



UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI BIOMEDICINA COMPARATA ED ALIMENTAZIONE (BCA)

TESI DI LAUREA IN BIOTECNOLOGIE PER L'ALIMENTAZIONE

Caratterizzazione molecolare di Crostacei Malacostraca di interesse alimentare

Relatore:

Dott. *Enrico Massimiliano Negrisola*

Laureando:

Andrea Basso

Matricola n. 1019444

ANNO ACCADEMICO 2012 – 2013

INDICE

Riassunto.....	1
Abstract.....	2
1. Introduzione	3
1.1. Il mitocondrio e il genoma mitocondriale	5
1.1.1. <i>Gene Order</i>	8
1.2. DNA barcoding	10
1.3. Malacostraca	12
1.3.1. Decapoda	12
1.3.2. Stomatopoda	14
2. Obiettivo della tesi.....	16
3. Materiali e Metodi.....	17
3.1. Campioni.....	17
3.2. Estrazione DNA	18
3.3. Determinazione della concentrazione del DNA estratto.....	19
3.4. Gel elettroforesi.....	20
3.5. Identificazione di specie	20
3.6. Sequenziamento.....	22
3.6.1. Correzione dell'elettroferogramma e creazione della sequenza consenso	22
3.6.2. <i>Next Generation sequencing</i> con tecnologia Illumina.....	23
3.6.3. Assemblaggio delle sequenze illumina.....	25
3.7. Reazioni di PCR di validazione	25
3.8. Analisi bioinformatiche	27
3.8.1. Dataset e annotazione dei genomi.....	27
3.8.2. Analisi dei genomi e dei <i>Gene Order</i>	30
3.8.3. Allineamenti multipli	30
3.8.4. Determinazione del segnale filogenetico.....	31
3.8.5. Costruzione degli alberi filogenetici.....	33
4. Risultati.....	34
4.1. Descrizione dei genomi e <i>Gene Order</i>	34

4.1.1.	Principali caratteristiche dei nuovi mtDNA.....	34
4.1.2.	Analisi dei <i>Gene Order</i>	35
4.1.3.	Confronto delle sequenze di taxa congenerici.....	46
4.2.	Determinazione del segnale filogenetico.....	47
4.3.	Analisi filogenetiche	51
5.	Discussione	60
5.1.	<i>Gene Order</i>	60
5.2.	Evoluzione dei PCG e loro valore filogenetico.....	61
5.3.	Individuazione di <i>tools</i> mitocondriali per vari tipi di applicazioni.....	62
6.	Bibliografia	64
	Ringraziamenti	77

Riassunto

Nel mercato agroalimentare, il settore ittico è uno dei più soggetti a erronee identificazioni di specie, con scopo fraudolento o semplicemente per mancanza di adeguata conoscenza tassonomica. Infatti, in seguito alla globalizzazione e alla maggiore disponibilità di prodotti a livello di *Grande Distribuzione Organizzata*, è soggetto ad una offerta estremamente vasta. Inoltre, il continuo cambiamento climatico, consente la migrazione volontaria di specie esotiche, nel bacino del Mediterraneo che possono essere incluse anche i nostri circuiti commerciali. Risulta quindi fondamentale identificare correttamente e in maniera più rapida possibile le specie ittiche di interesse alimentare. Tuttavia, l'unica metodologia attualmente valida in sede legale in Italia si basa sull'identificazione morfologica e quindi risulta inapplicabile a prodotti lavorati e trasformati i quali occupano una grande fetta del consumo domestico.

Lo scopo del presente lavoro è stato quindi quello sequenziare i genoma mitocondriali completi di 19 specie di crostacei Malacostraci, per confrontarli con gli altri già esistenti in modo da individuare vari tipi di marcatori utili per una identificazione molecolare univoca delle specie.

Il sequenziamento è avvenuto attraverso il metodo Illumina, un approccio molto efficace per la determinazione di interi genomi. I dati ottenuti hanno consentito di studiare A) il *gene order* delle specie in esame; B) i livelli di evoluzione molecolare dei diversi geni; C) le relazioni filogenetiche tra i principali gruppi di Decapodi.

L'identificazione di nuovi arrangiamenti come pure l'individuazione di tassi di cambiamento variabili tra i geni, aprono la possibilità di utilizzare queste caratteristiche per sviluppare marcatori molecolari capaci di identificare in maniera rapida e accurata le diverse specie anche quando queste sono state variamente lavorate, com'è spesso il caso dei prodotti ittici, e di sviluppare quindi migliori sistemi di controllo e sicurezza degli alimenti, potenziale identificazione della zona di cattura dell'animale, salvaguardia e tutela delle categorie di persone a rischio. Tali informazioni possono poi essere utilmente sfruttate per sviluppare marker utili per studi di filogeografia e genetica di popolazioni con possibilità di identificare l'origine geografica dei campioni analizzati e possono essere proficuamente impiegati nella gestione degli stock di pesca.

Parole chiave: Frodi alimentari; Illumina; Identificazione di specie; Filogenesi; Gene-order.

Abstract

In the food market, the fisheries sector is one of the most sensitive to the incorrect identification of species, with fraudulent intent, or simply due to lack of adequate knowledge of taxonomy. In fact, as result of globalization and the increased availability of products in Large Scale Distribution, is subject to a very wide offer. Moreover, climatic changes, allow voluntary migration of exotic species in the Mediterranean Sea, that can be included in our marketing networks. So it is essential to identify correctly and as quickly as possible, the species of food interest. However, the only legally accepted method in Italy, is the morphological identification that is not applicable to processed products, which occupy an important role in home consumption.

The aim of this study has been to sequence the whole mitochondrial genome of 19 species of crustaceans Malacostraca, to compare their genomes with others already available, in order to identify various types of useful molecular markers for the identification of species.

The sequencing was performed through the Illumina method, a very effective approach for the determination of entire genomes. The obtained data allowed to study A) the gene order of analysed species; B) the molecular evolution levels in different genes; C) the phylogenetic relationships among the major groups of Decapoda.

The identification of new arrangements as well as the identification of different substitution rates among genes, open the possibility to use these features to develop molecular markers capable to quickly and accurately identify the different species even if they have been variously worked, and then to develop the best systems of control and food safety.

They also allow to potential identify the capture areas, to preserve and protect categories of people at risk.

This information can be used to develop useful markers for phylogeography and population genetics studies with the possibility to identify the geographic origin of the samples analyzed and can be employed in the management of fisheries stocks.

Keywords: Food fraud; Illumina; Species identification, Phylogeny, Gene-order.

1.Introduzione

Negli ultimi anni c'è stata una richiesta sempre maggiore di prodotti ittici, dovuta principalmente: a) alle migliori proprietà nutritive degli stessi rispetto alle carni consumate tradizionalmente (Moretti & Busetto 2010); b) alla globalizzazione, che permette una maggiore disponibilità di prodotti a livello della grande distribuzione organizzata e favorisce gli scambi con il resto del mondo, in particolare con quello asiatico e africano, pur permettendo il commercio di prodotti freschi; c) alla crescente diversificazione di specie ittiche disponibili sul mercato.

Inoltre, a seguito del cambiamento climatico e del riscaldamento delle acque, si ha il passaggio volontario di alcune specie tropicali nel Mar Mediterraneo (specie che dal Mar Rosso e Oceano Indiano entrano nel bacino Mediterraneo attraverso il canale di Suez). Ciò ha permesso l'introduzione di specie alloctone e poco conosciute nei nostri circuiti commerciali (Busato 2010).

Tutto questo espone, in misura sempre maggiore, il settore ittico, a frodi commerciali; in particolare, la più frequente nel settore, risulta essere la "falsificazione", ossia la sostituzione di una specie pregiata con una di minor valore commerciale. Per tanto, ispettori sanitari e veterinari, sono chiamati a eseguire una corretta identificazione delle specie, presenti negli alimenti, sia per le implicazioni economiche, che per quelle sanitarie a tutela del consumatore, fermando la commercializzazione di specie potenzialmente tossiche o allergeniche (Busato 2010).

A livello di legislazione italiana, l'identificazione di specie deve avvenire tramite l'identificazione morfologica dell'animale, utilizzando le chiavi dicotomiche fornite dalla *Food and Agriculture Organization* (FAO). Ovviamente ciò non può avvenire su prodotti lavorati, i quali rappresentano quasi la metà dell'attuale consumo domestico (49 % nel 2011, dati Ismea).

Per ovviare a questo problema sono state testate diverse tecniche, come l'isoelettrofocalizzazione di proteine, ma si sono dimostrate di scarsa utilità a causa, ad esempio, della denaturazione delle proteine in seguito a trattamenti termici o alla presenza di isoforme in specie filogeneticamente simili (Rehnbein *et al.* 1995). Altre tecniche basate sul profilo di restrizione del genoma si sono rivelate migliori, ma sono potenzialmente soggette a problematiche dovute a mutazioni nei siti di restrizione causando un'alterazione nell'eventuale profilo (Busato 2010).

A livello commerciale tutto ciò si traduce nell'impossibilità di mettere a punto sistemi efficaci, rapidi ed economici per tracciare i prodotti di origine animale che, oltre a

fornire un sistema di controllo e sicurezza degli alimenti, permette di identificare la zona di cattura dell'animale, di salvaguardare categorie di persone a rischio e tutelare scelte alimentari individuali (Busato 2010).

L'approccio più efficace è di conseguenza, quello di utilizzare singoli/set di geni che consentano di identificare, in modo univoco a livello di specie, prodotti ittici variamente lavorati e quindi non più identificabili con le tradizionali chiavi dicotomiche (Nicolè *et al.* 2012).

Condizione necessaria e preliminare per lo sviluppo di tali marker è la caratterizzazione molecolare delle specie target per le quali spesso mancano completamente, o sono particolarmente limitati, i dati genetici.

Particolarmente efficace per il riconoscimento di specie animali risulta essere una porzione del gene mitocondriale *cox1* che viene utilizzato come *marker standard* per il *barcoding* animale (Hebert *et al.* 2003). Tuttavia questo singolo marcatore può essere in certi casi inefficace per il riconoscimento a livello specifico. Inoltre lo stesso consorzio internazionale per il *DNA barcoding* suggerisce, nelle sue linee guida, di suffragare l'identificazione ottenuta con *cox1* attraverso l'impiego di altri marcatori. Ciò è particolarmente importante se i risultati di tale identificazione devono essere utilizzati in sede legale (<http://www.boldsystems.org/>).

Quindi una caratterizzazione molecolare basata su un dataset di geni è il sistema più robusto per una univoca identificazione a livello specifico.

Per tale scopo sono disponibili i geni derivati da due distinti genomi: quello nucleare e quello mitocondriale. Quest'ultimo genoma presenta una serie di caratteristiche che lo rendono particolarmente idoneo allo scopo citato sopra e che sono illustrate in dettaglio nel paragrafo successivo.

La caratterizzazione molecolare di specie di interesse alimentare va al di là della mera identificazione di specie. Infatti una volta che sono stati individuati una serie di marcatori molecolari efficaci, questi possono essere impiegati anche per studi di filogeografia e per gestione degli stock di pesca. Inoltre in un ambito più generale, che esula dall'aspetto applicativo, la determinazione di genomi mitocondriali completi offre la possibilità di studi comparativi ed evolutivisti delle specie animali studiate.

1.1. Il mitocondrio e il genoma mitocondriale

Il mitocondrio è un organello cellulare, di forma allungata, costituito principalmente da una serie di membrane selettivamente permeabili in cui sono disposte una serie di proteine coinvolte nella fosforilazione ossidativa. Questi organelli sono presenti in alto numero nel citoplasma, nella maggior parte delle cellule eucariotiche per le quali fungono da centrale energetica della cellula (Campbell *et al.* 2006).

Il mitocondrio è anche coinvolto in altre molteplici funzioni, come: a) la regolazione del metabolismo cellulare (McBride *et al.* 2006); b) la regolazione del potenziale redox della cellula (Li *et al.* 2013, Voet *et al.* 2006); c) la sintesi del gruppo eme (Oh-hama 1997); d) la sintesi del colesterolo (Rossier 2006); e) la produzione di calore (Voet *et al.* 2006); f) l'induzione di apoptosi (Green 1998, Hajnóczky *et al.* 2006).

Il mitocondrio, inoltre, possiede un proprio genoma (mtDNA) che negli animali di solito include 37 geni, compresi quelli che codificano per le proteine necessarie alla fosforilazione ossidativa.

Questa caratteristica, più alcune altre similitudini con i batteri, ha portato alla formulazione della teoria endosimbiontica, in cui il mitocondrio deriva da un processo di fagocitosi e successiva trasformazione di un batterio, simile agli attuali α -proteobatteri e dotato di metabolismo ossidativo, da parte di un proto-eucariote, con conseguente formazione della primitiva cellula eucariote (Richards & Archibald 2013).

Successivamente questa simbiosi si sarebbe ulteriormente rafforzata e consolidata con il progressivo trasferimento da parte del batterio, di gran parte dei propri geni all'interno del nucleo del proto-eucariote, conferendo un vantaggio evolutivo a questi eucarioti ancestrali (Martin & Müller 2007; Saccone *et al.* 2008; Thrash *et al.* 2011).

Il genoma mitocondriale animale o mtDNA, è di solito una singola macromolecola a doppio filamento di circa 14-20 Kb, ma sono note eccezioni a tale arrangiamento standard (Boore 1999).

Restringendo la descrizione puntuale dell'mtDNA ai Crostacei Malacostraci, gruppo oggetto di questa tesi (Figura 1), si ha che il contenuto genico è quello usuale degli animali e cioè 37 geni privi di introni, suddivisi in: 22 geni che codificano per tRNA, 13 geni che codificano per le proteine mitocondriali coinvolte nella fosforilazione ossidativa e 2 rRNA (Boore 1999). Nei due filamenti complementari, i geni sono distribuiti in un numero differente. Per convenzione il filamento che contiene il gene

cox1 (Figura 1) viene chiamato *alpha* (α -)strand, mentre il suo complementare viene indicato come *beta* (β -)strand. I geni possono essere contigui sullo stesso filamento o sui filamenti complementari come pure sovrapporsi o essere separati da un numero variabile di nucleotidi (da 1 a alcune migliaia). Tali punti di separazione sono chiamati spaziatori intergenici (ISP, *Intergenic SPacers*) e sono variabili per numero, collocazione e dimensioni a seconda delle specie (Negrisolo *et al.* 2011). Lo spaziatore intergenico compreso tra *rrnS* e *trnI*, nel *gene order* comune ai Pancrustacea, viene indicato anche col nome di *AT-rich region* o regione di controllo (CR) in quanto è il sito in cui inizia la trascrizione e la replicazione del genoma stesso (Boore 1999) (Figura 1).

La CR è una porzione dell'mtDNA a rapida evoluzione ed è stata utilizzata come marcatore in studi di genetica di popolazione (es. Buccoli 2006; Babbucci *et al.* 2010).

Il codice genetico dell'mtDNA differisce da quello universale dei phyla animali e nel caso dei Crostacei si utilizza il codice genetico mitocondriale per gli invertebrati in cui il codone ATA codifica per la Metionina invece che per la Isoleucina, il codone TGA non è un codone di stop, ma codifica per il Triptofano, ed i codoni AGA e AGG codificano per la Serina (S1) invece che per l'Arginina. I PCG (*Protein-Coding Genes*) mitocondriali presentano oltre al codone di start canonico (ATG), codoni alternativi universalmente riconosciuti per gli invertebrati (ATN, GTG, TTG, GTT) oltre a codoni non canonici ristretti a particolari taxa (es. ACG per la *cox1* di molti crostacei). Analogamente i codoni di stop si possono raggruppare in: completi (TAA o TAG) e incompleti (T, TA). A quest'ultimi, in seguito alla trascrizione, verranno aggiunte una o più adenine mancanti completando il codone.

L'mtDNA è aploide e solitamente a trasmissione uniparentale materna; l'unica eccezione conosciuta, è rappresentata dai molluschi bivalvi in cui, a livello delle gonadi, è presente anche un genoma mitocondriale paterno (Boore 1999).

Il genoma mitocondriale, viene spesso utilizzato per l'identificazione di specie negli organismi animali in quanto, nella maggior parte dei casi, non presenta fenomeni di ricombinazione e quindi, a livello intraspecifico, può essere considerato come un clone. Inoltre, sebbene anche il DNA nucleare possa essere disponibile, viene usato preferibilmente l'mtDNA poiché è di dimensioni minori ed ha un tasso di sostituzione nucleotidica maggiore, presentando quindi maggiori zone polimorfiche a livello interspecifico, favorendo l'identificazione di specie strettamente correlate. Anche l'ordine dei geni (*gene order*), che può variare in gruppi di diverso rango tassonomico

(Boore 1999; Negrisolo *et al.* 2011), può essere utile come carattere diagnostico (come illustrato in seguito) sebbene il contenuto dell'mtDNA rimanga immutato (Negrisolo *et al.* 2011).

Un altro vantaggio dell'mtDNA è la resistenza alla denaturazione dovuta alla propria conformazione circolare che lo rende particolarmente adatto all'analisi su matrici alimentari complesse e sottoposte a trattamenti termici. Inoltre, essendo presente in diverse copie nella stessa cellula (da 100 a 10.000 copie), l'estrazione risulta facilitata anche con basse quantità di tessuto (Martin *et al.* 2009; Busato 2010).

Alla luce di tutto ciò, l'mtDNA risulta quindi molto efficace sia per l'identificazione molecolare di specie attraverso il *DNA barcoding* (Hebert *et al.* 2003), aspetto molto importante per prevenire/individuare frodi alimentari, sia per studi filogenetici, che grazie ai differenti tassi di sostituzione nucleotidica/amminoacidica dei differenti geni consentono di investigare con successo percorsi evolutivi sviluppatasi in periodi temporali anche molto diversi (Negrisolo *et al.* 2011).

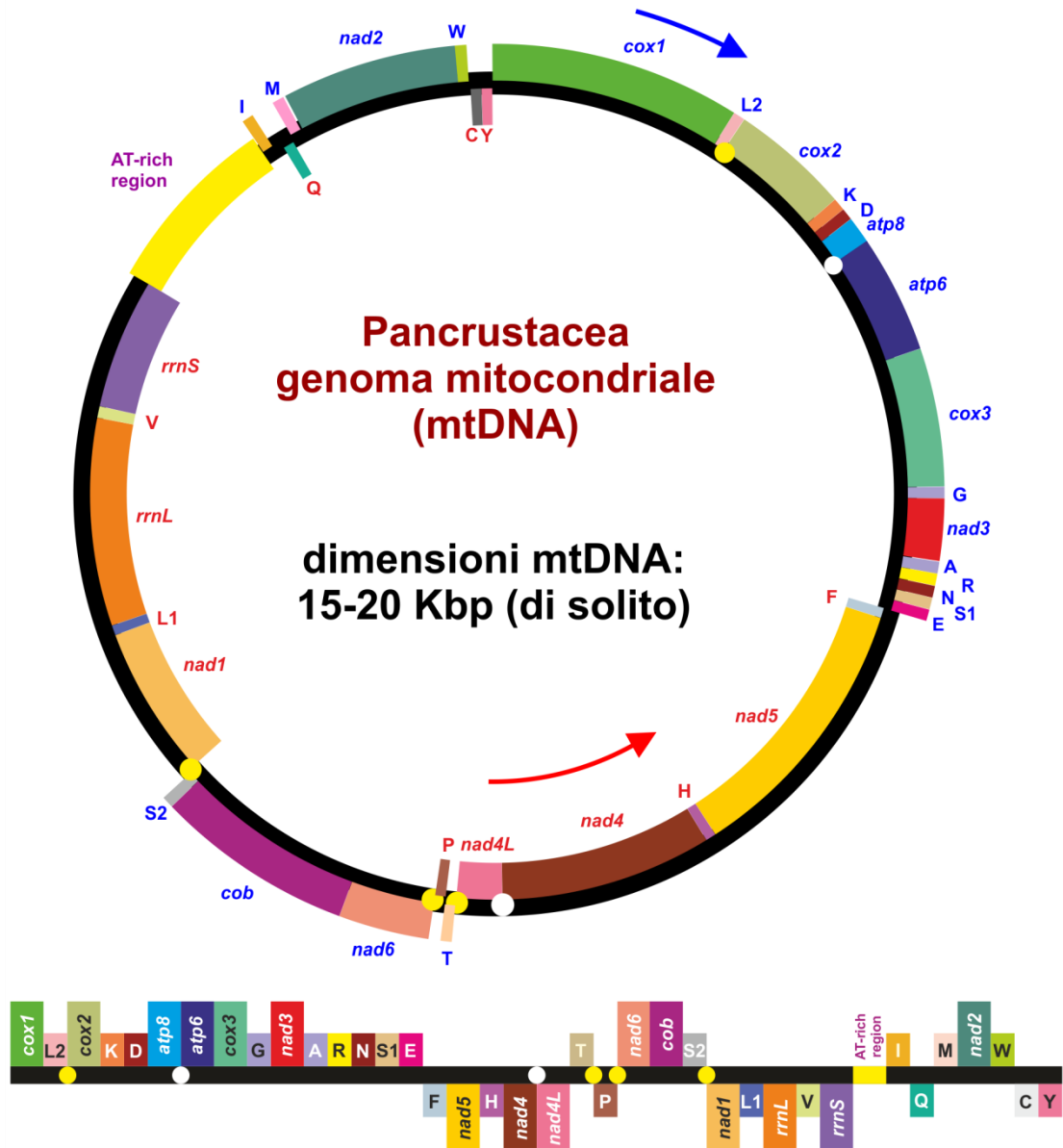


Figura 1: organizzazione standard del genoma mitocondriale dei Pancrustacea

1.1.1. Gene Order

Si definisce come *gene order* dell'mtDNA l'ordine in cui sono disposti i geni (posizione ed orientamento) nel genoma stesso. Trattandosi di una molecola circolare, per semplicità di confronto tra *gene order* diversi come pure per le analisi bioinformatiche, il genoma mitocondriale viene linearizzato a partire dal *cox1* (Figura 1). A fronte quindi di un contenuto genico che rimane invariato, l'ordine dei geni può cambiare sensibilmente nei diversi *taxa* anche se questa propensione al cambiamento non è uniformemente distribuita nei differenti gruppi animali (Krebs & Bastrop 2012). I possibili cambiamenti (Figura 2) che si possono osservare in un

gene order sono: (a) *trasposizione*, cioè lo spostamento di un gene rispetto alla sua collocazione in un *gene order* di riferimento, che mantiene però lo stesso orientamento; (b) *inversione*, cioè il cambiamento di orientamento di un gene che rimane nella stessa posizione ma passa nel filamento complementare; (c) la *trasposizione inversa* che combina i due processi descritti in (a) e (b). Gran parte dei cambiamenti che si osservano nei *gene order* possono essere spiegati applicando il modello “*Tandem Duplication and Random Loss*” (TDRL) (Boore 2000) (Figura 2a). Questo modello prevede un’iniziale duplicazione della porzione genomica in esame, principalmente dovuta ad un processo di *slippage* e conseguente *mispairing* della DNA polimerasi durante la replicazione, seguita dalla delezione casuale di uno dei due geni per ciascuna coppia duplicata (Boore 2000). L’inversione può essere spiegata con un processo di ricombinazione intramitocondriale (Dowton & Austin 1999) (Figura 2b). L’insieme dei cambiamenti che si verificano nella formazione di un nuovo *gene order* implica in molti casi la comparsa di nuove mutazioni deleterie che porteranno il mitocondrio alla morte. Tuttavia dato l’enorme numero di questi organelli per cellula, la probabilità che uno di questi cambiamenti sia neutrale e mantenga la funzionalità dell’organello, risulta consistente.

Sebbene il *gene order standard* dei Pancrustacea (Figura 1) sia quello maggiormente rappresentato nei Crostacei Malacostraci completamente sequenziati fino ad oggi, sono noti vari casi di riarrangiamento (es.: Bauza-Ribot *et al.* 2009; Ki *et al.* 2010; Krebs & Bastrop 2012). L’analisi di tali riarrangiamenti, oltre a consentire di studiare l’evoluzione strutturale dell’mtDNA, permette di individuare eventuali marcatori utili alla caratterizzazione e definizione dei gruppi animali a diverso rango tassonomico. In questo senso un peculiare *gene order* può diventare un vero e proprio marcatore genomico (Negrisolo *et al.* 2011).

Da un punto di vista applicativo, l’individuazione di un nuovo *gene order* può consentire di sfruttare la regione in cui è avvenuto il riarrangiamento a scopo diagnostico. Infatti, disegnando dei *primer ad hoc*, che consentano di amplificare solamente il DNA di quei taxa che possiedono quel peculiare arrangiamento, si può ottenere una identificazione certa, evitando nel contempo la necessità di sequenziare il prodotto di PCR. Il saggio si risolverà perciò in un test di presenza/assenza risultando nel contempo molto più rapido ed economico, degli attuali saggi per l’identificazione di specie.

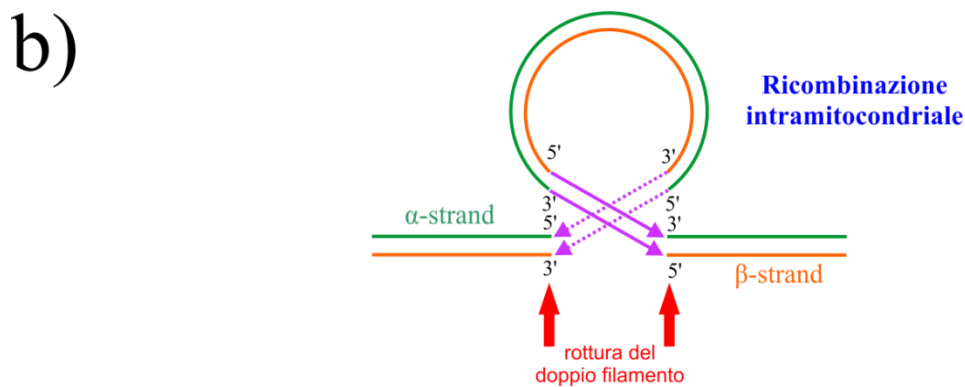
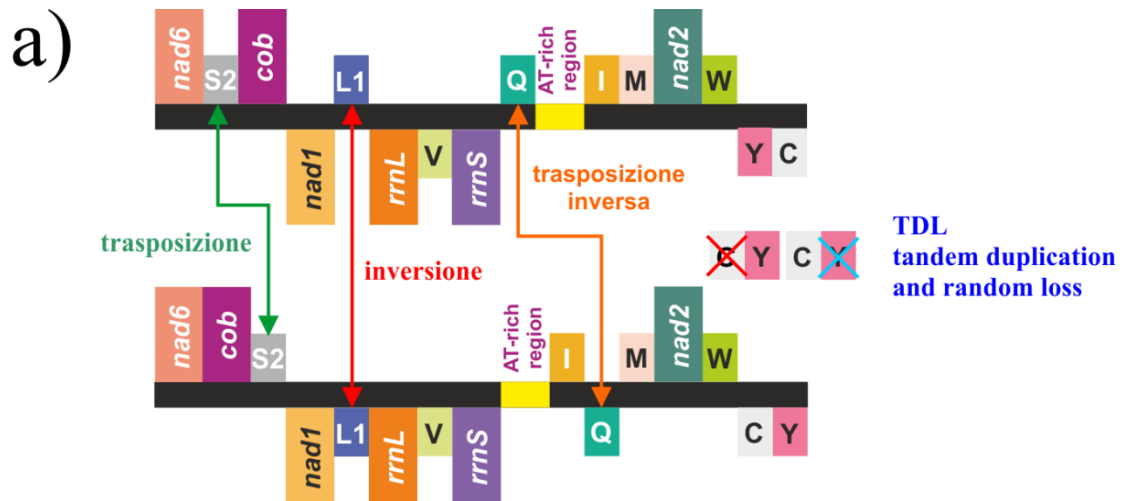


Figura 2: meccanismi che determinano la modifica del gene order standard

1.2. DNA barcoding

Il *DNA barcoding*, è una metodica molecolare che permette l'identificazione, a livello di specie, di organismi viventi, attraverso l'analisi di un marcatore molecolare. Questo consente di identificare con una ragionevole sicurezza l'organismo anche quando non è possibile l'identificazione morfologica, come nel caso in cui si abbia la mancanza di strutture caratterizzanti dovuta a: manipolazione, stadio giovanile diverso dall'adulto o quando non si abbia la possibilità di avvalersi di un tassonomo esperto nel corso dello studio, o quando gli esemplari in esame appartengano a specie morfologicamente molto simili, ecc. (Berardi 2011).

Essendo presenti queste limitazioni nell'identificazione morfologica, Herbert *et al.* (2003), proposero di utilizzare una porzione relativa al gene *cox1* dell'mtDNA per l'identificazione di specie; basandosi sull'assunto che, nel gene in esame, la variabilità interspecifica è nettamente superiore e non sovrapposta a quella intraspecifica. Molte delle ragioni per cui si è deciso di utilizzare l'mtDNA sono già

state citate precedentemente, ma quelle che hanno reso così popolare l'utilizzo di *cox1* a scopo diagnostico sono: (1) l'esistenza di *primer* universali (Zhang & Hewitt 1997, Herbert *et al.* 2003); (2) *cox1* sembra possedere il più alto segnale filogenetico di qualsiasi altro gene mitocondriale. Infatti, l'evoluzione di questo gene è sufficientemente rapida da consentire di discriminare sia specie strettamente correlate, ma anche gruppi distinti appartenenti alla stessa specie (Cox & Hebert 2001).

Il *DNA barcoding* si basa perciò sul sequenziamento di una porzione di *cox1*, di circa 700 paia di basi (bp), in prossimità dell'estremità 5' del gene stesso e sul confronto della sequenza ottenuta con un database pubblico di altre sequenze di riferimento depositate. Queste sequenze inoltre devono essere state ottenute da esemplari identificati da esperti tassonomi. Questo sistema di riconoscimento è diventato talmente proficuo e di uso comune da entrare a far parte del progetto BARCODING OF LIFE o BOLD (<http://www.boldsystems.org/>) in cui si possono confrontare le sequenze ottenute dai propri esemplari e spesso assegnarli in maniera certa ad una determinata specie (Ratnasingham *et al.* 2007).

L'assunto teorico che sta alla base dell'identificazione molecolare disponibile in BOLD, è che le sequenze ottenute dagli individui della stessa specie formano un gruppo monofiletico nettamente separato rispetto ad altre specie impiegando quindi il concetto di specie filogenetica. Numerosi progetti che hanno impiegato un approccio di *DNA barcoding* ne hanno dimostrato l'efficacia grazie alle corrette identificazioni ottenute per diversi gruppi di animali (uccelli, rettili, pesci, lepidotteri), piante e funghi (Lukhtanov *et al.* 2009). Non mancano però casi in cui il *DNA barcoding* si è dimostrato totalmente inefficace (Whitworth *et al.* 2007) e quindi il singolo utilizzo di una porzione di *cox1* non garantisce a priori di poter ottenere una corretta identificazione a livello specifico e supporta la richiesta della necessità di un set di marcatori per risolvere completamente questo tipo di necessità.

Nonostante i limiti ricordati sopra, negli ultimi anni il *DNA barcoding* ha assunto un'importanza sempre maggiore, sia per ricerche di base che applicate e sta diventando un ottimo metodo, almeno per una prima identificazione di diversi prodotti alimentari, soprattutto a base di pesce, crostacei e molluschi (es. Nicolè *et al.* 2012).

1.3. Malacostraca

Tutti gli esemplari studiati in questo lavoro appartengono ai Malacostraca, classe che contiene tutti i crostacei utilizzati come alimenti alle nostre latitudini. Le specie di interesse alimentare incluse ne due ordini Decapoda e Stomatopoda.

Generalmente i Malacostraca hanno un corpo costituito da 19-20 segmenti, che formano il cefalon o capo (5 segmenti), il torace (8 segmenti), l'addome o pleon (6-7 segmenti) più un segmento caudale denominato telson. Il capo può essere fuso con uno o più segmenti del torace il quale è ricoperto, in parte o totalmente, da un carapace di chitina diversamente mineralizzata. Le appendici presenti sono suddivise secondo i vari segmenti del corpo in: 0-3 paia di massillipedi (solitamente adibite alla manipolazione del cibo) più un paio di antennule e uno di antenne (con funzione sensoriale ed equilibratrice durante il nuoto) nel capo; 8 paia di toracopodi con funzione ambulacrale nel torace; nell'addome, invece, sono presenti 5-6 paia di pleopodi ed un paio di uropodi associati al telson con funzione principalmente natatoria o modificati in relazione alla respirazione o alla riproduzione. I Malacostraca possiedono inoltre un paio di occhi composti, pedunculati, situati all'estremità anteriore del cefalotorace (Brusca & Brusca 1996).

Sebbene questo schema sia molto variabile, basti pensare alla presenza di chele usate per la predazione o la difesa, il numero dei segmenti rimane principalmente invariato permettendo di assegnare una specie immediatamente a questa classe.

Lo sviluppo dei Malacostraca è spesso diretto, anche se in qualche caso, dall'uovo nasce una larva detta "*nauplius*" con corpo non segmentato, tre paia di appendici e un solo occhio mediano. La larva si allunga per la formazione di nuovi segmenti insieme con i quali si formano le nuove appendici. Attraverso mute successive, si passa allo stadio di "*zoea*" e successivamente di "*mysis*" raggiungendo infine lo stadio adulto (Martin & Davis 2001).

1.3.1. Decapoda

È l'ordine di Malacostraca più noto e più rappresentato commercialmente. Le specie incluse in questo taxon sono caratterizzate dalla presenza di 5 paia di arti deambulatori. Tra i decapodi troviamo sia animali dalla forma allungata, che tozza. Anche la taglia varia molto, si passa infatti dai pochi centimetri di certe specie di gambero, fino a 30-40 cm degli astici.

Oltre a tutte le caratteristiche generali della classe dei Malacostraca, i Decapoda, sono caratterizzati da un carapace ben sviluppato che racchiude anche le camere branchiali; inoltre si distinguono dagli altri ordini per la presenza di un paio di chelipedi (primo paio di arti toracici) che possono essere anche di grandi dimensioni e vengono usate principalmente per difesa o predazione (Martin & Davis 2001).

Le specie appartenenti all'ordine Decapoda mostrano una notevole diversità morfologica e possono essere suddivise in almeno quattro gruppi principali di morfotipi (Figura 3) brevemente descritti di seguito:

- Caridoide: aspetto di gamberetto, con una taglia media di circa 10 centimetri; rivestito da un carapace non molto coriaceo, generalmente suddiviso in cefalotorace e addome. Presenta anteriormente un paio di antenne molto sviluppate con funzione sensoriale e un paio di esili chele. (Figura 3a).
- Astacoide: aspetto di gambero, con taglia variabile da pochi centimetri fino ai 40 centimetri; caratterizzato da una forma più tozza e un carapace robusto. Gli animali più grandi sono predatori attivi che individuano la preda grazie alle antenne e la catturano con i chelipedi molto sviluppati. (Figura 3b).
- Paguroide: aspetto di paguro. Decapodi caratterizzati da un addome ricurvo e molle che utilizzano le conchiglie vuote di altri molluschi come riparo portandole sempre con sé. (Figura 3c).
- Cancroide: aspetto di granchio, dotato di un robusto carapace e un paio di potenti chele usate principalmente per la predazione e per difesa, l'addome è ripiegato ventralmente e mantiene solo una funzione riproduttiva in quanto viene usato per trattenere le uova. (Figura 3d).

Essendo un gruppo così altamente diversificato, sono riusciti a colonizzare ogni tipo di ambiente acquatico, in cui molti di essi trascorrono la maggior parte del ciclo vitale inoltre possono essere rinvenuti a tutte le profondità.. Diversi decapodi, infatti, sono pelagici, ma si possono trovare anche animali bentonici, sedentari, erranti o fossori. Un numero limitato di specie è completamente adattata alla terra ferma. Molto diverse sono anche le strategie di alimentazione, infatti, possiamo trovare, all'interno di questa classe, animali sospensivori, predatori, erbivori o spazzini (Brusca & Brusca 1996).



Figura 3: forme dei Decapoda

1.3.2. Stomatopoda

È il secondo ordine di crostacei di interesse alimentare per importanza, sebbene questo gruppo conti circa 350 specie di cui solo 7 sono presenti nel bacino del Mediterraneo. Sono caratterizzati da un corpo allungato e schiacciato sul dorso con un muscolo addominale molto sviluppato. Il secondo paio di toracopodi è modificato in appendici raptatorie molto simili alle zampe protoraciche della mantide religiosa (Figura 4a). Questo carattere risulta essere senza dubbio distintivo per gli animali di questo ordine sebbene gli arti possano variare per dimensione e forma. Infatti, le specie che principalmente cacciano animali sprovvisti di guscio, hanno modificato le appendici raptatorie in una forma allungata ad arpione (Figura 4b), provvista di spine coriacee per meglio trattenere la preda; mentre negli animali che si nutrono di crostacei o bivalvi a guscio duro, sono più tozze e massicce (Figura 4c) e vengono usate per spaccare i tegumenti mineralizzati della preda.

Gli occhi degli Stomatopoda sono pedunculati e molto sviluppati, tanto da consentire una visione binoculare e da distinguere i diversi colori (Marshall *et al.* 2007). Vivono

principalmente in tane e fessure presenti sul fondale marino, solitamente a pochi metri di profondità. Fondamentale quindi è il carapace flessibile e l'addome muscoloso che permette a questi animali di girarsi all'interno di anfratti piuttosto stretti. Tale abilità permette agli Stomatopoda sia una repentina azione di fuga in caso di attacco da parte di un altro animale, sia un rapido movimento verso la preda durante la caccia (Brusca & Brusca 1996).



a) stomatopoda



b) zampa ad "arpione"

c) zampa a "mazza"

Figura 4: a) esemplare di Stomatopoda; b) e c) particolare delle zampe raptatorie

2. Obiettivo della tesi

Questa tesi di laurea si è inserita in un progetto più ampio volto ad approfondire le conoscenze riguardanti la struttura e l'evoluzione del genoma mitocondriale di crostacei Malacostraci con particolare riferimento a Decapoda e Stomatopoda, nonché a definire le relazioni filogenetiche tra le principali linee filetiche di Decapoda. Infine altro importante obiettivo del progetto è stato quello di caratterizzare molecularmente diverse specie di crostacei di interesse alimentare sequenziandone l'intero genoma mitocondriale.

Lo scopo della mia tesi è stato quello di concorrere alla realizzazione di tale progetto, intervenendo sia nelle attività di laboratorio atte a preparare il DNA genomico, per il successivo sequenziamento presso centri specializzati attraverso il metodo Illumina, sia nel contribuire alle analisi bioinformatiche dei genomi ottenuti e alla validazione mediante tecniche di PCR classiche. Gli obiettivi delle analisi bioinformatiche sono stati volti allo studio dei *gene order*, al segnale filogenetico presente nei diversi geni, al conseguente possibile utilizzo quali marcatori filogenetici o diagnostici per l'identificazione di specie, nonché al confronto tra sister taxa che si è reso possibile grazie al sequenziamento ex novo di importanti specie di interesse alimentare e non.

3. Materiali e Metodi

Tutte le attività riguardanti la mia tesi di laurea sono state svolte presso i laboratori del dipartimento di BCA (Biomedicina Comparata e Alimentazione), situati nel polo di Agripolis (Viale dell'Università, 16 – 35020 Legnaro, PD), dell'Università degli Studi di Padova.

3.1. Campioni

Gli esemplari delle specie analizzate in questo lavoro, derivano da due campionamenti svolti nel 2012 rispettivamente nell'alto Adriatico e nel Tirreno.

In totale sono state analizzate 19 specie di crostacei (Tabella 1) Malacostraci; per ogni specie sono stati prelevati diversi esemplari, i quali sono stati conservati in etanolo 100% e mantenuti alla temperatura di -20°C.

Ogni esemplare è stato identificato a livello specifico dal Dr. Enrico M. Negrisola e dai raccoglitori Dr. Emilio Reginella (Dipartimento di Biologia, Università degli Studi di Padova) e Dr. Mario Sbrana (CIBM, Centro Interuniversitario di Biologia Marina ed Ecologia Applicata "G. Bacci", Livorno).

Per l'inquadramento sistematico di ogni specie ci si è riferiti a De Grave *et al.* (2009). La validità dell'uso attuale dei nomi di specie è stata ulteriormente controllata avvalendosi del database *WoRMS - World Register of Marine Species* (<http://www.marinespecies.org>) che viene costantemente aggiornato su questo aspetto.

	Superfamiglia	Famiglia	Specie	Denominazione italiana
Ordine: Decapoda;	Paguroidea	Paguridae	<i>Pagurus cuanensis</i>	
Infraordine: Anomura	Paguroidea	Paguridae	<i>Pagurus prideaux</i>	
Ordine: Decapoda;	Nephropoidea	Nephropidae	<i>Homarus gammarus</i>	Astice
Infraordine: Astacidea	Nephropoidea	Nephropidae	<i>Nephrops norvegicus</i>	Scampo
Ordine: Decapoda;	Portunoidea	Portunidae	<i>Carcinus aestuarii</i>	Granchio da moleca
Infraordine: Brachyura	Xanthoidea	Eriphiidae	<i>Eriphia verrucosa</i>	Favollo
	Goneplacoidea	Goneplacidae	<i>Goneplax rhomboides</i>	
	Majoidea	Majidae	<i>Maja crispata</i>	
	Majoidea	Majidae	<i>Maja squinado</i>	Granseola o granceola
Ordine: Decapoda;	Alpheoidea	Alpheidae	<i>Alpheus glaber</i>	
Infraordine: Caridea	Palaemonoidea	Palaemonidae	<i>Palaemon adspersus</i>	Gamberetto
	Pandaloidea	Pandalidae	<i>Plesionika gigliolii</i>	Gobetto
Ordine: Decapoda;	Penaeoidea	Aristeidae	<i>Aristaeomorpha foliacea</i>	Gambero rosso
Subordine:	Penaeoidea	Aristeidae	<i>Aristeus antennatus</i>	Gambero viola
Dendrobranchiata	Penaeoidea	Penaeidae	<i>Melicertus kerathurus</i>	Mazzancolla
	Penaeoidea	Penaeidae	<i>Parapenaeus longirostris</i>	Gambero rosa
	Sergestoidea	Sergestidae	<i>Sergia robusta</i>	
	Penaeoidea	Solenoceridae	<i>Solenocera membranacea</i>	Gambero
Ordine: Stomatopoda		Squillidae	<i>Rissoides desmaresti</i>	

Tabella 1: indice delle specie sequenziate ex novo

3.2. Estrazione DNA

L'estrazione del DNA totale dei campioni, è stata fatta mediante colonnine Midi (ZR *Genomic DNA™ -Tissue MidiPrep*), particolarmente efficaci nell'estrarre e purificare DNA da matrici complesse.

Prima di effettuare l'estrazione, le soluzioni sono state preparate, sciolte o diluite nel proprio *buffer*, secondo le specifiche del produttore:

- La proteinasi K è stata sciolta con il *Proteinasi K Storage Buffer* fino ad una concentrazione finale di 20 µg/µl, e conservata in freezer a -20°C.
- Il *Genomic Lysis Buffer* è stato diluito con β-mercaptoetanololo, fino ad una concentrazione finale di 0,5% v/v. Il β-mercaptoetanololo è utilizzato sia come agente denaturante delle proteine sia come gruppo protettivo che inibisce l'azione di aldeidi e chetoni che possono influenzare negativamente il processo di estrazione.
- Il *2X Digestion Buffer* e il *DNA Pre-Wash Buffer* sono stati solubilizzati a 37°C per alcuni minuti.
- Un sufficiente volume di *DNA Elution Buffer* è stato aliquotato e preventivamente riscaldato a 55°C per una resa migliore.

L'estrazione è avvenuta seguendo i passaggi descritti nel protocollo riportato sotto:

1. Tagliare e preparare la porzione di tessuto da cui estrarre il DNA (≤125 mg), avendo cura di scartare l'esoscheletro e asciugare con carta assorbente il campione il più possibile dall'etanolo.
2. Inserire il campione di tessuto in una provetta falcon da 15 ml e sminuzzarlo per migliorare l'efficacia della soluzione di estrazione.
3. Preparare la soluzione di estrazione aggiungendo, nella falcon, 475 µl di H₂O, 475 µl di *2X Digestion Buffer* e 55 µl di proteinasi K.
4. Incubare a 55-56°C per 3-3,5 ore in bagno termostatico, agitando vigorosamente sul vortex le falcon ad ogni ora circa.
5. Aggiungere alla soluzione di estrazione 4 ml di *Genomic Lysis Buffer* e agitare vigorosamente sul vortex.
6. Avvitare *Zymo-Spin™ V-E column* con *Zymo-Midi Filter™* e inserirlo in una falcon da 50 ml.
7. Trasferire la soluzione di estrazione nel colonnina appena assemblata e centrifugare per 5' ≥1000 x g (massimo 2000 x g).

8. Disconnettere *Zymo-Spin™ V-E column/Zymo-Midi Filter™* e trasferire *Zymo-Spin™ V-E column* in un *Collection Tube*. Centrifugare in microcentrifuga a 10.000 x *g* per 1', per rimuovere eventuali residui della soluzione d'estrazione e scartare il precipitato.
9. Aggiungere 300 µl di *DNA Pre-Wash Buffer* e centrifugare a 10.000 x *g* per 1'; scartare il precipitato.
10. Aggiungere 400 µl di *g-DNA Wash Buffer* e centrifugare a 10.000 x *g* per 1' e scartare il precipitato; ripetere questa operazione per pulire al meglio il DNA.
11. Trasferire la colonnina in una eppendorf da 1,5 ml e lavare con 150 µl di *DNA Elution Buffer* precedentemente riscaldato, lasciandolo agire a temperatura ambiente per 1', poi centrifugare a 10.000 x *g* per 1'.
12. Recuperare il precipitato avendo cura di aliquotare 7-10 µl per le successive quantificazioni.
13. Stoccare quindi il DNA totale estratto in freezer a -20°C.

Le operazioni con il β-mercaptoetanololo, sono state eseguite sotto cappa chimica, per limitare la tossicità e l'odore sgradevole del composto. Gli scarti inoltre sono stati trattati come rifiuti chimici, a norma di legge.

3.3. Determinazione della concentrazione del DNA estratto

La concentrazione del DNA estratto, è stata effettuata utilizzando il fluorimetro *Qubit® (Qubit® 1.0 Fluorometer e dsDNA HS Assay - Invitrogen™)* che permette di calcolare la concentrazione del DNA in base alla relazione tra i due *standards* usati per la curva di calibrazione. Tale metodo è più affidabile del normale sistema spettrofotometrico (es. Nanodrop) perché i fluorofori utilizzati si intercalano esclusivamente al DNA a doppio filamento evitando così fenomeni di sovrastima legati a acidi nucleici degradati, presenti di solito in campioni conservati in etanolo.

Il protocollo fornito dal produttore, prevede:

1. Creazione della *Quant-iT® working solution* mescolando una quantità in µl di *QUANT-iT® Reagent* pari al numero di misurazioni da effettuare, con una quantità in µl, di *QUANT-iT® Buffer* pari al numero di misurazioni moltiplicato per 199.
2. Creazione degli standards

- Aliquotare 190 μl di *working solution* in 2 eppendorf.
 - Aggiungere 10 μl di STANDARD KIT.
3. Creazione dei campioni
 - Aliquotare 198 μl di *working solution* in eppendorf pari al numero di campioni da quantificare.
 - Aggiungere 2 μl DNA estratto diluito 1:20.
 4. Agitare vigorosamente i campioni e gli *standards* sul vortex prima dell'analisi.
 5. Le misurazioni vengono eseguite dopo aver tarato lo strumento con i 2 *standards* e i valori forniti vengono annotati per il successivo calcolo della concentrazione.

Il *QUANT-iT[®] Reagent* è sensibile ai raggi UV e va mantenuto al riparo da quest'ultimi così come tutte le soluzioni in cui è presente.

Come provette vengono usate quelle da 0,5 ml fornite con lo strumento, che sono state create appositamente per permettere l'analisi fluorimetrica.

L'analisi al *Qubit[®]* inoltre è sensibile alla temperatura. È quindi necessario conservare tutti i reagenti a temperatura ambiente (condizioni standard) e inserire i reagenti o creare le soluzioni per l'analisi solo quando si esegue la misurazione.

3.4. Gel elettroforesi

La qualità del DNA estratto è stata valutata mediante elettroforesi in gel di agarosio. Ogni corsa elettroforetica è stata eseguita su gel di agarosio a 1% contenente 0,1% di *Syber Safe (InvitrogenTM, SYBR[®] Safe DNA gel stain)* in buffer TAE 1X. In ogni pozzetto sono stati caricati 5 μl di prodotto di amplificazione e 5 μl di una soluzione di colorante e glicerolo per monitorare la corsa e appesantire il DNA in modo da non farlo fuoriuscire dal pozzetto. Come marker sono stati utilizzati ogni volta 1,5 μl di marker 1 Kb (*InvitrogenTM, Lambda DNA - Hind III Digested*).

3.5. Identificazione di specie

Come riportato sopra, l'identificazione degli esemplari è stata fatta su base morfologica attraverso l'utilizzo di chiavi dicotomiche. A questo tipo di identificazione è stata successivamente associata un'identificazione di tipo molecolare utilizzando il marcatore mitocondriale *cox1* che è lo standard per il *DNA barcoding* animale (Ratnasingham *et al.* 2007).

Questo punto ha richiesto l'amplificazione tramite PCR ed il successivo sequenziamento di una regione di circa 700 bp contenuta nel gene *cox1* dell'mtDNA, con il seguente protocollo termico e la seguente mix di reazione (Tabella 2):

Temperatura	Tempo		
90°C	2'	Inizializzazione	
94°C	30"	Denaturazione	} x 5 cicli
40°C	1'	Annealing	
72°C	1'	Estensione	
94°C	30"	Denaturazione	} x 30 cicli
45°C	1'	Annealing	
72°C	1'	Estensione	
72°C	5'	Estensione finale	
15°C		Conservazione	

Mix di reazione per 1 campione	Concentrazione	Volume
Buffer	25 mM	4 µl
MgCl ₂	X5	2 µl
dNTPs	2,5 mM	0,2 µl
Primer FW	10 µM	0,25 µl
Primer RV	10 µM	0,25 µl
Taq	5 U/µl	0,2 µl
H ₂ O		11,1 µl
DNA	50 ng/µl	2 µl
Volume finale		20 µl

Tabella 2: mix reazione nelle PCR per *cox1*

I *primer* utilizzati sono stati scelti tra quelli universali disponibili in letteratura (Folmer *et al.* 1994).

Primer FW (forward): LCO1490 GGTCACAAATCATAAAGATATTGG
Primer RV (reverse): HCO2198 TAACTTCAGGGTGACCAAAAAATCA

I prodotti di PCR sono stati sottoposti ad elettroforesi a 100V per 40 minuti su gel d'agarosio 1% contenente lo 0,1% di Syber Safe in buffer TAE 1X. Anche in questo caso, per pozzetto, sono stati caricati 5µl di prodotto di amplificazione e 5 µl di una soluzione di colorante. Come marker sono stati utilizzati ogni volta 1,5 µl di marker 100 bp (*InvitrogenTM, DNA Ladder N3231S*).

I prodotti di PCR che hanno rivelato la presenza di una singola banda in elettroforesi, sono stati successivamente purificati tramite l'utilizzo di *Exosap (ExoSAP-IT® - USB*

corporation), un preparato che integra sia una fosfatasi alcalina per degradare i dNTPs rimasti, sia una esonucleasi 1 per degradare il DNA non amplificato. In seguito sono stati inviati all'azienda BMR Genomics (PD), per il sequenziamento automatico con metodo Sanger.

Una volta sequenziati entrambi i filamenti, i relativi elettroferogrammi sono stati controllati ed eventualmente corretti con il programma *Chromas Lite* (v2.1.1). La sequenza consenso è stata poi prodotta mediante il programma *DNASTAR*[®].

La sequenza consenso è stata infine confrontata con quelle disponibili nella banca dati BOLD per l'identificazione di specie.

3.6. Sequenziamento

Il sequenziamento dell'intero DNA mitocondriale delle specie è avvenuto tramite metodo Illumina inviando i campioni all' UC Davis Genome Center (CA); secondo le specifiche del centro di sequenziamento, i campioni sono stati inviati con una concentrazione DNA finale di 100 ng/μl. Inoltre, circa 0,5μg di DNA per ogni campione, sono stati conservati in freezer a -20°C per eventuali PCR, necessarie per una conferma o ri-sequenziamento di regioni particolarmente problematiche.

Per verificare che il risultato ottenuto dal sequenziamento mediante tecnologia Illumina fosse accurato sono state inoltre predisposte delle reazioni di PCR, ristrette a porzioni di mtDNA, ed i prodotti di tali reazioni sono stati successivamente sequenziati tramite metodo Sanger automatico inviando i campioni a BMR Genomics (PD).

3.6.1. Correzione dell'elettroferogramma e creazione della sequenza consenso

Gli elettroferogrammi ottenuti tramite sequenziamento Sanger, sono stati visualizzati con il software *Chromas Lite* (v2.1.1) per valutarne il segnale, la corretta chiamata delle singole basi rispetto ai picchi e correggerne le eventuali discrepanze. Le sequenze corrette, sono state processate con l'applicativo *SeqMan* (v5.01) contenuto nel pacchetto *DNASTAR*[®] (Lasergene - Madison, WI), per ricavarne la sequenza consenso.

3.6.2. Next Generation sequencing con tecnologia Illumina

È una metodica di *Next Generation Sequencing* introdotta nel 2006, che consente di sequenziare delle corte sequenze (*short reads*) di DNA. Si basa sulla sonicazione in soluzione del DNA stampo che genera frammenti casuali di circa 100-200 bp ai quali vengono legati degli adattatori. I frammenti così modificati vengono fatti ibridare a oligonucleotidi di ancoraggio complementari agli adattatori, fissati ad una superficie (*slide*) otticamente trasparente. A questo punto, i frammenti sono amplificati mediante una PCR “a ponte” che si basa sul ripiegamento ad arco del filamento del DNA, ibridandosi all’oligonucleotide di ancoraggio adiacente. Così facendo, cicli multipli di amplificazione creano dei *clusters* a partire da un singolo frammento di DNA. La reazione di sequenziamento, quindi, prevede l’ibridazione di questi filamenti con un *primer*, l’aggiunta di quattro nucleotidi terminatori “reversibili”, marcati con fluorofori differenti e di una DNA polimerasi (Shen *et al.* 2005).

Ad ogni ciclo la polimerasi, incorpora un nucleotide marcato che quando viene eccitato emette fluorescenza, inoltre permette di fermare reversibilmente la reazione consentendo alla macchina di registrare il segnale; nel ciclo successivo, la reazione riprende incorporando il nucleotide successivo. Questa procedura viene ripetuta per innumerevoli di cicli, fino al completamento della sequenza del DNA stampo (Figura 5) (Meyer & Kircher 2010).

Questa tecnica permette di sequenziare più di un miliardo di basi ad ogni seduta, attraverso la produzione di circa 50×10^6 clusters permettendo un’elevata copertura (*coverage*) del genoma in esame (Shen *et al.* 2005).

La piattaforma Illumina viene continuamente aggiornata e perfezionata e tali migliorie hanno consentito di aumentare la riproducibilità della frammentazione mediante sonicazione, aumentare l’efficienza della ligazione degli adattatori e ridurre i segnali di interferenza.

Il centro di Genomica di Davis, utilizza piattaforme HiSeq2500 e MiSeqs che consentono di sequenziare sia da filamento senso, che antisenso (*pair-ends sequencing*). Tale approccio aumenta notevolmente il *coverage* del DNA sequenziato, e facilita così l’assemblaggio delle *reads* con conseguente riduzione degli eventuali errori.

ILLUMINA

1. In seguito alla creazione dei clusters si prosegue con la reazione di sequenziamento in cui vengono incorporati i 4 nucleotidi marcati con differenti fluorofori
2. vengono lavati i componenti in eccesso
3. i fluorofori vengono eccitati e viene registrato il segnale
4. vengono rimossi i fluorofori e le porzioni terminatrici per continuare con l'incorporazione del nucleotide successivo

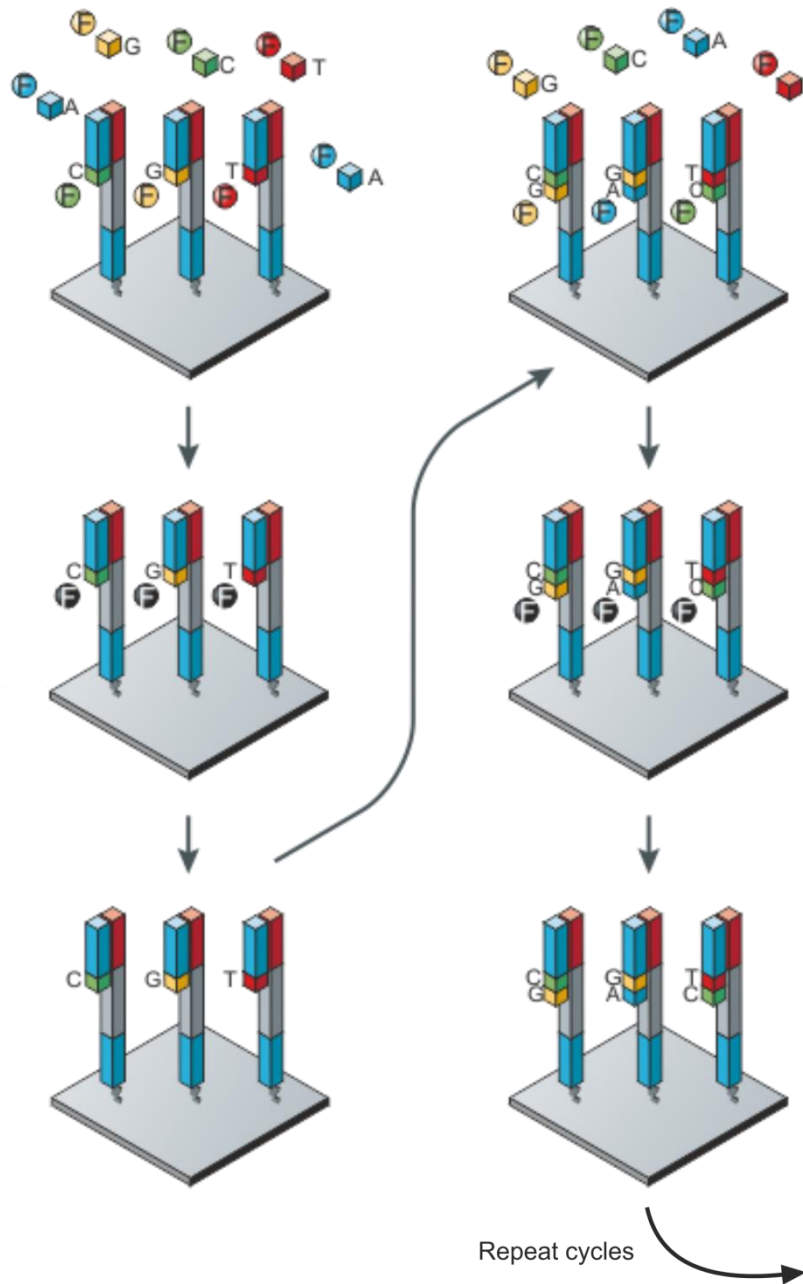


Figura 5: reazione di sequenziamento tramite tecnica Illumina

3.6.3. Assemblaggio delle sequenze illumina

L'assemblaggio delle sequenze *pair-ends* per ottenere il genoma mitocondriale completo/parziale delle 19 specie studiate in questa tesi, è stato eseguito con il software *MITObim* (<https://github.com/chrishah/MITObim>), che è un eccellente strumento bioinformatico per gestire genomi mitocondriali di organismi non-modello (Hahn *et al.* 2010). Il software permette infatti, di allineare le *reads* derivanti da “NGS” e ricostruire il genoma dell'organismo basandosi unicamente sul genoma di una specie filogeneticamente affine. *MITObim* consente una precisione superiore al 99,5 %, risultando, perciò, una valida alternativa alle tecniche tradizionali, e permette di produrre un assemblaggio accurato utilizzando *computer desktop standard* (Hahn *et al.* 2010).

Brevemente, il programma richiede un genoma mitocondriale di una specie correlata per iniziare a mappare le nuove *reads* su regioni conservate. Le *reads* vengono lette automaticamente dal software e allineate sulle regioni conservate andando così a formare uno *scaffold* di riferimento, sul quale possono attaccarsi ulteriori *reads* secondo un processo iterativo che può essere continuo o contenere gaps derivanti dalle regioni non conservate. A questo punto, *MITObim* allinea le *reads* che si sovrappongono parzialmente (con un minimo di 30 basi) o completamente alle diverse porzioni (*contigs*) dello *scaffold*, consentendo così la progressiva chiusura di gran parte/tutti i gap iniziali. Questo secondo passaggio viene replicato fino a quando sono terminate le *reads* che hanno un match con lo *scaffold*. Tale processo iterativo consente a *MITObim* di produrre l'assemblaggio completo o parziale dell'intero genoma riducendo di volta in volta i *gap* presenti tra i *contigs* e il quantitativo di memoria necessaria per il processo. Terminate le *reads* in lettura, il programma fornisce la sequenza consenso del genoma mitocondriale (Hahn *et al.* 2010). Il successo più o meno completo di questa procedura dipende dalla disponibilità di genomi ottenuti da specie filogeneticamente vicine e aventi un *gene order* non molto differente.

Le sequenze consenso ottenute sono state poi annotate ed analizzate secondo le procedure descritte nel punto 3.8.1.

3.7. Reazioni di PCR di validazione

Dopo il sequenziamento delle specie in esame e l'assemblaggio bioinformatico dei loro genomi mitocondriali, sono state predisposte delle reazioni di PCR per validare

le sequenze ottenute o per completare parti ambigue o non completamente sequenziate. I *primer* necessari per amplificare e successivamente sequenziare mediante il metodo di Sanger queste porzioni sono stati disegnati con il software *Primer3* (v. 0.4.0), reperibile online (www.bioinfo.ut.ee/primer3-0.4.0).

Le regioni sono state amplificate con il seguente protocollo termico e la seguente mix (Tabella 3).

Temperatura	Tempo		
90°C	2'	Inizializzazione	} x 30 cicli
94°C	45"	Denaturazione	
Gradiente	45"	Annealing	
72°C	1'-3'	Estensione	
72°C	5'	Estensione finale	
15°C		Conservazione	

Mix di reazione per 1 campione	Concentrazione	Volume
Buffer	25 mM	4 µl
MgCl ₂	X5	1,8 - 2,8 µl
dNTPs	2,5 mM	0,2 µl
Primer FW	10 µM	0,25 µl
Primer RV	10 µM	0,25 µl
Taq	5 U/µl	0,2 µl
H ₂ O		11,3 -10,3 µl
DNA	50 ng/µl	2 µl
Volume finale		20 µl

Tabella 3: mix di reazione nelle PCR di chiusura.

Sia la temperatura di *annealing* che il tempo di estensione del filamento, sono stati aggiustati di volta in volta per ottimizzare la PCR. Lo stesso è avvenuto con la concentrazione di MgCl₂ per favorire o meno la specificità della polimerasi.

Anche in questo caso, il risultato è stato controllato con una corsa elettroforetica con le stesse condizioni utilizzate per l'identificazione di specie. Il marker usato di volta in volta è stato scelto a seconda della prevista lunghezza degli ampliconi.

3.8. Analisi bioinformatiche

3.8.1. Dataset e annotazione dei genomi

Oltre ai 19 nuovi genomi mitocondriali sequenziati per questa tesi, sono stati utilizzati altri 67 genomi per la costituzione del dataset di analisi. Questi genomi, sono stati ottenuti dalla banca dati Genbank. Il dataset finale (Tabella 4) contiene tutte le sequenze complete di genomi mitocondriali disponibile in GenBank per quanto concerne gli ordini Decapoda, Stomatopoda e Euphausiacea.

I nuovi mtDNA, sono stati annotati seguendo un protocollo ben consolidato nel laboratorio nel quale ho svolto la mia tesi (Negrisolo *et al.* 2004a; Salvato *et al.* 2008; Negrisolo *et al.* 2011).

Inizialmente la sequenza dell' mtDNA è stata tradotta nelle proteine putative utilizzando il programma *Transeq* disponibile presso il sito dell'*EBI* (<http://www.ebi.ac.uk/>). L'identità di tali proteine è stata stabilita utilizzando il programma *BLAST* disponibile presso il sito dell'*NCBI* (<http://www.ncbi.nlm.nih.gov/>). Le estremità di ciascun gene codificante per proteine (PCG) sono state determinate come descritto di seguito. L'estremità 5' è stata identificata come posizionata sul primo codone in-frame della ORF (*Open Reading Frame*), accettato per l' mtDNA degli invertebrati (ATN, GTG, TTG, GTT) (Wolstenholme 1992), che non si sovrappone a monte con un PCG posto sullo stesso filamento. Le sole eccezioni sono rappresentate da *atp6* e *nad4* per i quali è stato dimostrato esistere una sovrapposizione rispettivamente con *atp8* e *nad4L* in numerosi genomi (Wolstenholme 1992). L'estremità 3' di ogni PCG è stata identificata come coincidente col primo codone di STOP in-frame. Quando il codone di stop è collocato all'interno di un secondo gene posto a valle e sullo stesso filamento del PCG in esame, è stato postulato un codone di stop incompleto (TA o T) adiacente al gene a valle. Il codone di stop incompleto viene poi poliadenilato a TAA nel processo di trascrizione (Wolstenholme 1992). Infine sono stati condotti allineamenti *pair-wise*, effettuati con *ClustalW* (Thompson *et al.* 1994), con geni ortologi in modo da definire in maniera ancora più completa e accurata i limiti di ciascun PCG. Indipendentemente dal codone di *start* si è sempre assunto che il primo amminoacido di ciascuna proteina fosse una Metionina in accordo con quanto dimostrato per altri genomi mitocondriali (Fearnley & Walker, 1987; Smith & Marcker 1968).

Gli RNA *transfer* sono stati identificati mediante il programma *tRNAscan-SE* (<http://lowelab.ucsc.edu/tRNAscan-SE/>) o identificati manualmente come sequenze

dotate del giusto anticodone e capaci di ripiegarsi e formare la classica struttura secondaria a quadrifoglio dei tRNA (Wolstenholme 1992). I limiti del gene ribosomiale *rrnL* sono stati definiti dalle estremità della coppia di tRNA *trnL1* e *trnV* o dalle estremità di *nad1* e *trnV* nel caso del genere *Pagurus*. L'estremità 3' di *rrnS* è stata identificata con l'inizio del gene *trnV* mentre per l'estremità 5' ci è avvalsi del confronto con geni ortologi di altre specie o con modelli di struttura secondaria sviluppati nel laboratorio in cui ho svolto la mia tesi (Negrisolo *et al.*, in preparazione). Nel caso del genere *Pagurus*, *rrnS* è stato prima identificato attraverso le estremità di *trnV* e *trnW* e poi verificato attraverso confronto e *modelling* della struttura secondaria.

Lo stesso approccio è stato applicato agli mtDNA ottenuti da GenBank, riannotandoli manualmente, in modo da avere annotazioni di alta qualità anche per questi genomi. Infatti, in molti casi, i genomi presenti in GenBank hanno una cattiva/pessima annotazione (Negrisolo *et al.*, in preparazione). Questa strategia ha comportato in numerosi casi la modifica dei confini dei geni stessi.

	Superfamiglia	Famiglia	Specie	Denominazione italiana	Accession No.
Ordine: Euphausiacea	Euphausiidae	Euphausiidae	<i>Euphausia pacifica</i>		NC_016184
Ordine: Stomatopoda		Gonodactylidae	<i>Gonodactylus chiragra</i>		NC_007442
		Lysiosquillidae	<i>Lysiosquilla maculata</i>		NC_007443
		Squillidae	<i>Harpisquilla harpax</i>		NC_006916
		Squillidae	<i>Oratosquilla oratoria</i>		NC_014342
		Squillidae	<i>Rissoidea desmaresti</i>	questa tesi	
		Squillidae	<i>Squilla empusa</i>		NC_007444
		Squillidae	<i>Squilla mantis</i>	Pannocchia o canocchia	NC_006081
Ordine: Decapoda; Infraordine: Achelata; Sottordine: Pleocyemata;	Palinuroidea	Palinuridae	<i>Panulirus homarus</i>	Aragosta tropicale	NC_016015
		Palinuridae	<i>Panulirus japonicus</i>	Aragosta tropicale	NC_004251
		Palinuridae	<i>Panulirus ornatus</i>	Aragosta tropicale	NC_014854
		Palinuridae	<i>Panulirus stimpsoni</i>	Aragosta tropicale	NC_014339
		Palinuridae	<i>Panulirus versicolor</i>	Aragosta tropicale	NC_017868
		Scyllaridae	<i>Scyllarides latus</i>	Cigala o Magnosa	NC_020022
	Synaxidae	<i>Palinurellus wieneckii</i>		KC847078	
Ordine: Decapoda; Infraordine: Anomura; Sottordine: Pleocyemata;	Galatheoidea	Porcellanidae	<i>Neopetrolisthes maculatus</i>		NC_020024
		Lithodidae	<i>Paralithodes brevipes</i>	Granchio reale	NC_021458
		Lithodidae	<i>Paralithodes camtschaticus</i>	Granchio reale	NC_020029
		Paguridae	<i>Pagurus cuanensis</i>	questa tesi	
		Paguridae	<i>Pagurus longicarpus</i>		AF150756
		Paguridae	<i>Pagurus prideaux</i>	questa tesi	
Ordine: Decapoda; Infraordine: Astacoidea; Sottordine: Pleocyemata;	Astacoidea	Cambaridae	<i>Cambaroides similis</i>		NC_016925
		Cambaridae	<i>Procambarus clarkii</i>	Gambero della Louisiana	NC_016926
		Cambaridae	<i>Procambarus fallax</i>		NC_020021
		Nephropidae	<i>Homarus americanus</i>	Astice americano	NC_015607
		Nephropidae	<i>Homarus gammarus</i>	Astice	questa tesi
		Nephropidae	<i>Nephrops norvegicus</i>	Scampo	questa tesi
Ordine: Decapoda; Infraordine: Brachyura; Sottordine: Pleocyemata;	Bythograeoidea	Bythograeidae	<i>Austinoagraea alayseeae</i>		NC_020314
		Bythograeidae	<i>Austinoagraea rodriguezensis</i>		NC_020312
		Bythograeidae	<i>Gandalfus yunohana</i>		NC_013713
		Goneplacidae	<i>Goneplax rhomboides</i>	questa tesi	
		Grapsidae	<i>Pachygrapsus crassipes</i>		KC878511
		Varunidae	<i>Eriocheir hepuensis</i>		NC_011598
		Varunidae	<i>Eriocheir japonica</i>		NC_011597
		Varunidae	<i>Eriocheir sinensis</i>	Granchio cinese	NC_006992
		Xenograpsidae	<i>Xenograpsus testudinatus</i>		NC_013480
		Majidae	<i>Chionoecetes japonicus pacificus</i>	Granchio giapponese	AB735678
		Majidae	<i>Maja crispata</i>	questa tesi	
		Majidae	<i>Maja squinado</i>	Granseola o granceola	questa tesi
		Dotillinae	<i>Ilyoplax deschampsii</i>		NC_020040
		Portunidae	<i>Callinectes sapidus</i>	Granchio nuotatore	NC_006281
		Portunidae	<i>Carcinus aestuarii</i>	Granchio da moleca	questa tesi
		Portunidae	<i>Charybdis japonica</i>		NC_013246
		Portunidae	<i>Portunus trituberculatus</i>	Granchio	NC_005037
		Portunidae	<i>Scylla olivacea</i>		NC_012569
		Portunidae	<i>Scylla paramamosain</i>		NC_012572
		Portunidae	<i>Scylla serrata</i>	Granchio indopacifico	NC_012565
		Portunidae	<i>Scylla tranquebarica</i>		NC_012567
		Potamidae	<i>Geothelphusa dehaani</i>		NC_007379
		Eriphiidae	<i>Eriphia verrucosa</i>	Favollo	questa tesi
		Eriphiidae	<i>Pseudocarcinus gigas</i>		NC_006891
Ordine: Decapoda; Infraordine: Caridea; Sottordine: Pleocyemata;	Alpheoidea	Alpheidae	<i>Alpheus distinguendus</i>		NC_014883
		Alpheidae	<i>Alpheus glaber</i>	questa tesi	
		Atyidae	<i>Halocaridina rubra</i>		NC_008413
		Alvinocarididae	<i>Alvinocaris chelys</i>		NC_018778
		Alvinocarididae	<i>Alvinocaris longirostris</i>		NC_020313
		Alvinocarididae	<i>Opaepele loihi</i>		NC_020311
		Alvinocarididae	<i>Rimicaris kairei</i>		NC_020310
		Palaemonidae	<i>Exopalaemon carinicauda</i>		NC_012566
		Palaemonidae	<i>Macrobrachium lanchesteri</i>		NC_012217
		Palaemonidae	<i>Macrobrachium nipponense</i>		NC_015073
		Palaemonidae	<i>Macrobrachium rosenbergii</i>	Gambero blu	NC_006880
		Palaemonidae	<i>Palaemon adspersus</i>	Gamberetto	questa tesi
		Pandalidae	<i>Plesionika giglioli</i>	Gobetto	questa tesi
Ordine: Decapoda; Infraordine: Polychelida; Sottordine: Pleocyemata;	Eryonoidea	Polychelidae	<i>Polycheles typhlops</i>		NC_020026
Ordine: Decapoda; Infraordine: Stenopodidae; Sottordine: Pleocyemata;		Stenopodidae	<i>Stenopus hispidus</i>		NC_018097
Ordine: Decapoda; Infraordine: Thalassinidea	Axiioidea	Strahlaxiidae	<i>Neaxius glyptocercus</i>		NC_019609
		Callianassidae	<i>Corallianassa coutierei</i>		NC_020025
		Callianassidae	<i>Nihonotrypaea japonica</i>		NC_020351
		Upogebiidae	<i>Austinogebina edulis</i>		NC_019606
		Upogebiidae	<i>Upogebia major</i>		NC_019607
		Upogebiidae	<i>Upogebia pusilla</i>		NC_020023
		Thalassinidae	<i>Thalassinia kelanang</i>		NC_019608
Ordine: Decapoda; Sottordine: Dendrobranchiata	Penaeroidea	Aristeidae	<i>Aristaeomorpha foliacea</i>	Gambero rosso	questa tesi
		Aristeidae	<i>Aristeus antennatus</i>	Gambero viola	questa tesi
		Penaeidae	<i>Farfantepenaeus californiensis</i>	Mazzancolla del Pacifico	NC_012738
		Penaeidae	<i>Fenneropenaeus chinensis</i>	Mazzancolla del Pacifico	NC_009679
		Penaeidae	<i>Litopenaeus stylirostris</i>	Mazzancolla del Pacifico	NC_012060
		Penaeidae	<i>Litopenaeus vannamei</i>	Mazzancolla tropicale	NC_009626
		Penaeidae	<i>Marsupenaeus japonicus</i>	Mazzancolla	NC_007010
		Penaeidae	<i>Melicertus kerathurus</i>	Mazzancolla	questa tesi
		Penaeidae	<i>Parapenaeus longirostris</i>	Gambero rosa	questa tesi
		Penaeidae	<i>Penaeus monodon</i>	Gambero gigante indopacifico	NC_002184
		Solenoceridae	<i>Solenocera membranacea</i>	Gambero	questa tesi
		Sergestidae	<i>Acetes chinensis</i>		NC_017600
		Sergestidae	<i>Sergia robusta</i>	questa tesi	

Tabella 4: specie utilizzate per creare i dataset di questa tesi; in rosso, le specie sequenziate ex novo.

3.8.2. Analisi dei genomi e dei *Gene Order*

Una volta creata la sequenza consenso ed effettuata l'annotazione secondo la strategia descritta sopra, per ciascun nuovo genoma è stata calcolata la composizione in basi, la percentuale di A+T e G+C, nonché gli $ATskew = (A-T)/(A+T)$ e $GC-skew = (G-C)/(G+C)$ che permettono di descrivere le asimmetrie composizionali (Perna & Kocher 1995). Tutti i calcoli si riferiscono all'*α-strand*. Le composizioni sono state determinate con il programma *EditSeq* contenuto nel pacchetto *DNASTAR*[®] mentre i vari calcoli sono stati effettuati con un programma di foglio elettronico (Excel - Microsoft Office).

Una volta identificato nei diversi genomi il *gene order*, quest'ultimo è stato studiato con il programma *CREx* (*Common interval Rearrangement Explorer*) (<http://pacosy.informatik.uni-leipzig.de/crex>) (Bernt *et al.* 2007). Questo software consente di analizzare l'evoluzione del *gene order* grazie all'utilizzo di algoritmi di evoluzione genomica basati sui modelli di TDRL e di inversione intramitocondriale (Bernt *et al.* 2007).

3.8.3. Allineamenti multipli

Per ciascuno dei 13 PCG è stato creato un dataset indipendente che poi è stato allineato utilizzando la *pipeline* implementata nel server *TranslatorX* (<http://translatorx.co.uk/>) (Abascal *et al.* 2010). Questo server consente di allineare le sequenze nucleotidiche utilizzando come *backbone* l'allineamento multiplo ottenuto dalle corrispondenti proteine. Gli allineamenti ottenuti con questo approccio producono risultati più accurati e biologicamente realistici nelle analisi filogenetiche e di evoluzione molecolare (Abascal *et al.* 2010).

Per evitare la possibile violazione del principio di *positional homology* (Abascal *et al.* 2010), gli allineamenti prodotti col programma *MAFFT* (Kato *et al.* 2005) implementato in *TranslatorX*, sono stati ulteriormente processati con il programma *GBlocks* (sempre disponibile nel pipeline) che seleziona blocchi di posizioni conservate, riducendo considerevolmente la possibilità di violazione del principio citato sopra. (Abascal *et al.* 2010).

Gli allineamenti così prodotti sono stati salvati in formato FASTA e caricati in *MEGA5* (Tamura *et al.* 2011) per le successive analisi descritte di seguito.

Sono stati inoltre creati dataset formati attraverso il concatenamento di un numero variabile di PCG. Questi ultimi dataset sono stati prodotti tenendo conto dei risultati relativi al segnale filogenetico dei singoli geni (vedi sotto).

3.8.4. Determinazione del segnale filogenetico

Il segnale filogenetico presente in un dataset è rappresentato dalla tendenza che hanno specie strettamente imparentate tra di loro di assomigliarsi, per i caratteri considerati, in modo nettamente più marcato di specie non correlate selezionate casualmente dallo stesso dataset (Munkemuller *et al.* 2012). Nel caso dei dataset molecolari questo significa che per utilizzare delle sequenze nucleotidiche/amminoacidiche al fine di dedurre i rapporti filogenetici tra gruppi di organismi, bisogna assicurarsi che queste non abbiano subito un processo di diversificazione così esteso che la diversità/somiglianza che osserviamo nel dataset stesso risulti dal puro effetto di sostituzioni casuali, impedendo così di tracciare la storia evolutiva delle sequenze a partire da un comune antenato (Guindon *et al.* 2010).

La stima del segnale filogenetico presente in ciascun PCG e corrispondente proteina è stata effettuata secondo la strategia descritta di seguito.

Un primo approccio, ha previsto il calcolo delle matrici delle distanze (fra coppie di campioni) basate su un modello di sostituzione nucleotidica/amminoacidica molto semplice che esprime il livello di divergenza tra una coppia di sequenze come la frazione di siti differenti rispetto alla lunghezza totale dell'allineamento (*p-distance*) implementato all'interno di *MEGA5* (Tamura *et al.* 2011), e su un modello di evoluzione molecolare complesso GTR/MtART + G, (nucleotidi/amminoacidi) implementato nel utilizzando il software *RaXML* (Stamatakis 2006). Questo secondo tipo di modelli tiene conto anche delle sostituzioni multiple che possono essersi verificate nel corso dell'evoluzione e fornisce quindi una descrizione più realistica del processo di evoluzione molecolare (Felsenstein, 2004).

Sono stati effettuati due tipi confronti tra le distanze calcolate. Il primo ha previsto il calcolo della differenza (DIF) tra distanze ottenute con modello complesso e distanze ottenute con modello semplice. Il secondo confronto è stato attuato facendo il rapporto tra distanze ottenute col modello semplice e distanze ottenute con modello complesso (vedi dettagli nella sezione risultati).

Inoltre, gli stessi allineamenti *gene by gene* sono stati analizzati con il software *TREEPUZZLE* (v5.2) (Schmidt *et al.* 2002), applicando il metodo di *likelihood*

mapping che permette di stimare la quantità di segnale filogenetico presente in dataset mediante un algoritmo di *quartet puzzling* (Strimmer & von Haeseler 1997). Il *likelihood mapping* opera secondo la strategia descritta brevemente di seguito. Date 4 sequenze (1 quartetto), esistono tre alberi senza radice che ne descrivono le relazioni filogenetiche. Questi tre alberi possono essere equiprobabili (quartetto irrisolto), 2 alberi possono essere migliori del terzo (quartetto parzialmente risolto), oppure un albero è nettamente migliore degli altri (quartetto completamente risolto). Queste tre possibilità si stabiliscono mediante il calcolo del valore di *likelihood* per ogni uno dei tre alberi. Ogni quartetto Q_i può quindi essere descritto come un vettore caratterizzato dai valori di *likelihood* dei tre alberi Q_i ($L1_i, L2_i, L3_i$). $L1_i, L2_i, L3_i$ possono essere considerati come coordinate baricentriche di Q_i e quindi il risultato dell'analisi condotta su ciascun quartetto Q_i può essere proiettato su un triangolo equilatero, sul qual si individuano 7 bacini principali di attrazione (Figura 6). I punti che mappano sui bacini $A_1, A_2,$ e A_3 indicano quartetti completamente risolti, i punti su ricadenti in $A_{12}, A_{23},$ e A_{13} quartetti parzialmente risolti, e i punti che mappano su (A_{123}) quartetti completamente irrisolti (Strimmer & von Haeseler 1997). La quantità di segnale filogenetico presente in un dataset è proporzionale alle percentuale di quartetti completamente risolti cioè quelli contenuti in $A_1, + A_2, + A_3$ (Strimmer & von Haeseler 1997). Pertanto si effettua il *likelihood mapping* per 10.000 quartetti e si valuta poi la percentuale di quartetti completamente risolti.

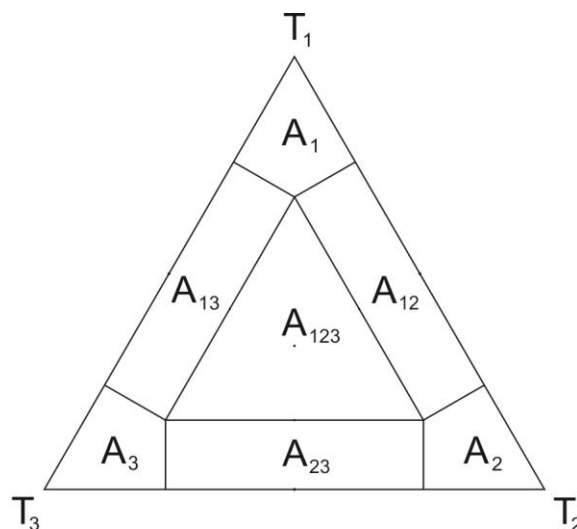


Figura 6: rappresentazione grafica dell'analisi con l'algoritmo quartet puzzling

L'identificazione del segnale filogenetico presente negli interi codoni, nelle singole posizioni dei codoni e nei corrispondenti allineamenti amminoacidici ha consentito di

identificare i dataset con miglior segnale e di concatenarli in differenti macro-allineamenti (vedi risultati).

3.8.5. Costruzione degli alberi filogenetici

Una volta identificato il segnale filogenetico di ciascun dataset sono stati creati diversi allineamenti (ALN) multigenici/multiproteici che sono elencati nella Tabella (Tabella 5) riportata sotto. Gli ALN 13.PCG e 13.PCG.aa contengono tutti i geni codificanti per proteine e i rispettivi polipeptidi. Gli altri ALN sono stati ottenuti da questi rimuovendo selettivamente geni/proteine e/o posizioni dei codoni tenendo conto dei risultati di segnale filogenetico (vedi risultati). Le analisi filogenetiche sono state effettuate con il software *RaXML* che è particolarmente efficace nelle analisi filogenetiche basate sul metodo di *maximum likelihood* (ML) applicate a dataset di grandi dimensioni (Stamatakis 2006). Il supporto statistico alle topologie è stato calcolato secondo il metodo di *bootstrap* (Felsenstein 1985). Per limitare i tempi computazionali sono state effettuate 100 repliche per ogni analisi. L'eterogeneità del processo di sostituzione per i vari siti dei differenti ALN è stata modellata con l'algoritmo CAT che è superiore alla funzione statistica gamma quando si analizzano dataset ad alta eterogeneità quali sono quelli composti da geni mitocondriali (Stamatakis 2006). Infine, gli alberi generati sono stati visualizzati con il software *Njplot* (Perrière & Gouy 1996).

Acronimo	Geni concatenati
13.PCG	<i>atp6, atp8, cob, cox1, cox2, cox3, nad1, nad2, nad3, nad4, nad4L, nad5, nad6</i>
8.PCG	<i>atp6, cob, cox1, cox2, cox3, nad1, nad4, nad5</i>

Acronimo	Posizioni considerate
.p2	analisi del segnale relativo alle seconde posizioni dei codoni
.p12	analisi del segnale relativo alle prime e seconde posizioni dei codoni
.p123	analisi del segnale relativo alle prime, seconde e terze posizioni dei codoni
.aa	analisi del segnale relativo agli amminoacidi

Tabella 5: nomenclatura dei dataset

4. Risultati

4.1. Descrizione dei genomi e *Gene Order*

4.1.1. Principali caratteristiche dei nuovi mtDNA

Per 18 dei 19 taxa studiati è stato possibile determinare la sequenza completa dell'mtDNA (Tabella 6). Unica eccezione è rappresentata da *Nephrops norvegicus* per il quale mancano ancora i dati relativi ad una piccola porzione (Tabella 6).

Il *coverage* per ciascuna sequenza consenso è risultato $\geq 40X$. Le sequenze consenso ottenute attraverso l'assemblaggio di *reads* derivate dall'approccio Illumina, sono state ulteriormente validate, sequenziando segmenti di genoma con il metodo *standard* di riferimento di *Sanger*. Questi confronti non hanno rivelato differenze.

Le dimensioni dei 18 mtDNA completi variano da 14.316 bp di *Homarus gammarus* a 16643 bp di *Alpheus glaber* (Tabella 6). *Nephrops norvegicus* seppure incompleto risulta il più lungo dei genomi sequenziati (16865 bp.). Tutti i genomi sono ricchi in A+T ma la percentuale varia dal 63.77% di *A. glaber* al 72.73% di *Pagurus prideaux*. Dieci dei 18 mtDNA completi presentano un *AT-skew* negativo, mentre per i restanti risulta positivo. In entrambi i casi, lo *skew* è limitato ($\leq |0.08|$). Tutte le specie eccetto *H. gammarus* presentano un *GC-skew* negativo. Il *GC-skew* risulta in valore assoluto maggiore del corrispettivo *AT-skew* in tutte le specie analizzate.

Specie	Lunghezza (bp)	A + T %	G + C %	A	G	T	C	AT skew	GC skew
<i>Pagurus cuanensis</i>	16.091	72,46	27,54	5915	1833	5745	2598	0,0145798	-0,1726473
<i>Pagurus prideaux</i>	16.252	72,73	27,27	5865	1842	5955	2590	-0,0076142	-0,1687726
<i>Homarus gammarus</i>	14.316	68,28	31,72	4508	2485	5312	2011	-0,0818737	0,105427
<i>Nephrops norvegicus</i>	16.865	71,49	28,51	5828	2073	6228	2736	-0,0331785	-0,1378665
<i>Carcinus aestuarii</i>	15.698	66,19	33,81	4894	2346	5496	2962	-0,0579403	-0,1160512
<i>Eriphia verrucosa</i>	15.605	68,39	31,61	5189	1737	5484	3195	-0,0276398	-0,2956204
<i>Goneplax rhomboides</i>	15.707	70,81	29,19	5534	1682	5588	2903	-0,0048552	-0,2663032
<i>Maja crispata</i>	16.592	70,30	29,70	5573	1847	6091	3081	-0,0444102	-0,2504058
<i>Maja squinado</i>	16.598	70,80	29,20	5601	1829	6151	3017	-0,0468005	-0,2451506
<i>Alpheus glaber</i>	16.643	63,77	36,23	5435	2320	5179	3709	0,0241191	-0,2303865
<i>Palaemon adspersus</i>	15.740	66,79	33,21	5407	2051	5105	3177	0,0287291	-0,2153787
<i>Plesionika giglioli</i>	15.909	64,62	35,38	5149	2308	5132	3320	0,0016535	-0,1798152
<i>Aristaeomorpha foliacea</i>	15.940	66,24	33,76	5239	2126	5319	3256	-0,0075772	-0,2099591
<i>Aristeus antennatus</i>	15.938	65,74	34,26	5328	2105	5149	3356	0,017085	-0,2290789
<i>Melicertus kerathurus</i>	15.976	66,50	33,50	5354	2049	5270	3303	0,0079066	-0,2343049
<i>Parapenaeus longirostris</i>	15.929	68,76	31,24	5733	1761	5219	3216	0,0469321	-0,2923448
<i>Sergia robusta</i>	16.067	68,66	31,34	5384	2158	5648	2877	-0,0239304	-0,1428004
<i>Solenocera membranacea</i>	15.910	69,21	30,79	5603	1864	5409	3034	0,0176171	-0,238873
<i>Rissoides desmaresti</i>	15.791	67,97	32,03	5342	2264	5391	2794	-0,0045654	-0,1047845

Tabella 6: caratteristiche degli mtDNA sequenziati *ex novo*

4.1.2. Analisi dei Gene Order

Una volta ottenute le sequenze consenso parziali/complete per i 19 mtDNA, e dopo aver effettuato la loro annotazione secondo i criteri descritti nei materiali e metodi, ho utilizzato la sequenza *cox1 barcoding* per verificare, se per le specie studiate, fosse possibile effettuare un'identificazione molecolare confrontando le nuove sequenze con quelle disponibili in BOLD. In Tabella 7 sono riportati i risultati di tali indagini. Per *Plesionika gigliolii* e *Rissoides desmaresti* non è stato possibile effettuare un'identificazione dato che non esistono sequenze di riferimento per queste specie in BOLD. Rimane quindi da verificare per questi due crostacei se il *DNA barcoding* funzioni oppure no. Nel caso di *Palaemon adspersus*, l'aplotipo del mio campione presenta una percentuale di identità che varia tra 97.82% e 97.55% con aplotipi ottenuti da altri esemplari della stessa specie. In tutti i casi siamo prossimi alla soglia di divergenza del 3% che di solito demarca la differenza tra specie diverse e di conseguenza BOLD non assegna con certezza il mio esemplare alla specie individuata con l'identificazione morfologica. Per tutte le altre specie invece il *DNA barcoding* è stato sufficiente per un'identificazione certa.

Le sequenze di tutti gli mtDNA sono state orientate in modo tale che il primo gene del filamento α coincidesse con *cox1*. Questo approccio mi ha permesso di confrontare i *gene order* individuati nei 19 mtDNA con quelli disponibili in letteratura.

Tutti i genomi completamente sequenziati presentano l'intero set di 37 geni che caratterizzano il mtDNA animale e cioè 13 geni codificanti per proteine, 22 geni codificanti tRNA e 2 geni codificanti rRNA (Figura 7). Nei genomi i geni possono essere sovrapposti, contigui, o spazati sullo stesso filamento o sui filamenti complementari (Figura 7).

Superfamiglia	Famiglia	Specie	Identificazione in BOLD	% di identità con aplotipi esistenti
Paguroidea	Paguridae	<i>Pagurus cuanensis</i> *, NEW	SI	100,00%
Paguroidea	Paguridae	<i>Pagurus prideaux</i> *, NEW	SI	100,00%
Nephropoidea	Nephropidae	<i>Homarus gammarus</i>	SI	99,53%
Nephropoidea	Nephropidae	<i>Nephrops norvegicus</i> *, NEW	SI	100,00%
Portunoidea	Portunidae	<i>Carcinus aestuarii</i> *	SI	100,00%
Xanthoidea	Eriphiidae	<i>Eriphia verrucosa</i> *	SI	100,00%
Goneplacoidea	Goneplacidae	<i>Goneplax rhomboides</i> *, NEW	SI	100,00%
Majoidea	Majidae	<i>Maja crispata</i> *, NEW	SI	99,85%
Majoidea	Majidae	<i>Maja squinado</i> *, NEW	SI	100,00%
Alpheoidea	Alpheidae	<i>Alpheus glaber</i>	SI	100,00%
Palaemonoidea	Palaemonidae	<i>Palaemon adspersus</i> *	SI	97,82% ✕
Pandaloidea	Pandalidae	<i>Plesionika gigliolii</i>	NO	n/d
Penaeoidea	Aristeidae	<i>Aristaeomorpha foliacea</i>	SI	100,00%
Penaeoidea	Aristeidae	<i>Aristeus antennatus</i>	SI	100,00%
Penaeoidea	Penaeidae	<i>Melicertus kerathurus</i>	SI	100,00%
Penaeoidea	Penaeidae	<i>Parapenaeus longirostris</i>	SI	99,84%
Sergestoidea	Sergestidae	<i>Sergia robusta</i>	SI	100,00%
Penaeoidea	Solenoceridae	<i>Solenocera membranacea</i>	SI	100,00%
	Squillidae	<i>Rissoides desmaresti</i>	NO	n/d

Tabella 7: identificazione in BOLD delle specie sequenziate ex novo; * indica le specie con gene order differente da quello dei Pancrustacea ma già noto in letteratura; NEW indica un nuovo gene order; ✕ indica specie per cui BOLD non fornisce un'identificazione certa.

Nel caso della presenza di spaziatori intergenici questi possono essere costituiti da pochi nucleotidi (es. 4 bp, ISP^{trnR-trnN} di *Rissoides desmaresti*) oppure estendersi per centinaia di bp al di fuori del tradizionale principale ISP rappresentato dalla *AT-rich region* (es. 926 bp, ISP^{trnS2-trnQ} di *Maja squinado*).

Homarus gammarus, *Alpheus glaber*, *Plesionika gigliolii*, *Aristaeomorpha foliacea*, *Aristeus antennatus*, *Melicertus kerathurus*, *Parapenaeus longirostris*, *Sergia robusta*, *Solenocera membranacea*, *Rissoides desmaresti*, 10 delle 19 specie studiate, possiedono un gene order che è identico a quello considerato standard per i Pancrustacea (Figure 1 e 7). Mentre le altre 9 specie (*Pagurus cuanensis*, *Pagurus prideaux*, *Carcinus aestuarii*, *Eriphia verrucosa*, *Goneplax rhomboides*, *Maja crispata*, *Maja squinado*, *Palaemon adspersus*) presentano gene order differenti che sono descritti in maggior dettaglio di seguito.

a) Stomatopoda

Squillidae

Rissoides desmaresti

Lng (bp) = 15.791; % A+T = 67,97; % G+C = 32,03; AT-skew = -0,0045654; GC-skew = -0,1047845;



b) Anomura

Paguridae

*Pagurus cuanensis**, NEW

Lng (bp) = 16.091; % A+T = 72,46; % G+C = 27,54; AT-skew = 0,0145798; GC-skew = -0,1726473;



*Pagurus prideaux**, NEW

Lng (bp) = 16.252; % A+T = 72,73; % G+C = 27,27; AT-skew = -0,0076142; GC-skew = -0,1687726;



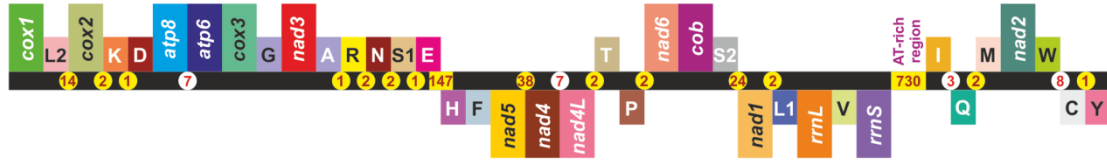
Figura 7: Organizzazione della struttura degli mtDNA sequenziati in questa tesi. Il codice colori per i diversi geni è identico a quello presentato in Figura 1. I numeri indicati lungo il genoma indicano, la lunghezza degli spaziatori intergenici (marrone su fondo giallo) o il numero di nucleotidi sovrapposti in geni contigui (rosso su fondo bianco). Lng, indica la lunghezza in paia di basi del genoma. *, gene order differente da quello standard per i Pancrustacea (vedi Figura 1); NEW, nuovo gene order (CONTINUA)

c) Brachyura

Portunidae

*Carcinus aestuarii**

Lng (bp) = 15.698; % A+T = 66,19; % G+C = 33,81; AT-skew = -0,0579403; GC-skew = -0,1160512;



Eriphiidae

*Eriphia verrucosa**

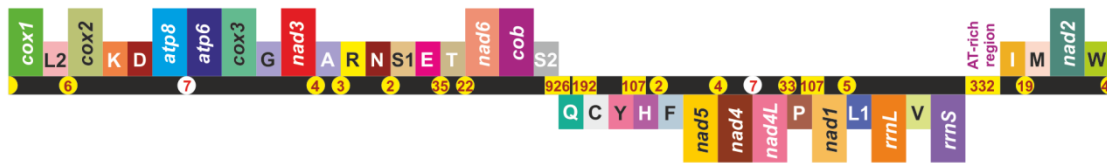
Lng (bp) = 15.605; % A+T = 68,39; % G+C = 31,61; AT-skew = -0,0276398; GC-skew = -0,2956204;



Majidae

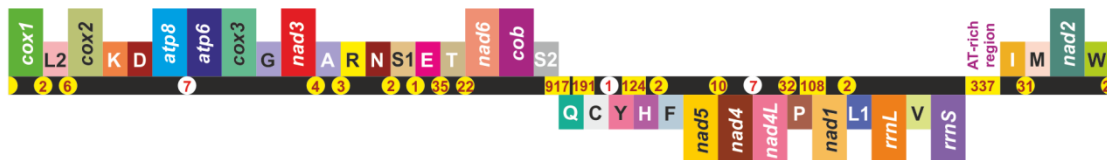
*Maja crispata**, NEW

Lng (bp) = 16.592; % A+T = 70,30; % G+C = 29,70; AT-skew = -0,0444102; GC-skew = -0,2504058;



*Maja squinado**, NEW

Lng (bp) = 16.598; % A+T = 70,80; % G+C = 29,20; AT-skew = -0,0468005; GC-skew = -0,2451506;



Goneplacidae

*Goneplax rhomboides**, NEW

Lng (bp) = 15.707; % A+T = 70,81; % G+C = 29,19; AT-skew = -0,0048552; GC-skew = -0,2663032;



Figura 7: Organizzazione della struttura degli mtDNA sequenziati in questa tesi. Il codice colori per i diversi geni è identico a quello presentato in Figura 1. I numeri indicati lungo il genoma indicano, la lunghezza degli spaziatori intergenici (marrone su fondo giallo) o il numero di nucleotidi sovrapposti in geni contigui (rosso su fondo bianco). Lng, indica la lunghezza in paia di basi del genoma. *, gene order differente da quello standard per i Pancrustacea (vedi Figura 1); NEW, nuovo gene order (CONTINUA)

d) Astacidea

Nephropidae

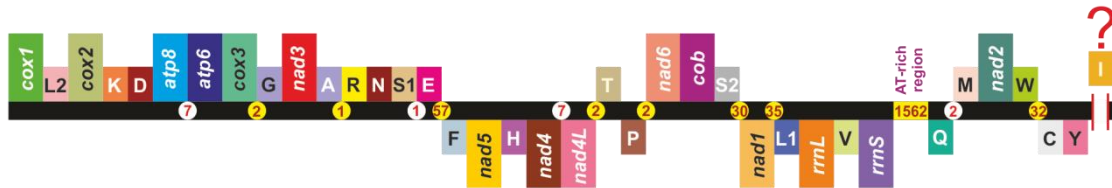
Homarus gammarus

Lng (bp) = 14.316; % A+T = 68,28; % G+C = 31,72; AT-skew = -0,0818737; GC-skew = 0,105427;



*Nephrops norvegicus**, NEW

Lng (bp) = 16.865; % A+T = 71,49; % G+C = 28,51; AT-skew = -0,0331785; GC-skew = -0,1378665;



e) Caridea

Palaemonidae

*Palaemon adspersus**

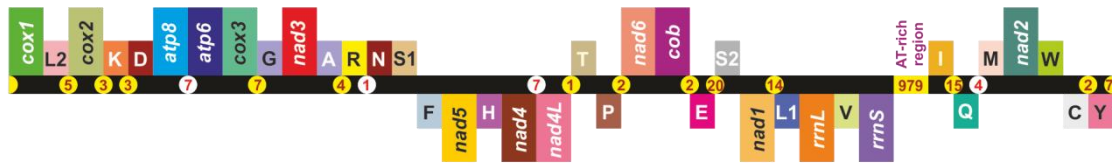
Lng (bp) = 15.740; % A+T = 66,79; % G+C = 33,21; AT-skew = 0,0287291; GC-skew = -0,2153787;



Alpheidae

Alpheus glaber

Lng (bp) = 16.643; % A+T = 63,77; % G+C = 36,23; AT-skew = 0,0241191; GC-skew = -0,2303865;



Pandalidae

Plesionika giuglioli

Lng (bp) = 15.909; % A+T = 64,62; % G+C = 35,38; AT-skew = 0,0016535; GC-skew = -0,1798152;



Figura 7: Organizzazione della struttura degli mtDNA sequenziati in questa tesi. Il codice colori per i diversi geni è identico a quello presentato in Figura 1. I numeri indicati lungo il genoma indicano la lunghezza degli spaziatori intergenici (marrone su fondo giallo) o il numero di nucleotidi sovrapposti in geni contigui (rosso su fondo bianco). Lng, indica la lunghezza in paia di basi del genoma. *, gene order differente da quello standard per i Pancrustacea (vedi Figura 1); NEW, nuovo gene order (CONTINUA)

f)

Dendrobranchiata

Aristeidae

Aristeus antennatus

Lng (bp) = 15.938; % A+T = 65,74; % G+C = 34,26; AT-skew = 0,017085; GC-skew = -0,2290789;

*Aristaeomorpha foliacea*

Lng (bp) = 15.940; % A+T = 66,24; % G+C = 33,76; AT-skew = -0,0075772; GC-skew = -0,2099591;



Solenoceridae

Solenocera membranacea

Ln (bp) = 15.910; % A+T = 69,21; % G+C = 30,79; AT-skew = 0,0176171; GC-skew = -0,238873;



Peneaeidae

Melicertus kerathurus

Lng (bp) = 15.976; % A+T = 66,50; % G+C = 33,50; AT-skew = 0,0079066; GC-skew = -0,2343049;

*Parapeneaus longirostris*

Lng (bp) = 15.929; % A+T = 68,76; % G+C = 31,24; AT-skew = 0,0469321; GC-skew = -0,2923448;



Sergestidae

Sergia robusta

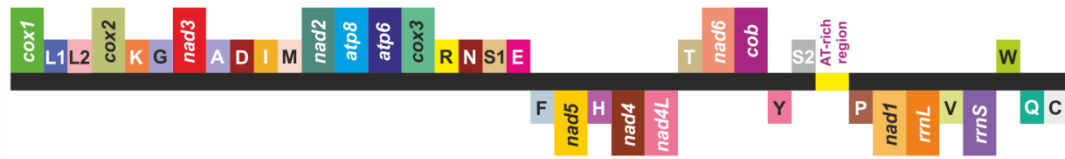
Lng (bp) = 16.067; % A+T = 68,66; % G+C = 31,34; AT-skew = -0,0239304; GC-skew = -0,1428004;



Figura 7: Organizzazione della struttura degli mtDNA sequenziati in questa tesi. Il codice colori per i diversi geni è identico a quello presentato in Figura 1. I numeri indicati lungo il genoma indicano, la lunghezza degli spaziatori intergenici (marrone su fondo giallo) o il numero di nucleotidi sovrapposti in geni contigui (rosso su fondo bianco). Lng, indica la lunghezza in paia di basi del genoma. *, gene order differente da quello standard per i Pancrustacea (vedi Figura 1); NEW, nuovo gene order

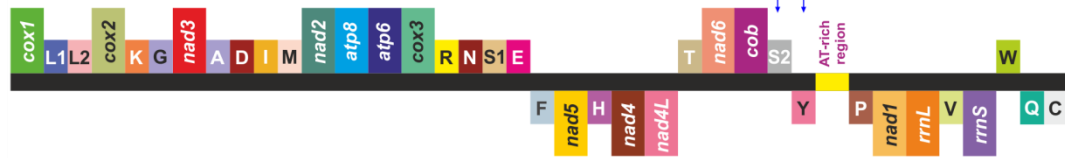
Pagurus cuanensis e *Pagurus prideaux*. Queste 2 specie presentano lo stesso *gene order*, che risulta differente rispetto a tutti i GO noti (Figura 8). Analizzando i riarrangiamenti con CREx è stato possibile tracciare un possibile scenario evolutivo in cui si è formato questo nuovo GO a partire dal GO standard dei Pancrustacea (Figura 8). Secondo quest'analisi è stata necessaria un'inversione intramitocondriale del *trnL1* seguita da tre successivi eventi TDRL per produrre il risultato finale osservato per *P. cuanensis* e *P. prideaux*. Si osservi che l'unica altra specie di *Pagurus* per cui è noto l'mtDNA completo (*P. longicarpus*) possiede un GO che differisce dal GO di *P. cuanensis* e *P. prideaux*, per la posizione di *trnS2*. Questo GO si può pensare derivato dal GO delle altre due specie di *Pagurus* per un singolo processo di TDRL. A favore di un riarrangiamento di questo tipo si ha che nel GO di *P. longicarpus* non sono più adiacenti *cob* e *trnS2*, un tratto tipico del GO standard dei Pancrustacea, mentre questa relazione è conservata in *P. cuanensis* e *P. prideaux*. Quindi il livello di riarrangiamento è maggiore in *P. longicarpus* rispetto a quanto si osserva per il GO di *P. cuanensis* e *P. prideaux*.

Pagurus longicarpus

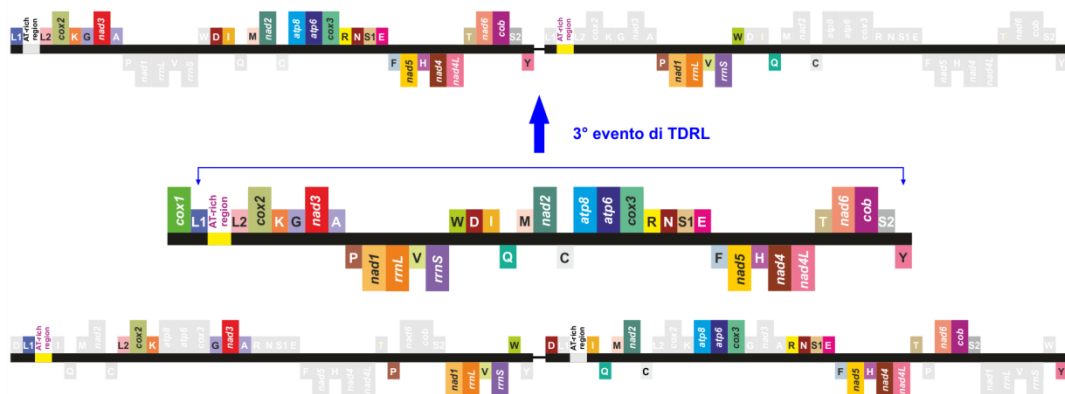


Tandem duplication and random loss

Pagurus prideaux, Pagurus cuanensis



3° evento di TDRL



2° evento di TDRL



1° evento di TDRL

Pancrustacea



Figura 8: nuovi gene order del genere *Pagurus* e eventi che ne hanno portato alla formazione. Il codice colori per i diversi geni è identico a quello presentato in Figura 1. I geni deleti successivamente alla duplicazione vengono indicati in bianco su sfondo grigio

Nephrops norvegicus. Sebbene l'mtDNA non sia completo, dall'analisi del GO dei geni noti si evince che nel corso dell'evoluzione di questa specie si è verificato lo spostamento del *trnI* dalla sua tradizionale collocazione tra *AT-rich region* e *trnQ*. Il completamento della sequenza (in fase di realizzazione) consentirà di capire se tale gene è solo traslocato o ha subito anche un processo di inversione (trasposizione inversa) (Figura 7d).

Carcinus aestuarii e *Eriphia verrucosa*. Queste due specie presentano il tipico GO dei Brachyura che differisce dal GO standard per il posizionamento di *trnH* (Figura 7c). Il GO dei Brachyura può essersi generato dal GO dei Pancrustacea attraverso un singolo evento TDRL (Figura 9).

Goneplax rhomboides. Questa specie di granchio presenta un GO precedentemente non noto, che differisce dal GO dei Brachyura per il posizionamento di *trnA*. Anche in questo caso un singolo evento TDRL spiega il passaggio al nuovo GO a partire dal GO dei Brachyura (Figura 9).

Goneplax rhomboides

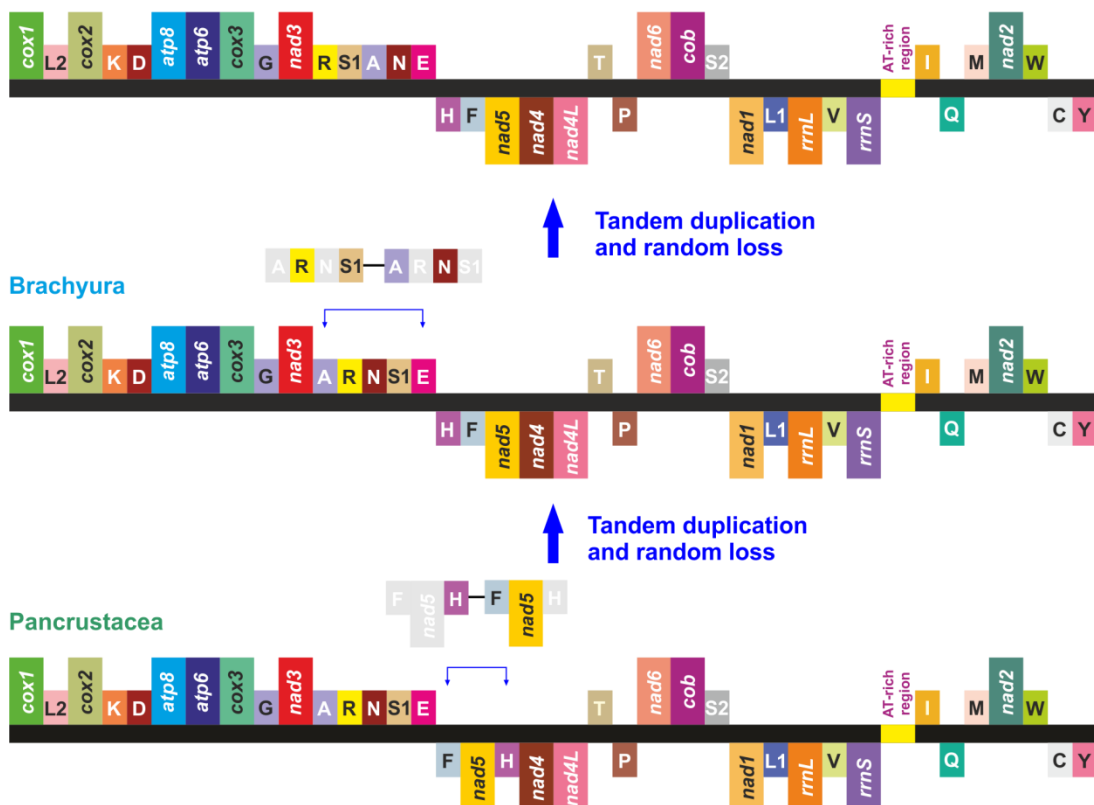


Figura 9: nuovo gene order di *Goneplax rhomboides* e eventi che ne hanno portato alla formazione. Il codice colori per i diversi geni è identico a quello presentato in Figura 1. I geni deleti successivamente alla duplicazione vengono indicati in bianco su sfondo grigio

Maja crispata e *Maja squinado*. Queste due specie possiedono un nuovo GO, che differisce vistosamente dal GO dei Brachyura (Figura 10). Anche in questo caso il nuovo GO può essere spiegato come il processo di un singolo evento TDRL che ha coinvolto però una porzione di mtDNA molto più grande di quella coinvolta nei GO degli altri rami descritti sopra. Questo GO non è comune a tutte le specie di Majidae. Infatti *Chionoecetes japonicus pacificus*, altra specie della famiglia per la quale è noto l'mtDNA, presenta il classico GO dei Brachyura.

Maja crispata, Maja squinado

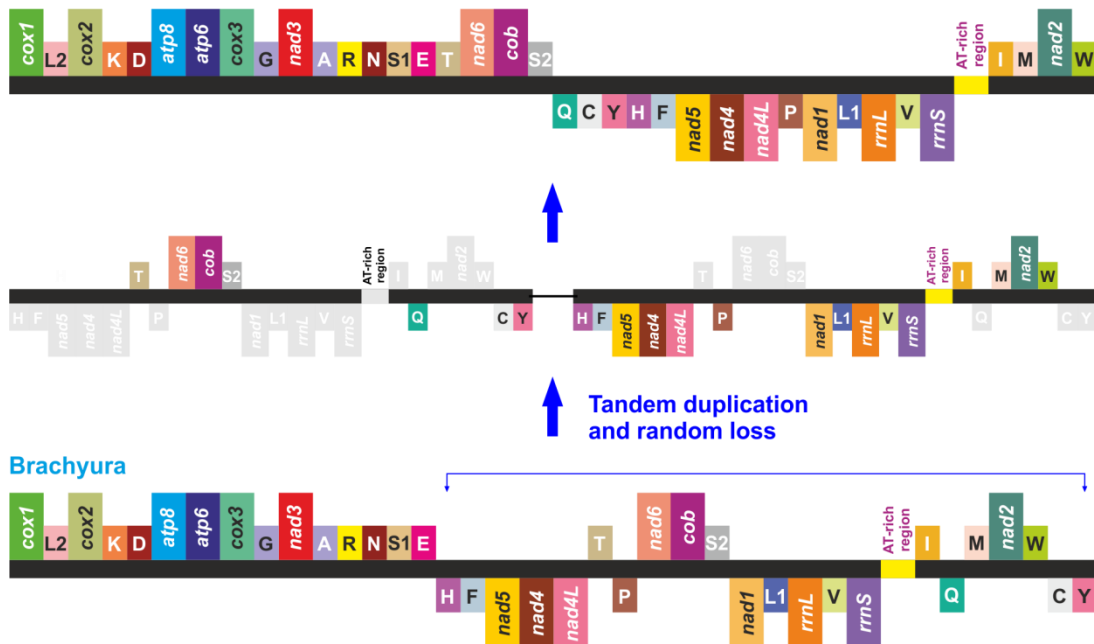


Figura 10: nuovi gene order del genere *Maja* e eventi che ne hanno portato alla formazione. Il codice colori per i diversi geni è identico a quello presentato in Figura 1. I geni deleti successivamente alla duplicazione vengono indicati in bianco su sfondo grigio

Palaemon adspersus. Questa specie presenta un GO che differisce dal GO dei Pancrustacea per il posizionamento di *trnT* e *trnP*. Questo GO è condiviso con *Exopalaemon carinicauda* mentre le specie del genere *Macrobrachium*, sempre membri della famiglia Palaemonidae, presentano il GO standard dei Pancrustacea. Anche per *P. adspersus* un singolo evento di TDRL spiega il passaggio dal GO dei Pancrustacea al nuovo GO (Figura 11).

Palaemon adspersus

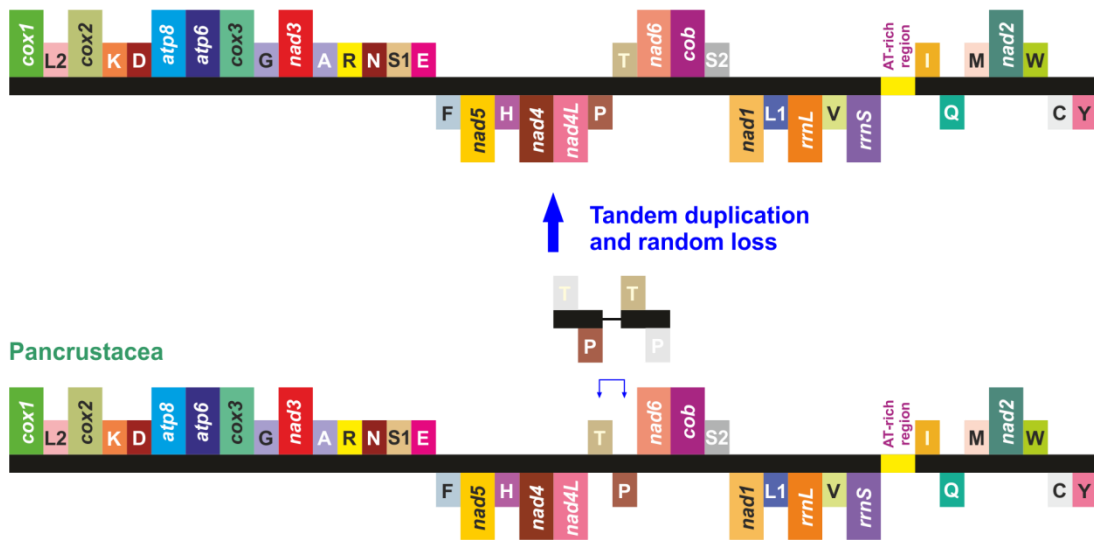


Figura 11: nuovo gene order di *Goneplax rhomboides* e eventi che ne hanno portato alla formazione. Il codice colori per i diversi geni è identico a quello presentato in Figura 1. I geni deleti successivamente alla duplicazione vengono indicati in bianco su sfondo grigio

4.1.3. Confronto delle sequenze di taxa congenerici

Il sequenziamento di specie appartenenti allo stesso genere per le quali esisteva, o è stato determinato in questa tesi l'intero mtDNA, ha consentito di effettuare alcune comparazioni *gene by gene* per verificare il livello di divergenza di ciascun marcatore ad un rango tassonomico basso. La distanza di riferimento calcolata in tutti i confronti è stata quella di Kimura 2 parametri, dato che essa è utilizzata come *gold standard* dagli algoritmi di *DNA barcoding*.

I risultati di questi confronti sono riportati in Tabella 8.

Il confronto *gene by gene* mostra come i 13 PCG hanno comportamenti differenti rispetto al livello di variabilità. Grossolanamente si possono dividere in due gruppi uno ad evoluzione più rapida che presenta distanze maggiori rispetto a quelle computate per i genomi interi ed uno ad evoluzione più lenta in cui invece le distanze sono più piccole. I geni *nad2*, *nad4* presentano sempre distanze maggiori di quelle genomiche. Lo stesso si osserva quasi sempre per *nad3*, *nad5* e *nad6* con una eccezione nel confronto tra specie di *Maja*. In questo caso però i genomi presentano grandi ISP al di fuori della *AT-rich region* che sono responsabili della maggiore divergenza delle sequenze genomiche. Anche *atp6* è collocato quasi sempre nel primo gruppo con eccezione del confronto *P. prideux* / *P. cuanensis*. Nel secondo gruppo rientrano sempre i geni *cox3*, *nad1* e *nad4L*. Anche *cox1*, che è lo standard di riferimento per il *DNA barcoding*, e *cox2* appartengono a questo secondo gruppo con una sola eccezione osservata nel confronto tra specie di *Homarus*. Molto variabile appare il comportamento di *cob* e *atp8*. Un comportamento simile ai PCG ad evoluzione più lenta si osserva nei geni ribosomali *rrnS* e *rrnL*. Infine, la *control region* (CR) appare particolarmente variabile e non è facilmente allineabile nel genere *Pagurus*.

	<i>atp6</i>	<i>atp8</i>	<i>cob</i>	<i>cox1</i>	<i>cox2</i>	<i>cox3</i>	<i>nad1</i>	<i>nad2</i>	<i>nad3</i>	<i>nad4</i>	<i>nad4L</i>	<i>nad5</i>	<i>nad6</i>	
<i>A. distinguendus</i>	0,322	0,502	0,270	0,260	0,285	0,248	0,274	0,405	0,367	0,354	0,282	0,316	0,448	<i>A. glaber</i>
<i>H. americanus</i>	0,076	0,077	0,065	0,071	0,084	0,063	0,065	0,088	0,054	0,079	0,050	0,070	0,090	<i>H. gammarus</i>
<i>M. crispata</i>	0,072	0,027	0,073	0,064	0,062	0,052	0,058	0,088	0,033	0,070	0,062	0,062	0,068	<i>M. squinado</i>
<i>P. cuanensis</i>	0,235	0,212	0,263	0,192	0,193	0,169	0,159	0,286	0,281	0,234	0,203	0,219	0,279	<i>P. longicarpus</i>
<i>P. longicarpus</i>	0,221	0,211	0,275	0,189	0,174	0,184	0,170	0,305	0,252	0,238	0,153	0,220	0,323	<i>P. prideaux</i>
<i>P. prideaux</i>	0,146	0,098	0,173	0,139	0,153	0,136	0,130	0,198	0,186	0,184	0,156	0,180	0,209	<i>P. cuanensis</i>

	concatenati	genomico	CR	<i>rrnL</i>	<i>rrnS</i>	
<i>A. distinguendus</i>	0,289	0,295	0,535	0,258	0,296	<i>A. glaber</i>
<i>H. americanus</i>	0,072	0,065	0,111	0,018	0,022	<i>H. gammarus</i>
<i>M. crispata</i>	0,065	0,068	0,210	0,028	0,025	<i>M. squinado</i>
<i>P. cuanensis</i>	0,208	0,212	>>1	0,139	0,137	<i>P. longicarpus</i>
<i>P. longicarpus</i>	0,210	0,215	>>1	0,139	0,139	<i>P. prideaux</i>
<i>P. prideaux</i>	0,159	0,162	0,411	0,080	0,087	<i>P. cuanensis</i>

Tabella 8: distanze calcolate con il modello Kimura 2 parametri, nel confronto a coppie tra congenerici

4.2. Determinazione del segnale filogenetico

I risultati delle distanze *pair-wise* calcolati per i differenti PCG, sono presentati nella Figura 12a,b e Figura 13a. L'analisi della Figura 12a, ci mostra come le terze posizioni (p3) di ogni PCG siano le più variabili, una affermazione sostenuta dal fatto che la *p-distance* media è sempre ≥ 0.50 . Sempre analizzando le singole posizioni dei codoni, si osserva che le seconde posizioni (p2) sono le più conservate in tutti i PCG (media ≤ 0.16) mentre il livello di variabilità delle prime posizioni (p1) (media ≤ 0.34) è intermedio tra quanto osservato per le seconde e le terze. Quando il confronto è esteso ai codoni interi (p123), oppure alle prime e seconde posizioni (p12), si ha che, nel primo caso le *p-distance* sono inferiori solo a quelle calcolate per le p3, mentre nel secondo caso, il livello di conservazione è molto più marcato con valori di *p-distance* superiori solo alle p2. Infine le sequenze amminoacidiche (aa) mostrano un livello di divergenza maggiore dei dataset riferiti a p12 e p2 mentre risultano simili alle p1 e meno divergenti dei codoni o delle p3. Quindi, ricapitolando, l'ordine di conservazione è p2, p12, p1 e aa, p123, p3.

Si osservi anche che i *pattern* di variazione non sono costanti per i diversi geni, con geni che mostrano un maggior livello di divergenza (*atp8*, *nad2-nad6*) e geni che sembrano essersi evoluti ad un ritmo di sostituzione più lento (*cob*, *cox1-cox3*).

Quando si prendono in considerazione le differenze (DIF) ottenute sottraendo alle distanze calcolate con un modello di evoluzione molecolare complesso GTR+G, che tiene conto delle sostituzioni multiple che possono essersi verificate nel corso dell'evoluzione, le distanze calcolate con un modello molto semplice (*p-distance*), in cui viene considerata solamente la variabilità osservata nell'allineamento multiplo (Figura 12b), emergono *pattern* di sostituzione molto differenti per le varie posizioni dei codoni, per le loro combinazioni e per gli amminoacidi.

In particolare le DIF $\gg 1$ osservate per le p3, certificano che le *p-distance* sottostimano enormemente le sostituzioni che sono avvenute nel corso del processo evolutivo. Inoltre gli elevati valori osservati mostrano che alla scala tassonomica a cui è stata condotta l'analisi (cioè a livello della classe Malacostraca) le p3 non hanno più segnale filogenetico. Anche i dataset costituiti da interi codoni mostrano DIF che in molti casi sono $\geq 0,80$. In questo caso la perdita di segnale filogenetico è elevata e dovuta in gran parte alle p3. Infatti, quando si considerano i dataset formati dalle p12 si vede che i valori di DIF diminuiscono drasticamente, fenomeno ancora più accentuato quando si considerano esclusivamente le p2. Gli amminoacidi (aa)

mostrano un comportamento simile alle p12 mentre le p1 presentano differenze maggiori con conseguente maggiore sottostima delle sostituzioni da parte delle *p-distanze*. Pertanto i valori di DIF mostrano che i dataset in cui il processo di sostituzione è stato meno marcato sono p2, p12 e aa.

Estendendo ulteriormente la comparazione al rapporto (Ratio) tra *p-distance* e la GTR+G distance (Figura 13a), si osserva come questo sia $\geq 0,80$ per gran parte delle p2 dei diversi geni, sia compreso tra 0,50 e 0,75 per p12 e aa di tutti i geni e scende al di sotto di 0,50 per le restanti combinazioni con alcune eccezioni per le p1 (*cox1-cox3*), raggiungendo i valori minimi per le p3. Questo dato corrobora il fatto che p12, p2, e aa presentano il maggior segnale filogenetico. Se da ultimo si considera anche la percentuale di quartetti completamente risolti, si osserva che *atp6*, *cob*, *cox1-cox3*, *nad1*, *nad4* e *nad5* presentano una percentuale $\geq 75\%$ per p2, p12, e aa mentre gli altri geni si collocano sotto questa soglia. Pertanto combinando tutti questi dati a partire dagli allineamenti concatenati 13.PCG e 13.PCG.aa sono stati prodotti dei sub set includenti le posizioni/geni a più alto segnale filogenetico per effettuare una analisi globale dei taxa studiati e verificare l'effetto dell'inclusione/esclusione dei diversi geni sugli alberi. I risultati di quest'analisi sono presentati nel paragrafo successivo.

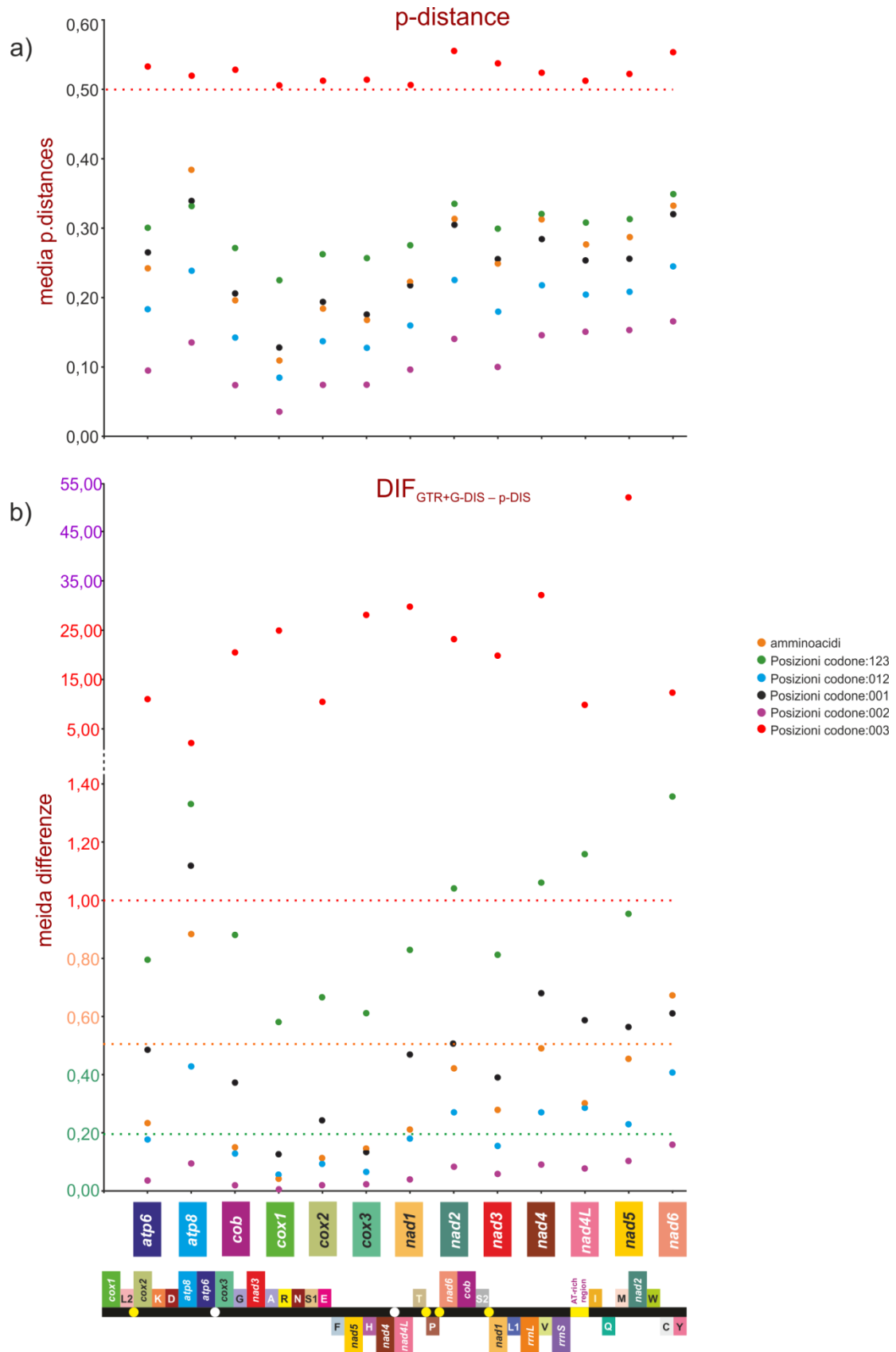


Figura 12: a) differenze GTR+G distance – p-distance; b) valori medi di p-distance. Il codice colori per i diversi geni è identico a quello presentato in Figura 1.

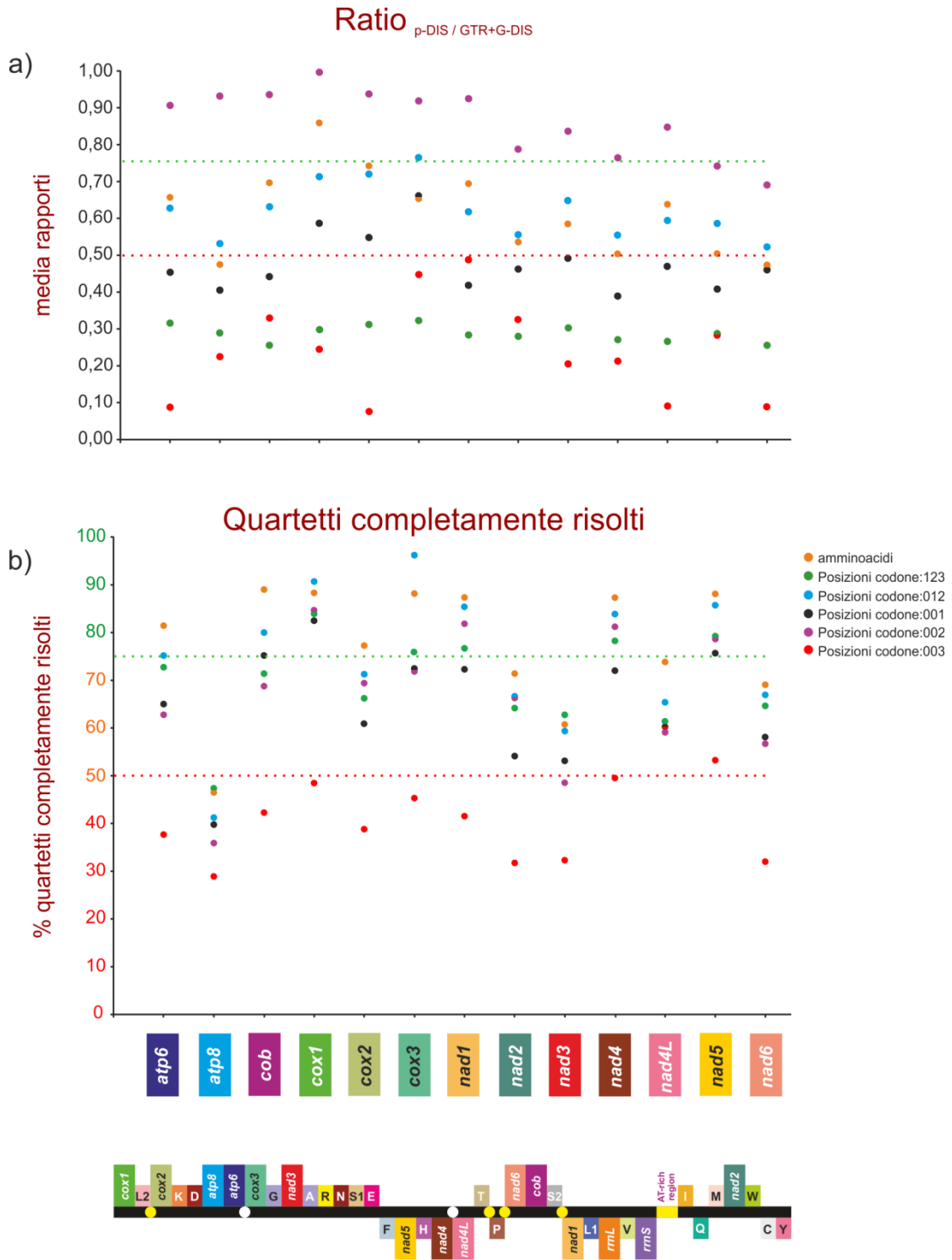


Figura 13: a) rapporto $p\text{-distance}/\text{GTR} + \text{G distance}$ b) percentuale di quartetti risolti. Il codice colori per i diversi geni è identico a quello presentato in Figura 1.

4.3. Analisi filogenetiche

Gli alberi ottenuti da 13.PCG.p123 (9636 posizioni), 13.PCG.p12 (6424 posizioni), 13.PCG.p2 (3212 posizioni), 13.PCG.aa (3212 posizioni), 8.PCG.p12 (5464 posizioni), 8.PCG.p2 (2731 posizioni), 8.PCG.aa (2731 posizioni), sono riportati nelle figure 14-20. Le diverse topologie sono congruenti in varie parti, ma presentano anche delle marcate differenze. Gran parte dei gruppi a rango tassonomico medio-alto (livello di famiglia e ranghi superiori), presenti nella sistematica tradizionale dei Decapoda, risultano monofiletici indipendentemente dal numero (8,13) e dal tipo (aa, p2, p12, p123) di marcatori utilizzati. Tale relazioni di monofilia sono supportate in gran parte delle analisi da valori di *bootstrap* > e >> 50%. Una prima eccezione di rilievo è rappresentata dall'infraordine Thalassinidea che risulta monofiletico solamente nell'analisi condotta su 13.PCG.p2 mentre in tutti gli altri casi è para/polifiletico, con la definizione di due cladi separati formati, uno dai rappresentanti delle famiglie Strahlaxiidae (*Neaxius glyptocercus*) e Callianassidae (*Corallianassa coutierei* + *Nihonotrypaea japonica*), e uno formato da Upogebiidae (*Austinogebia edulis*, specie del genere *Upogebia*) + Thalassinidae (*Thalassinia kelanang*). Questo comporta anche la polifilia della superfamiglia Callianassoidea (Callianassidae + Upogebiidae).

La seconda eccezione è data dall'infraordine Astacoidea che risulta monofiletico nelle analisi condotte su 13.PCG.aa, 13.PCG.p12, 13.PCG.p2 e 8.PCG.p2, mentre appare parafiletico negli altri alberi con i rappresentanti della superfamiglia Astacoidea (*Cambaroides* e *Procambarus*) e *sister group* di *Polycheles typhlops*, unico rappresentante dell'infraordine Polychelida incluso nell'analisi.

La terza eccezione è data dalla superfamiglia/famiglia Sergestoidea/Sergestidae (*Sergia robusta* + *Acetes chinensis*) che è monofiletica nelle topologie ottenute da 13.PCG.p123, 13.PCG.p12, 8.PCG.p12 e 8.PCG.p2 mentre risulta para/polifiletica negli altri alberi.

Particolarmente complessa risulta la serie di relazioni filogenetiche tra le varie linee filetiche del sottordine Pleocyemata. Come caratteristica generale si osserva che i nodi basali che descrivono gli eventi cladogenetici tra i principali gruppi di Pleocyemata non presentano alcun supporto di *bootstrap*. Inoltre i rapporti di parentela tra i cladi non risultano stabili nei differenti alberi. Come esempio rappresentativo di questo fenomeno citiamo qui Anomura e Brachyura che sono *sister taxa* negli alberi ottenuti da 13.PCG.p123, 13.PCG.p12, 13.PCG.p2 e

8.PCG.p12 mentre assumono posizioni differenti negli altri dataset. Esiste però un punto fermo anche all'interno dei Pleocyemata. Infatti i Caridea risultano sempre *sister group* di tutti gli altri *taxa* inclusi in questo sottordine che, con eccezione di *Stenopus hispidus* (Stenopodidea), formano un clade con supporto massimo di *bootstrap* (100%) in tutte le analisi effettuate.

Nei Dendrobranchiata oltre alle discrepanze citate sopra per la famiglia Sergestidae, risulta variabile anche la posizione di *Solenocera membranacea*, che appare alternativamente *sister species* delle famiglie Aristeidae (13.PCG.p123) e Penaeidae (tutti gli altri dataset) nelle differenti analisi. Le relazioni filogenetiche tra le specie contenute nel clade Achelata sono invariate in tutti gli alberi e gran parte dei nodi ricevono supporto di *bootstrap*. All'interno dei Caridea non è ben risolta la posizione di *Plesionika gigliolii*, anche se le differenti posizioni che assume questa specie all'interno del clade, non ricevono un robusto supporto di *bootstrap* ($\leq 81\%$). Nei Brachyura i rapporti filogenetici sono in gran parte stabili con eccezione di *Eriphia verrucosa* che è quasi sempre *sister taxon* di *Goneplax rhomboides*, con supporto di *bootstrap* variabile, ad eccezione dell'albero ottenuto da 8.PCG.p2. All'interno degli Anomura è controversa la situazione del genere *Pagurus*. Infatti, questo *taxon* risulta monofiletico solo nelle analisi condotte sulle seconde posizioni (13.PCG.p2, 8.PCG.p2). In tutti gli altri casi *Pagurus* appare parafiletico.

All'interno della famiglia Upogebiidae il genere *Upogebia* risulta sempre parafiletico. Infine la posizione di *Rissoides desmaresti* è variabile all'interno degli Stomatopoda. Infatti, questo crostaceo è *sister taxon* di tutti gli altri *taxa* della famiglia Squillidae negli alberi ottenuti dai dataset con 13 marcatori, mentre risulta *sister species* del genere *Squilla* negli alberi prodotti dagli allineamenti multipli con 8 marcatori.

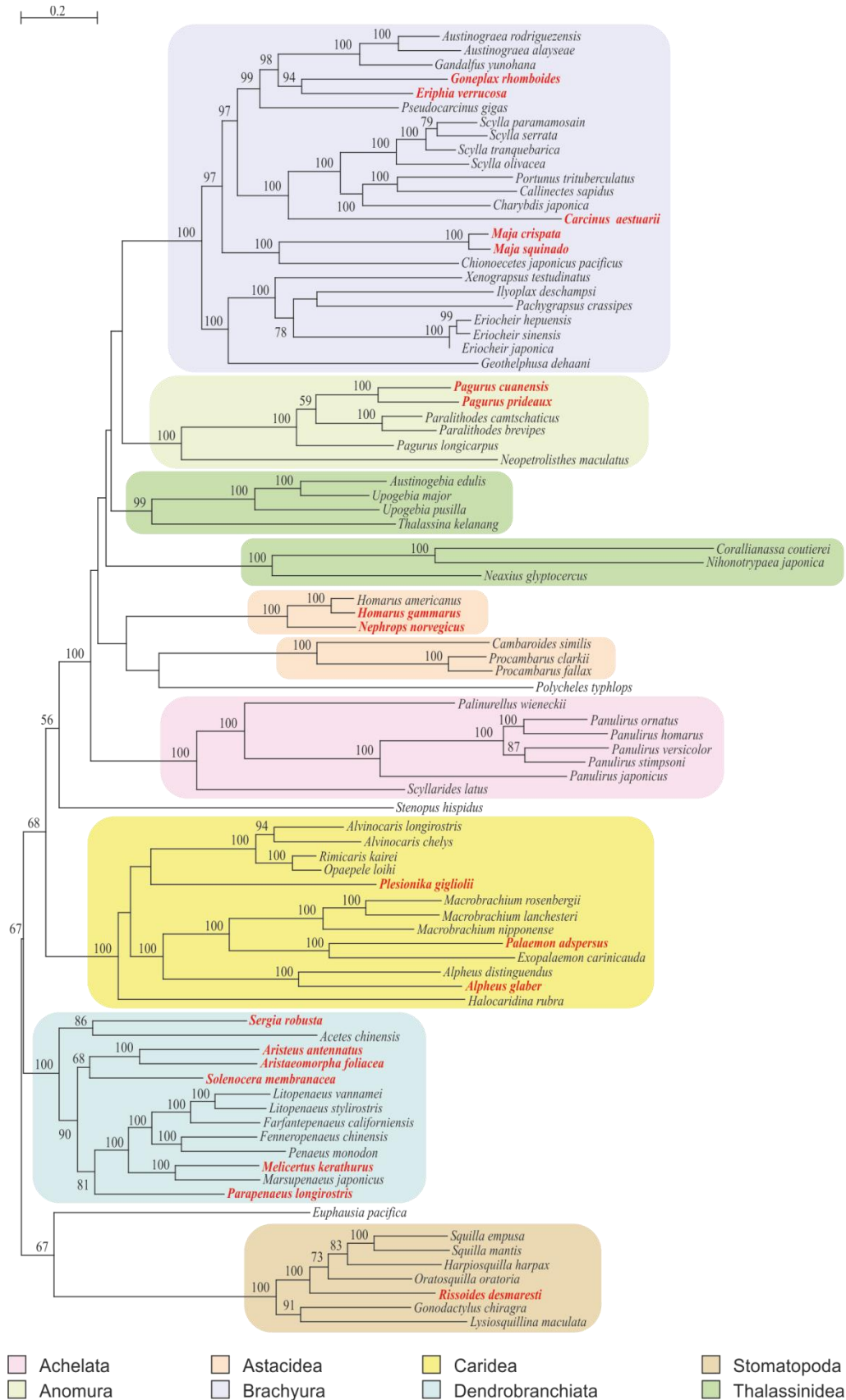


Figura 14: albero di Maximum Likelihood (-ln = 398968,532085) ottenuto dal data set 13.PCG.p123 utilizzando il programma RaxML. Vicino ai nodi sono riportati i valori di bootstrap, espressi in percentuale, che sono $\geq 50\%$. La barra indica 0.2 sostituzioni/posizione.

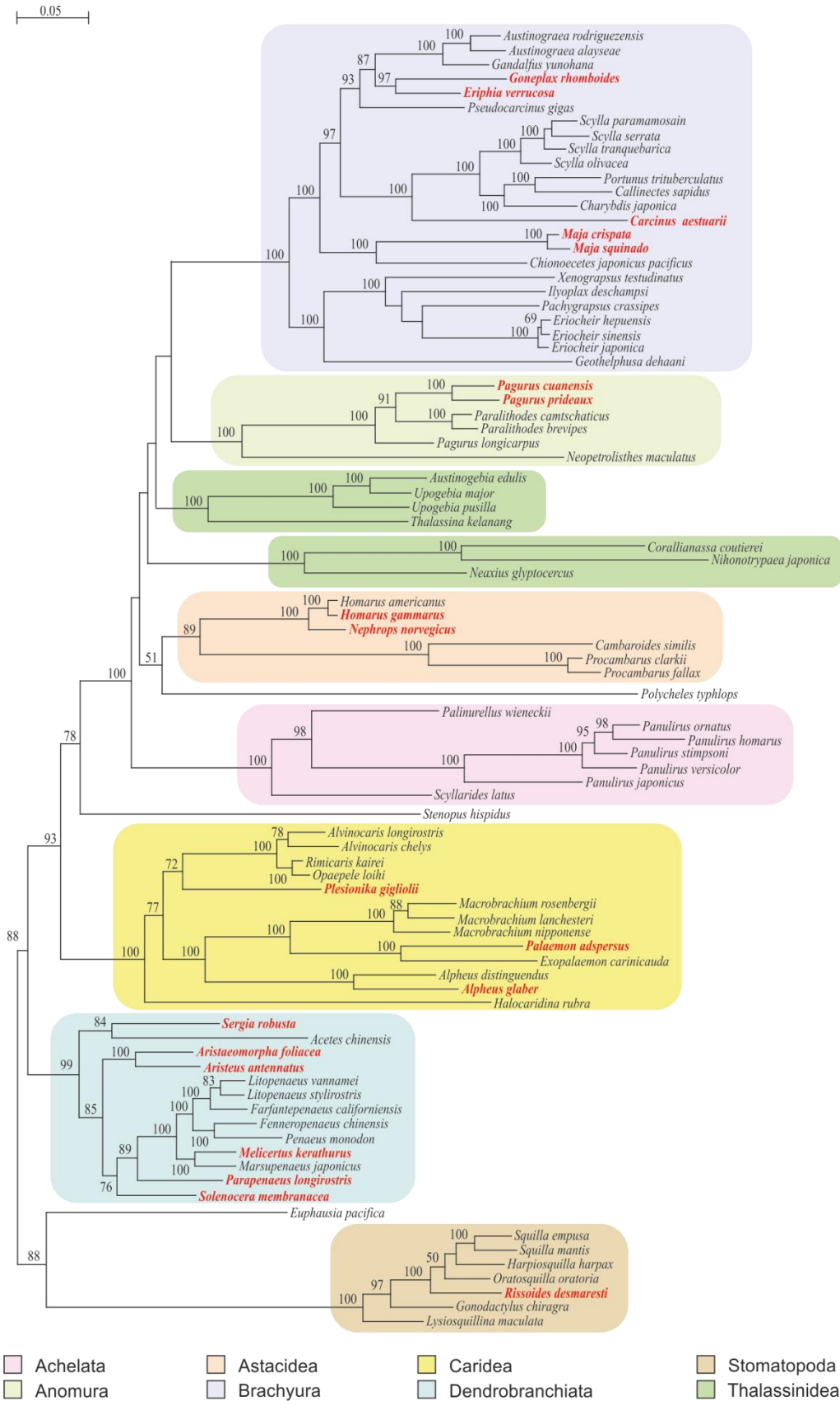


Figura 15: albero di Maximum Likelihood (-ln = 136967,678727) ottenuto dal data set 13.PCG.p12 utilizzando il programma RaxML. Vicino ai nodi sono riportati i valori di bootstrap, espressi in percentuale, che sono $\geq 50\%$. La barra indica 0.05 sostituzioni/posizione.

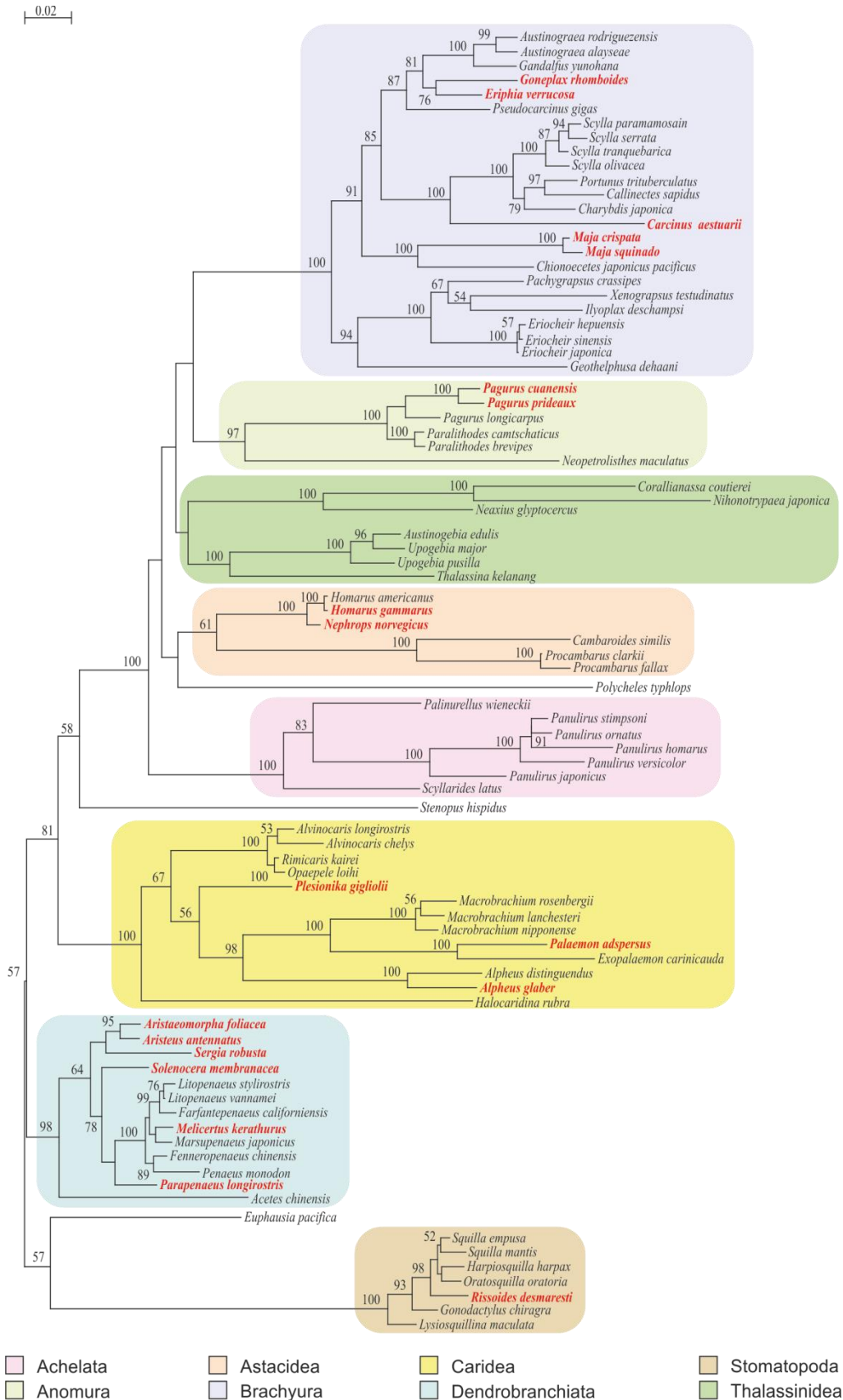


Figura 16: albero di Maximum Likelihood (-ln = 41163,406494) ottenuto dal data set 13.PCG.p2 utilizzando il programma RaxML. Vicino ai nodi sono riportati i valori di bootstrap, espressi in percentuale, che sono $\geq 50\%$. La barra indica 0.02 sostituzioni/posizione.

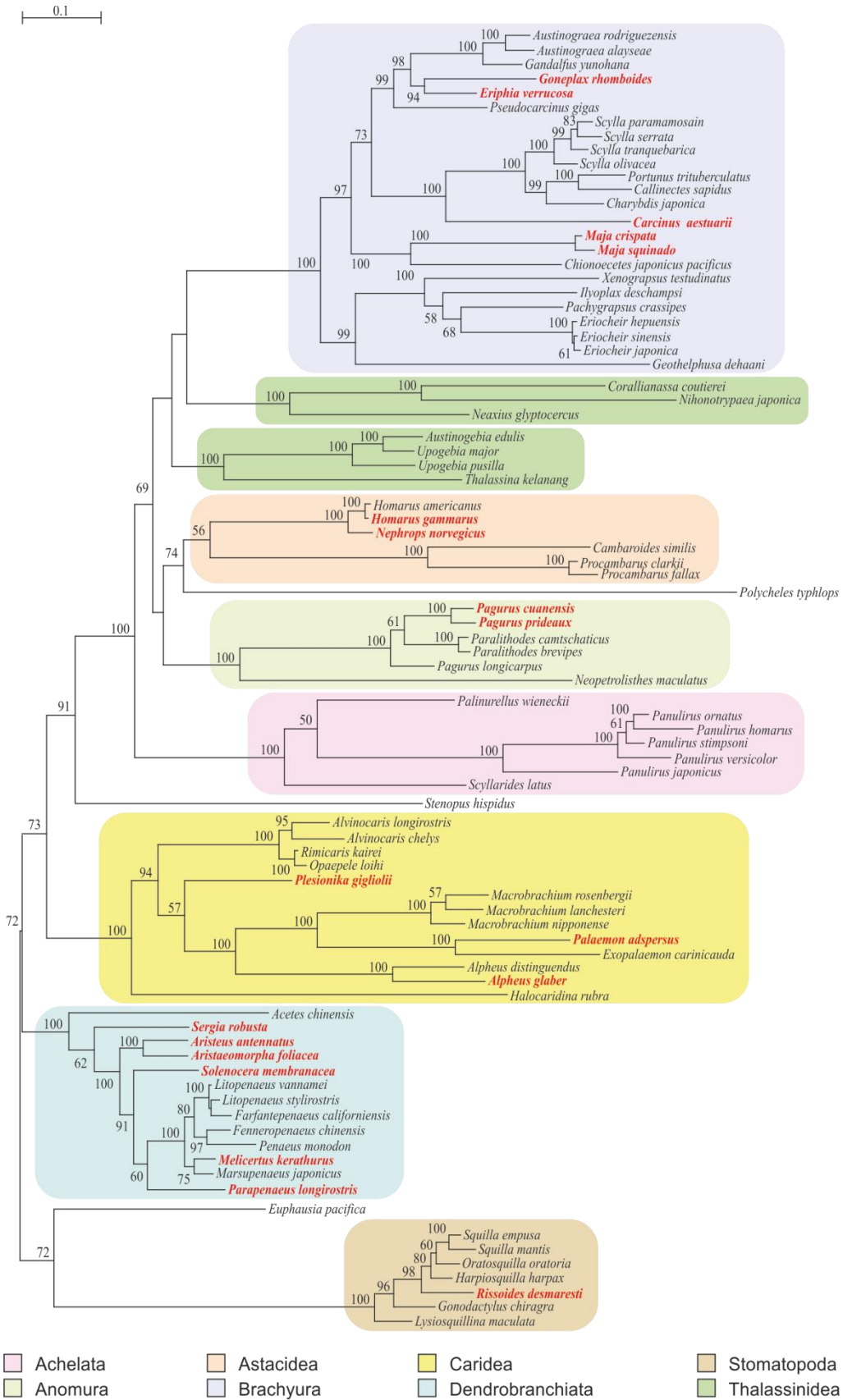


Figura 17: albero di Maximum Likelihood ($-ln = 110010,321344$) ottenuto dal data set 13.PCG.aa utilizzando il programma RaxML. Vicino ai nodi sono riportati i valori di bootstrap, espressi in percentuale, che sono $\geq 50\%$. La barra indica 0.1 sostituzioni/posizione.

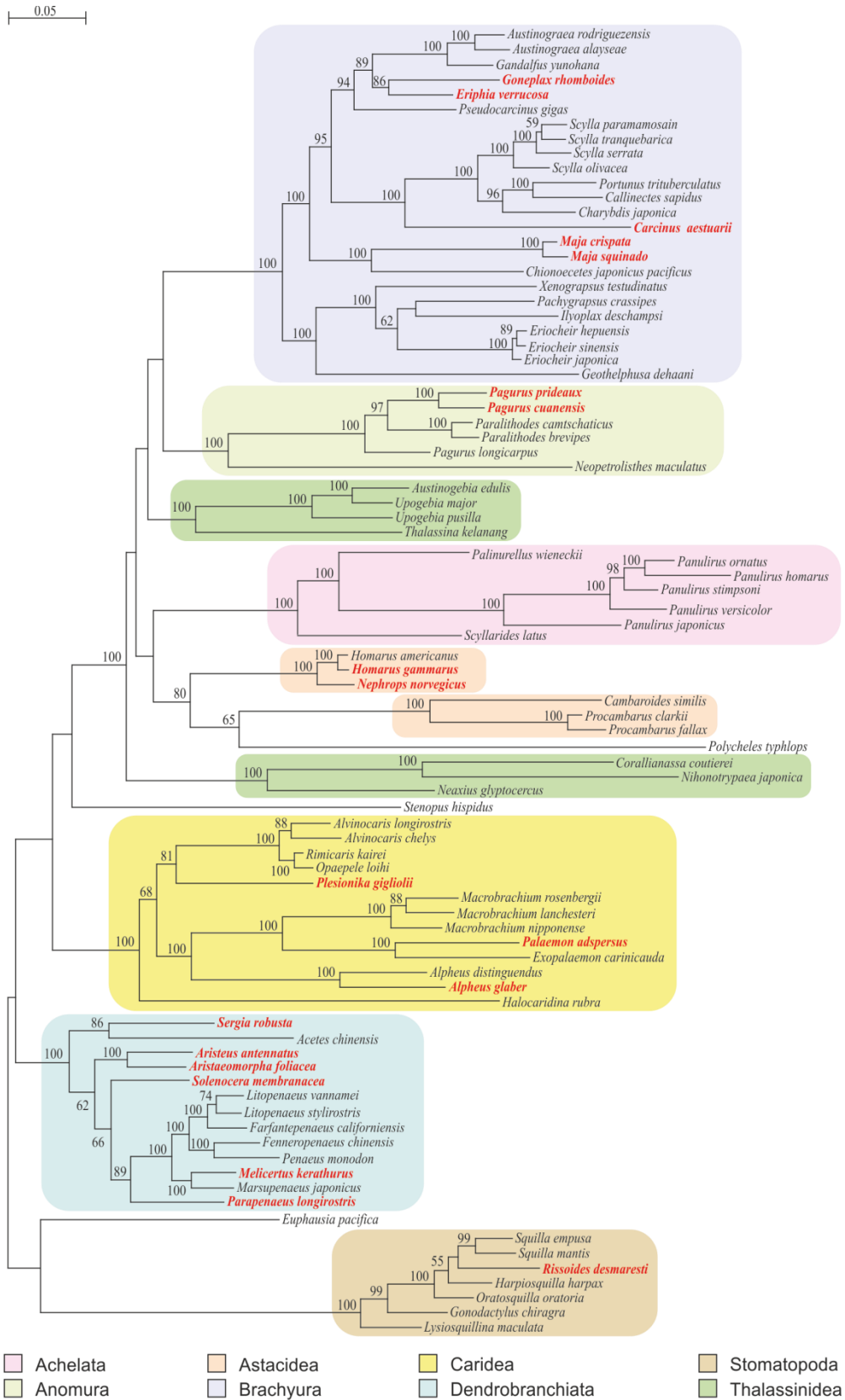


Figura 18: albero di *Maximum Likelihood* ($-ln = 110721,184607$) ottenuto dal data set 8.PCG.p12 utilizzando il programma RaxML. Vicino ai nodi sono riportati i valori di *bootstrap*, espressi in percentuale, che sono $\geq 50\%$. La barra indica 0.05 sostituzioni/posizione.

0.02

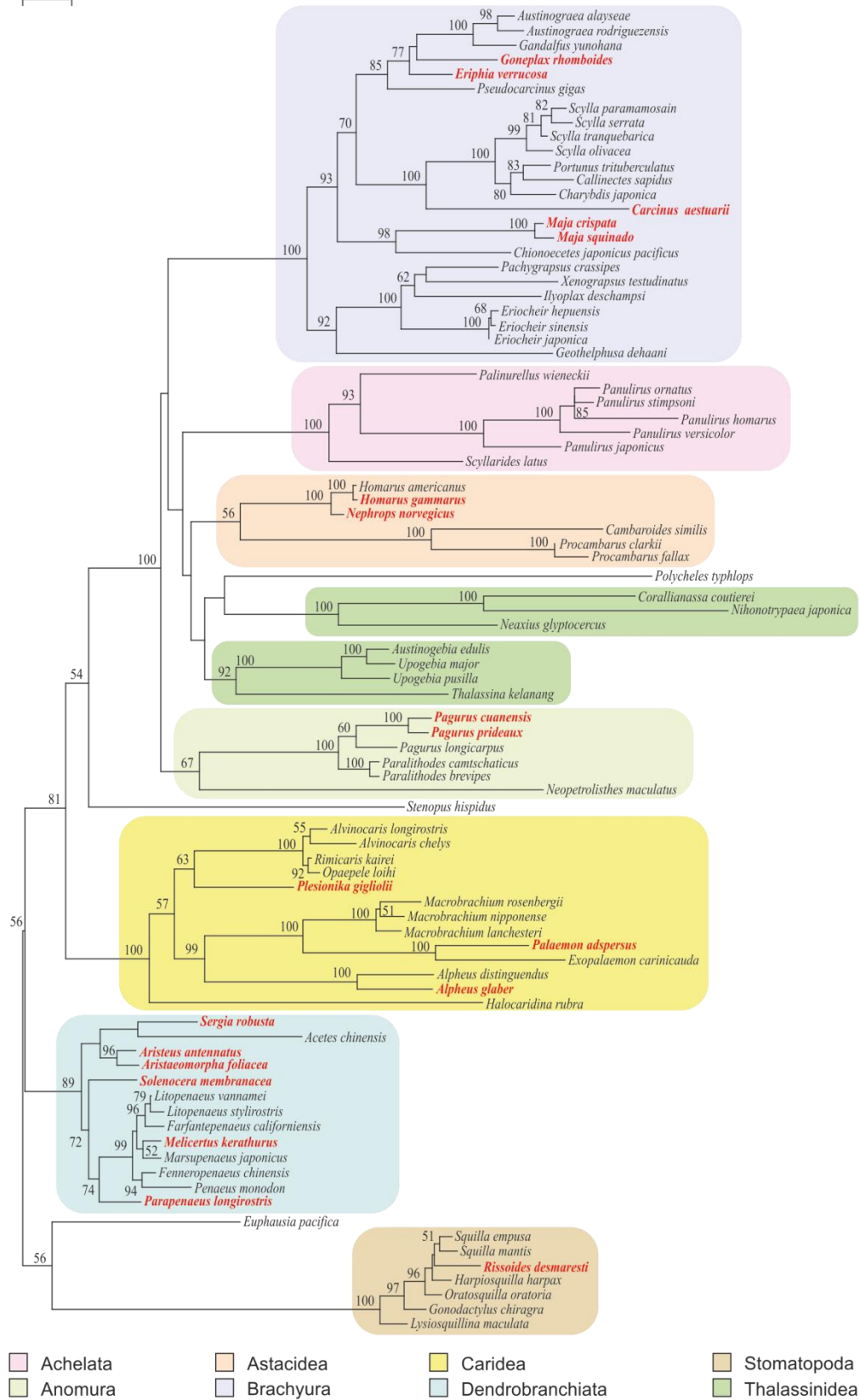


Figura 19: albero di Maximum Likelihood (-ln = 33403,663872) ottenuto dal data set 8.PCG.p2 utilizzando il programma RaxML. Vicino ai nodi sono riportati i valori di bootstrap, espressi in percentuale, che sono $\geq 50\%$. La barra indica 0.02 sostituzioni/posizione

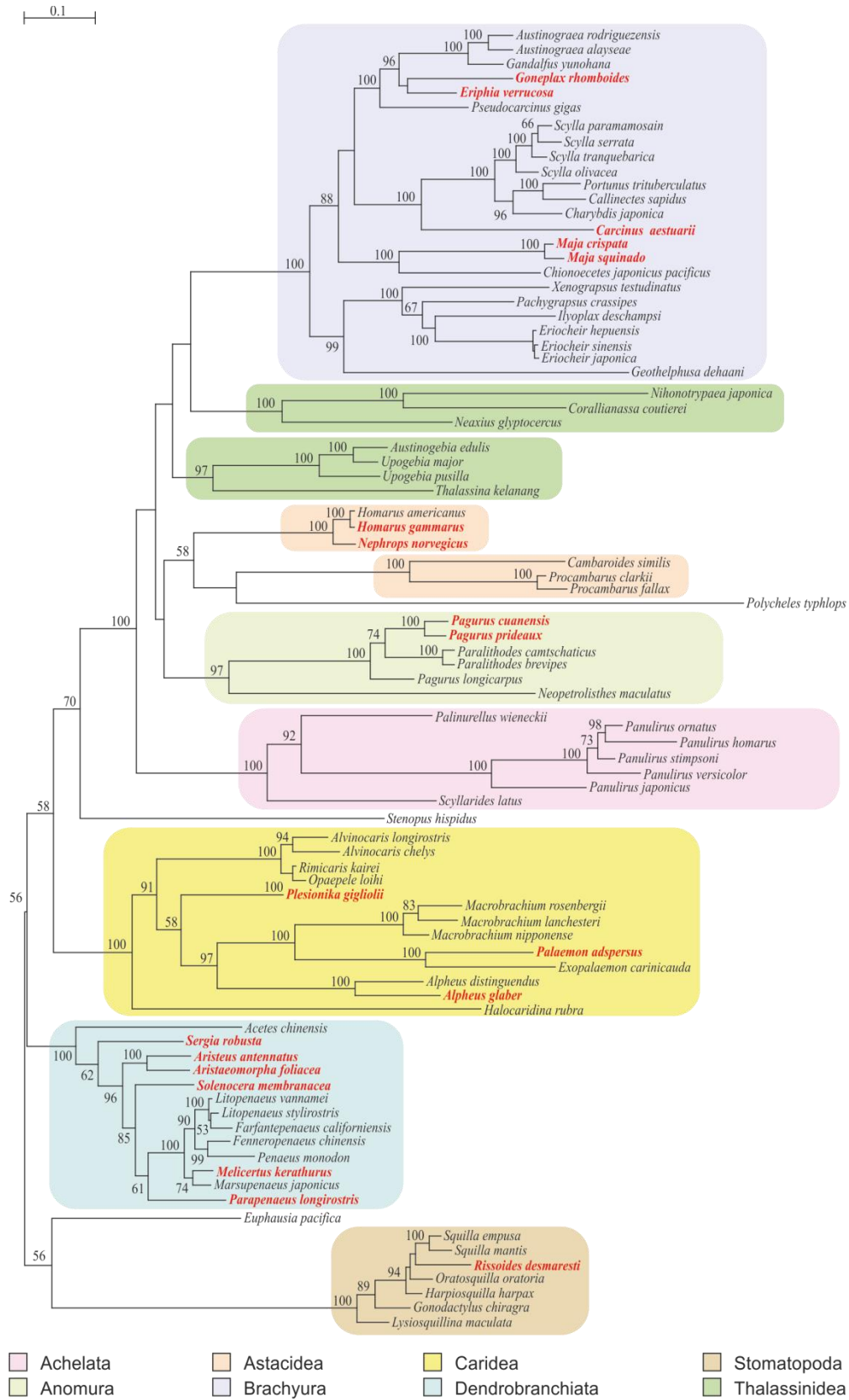


Figura 20: albero di *Maximum Likelihood* ($-\ln = 88793,423557$) ottenuto dal data set 8.PCG.aa utilizzando il programma RaxML. Vicino ai nodi sono riportati i valori di *bootstrap*, espressi in percentuale, che sono $\geq 50\%$. La barra indica 0.1 sostituzioni/posizione.

5. Discussione

5.1. Gene Order

I risultati relativi al GO rivelano come l'ordine Decapoda sia uno degli hot-spot, all'interno dei phylum degli Arthropoda, per quanto concerne la diversità di GO. Altri taxa di Arthropoda particolarmente variabili sono diverse linee filetiche di crostacei Malacostraca e non. Citiamo qui tra i malacostraci gli Amphipoda (Krebes & Bastrop 2012) e gli Isopoda (Kilpert & Podsiadlowski 2006) e tra gli altri crostacei i Copepoda (es. Ki *et al.* 2009). Anche diversi gruppi di Chelicerata presentano un notevole varietà di GO (Masta *et al.* 2008; Masta *et al.* 2010; Ovchinnikov & Masta 2012). Diversi GO si trovano anche nei Myriapoda (es. Lavrov *et al.* 2002; Negrisolo *et al.* 2004; Gai *et al.* 2008), e tra gli insetti citiamo almeno qui gli ordini Phthiraptera (Cameron *et al.* 2011) e Hymenoptera (Dowton & Austin 1999; Dowton *et al.* 2009). Si osservi però che non tutte le linee filetiche all'interno dei Decapoda si mostrano equamente variabili. Infatti in tutti i Dendrobranchiata sequenziati fino ad ora il GO rimane quello standard per i Pancrustacea, come ben evidenziato dai 6 nuovi genomi presentati in questa tesi (*A. foliacea*, *A. antennatus*, *M. kerathurus*, *P. longirostris*, *S. membranacea* e *S. robusta*) e dai dati di letteratura (es. Wilson *et al.* 2000). Analogamente tutti i rappresentanti degli Achelata presentano il GO tipico dei Pancrustacea (es. Liu & Cui 2011) anche se gran parte delle specie finora sequenziate appartengono al genere *Panulirus* e quindi questa tendenza, potrebbe trattarsi di un effetto dovuto al limitato *taxon sampling*. Negli Astacidea la situazione è più variegata, con i rappresentanti della famiglia Cambaridae (*Cambaroides* e *Procambarus*) che presentano lo stesso GO differente da quello standard (Kim *et al.* 2012), mentre nei membri della famiglia Nephopidae troviamo sia il GO standard (*Homarus*) o un GO del tutto nuovo (*Nephrops*), con uno spostamento di *trnI* (vedi Figura 7d). In quest'ambito si osservi che, diversamente da quanto riportato di recente da Shen *et al.* (2013), *H. gammarus* non ha perduto nessun gene nel proprio mtDNA, come chiaramente dimostrato dalla sequenza presentata in questa tesi. Quest'ultimo aspetto è molto importante perché mostra come l'approccio Illumina possa consentire di individuare artefatti nel sequenziamento tradizionale dei genomi mitocondriali ed apre la strada alla verifica di genomi "completi" mancanti di uno o più geni, riportati in letteratura (es. Yang & Yang 2008).

All'interno dei Caridea gran parte dei taxa possiede il GO standard con l'unica eccezione data dal nuovo genoma di *P. adspersus* e quello già noto di *E. carinicauda*, che presentano un comune cambiamento nella disposizione di *trnT* e *trnP*. Le due linee filetiche di Decapoda che sono dei veri *hot-spot* di diversità sono gli Anomura ed i Brachyura. Nessuna specie di Anomura fino ad oggi sequenziata ha un GO standard. Infatti, in tutti i taxa troviamo GO molto modificati (questa tesi; Negrisolo *et al.* in preparazione; Hickerson & Cunningham 2000; Shen *et al.* 2013). Analogamente nei Brachyura, gruppo per il quale abbiamo 24 mtDNA disponibili, nessuna specie possiede un GO standard, ed è anzi stato possibile fissare un GO di riferimento (vedi Liu & Cui 2010 e referenze citate) per questo infraordine (Figura 9) con ulteriori particolari riarrangiamenti e nuovi GO che si osservano in altri granchi man mano che il *taxon sampling* viene esteso, come bene esemplificato dalle specie di *Maja* studiate in questa tesi o da altri dati di letteratura (es. Sun *et al.* 2005).

5.2. Evoluzione dei PCG e loro valore filogenetico

Le analisi condotte sui differenti PCG hanno mostrato come questi evolvano con tassi di sostituzione differenti. Di conseguenza l'mtDNA presenta una suite di marcatori che sono potenzialmente utilizzabili a diversa scala tassonomica. Quest'aspetto si evince anche dalle analisi filogenetiche condotte su dataset contenenti sia un numero variabile di PCG, sia tipi differenti di caratteri (p123, p12, P2 e aa). Tra i sette allineamenti multipli analizzati in questa tesi, quello che produce una topologia perfettamente congruente con quelle che sono ritenute le relazioni filogenetiche più probabili tra i principali gruppi di Decapoda, è quello 13.PCG.p2 (Figura 16) (Shen *et al.* 2013). Questa topologia perfettamente congruente, non risulta però supportata in tutti i nodi con valori di *bootstrap* adeguati. In quest'albero, infatti, tutti i principali taxa sono monofiletici (compresi i Thalassinidea). Questo risultato è in perfetto accordo con i dati relativi ai tassi di sostituzione che vedono nelle p2, le posizioni con più elevato segnale filogenetico a questo livello tassonomico. Si osservi inoltre, che anche il genere *Pagurus* risulta monofiletico. L'unico punto critico è la posizione di *Sergia robusta* che risulta *sister taxon* della famiglia Aristeidae invece che *sister species* dell'altro sergestide *A. chinensis*. Tuttavia questa relazione non riceve supporto di *bootstrap* e si può considerare come una parte dell'albero non particolarmente ben risolta. Una topologia molto simile si ottiene anche da 13.PCG.p12 (Figura 15) dove i Thalassinidea non sono più monofiletici, eventualità questa, ammessa come possibile dalle ultime analisi filogenetiche sui Decapoda

(Shen *et al.* 2013). In questo caso la famiglia Sergestidae risulta monofiletica con un buon supporto di *bootstrap*, mentre *Pagurus* non è più monofiletico. 13.PCG.p123 produce un albero in cui oltre ai Thalassinidea, sono parafiletici anche gli Astacidea e ancora maggiori sono le discrepanze per 13.PCG.aa in cui i rapporti tra cladi principali differiscono sensibilmente. Infatti, in quest'ultimo caso, gli Anomura sono *sister group* di Astacidea + *Polycheles typhlops*. Infine le topologie ottenute dai dataset 8.PCG.p12, 8.PCG.p2 e 8.PCG.aa, formati solo da marcatori che avevano una quantità di quartetti completamente risolti $\geq 75\%$, mostrano topologie con numerose discrepanze rispetto a quanto postulato essere le relazioni filogenetiche nei Decapoda (Shen *et al.* 2013).

Quindi p2 o p12 dei 13 PCG appaiono i migliori marcatori ad alto rango tassonomico in perfetto accordo con quanto ricavato dalle analisi sui segnali filogenetici. I dataset formati da amminoacidi non appaiono particolarmente efficaci. Questo aspetto va ulteriormente approfondito, cercando di valutare diversi fattori (asimmetrie composizionali, presenza di *rogue taxa*, permanenza di polimorfismi ancestrali, ecc.) che per ragioni di tempo e perché esulavano dalla mia tesi, non sono stati opportunamente considerati.

5.3. Individuazione di *tools* mitocondriali per vari tipi di applicazioni

Come evidenziato dalle analisi condotte su BOLD, per gran parte delle specie, è possibile ottenere una identificazione certa semplicemente basandosi sul *DNA barcoding*. Tuttavia i dati ottenuti consentono di poter sviluppare una serie di strumenti molecolari che possono potenzialmente essere utilizzati proficuamente sul piano applicativo. L'esistenza di peculiari GO infatti permette di sfruttare questi ultimi per disegnare dei *primer* sui geni riarrangiati. Tali *primer* possono essere utilizzati, una volta messo a punto il protocollo di PCR e validata la procedura sincerandosi che siano specifici, per identificare le specie che possiedono un particolare GO, basandosi direttamente su i risultati di PCR e senza dover ricorrere al sequenziamento, che è sicuramente economico, ma richiede tempi ancora troppo lunghi, un aspetto questo di non secondaria importanza nel caso di prodotti ittici. Ovviamente per questo tipo di approccio è importante avere un'idea chiara su quanto un GO sia diffuso e se vi possano essere eventuali casi di convergenza (ad oggi non noti) in taxa filogeneticamente distanti. Quest'ultima evenienza peraltro potrebbe

essere superata disegnando in modo idoneo i *primer*, facendo sì che non funzionino in entrambi i taxa. Anche in taxa con un GO standard la conoscenza dell'intero genoma mitocondriale offre la possibilità di sviluppo di ulteriori marcatori ad alto valore diagnostico e rapida ed economica applicazione. La conoscenza delle *AT-rich region* spesso difficilmente allineabili anche tra specie congeneriche, consente di disegnare *primer* altamente mirati capaci di indentificare una determinata specie attraverso la presenza/assenza del prodotto di PCR, come già si fa per i kit diagnostici utilizzati per identificare le varie isoforme umane di HLA (*Human Leukocyte Antigen*).

La presenza di numerosi PCG che mostrano un tasso di sostituzione più elevato del *cox1*, fa sì che questi siano potenzialmente utili per studi di filogeografia/genetica di popolazioni e quindi nell'identificare l'origine geografica all'interno dell'areale di distribuzione della singola specie, con possibili conseguenti ricadute sulla gestione di stock di pesca.

In questa tesi non sono stati considerati gli spaziatori intergenici che in alcune specie di crostacei possono essere lunghi anche parecchie centinaia di basi. L'unico studio che finora li ha considerati in un contesto di *taxon sampling* denso, è riferito ad un genere di farfalle (*Thaumetopoea*) e ha dimostrato che questi possono essere particolarmente utili per identificare e caratterizzare le varie specie (Simonato *et al.* 2013). Quindi, anche questa categoria di marcatori è potenzialmente fonte di strumenti per lavorare proficuamente a basso rango tassonomico, ambito dove spesso può essere richiesto l'intervento del biotecnologo per l'alimentazione in relazione a problemi di frodi alimentari.

Quindi la caratterizzazione molecolare di specie di Crostacei di interesse alimentare e non basata sul sequenziamento parziale/completo di genomi mitocondriali, mi ha consentito non solo di ampliare le conoscenze di base su questi taxa, ma anche di ottenere tutta una serie di potenziali categorie di marcatori molecolari che una volta ottimizzati e validati possono avere rapide, efficaci e immediate applicazioni anche nel settore della sicurezza alimentare.

6. Bibliografia

- Abascal F., Zardoya R., Telford M. J., 2010. *TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations*. Nucleic Acids Res., 8: W7–13.
- Atti convegno. *La risorsa Crostacei nel Mediterraneo: ricerca, produzione e mercato*. CONARGA 2011.
- Babbucci M., Buccoli S., Cau A., Cannas R., Goñi R., Díaz D., Marcato S., Zane L., Patarnello T., 2010. *Population structure, demographic history, and selective processes: contrasting evidences from mitochondrial and nuclear markers in the European spiny lobster *Palinurus elephas* (Fabricius, 1787)*. Mol. Phylogenet. Evol., 56(3): 1040–1050.
- Bauza-Ribot M. M., Jaume D., Juan C., Pons J., 2009. *The complete mitochondrial genome of the subterranean crustacean *Metacrangonyx longipes* (Amphipoda): A unique gene order and extremely short control region*. Mitochondrial DNA, 20(4): 1–12.
- Berardi L., 2011 *Caratterizzazione molecolare di comunità di lepidotteri diurni presenti in ambienti alpini e montani (Parco naturale Paneveggio–Pale di San Martino)*. [Tesi di Laurea Magistrale in Biotecnologie per l’Alimentazione].
- Bernt M., Merkle D., Ramsch K., Fritzsche G., Perseke M., Bernhard D., Schlegel M., Stadler P. F., Middendorf M., 2007. *CREx: inferring genomic rearrangements based on common intervals*. Bioinformatics, 23(21): 2957–2958.
- Boni S., 2010. *Caratterizzazione della frazione lipidica dei prodotti ittici nell’ambito del progetto Last Minute Fish*. [Tesi di Dottorato di ricerca in Alimenti e Salute].
- Boore J. L., 1999. *Animal mitochondrial genomes*. Nucleic Acids Res., 27(8): 1767–1780.
- Boore J. L., 2000. *The duplication/random loss model for gene rearrangement exemplified by mitochondrial genomes of deuterostome animals*. In D.

SANKOFF and J. H. NADEAU, eds. Comparative genomics. Kluwer Academic Publishers, Dordrecht. 133–147.

- Boore J. L., Lavrov D. V., Brown W. M., 1998. *Gene translocation links insects and crustaceans*. *Nature*, 392: 667–668.
- Brusca G. J. & Brusca R. C., 1996. *Invertebrati*. Zanichelli Ed., Bologna, 505–531: 651–719.
- Buccoli S., 2006. *Analisi del differenziamento genetico tra popolazioni di *Palinurus elephas* (Fabricius, 1787) attraverso l'impiego di marcatori genetici mitocondriali e nucleari*. [Tesi di Dottorato di ricerca in Biologia Animale, Molecolare e dell'Uomo].
- Busato S., 2010. *Frodi alimentari per sostituzione di specie: identificazione di specie in prodotti ittici mediante tecnica pyrosequencing*. [Tesi di Laurea Specialistica in Biotecnologie per l'Alimentazione].
- Busonero. F., 2010. *Studi di associazione e sequenziamento estesi a tutto il genoma: il Progetto ProgeNIA*. [Tesi di Dottorato di ricerca in Terapia Pediatrica e Farmacologia dello Sviluppo].
- Cameron S. L., Yoshizawa K., Mizukoshi A., Whiting M. F., Johnson K. P., 2011. *Mitochondrial genome deletions and minicircles are common in lice (insecta: Phthiraptera)*. *BMC Genom.* 12: 394.
- Campbell N. A., Williamson B., Heyden R. J., 2006. *Biology: Exploring Life*. Boston, Massachusetts: Pearson Prentice Hall.
- Carucci F. 2011. *Filogenesi molecolare del genere *Cirsium* Mill. sect. *Eriolepis* (Cass.) Dumort.* [Tesi di Dottorato di ricerca in Biologia Avanzata].
- Cook C. E., 2005. *The complete mitochondrial genome of the stomatopod crustacean *Squilla mantis**. *BMC Genomics*, 6:105.
- Cox A. J., & Hebert P. D. N., 2001. *Colonization, extinction and phylogeographic patterning in a freshwater crustacean*. *Mol. Ecol.*, 10: 371–386.

- De Grave S., Pentcheff N. D., Ah Yong S. T., Chan T., Crandall K. A., Dworschak P. C., Felder D. L., Feldmann R. M., Fransen C. H. J. M., Goulding L. Y. D., Lemaitre R., Low M. E. Y., Martin J. W., Ng P. K. L., Schweitzer C. E., Tan S.H., Tshudy D., Wetzer R., 2009. *A classification of living and fossil genera of decapod crustaceans*. Raffles Bulletin of Zoology, Supplement No. 21: 1–109.
- Decreto ministeriale MiPAAF del 31 gennaio 2008 (G.U. n. 45 del 22 febbraio 2008) e successive integrazioni. *Elenco delle denominazioni in lingua italiana delle specie ittiche di interesse commerciale*.
- Downton M. & Austin A. D., 1999. *Evolutionary dynamics of a mitochondrial rearrangement “hot spot” in the Hymenoptera*. Mol. Biol. Evol., 16: 298–309.
- Downton M., Cameron S. L., Dowavic J. I., Austin A. D., Whiting M., 2009. *Characterization of 67 mitochondrial tRNA gene rearrangements in the hymenoptera suggests that mitochondrial tRNA gene position is selectively neutral*. Mol. Biol. Evol., 26: 1607–1617.
- Downton, M. & Austin, A. D., 1999. *Evolutionary dynamics of a mitochondrial rearrangement “hot spot” in the Hymenoptera*. Mol. Biol. Evol., 16: 298–309.
- Fearnley I. M. & Walker J. E., 1987. *Initiation codons in mammalian mitochondria: differences in genetic code in the organelle*. Biochemistry, 26: 8247–8251.
- Felsenstein J., 1985. *Confidence limits on phylogenies: an approach using bootstrap*. Evolution 39: 783–791.
- Felsenstein J., 2004. *Inferring Phylogenies*. University of Washington.
- Folmer O., Black M., Hoeh W., 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.
- Gai Y., Song D., Sun H., Yang Q., Zhou K., 2008. *The complete mitochondrial genome of Symphylella sp. (Myriapoda: Symphyla): extensive gene order*

rearrangement and evidence in favor of Progoneata. Mol. Phylogenet. Evol., 49: 574–585.

- Green D. R., 1998. *Apoptotic pathways: the roads to ruin*. Cell, 94(6): 695–8.
- Guindon S., Dufayard J. F., Lefort V., Anisimova M., Hordijk W., Gascuel O., 2010. *New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0*. Systematic Biology, 59(3): 307–21.
- Hajnóczky G., Csordás G., Das S., Garcia-Perez C., Saotome M., Sinha Roy S., Yi M., 2006. *Mitochondrial calcium signalling and cell death: approaches for assessing the role of mitochondrial Ca²⁺ uptake in apoptosis*. Cell Calcium, 40(5–6): 553–560. .
- Hebert P. D. N., Cywinska A., Ball S. L., deWaard J. R., 2003. *Biological identifications through DNA barcodes*. Proc. R. Soc. Lond. B, 270: 313–321.
- Hebert P. D. N., Penton E. H., Burns J. M., Janzen D. H., Hallwachs W., 2004. *Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator**. PNAS, 101: 14812–14817.
- Hebert P. D. N., Ratnasingham S., deWaard J. R., 2003b. *Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species*. Proc. R. Soc. Lond. B (Suppl.), 270: S96–S99.
- Hickerson M. J. & Cunningham C. W., 2000. *Dramatic mitochondrial gene rearrangements in the hermit crab *Pagurus longicarpus* (Crustacea, anomura)*. Mol. Biol. Evol., 17(4): 639-644.
- Holthuis L. B., 1991. *FAO SPECIES CATALOGUE: VOL, 13 MARINE LOBSTERS OF THE WORLD. An annotated and illustrated catalogue of species of interest to fisheries known to date*. FAO Fisheries Synopsis, 13(125): 292.
- Innocenti P., Morrow E. H., Dowling D. K., 2011. *Experimental Evidence Supports a Sex-Specific Selective Sieve in Mitochondrial Genome Evolution*. Science, 332(6031): 845–848.

- Ismea 2011. *Il mercato dei prodotti ittici in Italia, produzione, scambi e consumi*. EXPO RURALE, Firenze.
- Jondeung A., Karinthanyakit W., Kaewkhumsan J., 2012. *The complete mitochondrial genome of the black mud crab, Scylla serrata (Crustacea: Brachyura: Portunidae) and its phylogenetic position among (pan)crustaceans*. Mol. Biol. Rep., 39: 10921–10937.
- Katoh K., Kuma K., Toh H., Miyata T., 2005. *MAFFT version 5: improvement in accuracy of multiple sequence alignment*. Nucleic Acids Res., 33: 511–518.
- Ki J. S., Hop H., Kim S. J., Kim I. C., Park H. G., Lee J. S., 2010. *Complete mitochondrial genome sequence of the Arctic gammarid, Onisimus nanseni (Crustacea; Amphipoda): Novel gene structures and unusual control region features*. Comparative Biochemistry and Physiology, Part D 5: 105–115.
- Ki J. S., Park H. G., Lee J. S., 2009. *The complete mitochondrial genome of the cyclopoid copepod Paracyclopina nana: a highly divergent genome with novel gene order and atypical gene numbers*. Gene, 435: 13–22.
- Kilpert F. & Podsiadlowski L., 2006. *The complete mitochondrial genome of the common sea slater, Ligia oceanica (Crustacea, Isopoda) bears a novel gene order and unusual control region features*. BMC Genomics, 7: 241.
- Kim S., Park M. H., Jung J. H., Ahn D. H., Sultana T., Kim S., Park J. K., Choi H. G., Min G. S., 2012. *The mitochondrial genomes of Cambaroides similis and Procambarus clarkii (Decapoda: Astacidea: Cambaridae): the phylogenetic implications for Reptantia*. Zoologica Scripta, 41 281–292.
- Krebs L. & Bastrop R., 2012. *The mitogenome of Gammarus duebeni (Crustacea Amphipoda): a new gene order and non-neutral sequence evolution of tandem repeats in the control region*. Comparative Biochemistry and Physiology, Part D 7: 201–211.
- Lavrov D. V., Boore J. L., Brown, W. M., 2002. *Complete mtDNA sequences of two millipedes suggest a new model for mitochondrial gene rearrangements: duplication and nonrandom loss*. Mol. Biol. Evol., 19: 163–169.

- Li X., Fang P., Mai J., Choi E. T., Wang H., Yang X. F., 2013. *Targeting mitochondrial reactive oxygen species as novel therapy for inflammatory diseases and cancers*. J. Hematol. Oncol., 6(19): 19. .
- Liu Y. & Cui Z., 2011. *Complete mitochondrial genome of the Chinese spiny lobster Panulirus stimpsoni (Crustacea: Decapoda): genome characterization and phylogenetic considerations*. Mol. Biol. Rep., 38: 403–410.
- Liu Y. & Cui Z., 2010. *Complete mitochondrial genome of the Asian paddle crab Charybdis japonica (Crustacea: Decapoda: Portunidae): gene rearrangement of the marine brachyurans and phylogenetic considerations of the decapods*. Mol. Biol. Rep., 37(5): 2559–2569
- Liu Y., & Cui Z., 2010. *The complete mitochondrial genome of the mantid shrimp Oratosquilla oratoria (Crustacea: Malacostraca: Stomatopoda): Novel non-coding regions features and phylogenetic implications of the Stomatopoda*. Comparative Biochemistry and Physiology, Part D 5: 190–198.
- Liu Y., & Cui Z., 2011. *Complete mitochondrial genome of the Chinese spiny lobster Panulirus stimpsoni (Crustacea: Decapoda): genome characterization and phylogenetic considerations*. Mol. Biol. Rep., 38: 403–410.
- Lowe T. M. & Eddy S. R., 1997. *tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence*. Nucleic Acids Res, 25: 955–964.
- Lukhtanov V., Sourakov A., Zakharov E., Hebert P., 2009. *DNA barcoding Central Asian butterflies: increasing geographical dimension does not significantly reduce the success of species identification*. Mol. Ecol. Res., 9: 1302–1310.
- Marshall J, Cronin T.W., Kleinlogel S., 2007. *Stomatopod eye structure and function: A review*. Arthropod Structure & Development, 36: 420–448.
- Martin J. W., Crandall K. A., Felder D. L., 2009. *Decapod Crustacean Phylogenetics*. Taylor & Francis Group.

- Martin J.W., & Davis, G.E., 2001. *An Updated Classification of the Recent Crustacea*. Natural History Museum of Los Angeles County, Science Series 39.
- Martin W. F., & Müller M., 2007. *Origin of mitochondria and hydrogenosomes*. Springer Verlag, Heidelberg.
- Masta S. E., Klann A. E., Podsiadlowski L., 2008. *A comparison of the mitochondrial genomes from two families of Solifugae (Arthropoda: Chelicerata): Eremobatidae and Ammotrechidae*. *Gene*, 417: 35–42.
- Masta S. E., McCall A., Longhorn S. J., 2010. *Rare genomic changes and mitochondrial sequences provide independent support for congruent relationships among the sea spiders (Arthropoda, Pycnogonida)*. *Mol. Phylogenet. Evol.*, 57: 59–70.
- McBride H.M., Neuspiel M., Wasiak S., 2006. *Mitochondria: more than just a powerhouse*. *Curr. Biol.* 16(14): R551–560.
- Metzker M. L., 2010. *Sequencing technologies - the next generation*. *Nature Reviews Genetics*, 11: 31–46.
- Meyer M. & Kircher M., 2010. *Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing*. Cold Springs Harbor Protocols (6).
- Miller A.D. & Austin C.M. 2006, *The complete mitochondrial genome of the mantid shrimp Harpiosquilla harpax, and a phylogenetic investigation of the Decapoda using mitochondrial sequences*. *Mol. Phylogenet. Evol.*, 38: 565–574.
- Moretti V. M. & Busetto M., 2010. L., *Qualità e sicurezza delle carni di gambero*. CONVEGNO - La risorsa Crostacei nel Mediterraneo: ricerca, produzione e mercato.
- Morrison C. L., Harvey A. W, Lavery S., Tieu K., Huang Y., Cunningham C. W., 2002. *Mitochondrial gene rearrangements confirm the parallel evolution of the crab-like form*. *Proc. R. Soc. Lond. B*, 269: 345–350.

- Munkemuller T., Lavergne S., Bzeznik B., Dray S., Jombart T., Schiffrers K., Thuiller W., 2012. *How to measure and test phylogenetic signal*. *Methods in Ecology and Evolution*, 3: 743–756.
- Negrisol E. M., Maistro S., Incarbone M., Moro I., Dalla Valle L., Broady P. A., Andreoli C., 2004a. *Morphological convergence characterizes the evolution of Xanthophyceae (Heterokontophyta): evidence from nuclear SSU rDNA and plastidial rbcL genes*. *Mol. Phylogenet. Evol.*, 33(1): 156–170.
- Negrisol E. M., Babbucci M., Patarnello T., 2011. *The mitochondrial genome of the ascalaphidowfly Libelloides macaronius and comparative evolutionary mitochondriomics of neuropterid insects*. *BMC Genomics*, 12: 221.
- Negrisol E. M., Minelli A., Valle G., 2004b. *The mitochondrial genome of the house centipede Scutigera and myriapod monophyly vs. paraphyly*. *Mol. Biol. Evol.*, 21(4): 770–780.
- Nicolè S., Negrisol E. M., Eccher G., Mantovani R., Patarnello T., Erickson D. L., Kress W. J., Barcaccia G., 2012. *DNA Barcoding as a Reliable Method for the Authentication of Commercial Seafood Products*. *Food Technol. Biotechnol.*, 50(4): 387–398.
- Oh-hama T., 1997. *Evolutionary consideration on 5-aminolevulinic acid synthase in nature*. *Orig. Life Evol. Biosph.*, 27(4): 405–412.
- Ovchinnikov S. & Masta S. E., 2012. *Pseudoscorpion mitochondria show rearranged genes and genome-wide reductions of RNA gene sizes and inferred structures, yet typical nucleotide composition bias*. *BMC Evolutionary Biology*, 12: 31.
- Pedersen P. L., 1994. *ATP Synthase: The machine that makes ATP*. *Current Biology*, 4(12): 1138–1141.
- Perna N. T., Kocher T. D., 1995. *Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes*. *J. Mol. Evol.*, 41: 353–358.
- Perrière G. & Gouy M., 1996. *WWW-Query: An on line retrieval system for biological sequence banks*. *Biochimie*, 78: 364–369.

- Place A. R., Feng X., Steven C. R., Fourcade H. M, Boore J. L., 2005. *Genetic markers in blue crabs (Callinectes sapidus) II. Complete mitochondrial genome sequence and characterization of genetic variation*. J. Exp. Mar. Biol. Ecol., 319: 15–27.
- Ratnasingham S., Hebert P.D.N., 2007. BOLD: The Barcode of Life Data System (www.barcodinglife.org). Mol. Ecol. Notes, 7: 355–364.
- Raven, P., Johnson G., Mason K., Losos J., Singer S., 2010. *Biology*. McGraw-Hill 9th Ed.
- Rehbein H. 1995. *Electrophoretic techniques for species identification of fishery products*. Zeitschrift für Lebensmittel Uterssuchung und Forschung, 119: 1.
- Renon P., 2000. I prodotti della pesca. Agriteam Milano, Regione Lombardia.
- Richards T.A. & Archibald J.M., 2013. *Cell evolution: gene transfer agents and the origin of mitochondria*. Current Biology, 21: R12–R13.
- Rossier M.F., 2006. *T channels and steroid biosynthesis: in search of a link with mitochondria*. Cell Calcium, 40(2): 155–164.
- Saccone C., Gissi C., Reyes A., Larizza A., Sbisà E., Pesole G., 2002. *Mitochondrial DNA in metazoa: degree of freedom in a frozen event*. Gene, 286: 3–12.
- Salvato P., Simonato M., Battisti A., Negrisolo E., 2008. *The complete mitochondrial genome of the bag-shelter moth Ochrogaster lunifer (Lepidoptera, Notodontidae)*. BMC Genomics, 9: 331,1471–2164.
- Sanger F., Coulson A. R., 1975. *A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase*. J. Mol. Biol., 94(3): 441–448.
- Sanger F., Nicklen S., Coulson A. R., 1977. *DNA sequencing with chain-terminating inhibitors*. Proc. Natl. Acad. Sci. USA, 74(12): 5463–5467. .
- Schmidt H. A., Strimmer K., Vingron M., von Haeseler A., 2002. *TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing*. Bioinformatics, 18: 502–504.

- Sciuto K., 2008. *Cianobatteri e variazioni ambientali: effetti sui lipidi di membrana e sulle desaturasi*. [Tesi di Dottorato di ricerca in Biologia Evoluzionistica].
- Segawa R. D., Aotsuka T., 2005. *The mitochondrial genome of the Japanese freshwater crab, Geothelphusa dehaani (Crustacea: Brachyura): Evidence for its evolution via gene duplication*. *Gene*, 355: 28 – 39.
- Shen H., Braband A., Scholtz G., 2013. Mitogenomic analysis of decapod crustacean phylogeny corroborates traditional views on their relationships. *Mol. Phylogenet. Evol.*, 66(3): 776-789.
- Shen R., Fan J.B., Campbell D., Chang W., Chen J., Doucet D., Yeakley J., Bibikova M., Wickham Garcia E., McBride C., Steemers F., Garcia F., Kermani B.G., Gunderson K., Oliphant A., 2005. *High-throughput SNP genotyping on universal bead arrays*. *Mutat. Res.*, 573(1-2): 70–82.
- Shi H., Liu R., Sha Z., Ma j., 2012. *Complete mitochondrial DNA sequence of Stenopus hispidus (Crustacea: Decapoda: Stenopodidea) and a novel tRNA gene cluster*. *Marine Genomics* 6: 7–15.
- Simonato M., 2010. *Genetics and genomics of pine processionary moths and their parasitoids*. [Tesi di Dottorato di ricerca in Scienze delle Produzioni Vegetali].
- Simonato M., Battisti A., Kerdelhué C., Burban C., Lopez-Vaamonde C., Pivotto I., Salvato P., Negrisolo E. M. *Host and Phenology Shifts in the Evolution of the Social Moth Genus Thaumetopoea*. *PLoS ONE* 8(2): e57192
- Smith A. E. & Marcker K.A., 1968. *N-formylmethionyl transfer RNA in mitochondria from yeast and rat liver*. *J. Mol. Biol.*, 38: 241–243.
- Stamatakis, A. 2006. *RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models*. *Bioinformatics*, 22: 2688–2690.
- Strimmer K. & Von Haeseler A., 1997. *Likelihood-mapping: A simple method to visualize phylogenetic content of a sequence alignment*. *Proc. Natl. Acad. Sci. USA*, 94: 6815–6819.

- Sun H., Zhou K., Song D., 2005. *Mitochondrial genome of the Chinese mitten crab *Eriocheir japonica sinensis* (Brachyura: Thoracotremata: Grapsoidea) reveals a novel gene order and two target regions of gene rearrangements.* Gene, 349: 207–217.
- Tamura K., Peterson D., Peterson N., Stecher G., Nei M., Kumar S., 2011. *MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods.* Mol. Biol. Evol., 28(10): 2731–2739.
- Thompson J. D., Higgins D. G., Gibson T. J., 1994. *CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.* Nucleic Acids Res., 22: 4673–4680.
- Thrash C. J., Boyd A., Huggett M. J., Grote J., Carini P., Yoder R. J., Robbertse B., Spatafora J. W., Rappé M. S., Giovannoni J. S., 2011. *Phylogenomic evidence for a common ancestor of mitochondria and the SAR11 clade.* Nature, Scientific Reports, 1(13).
- Voet, D., Voet, J. G., Pratt, C. W., 2006. *Fundamentals of Biochemistry*, 2nd Edition. John Wiley and Sons, Inc., 547.
- Weinberg F. & Chandel N. S., 2009. *Mitochondrial Metabolism and Cancer.* Annals of the New York Academy of Sciences, 1177(1): 66–73.
- Whitworth T. L., Dawson R. D., Magalon H., Baudry E., 2007. *DNA barcoding cannot reliably identify species on the blowfly genus *Protophthora* (Diptera: Calliphoridae).* Proceeding of the royal society B, 274: 1731–1739.
- Will K. W., Rubinoff D., 2004. *Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification.* Cladistics, 20: 47–55.
- Wilson K., Cahill V., Ballment E., Benzie J., 2000. *The Complete Sequence of the Mitochondrial Genome of the Crustacean *Penaeus monodon*: are Malacostracan Crustaceans more closely related to Insects than to Branchiopods?* Mol. Biol. Evol., 17(6): 863–874.

- Wolstenholme D. R., 1992. *Animal mitochondrial DNA: structure and evolution*. Int. Rev. Cytol., 141: 173–216.
- Yamauchi M. M., Miya M. U., Nishida M., 2003. *Complete mitochondrial DNA sequence of the swimming crab, Portunus trituberculatus (Crustacea: Decapoda: Brachyura)*. Gene, 311: 129–135.
- Yang J. S. & Yang W. J., 2008. The complete mitochondrial genome sequence of the hydrothermal vent galatheid crab *Shinkaia crosnieri* (Crustacea: Decapoda: Anomura): a novel arrangement and incomplete tRNA suite. BMC Genomics, 9: 257.
- Zhang D. X., & Hewitt G. M., 1997. *Assessment of the universality and utility of a set of conserved mitochondrial primers in insects*. Insect Mol. Biol., 6: 143–150.

Ringraziamenti

Desidero innanzitutto ringraziare il Dr. Enrico Massimiliano Negrisolo, che con la sua semplicità e disponibilità mi ha permesso di fare questa esperienza e per avermi saputo consigliare durante tutte le analisi di laboratorio e la stesura della tesi.

Grazie a Laura per la pazienza nell'insegnarmi le tecniche di laboratorio e la precisione per i dettagli.

Grazie a Massimiliano per la simpatia e l'aiuto nell'assemblaggio dei genomi.

Grazie inoltre a tutto il dipartimento di BCA per avermi accolto e ospitato e per avermi sempre fatto sentire a mio agio.

Un profondo ringraziamento va a tutta la mia famiglia che con il loro sostegno morale ed economico hanno permesso di raggiungere questo risultato.

Grazie anche a tutti i miei amici, vecchi e nuovi che in questi anni mi hanno aiutato in tutte le nuove sfide.

Ed infine grazie a Valentina, per essermi accanto e l'immensa pazienza portata durante tutto questo periodo, sostenendomi sempre in tutte le mie scelte; grazie per credere in me e aiutarmi anche quando tutto "sembra perduto".