storici, come l'isolamento di popolazioni in rifugi durante il periodo glaciale, possano aver ricoperto un ruolo importante nell'origine di una struttura di popolazione negli organismi marini (Gysels *et al.*, 2004; Hoarau *et al.*, 2007).

Quindi, il grado di differenziamento intraspecifico non è a priori prevedibile, poiché è anche determinato dalla disponibilità di habitat che a sua volta è influenzata da fattori estrinseci di tipo geologico, fisico e climatico (modifica dell'habitat per movimenti tettonici, avanzamento o scioglimento dei ghiacci, variazioni del livello del mare), perdita di habitat ad opera dell'uomo e introduzione di specie alloctone.

La struttura genetica delle specie e, quindi, i meccanismi evolutivi e le strategie adattative da esse adottati possono essere conosciuti analizzando la distribuzione dei geni all'interno delle popolazioni. E' in questo contesto che si inserisce lo studio effettuato in questa tesi: è un'indagine sulla genetica di popolazione e sulla filogeografia di specie lagunari, con particolare riferimento al granchio verde

*Carcinus aestuarii* (Nardo, 1847) e alle relative popolazioni presenti nel Mare Adriatico.

Questa specie può essere considerata un valido modello di studio, essendo un tipico rappresentante della fauna lagunare del Mediterraneo: infatti, per l'elevato potere dispersivo larvale e per la facilità di campionamento, *C. aestuarii* potrebbe ricoprire un ruolo determinante nella comprensione delle possibili connessioni tra popolazioni di differenti lagune.

## 1.1 Organismo di studio: *Carcinus aestuarii*

*C. aestuarii* (figura 1.1), comunemente detto granchio verde, è un Decapode appartenente alla famiglia dei Portunidi, nativo del Mediterraneo ma che, circa due secoli fa, è stato accidentalmente introdotto in regioni diverse rispetto al suo luogo d'origine, quali Giappone e Sud Africa. Vive abitualmente nelle acque estuarine e nelle lagune privilegiando le aree dei fondali ricoperte da *Zostera* (Strabella, 1996).



Figura 1.1: esemplare adulto di *Carcinus aestuarii*.

È un organismo molto resistente e, rispetto alle altre specie di decapodi, possiede un'elevata capacità di adattamento all'ambiente: è più resistente all'inquinamento, tollera bene l'esposizione all'aria, periodi di digiuno e ampie escursioni di temperatura e salinità (euritermo e eurihalino). È una specie onnivora, che si nutre di una grande varietà di organismi diversi quali: vegetazione paludosa, alghe, crostacei, molluschi e pesci (Carlton and Choen, 2003).

Ha una vita media di circa 3-4 anni ed è tipicamente stanziale anche se è in grado di percorrere diversi chilometri al giorno se le condizioni ambientali ottimali vengono a mancare (Strabella, 1996).

Tutte queste caratteristiche lo rendono un organismo invasore: negli ultimi due secoli, infatti, la specie è stata accidentalmente introdotta in diverse regioni esterne al suo range nativo, a causa del commercio marittimo e delle navi da trasporto (Carlton and Choen, 2003). Può portare ad effetti nocivi a livello delle comunità marine in quanto altera le catene alimentari, distrugge gli habitat, induce le popolazioni autoctone allo spostamento e preda specie di valore commerciale quali vongole, cozze, ostriche e giovani granchi autoctoni (Behrens Yamada and Hauck, 2001).

*Carcinus aestuarii* presenta molte somiglianze con la specie atlantica *Carcinus maenas* con il quale in passato è stato spesso confuso fino a quando gli studi svolti da Nardo nel 1847 ne hanno accertato l'identità. *C. maenas* è una specie nativa dell'Oceano Atlantico e del nord-ovest dell'Africa e anch'essa è stata introdotta in ambienti diversi da quelli di origine come le coste est ed ovest del Nord America, Australia e Tasmania (Behrens Yamada and Hauck, 2001). Piccole differenze distinguono *C. maenas* da *C. aestuarii*: il rapporto altezza-larghezza del carapace è diverso; nella serie di dentelli antero-laterali del carapace, l'ultimo dentello è rivolto verso l'interno nella specie atlantica, mentre è puntato verso l'esterno in quella mediterranea; i pleopodi maschili sono paralleli e rettilinei in *C. aestuarii*, ricurvi in *C. maenas* (Behrens Yamada and Hauck, 2001). Per ciò che riguarda le popolazioni mediterranee e atlantiche, lo Stretto di Gibilterra, con la sua alta soglia, agisce come una parziale barriera per lo scambio larvale e genico (Behrens Yamada and Hauck, 2001).

A livello morfologico *Carcinus aestuarii* si presenta generalmente piccolo, con un carapace liscio a ventaglio stretto e spesso; ha simmetria bilaterale, la forma esagonale del corpo tende ad essere appiattita dorso-ventralmente. Il margine postero-laterale è pressoché dritto e più lungo di quello antero-laterale; in questa posizione, inoltre, presenta 5 dentelli sporgenti rivolti in avanti. Nella parte frontale protrudono tre lobi e si notano antenne ricoperte da peli; sempre nella parte frontale del carapace, collocati nelle rispettive cavità, si trovano i peduncoli oculari che sostengono l'occhio composto dell'invertebrato (Behrens Yamada e Hauck, 2001).

A livello ventrale, presenta 5 paia di arti ambulacrali e di questi, il quinto paio più lungo, viene utilizzato per il nuoto (Behrens Yamada and Hauck, 2001). La coppia di arti anteriore porta all'estremità due robuste chele utili a scopo offensivo, difensivo e di corteggiamento ma anche per la comunicazione visiva e sonora. Inoltre presenta particolari appendici trasformate, i pleopodi, collocate sotto l'addome, che nelle femmine servono per trattenere le uova e nei maschi come organi copulatori (Strabella, 1996).

La colorazione è verde olivastro superiormente e giallo crema inferiormente.

I sessi sono separati e facilmente determinabili: il maschio, il cui carapace può arrivare a 65 mm di larghezza, è più grande e presenta un addome triangolare mentre la femmina, più piccola (con un carapace che arriva a 45 mm), ha un addome convesso (Strabella, 1996, figura 1.2).

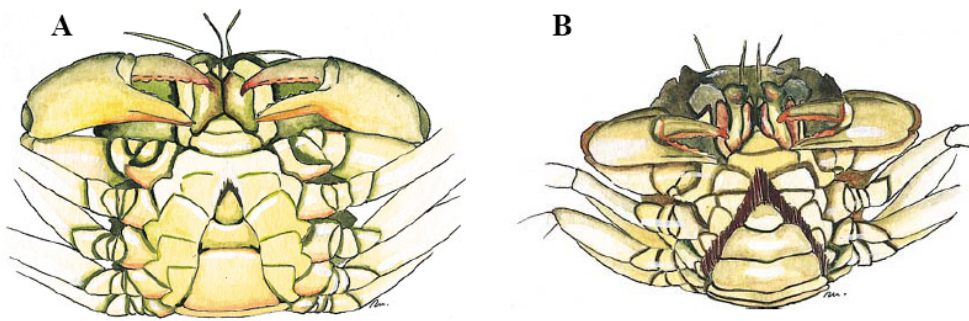


Figura 1.2: differente forma dell'addome nel maschio (A) e nella femmina (B) di *Carcinus aestuarii*.

Anche il periodo di muta varia tra i sessi: nel maschio avviene durante il periodo primaverile o autunnale, con due picchi tra marzo-aprile e ottobre-novembre, mentre nelle femmine durante il periodo estivo.

Il periodo della riproduzione comincia a primavera e si protrae sino all'autunno inoltrato, con un picco tra maggio e giugno, anche se la maturazione delle gonadi e la preparazione delle uova alla schiusa può inoltrarsi fino all'inizio della primavera successiva. Quindi la riproduzione avviene tra maschi in intermuta e femmine in prossima muta (Behrens Yamada e Hauck, 2001). La fecondazione è interna e avviene per trasferimento di spermatofores, tramite i pleopodi, direttamente nel gonodotto femminile. Le uova sono trattenute nell'addome della femmina fino alla liberazione delle larve che avviene in mare; sono pelagiche e soggette ad una dispersione passiva. Dall'uovo si schiude una prima larva (zoea) che, dopo l'ultimo stadio larvale (megalopa), muta per la prima volta in giovane granchio; da questo momento si susseguono più mute nell'arco di un anno, fino al raggiungimento dello stadio di granchio adulto (figura 1.3).

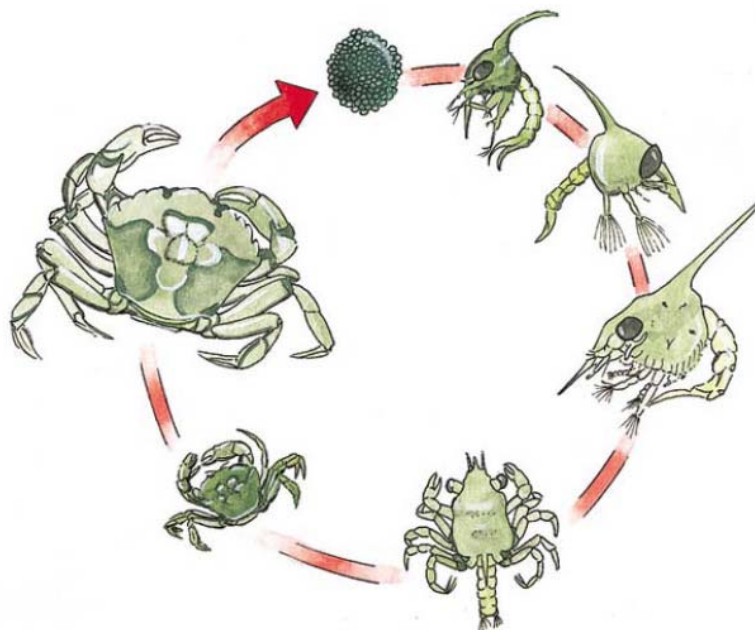


Figura 1.3: ciclo vitale di *Caricinus aestuarii*.

È una specie commestibile, viene pescato e commercializzato soprattutto in prossimità del periodo di muta e può anche essere allevato. La molechicoltura, la pesca cioè del granchio per la produzione delle “moleche o moeche” (il granchio privo di carapace) è un’attività strettamente locale ed esclusiva delle lagune venete, nelle quali ha una tradizione secolare e ancora oggi una certa rilevanza economica (Strabella, 1996).

Lo studio di questa specie, per le caratteristiche sopra descritte, potrebbe fornire delle preziose indicazioni per comprendere le possibili relazioni che sussistono tra popolazioni di diverse lagune. In altre parole, si potrebbe capire se tali popolazioni, e quindi i loro ambienti, siano interconnesse da fenomeni migratori asimmetrici, in cui prevalgano dinamiche di popolazione del tipo sorgente-pozzo o se, invece, predomini una situazione di autosostentamento, in cui ogni laguna con la propria popolazione costituisca un ambiente a sé, indipendente dalle altre.

## 1.2 Lagune e filogeografia

Le lagune costiere sono da sempre oggetto di studio perché caratterizzate da un’elevata biodiversità, tant’è che sono attualmente annoverate tra gli habitat di interesse comunitario per cui sono previste speciali misure di conservazione. Lo studio della struttura di popolazione delle specie caratteristiche di questi ambienti, oltre ad essere un indispensabile strumento per stabilire misure di gestione ai fini conservazionistici degli ambienti stessi, è altresì determinante per comprendere i meccanismi evolutivi alla base della differenziazione tra popolazioni e, di conseguenza, dei fenomeni microevolutivi (Maltagliati *et al.*, 2003). Le lagune costituiscono il 13% del litorale costiero mondiale e si estendono dalle zone tropicali sino ai poli; ricevono continui apporti sia dagli oceani che dai continenti e sono tra gli ecosistemi più produttivi al mondo. Infatti, in quanto ambienti di passaggio tra l’ecosistema marino e terrestre, presentano molti degli organismi e delle caratteristiche delle comunità di sovrapposizione oltre a caratteristiche e organismi propri e spesso, sia il numero di specie che la densità di popolazione sono maggiori rispetto alle comunità che le circondano (Critto and Marcomini, 2001). Inoltre, è importante sottolineare come gli ecotoni siano fortemente influenzati dagli ecosistemi che li delimitano e, almeno in parte, li influenzano a loro volta.

In diverse parti del mondo, inoltre, le lagune sono state fortemente influenzate dall’attività antropica che ne ha alterato le caratteristiche morfologiche e idrologiche, come nel caso della Laguna di Venezia, dove l’attività umana e i processi ecologici hanno raggiunto un livello di interazione profondo, completo e allo stesso tempo complesso (Lasserre and Marzollo, 2000).

Sulle specie animali e vegetali presenti in questi ambienti di transizione sono stati condotti diversi studi di genetica di popolazione e di filogeografia (e.g. Dawson *et al.*, 2002; Thompson *et al.*, 2005). Queste due discipline si occupano di processi inter-individuali e di popolazione che sono in funzione di meccanismi microevolutivi (flusso genico, mutazione, selezione naturale), che determinano la distribuzione spazio-temporale della diversità genetica di una specie. In particolare, la filogeografia studia i processi e i principi che regolano la distribuzione geografica di linee genealogiche all’interno delle specie e tra specie strettamente correlate (Avise, 2000).

Questa materia di studio nasce alla fine degli anni '70 in seguito all'introduzione delle analisi del DNA mitocondriale (mtDNA) nella genetica delle popolazioni animali e questo ha portato ad un cambiamento rivoluzionario nello studio della struttura di popolazione intraspecifica, introducendo una prospettiva storico-genealogica (Avise, 2004). Le sequenze di mtDNA presentano delle caratteristiche peculiari che le rendono particolarmente utili nella ricostruzione di alberi filogenetici: queste, infatti, evolvono generalmente con velocità piuttosto elevate, mostrano prevalentemente ereditarietà materna e non ricombinano. In questo modo, quindi, forniscono degli aplotipi la cui similarità genetica all'interno di una specie può essere utilizzata per costruire dei dendrogrammi che riflettono la componente matriarcale della genealogia di un organismo (Avise, 2004).

Sin dalle prime analisi delle popolazioni naturali è risultato chiaro che, in alcuni casi, gruppi di aplotipi identici o separati da poche mutazioni erano localizzati geograficamente, mentre aplotipi molto diversi venivano trovati in aree distinte (Avise *et al.*, 1979). Queste osservazioni hanno portato a coniare il termine "filogeografia" (Avise *et al.*, 1987). Questa disciplina focalizza esplicitamente l'attenzione sulle componenti storica e filogenetica della struttura di popolazione (includendo come queste possano essere state influenzate dalla deriva genetica, dal flusso genico, dalla selezione naturale o da altre forze evolutive). In termini più ampi, il contributo più importante della filogeografia è stato quello di enfatizzare gli aspetti di non-equilibrio della struttura di popolazione e della microevoluzione, chiarire le strette connessioni che esistono tra demografia di popolazione e genealogia e creare dei nessi tra campi di ricerca, precedentemente separati, della genetica di popolazione tradizionale e della biologia evolutiva. L'analisi, infatti, e l'interpretazione delle distribuzioni di linee genealogiche richiedono il contributo integrato di informazioni di genetica di popolazione e filogenesi, di demografia, di etologia e di paleoclimatologia.

Se da una parte il DNA mitocondriale si è rivelato estremamente utile per lo sviluppo di tali studi, dall'altra il suo potere informativo è circoscritto alla sola componente femminile-materna. Infatti, attraverso l'analisi dell'mtDNA è possibile risalire alla distribuzione filogeografica di una specie, ma ciò che si ricava è solo un'informazione parziale dovuta al fatto che si sta considerando un carattere che viene trasmesso, generalmente, dalla sola linea materna. Questo inconveniente è stato in parte risolto attraverso l'impiego dei marcatori nucleari, come ad esempio i microsattelliti (e.g. Gum *et al.*, 2005, Johnsen *et al.*, 2006), di più recente introduzione, che forniscono una visione completa di come sia distribuita la variabilità genetica in una specie, prendendo in considerazione anche quei caratteri ereditati dalla linea maschile e permettendo quindi di ricostruire i processi che avvengono a carico dell'intero genoma.

Nonostante questa disciplina sia ampiamente diffusa e applicata ad un vasto numero di specie animali e vegetali, lo studio della filogeografia relativo ad organismi di ambienti di transizione (ecotoni), come sono le lagune, è ancora ad uno stadio embrionale, specie per quanto riguarda il Mare Adriatico. Infatti, mentre in letteratura è presente una gran quantità di lavori che trattano di filogeografia applicata ad organismi di sistemi ecotonali, come ad esempio Gobidi (*Clevelandia ios* e *Eucyclogobius newberryi*, Dawson *et al.*, 2002) o Alfeidi (*Alpheus djeddensis*, Thompson *et al.*, 2005), per il Mare Adriatico esiste un numero limitato di lavori effettuati su poche specie tipicamente lagunari, come ad esempio i latterini (*Atherina boyeri*, Astolfi *et al.*, 2005).

### 1.3 Mar Adriatico, cenni storici

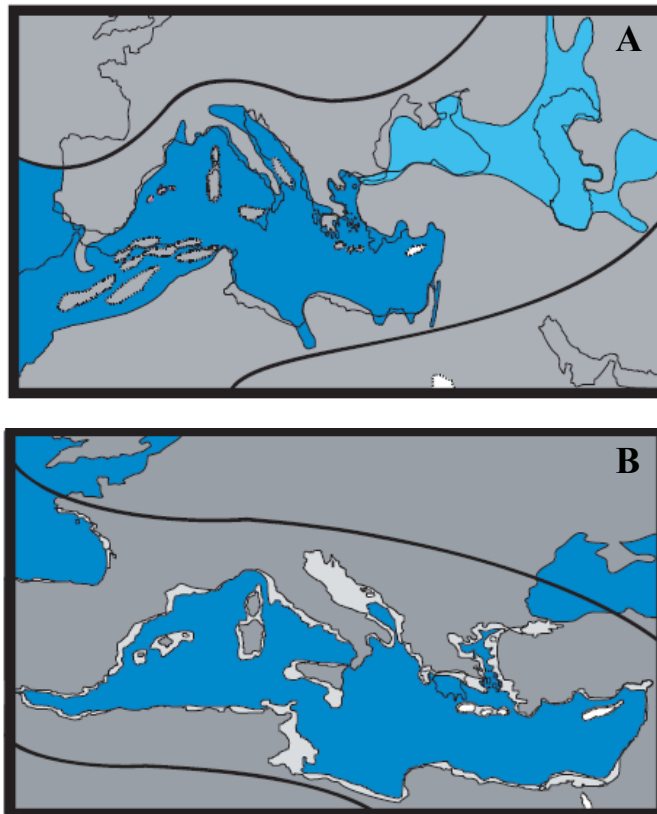
Assieme al Mar Nero, il Mar Adriatico è il bacino più continentale del Mediterraneo; è chiuso tra due catene montuose, gli Appennini a ovest e i Balcani a est (Franco *et al.*, 1982). Occupa una depressione della superficie terrestre che si sarebbe originata per sprofondamento di un'antica area continentale in seguito all'azione di compressioni convergenti esercitate dalle masse continentali presenti nel Mediterraneo. Lo sprofondamento dell'area sarebbe iniziato 200 milioni di anni fa e da allora grandi quantità di sedimenti sono andati ad accumularsi sui fondali del bacino. La formazione del Mediterraneo, nella sua conformazione attuale, inizia a determinarsi circa 10 milioni di anni fa, durante il Miocene con la formazione dell'istmo di Suez e quindi la separazione dall'Indopacifico e si conclude 6 milioni di anni fa (tardo Miocene: Messiniano) quando la connessione con l'Atlantico si chiude anch'essa progressivamente e il Mediterraneo diventa un mare chiuso. E' proprio con l'isolamento del Mediterraneo che si produce la cosiddetta crisi messiniana, durante la quale prevalsero fenomeni di evaporazione e il disseccamento di vaste aree di fondale (figura 1.4): nel corso di tale crisi si registrò l'estinzione, dovuta al forte innalzamento della salinità, di buona parte delle specie presenti nel Mediterraneo. Tuttavia, tale estinzione non fu totale, e probabilmente le specie di origine indopacifica, presenti attualmente nel Mediterraneo, sono sopravvissute in aree rifugio (Hrbek and Meyer, 2003) dove la salinità non ha costituito un fattore limitante.



**Figura 1.4: crisi del messiniano: conformazione del bacino mediterraneo 5.6-5 milioni di anni fa. Le zone grigie rappresentano le terre emerse; quelle azzurre contraddistinguono acqua con bassi valori di salinità mentre quelle blu acqua ipersalina (da Rogl and Steininger 1983; Dawson 1992; Rogl 1999; Patarnello *et al.*, 2007).**

Con la riapertura del collegamento con l'Atlantico, avvenuta nel Pliocene (5 milioni di anni fa), il Mediterraneo subisce una nuova invasione marina che porta ad un innalzamento del livello del mare e alla formazione di un primitivo Mar Adriatico.

Questo, essendo un bacino poco profondo e semi chiuso, presenta una morfologia e un'area che hanno subito variazioni di estensione e di livello in periodi geologici anche recenti (Asioli *et al.*, 2001): mentre nel Pliocene (3.5-3 milioni di anni fa) si estendeva al posto dell'attuale pianura padana, durante alcune fasi delle espansioni glaciali pleistoceniche (18000-13500 anni fa) si è ridimensionato entro limiti assai più ridotti degli attuali, in seguito all'abbassamento del livello dei mari (calcolato in circa 120 metri rispetto all'attuale, Zonneveld, 1996). Di conseguenza, l'Adriatico settentrionale risultava completamente emerso e trasformato in una vasta pianura alluvionale (figura 1.5).



**Figura 1.5:** variazioni di estensione dell'Adriatico nel Terziario (3.5-3 milioni di anni fa, figura A) e nel Quaternario (18000-13500 anni fa, figura B, da Rogl and Steininger 1983; Dawson 1992; Rogl 1999; Patarnello *et al.*, 2007).

Successivamente, con il ritiro dei ghiacci alla fine dell'ultima glaciazione, il livello del mare si è innalzato progressivamente sino alla posizione attuale, permettendo all'acqua di avanzare nuovamente verso nord e ricolmare quella pianura che si era venuta a formare. In questo modo, l'Adriatico ha acquisito via via la morfologia che lo contraddistingue attualmente e solo verso la fine di questo processo hanno cominciato a prendere forma tutti quegli ambienti di transizione, quali, ad esempio, le lagune, che lo caratterizzano.

Tutto questo permette di comprendere come le lagune adriatiche siano, quindi, degli ecosistemi di recentissima formazione. In particolare, datazioni al carbonio 14 di sedimenti lagunari testimoniano la formazione della Laguna di Venezia tra circa 10000 anni fa, per la laguna meridionale, e 5000 anni fa, per la laguna nord (Bondesan and Meneghel, 2003).

Lo studio filogeografico, quindi, di specie strettamente associate a ecosistemi di transizione potrebbe fornire preziose informazioni sulle possibili relazioni tra popolazioni di differenti lagune, sul ruolo che la storia recente dell'ambiente ricopre nel determinare la distribuzione della variabilità genetica (Astolfi *et al.*, 2005) e sui pattern di ricolonizzazione postglaciale; inoltre si potrebbero trarre delle possibili indicazioni per la conservazione della biodiversità.

A tal fine, sono state utilizzate due classi di marcatori molecolari: il DNA mitocondriale e i microsattelliti. Le analisi filogeografiche e filogenetiche sono state condotte su una regione del mtDNA, la citocromo c ossidasi I, altamente polimorfica nei crostacei. Per l'analisi della variabilità inter/intrapopolazionale e di genetica di popolazione, invece, sono stati utilizzati i *loci* microsattelliti, marcatori codominanti in grado di rivelare elevati livelli di polimorfismo, particolarmente adatti per valutare strutturazione e flusso genico.

## 1.4 Marcatori molecolari

In questo lavoro, è stato amplificato e sequenziato un frammento di 482 paia di basi del gene mitocondriale codificante per la subunità I della citocromo c ossidasi di *C. aestuarii*. Inoltre sono stati isolati *ex novo* 8 *loci* microsattelliti specie-specifici per *Carcinus aestuarii*, a cui sono stati affiancati 3 *loci* specifici per la specie atlantica *C. maenas*.

### 1.4.1 DNA mitocondriale (mtDNA)

Il DNA mitocondriale è da sempre considerato un utile marcatore per studi evolutivi finalizzati ad analizzare i rapporti che intercorrono tra i vari taxa a differenti livelli filogenetici, nonché i rapporti che intercorrono tra le popolazioni di una singola specie.

Il genoma mitocondriale (mtDNA) è costituito da piccole molecole circolari aploidi di DNA (16-20.000 bp nei vertebrati) contenute all'interno dei mitocondri, organelli citoplasmatici con un ruolo fondamentale nella sintesi dell'ATP e, quindi, nella produzione dell'energia necessaria all'attività cellulare. È estremamente compatto ed è costituito principalmente da regioni codificanti prive di introni e non separate tra loro da sequenze interposte, anche se è presente una regione non codificante (D-loop) che svolge un ruolo importante nel controllo della replicazione e della trascrizione del DNA mitocondriale (Machida *et al.*, 2004). In ogni cellula sono contenute numerose copie di mtDNA e generalmente queste copie sono identiche (omoplasmiche). A volte però ci possono essere due o tre genomi mitocondriali diversi (eteroplasmici) in uno stesso individuo; l'eteroplasmia può derivare da nuove mutazioni insorte a livello del mtDNA, dal "pool" di mitocondri presenti nella cellula uovo da cui deriva l'individuo (Moritz *et al.*, 1987) o, in alcune specie, può essere la conseguenza di una trasmissione biparentale del mtDNA (Zouros *et al.*, 1992).

Il mtDNA codifica per 37 geni: 13 proteine (COI, COII, COIII, ND1, ND2, ND3, ND4, ND4L, ND5, ND6, Cyt b, ATPasi3, ATPasi6), due RNA ribosomiali (12S rRNA e 16S rRNA) e 22 RNA di trasferimento (tRNA).

Molti studi dimostrano che il mtDNA è più soggetto a mutazione di quello nucleare a causa dell'assenza di meccanismi di riparazione e di ricombinazione, dell'elevato numero di replicazioni, della mancanza di istoni con funzione protettiva (presenti, invece, nel DNA nucleare), e della stretta associazione fisica con la membrana mitocondriale interna. Inoltre, è più esposto ad agenti mutageni: il mitocondrio è la sede della fosforilazione ossidativa ed è quindi a contatto con i radicali liberi dell'ossigeno prodotti durante alcune fasi della reazione (Richter *et al.*, 1988; Avise, 2000).

Diverse caratteristiche hanno reso il mtDNA uno dei marcatori maggiormente utilizzati per studi di filogenesi e di genetica di popolazione, sin dagli anni '70. Per il fatto di essere presente in elevate quantità per ogni cellula (ogni cellula contiene numerosi mitocondri e ogni mitocondrio contiene numerose copie di mtDNA) e di mostrare tassi di mutazione differenti a seconda delle regioni (regioni codificanti sono più conservate e presentano un tasso di mutazione più lento rispetto a regioni non codificanti), per la sua versatilità e affidabilità, il mtDNA si presta molto bene a indagini di tipo filogenetico e filogeografico. Nel presente studio è stata indagata la variabilità di una regione del mtDNA: la citocromo c ossidasi I (COI).

La citocromo c ossidasi è una proteina che fa parte della catena di trasporto degli elettroni; è l'ultimo enzima della catena ed è costituita da tre subunità. Di queste, la subunità I è la più grande, costituita da 12 eliche transmembrana (Castresana *et al.*, 1994).

## 1.4.2 Marcatori microsatellite

Per la parte di genetica di popolazione, sono stati utilizzati i marcatori microsatellite.

Questi possono ritenersi i marcatori molecolari più usati negli studi genetici grazie alla loro applicazione in molti settori, tra cui la genetica della conservazione, la genetica delle popolazioni e la genetica forense. Il motivo di questo largo impiego va ricercato nel fatto che i microsatelliti sono co-dominanti, multi-allelici, altamente riproducibili, hanno un'alta risoluzione e sono basati sulla reazione a catena della polimerasi (PCR). Nonostante all'inizio l'uso della tecnologia microsatellite sia stato limitato a causa dei costi eccessivi, oggi nuovi ed efficienti metodi di isolamento dei microsatelliti consentono l'applicazione di queste tecnologie in numerose specie. I marcatori molecolari precedenti mancavano di potenza statistica e accuratezza nella stima delle differenze genetiche tra le specie ed all'interno delle stesse, caratteristiche riscontrabili invece nei microsatelliti.

I microsatelliti, conosciuti anche come ripetizioni di sequenze semplici (*SSRs Simple Sequence Repeats*) o corte ripetizioni a tandem (*STR Short Tandem Repeat*), sono regioni non codificate ripetute nel DNA composte da piccoli motivi da 1 a 6 nucleotidi ripetuti in tandem per decine di volte fino ad una lunghezza massima della successione di 200 paia di basi (Rassman *et al.*, 1991).

Queste ripetizioni sono ampiamente diffuse sia nel genoma eucariotico che in quello procariotico (Field and Wills, 1996; Tóth *et al.*, 2000). Sono più lunghe ed abbondanti nei vertebrati che negli invertebrati e la maggior parte di esse è costituita da ripetizioni dinucleotidiche (Wang *et al.*, 1994).

Ogni *locus* microsatellite è composto da una sequenza unica di DNA (regione fiancheggiante), che definisce il *locus*, e da un motivo ripetuto che può essere comune a più *loci* microsatelliti (DeWoody and Avise 2000).

Le SSRs hanno un tasso di mutazione maggiore rispetto ad altre regioni del genoma e ciò le rende funzionali come marcatori molecolari (Jarne and Lagoda, 1996): varia da  $10^{-3}$  a  $10^{-6}$  eventi di mutazione per *locus* in ogni generazione (Sia *et al.*, 2000), con un tasso medio generalmente accettato nei pesci di  $5 \cdot 10^{-4}$  mutazioni per generazione (Lippe *et al.*, 2006). Le mutazioni avvengono come inserzione e/o delezione del motivo ripetuto, che causano variazioni nella lunghezza complessiva della sequenza (Goldstein and Schlötterer, 1999). Le cause delle mutazioni possono essere attribuite prevalentemente ad uno scorrimento della fase di lettura (polymerase slippage) durante la replicazione (Levinson and Gutman, 1987) o alla riparazione del DNA (Tautz, 1989, Strand *et al.*, 1993), o ad un evento di ricombinazione tra molecole di DNA.

Parallelamente, le regioni fiancheggianti tendono a mantenersi costanti, in quanto i tassi di mutazione puntiforme ( $10^{-9}$ - $10^{-10}$  mutazioni per *locus* per generazione, Hancock, 1999) sono molto più bassi di quelli del microsatellite vero e proprio. Ciò consente di disegnare dei *primer* specifici che permettono l'amplificazione via PCR (*Polymerase Chain Reaction*) del *locus* da diversi individui e funzionano spesso anche in specie filogeneticamente vicine (Goldstein *et al.*, 1995).

In base alla composizione, i microsatelliti si distinguono in perfetti, composti, interrotti e imperfetti (Jarne and Lagoda, 1996; tabella 1.1).

**Tabella 1.1: esempio dei quattro diversi tipi di sequenze microsatellite conosciute.**

| TIPO       | SEQUENZA TIPO        |
|------------|----------------------|
| Perfetti   | CACACACACACACACACA   |
| Composti   | CACACACAGAGAGAGAGA   |
| Interrotti | GATGATGATGATTGAATGAT |
| Imperfetti | ACACACATACACACACACAC |

Le ripetizioni più diffuse sono i di-, tri- e tetra-nucleotidi. I *loci* microsatelliti con motivo di-nucleotidico sono usati con maggior frequenza, poiché risultano di più facile isolamento, data la maggiore densità di presenza (in media si trova una ripetizione di dinucleotidi ogni 30-50 kb), e hanno una media di ripetizioni per *locus* minore di 30 (Jarne and Lagoda, 1996). Le ripetizioni dinucleotidiche più frequenti sono CA (particolarmente diffusa negli animali) e TG (Beckmann and Weber, 1992), che mostrano una frequenza doppia rispetto al dinucleotide AT e tripla rispetto alla coppia AG. Tra i trinucleotidi, invece, i motivi più comuni sono CAG e AAT; infine GATA e GACA sono gli unici motivi di tetranucleotidi finora individuati (Jarne and Lagoda, 1996).

Sono marcatori nucleari ipervariabili: sono stati identificati fino a 50 alleli per un singolo *locus*; l'eterozigosità media è del 50% e per alcuni *loci* raggiunge picchi del 100%.

Scoperti originariamente nel genoma degli eucarioti intorno alla metà degli anni '60, hanno assunto un ruolo primario solo alla fine degli anni '80, quando, con l'avvento della PCR, ne è stata messa in luce la versatilità (Jarne and Lagoda, 1996). Il progresso nell'uso dei microsatelliti ha incontrato ritardi anche a causa degli alti costi richiesti per lo sviluppo di specifici *primer*. Tuttavia molti studi

hanno mostrato che il paio di *primer* disegnati per una specie può essere usato anche per altre specie dello stesso genere (Isagi and Suhandono, 1997) o addirittura per differenti generi della stessa famiglia (White and Powell, 1997). Questa caratteristica è conosciuta come trasferibilità o amplificazione cross-specie ed è imputabile alla natura omologa delle sequenze di DNA delle regioni fiancheggianti i microsatelliti. La trasferibilità può rappresentare un fattore importante nell'agevolare l'uso dei microsatelliti difficili da isolare e di quelli con bassa frequenza. Tuttavia, il tasso di successo dell'amplificazione diminuisce con l'aumentare delle differenze genetiche tra le specie (Primmer and Merilä, 2002).

Negli ultimi anni l'interesse dei ricercatori verso i microsatelliti è aumentato considerevolmente per la loro utilità nella costruzione delle mappe genetiche, per la relazione esistente tra variabilità nel numero di ripetizioni e le malattie genetiche umane (Mahadevan *et al.*, 1992; O'Donnell and Warren, 2002), per il loro ampio uso negli studi di genetica delle popolazioni e per le analisi genotipiche e di paternità.

Le metodologie di analisi che si basano sull'impiego di tali marcatori, tuttavia, possono presentare alcuni inconvenienti, che portano a sottostimare le differenze genetiche: il fenomeno dell'omoplasia, ad esempio, non permette di distinguere, mediante elettroforesi, due alleli non omologhi, perché di discendenza diversa, ma di uguale lunghezza. Un altro caso è quello degli alleli nulli che si presentano quando una mutazione, nella regione di appaiamento dei *primer*, impedisce il corretto posizionamento dell'oligonucleotide per l'amplificazione. Questa non può più avvenire: se in una popolazione sono presenti alleli nulli, quindi, ci sarà una sottostima del numero di eterozigoti rispetto all'atteso, poiché negli eterozigoti per l'allele nullo si amplifica uno solo degli alleli e, di conseguenza, questi individui verrebbero erroneamente considerati omozigoti.

## 1.5 Scopo della tesi

In questa tesi, si è voluto indagare la struttura genetica e la filogeografia di *Carcinus aestuarii*. Sono stati analizzati complessivamente 870 individui. In particolare, il lavoro sperimentale ha previsto:

- l'acquisizione dei dati genetici per 11 *loci* microsatellite di 8 campioni di popolazione provenienti dai siti di Marano, Venezia, Goro, Ravenna, Lesina, Aquatina, Taranto e Cagliari;
- l'applicazione *ex novo* di marcatori mitocondriali, con l'amplificazione e sequenziamento di una porzione del gene mitocondriale codificante per la subunità I della citocromo c ossidasi.

Lo scopo di questa tesi di dottorato consiste nello stabilire se i campioni di popolazioni di *C. aestuarii*, con particolare interesse per quelli adriatici, presentano struttura genetica e cercare di ottenere informazioni sulle possibili relazioni tra popolazioni di diversi siti.

Inoltre, si cercherà di capire il ruolo che possono aver ricoperto alcune caratteristiche di *life history* della specie e la storia bio-geografica del Mar Adriatico nel determinare la distribuzione della variabilità genetica in *C. aestuarii*.



## ***2. MATERIALI E METODI***



Il lavoro svolto in questa tesi può essere suddiviso in due parti: isolamento di marcatori molecolari e applicazione degli stessi su campioni di popolazione Mediterranei del crostaceo portunide *Carcinus aestuarii* (Nardo, 1847).

## 2.1 Estrazione del DNA genomico

Il DNA è stato estratto da 870 individui, per i quali sono state prelevate e conservate in freezer a  $-80\text{ }^{\circ}\text{C}$  una o entrambe le chele (a seconda della dimensione dell'organismo), seguendo il protocollo di estrazione *Salting out* (Patwary *et al.*, 1994) che prevede l'isolamento del DNA dopo l'eliminazione delle proteine per precipitazione.

Generalmente questa tecnica viene utilizzata in chimica organica per separare un composto organico da un mezzo acquoso. A tal fine viene ridotta l'affinità di tale composto per l'acqua tramite aggiunta di una soluzione ionica: la presenza di ioni  $\text{Na}^+$  e  $\text{Cl}^-$  in soluzione incrementa la forza ionica dell'acqua e ciò consente al composto di precipitare nella fase acquosa.

Il protocollo prevede inizialmente una distruzione e lisi del materiale di partenza seguito da una rimozione delle proteine ed altri contaminanti ed il recupero finale del DNA. La rimozione delle proteine è ottenuta attraverso digestione con proteinasi K e successiva precipitazione in etanolo o isopropanolo.

Da ogni chela è stata prelevata, con pinze sterili, una piccola porzione di tessuto muscolare che è stata successivamente posta in *ependorf* da 2 ml. Il tessuto è stato quindi trattato con 300  $\mu\text{l}$  di *buffer* di lisi (0.1 M EDTA, 0.05 M TRIS HCl a pH 8), 37.5  $\mu\text{l}$  di tampone di lisi cellulare (SDS 10%) e 12  $\mu\text{l}$  di enzima di lisi (proteinasi K 20 mg/ml). Quindi, le *ependorf* contenenti i reagenti e il tessuto prelevato sono state incubate in un bagnetto termico a  $55\text{ }^{\circ}\text{C}$  per un intervallo di tempo compreso tra 1 e 3 ore (ma mai per più di tre ore), agitandole tramite *vortex* ogni 30 minuti, fino a completa lisi del tessuto.

L'EDTA (acido etilendiamminicotetracetico) è un agente chelante ioni  $\text{Mg}^{2+}$ , riduce la loro disponibilità per le nucleasi e in questo modo protegge il DNA durante i vari passaggi.

L'SDS (sodio dodecilsolfato) è un detergente anionico che destabilizza le membrane cellulari, rompe i legami non covalenti delle proteine provocandone la denaturazione; la proteinasi K è un enzima proteolitico che catalizza l'idrolisi di legami peptidici, favorendo così l'eliminazione delle proteine.

Al termine della digestione, si è proceduto con la precipitazione delle proteine mediante l'aggiunta di 90  $\mu\text{l}$  di soluzione satura di NaCl ( $>6\text{ M}$ ) e successiva agitazione tramite *vortex* per 20 minuti. Mediante centrifugazione a freddo ( $4\text{ }^{\circ}\text{C}$ ) alla massima velocità (14000 RPM), per 30 minuti, si è proceduto alla precipitazione delle proteine e del tessuto non degradato. Il sovrantante, contenente il DNA, è stato prelevato e trasferito in una *ependorf* da 1.5 ml, contenente un volume di isopropanolo freddo, e centrifugato a freddo per 20 minuti ad una velocità di 14000 RPM.

Durante questo passaggio il DNA precipita e, dopo aver eliminato il sovrantante, il pellet contenente il DNA precipitato è stato risospeso in 500  $\mu\text{l}$  di etanolo al 70% per eliminare sali coprecipitati, centrifugando per 5 minuti a temperatura ambiente. Infine, l'etanolo viene eliminato e le *ependorf* messe in stufa a  $37\text{ }^{\circ}\text{C}$  al fine di essiccare completamente il pellet. Questo, essiccato, viene risospeso in

100  $\mu$ l di acqua sterile e autoclavata. Le *ependorf* contenenti il DNA sono state quindi conservate a -20 °C.

L'esito di ogni estrazione è stato verificato mediante elettroforesi su gel di agarosio all'1% in tampone TAE 0.5 X: si scioglie 1 g di agarosio in 100 ml di TAE 0.5X (TRIS 20 mM, acido acetico 10 mM, EDTA 0.5 mM, pH 8.0) e si riscalda poi la soluzione in microonde agitandola ad intervalli di 2 minuti, per favorirne il completo e omogeneo scioglimento. L'agarosio è un polisaccaride che solidifica a una temperatura inferiore ai 42 °C formando una rete irregolare con maglie della dimensione adatta alla separazione delle macromolecole organiche.

Dopo parziale raffreddamento, è stato aggiunto bromuro d'etidio (Sigma, 10 mg/ml; 5  $\mu$ l in 100 ml di TAE), un intercalante che si interpone tra due paia di basi adiacenti delle molecole di DNA, aumentandone la lunghezza, riducendone la densità e rendendole visibili alla luce UV.

La soluzione quindi è stata fatta raffreddare in vaschette con pettini per la formazione di pozzetti in cui caricare 5  $\mu$ l di DNA estratto mescolato a 4  $\mu$ l di *loading dye*. Il *loading dye* è un tampone di carica che contiene glicerolo, per aumentare la densità della soluzione di DNA e permetterne la deposizione nei pozzetti, ed un colorante carico negativamente, l' *Orange G*, per seguire visivamente il fronte di migrazione durante la corsa elettroforetica. Questa è stata effettuata applicando, tra le due estremità del gel, una differenza di potenziale (d.d.p.) di 100 mV (3 V/cm), che induce la migrazione del DNA, carico negativamente, verso il polo positivo (anodo). In un pozzetto sono stati caricati 2  $\mu$ l di un marcatore, le cui bande hanno dimensione nota ( $\lambda$ HindIII, Promega): infatti, confrontando le bande di DNA estratto con quelle del marcatore, è possibile stimare il peso molecolare del DNA estratto e l'eventuale presenza di prodotti aspecifici. Per accertare l'avvenuta estrazione, il gel viene visualizzato ad un transilluminatore a luce UVB (302 nm) e quindi fotografato con il programma *Quantity One* (BioRad, Chemi Doc).

Per verificare se i campioni a disposizione fossero effettivamente appartenenti alla specie *C. aestuarii* è stata effettuata la procedura di restrizione con enzima *AluI* (Promega). Si è amplificata una porzione mitocondriale del gene 16S rRNA: l'enzima *AluI* riconosce una sequenza specifica (AGCT), diagnostica per *C. aestuarii*, presente nella suddetta regione (Geller *et al.*, 1997).

Il volume totale di reazione per un campione è 10  $\mu$ l, di cui 3  $\mu$ l di DNA amplificato 16S e 7  $\mu$ l di mix costituita dall'enzima, il suo buffer di reazione e acqua. La soluzione è stata caricata in piastre da 96 pozzetti, incubata a 37 °C per 1 ora e successivamente a 65 °C per 20 minuti per l'inattivazione dell'enzima. L'avvenuta digestione è stata verificata tramite elettroforesi su gel di agarosio all'1.8%.

## 2.2 Amplificazione di marcatori mitocondriali

Un frammento di 482 paia di basi del gene mitocondriale di *C. aestuarii*, codificante per la subunità I della citocromo c ossidasi (COI), è stato amplificato mediante PCR (*Polymerase Chain Reaction*), una tecnica che consente la sintesi

*in vitro* di acidi nucleici. Essa facilita l'identificazione di segmenti precisi di DNA e permette la sintesi *in vitro* di milioni di copie di tali frammenti.

Gli elementi indispensabili per portare a termine la reazione di PCR sono:

- una coppia di oligonucleotidi o *primer*, cioè di brevi sequenze (18-28 basi circa) di DNA a singolo filamento che risultano complementari alle estremità di una definita sequenza del DNA di interesse. Gli oligonucleotidi sono utilizzati in concentrazioni tali da restare in eccesso per tutta la reazione, in quanto non devono mai costituire un fattore limitante. In genere, i *primer* sono disegnati cercando di ottenere un contenuto in G e C attorno al 50-60% e cercando di evitare regioni a sequenza inusuale, come ad esempio tratti di polipurine, polipirimidine e di motivi ripetuti, o regioni che formano strutture secondarie, come i *loop* interni. La composizione viene considerata anche in funzione della temperatura di *melting* ( $T_m$ ) o di denaturazione dei due frammenti di DNA, che dipende dalla percentuale delle diverse basi e deve essere tale da permettere l'appaiamento del *primer* alla sequenza bersaglio ad una temperatura compresa fra 50 °C e 62 °C. La  $T_m$  è la temperatura alla quale circa il 50% di acido nucleico è nella forma a singolo filamento e il 50% è appaiato; può essere approssimativamente calcolata mediante la reazione:

$$T_m = (A+T) \times 2 \text{ °C} + (C+G) \times 4 \text{ °C}$$

- una DNA polimerasi termostabile, per facilitare l'automatizzazione del processo, di solito la *Taq* polimerasi derivante dal batterio *Thermus aquaticus* (Saiki *et al.*, 1988), che facilita il processo in quanto può essere aggiunta in un'aliquota all'inizio dell'amplificazione. Questo enzima catalizza la sintesi di lunghe catene polinucleotidiche utilizzando desossinucleotidi trifosfati, estendendo i *primer* e usando come stampo uno dei filamenti originali. La *Taq* polimerasi (Promega) utilizzata in questo studio ha un tasso di estensione di 75 nucleotidi/secondo, se le condizioni di reazione sono ottimali;
- DNA genomico come stampo;
- desossinucleotidi trifosfato (*dNTPs*): *dATP*, *dGTP*, *dTTP*, *dCTP*, che devono essere dosati in quantità equimolare con concentrazione compresa fra 50 e 200  $\mu$ M;
- $MgCl_2$ , la cui concentrazione può variare fra 0.5-5.0 mM. Gli ioni  $Mg^{2+}$  sono cofattore della *Taq* polimerasi, la cui concentrazione deve essere ottimizzata per ogni coppia di *primer*. Lo ione  $Mg^{2+}$  è infatti un fattore vincolante l'efficienza della reazione in quanto influenza la temperatura di dissociazione degli oligonucleotidi e la specificità del prodotto. Infatti interagendo coi gruppi fosfato di DNA o oligonucleotidi può dare stabilità agli appaiamenti, ma se presente in eccesso può aumentare la frequenza di appaiamenti dei *primer* con sequenze non perfettamente complementari. Gli ioni  $Mg^{2+}$  formano un complesso solubile con i *dNTPs*, che è essenziale alla loro incorporazione; per questo alla concentrazione calcolata va sempre addizionato un eccesso di 0.5-2.5 mM al di sopra della concentrazione dei *dNTPs*, inoltre incrementano la  $T_m$  della doppia elica di DNA e l'interazione *primer*/stampo. La concentrazione di  $MgCl_2$  influisce

sulla specificità del prodotto, per cui una quantità insufficiente porta a scarsi risultati, mentre un eccesso causa un accumulo di prodotti aspecifici;

- *Buffer* (tampone) di reazione, che determina la forza ionica e il pH ottimale della reazione;
- Acqua.

L'amplificazione dei filamenti di DNA si ottiene grazie alla ripetizione di cicli secondo un profilo termico che prevede tre stadi a differenti temperature:

- **Denaturazione** del DNA ad una temperatura superiore alla sua  $T_m$ , di solito comprese tra 93 °C e 95 °C, con conseguente separazione delle due eliche;
- **Annealing**, in cui avviene l'appaiamento specifico dei *primer* alle estremità complementari della sequenza da amplificare, a temperature comprese tra 45 °C e 60 °C;
- **Extension**, in cui avviene la sintesi di DNA ad opera della polimerasi, che catalizza l'estensione dei *primer* e la formazione dei nuovi filamenti. Questa fase avviene quasi sempre a 72 °C, temperatura ottimale di sintesi per la *Taq* polimerasi.

Ripetuti cicli consentono un'amplificazione esponenziale della sequenza di DNA compresa fra i due *primer*: infatti, il numero di copie di DNA (N) può essere espresso in funzione del numero di cicli (n) e del numero di copie di DNA stampo (X) dalla seguente relazione:

$$N = (2^n - 2n) X$$

In realtà, l'andamento esponenziale della reazione è solamente teorico e non raggiunge l'efficienza del 100%, poiché la concentrazione dell'enzima diviene fattore limitante all'aumentare della concentrazione di substrato. Inoltre, per quanto la *Taq* polimerasi sia un enzima termoresistente, la sua attività si riduce nel tempo a causa dei ripetuti cicli di denaturazione. Nei cicli più tardivi, l'equilibrio della reazione tende a spostarsi a sfavore della polimerizzazione sia per l'accumulo di pirofosfato e la riduzione dei reagenti (*primer* e *dNTPs*), sia per il cosiddetto effetto *plateau*, dovuto all'elevata quantità di DNA presente, che facilita la rinaturazione delle catene omologhe. Di conseguenza, pur aumentando il numero di cicli, non aumenta il numero di copie di DNA.

Si riportano di seguito le condizioni standard di PCR utilizzate in questo studio. Le quantità e le concentrazioni dei reagenti per ogni singolo campione amplificato in 20 µl di volume totale sono:

- 4 µl *buffer* di reazione (5X, Promega)
- 1.1 µl MgCl<sub>2</sub> (25 mM)
- 1 µl *dNTPs* (2 mM)
- 0.5 µl *primer* (forward e reverse 10 µM)
- 0.2 µl *Taq* polimerasi (5 unità/µl)
- 2 µl DNA

Per ogni amplificazione è stato incluso un controllo negativo contenente la sola miscela di reazione (escluso il DNA). Al fine di minimizzare la possibilità di contaminazione, le miscele sono state preparate sotto cappa sterile a flusso laminare, utilizzando come di consueto materiale e reagenti sterili. Tutte le reazioni di PCR di questa tesi sono state condotte mediante il *Thermal Cycler GeneAmp 2700* (Applied Biosystems).

## 2.2.1 Condizioni di amplificazione

Per l'amplificazione del frammento di 482 paia di basi della citocromo c ossidasi I sono stati utilizzati i *primer* (rinominati in questa tesi COI1 e COI2, tabella 2.1) riportati in Roman and Palumbi (2004):

**Tabella 2.1: sequenze degli oligonucleotidi utilizzati per l'amplificazione del mtDNA. Si riportano la sequenza dei *primer* e la lunghezza in paia di basi.**

| Nome | Sequenza 5'-3'                | Dimensione |
|------|-------------------------------|------------|
| COI1 | 5'- GCTTGAGCTGGCATAGTAGG -3'  | 20         |
| COI2 | 5'- GAATGAGGTGTTTAGATTTCG -3' | 21         |

Questi *primer*, specifici per *Carcinus maenas*, sono stati ottenuti dagli autori a partire da un frammento più lungo, amplificato con i *primer* universali riportati in Folmer *et al.* (1994).

La reazione di PCR prevede 30 cicli con fase di denaturazione di 1 minuto a 94 °C, fase di *annealing* di 1 minuto a 50 °C e fase di estensione di 1 minuto a 72 °C come riportato in Roman and Palumbi (2004).

I risultati dell'amplificazione sono stati esaminati con una corsa elettroforetica su gel di agarosio all'1.8% in tampone TAE 0.5X.

## 2.2.2 Purificazione e sequenziamento

Gli amplificati di ogni individuo sono stati purificati con EXO-SAP (Exonucleasi I di *E. coli* e Shrimp Alkaline Phosphatase; USB).

L'esonucleasi elimina i *primer* in eccesso ed eventuale DNA a singolo filamento prodotto in PCR, mentre la fosfatasi elimina i *dNTPs* rimanenti. A 5 µl di prodotto di PCR sono stati aggiunti 0.5 µl di ciascuno dei due enzimi e sono stati incubati per 15 minuti a 37 °C e per 15 minuti a 80 °C (per inattivare gli enzimi). Di questi, è stato prelevato 1 µl o 1.5 µl, in base all'intensità della banda di DNA amplificato, ai quali sono stati addizionati 2 µl di *primer* COI1; a questo punto si è lasciata seccare la miscela in termociclatore a 70°C.

Terminata questa procedura, il DNA purificato è stato sequenziato dal BMR Genomics.

## 2.3 Isolamento e messa a punto di marcatori microsatellite

### 2.3.1 Costruzione di una *library* genomica parziale arricchita

L'isolamento *ex novo* degli 8 *loci* (Cae01, Cae07, Cae14, Cae17, Cae30, Cae33, Cae71 e Cae86), riportato in Marino *et al.* (2008), è stato realizzato attraverso il metodo FIASCO (*Fast Isolation by AFLP of Sequences COntaining repeats*, Zane *et al.*, 2002) che ha permesso la creazione di una *library* arricchita per i *repeat* AC<sub>17</sub>, AG<sub>17</sub>, CAG<sub>11</sub>, AAC<sub>12</sub>, AAT<sub>12</sub>, GATA<sub>9</sub> e CACC<sub>8</sub>. Secondo il protocollo, il DNA viene digerito con un appropriato enzima di restrizione, legato a degli *adapter* specifici ed amplificato mediante *PCR* per ottenere una quantità di templato sufficiente per i passaggi successivi.

L'arricchimento viene effettuato mediante ibridizzazione selettiva con un oligonucleotide biotinilato (sonda), specifico per i *repeat* d'interesse.

Il DNA, ibridizzato con le sonde, può essere recuperato dalla fase acquosa utilizzando delle sferule paramagnetiche ricoperte di streptavidina (molecola con sito di legame ad alta affinità per la biotina) che lega, pressoché irreversibilmente, la biotina della sonda. Dopo una serie di lavaggi di stringenza per l'eliminazione dei frammenti non ibridizzati (e quindi non contenenti *repeat*), il DNA viene eluito dalla sonda ed amplificato mediante *PCR*. In generale, quindi, la procedura seleziona preferenzialmente DNA che contiene i *repeat* d'interesse. Il DNA, ottenuto dopo questa fase del protocollo, viene successivamente clonato per creare una *library* arricchita.

Dalla *library* possono essere identificati e selezionati dei cloni effettivamente contenenti i microsatelliti cercati; tali cloni vengono sequenziati e utilizzati per disegnare dei *primer* specifici per i *loci* microsatellite al fine di amplificarli nel DNA genomico degli individui appartenenti alle popolazioni di interesse.

#### Arricchimento

Il DNA è stato digerito con l'enzima di restrizione *MseI* (Biolabs) e legato a degli *adapter* specifici (tabella 2.2, Vos *et al.*, 1995) secondo le condizioni descritte di seguito.

Tabella 2.2: sequenze degli *adapter* utilizzati dopo la digestione con *MseI*.

| <i>ADAPTER</i>                | Sequenza 5'-3'         |
|-------------------------------|------------------------|
| <i>MseI AFLP adaptor up</i>   | 5'-TACTCAGGACTCAT-3'   |
| <i>MseI AFLP adaptor down</i> | 5'-GACGATGAGTCCTGAG-3' |

Sono stati impiegati 25-250 ng di DNA genomico ed aggiunti ad una miscela di *ligation* con gli *adapter* contenente:

- 2.5 µl *Buffer* di reazione (*OnePhorAll Pharmacia*) 1X
- 1.25 µl DTT 100 mM
- 0.125 µl Bsa 10 µg/ml
- 0.250 µl *MseI* (*New England Biolabs*) 2.5 unità
- 0.5 µl *Adaptor* per *MseI* 50 µM
- 0.25 µl ATP 20 mM
- 1 µl T4 DNA *ligase* (*Amersham-Pharmacia*) 1 unità/ µl
- 5 µl DNA genomico
- 14.125 µl H<sub>2</sub>O

per un volume totale di 25 µl.

La miscela di reazione è stata incubata per 3 ore a 37 °C.

Sono stati impiegati 5 µl del diluito 1:10 della miscela di reazione di *ligation* per una PCR che utilizza i quattro *primer AFLP* specifici per gli *MseI AFLP adapter*, diversi per il tipo di base terminale (A, T, G, C).

La miscela di reazione per l'amplificazione contiene in 20 µl totali:

- 2 µl *buffer* (Promega) 1X
- 1.2 µl MgCl<sub>2</sub> 25 mM
- 4 µl *primer AFLP* 7.5 ng/ µl (1 µl per ciascun *primer*)
- 2 µl *dNTPs* 2 mM
- 0.08 µl *Taq* polimerasi (Promega 5 unità/µl)
- 5.72 µl H<sub>2</sub>O

La reazione di PCR è stata effettuata in un *Thermal Cycler* (*GeneAmp PCR System 3700 Applied Biosystems*) secondo il seguente profilo termico:

- |  |   |          |
|--|---|----------|
| <ul style="list-style-type: none"> <li>- denaturazione 94 °C per 30 secondi</li> <li>- <i>annealing</i> 53 °C per 1 minuto</li> <li>- <i>extension</i> 72 °C per 1 minuto</li> </ul> | } | 22 cicli |
|--|---|----------|

Sono state eseguite 8 repliche della PCR che ha dato prodotto per raggiungere una quantità sufficiente di templatò per il successivo arricchimento.

L'arricchimento è stato realizzato ibridizzando il DNA con una mix di sonde biotinate in 5' contenenti i *repeat* (AC)<sub>17</sub>, (AG)<sub>17</sub>, (CAG)<sub>11</sub>, (AAC)<sub>12</sub>, (AAT)<sub>12</sub>, (GATA)<sub>9</sub> e (CACC)<sub>8</sub>.

Circa 250 ng di DNA (quantificato dal gel d'agarosio), ottenuto dalle precedenti amplificazioni, è stato mescolato, in 25 µl di volume totale, a 80 pmol di ciascuna sonda biotinilata in un volume di 100 µl di SSC 4.2X, SDS 0.07%. È stata successivamente realizzata una denaturazione a 95 °C per 3 minuti ed un seguente raffreddamento fino a 50 °C per 30 minuti per consentire l'*annealing* tra le sonde e il DNA. La soluzione è stata quindi diluita in 300 µl di TEN<sub>100</sub>.

Per isolare il DNA ibridizzato dai frammenti non selezionati, sono state utilizzate delle sferule paramagnetiche di streptavidina (*Streptavidin Coated Magnetic Particles, Boehringer*).

A tale scopo, 100  $\mu$ l di sferule (pari a 1 mg) sono stati più volte lavati in TEN<sub>100</sub> (10 mM Tris-HCl, 1 mM EDTA, 100 mM NaCl pH 7.5) e risospesi in 80  $\mu$ l dello stesso *buffer*. Quindi, alle sferule di streptavidina è stato aggiunto un eccesso di tRNA (5  $\mu$ l di una soluzione 100 mg/ml) per minimizzare i legami non specifici. Infine, le sferule sono state mescolate con il DNA precedentemente ibridizzato con le sonde e diluito con 300  $\mu$ l di TEN<sub>100</sub>. La miscela finale è stata lasciata a temperatura ambiente per 30 minuti.

La soluzione finale è stata sottoposta a tre lavaggi non stringenti con 400  $\mu$ l di TEN<sub>1000</sub> (10 mM Tris-HCl, 1 mM EDTA, 1 M NaCl pH 7.5) e a tre successivi lavaggi stringenti (con 400  $\mu$ l di tampone SSC 0.2X, SDS 0.1%) per eliminare il DNA non specifico. Ogni lavaggio è stato condotto a temperatura ambiente, agitando con cautela le provette per 5 minuti. Sia l'ultimo lavaggio non stringente sia quello stringente sono stati conservati. Per prelevare il sovranatante, il complesso DNA-sonda biotinilata-streptavidina è stato isolato applicando un campo magnetico alla soluzione.

Due successive denaturazioni hanno permesso di separare le sferule con streptavidina dalla biotina. La prima denaturazione avviene a 95 °C per 5 minuti con 50  $\mu$ l di TE (10 mM Tris-HCl, 1 mM EDTA, pH 8). Il sovranatante contenente il DNA d'interesse è stato rapidamente prelevato.

La seconda denaturazione è stata condotta trattando le sferule con 12  $\mu$ l di NaOH 0.15 M neutralizzato aggiungendo acido acetico 0.166 M. Alla soluzione è stato aggiunto TE per un volume finale di 100  $\mu$ l.

Il DNA, recuperato dai lavaggi e dalle denaturazioni, è stato fatto precipitare con un volume di isopropanolo freddo e sodioacetato (concentrazione finale di 0.15 M) e risospeso in 100  $\mu$ l di acqua dH<sub>2</sub>O sterile.

Per l'amplificazione sono stati utilizzati 5  $\mu$ l di DNA arricchito e le concentrazioni per la miscela di reazione rimangono le stesse utilizzate precedentemente con i *primer AFLP* specifici per gli *adapter MseI*.

Il profilo termico, con *annealing* a 53 °C, è stato ripetuto per 30 cicli.

I prodotti ottenuti dall'amplificazione del DNA della frazione più stringente sono stati considerati adatti per la preparazione di una *library* arricchita per frammenti contenenti i *repeat* di interesse.

## **Clonaggio**

La *library* dei frammenti, ottenuti attraverso i passaggi sopra descritti, è stata realizzata impiegando il kit *TOPO TA (Cloning Invitrogen)*. Questo rappresenta uno strumento altamente efficiente per inserire frammenti di prodotti di PCR nel vettore plasmidico pCR<sup>®</sup> II TOPO<sup>®</sup> (Versione P) e trasformarlo in cellule competenti di *Escherichia coli* selezionate per la resistenza all'antibiotico ampicillina.

Il vettore plasmidico viene fornito con l'enzima Topoisomerasi I che si lega covalentemente alla sequenza del plasmide. Si utilizza un vettore linearizzato (aperto), con una timidina sporgente all'estremità 3' per il legame con la Topoisomerasi I e con l'inserito, che deve quindi possedere un'adenina sporgente all'estremità 3' (inserita normalmente per attività della *Taq* polimerasi).

In 6  $\mu$ l totali, è stato mescolato 1  $\mu$ l di DNA (prodotto di PCR), 1  $\mu$ l di soluzione salina (1.2 M NaCl, 0.06 M MgCl<sub>2</sub>) e 1  $\mu$ l di vettore plasmidico pCR<sup>®</sup> II TOPO<sup>®</sup>.

La miscela di reazione è stata incubata per 5 minuti a temperatura ambiente e, prima di metterla in ghiaccio, è stato aggiunto 1  $\mu$ l di *TOPO Cloning Stop Solution* 6X. Due microlitri della miscela *TOPO Cloning* sono stati aggiunti, in una provetta, a cellule di *E. coli* rese chimicamente competenti e sottoposte a *shock* termico. Le cellule di *E. coli* sono state prima incubate in ghiaccio per 5-30 minuti e poi immerse in un bagno a 42 °C per 30 secondi senza scuoterle. Subito dopo averle sottoposte allo *shock* termico, le cellule sono state nuovamente riposte in ghiaccio e sono stati aggiunti 250  $\mu$ l di SOC (950 ml H<sub>2</sub>O deionizzata, 20 g Bacto-tryptone, 5 g Bacto-yeast extract, 0.5 g di NaCl, 20 mM glucosio) a temperatura ambiente. Le cellule sono state lasciate crescere mentre la provetta veniva tenuta in posizione orizzontale e agitata gentilmente per 1 ora a 37 °C.

Nel frattempo si è proceduto all'aggiunta di 40  $\mu$ l di X-Gal (40 mg/ml) (substrato della  $\beta$ -galattosidasi) su piastre *petri* preparate in precedenza contenenti terreno solido LB e 50  $\mu$ g/ml ampicillina. Il terreno solido LB (Luria Bertani) delle piastre è stato preparato con 1% di triptone, 0.5% estratto di lievito, 1% NaCl a pH 7.15 g/l di agar (aggiunto prima di autoclavare la soluzione per 20 minuti).

Le piastre *petri* così composte sono state incubate a 37 °C per 30 minuti.

A questo punto sono stati realizzati due trasferimenti distinti per ogni trasformazione, con volumi differenti di cellule (40  $\mu$ l e 20  $\mu$ l), in modo da ottenere, almeno per una piastra, una spaziatura delle colonie che ne facilitasse il prelievo. Le piastre sono state lasciate ad incubare a 37 °C per tutta la notte per far crescere le colonie.

Poiché il terreno utilizzato contiene ampicillina, possono crescere su di esso solo le cellule trasformate che hanno incorporato il vettore plasmidico. Inoltre in virtù del fatto che il vettore plasmidico contiene il gene  $\beta$ -galattosidasi è possibile discriminare le colonie che contengono l'inserito da quelle che non lo contengono. Nei vettori in cui l'inserito non è presente, infatti, il gene per la  $\beta$ -galattosidasi è funzionante e metabolizza l'X-Gal dando un prodotto di colore blu; nei vettori contenenti l'inserito, invece, tale gene viene interrotto in seguito all'inserimento del frammento di PCR nel plasmide *TOPO*. In questo modo il gene perde la capacità di metabolizzare l'X-Gal e, quindi, le cellule non manifesteranno la colorazione blu tipica delle cellule contenenti il vettore circolarizzato e privo di inserito, ma una colorazione bianca o blu chiaro.

### **Amplificazione dell'inserito**

Per amplificare l'inserito, sono stati selezionati dei cloni da impiegare per una PCR con la coppia di *primer* universali T7 Promoter e M13REV (-20, tabella 2.3).

**Tabella 2.3: sequenze degli oligonucleotidi M13REV (-20) e T7 Promoter. Sono riportate la sequenza dei due *primer*, la temperatura di *annealing* ( $T_a$ ), calcolata secondo la composizione in basi e la lunghezza in paia di basi.**

| Nome                | Sequenza 5'-3'           | $T_a$ | Dimensione |
|---------------------|--------------------------|-------|------------|
| <b>M13REV (-20)</b> | 5'GTAAAACGACGGCCAG3'     | 50 °C | 17         |
| <b>T7 Promoter</b>  | 5'TAATACGACTCACTATAGGG3' | 56 °C | 20         |

La colonia è stata lisata a 96 °C per 6 minuti in 12 µl di dH<sub>2</sub>O. Cinque microlitri di lisato sono stati impiegati per la PCR, con le seguenti quantità di reagenti per un volume totale di 20 µl:

- 2 µl *buffer* Promega 1X
- 2 µl MgCl<sub>2</sub> (25mM)
- 2 µl dNTPs (2mM)
- 1 µl M13 (10µM)
- 1 µl T7 (10 µM)
- µl *Taq* Promega (5 unità/µl)
- 6.9 µl H<sub>2</sub>O

Il profilo termico utilizzato è:

- |   |   |          |
|---|---|----------|
| <ul style="list-style-type: none"> <li>- 95 °C per 2 minuti</li> <li>- 95 °C per 30 secondi</li> <li>- 56 °C per 1 minuto</li> <li>- 72 °C per 1 minuto e 30 secondi</li> <li>- 72 °C per 5 minuti</li> </ul> | } | 33 cicli |
|---|---|----------|

Al termine della PCR gli amplificati sono stati caricati su di un gel di agarosio all'1.8%, per verificare l'avvenuta amplificazione. La procedura di verifica è molto simile a quella utilizzata per il DNA estratto fatta eccezione per il quantitativo di agarosio che in questo caso è di 1.8 g su 100 ml di TAE 0.5X e per il marcatore di peso molecolare che non è più λHindIII (Promega) ma 1 Kb plus (Invitrogen) che meglio si presta a visualizzare il dimensionamento degli amplificati.

### 2.3.2 Purificazione e sequenziamento del DNA

Dalla foto del gel, sono stati scelti i migliori amplificati e sono stati purificati con EXO-SAP (Exonucleasi I di E. coli e Shrimp Alkaline Phosphatase; USB). Per quanto riguarda il metodo citato si rimanda a quanto descritto nella sezione 2.2.2 (Isolamento dei marcatori mitocondriali, pag. 5).

Mediante il sequenziamento sono state identificate le colonie realmente contenenti microsatelliti che sono state utilizzate per il disegno dei *primer*.

### 2.3.3 Disegno dei *primer*

Le sequenze (in formato ABI) sono state analizzate con il programma *Chromas* 2.12. Una volta individuate le ripetizioni corrispondenti ai microsatelliti e la presenza di sufficiente sequenza fiancheggiante, sono state disegnate le coppie di *primer* al fine di poter amplificare i *loci* microsatellite da tutti i campioni a disposizione. In molti casi è stato necessario sequenziare i prodotti di PCR anche con il *primer* T7 per poter avere la sequenza completa: le sequenze così ottenute

dai due filamenti (forward e reverse) sono state allineate con il software *Clustal X* (Thompson *et al.*, 1997), per ottenere la sequenza completa di ogni clone.

La sequenza del *primer* è stata individuata utilizzando prima il programma on line *Primer3* ([http://www-genome.wi.mit.edu/cgi-bin/primer/primer3\\_www.cgi](http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi)) e, successivamente, il software *Oligo Explorer* versione 1.0.3 (<http://www.uku.fi/~kuulasma/OligoSoftware>).

*Primer3* permette di impostare le caratteristiche alle quali devono rispondere i *primer*, quali ad esempio dimensioni del frammento amplificato, temperatura di *melting* della coppia di *primer*, il loro contenuto in basi C e G, la loro lunghezza, la complementarietà presente al 3' tra coppie di *primer*. *Oligo Explorer* consente, invece, di verificare le strutture secondarie e i dimeri che i *primer* avrebbero potuto formare tra di loro, evitando così eventuali appaiamenti alle estremità 3', che potrebbero promuovere la formazione di artefatti durante l'amplificazione.

### 2.3.4 Condizioni di amplificazione

Per raggiungere delle condizioni ottimali di amplificazione per i *primer* disegnati, sono state eseguite varie prove per ciascuno dei *loci* su 10 individui, partendo da concentrazioni e profilo termico comuni. In questa sede sono state messe a punto anche le condizioni ottimali di amplificazione per 3 coppie di *primer* (Cma02EPA, Cma04EPA e Cma14EPA) provenienti da studi condotti su *C. maenas* (Tepolt *et al.*, 2006): anche in questo caso sono state eseguite delle prove per ciascuno dei 3 *loci* su 10 individui a partire, però, dalle concentrazioni e dai profili termici citati nel lavoro originale.

Per l'amplificazione di ogni *locus* sono stati impiegati da 2 a 4 µl di DNA in 20 µl di volume totale di miscela di reazione contenente:

- 4 µl *buffer* Promega (5X)
- 1.1 µl MgCl<sub>2</sub> (25mM)
- 1 µl *d*NTPs (2mM)
- 0.2 µl *Taq* Promega(5u/µl)
- 1 µl *primer* (10µM) (0.5µl For; 0.5µl Rev)
- 2 µl DNA
- 10.7 µl dH<sub>2</sub>O

Per l'ottimizzazione del profilo termico sono state sperimentate diverse temperature di *annealing* in funzione delle caratteristiche delle coppie di *primer* disegnati.

Il profilo base prevede:

- uno step iniziale di pre-denaturazione del DNA a 94 °C per 1 minuto,
- seguito da 30-35 cicli, in cui si susseguono:
  - a. uno step di denaturazione a 94 °C per 30 secondi;
  - b. uno step di *annealing* a temperatura variabile dipendente dalla coppia di *primer* per 40 secondi;
  - c. uno step di *extension* a 72 °C per 1 minuto;
- uno step di estensione finale a 72 °C per 5 minuti.

Il successo della reazione di PCR è stato verificato mediante corsa elettroforetica in gel di agarosio all'1.8 %.

Per ogni coppia uno dei due *primer* è stato marcato all'estremità 5' con un fluoroforo, che emette fluorescenza ad una specifica lunghezza d'onda se eccitato da una luce *laser*. In questo modo è possibile individuare i frammenti amplificati durante la migrazione elettroforetica in un sequenziatore automatico. Nel caso specifico, sono stati utilizzati tre diversi fluorofori: TAMRA (giallo), HEX (verde), 6-FAM (blu, tabella 2.4). Grazie al fatto che le coppie di *primer* sono state marcate con fluorofori diversi, è stato possibile analizzare contemporaneamente sino a tre *loci* per individuo: utilizzando piastre da PCR (96 pozzetti da 0.2 ml), in ogni pozzetto sono stati riuniti gli amplificati di 3 *loci* diversi. Inoltre è stato possibile incrementare il numero di *loci* analizzabili in una singola corsa differenziandoli ulteriormente per dimensioni.

Per ciascun campione è stata utilizzata una quantità variabile di DNA amplificato determinata di volta in volta in base alla foto del gel in questione. Le piastre così preparate sono state analizzate mediante elettroforesi denaturante su capillare dal servizio di *genotyping* del BMR Genomics, che ha utilizzato sequenziatori *ABI PRISM DNA Analyzer 3100/3700*. Il dimensionamento dei singoli alleli è stato ottenuto tramite l'aggiunta di uno standard interno (*GS 400 Hd ROX*) ed effettuato dal BMR Genomics.

**Tabella 2.4:** elenco dei *primer* utilizzati in questa tesi. Sono riportati il numero di *repeat*, la sequenza dei *primer*, il tipo di fluoroforo con il quale sono stati marcati, la temperatura di *annealing* e il numero di cicli utilizzati.

| Locus                 | Repeat   | Sequenza del primer (5'-3')                                  | Fluoroforo | T <sub>a</sub> (°C); cycles |
|-----------------------|--|--|------------|-----------------------------|
| Cae01 <sup>1</sup>    | (CAT) <sub>12</sub> CGT (CAT) <sub>8</sub>                           | F: CCTAACCTATCTCAACAGTG<br>R: CTGCTTGCCTTGTTCGTA             | 6-FAM      | 53 (35)                     |
| Cae07 <sup>1</sup>    | (GA) <sub>28</sub>   | F: TGGCTCTTTTCACCTTCTGT<br>R: CTATCGGAAGAGACGCAAGA           | HEX        | 50 (35)                     |
| Cae14 <sup>1</sup>    | (CACAGA) <sub>9</sub> CAGA<br>(CACAGA) <sub>2</sub> GAGA<br>(CACAGA) | F: AACACATCAGTGAGTGCCAC<br>R: CTAAGCCATCTGGAATCACG           | TAMRA      | 57 (35)                     |
| Cae17 <sup>1</sup>    | (CA) <sub>4</sub> GC(CA) <sub>13</sub> GC(CA) <sub>8</sub>           | F: ACTCACCTCCGTCATCACC<br>R: TCCTGAGCTCCATCAAGACT            | 6-FAM      | 55 (30)                     |
| Cae30 <sup>1</sup>    | (GT) <sub>9</sub>  | F: TACAGGGTAAGCCACAGTCC<br>R: GCAGAAGAGGTCACGAGAAG           | TAMRA      | 55 (30)                     |
| Cae33 <sup>1</sup>    | (AC) <sub>13</sub> TTTT(AC) <sub>9</sub>                             | F: CCATACAGGGCTGTACAGAAA<br>R: CGGAATCTCGGTTATATGGCTT        | 6-FAM      | 52 (30)                     |
| Cae71 <sup>1</sup>    | (GT) <sub>17</sub>   | F: CCAGGACTAACAGTGGGTGA<br>R: TGCATGTTAGCCACTCCTCT           | HEX        | 56 (35)                     |
| Cae86 <sup>1</sup>    | (GA) <sub>22</sub>   | F: GCGGTGAAGAAGATGGAGTA<br>R: GCGAGCTCAAATGAATAATGGT         | TAMRA      | 51 (35)                     |
| CmaEPA02 <sup>2</sup> | (TAGA) <sub>7</sub>  | F: TTGATCCATATATCTATTCTTTCTTGC<br>R: TTTGGGGGAAATACTTATCTCTT | 6-FAM      | 52 (30)                     |
| CmaEPA04 <sup>2</sup> | (TAGA) <sub>10</sub>   | F: GAGCTCCAGGAAACTGTATCTGA<br>R: GCCCTCTATCTCGCTTTATATCTC    | HEX        | 56 (30)                     |
| CmaEPA14 <sup>2</sup> | (GGT) <sub>8</sub>   | F: ACGGCTCACCTACGTGCACT<br>R: GGCTGTGGTCTGTGTTTCATT          | TAMRA      | 59 (35)                     |

1 - isolamento da *C. aestuarii* riportato in Marino *et al.*, 2008.

2 - isolamento da *C. maenas* riportato in Tepolt *et al.*, 2006.

## 2.4 Applicazione dei marcatori

### 2.4.1 Campionamento

I campioni di popolazione di *Carcinus aestuarii*, analizzati in questo studio, sono stati raccolti in 8 siti interni al Mar Mediterraneo (figura 2.1) in un periodo di tempo compreso tra il 2005 e il 2008. Sei degli otto siti sono situati nel Mar Adriatico (Laguna di Marano, Laguna di Venezia, Sacca di Goro, Ravenna, Laguna di Lesina e Laguna di Aquatina), uno nel Mar Ionio (Taranto) e uno nel Mar Tirreno (Cagliari).

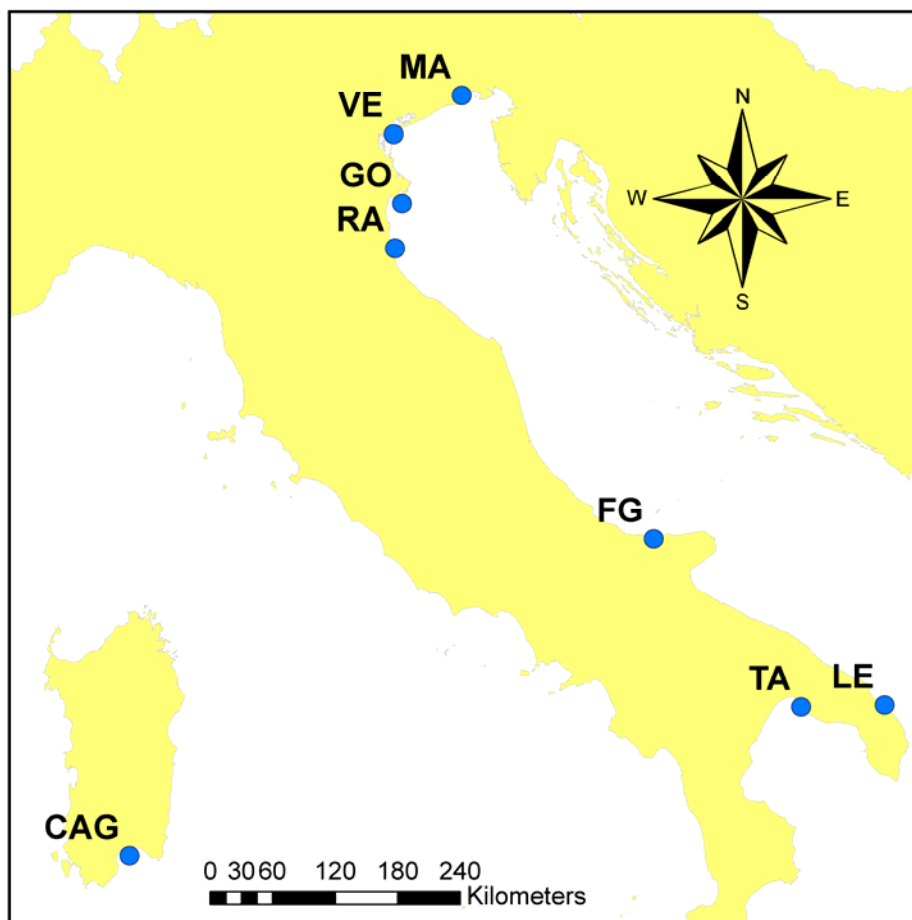


Figura 2.1: mappa rappresentativa dei siti di campionamento.

I campioni provenienti dalla Laguna di Venezia sono stati raccolti in 7 siti interni alla laguna (Ca' Zane, Saline [ex Saline San Felice ed ex Salina Bossolaro], Paludi della Centrega, Campalto, San Clemente, Fusina, Ca'Roman interna, figura 2.2) e uno esterno alla stessa (Ca'Roman esterna), in tre diverse campagne di campionamento avvenute tra il 2005 e il 2007, per un numero complessivo di individui pari a 486.



- |                         |                |                     |
|-------------------------|----------------|---------------------|
| □ Ca' Zane              | □ Campalto     | □ Ca' Roman interna |
| □ Saline                | □ San Clemente | □ Ca' Roman esterna |
| □ Paludi della Centrega | □ Fusina       |                     |

**Figura 2.2: mappa rappresentativa dei siti di campionamento interni alla Laguna di Venezia.**

Per quanto riguarda i siti restanti, è stata campionata una cinquantina di organismi ciascuno (tabella 2.5): più precisamente 55 individui a Marano (Ma), 60 a Goro (Go), 50 a Ravenna (Ra), 54 a Lesina (Fg), 67 ad Aquatina (Le), 48 a Taranto (Ta) e 50 a Cagliari (Cag).

Complessivamente, quindi, le analisi molecolari sono state condotte su 870 individui.

**Tabella 2.5:** in tabella vengono riportati i siti di campionamento e il numero di individui catturati per sito e per anno.

| Sito di campionamento              |                              | n° individui catturati nel 2005 | n° individui catturati nel 2006 | n° individui catturati nel 2007 | n° individui catturati nel 2008 |
|------------------------------------|------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| <b>Marano (Ma)</b>                 |                              |                                 | 55                              |                                 |                                 |
| <b>Venezia (Ve)</b>                | <b>Ca'Zane</b>               |                                 | 47                              |                                 |                                 |
|                                    | <b>Saline</b>                |                                 | 40                              |                                 |                                 |
|                                    | <b>Paludi della Centrega</b> |                                 | 20                              |                                 |                                 |
|                                    | <b>Campalto</b>              |                                 |                                 | 25                              |                                 |
|                                    | <b>San Clemente</b>          |                                 |                                 | 25                              |                                 |
|                                    | <b>Fusina</b>                | 50                              | 47                              | 25                              |                                 |
|                                    | <b>Ca'Roman interna</b>      | 32                              | 44                              | 25                              |                                 |
| <b>Ca'Roman esterna</b>            | 37                           | 69                              |                                 |                                 |                                 |
| <b>Goro (Go)</b>                   |                              |                                 | 60                              |                                 |                                 |
| <b>Ravenna (Ra)</b>                |                              |                                 |                                 | 50                              |                                 |
| <b>Lesina (Fg)</b>                 |                              |                                 |                                 |                                 | 54                              |
| <b>Aquatina (Le)</b>               |                              |                                 |                                 | 67                              |                                 |
| <b>Taranto (Ta)</b>                |                              |                                 |                                 |                                 | 48                              |
| <b>Cagliari (Cag)</b>              |                              |                                 |                                 | 50                              |                                 |
| <b>Totale individui campionati</b> |                              |                                 | <b>870</b>                      |                                 |                                 |

## 2.4.2 Analisi dati

### 2.4.2.1 DNA mitocondriale

Per quanto riguarda i marcatori mitocondriali, sono stati analizzati 32 individui (tranne per il campione di Ravenna, in cui gli individui analizzati sono 31) provenienti dagli 8 campioni di popolazione, ottenendo così 255 sequenze. Le sequenze sono state allineate mediante il programma *Clustal X* (Thompson *et al.*, 1997) e analizzate attraverso il programma *MEGA* versione 4.0 (Tamura *et al.*, 2007) al fine di calcolare la composizione nucleotidica (la percentuale dei quattro nucleotidi), il numero di siti variabili e il numero di siti informativi per la parsimonia.

In una prima fase le analisi sono state condotte su tutte le sequenze ottenute; successivamente, per chiarezza grafica nell'esposizione dei risultati, sono stati individuati i differenti aplotipi rappresentativi di tutte le sequenze mediante il programma *TCS* versione 1.21 (Clement *et al.*, 2000), e le analisi sono state condotte sugli aplotipi. Nell'analisi sono state incluse, inoltre, 13 sequenze, presenti in GenBank (AY616437-AY616444, DQ523682-523686, Roman and

Palumbi, 2004; Roman, 2006), del frammento mitocondriale della citocromo c ossidasi I di *Carcinus maenas* e una sequenza di *C. aestuarii* (AY616445, Roman and Palumbi, 2004).

Attraverso il programma *ModelTest 3.7* (Posada and Crandall, 1998), è stata effettuata la ricostruzione delle relazioni filogenetiche, mediante il metodo della *Maximum Likelihood* (ML). Questo metodo si basa sulla ricerca dell'albero con la topologia che massimizza la verosimiglianza, scegliendo per ogni albero la lunghezza delle ramificazioni che meglio si adatta al modello evolutivo specificato a priori. In questo caso il modello utilizzato è l'HKY+G (Hasegawa-Kishino-Yano + il modello di sostituzione G) come il modello più appropriato di evoluzione di sequenza suggerito dal programma *ModelTest 3.7*. Questo *software* permette infatti di ricavare il modello di sostituzione nucleotidica più appropriato per i dati di sequenza in esame, attraverso il metodo *hLRTs* (*Hierarchical Likelihood Ratio Test*). Sulla base del modello di sostituzione selezionato, è stato ricostruito, attraverso il *software PAUP\** 4.0 (Swofford, 1998), un albero *Maximum Likelihood*. Quindi, mediante lo stesso programma è stata calcolata la likelihood di quest'albero, assumendo o meno un *molecular clock* ed è stata verificata la significatività di  $\Delta L$  secondo i due modelli attraverso un *LRT*. In questo caso, è stata verificata la presenza di orologio molecolare.

Sulla base di questi risultati, dell'albero *Maximum Likelihood* e del modello di sostituzione selezionati, è stato possibile calibrare l'orologio molecolare fissando l'altezza totale dell'albero a  $5.3 \cdot 10^6$  anni fa: la calibrazione è basata sul momento in cui si ritiene sia avvenuta la separazione tra la specie *Carcinus maenas* e la specie *C. aestuarii*. Attraverso tale calibrazione è stato possibile, utilizzando il programma *BEAST 1.4.8* (Drummond and Rambaut, 2007), ottenere un albero orientato nel tempo e riportante la diversificazione degli aplotipi in milioni di anni; inoltre si è potuto stimare il tasso di sostituzione di  $1.058 \cdot 10^{-2}$  mutazioni per *locus* per anno.

### Indici di diversità

La variabilità genetica dei campioni di popolazione, stimata attraverso il *software Arlequin 3.1* (Excoffier *et al.*, 2005), è stata descritta mediante due statistiche, l'haplotype diversity ( $h$ ) e la nucleotide diversity ( $\pi$ ).

L'haplotype diversity ( $h$ ) è una misura della variabilità genetica, corrisponde concettualmente alla proporzione attesa di individui eterozigoti per il *locus* considerato ed è definita come:

$$h = 1 - \sum_{i=1}^k p_i^2$$

dove  $k$  è il numero totale di aplotipi e  $p$  la frequenza dell' $i$ -esimo aplotipo presente nella popolazione.

Questa statistica varia tra zero, in caso di completa omogeneità, e uno, nel caso in cui tutti gli individui presenti nella popolazione abbiano aplotipi diversi e corrisponde alla probabilità di campionare due sequenze diverse estraendo a caso due individui dalla popolazione.

La nucleotide diversity ( $\pi$ ), invece, misura il grado di polimorfismo in una popolazione e viene definita come:

$$\pi = \sum_{ij}^k p_i p_j d_{ij}$$

dove  $k$  è il numero totale di aplotipi,  $p_i$  è la frequenza dell'aplotipo  $i$ ,  $p_j$  è la frequenza dell'aplotipo  $j$  e  $d_{ij}$  è il numero medio di differenze nucleotidiche di due individui estratti a caso dalla popolazione.

### Analisi della struttura di popolazione

Mediante il *software Arlequin 3.1* è stato possibile verificare la differenziazione tra i singoli campioni di popolazione attraverso le F-statistiche, coefficienti sviluppati da Wright (1951). Queste misurano il grado di differenziamento genetico tra popolazioni; in particolare l'indice  $F_{st}$  corrisponde alla distanza genetica calcolata tra le singole coppie di popolazioni. Il valore di tale indice è pari a 0 in caso di omogeneità e tende a 1 in caso di differenziamento completo; si calcola confrontando l'eterozigosità attesa all'interno di una subpopolazione con quella della popolazione totale, utilizzando la formula

$$F_{st} = \frac{(H_t - H_s)}{H_t}$$

dove  $H_t$  è l'eterozigosità attesa della popolazione totale, mentre  $H_s$  è l'eterozigosità media attesa della subpopolazione.

Per calcolare gli  $F_{st}$ , il software si avvale delle formule di Weir and Cockerharm (1984)

$$\theta = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_b^2 + \sigma_w^2}$$

dove  $\theta$  corrisponde all'indice  $F_{st}$  di Wright,  $\sigma$  rappresenta la varianza tra campioni ( $\sigma_a$ ), tra individui all'interno di uno stesso campione ( $\sigma_b$ ) e tra individui ( $\sigma_w$ ), sotto un modello ad isola in cui tutte le popolazioni sono equivalenti e vi è equilibrio tra migrazione e deriva genetica. La significatività viene calcolata mediante permutazioni (in questo caso 10000) del *dataset* originale.

È stato possibile inoltre calcolare un equivalente degli indici  $F_{st}$ , attraverso l'analisi molecolare della varianza (*AMOVA*, Excoffier *et al.*, 1992), effettuando un confronto globale tra tutte le popolazioni. L'*AMOVA* è uno strumento utile per verificare l'eterogeneità della distribuzione della variabilità genetica tra popolazioni, ed è caratterizzata dalla possibilità di utilizzare le distanze molecolari tra i diversi aplotipi. Viene calcolata la varianza per diversi livelli gerarchici, e viene poi verificato quanto della varianza totale sia spiegata dalle differenze tra individui di una popolazione e quanta, invece, derivi dalla suddivisione in popolazioni diverse. L'indice che riassume questa relazione è stato originariamente denominato  $\Phi_{st}$  (Excoffier *et al.*, 1992). La significatività viene calcolata mediante permutazioni (10000), in cui gli individui sono riassegnati casualmente alle diverse popolazioni: in questo modo si calcola la distribuzione di  $\Phi_{st}$  attesa in assenza di suddivisione in popolazioni. Si calcola poi un valore di probabilità (P) che rappresenta la frequenza con cui è possibile ottenere un valore

di  $\Phi_{st}$  maggiore di quello osservato per solo effetto del caso, e dunque la probabilità che il valore osservato sia frutto di errori di campionamento.

### Test di neutralità

I test di neutralità permettono di trarre indicazioni sui processi evolutivi che agiscono sulle popolazioni studiate e di investigare la storia dell'espansione demografica. I test di neutralità di Tajima e di Fu richiedono l'assunzione di equilibrio tra mutazione e deriva: la loro significatività è spesso interpretata come mancanza di equilibrio di popolazione.

Il test di neutralità di Tajima (Tajima, 1989) assume un modello a siti infiniti senza ricombinazione e verifica lo scostamento dallo zero delle differenze standardizzate ( $D$ ) tra due stime di  $M$  ( $2Nu$  per un sistema aploide, dove  $N$  rappresenta la dimensione della popolazione e  $u$  il tasso di mutazione), calcolate dalla media delle *pairwise differences* (numero di differenze tra diverse sequenze prese a coppie) e dal numero di siti segreganti.  $M$ , quindi, è un parametro che determina il livello di variabilità all'interno della popolazione all'equilibrio. Processi non neutrali e di non-equilibrio spostano i valori di  $D$  verso valori positivi o negativi. La significatività dello scostamento è misurata mediante il confronto con una distribuzione  $\beta$ .

Il test di Fu (Fu, 1997), come quello di Tajima, è basato sul modello a siti infiniti senza ricombinazione ed è appropriato per brevi sequenze di DNA o aplotipi RFLP. Questo test valuta la probabilità di osservare un campione casuale, che risponde alle caratteristiche di neutralità, con un numero di alleli simile o minore rispetto al valore osservato dato dal numero di *pairwise differences* osservato, considerato come parametro che stima  $\theta$ . Il test di Fu è particolarmente sensibile all'espansione demografica di popolazione, restituendo generalmente valori di  $F_s$  molto negativi.

### Mismatch distribution

Rogers and Harpending (1992) hanno proposto un modello per ricostruire, a partire dalla distribuzione delle differenze medie tra sequenze (*mismatch distribution*), alcuni aspetti della storia evolutiva delle popolazioni. Secondo le simulazioni condotte, il tipo di distribuzione assunta dal conteggio del numero di differenze nei confronti a coppie dipende dai processi a cui la popolazione in esame è andata incontro. In particolare, per una popolazione di dimensione costante la distribuzione tende verso una curva di *Poisson* (Avise *et al.*, 1988). Rogers and Harpending hanno proposto un modello per una popolazione non in equilibrio che abbia subito un'espansione o un declino: la distribuzione delle differenze medie tra sequenze assume una forma a campana, unimodale. Tale modello prevede una crescita o un declino istantanei da una dimensione effettiva  $\theta_0$  della popolazione al tempo 0 ( $\theta_0 = 2N_0u$ ) ad una dimensione finale  $\theta_1$  ( $\theta_1 = 2N_1u$ ), dove  $N$  è la dimensione della popolazione e  $u$  il tasso di mutazione per sito per anno; si assume che la crescita inizi  $\tau$  unità di tempo di mutazione fa ( $\tau = 2ut$ ). Si può quindi confrontare qualitativamente se la propria distribuzione osservata è più in accordo con una distribuzione *poissoniana* o con quella prevista in caso di dimensione di popolazione variabile. Inoltre, se si possiede una stima di  $u$ , è

possibile datare approssimativamente il momento iniziale dell'espansione, utilizzando la formula:

$$t = \frac{\tau}{2u}$$

In questa tesi è stato utilizzato un valore di  $u = 5.1 \cdot 10^{-6}$  mutazioni per anno per *locus*, che risulta considerando un tasso di sostituzione per nucleotide ( $2u$ ) dello 0.058% per milione di anni ed una lunghezza del *locus* di *C. aestuarii* di 482 bp. Poiché la struttura di popolazione potrebbe avere un effetto limitato sulla *mismatch distribution* (Rogers, 1995; Van Houdt *et al.*, 2003), il modello di espansione improvvisa è stato testato a differenti livelli di raggruppamento.

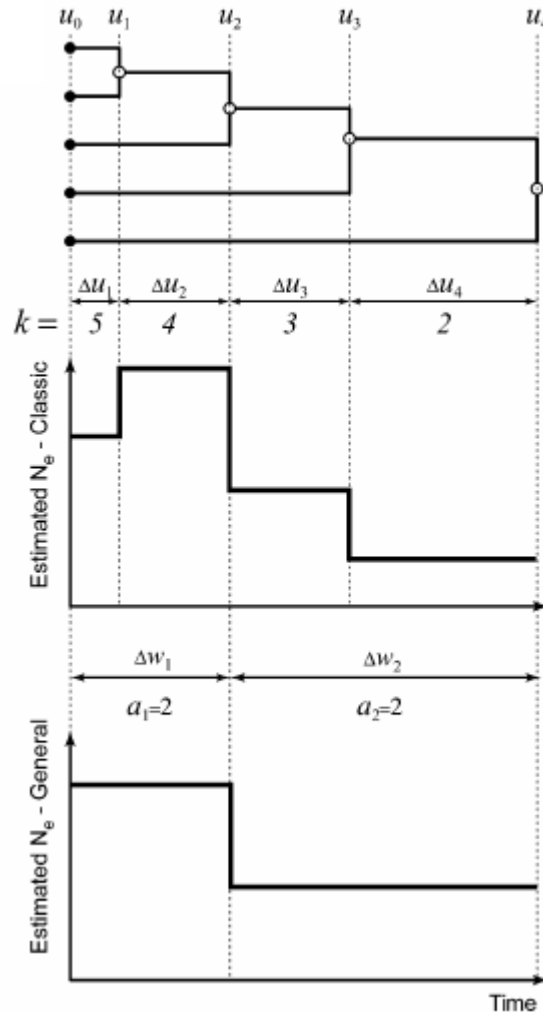
### Dinamiche di popolazione nel tempo

La stima della dimensione effettiva di popolazione, come pure il tasso di cambiamento nel tempo, possono fornire delle utili informazioni sulla storia evolutiva e demografica di una popolazione. Molti dei metodi sviluppati a questo scopo si basano sulla teoria della coalescenza (Hudson, 1990; Nordborg, 2001): tipicamente, sono basati su un modello parametrico  $N(t)$  che descrive la dimensione effettiva di popolazione nel tempo. Questo è pari a zero per il presente e aumenta a mano a mano che si va indietro nel tempo, quindi  $N(0)$  rappresenta la dimensione effettiva di popolazione attuale. I modelli demografici più frequentemente utilizzati sono due: un modello in cui la dimensione di popolazione è costante nel tempo,  $N(t)=N(0)$ ; e un modello di crescita esponenziale,  $N(t)=N(0)e^{rt}$ . A questi, si sono quindi affiancati dei modelli non-parametrici, ricoprendo così un importante ruolo nel trarre informazioni sulla storia di una popolazione a partire da dati di sequenza.

Nee e collaboratori (1995) hanno proposto il *lineare through time plot (LTT)* per indagare graficamente la storia demografica di sequenze geniche: questo modello mostra il tasso di coalescenza nel tempo in una genealogia ricostruita a partire dall'allineamento di sequenze omologhe. Successivamente, Pybus, Rambaut e Harvey (2000) hanno convertito il tasso di coalescenza in un plot in cui viene stimata la dimensione effettiva di popolazione nel tempo, chiamato *classic skyline plot*. A questo è stato, poi, affiancato il modello *generalized skyline plot*, che ammette che solo alcuni eventi di coalescenza siano stati accompagnati da una variazione della dimensione effettiva di popolazione.

Per investigare, quindi, la storia dell'espansione demografica dei campioni di popolazione di *C. aestuarii*, è stato utilizzato l'approccio bayesiano dello *skyline plot* (Drummond *et al.*, 2005) implementato nel programma *BEAST* 1.4.8 (Drummond and Rambaut, 2007). Questo metodo utilizza procedure di campionamento basate sull'algoritmo denominato *Markov Chain Monte Carlo* (MCMC) per stimare la dimensione effettiva di popolazione nel tempo, a partire da un modello di sostituzione nucleotidica specificato (Drummond *et al.*, 2002). Il dato di partenza è rappresentato da un albero filogenetico, stimato a partire da  $n$  sequenze, e la cui lunghezza dei rami sia stata fissata; i nodi interni devono essere datati, in accordo con la scala temporale stabilita (distanza genetica o tempo) e definiti come  $n-1$  volte in cui si verifica un evento di coalescenza  $\mathbf{u} = \{u_1, u_2, \dots, u_{n-1}\}$ . Gli intervalli di tempo tra eventi coalescenti sono definiti come  $\Delta u_i = u_i - u_{i-1}$ .

$i$ ; il numero di discendenti presenti durante ogni  $\Delta u_i$  definisce una serie corrispondente di valori, indicati con  $k_1, k_2, \dots, k_{n-1}$ . Questo modello suppone che la dimensione effettiva di popolazione sia costante tra eventi coalescenti ma che possa cambiare al momento di un evento  $u$  (figura 2.3, Drummond *et al.*, 2005). Il modello classico assume che ogni intervallo di coalescenza abbia una differente dimensione effettiva di popolazione, mentre quello generalizzato permette di raggruppare intervalli di coalescenza adiacenti e, inoltre, assume la stessa dimensione effettiva di popolazione.



**Figura 2.3:** genealogia di cinque individui campionati contemporaneamente associata allo *skyline plot* classico (grafico centrale) e generalizzato (grafico in basso). Nel modello classico di *skyline plot*, i cambiamenti nella dimensione effettiva di popolazione coincidono con eventi di coalescenza; in quello generalizzato, i cambiamenti coincidono con alcuni, ma non necessariamente tutti, eventi di coalescenza (da Drummond *et al.*, 2005).

In questa tesi è stato utilizzato il modello HKY+G (Hasegawa-Kishino-Yano e il modello di sostituzione G) come il modello più appropriato di evoluzione di sequenza suggerito dal *software ModelTest 3.7* (Posada and Crandall, 1998).

Ogni subcampione del dataset è stato analizzato mediante l'algoritmo MCMC per dieci milioni di step sotto i modelli HKY+G e stepwise skyline con cinque separati intervalli di tempo. Il tasso di mutazione è stato fissato a 0.01058 ed è stato selezionato un modello esatto di orologio molecolare. I risultati delle corse sono stati in seguito visualizzati attraverso il programma *Tracer* 1.4 (Rambaut and Drummond, 2007) per verificare che fossero arrivati a convergenza.

### 2.4.2.2 Marcatori microsatellite

I risultati dell'analisi dei frammenti sono stati analizzati mediante il *software Genotyper* 3.7 (Applied Biosystems) che fornisce per ogni *locus* la dimensione e l'intensità del frammento di DNA amplificato di ogni individuo. Una volta selezionati i picchi corrispondenti agli alleli, i profili sono stati esportati come *file* di testo, importati in *Excel* e infine sono stati ricostruiti i genotipi mediante la macro *Flexibin* (Amos *et al.*, 2007).

Il programma *Genepop* 3.4 (Rousset and Raymond, 1995) è stato utilizzato per verificare la frequenza allelica, l'equilibrio di Hardy-Weinberg (HWE), il *linkage disequilibrium* e il differenziamento genico. La variabilità genetica e gli indici  $F_{st}$  sono stati stimati tramite il programma *Genetix* (Belkir *et al.*, 2001), mentre l'*AMOVA* e l'isolamento per distanza sono stati testati statisticamente attraverso il *software Arlequin* 3.1.

#### Frequenze alleliche

La frequenza con cui in una popolazione si manifesta un determinato allele in rapporto a quella complessiva viene definita frequenza allelica. Solitamente utilizzata per descrivere un *pool* genico, viene calcolata attraverso il numero di genotipi differenti osservati. La frequenza  $p_i$  dell'allele *i*-esimo di un qualsiasi *locus* è calcolata mediante la formula:

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

dove  $n_1$  rappresenta il numero di omozigoti per l'allele *i*-esimo;  $n_2$  il numero di eterozigoti per l'allele *i*-esimo ed  $N$  il numero di individui.

Nota la frequenza allelica è possibile valutare se un *locus* è polimorfo o monomorfo ricorrendo a due criteri: quello dell'1% che consente di classificare come polimorfo quel particolare *locus* che presenta frequenza inferiore allo 0.99 per l'allele più comune; quello del 5% in cui la soglia della frequenza genica dell'allele più comune viene fissata allo 0.95.

Comunque il criterio dell'1% è ritenuto affidabile quando i campioni hanno una numerosità superiore a 50 (Nei, 1987).

#### Eterozigosità attesa ( $H_e$ ) ed eterozigosità osservata ( $H_o$ )

Il più utilizzato descrittore della variabilità genetica presente all'interno delle popolazioni è l'eterozigosità attesa ( $H_e$ ). Per un *locus*, essa è la frazione di genotipi eterozigoti attesa in base all'equilibrio di Hardy-Weinberg ed è data dall'equazione:

$$He = \frac{\sum_{j=1}^n \left( 1 - \sum_{i=1}^m p_i^2 \right)_j}{n}$$

dove  $p_i$  è la frequenza dell'allele  $i$ -esimo;  $m$  il numero totale di alleli per il  $j$ -esimo *locus* ed  $n$  il numero di *loci* esaminati.

L'eterozigosità osservata è la frequenza di individui eterozigoti per un determinato *locus* osservato nei campioni. Per ottenere l'eterozigosità osservata di una popolazione si esegue il seguente calcolo:

$$Ho = \frac{\sum_{j=1}^n \left( \frac{E}{E+O} \right)_j}{n}$$

dove  $E$  rappresenta il numero di eterozigoti osservati per il  $j$ -esimo *locus*;  $O$  il numero di omozigoti osservati per il  $j$ -esimo *locus* ed  $n$  il numero di *loci* analizzati.

Sia la proporzione di *loci* polimorfici che l'eterozigosità attesa rappresentano due parametri in grado di dare una stima dell'entità della variabilità genetica presente all'interno di una popolazione.

### Equilibrio di Hardy-Weinberg

La legge di Hardy-Weinberg descrive cosa accade alle frequenze alleliche e genotipiche di una vasta popolazione quando i gameti si incontrano in maniera casuale e in assenza di mutazione, migrazione e selezione naturale. Se queste condizioni sono soddisfatte, la popolazione è definita in equilibrio, le frequenze alleliche non variano di generazione in generazione e le frequenze genotipiche per un *locus* autosomico con gli alleli  $A_1, A_2, \dots, A_m$  si stabilizzano dopo una generazione secondo queste proporzioni:

$$\text{frequenza omozigoti} = \sum p_i^2 A_j A_j$$

$$\text{frequenza eterozigoti} = \sum 2p_i^2 A_j A_j$$

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

Per la verifica dell'equilibrio di Hardy-Weinberg (HWE) è stato eseguito un test esatto di probabilità (Guo and Thompson, 1992) che utilizza un approccio permutazionale e si avvale del metodo delle catene di Markov: la probabilità di accettare l'ipotesi  $H_0$  (unione casuale dei gameti) corrisponde alla somma cumulata delle probabilità di tutte le tabelle aventi lo stesso numero di alleli e con una differenza tra le medie delle distribuzioni minore o uguale a quella osservata.

L'ipotesi di equilibrio HWE viene rigettata valutando il grado di deviazione tra frequenze genotipiche attese ed osservate mediante il test statistico del  $\chi^2$ . Nel

caso dei microsattelliti, la difficoltà di applicare questo tipo di test è che ci sono troppi genotipi possibili nel caso di un gran numero di alleli per *locus*, che si traducono in un numero elevato di celle vuote.

Secondo HWE, gli alleli sono distribuiti indipendentemente all'interno dei possibili genotipi, così che il probabile genotipo trovato per permutazione degli  $A_1, A_2, \dots, A_m$  alleli presenti nel campione, corrisponde ad uno di quelli possibili sotto HWE. Si conduce perciò un elevato numero di permutazioni alleliche (10000), che permettono di stimare la probabilità di ottenere l'eterozigosità osservata rispetto a quella possibile in caso di equilibrio. Al test è assegnato un livello di significatività  $\alpha$  ( $\alpha = 0.05$ ), corrispondente alla probabilità massima di errore che si è disposti ad accettare nel rigettare l'ipotesi  $H_0$  (Waples and Gaggiotti, 2006).

### Linkage disequilibrium

Quando all'interno di una popolazione l'accoppiamento è casuale e gli alleli di un qualsiasi *locus*, se in condizione di indipendenza tra loro, raggiungono rapidamente l'associazione casuale nei genotipi, allora i geni si dicono in *linkage equilibrium*. Se questa situazione non si verifica, i geni vengono detti in *linkage disequilibrium*. *Genepop* si avvale di tabelle di contingenza corrispondenti a tutte le possibili coppie di *loci* in ciascuna popolazione e, partendo dall'ipotesi zero che gli alleli di un *locus* siano indipendenti dagli alleli di un altro *locus*, esegue un test di probabilità (o test esatto di Fisher) per ciascuna tavola usando le catene di Markov e stima così l'esatto *p-value*.

### Differenziamento di popolazioni: test esatto di Fisher e indici $F_{st}$

Per l'analisi di differenziamento di popolazioni il programma *Genepop* 3.4 fornisce sia una tabella contenente il numero di alleli per quel *locus* in tutte le popolazioni che una probabilità globale di differenziamento. Il programma utilizza il test esatto di Fisher per tabelle di contingenza, applicando un metodo di simulazione che si basa sulle catene di Markov per ottenere delle stime non distorte della probabilità di omogeneità. Per ogni *locus*, si analizza una tabella di contingenza che presenta su ogni riga una popolazione e su ogni colonna il numero di volte in cui un allele compare nella popolazione. L'ipotesi nulla  $H_0$  che viene testata corrisponde all'assenza di differenziamento tra popolazioni; in altri termini, il test verifica l'ipotesi di distribuzione allelica identica per tutti i campioni (Rousset and Raymond, 1995). I valori di probabilità, ottenuti per ogni *locus*, sono combinati in un valore complessivo mediante il metodo di Fisher (Manly, 1984). Inoltre è possibile creare una tabella di contingenza che effettua i confronti a coppie di popolazioni per ogni *locus* e permette di verificare a coppie il differenziamento tra popolazioni.

Anche per i marcatori microsatellite sono stati stimati gli indici  $F_{st}$  e l'*AMOVA*, già descritti nella sezione 2.4.2.1 (analisi dati, parte DNA mitocondriale pag. 15).

### Isolamento per distanza e Mantel test

L'isolamento per distanza (Wright, 1938, 1940) è stato utilizzato per la prima volta per descrivere i pattern di variazione genetica di una popolazione che derivano da un limitato flusso genico a livello spaziale; è definito come riduzione nella similarità genetica tra popolazioni in rapporto all'incremento della loro distanza geografica.

Attraverso il programma *Arlequin* 3.1 è stato testato statisticamente l'isolamento per distanza utilizzando il test di Mantel, mettendo in relazione la distanza genetica (rappresenta dagli indici  $F_{st}$ ) con la distanza geografica (espressa in km) in modo da effettuare una regressione lineare e verificare se vi fosse correlazione tra le matrici di distanze genetiche e le matrici di distanze geografiche.

### Stima dei tassi di migrazione

Stime di maximum likelihood, basate sull'algoritmo *Markov Chain Monte Carlo* (MCMC), della dimensione effettiva di popolazione ( $N_e$ ) e dei tassi di migrazione ( $m$ ) sono stati ottenuti utilizzando il programma *Migrate-n* 2.3 (Beerli and Felsenstein, 2001), che stima simultaneamente i parametri  $\theta_i$  ( $\theta = 4N_e\mu$ ) e  $M_{j \rightarrow i}$  ( $M = m/\mu$ ) esplorando tutte le possibili genealogie per un set di campioni di popolazione, posto che le popolazioni siano in equilibrio tra mutazione e deriva e che abbiano mantenuto una dimensione costante nel tempo (Beerli and Felsenstein, 1999).

Sono stati testati due specifici modelli di struttura di popolazione, un modello a isole (ipotesi nulla: migrazione tra tutte le popolazioni, indipendentemente dalla distanza geografica, permettendo un flusso asimmetrico tra le stesse) e un modello *stepping stone* modificato (ipotesi nulla: migrazione limitata a popolazioni adiacenti). Per entrambi i modelli, i valori iniziali di theta e di migrazione sono stati generati a partire dal calcolo degli indici  $F_{st}$ ; sono state fatte correre 10 catene corte con  $10^4$  generazioni ciascuna seguite da 3 catene lunghe con  $10^5$  generazioni. Infine le stime di maximum-likelihood di due o più corse indipendenti sono state confrontate per entrambi i modelli testati.

### ***3. RISULTATI***



### 3.1 Estrazione del DNA genomico

Il protocollo di estrazione *Salting out* (Patwary *et al.*, 1994) si è dimostrato affidabile e ha permesso di estrarre con successo il DNA da tutti gli 870 individui provenienti dagli 8 siti di campionamento. Questo protocollo ha fornito DNA amplificabile sia per il *locus* mitocondriale che per i *loci* microsatellite. Le analisi su gel di agarosio all'1% dei prodotti di estrazione hanno mostrato una banda ad alto peso molecolare (D), costituita da DNA genomico, ed uno *smear* (S), costituito da RNA e DNA degradati. (figura 3.1).

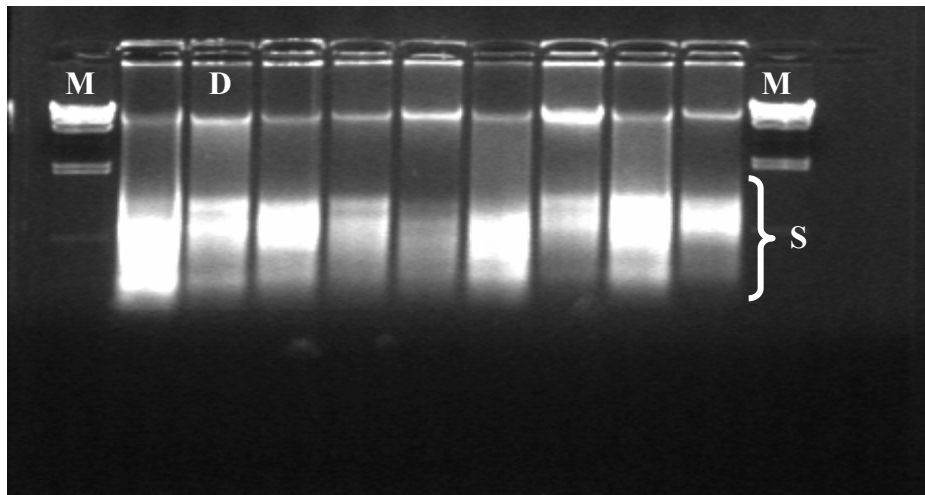


Figura 3.1: elettroforesi su gel di agarosio all'1% dei prodotti di estrazione del DNA di 9 individui di *Aquatina* (Le). Con "D" è contrassegnato il DNA estratto, con "S" lo *smear*, mentre con "M" il marcatore di peso molecolare ( $\lambda$  *HindIII*, Promega).

Attraverso la procedura di restrizione con enzima *AluI* è stato verificato se la specie in esame fosse effettivamente *Carcinus aestuarii*. L'elettroforesi su gel di agarosio all'1.8% (figura 3.2) dei prodotti di restrizione ha confermato che tutti gli 870 individui campionati appartengono a suddetta specie, in quanto tutti i frammenti sono stati tagliati in due dall'enzima di restrizione.

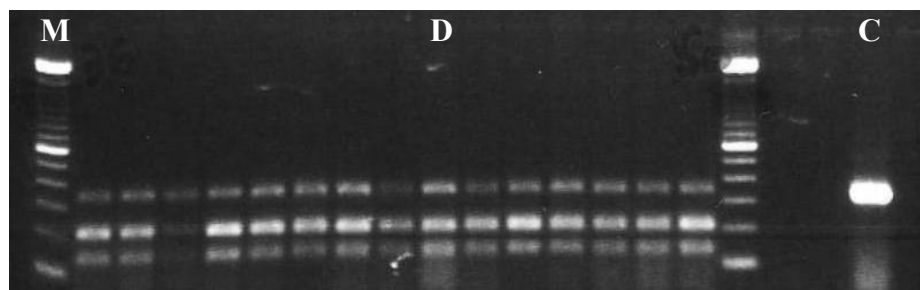


Figura 3.2: elettroforesi su gel di agarosio all'1.8% dei prodotti di restrizione con enzima *AluI* di individui provenienti da Cagliari (Cag). Con la "D" sono contrassegnati i frammenti tagliati in due bande dall'enzima, con la "C" il controllo negativo, mentre con la "M" il marcatore di peso molecolare (100bp, Promega).

### 3.2 Amplificazione del DNA mitocondriale

Per quanto riguarda l'amplificazione del DNA mitocondriale, per ognuno dei 32 individui selezionati dagli 8 campioni di popolazione è stato amplificato con successo il frammento di 482 paia di basi del gene mitocondriale della citocromo c ossidasi I e l'esito della PCR è stato visualizzato mediante elettroforesi su gel di agarosio all'1.8% (figura 3.3).

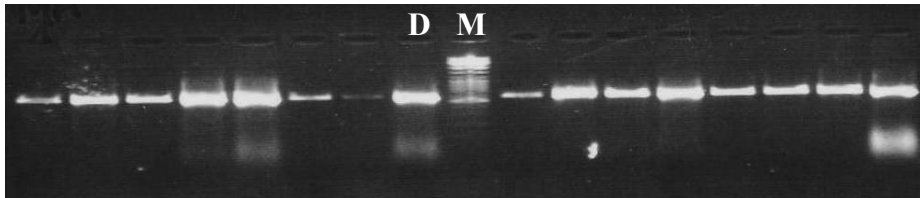


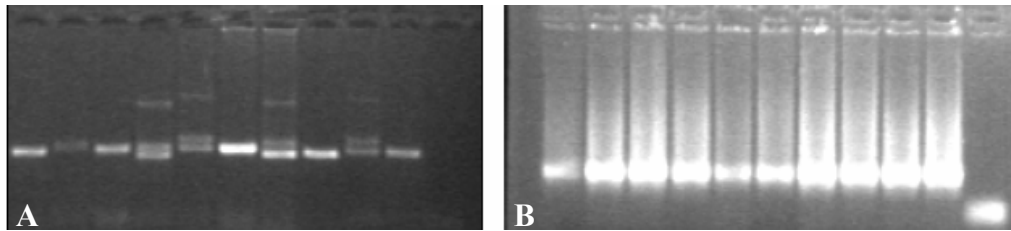
Figura 3.3: elettroforesi su gel di agarosio all'1.8% del DNA mitocondriale amplificato di individui provenienti da Marano (Ma). Con la lettera D si contrassegna il pozzetto contenente il DNA amplificato, con la lettera M il marcatore di peso molecolare (100bp, Promega).

### 3.3 Isolamento e messa a punto di marcatori microsatellite

Il protocollo di isolamento dei marcatori nucleari che è stato adottato ha fornito materiale sufficiente ad individuare 8 *loci* polimorfici. Dalle colonie prelevate dalla libreria plasmidica e sottoposte ad amplificazione, 107 cloni sono stati sequenziati previa purificazione. Di questi, 61 contenevano effettivamente il microsatellite; tra le sequenze contenenti regioni ripetute, ne sono state selezionate 30, aventi sufficiente sequenza fiancheggiante per disegnare una coppia di *primer* per l'amplificazione del *locus* da DNA genomico.

Ogni coppia di *primer* è stata poi testata su alcuni individui ( $n = 10$ ) di *Carcinus aestuarii*. Per il raggiungimento delle condizioni ottimali di PCR sono state eseguite alcune prove di amplificazione con profili termici diversi, determinati prendendo in considerazione le caratteristiche degli oligonucleotidi disegnati. I prodotti di PCR sono stati quindi lasciati correre su gel di agarosio all'1.8% per verificarne l'effettiva amplificazione.

Delle 30 coppie di *primer* disegnate, solo 8 sono risultate funzionanti e polimorfe, e quindi impiegabili per lo studio di genetica di popolazione; le restanti 22 sono state scartate o perché monomorfe o perché dimostravano effettive difficoltà di amplificazione (figura 3.4).



**Figura 3.4:** elettroforesi su gel di agarosio all'1.8%. La foto A mostra come si presenta alla lettura UV un *locus* polimorfo (Cae71), mentre la foto B mostra come si presenta alla lettura UV un *locus* monomorfo.

Durante la fase di messa a punto sono state raggiunte le condizioni ottimali di amplificazione anche per le 3 coppie di *primer* (Cma02EPA, Cma04EPA e Cma14EPA) provenienti da studi condotti su *C. maenas* (Tepolt *et al.*, 2006), allo stesso modo delle precedenti.

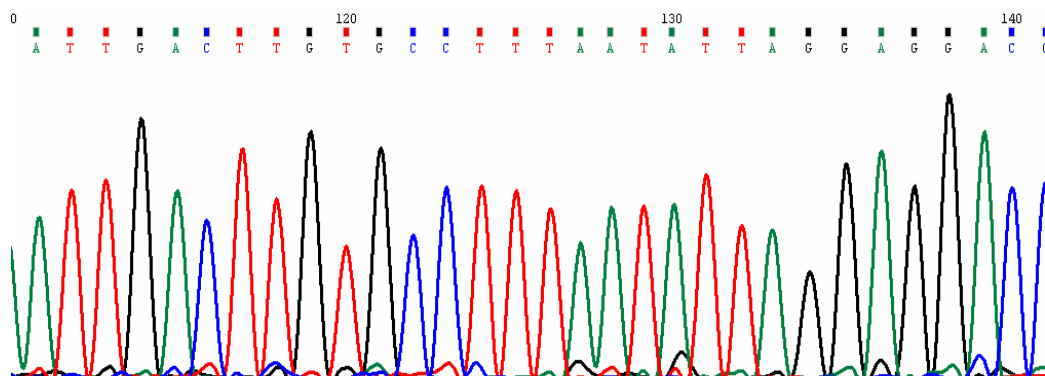
Per ogni coppia di *primer*, inoltre, è stato marcato un oligonucleotide con un fluoroforo appropriato che ne ha permesso la visualizzazione su un sequenziatore automatico ABI.

Inizialmente, sono stati genotipizzati pochi individui (n=16) per ottimizzare le condizioni di corsa al sequenziatore, verificare la qualità dei singoli *loci* microsatellite e testare quali tra questi presentassero, ad una prima analisi, alleli differenti chiaramente risolvibili e quindi adatti agli obiettivi di questa tesi.

## 3.4 Analisi dati

### 3.4.1 Analisi del DNA mitocondriale

Complessivamente sono stati sequenziati 255 individui, i cui cromatogrammi sono stati visualizzati con il *software* CHROMAS 2.12 (Technelysium Pty Ltd., Helensvale, Australia; reperibile sul sito <http://www.technelysium.com.au/chromas.html>) (figura 3.5). Delle 482 paia di basi, 70 sono siti variabili dei quali 48 sono siti informativi per la parsimonia. La composizione in basi è: A 26.3%, T 37.6%, C 18% e G 18.2%.

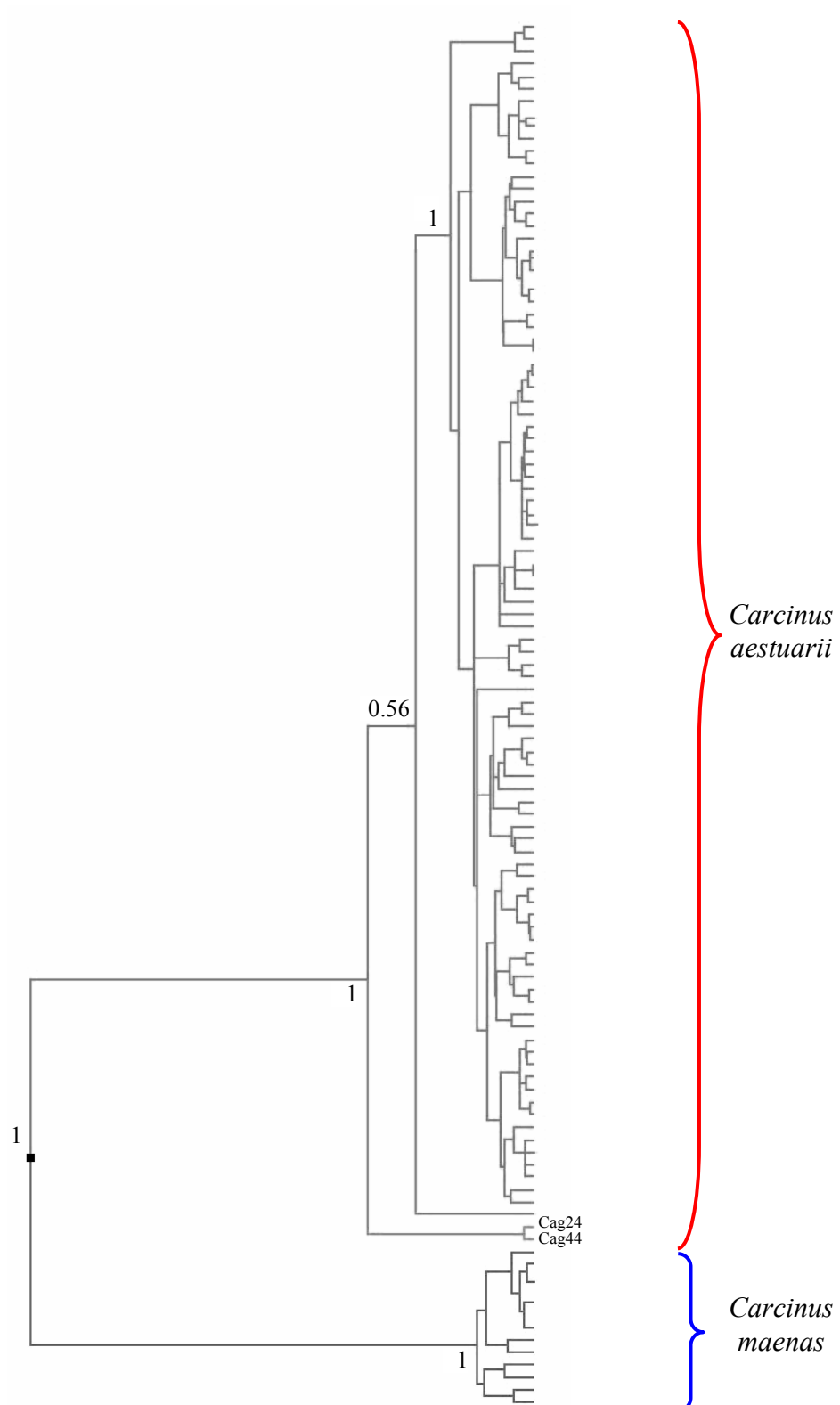


**Figura 3.5:** cromatogramma dell'individuo 29 proveniente dal sito di Cagliari (Cag). Ogni picco rappresenta una base che è visualizzata nella riga in alto e contrassegnata da un diverso colore (ATCG).

Attraverso *ModelTest* 3.7 è stato, innanzitutto, selezionato il modello di sostituzione più appropriato per i dati di sequenza in esame, il modello HKY+G. Tale modello è stato utilizzato per ricostruire, attraverso il software *PAUP\** 4.0 (Swofford, 1998), un albero *Maximum Likelihood* assumendo o meno un *molecular clock*. Si è quindi eseguito un *LRT (Likelihood Ratio Test)* per calcolare la probabilità associata alle diverse ipotesi (presenza o assenza di orologio molecolare): in questo caso, è stata verificata la presenza di orologio molecolare. Il *molecular clock* è stato così calibrato a  $5.3 \cdot 10^6$  anni fa, momento in cui si ritiene sia avvenuta la separazione delle specie *Carcinus maenas* e *C. aestuarii*. Sulla base di questo risultato, attraverso il programma *BEAST*, è stato costruito un albero consenso (figura 3.6), il più probabile a partire dai dati osservati e supportato dai valori di *posterior probability*, mediante il quale è stato possibile stimare il tasso di sostituzione di  $1.058 \cdot 10^{-2}$  mutazioni per *locus* per anno.

Dall'albero è possibile osservare la presenza di due cluster ben differenziati e supportati dai valori di *posterior probability*, corrispondenti alle due specie di *Carcinus*. Nel cluster corrispondente alla specie *C. aestuarii*, si può notare l'elevata variabilità degli aplotipi analizzati, tuttavia, due singole varianti di sequenza (Cag24 e Cag44) si collocano in posizione basale rispetto al resto del cluster, collocandosi esattamente tra i due gruppi di specie. Questi due aplotipi, infatti, pur appartenendo alla specie *C. aestuarii* (verificato tramite procedura di restrizione con enzima *AluI*), differiscono di molto, a livello di sequenza, dal resto degli aplotipi analizzati.

Complessivamente sono state trovate 164 diverse varianti di sequenza (aplotipi). Il numero di aplotipi per popolazione (tabella 3.1) è risultato essere elevato, variando da 13 per Venezia (Ve) a 29 per Goro (Go) su un campione di 32 individui per sito (fatta eccezione per il campione di Ravenna (Ra) per il quale sono stati analizzati 31 individui). Il valore di nucleotide diversity ( $\pi$ ) varia da 0.003760 per il campione di Aquatina (Le) a 0.006981 per il campione di Goro (Go) e mediamente risulta pari a 0.004992 a significare che due sequenze differiscono mediamente per lo 0.4%. Il valore di haplotype diversity ( $h$ ), invece, varia da 0.6431 per il campione di Cagliari (Cag) a 0.9879 per il campione di Goro (Go) e risulta essere in media di 0.87: questo significa che due sequenze estratte casualmente dalla popolazione totale differiscono nell'87% dei casi.



**Figura 3.6:** *Citocromo c ossidasi I*. Albero consenso ottenuto mediante analisi bayesiana. Sono riportate le *posterior probability* dei nodi principali.

**Tabella 3.1: indici di diversità. Sono riportati i siti di campionamento, il numero di individui analizzati per popolazione, il numero di aplotipi, i valori di nucleotide diversity ( $\pi$ ) e di haplotype diversity ( $h$ ) con i rispettivi errori standard.**

| Siti       | N° individui | N° aplotipi | $\pi$                 | $h$               |
|------------|--------------|-------------|-----------------------|-------------------|
| <b>Cag</b> | 32           | 24          | 0.005860 +/- 0.003530 | 0.6431 +/- 0.0942 |
| <b>Ta</b>  | 32           | 16          | 0.003848 +/- 0.002519 | 0.8246 +/- 0.0624 |
| <b>Le</b>  | 32           | 17          | 0.003760 +/- 0.002474 | 0.8690 +/- 0.0502 |
| <b>Fg</b>  | 32           | 20          | 0.005049 +/- 0.003125 | 0.9113 +/- 0.0399 |
| <b>Ra</b>  | 31           | 22          | 0.005024 +/- 0.003116 | 0.9376 +/- 0.0371 |
| <b>Go</b>  | 32           | 29          | 0.006981 +/- 0.004087 | 0.9879 +/- 0.0115 |
| <b>Ve</b>  | 32           | 13          | 0.003898 +/- 0.002545 | 0.9093 +/- 0.0334 |
| <b>Ma</b>  | 32           | 23          | 0.005521 +/- 0.003361 | 0.8972 +/- 0.0457 |

Il software *Arlequin* ha permesso di calcolare gli indici  $F_{st}$ , corrispondenti alla distanza genetica calcolata tra le singole coppie di popolazioni, e i valori ottenuti sono riportati in tabella 3.2. Il sito di Cagliari presenta delle differenze significative nei confronti di tutti gli altri campioni: infatti è possibile osservare come i valori più elevati di  $F_{st}$  siano proprio quelli associati al campione cagliaritano. Per quanto riguarda i campioni adriatici e ionico, non si riscontra alcun differenziamento, con valori di  $F_{st}$  negativi (ad indicare una completa omogeneità) e non significativi.

**Tabella 3.2: indici  $F_{st}$ . Sono riportati sotto la diagonale gli indici  $F_{st}$  calcolati a coppie e sopra la diagonale i corrispondenti valori di significatività.**

|            | <b>Cag</b> | <b>Ta</b> | <b>Le</b> | <b>Fg</b> | <b>Ra</b> | <b>Go</b> | <b>Ve</b> | <b>Ma</b> |
|------------|------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| <b>Cag</b> |            | 0.00586   | 0.06934   | 0.01953   | 0.04590   | 0.00488   | 0.01562   | 0.04395   |
| <b>Ta</b>  | 0.04471    |           | 0.35742   | 0.49707   | 0.29688   | 0.27344   | 0.45703   | 0.41699   |
| <b>Le</b>  | 0.02216    | 0.00230   |           | 0.35156   | 0.12695   | 0.02832   | 0.31836   | 0.48828   |
| <b>Fg</b>  | 0.03091    | -0.00165  | 0.00278   |           | 0.86523   | 0.72852   | 0.67578   | 0.98535   |
| <b>Ra</b>  | 0.02607    | 0.00434   | 0.01190   | -0.01137  |           | 0.64551   | 0.23047   | 0.88379   |
| <b>Go</b>  | 0.04518    | 0.00465   | 0.02399   | -0.00773  | -0.00649  |           | 0.09180   | 0.74316   |
| <b>Ve</b>  | 0.03670    | -0.00196  | 0.00418   | -0.00638  | 0.00617   | 0.01652   |           | 0.78125   |
| <b>Ma</b>  | 0.02439    | -0.00098  | -0.00113  | -0.01736  | -0.01292  | -0.00765  | -0.00984  |           |

Si è proceduto, quindi, all'analisi in dettaglio del differenziamento tra popolazioni, attraverso l'analisi molecolare della varianza (*AMOVA*), i cui risultati sono riportati in tabella 3.3. Le popolazioni analizzate sono state suddivise in due gruppi: un primo gruppo composto dalle popolazioni dell'Adriatico e dello Ionio

(Ma, Ve, Go, Ra, Fg, Le, Ta) e un secondo gruppo composto solo dalla popolazione di Cagliari (Cag).

**Tabella 3.3: analisi molecolare della varianza (AMOVA). La significatività è stata stimata mediante 10000 permutazioni.**

| Sorgente di variazione              | d.f. | Somma dei quadrati | Componenti della varianza | Percentuale della variazione | <i>p-value</i> |
|-------------------------------------|------|--------------------|---------------------------|------------------------------|----------------|
| <b>Tra gruppi</b>                   | 1    | 3.884              | 0.04881                   | 3.90                         | <0.00001       |
| <b>Tra popolazioni entro gruppi</b> | 6    | 6.911              | -0.00161                  | -0.13                        | 0.49604        |
| <b>Entro popolazioni</b>            | 247  | 297.198            | 1.20323                   | 96.23                        |                |
| <b>Totale</b>                       | 254  | 307.922            | 1.25043                   |                              |                |

Tale suddivisione deriva da vari di tentativi di raggruppamento (si è cercato, ad esempio, di raggruppare assieme Taranto e Cagliari), ma questa è risultata la migliore suddivisione che massimizza la percentuale di variazione.

La maggior parte della varianza (96.23%), che spiega buona parte della variabilità, è stata trovata entro popolazioni; tuttavia, è riscontrabile una quota di variabilità (3.90%) derivante dalla suddivisione in gruppi. Questa quota di variabilità risulta essere significativa, avendo un valore di  $P < 0.00001$ . Un valore di variabilità negativo e non significativo è stato invece riscontrato tra popolazioni entro gruppi, a conferma dell'assenza di differenziamento tra i campioni adriatici e ionico: ciò sta ad indicare che individui provenienti da popolazioni diverse sono più simili tra loro che individui provenienti da una stessa popolazione.

Per confermare questo risultato, l'AMOVA è stata effettuata escludendo il campione di Cagliari dall'analisi: la percentuale di variazione tra popolazioni ha restituito, in accordo con i valori ottenuti dall'analisi complessiva, un valore negativo e non significativo. L'indagine tramite marcatore mitocondriale evidenzia, quindi, solo un differenziamento tra il campione tirrenico e le restanti popolazioni adriatiche e ionica, ma non mostra alcun differenziamento all'interno di queste ultime.

Un ulteriore livello di analisi dei dati, consentito da dati molecolari di sequenza, è consistito nell'analisi dei processi evolutivi alla base della variabilità molecolare a livello di specie. Infatti, dal confronto tra i valori osservati di alcune statistiche con i valori attesi da alcuni modelli teorici è possibile trarre indicazioni riguardo ai fenomeni evolutivi degli aplotipi in esame.

In tabella 3.4 sono riportati i risultati ottenuti con i test di neutralità selettiva di Tajima e Fu: questi test, nonostante siano stati introdotti allo scopo di confermare l'ipotesi di neutralità e la presenza di selezione, verificano anche l'ipotesi di equilibrio tra mutazione e deriva, come descritto in materiali e metodi. Valori negativi significativi di  $D$  e  $F_s$  possono essere interpretati come un'espansione di popolazione.

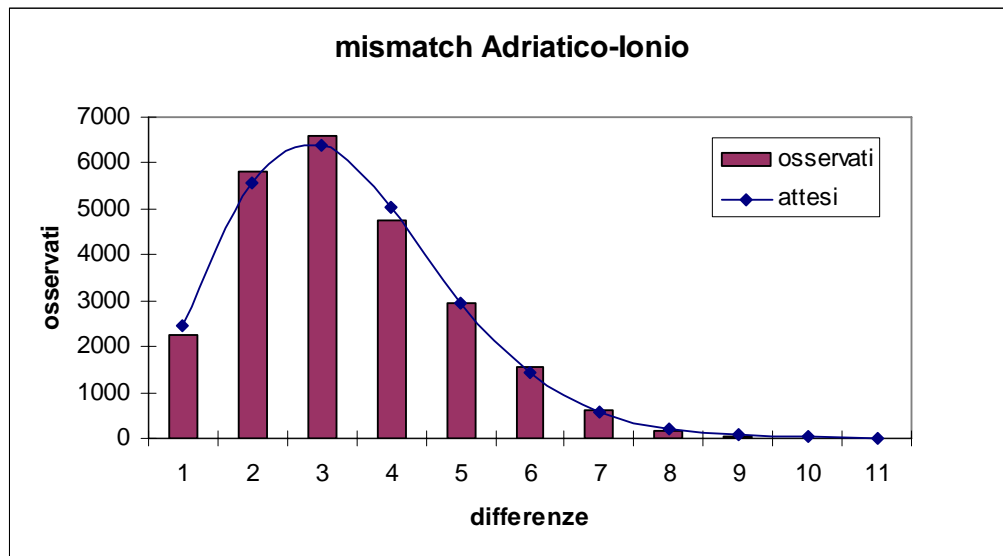
È possibile notare come le statistiche  $D$  e  $F_s$  siano negative per tutte le popolazioni analizzate. Il test  $D$  risulta significativo per tutte le popolazioni, tranne che per Venezia (Ve,  $D = -1.36173$ ,  $P = 0.086$ ), mentre il test  $F_s$  risulta statisticamente significativo per tutte le popolazioni, tranne che per Cagliari (Cag,

$F_s = -2.35731$ ,  $P=0.144$ ). Complessivamente questi test permettono di rigettare l'ipotesi di neutralità ed equilibrio per i campioni analizzati. Simulazioni al computer hanno dimostrato che valori negativi delle statistiche  $D$  di Tajima e  $F_s$  di Fu sono attribuibili a fenomeni demografici, quali una espansione di popolazione avvenuta nel passato (Tajima, 1983; Fu 1997).

**Tabella 3.4:** test di neutralità calcolati per ogni singola popolazione. Per il test di Tajima è riportata la statistica  $D$ , per il test di Fu è riportata la statistica  $F_s$ . Per ognuno dei test è riportata la probabilità associata ( $p$ -value) che corrisponde alla probabilità di osservare il valore riportato nel caso di neutralità ed equilibrio.

| Statistics              | Cag     | Ta      | Le       | Fg       | Ra       | Go       | Ve       | Ma       |
|-------------------------|---------|---------|----------|----------|----------|----------|----------|----------|
| Tajima's $D$            | -1.8438 | -1.7876 | -1.9278  | -1.7553  | -1.9590  | -1.898   | -1.3617  | -1.8643  |
| Tajima's $D$ $p$ -value | 0.013   | 0.024   | 0.014    | 0.016    | 0.011    | 0.015    | 0.086    | 0.011    |
| $F_s$                   | -2.3573 | -8.6540 | -10.5583 | -15.0580 | -25.0001 | -26.0968 | -12.0408 | -12.0852 |
| $F_s$ $p$ -value        | 0.144   | <0.001  | <0.001   | <0.001   | <0.001   | <0.001   | <0.001   | <0.001   |

A supporto di questa ipotesi si riporta in figura 3.7 la distribuzione osservata del numero di differenze per tutti i confronti a coppie tra le sequenze (*mismatch distribution*). Questa è stata testata a diversi livelli di raggruppamento: in una prima analisi sono state raggruppate le popolazioni dell'Adriatico e dello Ionio, escludendo l'unica popolazione del Tirreno (Cag). La distribuzione risulta avere una media di 2.345 differenze e una varianza di 2.353 (tabella 3.5). È possibile notare come le frequenze che differiscono per due e tre mutazioni siano predominanti con un andamento "a campana"; nella stessa figura è riportato l'andamento atteso secondo il modello di Rogers and Harpending (1992) per una popolazione in espansione e si può osservare come questo modello predica una curva con una distribuzione teorica simile a quella osservata.



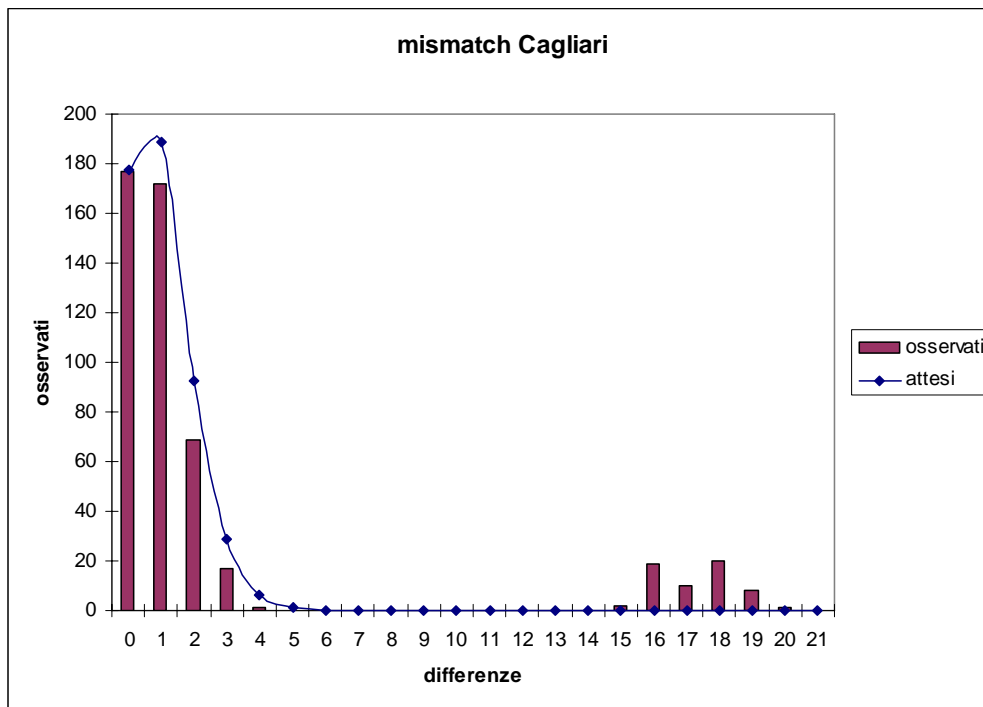
**Figura 3.7:** distribuzione osservata del numero di differenze per tutti i possibili confronti tra le sequenze a coppie. La linea continua rappresenta la curva prevista per una popolazione che ha subito un'espansione demografica nel passato.

I parametri su cui è basata la curva della distribuzione osservata sono riportati in tabella 3.5: il valore di  $\tau = 2.041$  (valore compreso tra 0.83 e 3.83) corrisponde, utilizzando un tasso di mutazione di  $5.1 \cdot 10^{-6}$  mutazioni per *locus* per anno (materiali e metodi, sezione 2.4.2.1 *Mismatch distribution*), ad una espansione avvenuta 200115 anni fa (valore medio compreso tra 375522 e 81379 anni fa in base al valore assunto da  $\tau$ ).

**Tabella 3.5:** parametri della *mismatch distribution* raggruppando le popolazioni dall'Adriatico e dello Ionio.

|                                     |       |
|-------------------------------------|-------|
| Media delle differenze osservate    | 2.345 |
| Varianza delle differenze osservate | 2.353 |
| $\tau$                              | 2.041 |
| $\theta_0$                          | 0.283 |
| $\theta_1$                          | 99999 |

In una seconda analisi è stata presa in esame solo la popolazione di Cagliari (Cag). In figura 3.8 viene riportata la distribuzione di tale popolazione, che risulta avere una media di 2.825 e una varianza di 29.559 (tabella 3.6). Nella stessa figura è riportato l'andamento atteso secondo il modello di Rogers e Harpending (1992) per una popolazione in espansione: in questo caso, però, la curva che descrive le differenze osservate risulta bimodale, caratteristica di una popolazione costante.



**Figura 3.8:** distribuzione osservata del numero di differenze per tutti i possibili confronti tra le sequenze a coppie. La linea continua rappresenta la curva prevista per una popolazione che ha subito una espansione demografica nel passato.

Tale distribuzione, in realtà è dovuta esclusivamente alla presenza di due aplotipi rari (Cag24 e Cag44) molto divergenti a livello di sequenza, che si collocano a livello basale nel cluster corrispondente alla specie *C. aestuarii* nell'albero consenso di figura 3.6.

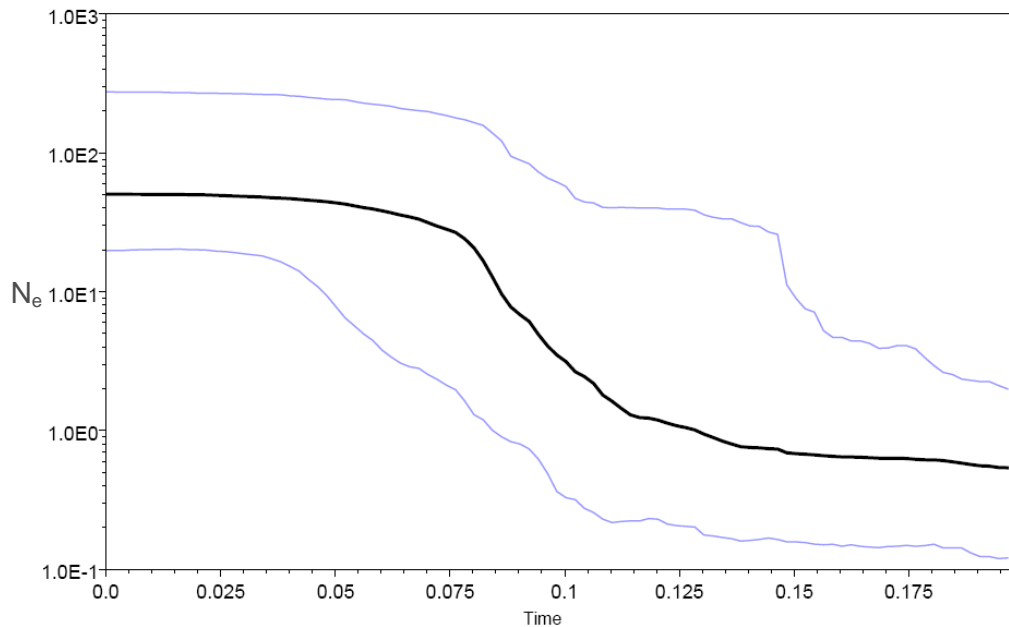
**Tabella 3.6:** parametri della *mismatch distribution* della popolazione di Cagliari.

|                                     |        |
|-------------------------------------|--------|
| Media delle differenze osservate    | 2.825  |
| Varianza delle differenze osservate | 29.559 |
| $\tau$                              | 0.957  |
| $\theta_0$                          | 0.000  |
| $\theta_1$                          | 99999  |

Per chiarire meglio il tipo di variazione demografica e i processi evolutivi subiti dai campioni adriatici e ionico, ma soprattutto dal campione tirrenico, è stato utilizzato il *software BEAST 1.4.8* (Drummond and Rambaut, 2007), che ha permesso di introdurre dei modelli demografici più realistici.

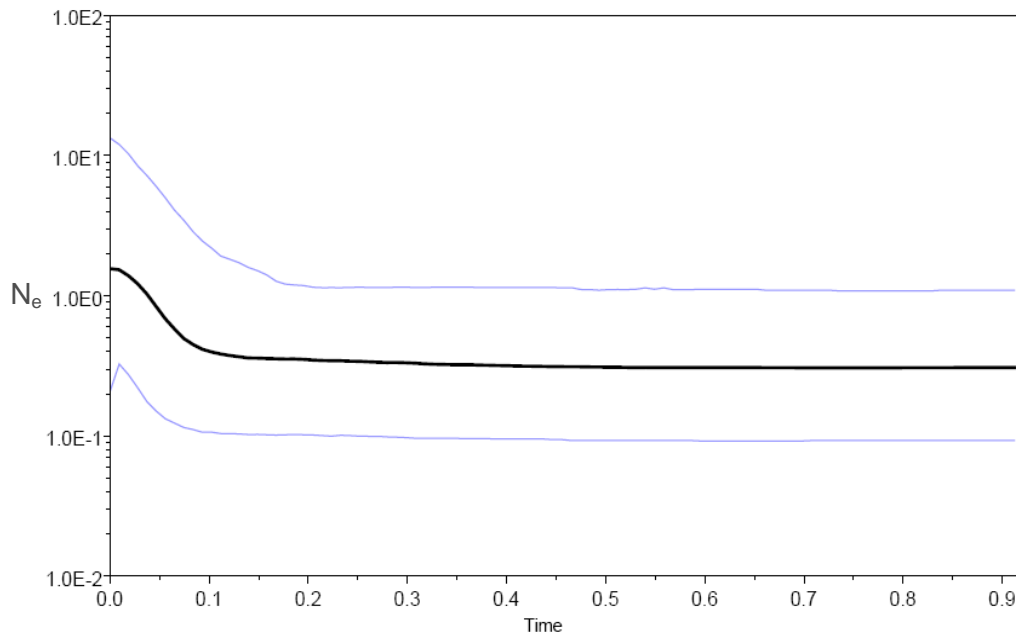
I *bayesian skyline plot*, ottenuti con *BEAST* per stimare i cambiamenti demografici nel tempo e la loro durata, hanno confermato l'espansione demografica subita dalle popolazioni adriatiche e ionica (figura 3.9), che si colloca approssimativamente in un intervallo di tempo (compreso tra 200000 e 100000 anni fa) che coincide con l'inizio dell'espansione stimato dalla *mismatch*

*distribution*. Se queste stime sono realistiche, anche considerando l'errore associato, un evento demografico con un forte impatto sulla diversità genetica di *C. aestuarii* dovrebbe essere avvenuto prima che il Mar Adriatico assumesse la morfologia che lo contraddistingue attualmente.



**Figura 3.9:** *bayesian skyline plot* calcolato per le popolazioni dell'Adriatico e dello Ionio mediante una corsa di 1000000 di step sotto i modelli HKY+G e *stepwise skyline* con 5 intervalli di tempo stabiliti a priori. In ascissa è riportato il tempo in milioni di anni, mentre in ordinata è riportata la dimensione di popolazione in scala logaritmica.

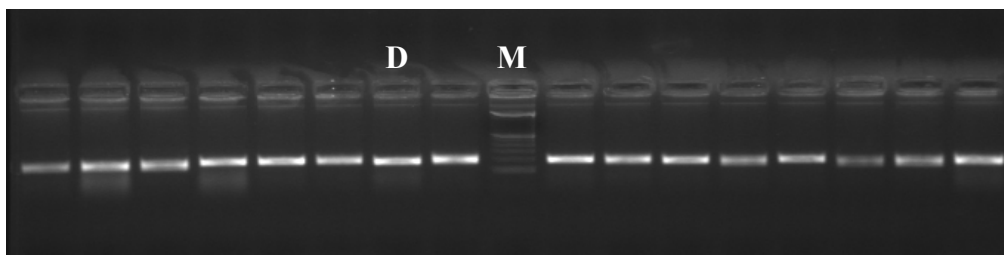
Per quanto riguarda il campione di Cagliari, invece, la situazione risulta diversa: infatti, dai *bayesian skyline plot* risulta che la popolazione è rimasta costante per un intervallo di tempo molto lungo (da 900000 a 100000 anni fa) e solo recentemente abbia subito un'espansione demografica (figura 3.10). Questo risultato collocherebbe, quindi, l'espansione subita dai campioni cagliaritari in un periodo di tempo che coincide con l'ultima glaciazione pleistocenica.



**Figura 3.10:** *bayesian skyline plot* calcolato per la popolazione di Cagliari mediante una corsa di 10000000 di step sotto i modelli HKY+G e stepwise skyline con 5 intervalli di tempo stabiliti a priori. In ascissa è riportato il tempo, mentre in ordinata è riportata la dimensione di popolazione in scala logaritmica.

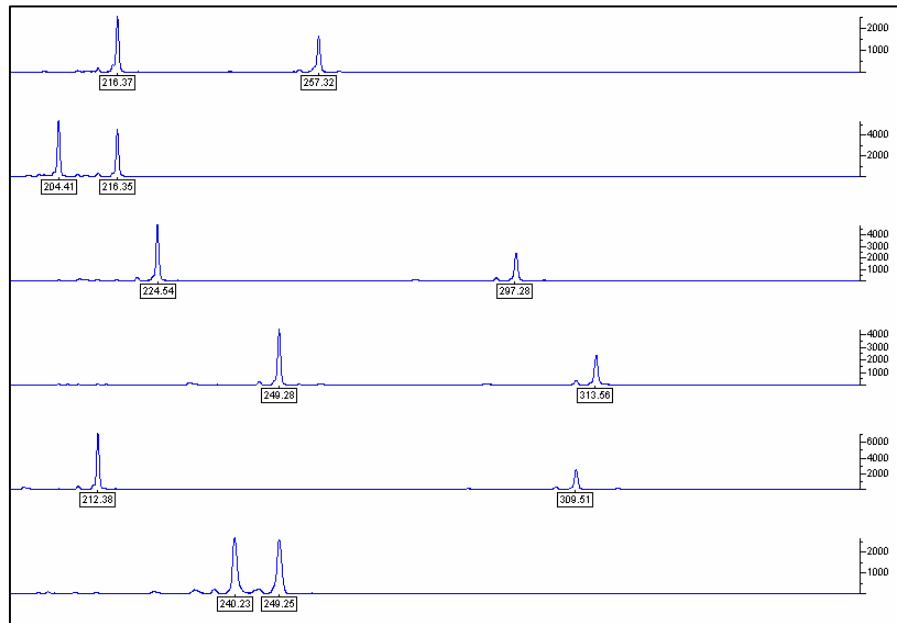
### 3.4.2 Analisi dei marcatori microsatellite

Le 11 coppie di *primer* così ottenute (8 disegnate per i *loci* specie-specifici di *Carcinus aestuarii*, 3 per i *loci* specie-specifici di *C. maenas*), sono state utilizzate per l'amplificazione di tutti gli 870 individui raccolti. Anche per i marcatori microsatellite, l'esito della PCR è stato visualizzato mediante elettroforesi su gel di agarosio all'1.8% (figura 3.11).



**Figura 3.11:** elettroforesi su gel di agarosio all'1.8% del DNA nucleare amplificato di individui prelevati a Ravenna (Ra) con il *locus* Cae01. Con la lettera D si contrassegna il pozzetto contenente il DNA amplificato, con la M il marcatore di peso molecolare (100bp, Promega).

I risultati dell'analisi dei frammenti sono stati analizzati con il *software Genotyper* 3.7 (Applied Biosystems) che ha fornito per ogni *locus* la dimensione e l'intensità del frammento di DNA amplificato di ogni individuo (figura 3.12).



**Figura 3.12: visualizzazione mediante *Genotyper* 3.7 (Applied Biosystems) di genotipi microsatellite di 6 individui per il *locus* CmaEPA02. Ogni riga rappresenta un individuo ed ogni picco rappresenta un allele. I valori riportati in ordinata rappresentano la scala di intensità di fluorescenza che dipende dalla concentrazione del DNA amplificato.**

L'analisi statistica è stata condotta su tutti gli 870 individui campionati e per tutti gli 11 *loci* microsatellite.

I genotipi ottenuti, sono stati analizzati attraverso l'ausilio del programma *Genepop* 3.4 (Rousset and Raymond, 1995) per verificare l'equilibrio di Hardy-Weinberg (HWE), il *linkage disequilibrium* e la frequenza allelica, e attraverso il programma *Genetix* (Belkir *et al.*, 2001) per calcolare la variabilità genetica (tabella 3.7 e 3.8).

I risultati del test di *linkage disequilibrium* hanno confermato che tutti gli 11 *loci* si trovano in condizioni di indipendenza, situazione che ha permesso di utilizzare i dati per quei test statistici che esigono tale assunzione.

Dall'analisi dei dati si è potuto riscontrare che tutti i campioni di popolazione sono caratterizzati da elevati livelli di variabilità: il numero medio di alleli va da un minimo di 16.7273 (per il campione di Taranto, Ta) ad un massimo di 28.8182 (per il campione di Venezia, Ve). Il livello di variabilità risulta piuttosto elevato: infatti l'eterozigosità media attesa varia da 0.7777 per il campione di Cagliari (Cag) a 0.8071 per il campione di Taranto (Ta). Per quanto riguarda, invece, l'eterozigosità media osservata, questa varia tra 0.7043 (per il campione di Lesina, Fg) e 0.7417 (per il campione di Taranto, Ta). Tutti i campioni di popolazione, inoltre, sono risultati in disequilibrio di Hardy-Weinberg, e ciò è da attribuire ad un eccesso di omozigoti.

Complessivamente tutti i *loci* si presentano polimorfici: infatti, in tutti i casi l'allele più comune ha una frequenza più bassa del 99% e il numero degli alleli va da un minimo di 7 (per il *locus* Cae30) ad un massimo di 72 (per il *locus* Cae07). Il livello di variabilità risulta piuttosto elevato: infatti, l'eterozigosità attesa varia da 0.1010 per il *locus* Cae30 a 0.9622 per il *locus* Cae07. Escludendo i *loci* Cae30 e Cae14, in generale l'eterozigosità attesa risulta superiore ad un valore di 0.7658: questo significa che la probabilità di ottenere in un individuo diploide due alleli diversi è del 76.58%. L'eterozigosità osservata, invece, varia da 0.0974 per il *locus* Cae30 a 0.9251 per il *locus* Cae71. È importante notare, comunque, come l'eterozigosità osservata sia generalmente più bassa di quella attesa, situazione che è stata rilevata anche dal test di equilibrio di Hardy-Weinberg: infatti, sei test su undici risultano altamente significativi e tale disequilibrio è da imputare ad un eccesso di omozigoti. Una simile situazione, del resto, è piuttosto comune e diffusa in organismi marini (Burton, 1983; Ayre *et al.*, 1997; Dias *et al.*, 2006) e può essere ascritta a diversi motivi, come selezione contro gli eterozigoti, presenza di alleli nulli, effetto Wahlund, inbreeding. In questo studio, tuttavia, si ritiene che il motivo principale di un tale deficit di eterozigoti sia da attribuire a fenomeni di inincrocio, come confermato anche dai valori positivi di  $F_{is}$  osservati per tutti gli 11 *loci* (valori compresi tra 0.004 e 0.247). Il fatto, poi, che questa carenza di eterozigoti sia osservata in tutte le popolazioni e con tutti i *loci* fa sospettare la presenza di un reale (biologico) disequilibrio di Hardy-Weinberg.

**Tabella 3.7:** valori di variabilità genetica per gli otto campioni di popolazione. Sono riportati la sigla della popolazione, il numero medio di alleli (N), l'eterozigosità media attesa (H exp) e media osservata (H obs), la probabilità dell'equilibrio di Hardy-Weinberg.

| Pop | N° medio alleli | H media exp. | H media obs. | HWE     |
|-----|-----------------|--------------|--------------|---------|
| Cag | 18.0909         | 0.7777       | 0.7267       | <0.0001 |
| Ta  | 16.7273         | 0.8071       | 0.7417       | <0.0001 |
| Le  | 18.1818         | 0.8059       | 0.7129       | <0.0001 |
| Fg  | 17.6364         | 0.7793       | 0.7043       | <0.0001 |
| Ra  | 17.9091         | 0.8017       | 0.7152       | <0.0001 |
| Go  | 18.8182         | 0.7858       | 0.7077       | <0.0001 |
| Ve  | 28.8182         | 0.7963       | 0.7332       | <0.0001 |
| Ma  | 17.8182         | 0.8044       | 0.7140       | <0.0001 |

**Tabella 3.8:** valori di variabilità genetica per gli undici *loci* microsatellite. Sono riportati il nome del *locus*, il numero medio di alleli (N), l'eterozigosità attesa (H exp) e osservata (H obs), la probabilità dell'equilibrio di Hardy-Weinberg.

| loci      | N° alleli | H exp. | H obs. | HWE     |
|-----------|-----------|--------|--------|---------|
| Cae01     | 25        | 0.8376 | 0.8231 | <0.0001 |
| Cae07     | 72        | 0.9622 | 0.8179 | <0.0001 |
| Cae14     | 22        | 0.6445 | 0.5623 | <0.0001 |
| Cae17     | 24        | 0.8664 | 0.8630 | 0.1475  |
| Cae30     | 7         | 0.1010 | 0.0974 | 0.6747  |
| Cae33     | 59        | 0.9596 | 0.7242 | <0.0001 |
| Cae71     | 42        | 0.9483 | 0.9251 | 0.0311  |
| Cae86     | 26        | 0.9284 | 0.9148 | 0.0249  |
| CmaEPA02  | 43        | 0.9026 | 0.7218 | <0.0001 |
| CmaEPA04  | 28        | 0.8263 | 0.8117 | 0.0021  |
| CmaEPA14m | 17        | 0.7658 | 0.6531 | <0.0001 |

Il programma *Genepop* ha permesso inoltre di effettuare l'analisi di differenziamento di popolazioni a diverse scale geografiche.

In primo luogo si è voluta testare l'appartenenza ad un'unica popolazione omogenea (popolazione panmittica) per i campioni interni alla Laguna di Venezia (N=486). Il confronto tra i campioni raccolti è risultato significativo ( $P < 0.0001$ ), permettendo quindi di rigettare l'ipotesi di omogeneità. A conferma di quanto ottenuto dal test di differenziamento totale, è stato condotto un test tra tutte le possibili coppie di popolazioni veneziane (tabella 3.9).

**Tabella 3.9: risultati del test esatto di differenziamento genico tra tutte le possibili coppie di popolazioni veneziane. In grassetto sono evidenziati i valori di probabilità significativi in base alla soglia dello 0.000641 (soglia con correzione di Bonferroni).**

|              | <b>Ce05</b> | <b>Ci05</b> | <b>Fu05</b>    | <b>Ce06</b> | <b>Ci06</b> | <b>Cz06</b> | <b>Fu06</b> | <b>Pc06</b>    | <b>Sal06</b> | <b>Ca07</b> | <b>Ci07</b> | <b>Fu07</b> | <b>Sc07</b> |
|--------------|-------------|-------------|----------------|-------------|-------------|-------------|-------------|----------------|--------------|-------------|-------------|-------------|-------------|
| <b>Ce05</b>  |             |             |                |             |             |             |             |                |              |             |             |             |             |
| <b>Ci05</b>  | 0.00244     |             |                |             |             |             |             |                |              |             |             |             |             |
| <b>Fu05</b>  | 0.00445     | 0.00071     |                |             |             |             |             |                |              |             |             |             |             |
| <b>Ce06</b>  | 0.09476     | 0.09880     | 0.00830        |             |             |             |             |                |              |             |             |             |             |
| <b>Ci06</b>  | 0.53851     | 0.14432     | 0.10194        | 0.47561     |             |             |             |                |              |             |             |             |             |
| <b>Cz06</b>  | 0.13738     | 0.12916     | 0.03489        | 0.00579     | 0.06396     |             |             |                |              |             |             |             |             |
| <b>Fu06</b>  | 0.04252     | 0.00066     | 0.32853        | 0.20529     | 0.53861     | 0.19208     |             |                |              |             |             |             |             |
| <b>Pc06</b>  | 0.00531     | 0.00885     | <b>0.00050</b> | 0.01017     | 0.19103     | 0.00628     | 0.00410     |                |              |             |             |             |             |
| <b>Sal06</b> | 0.03037     | 0.01118     | 0.09101        | 0.20131     | 0.54952     | 0.21716     | 0.26387     | 0.08192        |              |             |             |             |             |
| <b>Ca07</b>  | 0.00209     | 0.09374     | 0.05173        | 0.47281     | 0.91566     | 0.00972     | 0.06857     | <b>0.00014</b> | 0.25358      |             |             |             |             |
| <b>Ci07</b>  | 0.10547     | 0.01644     | 0.05869        | 0.82216     | 0.97330     | 0.20224     | 0.31129     | 0.00276        | 0.38094      | 0.55568     |             |             |             |
| <b>Fu07</b>  | 0.16070     | 0.00598     | 0.00500        | 0.31931     | 0.56367     | 0.21243     | 0.38161     | 0.00361        | 0.59248      | 0.45087     | 0.43982     |             |             |
| <b>Sc07</b>  | 0.04850     | 0.12266     | 0.04513        | 0.62816     | 0.68670     | 0.06647     | 0.44813     | 0.00577        | 0.15011      | 0.48887     | 0.38015     | 0.29637     |             |

Dalla matrice, appare evidente che la non omogeneità complessiva derivi dalla significatività di 2 singoli confronti (Paludi della Centrega 2006-Fusina 2005,  $P=0.00050$ ; Paludi della Centrega 2006-Campalto 2007,  $P=0.00014$ , soglia dello 0.05 con correzione di Bonferroni) e da numerosi confronti significativi per una soglia dello 0.05 senza correzione di Bonferroni, ma non vi è alcuna popolazione che risulti completamente diversa da tutte le altre.

Per questo motivo e per i valori molto bassi e non significativi degli indici  $F_{st}$  (si veda tabella 3.12) calcolati a coppie, si è ritenuto opportuno accorpate i campioni provenienti dalla laguna di Venezia in un unico gruppo di 486 individui e considerarli, nelle analisi successive, come un singolo campione di popolazione.

Si è voluta, quindi, testare la presenza di differenziamento all'interno dei campioni di popolazione dell'Adriatico settentrionale (Marano, Venezia, Goro e Ravenna). Dal test di differenziamento effettuato a coppie appare evidente come i campioni presi in esame siano caratterizzati da una sostanziale omogeneità (tabella 3.10), in quanto i valori di probabilità di omogeneità risultano ben al di sopra della soglia di significatività, fatta eccezione per due singoli confronti che interessano il campione di Marano (Ma) e che risultano statisticamente significativi. Questo dato permette, quindi, di affermare che, complessivamente, l'Adriatico settentrionale presenta livelli di variabilità genetica piuttosto bassi che si traducono in una sostanziale omogeneità tra i campioni presi in considerazione.

**Tabella 3.10: risultati del test di differenziamento genico a coppie per i campioni di popolazione dell'Adriatico settentrionale. In grassetto sono evidenziati i valori di probabilità significativi in base alla soglia dello 0.008333 (soglia con correzione di Bonferroni).**

|           | <b>Ma</b>       | <b>Ve</b> | <b>Go</b> | <b>Ra</b> |
|-----------|-----------------|-----------|-----------|-----------|
| <b>Ma</b> |                 |           |           |           |
| <b>Ve</b> | <b>0.000022</b> |           |           |           |
| <b>Go</b> | <b>0.004551</b> | 0.148404  |           |           |
| <b>Ra</b> | 0.009393        | 0.032734  | 0.903182  |           |

Infine si è voluto estendere l'indagine a tutti i campioni di popolazione adriatici e ai campioni dello Ionio e del Tirreno a disposizione. Dal test di differenziamento genico a coppie condotto tra gli otto campioni è apparso chiaro come il sito di Cagliari risulti avere valori di variabilità genetica altamente significativi nei confronti di tutti gli altri siti di campionamento (sia che si applichi o meno la correzione di Bonferroni per test multipli, tabella 3.11): questo risultato sta ad indicare che i campioni cagliaritari sono significativamente diversi da quelli dell'Adriatico e dello Ionio. Tuttavia, anche tra i campioni adriatici e ionico e all'interno dei soli campioni adriatici sono presenti dei confronti significativi: ad esempio, il sito di Taranto (Ta) risulta significativamente diverso da tutti i campioni adriatici (per una soglia dello 0.05 senza correzione di Bonferroni), ad eccezione del confronto con il sito di Aquatina (Le), che geograficamente è il più vicino ad esso (se si applica la correzione per test multipli). All'interno dell'Adriatico, Aquatina (Le) risulta essere il campione che maggiormente si differenzia dagli altri: infatti per una soglia dello 0.05 il campione leccese risulta significativamente diverso dal resto dei campioni adriatici. Applicando la correzione di Bonferroni, 4 confronti su 5 rimangono comunque significativi.

**Tabella 3.11: risultati del test esatto di differenziamento genico tra tutte le possibili coppie di campioni. In grassetto sono evidenziati i valori di probabilità significativi in base alla soglia dello 0.00178571 (soglia con correzione di Bonferroni).**

|            | <b>Cag</b>         | <b>Ta</b>          | <b>Le</b>          | <b>Fg</b>      | <b>Ra</b> | <b>Go</b> | <b>Ve</b>      | <b>Ma</b> |
|------------|--------------------|--------------------|--------------------|----------------|-----------|-----------|----------------|-----------|
| <b>Cag</b> |                    |                    |                    |                |           |           |                |           |
| <b>Ta</b>  | <b>&lt;0.00001</b> |                    |                    |                |           |           |                |           |
| <b>Le</b>  | <b>&lt;0.00001</b> | 0.00391            |                    |                |           |           |                |           |
| <b>Fg</b>  | <b>&lt;0.00001</b> | <b>&lt;0.00001</b> | <b>0.00001</b>     |                |           |           |                |           |
| <b>Ra</b>  | <b>&lt;0.00001</b> | <b>0.00004</b>     | 0.00635            | 0.02451        |           |           |                |           |
| <b>Go</b>  | <b>&lt;0.00001</b> | <b>0.00039</b>     | <b>0.00030</b>     | 0.02011        | 0.90416   |           |                |           |
| <b>Ve</b>  | <b>&lt;0.00001</b> | <b>&lt;0.00001</b> | <b>&lt;0.00001</b> | 0.00199        | 0.03377   | 0.14729   |                |           |
| <b>Ma</b>  | <b>&lt;0.00001</b> | <b>&lt;0.00001</b> | <b>&lt;0.00001</b> | <b>0.00124</b> | 0.00850   | 0.00570   | <b>0.00002</b> |           |

Da questa analisi, quindi, è possibile osservare la presenza di differenziamento tra organismi che provengono da siti di campionamento geograficamente separati: in generale, maggiore è la distanza geografica e maggiore è il differenziamento. In tabella 3.11 è possibile, tuttavia, osservare delle eccezioni che non rispettano tale tendenza: ad esempio, oltre al confronto tra Marano (Ma) e Venezia (Ve) ( $P=0.00002$ ), che conferma quanto ottenuto dal test di differenziamento per i campioni di popolazione dell'Adriatico settentrionale, vi è anche il confronto tra il sito di Lesina (Fg) e di Aquatina (Le) a presentare un valore di differenziamento significativo ( $P=0.00001$ ) nonostante la vicinanza geografica.

Per quantificare le differenze geniche tra i campioni della laguna di Venezia, i campioni dell'Adriatico settentrionale e gli 8 campioni di popolazione, sono stati calcolati i rispettivi indici  $F_{st}$  con il programma *Genetix*. I valori ottenuti e le probabilità ad essi associate sono riportati nelle tabelle 3.12 (indici  $F_{st}$  per la laguna di Venezia), 3.13 (indici  $F_{st}$  per l'Adriatico settentrionale), 3.14 (indici  $F_{st}$  per gli 8 campioni di popolazione).

**Tabella 3.12: indici  $F_{st}$  per la laguna di Venezia. Sono riportati sotto la diagonale gli indici  $F_{st}$  calcolati a coppie e sopra la diagonale i corrispondenti valori di significatività. La soglia di significatività è stata fissata a 0.000641, in seguito a correzione di Bonferroni per test multipli.**

|              | <b>CE05</b> | <b>CI05</b> | <b>FU05</b> | <b>CE06</b> | <b>CI06</b> | <b>FU06</b> | <b>CZ06</b> | <b>PC06</b> | <b>SAL06</b> | <b>CA07</b> | <b>CI07</b> | <b>FU07</b> | <b>SC07</b> |
|--------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|--------------|-------------|-------------|-------------|-------------|
| <b>CE05</b>  |             | 0.480       | 0.260       | 0.332       | 0.886       | 0.746       | 0.484       | 0.105       | 0.796        | 0.184       | 0.488       | 0.601       | 0.647       |
| <b>CI05</b>  | 0.00014     |             | 0.192       | 0.314       | 0.752       | 0.366       | 0.656       | 0.087       | 0.689        | 0.504       | 0.265       | 0.164       | 0.751       |
| <b>FU05</b>  | 0.00134     | 0.00196     |             | 0.098       | 0.703       | 0.615       | 0.057       | 0.010       | 0.683        | 0.482       | 0.367       | 0.068       | 0.391       |
| <b>CE06</b>  | 0.00073     | 0.00126     | 0.00247     |             | 0.343       | 0.203       | 0.001       | 0.079       | 0.382        | 0.200       | 0.446       | 0.136       | 0.330       |
| <b>CI06</b>  | -0.00276    | -0.00194    | -0.00102    | 0.00069     |             | 0.913       | 0.509       | 0.249       | 0.984        | 0.908       | 0.914       | 0.558       | 0.755       |
| <b>FU06</b>  | -0.00125    | 0.00080     | -0.00061    | 0.00146     | -0.00234    |             | 0.546       | 0.015       | 0.906        | 0.359       | 0.256       | 0.563       | 0.907       |
| <b>CZ06</b>  | -0.00014    | -0.00099    | 0.00314     | 0.00592     | -0.00015    | -0.00043    |             | 0.011       | 0.610        | 0.087       | 0.210       | 0.508       | 0.449       |
| <b>PC06</b>  | 0.00442     | 0.00588     | 0.00820     | 0.00908     | 0.00309     | 0.00829     | 0.00931     |             | 0.139        | 0.003       | 0.020       | 0.003       | 0.042       |
| <b>SAL06</b> | -0.00193    | -0.00118    | -0.00096    | 0.00042     | -0.00427    | -0.00253    | -0.00075    | 0.00383     |              | 0.792       | 0.852       | 0.781       | 0.781       |
| <b>CA07</b>  | 0.00235     | -0.00009    | 0.00006     | 0.00241     | -0.00370    | 0.00065     | 0.00401     | 0.01461     | -0.00208     |             | 0.769       | 0.298       | 0.760       |
| <b>CI07</b>  | -0.00007    | 0.00198     | 0.00060     | 0.00019     | -0.00403    | 0.00164     | 0.00215     | 0.01024     | -0.00290     | -0.00233    |             | 0.492       | 0.793       |
| <b>FU07</b>  | -0.00090    | 0.00264     | 0.00394     | 0.00348     | -0.00069    | -0.00079    | -0.00016    | 0.01400     | -0.00220     | 0.00151     | -0.00017    |             | 0.601       |
| <b>SC07</b>  | -0.00110    | -0.00196    | 0.00054     | 0.00099     | -0.00213    | -0.00308    | 0.00012     | 0.00768     | -0.00216     | -0.00201    | -0.00272    | -0.00110    |             |

Dalla matrice riportata in tabella 3.12 appare evidente come, nei confronti a coppie tra i diversi campioni della laguna veneta, non ci siano valori di  $F_{st}$  statisticamente significativi (per una soglia dello 0.000641, con correzione di Bonferroni); tuttavia, come per il test di differenziamento, è possibile osservare sporadicamente valori di  $F_{st}$  molto elevati (ad esempio Paludi della Centrega 2006 e Fusina 2006 con  $F_{st}=0.00829$ ; Paludi della Centrega 2006 e Ca'Zane 2006 con  $F_{st}=0.00931$ ), nonostante la ridotta scala geografica presa in considerazione. Tuttavia, la presenza in generale di valori di  $F_{st}$  non statisticamente significativi ha fatto sì che i campioni della laguna di Venezia fossero riuniti in un unico gruppo di 486 individui, come era stato effettuato per i test di differenziamento.

Anche per quanto riguarda l'Adriatico settentrionale, in generale, non si riscontrano differenze significative (tabella 3.13), fatta eccezione per il confronto tra Marano (Ma) e Venezia (Ve,  $F_{st}=0.00278$  con  $P=0.004$ ). Questi dati risultano in perfetto accordo con quanto ottenuto dal test di differenziamento a coppie: il calcolo degli indici  $F_{st}$ , quindi, conferma la situazione di complessiva omogeneità dei campioni dell'Adriatico settentrionale.

**Tabella 3.13: indici  $F_{st}$  per l'Adriatico settentrionale. Sono riportati sotto la diagonale gli indici  $F_{st}$  calcolati a coppie e sopra la diagonale i corrispondenti valori di significatività. In grassetto sono evidenziati i valori  $F_{st}$  statisticamente significativi per una soglia dello 0.008333 (soglia con correzione di Bonferroni).**

|    | Ma             | Ve       | Go       | Ra    |
|----|----------------|----------|----------|-------|
| Ma |                | 0.004    | 0.105    | 0.154 |
| Ve | <b>0.00278</b> |          | 0.750    | 0.716 |
| Go | 0.00192        | -0.00055 |          | 0.795 |
| Ra | 0.00160        | -0.00057 | -0.00128 |       |

Infine, per quanto riguarda il calcolo degli  $F_{st}$  per gli otto campioni di popolazione (Marano, Venezia, Goro, Ravenna, Lesina, Aquatina, Taranto e Cagliari) risulta evidente che i valori più elevati e significativi sono quelli associati al sito di Cagliari (Cag): infatti, questo sito mostra differenze altamente significative (tabella 3.14) nei confronti di tutti gli altri siti considerati ( $P<0.0001$ ), dato che conferma quanto ottenuto dal test di differenziamento a coppie. Inoltre, se si considera una soglia dello 0.05 senza correzione di Bonferroni, è possibile osservare valori di  $F_{st}$  significativi anche tra i campioni ionico e adriatici e all'interno dei soli campioni adriatici. Infatti, sebbene i valori osservati siano sensibilmente più bassi, sono confermate le differenze, ad esempio, tra Taranto (Ta) e tre dei campioni adriatici (Lesina, Fg; Venezia, Ve e Marano, Ma) e tra Aquatina (Le), Venezia (Ve) e Marano (Ma). Inoltre sono confermate le differenze tra Aquatina (Le) e Lesina (Fg) e tra Marano (Ma) e Venezia (Ve), come già osservato nel test di differenziamento.

**Tabella 3.14:** indici  $F_{st}$  per gli 8 campioni di popolazione. Sono riportati sotto la diagonale gli indici  $F_{st}$  calcolati a coppie e sopra la diagonale i corrispondenti valori di significatività. In grassetto sono evidenziati i valori  $F_{st}$  statisticamente significativi per una soglia dello 0.001786 (soglia con correzione di

|            | <b>Cag</b>     | <b>Ta</b> | <b>Le</b> | <b>Fg</b> | <b>Ra</b> | <b>Go</b> | <b>Ve</b> | <b>Ma</b> |
|------------|----------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| <b>Cag</b> |                | <0.0001   | <0.0001   | <0.0001   | <0.0001   | <0.0001   | <0.0001   | <0.0001   |
| <b>Ta</b>  | <b>0.02773</b> |           | 0.3470    | 0.0050    | 0.0600    | 0.1650    | 0.0060    | 0.0290    |
| <b>Le</b>  | <b>0.02719</b> | 0.00044   |           | 0.0170    | 0.2630    | 0.1060    | 0.0090    | 0.0230    |
| <b>Fg</b>  | <b>0.03781</b> | 0.00479   | 0.00299   |           | 0.1600    | 0.5050    | 0.0510    | 0.0120    |
| <b>Ra</b>  | <b>0.03512</b> | 0.00262   | 0.00086   | 0.00154   |           | 0.7730    | 0.7050    | 0.1640    |
| <b>Go</b>  | <b>0.03828</b> | 0.00153   | 0.00162   | -0.00019  | -0.00128  |           | 0.7250    | 0.1100    |
| <b>Ve</b>  | <b>0.03800</b> | 0.00255   | 0.00200   | 0.00167   | -0.00057  | -0.00055  |           | 0.0050    |
| <b>Ma</b>  | <b>0.03066</b> | 0.00326   | 0.00309   | 0.00432   | 0.00160   | 0.00192   | 0.00278   |           |

In tutti i casi, indici  $F_{st}$  con valori negativi stanno ad indicare che individui provenienti da popolazioni diverse sono più simili tra loro che individui provenienti da una stessa popolazione. Ciò dimostra ancora una volta che i campioni di popolazione adriatici analizzati sono geneticamente molto simili, sebbene sia presente un debole differenziamento al loro interno.

Si è proceduto, quindi, tramite il programma *Arlequin*, all'analisi in dettaglio del differenziamento tra popolazioni, attraverso l'analisi molecolare della varianza (*AMOVA*), i cui risultati sono riportati in tabella 3.15. In questo caso il calcolo è stato effettuato solo per gli otto campioni di popolazione. Le popolazioni analizzate sono state suddivise in due gruppi: un primo gruppo composto dalle popolazioni dell'Adriatico e dello Ionio (Ma, Ve, Go, Ra, Fg, Le, Ta) e un secondo gruppo composto solo dalla popolazione di Cagliari (Cag).

**Tabella 3.15:** analisi molecolare della varianza (*AMOVA*). La significatività è stata stimata mediante 10000 permutazioni.

| Sorgente di varianza                | d.f. | Somma dei quadrati | Componenti della varianza | Percentuale di variazione | P-value  |
|-------------------------------------|------|--------------------|---------------------------|---------------------------|----------|
| <b>Tra gruppi</b>                   | 1    | 32.726             | 0.14934                   | 3.82                      | <0.00001 |
| <b>Tra popolazioni entro gruppi</b> | 6    | 28.948             | 0.00637                   | 0.16                      | 0.00188  |
| <b>Entro popolazioni</b>            | 1726 | 6475.133           | 3.75123                   | 96.01                     |          |
| <b>Totale</b>                       | 1733 | 6536.807           | 3.90723                   |                           |          |

Tale suddivisione deriva da vari tentativi di raggruppamento (si è cercato, ad esempio, di raggruppare assieme Taranto e Cagliari), ma questa è risultata la migliore suddivisione che massimizza la percentuale di variazione.

La maggior parte della varianza (96.01%), che spiega quasi tutta la variabilità, è stata trovata entro popolazioni; tuttavia, una quota della variazione (3.82%) è da attribuire proprio alla divisione in gruppi. Questa parte della variabilità associata alla suddivisione in gruppi risulta statisticamente significativa, con un valore di  $P < 0.00001$ . Questo dato rappresenta, quindi, un'ulteriore conferma ai risultati ottenuti dal calcolo degli indici  $F_{st}$  a coppie: Cagliari (Cag) risulta essere il campione che maggiormente si differenzia dal resto dei campioni analizzati.

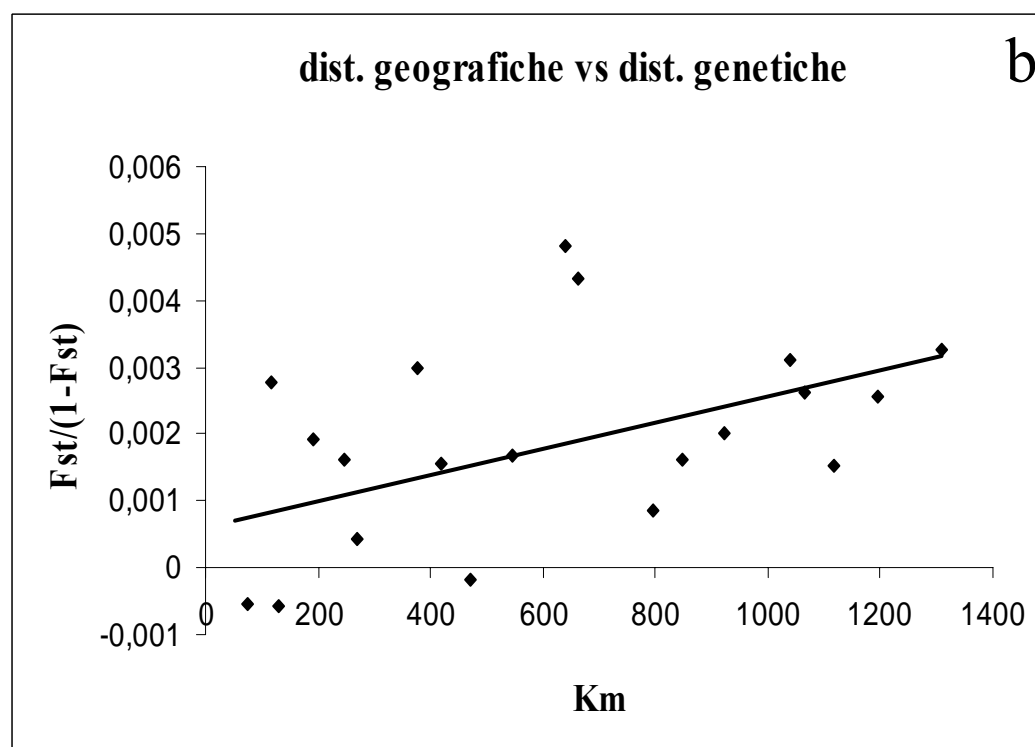
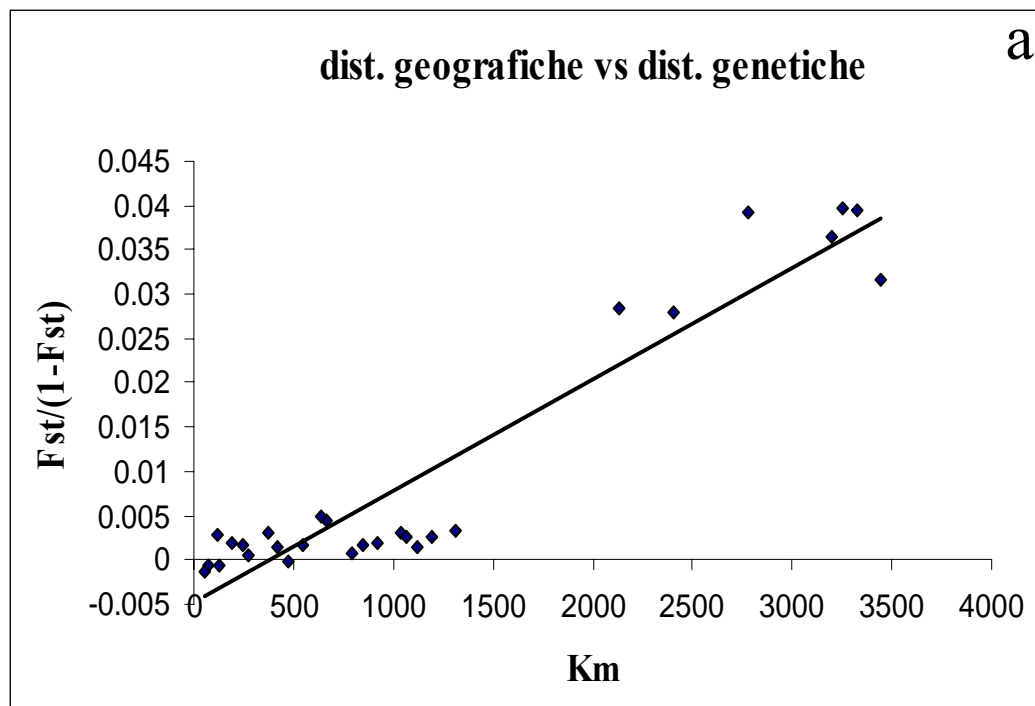
Un valore di variabilità estremamente basso ma significativo è stato inoltre riscontrato tra popolazioni entro gruppi, a conferma di un debole differenziamento tra i campioni adriatici e ionico.

Per rafforzare ulteriormente questo risultato, l'*AMOVA* è stata quindi effettuata escludendo dall'analisi il campione di Cagliari: la percentuale di variazione tra popolazioni ha restituito, in accordo con i valori ottenuti dall'analisi complessiva, un valore molto basso (0.17% con  $P < 0.05$ ) ma significativo. Questo dato rappresenta un risultato molto importante in questa fase di analisi: infatti ha evidenziato un seppur debole differenziamento che con il marcatore mitocondriale non era stato rilevato. Il motivo di una tale discordanza potrebbe risiedere nel diverso potere risolutivo dei due marcatori: in questo caso, i microsatelliti si sono dimostrati di fondamentale importanza proprio per il loro elevato potere risolutivo che ha permesso di evidenziare quelle minime differenze presenti tra i campioni adriatici e ionico, altrimenti non riscontrabili con il solo utilizzo del marcatore mitocondriale.

Il programma *Arlequin* ha, inoltre, permesso di testare statisticamente l'isolamento per distanza attraverso il test di Mantel. Il grafico ottenuto (figura 3.13a) ha evidenziato una correlazione positiva tra distanza genetica e distanza geografica, confermando in questo modo i risultati ottenuti dal test di differenziamento genico, dagli indici  $F_{st}$  e dall'*AMOVA*: popolazioni geograficamente più distanti sono anche geneticamente più diverse tra loro.

Il coefficiente di correlazione restituito dal test, pari a 0.932506, risulta statisticamente significativo con un *p-value* di 0.0111: questo dato conferma ulteriormente la presenza di isolamento per distanza.

Tuttavia, è necessario sottolineare come la significatività riscontrata dal test sia dovuta in larga misura ad un campione geograficamente molto distante (Cagliari): escludendo, infatti, tale campione, il coefficiente di correlazione diminuisce notevolmente (0.367524), perdendo di significatività (*p-value* 0.0592, figura 3.13b).

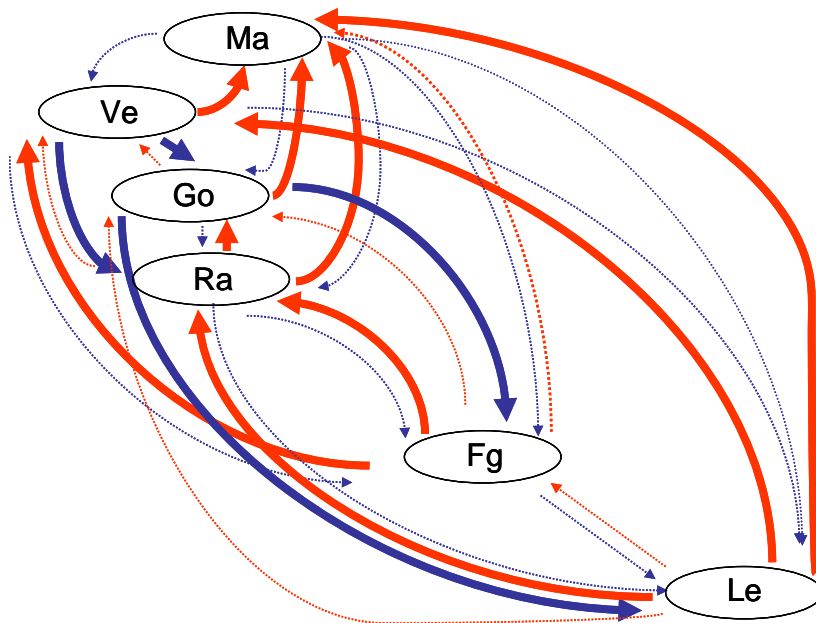


**Figura 3.13: isolamento per distanza. Il grafico a) mostra la regressione lineare tra distanza genetica e distanza geografica degli otto campioni di popolazione. Il grafico b) mostra la regressione lineare tra distanza genetica e distanza geografica dei campioni di popolazione adriatici e ionico. Nell'asse delle ascisse viene riportata la distanza geografica espressa in km; nell'asse delle ordinate viene riportata la distanza genetica espressa in  $F_{st}/(1-F_{st})$ .**

Ciò sta ad indicare che la distanza geografica ha sì un effetto sul differenziamento, ma non rappresenta l'unico fattore determinante: altri fattori, come ad esempio le correnti oceanografiche o i flussi migratori, potrebbero intervenire e mantenere un certo livello di variabilità tra i campioni presi in esame.

A tale fine, sono state effettuate, attraverso il programma *Migrate 2.3* (Beerli and Felsenstein, 2001), delle indagini volte a stimare i tassi di migrazione tra i 6 campioni di popolazione adriatici.

In un primo momento si è voluto testare l'ipotesi del modello a isole. Nella sua implementazione, questo prevede che la migrazione avvenga tra tutte le popolazioni, indipendentemente dalla distanza geografica e permette un flusso asimmetrico tra le stesse. Il quadro che ne deriva risulta piuttosto complesso (figura 3.14): i flussi migratori, infatti, tendono generalmente ad andare da sud verso nord, anche se sono presenti numerose eccezioni. In particolare, i campioni di popolazione di Venezia, Goro e Ravenna scambiano tutti e tre un numero piuttosto elevato di individui in direzione di Marano, mentre la situazione si inverte quando si osservano gli scambi tra Venezia e Goro e tra Venezia e Ravenna. In questi casi, infatti, il flusso migratorio va da nord a sud, sfruttando le correnti superficiali cicloniche (antiorarie) dell'Adriatico. Inoltre, è possibile osservare come i campioni meridionali (Lesina e Aquatina) rappresentino un importante contributo nei confronti delle popolazioni settentrionali, scambiando un numero consistente di individui soprattutto con Marano, Venezia e Ravenna (tabella 3.16), risalendo lungo la costa croata attraverso le correnti superficiali dell'Adriatico. Tuttavia, il flusso si inverte nuovamente da Goro a Lesina e da Goro ad Aquatina.



**Figura 3.14:** rappresentazione grafica dei risultati ottenuti con il programma *Migrate 2.3* sui tassi di migrazione secondo il modello a isole. Con le frecce rosse sono indicati i flussi da sud a nord, in blu quelli da nord a sud; maggiore è lo spessore della freccia e più intenso è il flusso migratorio.

Quindi, utilizzando un modello a isole, ciò che si ricava è una situazione di non facile interpretazione, anche se in prevalenza il flusso migratorio tende ad andare in un'unica direzione, da sud a nord.

Nel tentativo di rendere più comprensibile i meccanismi alla base del flusso di migrazione, si è cercato di introdurre un modello più semplice e realistico, che spiegasse in modo più chiaro la situazione presente in Adriatico. Si è quindi passati a testare l'ipotesi del modello *stepping stone*. Questo modello, infatti, prevede che la migrazione sia limitata solo a popolazioni adiacenti. In realtà, in questo lavoro il modello adottato è stato in parte modificato, prendendo in considerazione non solo i possibili scambi tra sottobacini, ma consentendo inoltre un contributo da parte delle popolazioni meridionali. In questo contesto, il flusso migratorio assume un andamento ciclonico (antiorario), in accordo con le correnti presenti in Adriatico, con una tendenza che va da nord a sud. Un'unica eccezione al generale flusso antiorario, in questo quadro estremamente semplificato ma di più facile interpretazione, è rappresentata da Venezia che scambia un numero maggiore di individui in direzione di Marano (figura 3.15). I campioni di popolazione meridionali (Lesina, Fg e Aquatina, Le) contribuiscono effettivamente a mantenere il flusso ciclonico, scambiando una quantità consistente di individui con i campioni settentrionali, rappresentati soprattutto dal campione di Marano (tabella 3.17a).

Parallelamente a questo tipo di modello, si è voluta testare l'ipotesi che prevede il contributo da parte del solo campione di Aquatina (fig. 3.16). Dalla tabella 3.17b è possibile osservare come i valori dei tassi di migrazione siano del tutto paragonabili a quelli ottenuti con il modello *stepping stone*, fatta eccezione per i campioni di Venezia e Goro, per i quali sembra ci sia un flusso uniforme in entrambe le direzioni. Tuttavia, questi risultati permettono non solo di confermare l'andamento antiorario del flusso migratorio, da nord a sud, ma inoltre, di rafforzare l'ipotesi di un reale contributo dei campioni di popolazione meridionali verso nord.

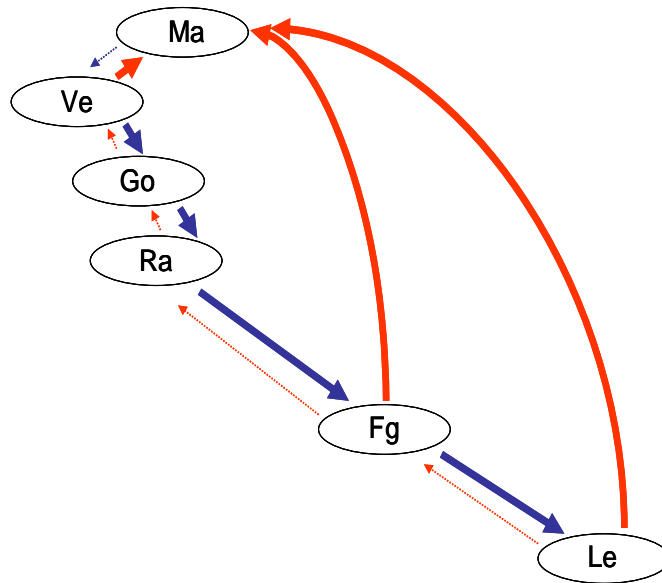


Figura 3.15: rappresentazione grafica dei risultati ottenuti con il programma *Migrate* 2.3 sui tassi di migrazione secondo il modello *stepping stone* modificato. Con le frecce rosse sono indicati i flussi da sud a nord, in blu quelli da nord a sud; maggiore è lo spessore delle frecce e più intenso è il flusso migratorio.

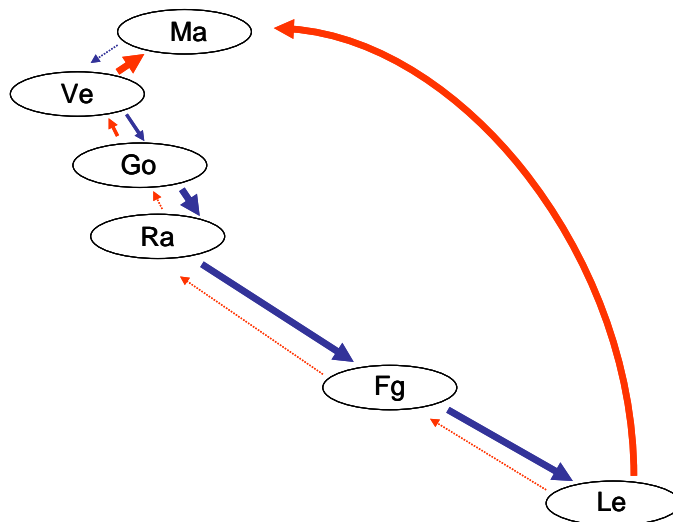


Figura 3.17: rappresentazione grafica dei risultati ottenuti con il programma *Migrate* 2.3 sui tassi di migrazione secondo il modello *stepping stone*, con contributo da parte del solo campione di Aquatina. Con le frecce rosse sono indicati i flussi da sud a nord, in blu quelli da nord a sud; maggiore è lo spessore delle frecce e più intenso è il flusso migratorio.

**Tabella 3.16: tassi di migrazione secondo il modello a isole. In matrice vengono riportati i valori di M (tasso di migrazione) con i rispettivi intervalli di confidenza ottenuti con il programma *Migrate* 2.3. In rosso sono evidenziati i flussi da sud a nord, mentre in blu i flussi da nord a sud. I numeri riportati a pedice di ogni indice M corrispondono rispettivamente a 1) Marano, 2) Venezia, 3) Goro, 4) Ravenna, 5) Lesina, 6) Aquatina.**

|   |  |   |  |   |
|---|--|---|--|---|
| Marano - Venezia<br>M <sub>12</sub> 1.36861(1.26147-1.48160)<br>M <sub>21</sub> 1.87657(1.74482-2.01475)  |  |   |  |   |
| Marano - Goro<br>M <sub>13</sub> 1.08913(0.99413-1.18997)<br>M <sub>31</sub> 1.50358(1.38619-1.62741)     | Venezia - Goro<br>M <sub>23</sub> 1.44676(1.33655-1.56274)<br>M <sub>32</sub> 1.13227(1.03492-1.23554)     |   |  |   |
| Marano - Ravenna<br>M <sub>14</sub> 1.20041(1.10098-1.30566)<br>M <sub>41</sub> 1.68064(1.55624-1.81142)  | Venezia - Ravenna<br>M <sub>24</sub> 1.50568(1.39349-1.62363)<br>M <sub>42</sub> 1.17763(1.07835-1.28300)  | Goro - Ravenna<br>M <sub>34</sub> 1.04227(0.94960-1.14070)<br>M <sub>43</sub> 2.09608(1.96268-2.23550)  |  |   |
| Marano - Lesina<br>M <sub>15</sub> 1.44170(1.33761-1.55098)<br>M <sub>51</sub> 1.59800(1.47695-1.72548)   | Venezia - Lesina<br>M <sub>25</sub> 1.63836(1.52751-1.75439)<br>M <sub>52</sub> 2.47695(2.33140-2.62847)   | Goro - Lesina<br>M <sub>35</sub> 1.31790(1.21859-1.42237)<br>M <sub>53</sub> 0.92791(0.84048-1.02125)   | Ravenna - Lesina<br>M <sub>45</sub> 1.19964(1.10525-1.29926)<br>M <sub>54</sub> 1.84698(1.72226-1.97769)   |   |
| Marano - Aquatina<br>M <sub>16</sub> 1.00882(0.90605-1.11911)<br>M <sub>61</sub> 1.25548(1.15349-1.36333) | Venezia - Aquatina<br>M <sub>26</sub> 0.90923(0.82482-0.99915)<br>M <sub>62</sub> 1.24517(1.14229-1.35406) | Goro - Aquatina<br>M <sub>36</sub> 1.45301(1.34583-1.56571)<br>M <sub>63</sub> 1.27896(1.17530-1.38851) | Ravenna - Aquatina<br>M <sub>46</sub> 1.26418(1.16338-1.37065)<br>M <sub>64</sub> 1.58717(1.47222-1.70802) | Lesina - Aquatina<br>M <sub>56</sub> 1.05111(0.96018-1.14758)<br>M <sub>65</sub> 1.10809(1.01711-1.20428) |

**Tabella 3.17:** tassi di migrazione secondo a) il modello *stepping stone* modificato; b) il modello *stepping stone*, con contributo da parte del solo campione di Aquatina. In matrice vengono riportati i valori di M (tasso di migrazione) con i rispettivi intervalli di confidenza ottenuti con il programma *Migrate 2.3*. In rosso sono evidenziati i flussi da sud a nord, mentre in blu i flussi da nord a sud. I numeri riportati a pedice di ogni indice M corrispondono rispettivamente a 1) Marano, 2) Venezia, 3) Goro, 4) Ravenna, 5) Lesina, 6) Aquatina.

| a) | Marano – Venezia                          | b) | Marano – Venezia                          |
|----|---|----|---|
|    | M <sub>12</sub> 1.71597 (1.60608-1.83083) |    | M <sub>12</sub> 1.94994 (1.82840-2.07578) |
|    | M <sub>21</sub> 2.24271 (2.04831-2.44903) |    | M <sub>21</sub> 2.49565 (2.31296-2.68775) |
|    | Venezia – Goro                            |    | Venezia – Goro                            |
|    | M <sub>23</sub> 2.77326 (2.63160-2.91962) |    | M <sub>23</sub> 1.73601 (1.63422-1.84142) |
|    | M <sub>32</sub> 1.72179 (1.61034-1.83823) |    | M <sub>32</sub> 1.72550 (1.61203-1.84503) |
|    | Goro – Ravenna                            |    | Goro – Ravenna                            |
|    | M <sub>34</sub> 2.28999 (2.16102-2.42298) |    | M <sub>34</sub> 1.92482 (1.81913-2.03447) |
|    | M <sub>43</sub> 1.24459 (1.15128-1.34219) |    | M <sub>43</sub> 1.43191 (1.34115-1.52705) |
|    | Ravenna – Lesina                          |    | Ravenna – Lesina                          |
|    | M <sub>45</sub> 2.68786 (2.55646-2.82365) |    | M <sub>45</sub> 2.17600 (2.05098-2.30854) |
|    | M <sub>54</sub> 0.77316 (0.69990-0.85139) |    | M <sub>54</sub> 1.44957 (1.35191-1.55051) |
|    | Lesina – Aquatina                         |    | Lesina – Aquatina                         |
|    | M <sub>56</sub> 2.61926 (2.48620-2.75668) |    | M <sub>56</sub> 1.94150 (1.82791-2.05925) |
|    | M <sub>65</sub> 0.98157 (0.90287-1.06465) |    | M <sub>65</sub> 1.39554 (1.29858-1.49852) |
|    | Aquatina – Marano                         |    | Aquatina – Marano                         |
|    | M <sub>61</sub> 1.87491 (1.69701-2.06556) |    | M <sub>61</sub> 1.86008 (1.70269-2.02785) |
|    | Lesina - Marano                           |    |   |
|    | M <sub>51</sub> 1.90293 (1.72194-2.09518) |    |   |

## ***4. DISCUSSIONE***



Nel corso di questo lavoro sono stati applicati due tipi di marcatori molecolari, il DNA mitocondriale e i microsattelliti, per analizzare la struttura genetica di campioni di popolazione adriatici di *Carcinus aestuarii*. Questo ha implicato l'amplificazione di un frammento di 482 paia di basi del gene mitocondriale codificante per la subunità I della citocromo c ossidasi (COI) e l'isolamento *ex novo* di 8 marcatori microsatellite specie-specifici per *C. aestuarii*, a cui sono stati affiancati 3 *loci* specifici per la specie atlantica *C. maenas*.

I risultati ottenuti forniscono una prima indicazione della struttura genetica del granchio verde nell'Adriatico e potrebbero permettere una prima analisi di come questa struttura sia stata influenzata dalle caratteristiche geografiche e oceanografiche dell'habitat della specie analizzata e da eventi storici e demografici passati.

Complessivamente è stato evidenziato un basso ma significativo livello di differenziamento genico tra i campioni di popolazione analizzati, condizione che permette di affermare la presenza di una lieve strutturazione genetica e di rigettare l'ipotesi di panmissia. In particolare, l'utilizzo dei marcatori microsatellite si è dimostrato di fondamentale importanza nel rilevare quelle piccole differenze presenti tra i campioni dell'Adriatico e dello Ionio: ciò è dovuto principalmente all'elevato potere risolutivo di suddetti marcatori, in grado di individuare anche le minime differenze genetiche che possono manifestarsi all'interno e tra popolazioni simili, che altrimenti non potrebbero essere riscontrate con il solo impiego dei marcatori mitocondriali. Infatti, negli organismi marini, generalmente caratterizzati da bassi livelli di differenziamento rispetto ad organismi dulciacquicoli e terrestri (Ward *et al.*, 1994; Waples, 1998), i microsattelliti hanno permesso di rilevare differenze anche molto piccole tra popolazioni di diverse specie. In particolare, con questi marcatori è stato possibile rilevare differenze tra campioni di individui di isole adiacenti, come nel caso del gobide *Elacatinus evelynae* (Taylor and Hellberg, 2003), di estuari vicini, com'è stato osservato per *Aphanius fasciatus* e *Plecoglossus altivelis ryukyuensis* (Maltagliati *et al.*, 2003; Ikeda *et al.*, 2003), all'interno di una stessa laguna, com'è stato riscontrato per le specie *Stegastes nigricans* e *Acanthurus triostegus* (Planes *et al.*, 1998) e persino nella stessa area di riproduzione, come nel caso del merluzzo *Gadus morhua* (Ruzzante *et al.*, 1998).

Nel presente lavoro, come già anticipato precedentemente, si è voluta indagare la struttura genetica della specie lagunare *Carcinus aestuarii*, utilizzando sia marcatori nucleari che mitocondriali. Per quanto riguarda questi ultimi, l'analisi è stata condotta su sei differenti campioni di popolazione provenienti da altrettante aree lagunari dell'Adriatico. Inoltre, sono stati analizzati, come siti di controllo, due campioni di popolazione esterni all'Adriatico: uno proveniente dal Mar Ionio e uno dal Mar Tirreno. Sono stati analizzati complessivamente 255 individui, per i quali è stato sequenziato un frammento di 482 paia di basi del gene mitocondriale codificante per la subunità I della citocromo c ossidasi (COI).

Si tratta della prima applicazione di tale marcatore in *C. aestuarii* su ampia scala geografica. La scelta di utilizzare il gene codificante per COI deriva dal diffuso impiego di questo marcatore molecolare in analisi di genetica di popolazione, come ampiamente riportato in letteratura. Infatti, studi condotti su vari invertebrati marini, in cui era previsto l'uso del gene codificante per la subunità I della citocromo c ossidasi, hanno rivelato la presenza di suddivisione di popolazione in specie europee quali ad esempio i gasteropodi *Hydrobia acuta* e *H. glyca* (Wilke and Pfenninger, 2002) o il bivalve *Macoma balthica* (Luttikhuisen *et al.*, 2003),

permettendo di affermare che nelle popolazioni analizzate lo scambio genico è piuttosto limitato. Inoltre, l'analisi di tale sequenza si è spesso rivelata un utile strumento per indagare la genetica di popolazione di altri invertebrati marini, come ricci di mare (Lessios *et al.*, 2001, 2003), gamberetti del genere *Penaeus* (Baldwin *et al.*, 1998) e varie specie di granchi, quali *Callinectes bellicosus* (Pfeiler *et al.*, 2005) e *Pachygrapsus crassipes* (Cassone and Boulding, 2006). Infine, per i suoi elevati livelli di polimorfismo, la subunità I della citocromo c ossidasi si è dimostrata un valido marcatore per condurre analisi di filogeografia soprattutto nei crostacei (Lejeusne and Chevaldonné, 2006).

In accordo con quanto affermato sino adesso, in *C. aestuarii* è stata riscontrata un'elevata variabilità di sequenza con valori di haplotype diversity ( $h$ ) che variano tra 0.6431 (per il campione di popolazione di Cagliari) e 0.9879 (per il campione di popolazione di Goro). Questo indice di diversità è una vera e propria misura di variabilità genetica e corrisponde concettualmente alla proporzione attesa di individui eterozigoti per il *locus* considerato, in quanto rappresenta la probabilità di campionare due sequenze diverse estraendo a caso due individui dalla popolazione. Il fatto che per Cagliari questo indice non sia così elevato, come per gli altri campioni, suggerisce già un primo differenziamento: infatti, dall'analisi degli indici  $F_{st}$  calcolata a coppie tra campioni di popolazione, è emerso un pronunciato livello di diversità nei confronti di tutti i campioni dell'Adriatico e dello Ionio. Inoltre, i valori più elevati sono stati riscontrati proprio nei confronti che riguardano il campione cagliaritano, con un valore medio di 0.03287. Il differenziamento di Cagliari è stato inoltre confermato dall'analisi molecolare della varianza (*AMOVA*), che ha permesso di effettuare un confronto globale tra tutte le popolazioni, considerando diversi livelli gerarchici. Da tale confronto è emersa la presenza di una lieve strutturazione genetica in questa specie. Dividendo, infatti, i campioni di popolazione analizzati in due gruppi (adriatici e ionico in un gruppo e tirrenico nell'altro) è emerso che una quota significativa di variabilità, pari a 0.03904, è attribuibile proprio a tale suddivisione.

Viceversa, per i campioni adriatici e ionico non è stato possibile evidenziare alcun grado di differenziamento: i valori degli indici  $F_{st}$  sono risultati molto bassi e nella maggior parte dei casi negativi e non significativi.

L'indice  $F_{st}$  è una vera e propria misura di distanza tra campioni di individui e varia da 0 in caso di completa omogeneità, a 1 in caso di completo differenziamento. Il fatto che nei campioni dell'Adriatico e dello Ionio siano presenti valori di  $F_{st}$  negativi sta ad indicare che tali campioni sono geneticamente molto simili e che individui provenienti da popolazioni diverse sono molto più simili tra loro che individui provenienti da una stessa popolazione. Anche in questo caso, l'*AMOVA* ha confermato questa generale omogeneità, restituendo un valore di percentuale di variazione tra popolazioni entro gruppi negativo e non significativo (-0.00134,  $P=0.49604$ ).

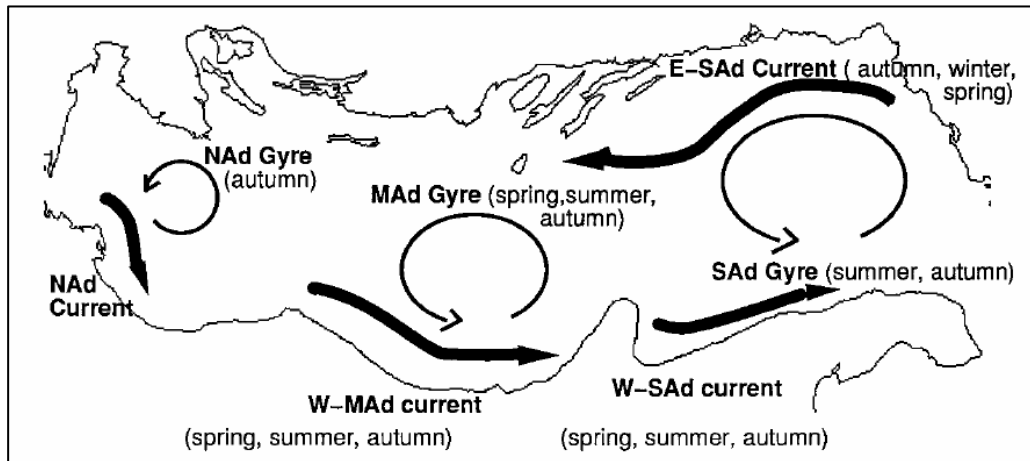
La stessa analisi effettuata, invece, con i marcatori microsatellite (per i quali sono stati amplificati 870 individui per un totale di 11 *loci*), ha evidenziato sì una debole strutturazione genetica, ma ha inoltre permesso di riscontrare un seppur debole differenziamento tra i campioni adriatici e ionico. È soprattutto a livello di analisi gerarchica che emergono delle differenze rispetto al DNA mitocondriale. Più precisamente, viene confermata la strutturazione genetica attribuibile alla suddivisione in gruppi, con un valore di variazione pari a 0.03822 (molto simile a quello ottenuto con l'analisi mitocondriale), ma soprattutto viene evidenziato un valore di variabilità estremamente basso ma significativo (0.0170,  $P=0.00188$ ) tra

popolazioni entro gruppi. Tale risultato, riscontrato solo con i microsatelliti, è molto importante, in quanto permette di confermare la presenza di differenziamento anche tra i campioni di popolazione dell'Adriatico e dello Ionio, che altrimenti non sarebbe stato possibile osservare con il solo marcatore mitocondriale. Il fatto che i due marcatori abbiano dato risultati diversi non deve sorprendere: i microsatelliti, rispetto al DNA mitocondriale, sono dotati di un maggior potere risolutivo e probabilmente risiede proprio in questa diversa capacità il motivo di tale incongruenza.

Tuttavia, il debole differenziamento tra Adriatico e Ionio evidenziato dai microsatelliti può essere in parte atteso se si prendono in considerazione alcuni aspetti e caratteri di *life history* di *C. aestuarii*. Questa specie, infatti, è caratterizzata da una fase larvale pelagica altamente dispersiva. Spesso, negli organismi marini, bassi livelli di differenziamento genetico sono associati ad un'ampia dispersione legata proprio agli stadi larvali pelagici e questo è stato riscontrato in molte specie in cui gli stadi larvali durevoli nel tempo permettono un'ampia dispersione geografica degli individui (Avise, 2004). In particolare, recentemente sono stati effettuati diversi studi sulla struttura di popolazione di alcune specie di granchi e i risultati ottenuti in questa tesi sembrano correlare con quanto riportato in letteratura. Ad esempio, Roman e Palumbi (2004) hanno analizzato sequenze di mtDNA della specie atlantica *Carcinus maenas* per verificare la presenza di struttura genetica in popolazioni dell'Atlantico nord-orientale. In questa specie, filogeneticamente vicina a *C. aestuarii* con la quale condivide molte caratteristiche morfologiche ed ecologiche, gli autori hanno riscontrato un valore medio di  $F_{st}$  di 0.012 tra popolazioni separate mediamente da circa 500 km, dato indicativo di un debole differenziamento genico. Un risultato simile è stato ottenuto anche nel granchio *Erimacrus isenbeckii* (Azuma *et al.*, 2008): livelli molto bassi di differenziamento sono stati riscontrati tra campioni prelevati lungo la costa nord-orientale dell'isola di Hokkaido. Uno studio, inoltre, volto ad indagare la genetica di popolazione del granchio delle mangrovie brasiliane *Ucides cordatus*, utilizzando la regione di controllo del DNA mitocondriale, non ha evidenziato alcuna struttura geografica, nonostante la distanza tra i siti di campionamento più lontani superasse i 3000 km (Oliveira-Neto *et al.*, 2007). Anche nel granchio *Eriocheir sinensis* sono stati riscontrati livelli di differenziamento molto bassi: utilizzando 6 loci microsatellite, Herborg e collaboratori (2007) hanno evidenziato valori di  $F_{st}$  del tutto comparabili con quelli ottenuti in questa tesi per le popolazioni adriatiche e ionica. In particolare, sono stati osservati indici  $F_{st}$  pari a zero per campioni di popolazione distanti circa 50 km, pari a 0.005 per campioni separati da circa 1500 km.

In base ai risultati ottenuti in questa tesi e negli studi sopra citati, come già detto in precedenza, si può ritenere che, in generale, i granchi siano caratterizzati da bassi livelli di struttura genetica e che questo possa essere determinato da alcune caratteristiche di *life history* che li accomunano. Infatti, *C. aestuarii* e le altre specie di granchi citate sono caratterizzati da una fase larvale altamente dispersiva. Studi effettuati in laboratorio su *C. maenas* hanno evidenziato come il periodo larvale di questa specie vari da 17-25 giorni a 25° C a 44-80 giorni a 12° C (Williams, 1968). In letteratura non sono presenti lavori simili relativi a *C. aestuarii*, ma si può ipotizzare che le larve di questa specie posseggano tempi di sopravvivenza simili a quelli riscontrati in *C. maenas*. Ovviamente, se questa ipotesi fosse vera, bisognerebbe tener presente che si tratta pur sempre di dati ottenuti in laboratorio, non considerando tutte le variabili che potrebbero entrare

in gioco in ambiente naturale. Generalmente, la dispersione larvale è correlata positivamente con la durata dello stadio larvale e di conseguenza anche con il flusso genico di una specie. Migrazioni larvali di *C. aestuarii* potrebbero realizzarsi passivamente mediante lo sfruttamento delle correnti presenti in Adriatico (figura 4.1).



**Figura 4.1:** rappresentazione schematica della circolazione superficiale del Mar Adriatico. (NAd=Northern Adriatic; MAd=Middle Adriatic; SAd=Southern Adriatic; E=Eastern; W=Western, da Artegiani *et al.*, 1997).

Prendendo in considerazione le caratteristiche oceanografiche del Mar Adriatico si può notare come la circolazione generale del bacino sia caratterizzata da un andamento di tipo antiorario (Zavatarelli *et al.*, 1998). Infatti, il deflusso del Po richiama acque ad elevata salinità dalla soglia di Otranto; queste scorrono lungo il margine della ex-Jugoslavia, giungono nel bacino nord per poi defluire lungo il margine della penisola italiana (circolazione ciclonica). La circolazione presente in Adriatico e l'alto grado di dispersione larvale, quindi, potrebbero spiegare lo scarso differenziamento genetico tra le popolazioni adriatiche e ioniche esaminate in questa tesi.

A tale scopo, sono state effettuate varie simulazioni con il programma *Migrate 2.3* (Beerli and Felsenstein, 2001) volte ad indagare e capire il tipo di flussi migratori, e i relativi tassi, presenti tra i campioni adriatici analizzati. Ne è emerso un quadro piuttosto complesso. Infatti, partendo da un modello a isole in cui si assume che la migrazione possa avvenire tra tutte le popolazioni, senza tener conto della loro disposizione spaziale, si osserva che i flussi migratori vanno tendenzialmente da sud verso nord, come ad esempio accade nell'alto Adriatico. Qui, infatti, prevale una migrazione di tipo antiorario, con i campioni di Venezia, Goro e Ravenna che scambiano un numero consistente di individui in direzione di Marano (tabella 3.16 dei risultati). Un tale risultato appare, però, in contrasto con quanto atteso dal pattern di circolazione presente in Adriatico, in cui dovrebbe prevalere un flusso nord-sud. Solo in alcuni casi gli scambi tendono ad invertirsi.

Quello che si può ipotizzare è che nella parte settentrionale del bacino predomini la corrente ciclonica del NAd gyre che favorirebbe così un ritorno dei campioni verso il sito più a nord (Marano), piuttosto che un flusso da nord a sud lungo la costa. Inoltre, sempre secondo questo modello, è apparso evidente come i campioni meridionali (Lesina e Aquatina) contribuiscano in maniera

considerevole a mantenere il flusso sud-nord, scambiando un numero cospicuo di individui con Marano, Venezia e Ravenna. Nel fare questo, potrebbe essere sfruttata la corrente superficiale adriatica, che, risalendo lungo la costa croata, contribuirebbe a trasportare i campioni fino alle lagune settentrionali.

In ogni caso, sebbene sia emersa una migrazione asimmetrica tra i campioni analizzati, una situazione come quella proposta dal modello a isole appare inattesa e soprattutto irrealistica, in quanto non si tiene conto della distanza geografica tra i singoli campioni di popolazione.

Nel tentativo, quindi, di semplificare e di rendere più comprensibili i meccanismi alla base del flusso di migrazione, si è passati a testare un modello modificato di *stepping stone*. Infatti, oltre a considerare una migrazione solo tra popolazioni adiacenti, si è permesso che gli scambi si realizzassero, in maniera unidirezionale, anche tra i campioni meridionali e quelli di Marano. Secondo questo modello i flussi di migrazione tra i campioni di popolazione analizzati coincidono con le correnti superficiali antiorarie dell'Adriatico, andando da nord a sud. Inoltre, non solo è confermata la situazione presente nell'alto Adriatico, con Marano che riceve un elevato numero di individui provenienti da Venezia (come osservato nel modello a isole), ma pare che le popolazioni meridionali rappresentino effettivamente un contributo per le popolazioni settentrionali, scambiando, in maniera cospicua, individui con Marano. Ciò sarebbe dovuto alle correnti cicloniche adriatiche che, risalendo lungo la costa croata, porterebbero nuovi individui alla popolazione di Marano. Questa stessa situazione viene confermata anche considerando il solo contributo da parte di Aquatina: non solo sembra confermata la tendenza generale antioraria del flusso migratorio, ma pare sia fondata l'ipotesi di un reale contributo delle popolazioni meridionali verso quelle a nord.

I risultati ottenuti con il programma *Migrate*, seppure non siano di facile interpretazione, sembrano quindi avvalorare l'ipotesi secondo cui le migrazioni larvali di *C. aestuarii* possano realizzarsi passivamente sfruttando le correnti presenti in Adriatico: questi due fattori combinati assieme potrebbero spiegare il debole differenziamento riscontrato tra le popolazioni adriatiche e ionica esaminate in questo lavoro.

Uno scenario simile è stato proposto per spiegare la ridotta struttura genetica presente in popolazioni adriatiche e ioniche di *Atherina boyeri* (Congiu *et al.*, 2002), come pure il basso livello di differenziamento genico di popolazioni adriatiche di triglia rossa (*Mullus barbatus*) (Garoia *et al.*, 2004a): in entrambi i casi, si ipotizza che la dispersione passiva degli stadi larvali, sfruttando le correnti oceanografiche, possa rappresentare il motivo preponderante della debole strutturazione genetica di tali specie. Garoia e i suoi collaboratori (2004b), inoltre, hanno evidenziato come l'Adriatico centro-settentrionale sia caratterizzato da elevati livelli di omogeneità idrogeografica e che la scarpata continentale presente nella parte meridionale di tale bacino non rappresenti un ostacolo al flusso genico delle popolazioni di cefalopodi da loro analizzate (*Loligo vulgaris* e *Sepia officinalis*).

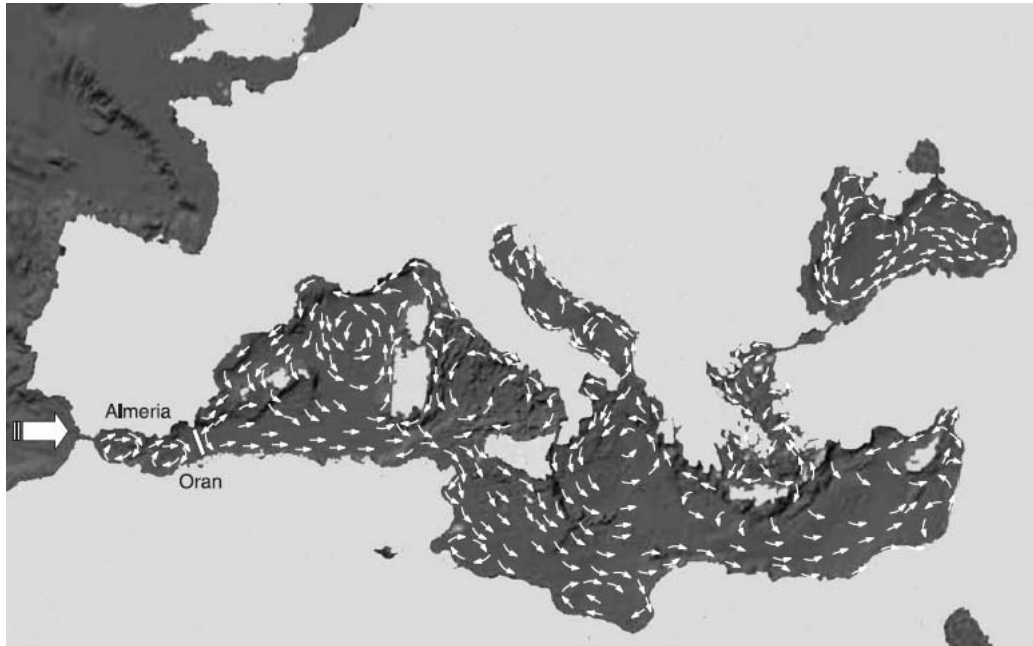
È importante comunque sottolineare che una eventuale dispersione passiva delle larve con le correnti non rende completamente omogenee le popolazioni adriatiche e ioniche di *C. aestuarii* studiate in questa tesi. Infatti, poiché sono state trovate lievi differenze tra tali popolazioni, sia con i test di differenziamento genico che con il calcolo degli indici  $F_{st}$  e dell'*AMOVA*, bisogna supporre che altri fattori intervengano a mantenere un'identità precisa per ogni popolazione.

Fenomeni di ritenzione larvale, ad esempio, potrebbero essere presenti in tale specie, facendo in modo che la percentuale maggiore di larve ritorni al sito di origine e solo una piccola parte venga trasportata passivamente dalle correnti. Come anticipato nell'introduzione di questa tesi, il fenomeno della ritenzione larvale è ampiamente diffuso negli organismi marini: lavori condotti sia su vertebrati che invertebrati marini (Knowlton and Keller, 1986; Jones *et al.*, 1999; Swearer *et al.*, 1999; Taylor and Hellberg, 2003) hanno dimostrato che una buona percentuale degli stadi larvali e giovanili può reclutare in prossimità del sito di origine delle popolazioni natali.

Inoltre, i *gyre* presenti nei sottobacini dell'Adriatico potrebbero in parte limitare il flusso genico delle popolazioni analizzate, costringendo una percentuale delle larve a restare vicine al sito d'origine.

Per quanto riguarda il campione di Cagliari, invece, è stato già evidenziato il marcato differenziamento nei confronti dei restanti campioni di popolazione. Tale risultato è stato confermato sia dall'analisi effettuata con il mtDNA sia da quella con i microsatelliti. Si potrebbe ipotizzare che anche per il campione cagliaritano la circolazione delle acque tirreniche (figura 4.2) ricopra un ruolo determinante, impedendo in questo modo che si realizzi un elevato flusso genico con le popolazioni adriatiche e ionica. Infatti, la corrente atlantica, che entra da Gibilterra, continua la sua corsa verso est in quella che è chiamata corrente Algerina; questa, in corrispondenza delle coste tunisine, si amplia e si divide in due rami. Il primo ramo risale tutto il Tirreno fino al Mar Ligure per poi procedere verso ovest; il secondo prosegue la sua corsa verso il Mediterraneo orientale attraverso il canale di Sicilia.

Se realmente la circolazione mediterranea ricopre un ruolo nel determinare il differenziamento tra il campione tirrenico e il resto delle popolazioni analizzate, allora le larve della popolazione cagliaritano di *C. aestuarii* potrebbero essere influenzate dalla corrente che sale verso nord e rimanere confinate in un moto circolare per l'intera durata della loro vita planctonica determinando un relativo isolamento di tale popolazione e conseguente differenziamento genetico con le altre popolazioni adriatiche e ionica.



**Figura 4.2:** rappresentazione schematica della circolazione del Mar Mediterraneo (da Patarnello *et al.*, 2007).

Anche nella specie atlantica *C. maenas* è stata riscontrata una simile situazione (Roman and Palumbi, 2004): è stato evidenziato che popolazioni presenti nelle Isole Faeroe e in Islanda risultano geneticamente separate dalle popolazioni continentali dell'Atlantico orientale. Gli autori ipotizzano che la causa di un tale differenziamento sia da attribuire alla barriera d'acqua profonda e alla direzione delle correnti che impediscono il flusso genico tra le popolazioni isolate e quelle continentali. Inoltre, i risultati di un'analisi effettuata sul granchio *Erimacrus isenbeckii* (Azuma *et al.*, 2008) mostrano differenze tra popolazioni del Mar del Giappone e popolazioni campionate lungo la costa pacifica del Giappone, probabilmente causate dalle correnti marine che ostacolano il flusso genico tra tali popolazioni.

In accordo con l'ipotesi di un differenziamento associato alle correnti oceanografiche del Mediterraneo, è stato condotto uno studio sulla struttura genetica del latterino (*Atherina boyeri*) in Mediterraneo. Tale lavoro ha evidenziato come popolazioni della Sardegna mostrano una rilevante distanza genetica nei confronti delle popolazioni adriatico-ioniche ma anche nei confronti di popolazioni situate lungo la costa continentale tirrenica italiana, probabilmente proprio a causa della barriera al flusso genico determinata dalle correnti (Congiu *et al.*, 2002).

In base ai risultati ottenuti in questa tesi per la parte di genetica di popolazione, come già ampiamente riportato, si è osservato un debole ma significativo differenziamento tra campioni di popolazione che può essere attribuito all'elevato potere dispersivo dello stadio larvale e alle correnti oceanografiche dell'Adriatico e del Mediterraneo. Non è da escludere, comunque, che fattori come la ritenzione larvale possano contribuire a mantenere un'identità precisa per ogni campione di popolazione.

Per quanto riguarda gli aspetti di filogeografia e di demografia storica, le analisi effettuate in questo lavoro hanno evidenziato come tutte le popolazioni di *C. aestuarii* analizzate abbiano subito, in passato, fenomeni di espansione. In particolare, i test di neutralità-equilibrio, le *mismatch distribution* e i *bayesian skyline plot* hanno permesso di trarre indicazioni riguardo ai fenomeni demografici avvenuti in passato. Infatti i test di neutralità di Tajima ( $D$ , Tajima, 1989) e Fu ( $F_s$ , Fu, 1997), inizialmente introdotti allo scopo di confermare l'ipotesi di neutralità e la presenza di selezione, permettono di verificare l'ipotesi di equilibrio tra mutazione e deriva, di trarre indicazioni sui processi evolutivi che agiscono sulle popolazioni studiate e di investigare la storia dell'espansione demografica. Valori negativi statisticamente significativi di  $D$  e  $F_s$  possono essere interpretati come un segnale di espansione di popolazione.

Le statistiche  $D$  e  $F_s$  stimate per gli 8 campioni di popolazione analizzati in questa tesi sono risultate, in tutti i casi, negative e significative, il che ha permesso di rigettare l'ipotesi di neutralità selettiva e di popolazione in equilibrio per gli aplotipi esaminati.

Tali statistiche possono deviare dal valore atteso di zero a causa di diversi fattori che possono influenzare differenzialmente i due parametri da cui vengono calcolate, la media delle differenze a coppie e il numero di siti segreganti. Alcuni fattori, quali l'omoplasia (Bertorelle and Slatkin, 1995) e la suddivisione in popolazioni (Simonsen *et al.*, 1995) tendono a spostare i valori di  $D$  e  $F_s$  verso valori positivi. In *C. aestuarii*, queste statistiche assumono sempre valori negativi, per cui si può escludere che questi fenomeni siano determinanti nell'evoluzione di tale specie. Rimane come possibile causa una variazione delle dimensioni di popolazione.

Secondo il modello di Rogers e Harpending (1992), un'espansione demografica di popolazione produce una distribuzione unimodale delle differenze a coppie. In effetti, come riportato nei risultati di questa tesi, la distribuzione osservata per le popolazioni adriatiche e ionica risulta avere un andamento di questo tipo. Il valore medio di 2.345 differenze e la varianza di 2.353 rendono la distribuzione osservata molto diversa da quella di Poisson, attesa per una popolazione di dimensioni costanti. Seguendo il modello di Rogers ed Harpending, l'espansione demografica di queste popolazioni di *C. aestuarii* dovrebbe avere avuto inizio circa 200000 anni fa.

La *mismatch distribution* relativa al campione di Cagliari, invece, ha prodotto una curva leggermente diversa: infatti, il valore medio di 2.825 differenze e la varianza di 29.559 rendono tale curva bimodale e potrebbero indurre a interpretarla come un segnale di popolazione costante. Tale distribuzione, in realtà, è dovuta esclusivamente alla presenza di due aplotipi rari (Cag24 e Cag44) molto divergenti a livello di sequenza, che si collocano a livello basale nel cluster corrispondente alla specie *C. aestuarii* nell'albero consenso (figura 3.6) riportato nei risultati di questa tesi.

I *bayesian skyline plot*, ottenuti con il programma *BEAST*, si sono rivelati un valido strumento per chiarire il tipo di variazione demografica e i processi evolutivi subiti dai campioni adriatici e ionico, ma soprattutto dal campione tirrenico. Sviluppati da Pybus e collaboratori nel 2000, sono attualmente il metodo più frequentemente utilizzato per investigare la storia delle espansioni demografiche di vari organismi (Crandall *et al.*, 2008; Nicolas *et al.*, 2008; Zarza *et al.*, 2008).

Per quanto riguarda i campioni dell'Adriatico e dello Ionio, i *bayesian skyline plot* hanno confermato l'espansione demografica subita da tali popolazioni, che si colloca approssimativamente in un intervallo di tempo (compreso tra 200000 e 100000 anni fa) che coincide con l'inizio dell'espansione stimato dalla *mismatch distribution*.

Se queste stime sono realistiche, anche considerando l'errore associato, un evento demografico con un forte impatto sulla diversità genetica di *C. aestuarii* dovrebbe essere avvenuto prima che il Mar Adriatico assumesse la morfologia che attualmente lo contraddistingue. Infatti, come riportato in introduzione, le lagune che oggi si possono osservare in Adriatico dovrebbero aver preso forma alla fine della glaciazione di Wurm (ultima glaciazione pleistocenica), circa 10000 anni fa. Da questo si potrebbe dedurre che le popolazioni di *C. aestuarii* non abbiano risentito dal punto di vista genetico della storia recente e dei cambiamenti climatici avvenuti nel Mediterraneo.

Questi valori sono in linea con quelli ottenuti analizzando altre specie del Mediterraneo. Patarnello e collaboratori (2007) riportano che in *Pomatoschistus microps* una probabile espansione sia avvenuta mediamente 67000 anni fa (valore compreso tra 31000 e 102000 anni fa); in *Pagellus bogaroveo* dovrebbe essere avvenuta mediamente 152000 anni fa (valore compreso tra 47000 e 258000 anni fa); in *Sagitta setosa* circa 81000 anni fa (valore compreso tra 2000 e 161000 anni fa).

Per il campione di Cagliari, invece, la situazione risulta diversa: infatti, dai *bayesian skyline plot* risulta che la popolazione è rimasta costante per un intervallo di tempo molto lungo (da 900000 a 100000 anni fa) e solo recentemente abbia subito un'espansione demografica. Questo risultato collocherebbe, quindi, l'espansione subita dai campioni cagliaritari in un periodo di tempo che coincide con l'ultima glaciazione pleistocenica.

In ogni caso, le analisi effettuate in questo lavoro indicano la presenza di espansioni, anche al di fuori del bacino Adriatico.

È possibile, quindi, che in generale le popolazioni di organismi marini mediterranei non siano state influenzate in modo significativo dalle ultime glaciazioni, ma rechino traccia di eventi più antichi.

Si può ipotizzare che le popolazioni adriatiche di *C. aestuarii*, durante l'ultima glaciazione, siano migrate più a sud per far fronte all'avanzamento dei ghiacci e progressivamente abbiano ricolonizzato i loro abituali areali in seguito al miglioramento delle condizioni ambientali. È plausibile supporre che tale ricolonizzazione sia avvenuta in maniera graduale e massiva, senza necessariamente aver subito fenomeni di colli di bottiglia legati ad una drastica riduzione delle dimensioni della popolazione.

Resta da capire, invece, che tipo di processi evolutivi e di storia demografica hanno potuto influire sul campione cagliaritano. Presumibilmente, visti i risultati ottenuti, una pronunciata variazione nelle dimensioni di popolazione può essere avvenuta durante l'ultima espansione glaciale pleistocenica (glaciazione di Wurm, 110000-10000 anni fa). Fino a quel momento sembra che il campione si sia mantenuto di dimensioni costanti, dato che può essere ritenuto verosimile se si considera la storia geologica della Sardegna. Questa, infatti, assieme alla Corsica, si è separata dalla Penisola Iberica circa 29 milioni di anni fa, mentre la loro rotazione e successiva divisione sono avvenute in un intervallo di tempo compreso tra 15 e 9 milioni di anni fa (Boccaletti *et al.*, 1990; Ketmaier *et al.*, 2006). In seguito, circa 5 milioni di anni fa come suggerito da Meulenkamp e

Sissingh (2003), sembra che la microplacca Sardo-corsa sia rimasta in connessione con l'Europa attraverso un ponte di terra che la collegava alla Toscana. Da quel momento, le due isole si sono definitivamente separate e hanno mantenuto la morfologia che le contraddistingue tuttora, con qualche variazione di estensione associata al cambiamento del livello del mare.

È plausibile, quindi, che il campione di Cagliari rechi traccia di questi eventi geologici e abbia mantenuto una dimensione costante di popolazione proprio per tali motivi. Successivamente, con l'avanzamento dei ghiacci durante l'ultima glaciazione, si può supporre che tale campione si andato incontro ad un drastico declino demografico seguito da una rapida espansione. Una possibile spiegazione può essere ricercata nell'assenza, più a sud, di adeguati areali in cui potersi rifugiare: se realmente si è realizzata una situazione del genere, è probabile che il campione cagliaritano abbia subito successivi fenomeni di collo di bottiglia, che ne hanno ridotto drasticamente la dimensione di popolazione. Con il miglioramento delle condizioni ambientali, poi, il campione avrebbe ricolonizzato le coste meridionali della Sardegna, andando incontro ad una rapida espansione di popolazione.

Sulla base dei risultati ottenuti per la parte di filogeografia e demografia storica, come largamente descritto, sono stati osservati fenomeni di espansione in tutti i campioni di popolazione considerati. Tali fenomeni risultano antecedenti alle ultime glaciazioni, e quindi anche alla formazione delle lagune adriatiche nel caso dei campioni adriatici e ionico, mentre sembrano coincidere con l'ultima espansione glaciale nel caso del campione tirrenico.

Complessivamente, i risultati conseguiti durante i tre anni di dottorato suggeriscono che la distribuzione della variabilità genetica nei campioni di popolazione analizzati sia fortemente influenzata dalle correnti oceanografiche e dalle barriere presenti in Adriatico, e più in generale nel Mediterraneo, piuttosto che dalla storia recente del Mediterraneo e dai cambiamenti climatici avvenuti nel passato.

## 4.1 Analisi comparata

I risultati di differenziamento ottenuti in questa tesi nella specie *Carcinus aestuarii* possono essere messi a confronto con quelli di altri organismi lagunari presenti comunemente in Adriatico e, più in generale, in Mediterraneo.

In questa sede, verranno analizzati alcuni aspetti della genetica di popolazione di due specie di teleostei tipicamente lagunari: il go (*Zosterisessor ophiocephalus*, Pallas, 1814) e il latterino (*Atherina boyeri*, Risso, 1810).

Il go, o ghiozzo, rappresenta l'unica specie del genere *Zosterisessor* appartenente alla famiglia Gobiidae. Specie eurialina, diffusa in acque marine costiere e salmastre, è caratteristica delle lagune e degli estuari del Mediterraneo, ma anche del Mar Nero e del Mare d'Azov (Miller, 1979). Non tollera acque eccessivamente dolci ma sopporta bene le basse temperature (euritermia). L'habitat tipico è costituito da fondali a substrato misto di sabbia e fango, dove sono presenti praterie di fanerogame marine (genere *Zostera* e *Posidonia*). Ha un ciclo vitale pluriennale (in media circa 3 anni, Mazzoldi *et al.*, 2000), che si svolge interamente all'interno della laguna e, nella laguna di Venezia, dove rappresenta il gobide più grande e abbondantemente distribuito, sembra diventare sessualmente maturo ad un anno di età. Durante il periodo riproduttivo, compreso tra marzo e maggio, i maschi più grandi (tipicamente individui di 2-3 anni) costruiscono una tana in cui una o più femmine depongono le uova.

Molte delle caratteristiche sopra descritte accomunano *Zosterisessor ophiocephalus* a *Carcinus aestuarii*: anche nel teleosteo gli adulti sono relativamente poco mobili, mentre lo stadio larvale, della durata di circa 28 giorni, è caratterizzato da un elevato potere dispersivo.

Per indagare la variabilità genetica di questa specie, analogamente a quanto fatto per *C. aestuarii*, è stato effettuato uno studio utilizzando i marcatori microsatellite (Bernardinelli, 2004). In tale lavoro, le analisi molecolari, che hanno previsto l'impiego di 13 loci microsatellite, sono state condotte su un numero complessivo di 236 individui, prelevati nelle lagune di Marano, Venezia e Goro.

I test di differenziamento e gli indici  $F_{st}$  calcolati per i tre campioni di popolazione hanno evidenziato una totale assenza di diversità genetica: le probabilità di omogeneità stimate per Marano, Venezia e Goro, infatti, sono risultate tutte non significative, come pure le probabilità associate agli indici  $F_{st}$ .

Al fine di comprendere il motivo di una tale omogeneità del campione, in questa tesi di dottorato è stato rianalizzato l'intero dataset di go, includendo nell'analisi anche il calcolo dell'AMOVA e dei tassi di migrazione. L'analisi molecolare della varianza, condotta per i tre campioni raggruppati assieme, ha restituito un valore di variazione entro popolazione pari a 1, mentre un valore negativo e non significativo (-0.00048,  $P=0.53$ ) è stato ottenuto per la variazione tra popolazioni. Questi risultati, quindi, dimostrano come effettivamente i campioni provenienti dalle diverse lagune siano caratterizzati da una complessiva omogeneità, dato che conferma quanto riscontrato in *C. aestuarii* per l'alto Adriatico.

Per quanto riguarda il tipo di migrazione e i relativi tassi, dai risultati ottenuti con *Migrate 2.3* è emerso un quadro leggermente diverso rispetto a quello ottenuto nel granchio verde, sia utilizzando il modello a isole che quello *stepping stone* modificato. Pare, infatti che per *Z. ophiocephalus* la tendenza generale dei flussi migratori sia da nord a sud, con un'unica eccezione rappresentata dal campione di Goro che scambia un numero maggiore di individui in direzione di Venezia.

Colpisce soprattutto il fatto che nel go non si osserva quel flusso da Venezia a Marano che invece caratterizza le migrazioni di *C. aestuarii*. Il motivo di un tale risultato potrebbe essere ricondotto alla diversa durata della fase larvale di questa specie rispetto al granchio verde.

In ogni caso pare che, anche per il go, la dispersione larvale e le correnti oceanografiche presenti nell'alto Adriatico possano rendere conto del rimescolamento degli individui delle varie lagune, minimizzando in questo modo ogni differenziamento genetico.

Il latterino, *Atherina boyeri*, può essere considerato un valido modello di studio delle lagune, essendo un tipico rappresentante della fauna ittica di questi ambienti in tutto il Mediterraneo. È una specie appartenente alla famiglia Atherinidae che comprende 60 specie, raggruppate in 12 generi (Nelson, 2006). È un organismo gregario caratterizzato da un'ampia valenza ecologica, grazie alla quale è in grado di tollerare differenze, anche rilevanti, sia nella temperatura che nella salinità, e quindi di colonizzare diversi tipi di ambienti: acque marine costiere, estuari e laghi costieri salmastri, tratti bassi dei corsi d'acqua, laghi interni e laghi ipersalini (Henderson and Bamber, 1987). In generale, la preferenza va alle acque poco correnti e ricche di vegetazione, sufficientemente limpide e ossigenate (Zerunian, 2002). Trascorre buona parte del ciclo vitale all'interno dell'ambiente lagunare (Henderson and Bamber, 1987): infatti, sia la fase riproduttiva che quella larvale si svolgono in ambiente lagunare, mentre, durante la fase adulta, nei periodi invernali, alcuni individui escono in mare (Henderson and Bamber, 1987; Henderson *et al.*, 1988; Mistri and Colombo, 1988; Bardin and Pont, 2002). Tuttavia la maggior parte di essi rimane in prossimità della laguna di origine, in cui rientrano per il periodo riproduttivo, come dimostrato da studi morfologici (Berrebi and Britton-Davidian, 1980), per cui la specie mostra una bassa capacità di dispersione.

L'interesse a capire quanto le popolazioni lagunari di latterino siano tra loro differenziate e isolate geneticamente ha portato allo sviluppo di numerosi lavori: analisi molecolari, ad esempio, condotte lungo le coste italiane (Congiu *et al.*, 2002) e all'interno di tutto il bacino Mediterraneo (Astolfi *et al.*, 2005) hanno riportato la presenza di una elevata strutturazione genetica tra le diverse popolazioni, tra le quali tuttavia permane un basso livello di flusso genico che gli autori correlano alle abitudini anadrome della specie. Simili risultati sono stati ottenuti, inoltre, in uno studio morfologico/molecolare condotto lungo le coste mediterranee di Francia, Corsica e Tunisia (Trabelsi *et al.*, 2004). Gli autori identificano la presenza di due gruppi di popolazioni ben differenziati all'interno del bacino mediterraneo, che suggeriscono di interpretare come specie sorelle o almeno come sottospecie o semispecie all'interno della forma lagunare di *A. boyeri*.

Recentemente, è stato effettuato uno studio volto ad indagare a) la filogenesi di *A. boyeri*, allo scopo di chiarire l'esistenza di tre presunte forme/specie proposte per il latterino e b) la filogeografia e la variabilità genetica tra le popolazioni della forma lagunare, per indagare le relazioni esistenti tra di esse e verificarne il grado di strutturazione (Milana, 2008). Sebbene in questo lavoro sia stato preso in considerazione un numero maggiore di sito di campionamento, che comprendevano ad esempio lo Stagnone di Marsala e varie lagune della costa laziale, anche in questo caso, analogamente a quanto effettuato con *C. aestuarii*, è stato previsto l'impiego di due marcatori molecolari, il DNA mitocondriale, per il

quale sono state analizzate tre differenti regioni (citocromo b, D-loop e un'inserzione), e i microsatelliti.

Sia i risultati del calcolo degli  $F_{st}$  a coppie, che del test di differenziamento evidenziano la presenza di strutturazione. In particolare, le statistiche F di Wright (1951) mostrano indici di strutturazione piuttosto elevati, con un valore significativo ( $P < 0.001$ ) di  $F_{st}$  totale per tutte le popolazioni di 0.0633. I valori ottenuti dal confronto a coppie sono compresi tra un minimo di 0.0114 per Chioggia-Lesina, paragonabile ad alcuni confronti riscontrati in *C. aestuarii* su distanze geografiche simili, e un massimo di 0.1102 per Trasimeno-Stagnone. Anche per quanto riguarda i test di differenziamento genico e genotipico, si ottengono valori di probabilità altamente significativi per tutti i confronti considerati ( $P = 0.0000$ ), ad indicare una disomogeneità nella distribuzione delle frequenze alleliche e genotipiche. Inoltre, in accordo con i risultati ottenuti in questa tesi per il granchio verde, è stata verificata la presenza di isolamento per distanza: il test di Mantel, infatti, ha restituito un coefficiente di correlazione pari a 0.53 statisticamente significativo ( $P < 0.005$ ), dimostrando ulteriormente la presenza di differenziamento tra i campioni analizzati. Anche l'analisi molecolare della varianza (AMOVA) ha confermato quanto già ottenuto dai precedenti test: una quota significativa della variazione è proprio dovuta alla suddivisione in gruppi. Complessivamente, quindi, è stato chiaramente evidenziato un elevato livello di differenziamento tra le popolazioni delle diverse lagune, come già riportato in lavori precedenti (Focant *et al.*, 1999; Congiu *et al.*, 2002; Astolfi *et al.*, 2005).

Tale differenziamento, in accordo a quanto proposto da Astolfi e collaboratori (2005), potrebbe essere spiegato dalla presenza di un basso flusso genico tra le popolazioni, dovuto ad una bassa capacità di dispersione, che comporta un rientro massiccio di individui nella laguna di origine (Berrebi and Britton-Davidian, 1980).

Dai risultati ottenuti sia in *Z. ophiocephalus* che in *A. boyeri* è possibile affermare che, come per *C. aestuarii*, un ruolo importante nella determinazione di un differenziamento, più o meno pronunciato, è ricoperto dagli stadi larvali di queste due specie lagunari e dalle correnti oceanografiche presenti in Mediterraneo.

## 4.2 Conclusioni

In questa tesi di dottorato, attraverso l'impiego di marcatori mitocondriali e microsatellite sono stati analizzati campioni di *Carcinus aestuarii* provenienti da sei differenti siti lagunari dell'Adriatico, un sito ionico e uno tirrenico.

Le analisi effettuate hanno evidenziato la presenza di una debole ma significativa strutturazione genetica all'interno di questa specie; il basso livello di differenziamento è stato attribuito principalmente all'elevato potere dispersivo dello stadio larvale e alle correnti oceanografiche che caratterizzano l'Adriatico e, più in generale, il Mediterraneo. Tuttavia, non è da escludere che altri fattori, come ad esempio fenomeni di ritenzione larvale, possano intervenire a spiegare tale risultato. In ogni caso, è importante sottolineare come il debole differenziamento riscontrato in *C. aestuarii* sia comunque in linea con quanto riportato in letteratura su specie affini, ma anche su altri organismi lagunari, come nel caso del gobide *Zosterisessor ophiocephalus* e del latterino *Atherina boyeri*.

È stato inoltre dimostrato come i campioni di popolazione analizzati in questa tesi abbiano subito, in passato, fenomeni di espansione demografica: questi, sulla base dei risultati ottenuti, dovrebbero collocarsi in un periodo di tempo di molto antecedente l'ultima glaciazione pleistocenica. Evidentemente, ciò che si osserva ora a livello di DNA è, con molta probabilità, una traccia di qualche evento ancora più antico: la storia recente del Mediterraneo, quindi, non sembra avere un ruolo determinante nella distribuzione della variabilità genetica di *C. aestuarii*. Si ritiene, piuttosto, che siano state le correnti oceanografiche e alcune delle caratteristiche di *life history* della specie a influenzare maggiormente questo fattore.

Sarebbe interessante, in futuro, riuscire ad estendere l'analisi a campioni provenienti dalla costa croata e albanese dell'Adriatico, al fine di chiarire i meccanismi che rendono estremamente basso il differenziamento genico tra le popolazioni adriatiche e ioniche, così da avere una visione generale della genetica di popolazione di *C. aestuarii* per l'intero bacino.

Inoltre, sarebbe opportuno analizzare anche campioni provenienti dalla costa tirrenica italiana per capire se effettivamente vi è omogeneità tra le popolazioni del Tirreno o se invece esista una barriera al flusso genico tra i campioni isolani e quelli continentali di *C. aestuarii*. Tale analisi permetterebbe inoltre di far luce sui fenomeni di espansione avvenuti in questo bacino, chiarendo così se la situazione osservata per il campione di Cagliari è comune a tutti i campioni tirrenici o se sia, invece, esclusiva dell'isola.

Per quanto riguarda le altre specie lagunari prese in considerazione, sarebbe interessante ottenere campioni di *Z. ophiocephalus* dalle stesse lagune da cui sono stati raccolti gli esemplari di *C. aestuarii* per poter verificare che tipo di migrazioni sussistono e se corrispondono a quelle riscontrate nel crostaceo.

Anche nel caso di *A. boyeri* si potrebbero condurre dei test volti a verificare e quantificare i tassi di migrazione in modo tale da avere un quadro più completo della genetica di popolazione di specie che occupano abitualmente le lagune costiere del Mediterraneo e trarre eventuali indicazioni per la conservazione della biodiversità.

## ***5. BIBLIOGRAFIA***



- Amos W, Hoffman JI, Frodsham A, Zhang L, Best S, Hill AVS, (2007). Automated binning of microsatellite alleles: problems and solutions. *Molecular Ecology Notes*, **7**, 10-14.
- Artegiani A, Bregant D, Paschini E, Pinardi N, Raicich F, Russo A, (1997). The Adriatic sea general circulation. Part II: Baroclinic circulation structure. *Journal of Physical Oceanography*, **27**, 1515-1532.
- Asioli A, Trincardi F, Lowe JJ, Ariztegui D, Langone L, Oldfield F, (2001). Sub-millennial scale climatic oscillations in the central Adriatic during the Lateglacial: paleoceanographic implications. *Quaternary Science Reviews*, **20**, 1201-1221.
- Astolfi L, Dupanloup I, Rossi R, Bisol PM, Faure E, Congiu L, (2005). Mitochondrial variability of sand smelt *Atherina boyeri* populations from north Mediterranean coastal lagoons. *Marine Ecology Progress Series*, **297**, 233-243.
- Avise JC, (1987). Identification and interpretation of mitochondrial DNA stocks in marine species. In *Proceedings of the Stock Identification Workshop*, H. Kumpf and E. L. Nakamura (eds.). Plub. National Oceanographic and Atmospheric Administration, Panama City, FL, pp. 105-136.
- Avise JC, (2000). *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge (Massachusetts), London (England).
- Avise JC, (2004). *Molecular Markers, Natural History and Evolution. Second edition*.
- Avise JC, Arnold J, Ball RM Jr, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC, (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489-522.
- Avise JC, Ball RM, Arnold J, (1988). Current versus historical population size in vertebrate species with high gene flow: a comparison based on mitochondrial DNA lineages and inbreeding theory for neutral mutations. *Molecular Biology and Evolution*, **5**, 331-344.
- Avise JC, Giblin-Davidson C, Learn J, Patton JC, Lansman RA, (1979). Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, *Geomys pinetis*. *Proceedings of the National Academy of Sciences USA*, **76**, 6694-6698.
- Ayre DJ, Davis AR, Billingham M, Llorens T, Styan C, (1997). Genetic evidence for contrasting patterns of dispersal in solitary and colonial ascidians. *Marine Biology*, **130**, 51-61.
- Azuma N, Kunihiro Y, Sasaki J, Mihara E, Mihara Y, Yasunaga T, Jin DH, Abe S, (2008). Genetic variation and population structure of hair crab (*Erimacrus*

- isenbeckii*) in Japan inferred from mitochondrial DNA sequence analysis. *Marine Biotechnology*, **10**, 39-48.
- Bagley MJ, Franson SE, Christ SA, Waits ER, Toth GP, (2002). Genetic Diversity as an Indicator of Ecosystem Condition and Sustainability: Utility for Regional Assessments of Stream Condition in the Eastern United States. U.S. Environmental Protection Agency, Cincinnati, OH.
- Baldwin JD, Bass AL, Bowen BW, Clark WC, (1998). Molecular phylogeny and biogeography of the marine shrimp *Penaeus*. *Molecular Phylogenetics and Evolution*, **10**, 399-407.
- Bardin O, Pont D, (2002). Environmental factors controlling the spring immigration of two estuarine fishes *Atherina boyeri* and *Pomatoschistus* spp. into a Mediterranean lagoon. *Journal of Fish Biology*, **61**, 560-578.
- Beckmann JS, Weber JL, (1992). Survey of human and rat microsatellites. *Genomics*, **12**, 627-631.
- Beerli P, Felsenstein J, (1999). Maximum likelihood estimation of a migration matrix and effective population numbers in two populations by using a coalescent approach. *Genetics*, **152**, 763-773.
- Beerli P, Felsenstein J, (2001). Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences USA*, **98**, 4563-4568.
- Behrens Yamada S, Hauck L, (2001). Field identification of the european green crab species: *Carcinus maenas* and *Carcinus aestuarii*. *Journal of Shellfish Research*, **20**, 905-912.
- Belkir 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.
- Bernardinelli E, (2004). Variabilità genetica in *Zosterisessor ophiocephalus* di tre lagune dell'alto Adriatico. Tesi di Laurea in Scienze Naturali. Università degli Studi di Padova.
- Berrebi P, Britton-Davidian J, (1980). Enzymatic survey of four populations of *Atherina boyeri* based on electrophoresis and the occurrence of a microsporidiosis. *Journal of Fish Biology*, **16**, 149-157.
- Bertorelle G, Slatkin M, (1995). The number of segregating sites in expanding human-populations, with implications for estimates of demographic parameters. *Molecular Biology and Evolution*, **12**, 887-892.
- Boccaletti M, Ciaranfi N, Cosentino D, Deiana G, Gelati R, Lentini F, Massari F, Moratti G, Pescatore T, Ricci Lucchi F, Tortrici L, (1990). Palinspatic

- restoration and paleogeographic reconstruction of the peri-Tyrrhenian area during the Neogene. *Palaeogeography, Palaeoclimatology and Palaeoecology*, **77**, 41–50.
- Bondesan A, Meneghel M, (2003). Atti del Convegno "Studio Geoambientale del Territorio Provinciale di Venezia - parte centrale", Venezia 26 settembre 2003.
- Burton RS, (1983). Protein polymorphisms and genetic differentiation of marine invertebrate populations. *Marine Biology Letters*, **4**, 193-206.
- Burton RS, (1986). Evolutionary consequences of restricted gene flow among natural populations of the copepod, *Tigriopus californicus*. *Bulletin of Marine Science*, **39**, 526-535.
- Carlton JT, Choen AN, (2003). Episodic global dispersal in shallow water marine organisms: the case history of the European shore crabs *Carcinus maenas* and *C. aestuarii*. *Journal of Biogeography*, **30**, 1809-1820.
- Cassone BJ, Boulding EG, (2006). Genetic structure and phylogeography of the lined shore crab, *Pachygrapsus crassipes*, along the northeastern and western Pacific coasts. *Marine Biology*, **149**, 213-226.
- Castresana J, Lübben M, Saraste M, Higgins DG, (1994). Evolution of cytochrome oxidase, an enzyme older than atmospheric oxygen. *EMBO Journal*, **13**, 2516-2525.
- Clement M, Posada D, Crandall KA, (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657-1659.
- Congiu L, Rossi R, Colombo G, (2002). Population analysis of the sand smelt *Atherina boyeri* (Teleostei Atherinidae), from Italian coastal lagoons by random amplified polymorphic DNA. *Marine Ecology Progress Series*, **229**, 279-289.
- Craddock C, Hoeh WR, Lutz RA, Vrijenhoek RC, (1995). Extensive gene flow among mytilid (*Bathymodiolus thermophilus*) populations from hydrothermal vents of the eastern Pacific. *Marine Biology*, **124**, 137-146.
- Crandall ED, Jones ME, Muñoz MM, Akinronbi B, Erdmann MV, Barber PH, (2008). Comparative phylogeography of two seastars and their ectosymbionts within the Coral Triangle. *Molecular Ecology*, **17**, 5276-5290.
- Critto A, Marcomini A, (2001). Rischio ecologico e inquinamento chimico lagunare. *Libreria Editrice Cafoscarina*. Milano.
- Dawson AG, (1992). *Ice Age Earth: Late Quaternary Geology and Climate*. Routledge, London.

- Dawson MN, Louie KD, Barlow M, Jakobs DK, Swift CC, (2002). Comparative phylogeography of sympatric sister species, *Clevelandia ios* and *Eucyclogobius newberryi* (Teleostei, Gobiidae), across the California Transition Zone. *Molecular Ecology*, **11**, 1065-1075.
- DeWoody JA, Avise JC, (2000). Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *Journal of Fish Biology*, **56**, 461-473.
- Dias GM, Duarte LFL, Solferini VL, (2006). Low genetic differentiation between isolated populations of the colonial ascidian *Symplesma rubra* Monniot C, 1972. *Marine Biology*, **148**, 807-815.
- Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W, (2002). Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics*, **161**, 1307-1320.
- Drummond AJ, Rambaut A, (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.
- Drummond AJ, Rambaut A, Shapiro B, Pybus OG, (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, **22**, 1185-1192.
- Excoffier L, Laval G, Schneider S, (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47-50.
- Excoffier L, Smouse PE, Quattro JM, (1992). Analysis of molecular variance inferred from metric distances among DNA haplotype: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479-491.
- Field D, Wills C, (1996). Long polymorphic microsatellites in simple organisms. *Proceeding of the Royal Society of London, Series B*, **263**, 209-215.
- Focant B, Rosecchi E, Crivelli AJ, (1999). Attempt at biochemical characterization of sand smelt *Atherina boyeri* Risso 1810 (Pisces, Atherinidae) populations from the Camargue (Rhône delta, France). *Comparative Biochemistry and Physiology, part B*, **122**, 261-267.
- Folmer O, Black M, Hoeh R, Lutz R, Vrijenhoek R, (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**, 294-299.
- Forbes AT, Demetriades NT, Benzie JAH, Ballment E, (1999). Allozyme frequencies indicate little geographic variation among stocks of giant tiger prawn *Penaeus monodon* in the south-west Indian Ocean. *South African Journal of Marine Science* **21**, 271-277.

- Franco P, Jeftic L, Malanotte Rizzoli P, Nichelato A, Orlic M, (1982). Descriptive model of the Northern Adriatic. *Oceanologica Acta*, **5**, 379-389.
- Fu YX, (1997). Statistical test of neutrality of mutations against population growth hitchhiking and background selection. *Genetics*, **147**, 915-925.
- Garofa F, Guarniero I, Piccinetti C, Tinti F, (2004a). First microsatellite *loci* of red mullet (*Mullus barbatus*) and their application to genetic structure analysis of Adriatic shared stock. *Marine Biotechnology*, **6**, 446-452.
- Garofa F, Guarniero I, Ramšak A, Ungaro N, Landi M, Piccinetti C, Mannini P, Tinti F, (2004b). Microsatellite DNA variation reveals high gene flow and panmictic populations in the Adriatic shared stocks of the European squid and cuttlefish (Cephalopoda). *Heredity*, **93**, 166-174.
- Geller JB, Walton W, Grosholz ED, Ruiz GM, (1997). Cryptic invasions of the crab *Carcinus* detected by molecular phylogeography. *Molecular Ecology*, **6**, 901-906.
- Goldstein DB, Ruiz-Linares A, Cavalli-Sforza LL, Feldman MW, (1995). An evaluation of genetic distances for use with microsatellite *loci*. *Genetics*, **139**, 463-471.
- Goldstein DB, Schlötterer C, (1999). Microsatellites: Evolution and Applications. Oxford University Press, New York, pp. 352.
- Gum B, Gross R, Kuehn R, (2005). Mitochondrial and nuclear DNA phylogeography of European grayling (*Thymallus thymallus*): evidence for secondary contact zones in central Europe. *Molecular Ecology*, **14**, 1707-1725.
- Guo S, Thompson E, (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361-372.
- Gysels ES, Hellemans B, Pampoulie C, Volckaert FAM, (2004). Phylogeography of the common goby, *Pomatoschistus microps*, with particular emphasis on the colonization of the Mediterranean and the North Sea. *Molecular Ecology*, **13**, 403-417.
- Hancock JM, (1999). Microsatellites and other simple sequences: Genomic context and mutational mechanisms. In *Microsatellites: Evolution and Applications*, Goldstein DB and Schlötterer C (eds). Oxford University Press, New York, pp. 1-9.
- Hedgecock D, (1986). Is gene flow from pelagic larval dispersal important in the adaptation and evolution of marine invertebrates? *Bulletin of Marine Science*, **39**, 550-564.

- Henderson PA, Bamber RN, (1987). On the reproductive biology of the sand smelt *Atherina boyeri* Risso (Pisces: Atherinidae) and its evolutionary potential. *Biological Journal of the Linnean Society*, **32**, 395-415.
- Henderson PA, Holmes RHA, Bamber RN, (1988). Size-selective overwintering mortality in the sand smelt, *Atherina boyeri* Risso, and its roles in population regulation. *Journal of Fish Biology*, **33**, 221-233.
- Herborg LM, Weetman DW, Van Oosterhout C, Hänfling B, (2007). Genetic population structure and contemporary dispersal patterns of a recent European invader, the Chinese mitten crab, *Eriocheir sinensis*. *Molecular Ecology*, **16**, 231-242.
- Hoarau G, Coyer JA, Veldsink JH, Stam WT, Olsen JL, (2007). Glacial refugia and recolonization pathways in the brown seaweed *Fucus serratus*. *Molecular Ecology*, **16**, 3606-3616.
- Hrbek T, Meyer A, (2003). Closing of the Tethys Sea and the phylogeny of Eurasian killifishes (Cyprinodontiformes: Cyprinodontidae). *Journal of Evolutionary Biology*, **16**, 17-36.
- Hudson RR, (1990). Gene genealogies and the coalescent process. *Oxford Surveys in Evolutionary Biology*, **9**, 1-44.
- Ikeda M, Nunokawa M, Taniguchi N, (2003). Lack of mitochondrial gene flow between populations of the endangered amphidromous fish *Plecoglossus altivelis ryukyuensis* inhabiting Amami-oshima Island. *Fishery Sciences*, **69**, 1162-1168.
- Isagi Y, Suhandono S, (1997). PCR primers amplifying microsatellite *loci* of *Quercus myrsinifolia* Blume and their conservation between oak species. *Molecular Ecology*, **6**, 897-899.
- Jarne P, Lagoda PJJ, (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution*, **11**, 424-429.
- Johnsen A, Andersson S, Fernandez JG, Kempnaers B, Pavel V, Questiau S, Raess M, Rindal E, Lifjeld JT, (2006). Molecular and phenotypic divergence in the bluethroat (*Luscinia svecica*) subspecies complex. *Molecular Ecology*, **15**, 4033-4047.
- Jones GP, Milicich MJ, Emslie MJ, Lunow C, (1999). Self-recruitment in a coral reef fish population. *Nature*, **402**, 802-804.
- Ketmaier V, Giusti F, Caccone A, (2006). Molecular phylogeny and historical biogeography of the land snail genus *Solatopupa* (Pulmonata) in the peri-Tyrrhenian area. *Molecular Phylogenetics and Evolution*, **39**, 439-451.

- Knowlton N, Keller BD, (1986). Larvae which fall short of their potential: Highly localized recruitment in an alpheid shrimp with extended larval development. *Bulletin of Marine Science*, **39**, 213-223.
- Lacson JM, (1992). Minimal genetic variation among samples of six species of coral reef fishes collected at La Parguera, Puerto Rico, and Discovery Bay, Jamaica. *Marine Biology*, **112**, 327-331.
- Lasserre P, Marzollo A, (2000). The Venice lagoon ecosystem project: genesis, goals and overview. In *the Venice Lagoon ecosystem: inputs and interactions between land and sea*, Lasserre P and Marzollo A (editors), pp. 1-19.
- Lejeusne C, Chevaldonné P, (2006). Brooding crustaceans in a highly fragmented habitat: the genetic structure of Mediterranean marine cave-dwelling mysid populations. *Molecular Ecology*, **15**, 4123-4140.
- Lessios HA, Kane J, Robertson DR, (2003). Phylogeography of the pantropical sea urchin *Tripneustes*: contrasting patterns of population structure between oceans. *Evolution*, **57**, 2026-2036.
- Lessios HA, Kessing BD, Pearse JS, (2001). Population structure and speciation in tropical seas: global phylogeography of the sea urchin *diadema*. *Evolution*, **55**, 955-975.
- Lessios HA, Kessing BD, Robinson DR, (1998). Massive gene flow across the world's most potent marine biogeographic barrier. *Proceeding of the Royal Society of London, Series B*, **265**, 583-588.
- Levinson G, Gutman GA, (1987). Slippedstrand mispairing: A major mechanism for DNA sequence evolution. *Molecular Biology and Evolution*, **4**, 203-221.
- Lippe C, Dumont P, Bernatchez L, (2006). High genetic diversity and no inbreeding in the endangered copper redhorse, *Moxostoma hubbsi* (Catostomidae, Pisces): the positive sides of a long generation time. *Molecular Ecology*, **15**, 1769-1780.
- Luttikhuizen PC, Drent J, Baker AJ, (2003) Disjunct distribution of highly diverged mitochondrial lineage clade and population subdivision in a marine bivalve with pelagic larvae. *Molecular Ecology*, **12**, 2215-2229.
- Machida RJ, Miya MU, Nishida N, Nishida S, (2004). Large-scale gene rearrangements in the mitochondrial genomes of two calanoid copepods *Eucalanus bungii* and *Neocalanus cristatus* (Crustacea) with notes on new versatile primers for the srRNA and COI genes. *Gene*, **332**, 71-78.
- Mahadevan M, Tsilfidis C, Sabourin L, Shutler G, Amemiya C, Jansen G, Neville C, Narang M, Barcelo J, O'Hoy K, Leblond S, Earle Macdonald J, De Jong J, Wieringa B, (1992). Myotonic dystrophy mutation: An unstable CTG repeat in the 38 untranslated region of the gene. *Science*, **255**, 1253-1258.

- Maltagliati F, Domenica P, Fosch CF, Cossu P, Casu M, Castelli A, (2003). Small-scale morphological and genetic differentiation in the Mediterranean killifish *Aphanius fasciatus* (Cyprinodontidae) from a coastal brackish-water pond and an adjacent pool in northern Sardinia. *Oceanologica Acta*, **26**, 111-119.
- Manly BF, (1984). The statistics of natural selection. Chapman & Hall, New York.
- Marino IAM, Barbisan F, Gennari M, Bisol PM, Zane L, (2008). Isolation and characterization of microsatellite *loci* in the Mediterranean shore crab *Carcinus aestuarii* (Decapoda, Portunidae). *Molecular Ecology Resources*, **8**, 370-372.
- Mazzoldi C, Scaggiante M, Ambrosin E, Rasotto MB, (2000). Mating system and alternative male mating tactic in the grass goby *Zosterisessor ophiocephalus* (Teleostei: Gobiidae). *Marine Biology*, **137**, 1041-1048.
- Meulenkamp JE, Sissingh W, (2003). Tertiary palaeogeography and tectonostratigraphic evolution of the Northern and Southern Peri-Tethys platforms and the intermediate domains of the African–Eurasian convergent plate boundary zone. *Palaeogeography, Palaeoclimatology and Palaeoecology*, **196**, 209-228.
- Milana V, (2008). Caratterizzazione genetica di *Atherina boyeri* (Teleostei, Atherinidae), mediante l'utilizzo di marcatori molecolari: microsatelliti e DNA mitocondriale. Tesi di dottorato in Biologia Animale. Università degli Studi di Roma.
- Miller PJ, (1979). Gobiidae. In *Check List of the Fishes of the North-Eastern Atlantic and of the Mediterranean (CLOFNAM)*. Hureau JC, Monod Th eds., Unesco, Paris. Vol. 1, pp. 483-515.
- Mistri M, Colombo G, (1988). Morphometric variability in sandsmelt, *Atherina boyeri* Risso 1810, population from different Italian sites. *Bollettino di Zoologia*, **3**, 129-132.
- Moritz C, (1995). Uses of molecular phylogenies for conservation. *Philosophical Transactions of the Royal Society of London, Series B*, **349**, 113-118.
- Moritz C, Dowling TE, Brown WM, (1987). Evolution of animal mitochondrial DNA: implications for population biology and systematics. *Annual Review of Ecology and Systematics*, **18**, 269-292.
- Nee S, Holmes EC, Rambaut A, Harvey PH, (1995). Inferring population history from molecular phylogenies. *Philosophical Transactions of the Royal Society, Series B*, **349**, 25-31.
- Nei M, (1987). *Molecular Evolutionary Genetics*. Columbia University Press, New York.

- Nelson JS, (2006). *Fishes of the world*. Fourth Edition. Hoboken, New Jersey: John Wiley & Sons, Inc.
- Nicolas V, Bryja J, Akpatou B, Konecny A, Lecompte E, Colyn M, Lalis A, Couloux A, Denys C, Granjon L, (2008). Comparative phylogeography of two sibling species of forest-dwelling rodent (*Praomys rostratus* and *P. tullbergi*) in West Africa: different reactions to past forest fragmentation. *Molecular Ecology*, **17**, 5118-5134.
- Nordborg M, (2001). Coalescent theory. In *Handbook of statistical genetics*, D.Balding, M. Bishop, and C. Cannings, eds. Wiley, Chichester, England, pp. 179-212.
- O'Donnell WT, Warren ST, (2002). A decade of molecular studies of fragile X syndrome. *Neurosciences*, **25**, 315-338.
- Oliveira-Neto JF, Pie MR, Boeger WA, Ostrensky A, Baggio RA, (2007). Population genetics and evolutionary demography of *Ucides cordatus* (Decapoda: Ocypodidae). *Marine Ecology*, **28**, 460-469.
- Ovenden JR, Brasher DJ, White RWG, (1992). Mitochondrial DNA analyses of the red rock lobster *Jasus edwardsii* supports an apparent absence of population subdivision throughout Australasia. *Marine Biology*, **112**, 319-326.
- Palumbi SR, (1996). Macrospatial genetic structure and speciation in marine taxa with high dispersal abilities. In *Molecular Zoology*, J. D. Ferraris and S. R. Palumbi (eds.). Wiley-Liss, New York, pp. 101-117.
- Palumbi SR, Wilson AC, (1990). Mitochondrial DNA diversity in the sea urchins *Strongylocentrotus purpuratus* and *S. droebachiensis*. *Evolution*, **44**, 403-415.
- Pampoulie C, Stefánsson MO, Jörundsdottir TD, Danilowicz BS, Daniélsdottir AK, (2008). Recolonization history and large-scale dispersal in the open sea: the case study of the North Atlantic cod, *Gadus morhua* L. *Biological Journal of the Linnean Society*, **94**, 315-329.
- Patarnello T, Volckaert FAMJ, Castilho R, (2007). Pillars of Hercules: is the Atlantic-Mediterranean transition a phylogeographical break? *Molecular Ecology*, **16**, 4426-4444.
- Patwary MU, Kenchington EL, Bird CJ, Zouros E, (1994). The use of random amplified polymorphic DNA (RAPD) markers in genetic studies of the sea scallop *Placopecten magellanicus* (Gmelin, 1791). *Journal of Shellfish Research*, **13**, 547-553.
- Pérrin C, Wing SR, Roy MS, (2004). Effect of hydrographic barriers on population genetic structure of the sea star *Coscinasterias mircata*

- (Echinodermata, Asteroidea) in the New Zealand fiords. *Molecular Ecology*, **13**, 2183-2195.
- Pfeiler E, Hurtado LA, Knowles LL, Torre-Cosío J, Bourillón-Moreno L, Márquez-Farías JF, Montemayor-López G, (2005). Population genetics of the swimming crab *Callinectes bellicosus* (Brachyura: Portunidae) from the eastern Pacific Ocean. *Marine Biology*, **146**, 559-569.
- Planes S, Parroni M, Chauvet C, (1998). Evidence of limited gene flow in three species of coral reef fishes in the lagoon of New Caledonia. *Marine Biology*, **130**, 361-368.
- Posada D, Crandall KA, (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics*, **14**, 817-818.
- Primmer CR, Merilä J, (2002). A low rate of cross-species microsatellite amplification success in Ranid frogs. *Conservation Genetics*, **3**, 445-449.
- Procaccini G, Maltagliati F, (2003). Approcci metodologici all'analisi della diversità genetica di organismi bentonici. *Biologia Marina Mediterranea*, **10**, 509-548.
- Pybus OG, Rambaut A, Harvey PH, (2000). An integrated framework for the inference of viral population history from reconstructed genealogies. *Genetics*, **155**, 1429-1437.
- Raimondi PT, Keough MJ, (1990). Behavioural variability in marine larvae. *Australian Journal of Ecology*, **15**, 427-427.
- Rambaut A, Drummond A, (2007). Tracer v1.4. Available from <http://beast.bio.ed.ac.uk/>.
- Rassman K, Schlotter C, Tautz D, (1991). Isolation of simple sequence *loci* for use in polimerase chain reaction-base DNA fingerprinting. *Electrophoresis*, **12**, 113-118.
- Richter C, Park JW, Ames BN, (1988). Normal oxidative damage to mitochondrial and nuclear DNA is extensive. *Proceedings of the National Academy of Sciences USA*, **85**, 6465-6467.
- Rogers AR, (1995). Genetic evidence for a Pleistocene population explosion. *Evolution*, **49**, 608-615.
- Rogers AR, Harpending H, (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552-569.
- Rögl F, (1999). Mediterranean and Paratethys. Fact and hypotheses of an Oligocene to Miocene paleogeography (short overview). *Geologica Carpathica*, **50**, 339-349.

- Rögl F, Steininger F, (1983). Vom Zerfall der Tethys zu Mediterran und Paratethys. *Annalen Des Naturhistorischen Museums in Wien*, **85A**, 135-163.
- Roman J, (2006). Diluting the founder effect: cryptic invasions expand a marine invader's range. *Proceedings of the Royal Society of London, Series B*, **273**, 2453-2459.
- Roman J, Palumbi SR, (2004). A global invader at home: population structure of the green crab, *Carcinus maenas*, in Europe. *Molecular Ecology*, **13**, 2891-2898.
- Rousset F, Raymond M, (1995). Testing heterozygote excess and deficiency. *Genetics*, **140**, 1413-1419.
- Rudloe A, (1979). *Limulus polyphemus*: A review of the ecologically significant literature. In *Biomedical Applications of the Horseshoe Crab (Limulidae)*, E. Cohen (ed.). Alan R. Liss, New York, pp. 27-35.
- Russel PJ, (2003). GENETICA. Terza edizione, EdiSES s.r.l, Napoli.
- Ruzzante DE, Taggart CT, Cook D, (1998). A nuclear DNA basis for shelf- and bank-scale population structure in northwest Atlantic cod (*Gadus morhua*): Labrador to Georges Bank. *Molecular Ecology*, **7**, 1663-1680.
- Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Hron GT, Mullis KB, Erlich HA, (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, **239**, 487-491.
- Saunders NC, Kessler LG, Avise JC, (1986). Genetic variation and geographic differentiation in mitochondrial DNA of the horseshoe crab, *Limulus polyphemus*. *Genetics*, **112**, 613-627.
- Shaklee JB, (1984). Genetic variation and population structure in the damselfish, *Stegastes fasciolatus*, throughout the Hawaiian archipelago. *Copeia*, **1984**, 629-640.
- Shaw IW, Arkhipin AI, Al-Khairulla H, (2004). Genetic structuring of Patagonian toothfish populations in the Southwest Atlantic Ocean: the effect of the Antarctic Polar Front and deep-water troughs as barrier to genetic exchange. *Molecular Ecology*, **13**, 3293-3303.
- Shields GF, Gust JR, (1995). Lack of geographic structure in mitochondrial DNA sequences of Bering Sea walleye pollack, *Theragra chalcogramma*. *Molecular Marine Biology and Biotechnology*, **4**, 69-82.
- Shulman MJ, Bermingham E, (1995). Early life histories, ocean currents, and the population genetics of Caribbean reef fishes. *Evolution*, **49**, 897-910.

- Sia EA, Butler CA, Dominska M, Greenwell P, Fox TD, Petes TD, (2000). Analysis of microsatellite mutations in the mitochondrial DNA of *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences USA*, **97**, 250-255.
- Simonsen KL, Churchill GA, Aquadro CF, (1995). Properties of statistical tests of neutrality for DNA polymorphism data. *Genetics*, **141**, 413-429.
- Soule ME, (1987). *Viable Populations for Conservation*. Cambridge University Press, Cambridge, U.K.
- Strabella S, (1996). Produzione di moleche in laguna. Laguna – Produzione e mercato, pp. 32-35.
- Strand M, Prolla TA, Liskay RM, Petes TD, (1993). Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature*, **365**, 274-276.
- Swearer SE, Caselle JE, Lea DW, Warner RR, (1999). Larval retention and recruitment in an island population of a coral-reef fish. *Nature*, **402**, 799-802.
- Swofford DL, (1998). PAUP: Phylogenetic Analysis Using Parsimony (\*and Other Methods), Version 4.0d64 ed, Sinauer Associates, Inc. Publishers, Sunderland, MA.
- Tajima F, (1989). Statistical methods for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585-595.
- Tamura K, Dudley J, Nei M, Kumar S, (2007). *MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0*. *Molecular Biology and Evolution*. **24**, 1596-1599.
- Tautz D, (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research*, **17**, 6463-6471.
- Taylor MS, Hellberg ME, (2003). Genetic evidence for local retention of pelagic larvae in a Caribbean reef fish. *Science*, **299**, 107-109.
- Tepolt CK, Bagley MJ, Geller JB, Blum MJ, (2006). Characterization of microsatellite *loci* in the European green crab (*Carcinus maenas*). *Molecular Ecology Notes*, **6**, 343-345.
- Thompson AR, Thacker CE, Shaw EY, (2005). Phylogeography of marine mutualists: parallel patterns of genetic structure between obligate goby and shrimp partners. *Molecular Ecology*, **14**, 3557-3572.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG, (1997). The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **24**, 4876-4882.

- Tóth G, Gáspari Z, Jurka J, (2000). Microsatellites in different eukaryotic genomes: Survey and analysis. *Genome Research*, **10**, 967-981.
- Trabelsi M, Maamouri F, Quignard JP, Boussaïd M, Faure E, (2004). Morphometric or morpho-anatomical or genetic investigations highlight allopatric speciation in Western Mediterranean lagoons within the *Atherina lagunae* species (Teleostei, Atherinidae). *Estuarine, Coastal and Shelf Science*, **61**, 713-723.
- Van Houdt JK, Hellemans B, Volckaert FAM, (2003). Phylogenetic relationships among Palearctic and Nearctic burbot (*Lota lota*): Pleistocene extinctions and recolonization. *Molecular Phylogenetics and Evolution*, **29**, 599-612.
- Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Friters A, Pot J, Paleman J, Kuiper M, Zabeau M, (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, **23**, 4407-4414.
- Wang L, Juji T, Tokunaga K, Takahashi K, Kuwata S, Ukida S, Tadokoro K, Takai K, (1994). Polymorphic microsatellite markers for the diagnosis of Graft-versus-Host disease. *The New England Journal of Medicine*, **330**, 398-401.
- Waples RS, (1998). Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, **89**, 438-450.
- Waples RS, Gaggiotti O, (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology* **15**, 1419-1439.
- Ward RD, Woodward M, Skibinski DOF, (1994). A comparison of genetic diversity levels in marine, freshwater, and anadromus fishes. *Journal of Fish Biology*, **44**, 213-232.
- Weir BS, (1990). Intraspecific differentiation. In *Molecular systematics*. Hillis D. M. & Moritz C. (eds). Sinauer, Sunderland, Massachusetts, pp. 373-410.
- Weir BS, Cockerham CC, (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370.
- White G, Powell W, (1997). Isolation and characterization of microsatellite *loci* in *Swietenia humilis* (Meliaceae): An endangered tropical hardwood species. *Molecular Ecology*, **6**, 851-860.
- Wilke T, Pfenninger M, (2002) Separating historic events from recurrent processes in cryptic species: phylogeography of mud snails (*Hydrobia* spp.). *Molecular Ecology*, **11**, 1439-1451.
- Williams BG, (1968). Laboratory rearing of the larval stage of *Carcinus maenas*. *Journal of Natural History*, **2**, 121-126.

- Wright S, (1938). Size of population and breeding structure in relation to evolution. *Science*, **87**, 430-431.
- Wright S, (1940). Breeding structure of populations in relation to speciation. *American Naturalist*, **74**, 232-248.
- Wright S, (1951). The genetic structure of population. *Annals of Eugenics*, **15**, 323-354.
- Zane L, Bargelloni L, Paternello T, (2002). Strategies for microsatellite isolation: a review. *Molecular Ecology*, **11**, 1-16.
- Zarza E, Renoso VH, Emerson BC, (2008). Diversification in the northern neotropics: mitochondrial and nuclear DNA phylogeography of the iguana *Ctenosaura pectinata* and related species. *Molecular Ecology*, **17**, 3259-3275.
- Zavatarelli M, Raicich F, Bregant D, Russo A, Artegiani A, (1998). Climatological biogeochemical characteristics of the Adriatic Sea. *Journal of Marine Systems*, **18**, 227-263.
- Zerunian S, (2002). Condannati all'estinzione? Biodiversità, biologia, minacce e strategie di conservazione dei Pesci d'acqua dolce indigeni in Italia. Edagricole. Bologna, pp. 220.
- Zonneveld KAF, (1996). Paleoclimatic reconstruction of the last deglaciation (18-8 ka BP) in the Adriatic Sea region; a land-sea correlation based on palynological evidence. *Paleogeography, Paleoclimatology, Paleoecology*, **122**, 89-106.
- Zouros E, Freeman KR, Oberhauser-Ball A, Pogson GH, (1992). Direct evidence for extensive paternal mitochondrial DNA inheritance in the marine mussel *Mytilus*. *Nature*, **359**, 412-414.