



UNIVERSITÀ POLITECNICA DELLE MARCHE
DIPARTIMENTO SCIENZE DELLA VITA E DELL'AMBIENTE

Corso di Laurea Magistrale
BIOLOGIA MOLECOLARE E APPLICATA

Analisi trascrittomiche e di microscopia elettronica rivelano differenze nella biomineralizzazione tra individui del bivalve *Chamelea gallina* provenienti da due diversi siti del Mar Adriatico.

Transcriptomic and electron microscopy analyses revealed differences in the biomineralization between individuals of the bivalve *Chamelea gallina* sampled from two different sites located in the Adriatic Sea.

Tesi di laurea magistrale di:

Alessandro Mosca

A. Mosca

Relatore Chiar.ma Prof.ssa

Maria Assunta Biscotti

Maria Assunta Biscotti
Correlatore:

Dott.ssa Federica Carducci

Federica

Sessione Estiva (Luglio 2021)

Anno Accademico 2020/2021

Sommario

Capitolo primo.....	5
1. INTRODUZIONE.....	5
<i>1.1 Chamelea gallina.....</i>	<i>5</i>
<i>1.1.1 Distribuzione e habitat</i>	<i>5</i>
<i>1.1.2 Biologia</i>	<i>6</i>
<i>1.1.3 Ecologia e commercio</i>	<i>9</i>
<i>1.2 La biomineralizzazione nei molluschi</i>	<i>11</i>
<i>1.2.1 Gli aspetti generali.....</i>	<i>11</i>
<i>1.2.2 Gli aspetti precoci durante lo sviluppo larvale</i>	<i>15</i>
<i>1.2.3 Le tipologie cellulari partecipanti al processo</i>	<i>16</i>
<i>1.2.4 Geni e prodotti genici coinvolti</i>	<i>19</i>
<i>1.2.5 Un potenziale modello per la biomineralizzazione</i>	<i>23</i>
<i>1.3 Influenza dei parametri ambientali sul processo di biomineralizzazione</i>	<i>26</i>
<i>1.3.1 La plasticità fenotipica degli organismi.....</i>	<i>26</i>
<i>1.3.2 I parametri principali</i>	<i>27</i>
<i>1.3.3 Un confronto in letteratura</i>	<i>29</i>
<i>1.4 Applicazioni del sequenziamento NGS nel campo della biologia della conservazione.....</i>	<i>31</i>
Capitolo secondo.....	33
2. MATERIALI E METODI.....	33
<i>2.1 Campionamento</i>	<i>33</i>
<i>2.2 La conservazione dei campioni</i>	<i>37</i>
<i>2.3 Il disegno sperimentale</i>	<i>38</i>
<i>2.3.1 Le metodiche sperimentali</i>	<i>38</i>
<i>2.3.2 Il sequenziamento</i>	<i>38</i>
<i>2.3.3 Le analisi bioinformatiche</i>	<i>39</i>
<i>2.3.4 SEM, microanalisi e diffrazione a raggi X.....</i>	<i>41</i>
<i>2.4 Purificazione dell'RNA totale</i>	<i>42</i>
<i>2.5 Elettroforesi su gel di agarosio</i>	<i>44</i>
<i>2.6 Quantificazione dell'acido nucleico e valutazione degli indici di purezza</i>	<i>45</i>
<i>2.6.1 La quantificazione dell'RNA</i>	<i>45</i>

2.6.2 <i>La valutazione degli indici di purezza</i>	46
2.7 <i>Costituzione dei pool</i>	49
2.8 <i>Sequenziamento NGS</i>	50
2.8.1 <i>Le valutazioni preliminari</i>	50
2.8.2 <i>La preparazione della libreria</i>	52
2.8.3 <i>Il sequenziamento</i>	53
2.8.4 <i>I risultati del sequenziamento</i>	55
2.9 <i>Le analisi bioinformatiche dei dati</i>	56
2.9.1 <i>Valutazione della qualità delle reads</i>	56
2.9.2 <i>Trimming dei dati</i>	59
2.9.3 <i>Assemblaggio delle reads filtrate</i>	60
2.9.4 <i>Valutazione della qualità dell'assemblaggio</i>	62
2.9.5 <i>Determinazione dei geni differenzialmente espressi (DEGs)</i>	63
2.9.6 <i>Annotazione dei DEGs</i>	68
2.10 <i>SEM, microanalisi e diffrazione a raggi X</i>	70
2.10.1 <i>SEM</i>	70
2.10.2 <i>Microanalisi</i>	71
2.10.3 <i>Diffrazione a raggi X</i>	71
Capitolo terzo	73
3. RISULTATI	73
3.1 <i>Valutazione della qualità delle reads</i>	73
3.2 <i>Trimming dei dati</i>	77
3.2.1 <i>Confronto qualità tra reads grezze e trimmed</i>	78
3.3 <i>Assemblaggio delle reads filtrate</i>	81
3.4 <i>Valutazione della qualità dell'assemblaggio</i>	82
3.5 <i>Determinazione dei geni differenzialmente espressi (DEGs)</i>	83
3.6 <i>Annotazione dei DEGs</i>	89
3.6.1 <i>Up-regolati Silvi Marina (o down-regolati Senigallia)</i>	89
3.6.2 <i>Down-regolati Silvi Marina (o up-regolati Senigallia)</i>	91
3.7 <i>SEM, microanalisi e diffrazione a raggi X</i>	93
3.7.1 <i>SEM</i>	93
3.7.2 <i>Microanalisi</i>	100

3.7.3 Diffrazione a raggi X	101
Capitolo quarto	102
4. DISCUSSIONE	102
Capitolo quinto	114
5. CONCLUSIONE	114
Capitolo sesto	116
6. BIBLIOGRAFIA	116

Capitolo primo

1. INTRODUZIONE

1.1 Chamelea gallina

1.1.1 Distribuzione e habitat

La *Chamelea gallina* (Linnaeus, 1758; Fig. 1) è un bivalve appartenente alla famiglia Veneridae ed un filtratore che vive infossato nei primi strati del substrato sabbioso a 3-7 m di profondità. Questa specie riveste una grande rilevanza commerciale per il settore della pesca (Froggia, 1989; Ramón & Richardson, 1992).



Fig. 1. Chamelea gallina. Da: www.marinespecies.org.

È ampiamente distribuita lungo la costa atlantica orientale, quindi presente in Norvegia e nelle Isole Britanniche fino alla penisola iberica, Marocco, Madera e alle Canarie (Gaspar & Monteiro, 1998), è anche presente nel Mar Nero, nel Mar Mediterraneo (Poppe & Goto, 1993) ed in particolare nell'Adriatico (Orban et al., 2007).

1.1.2 Biologia

Per quanto riguarda la conchiglia, tutti i bivalvi sono caratterizzati da due valve, nella maggior parte dei casi simmetriche come in *C. gallina*, connesse da una cerniera che comprende un legamento coriaceo e una serie di denti calcificati (Marin et al., 2012). In *C. gallina* la cerniera è costituita da incastri con 3 denti cardine in ciascuna valva e da legamenti. Il margine interno della conchiglia è seghettato finemente (Alemany, 1986). La colorazione esterna della conchiglia varia dal bianco al grigio chiaro o bluastra. Diversi esemplari presentano fasce radiali bianche e/o linee a zig-zag scure. All'interno delle valve (Fig. 2), il mantello racchiude i vari organi e strutture interne: branchie, sifoni inalante ed esalante, cuore, centri nervosi, muscoli adduttori, organi riproduttivi, palpi labiali, stomaco, intestino, ecc. (Cattaneo & Bernardi, 2008).

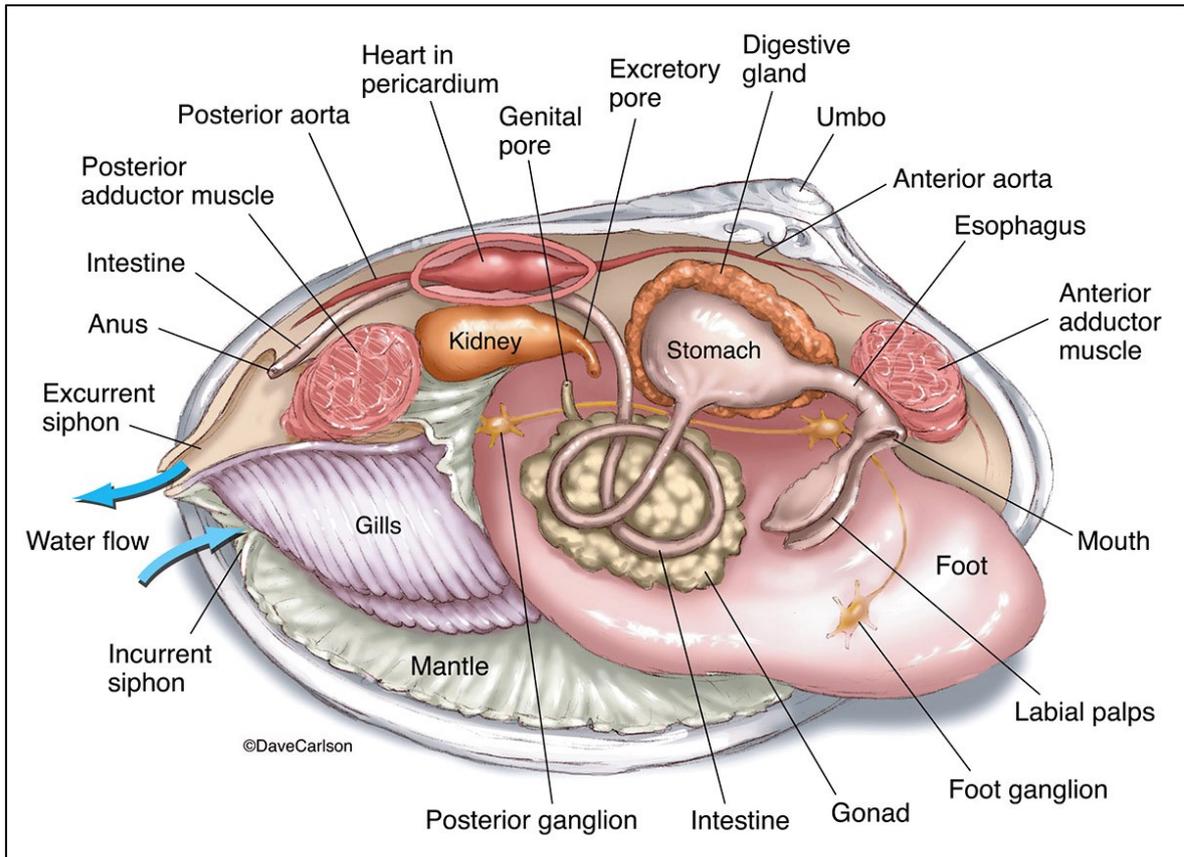


Fig. 2. Anatomia di *Chamelea gallina*. Da: www.anatomynote.com.

I muscoli adduttori hanno la funzione di tener chiuse le valve opponendo una forza contraria a quella esercitata dal legamento coriaceo della cerniera, che invece tende a far aprire le due valve. I due sifoni sono piuttosto corti e risultano fusi assieme (Morton, 2020). Le ciglia vibratili, di cui sono dotati gli epitelii branchiali e del mantello, generano correnti in grado di pompare acqua dall'esterno attraverso il sifone inalante, che poi passa all'interno attraverso la cavità del mantello e, una volta filtrata dalle branchie, viene espulsa attraverso il sifone esalante. L'acqua permette la respirazione: gli scambi ossigeno-anidride carbonica avvengono in prevalenza attraverso il mantello e le branchie

(Morton, 2020). Le branchie, oltre alla funzione respiratoria, fungono da filtro per le particelle alimentari che vengono inglobate in una sostanza mucosa secreta dall'epitelio branchiale e poi convogliate, tramite un sistema di ciglia, verso i palpi labiali situati sul margine ventrale, e quindi alla bocca. Il materiale più grossolano viene accumulato nella regione posteriore ed espulso mediante contrazione dei muscoli adduttori, mentre la rimanente parte prosegue lungo l'apparato digerente (Morton, 2020). Lo stomaco è un sacco circondato da una grossa ghiandola digestiva, le cui cellule inglobano il materiale minuto che arriva allo stomaco e lo digeriscono mediante i loro enzimi endocellulari. L'intestino decorre nel piede e dopo una o più circonvoluzioni, attraversa il pericardio e termina nella cavità posteriore del mantello (Morton, 2020). Il meccanismo di filtrazione dei bivalvi li porta ad accumulare nei tessuti le sostanze disciolte nell'acqua e quindi, in presenza di contaminanti, anche quest'ultimi: questa caratteristica rende la classe Bivalvia di particolare interesse per lo studio sulla presenza e sull'effetto di inquinanti ambientali. Questi organismi hanno, inoltre, un sistema circolatorio aperto e il sangue passa dai vasi a un sistema di lacune. Il sangue contiene un pigmento respiratorio (emocianina) che si ossigena a livello delle branchie (Morton, 2020). *C. gallina* è una specie gonocorica con un lungo periodo di emissione dei gameti, da aprile ad agosto (Casali, 1984; Frogliani, 1975; Keller et al., 2002; Valli et al., 1985;

Dalgıç et al., 2009; Delgado et al., 2013; Romanelli et al., 2009), durante il quale l'emissione avviene ad intervalli ed è sincronizzato al fine di liberare entrambi i tipi di gameti (spermi diflagellati e uova) per la fecondazione esterna. Le uova fecondate affrontano uno sviluppo indiretto: la prima fase prevede una larva *trocofora* natante, a cui fa seguito un *veliger* munito di *velum* che consente di nuotare ed alimentarsi. Dopo un periodo di vita planctonica, il *veliger* discende nel substrato per affrontare la metamorfosi in giovanile, in cui il *velum* si perde e si sviluppano il piede, la ghiandola digestiva e gli organi riproduttivi (Marin et al., 2012).

1.1.3 Ecologia e commercio

C. gallina, similmente ad altri bivalvi, ha una notevole importanza ecologica negli ecosistemi, e nello specifico nel ciclo dei nutrienti, nella creazione e modificazione degli habitat e nella rete trofica (Vaughn & Hoellein, 2018). Inoltre, grazie all'alimentazione per filtrazione, stile di vita sessile, ampia distribuzione, bioaccumulo di sostanze nei tessuti e alla possibilità di studiarne differenti stadi del ciclo vitale, molti bivalvi possono essere considerati biosensori per l'inquinamento e più in generale per i cambiamenti nelle acque costiere, permettendo il monitoraggio della qualità delle zone intertidali (Fitzer et al., 2014; Li et al., 2013; Luchmann et al., 2011; Philipp et al., 2012; Alhejoj et al., 2017; González-Domínguez et al., 2016; La Guardia et al., 2012; Noh et

al., 2018; Rittschof & McClellan-Green, 2005; Teixeira et al., 2017). Diversi fattori, che includono la variazione della temperatura, della salinità, la modifica del sedimento e della concentrazione di ossigeno in acqua (Ezgeta-Balic et al., 2011; Ivanina et al., 2012; Matozzo et al., 2012; Monari et al., 2011; Sobral & Widdows, 1997), la presenza di sostanze tossiche di origine antropogenica (Visciano et al., 2015) e la pesca eccessiva (Moschino et al., 2003) potrebbero aver causato una diminuzione di questa risorsa ittica (Moschino et al., 2003). Infatti, dagli anni '70, lo sviluppo della pesca delle vongole basata sulle draghe idrauliche ha portato ad uno sfruttamento intensivo di questa risorsa provocando una forte diminuzione nella densità di popolazione, associata alla riduzione del numero di individui oltre i 25mm, ovvero la taglia minima commerciale (Froggia, 1989). Negli ultimi anni, c'è stata una preoccupazione crescente per la sopravvivenza delle comunità dei bivalvi, ad esempio a causa delle grandi fluttuazioni annuali nell'abbondanza dei banchi e di irregolari eventi di mortalità, presumibilmente dovuti ad anossia, esondazione dei fiumi, inquinamento, infezioni batteriche e/o virali (Romanelli et al., 2009; Milan et al., 2016; Milan et al., 2019). Tutto ciò minaccia la sostenibilità economica e biologica di questa pesca, specialmente nel Mar Adriatico (Ramón & Richardson, 1992; Romanelli et al., 2009), in cui rappresenta una delle maggiori risorse con una resa annua di circa 100 milioni di euro (Ministerial

decreto n.27 del 17/6/2019). La Food and Agriculture Organization (FAO) riporta una media annuale di pesca totale di 60k tonnellate (2004-2013) nell'Atlantico, Mediterraneo e Mar Nero, con la pesca maggiore in Italia (22k) e Turchia (33k).

1.2 La biomineralizzazione nei molluschi

1.2.1 Gli aspetti generali

I molluschi sono animali in cui è assente sia uno scheletro esterno che uno interno e a causa di questo molte specie hanno adottato una complessa strategia per preservare i propri tessuti, proteggersi dalla predazione ed evitare l'essiccamento. Questa strategia prevede la costruzione di una struttura calcificata esterna, cioè la conchiglia. Questa struttura ed il suo processo di formazione rientrano nella categoria delle mineralizzazioni biologicamente controllate (al contrario di quelle biologicamente indotte, tipiche dei batteri). Questa categoria è stata definita da Mann (1983) e può essere schematizzata nel seguente modo:

A) la fabbricazione della conchiglia richiede uno speciale macchinario biologico, sia intracellulare che extracellulare, ne consegue che la formazione della conchiglia è sotto il controllo di cascate geniche;

B) i minerali formati per costruire la conchiglia sono lontani dall'equilibrio con l'ambiente, ciò significa che alcuni di essi, instabili in condizioni naturali, possono comunque essere sintetizzati;

C) i minerali prodotti differiscono nella forma e nella taglia rispetto alle loro controparti formate inorganicamente, inoltre al contrario di queste ultime si assemblano seguendo livelli gerarchici diversi;

D) i minerali prodotti sono formati in uno spazio delimitato, non a diretto contatto con l'esterno;

E) il processo di costruzione è organizzato da una matrice organica extracellulare, una parte della quale viene occlusa nella conchiglia durante la calcificazione. La matrice organica è secreta dall'epitelio del mantello e contiene proteine, peptidi, lipidi e carboidrati (Lowenstam & Weiner, 1989; Addadi et al., 2006).

In linea con i risultati ottenuti nei vari studi presenti in letteratura, le cellule ed i loro prodotti cooperano per produrre una conchiglia organica minerale, composta da strutture minerali e componenti macromolecolari organiche.

In media, la conchiglia dei molluschi è composta dal 95% di carbonato di calcio (CaCO_3) e dal 5% di matrice organica, il CaCO_3 in condizioni naturali esiste come differenti polimorfi cristallini (es. calcite, aragonite, vaterite). I bivalvi generalmente sequestrano il calcio e il bicarbonato dall'ambiente e lo

metabolizzano per formare il CaCO_3 che si ritrova nella conchiglia, e questo viene organizzato in strati con schemi distintivi per formare le microstrutture della conchiglia. Più di trenta diverse microstrutture del CaCO_3 sono state ritrovate nelle conchiglie, come quelle madreperlacee, fogliate, prismatiche, lamellari incrociate, omogenee, ciò grazie ad osservazioni al microscopio elettronico a scansione (Chateigner et al., 2000; Kouchinsky, 2000; Furuhashi et al., 2009). Una tipica conchiglia dei molluschi mostra una struttura tripartita (Fig. 3): lo strato più esterno, ovvero il periostraco (un sottile strato coriaceo organico) e due strati calcificati, di cui uno esterno prismatico ed uno interno madreperlaceo (Saruwatari et al., 2009).

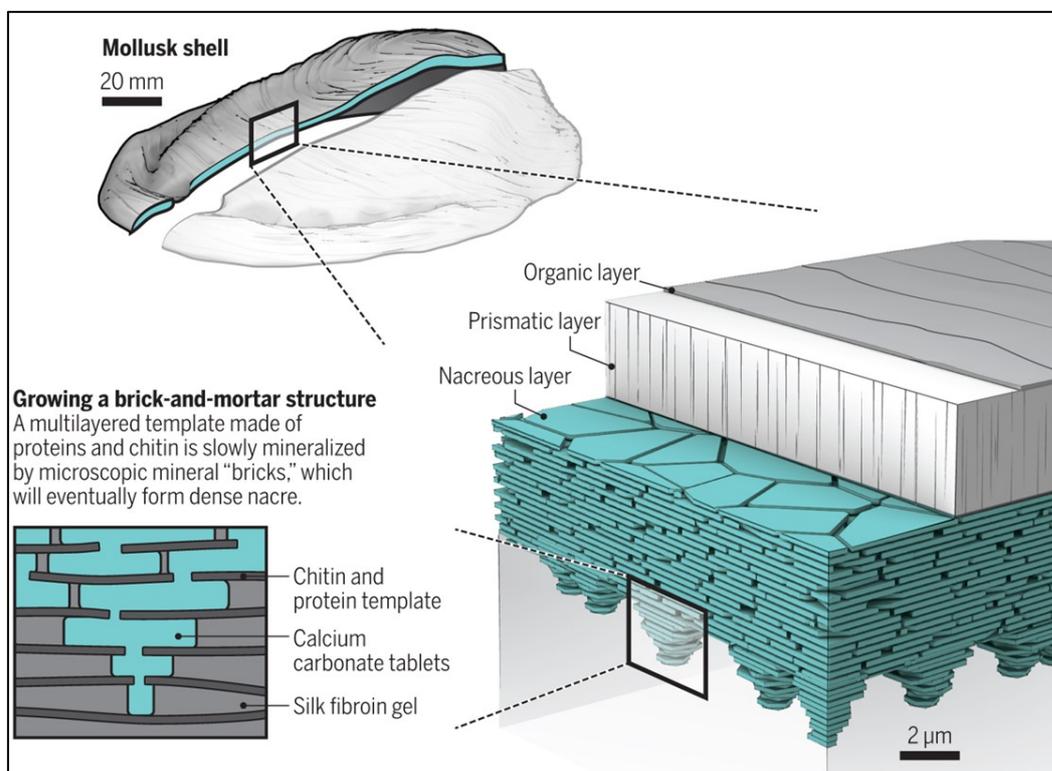


Fig. 3. Struttura tripartita del guscio. Modificata da: Barthelat, F. Growing a synthetic mollusk shell. *Science* 354, 6308 (2016).

Lo strato prismatico è composto da cristalli di calcite allungati formanti prismi perpendicolari al periostraco; lo strato madreperlaceo è composto da strutture laminari di cristalli di aragonite, organizzate in una struttura a mattoni (Marie et al., 2012). La madreperla è una microstruttura interessante a causa delle sue proprietà anti-frattura, è una struttura rigida che si spezza se piegata. Nella conchiglia questo aspetto viene bilanciato dal periostraco e dallo strato prismatico che offrono flessibilità alla struttura totale (Marin et al., 2012). Nel lavoro pubblicato da Gizzi e colleghi (2016) sono state ottenute informazioni specifiche per quanto riguarda la microstruttura della conchiglia di *C. gallina*, tramite osservazioni al microscopio elettronico. Queste mostrano che, in accordo a quanto osservato nella maggior parte dei Veneridae (Popov, 1986), *C. gallina* contiene due strati principali: in quello esterno si osservano prismi composti, formati dall'assemblaggio compatto di grani, lo strato interno è omogeneo e formato da granuli irregolari e intorno a questi granuli (di taglia 1 micron) si trovano strati di materiale disperso. È poi presente uno strato intermedio, cioè una zona di transizione tra i due strati principali, caratterizzato dalla presenza di granuli più grandi di quelli osservati nello strato interno (intorno ai 3 micron), e dai quali originano i prismi. I granuli osservati nello strato interno ed esterno sono abbastanza simili in taglia e forma. L'analisi della fase inorganica ottenuta dai risultati della diffrazione a raggi X e dalla

spettroscopia a infrarossi mostra che le conchiglie di *C. gallina* sono composte da sola aragonite (Gizzi et al., 2016). Invece, per quanto riguarda il contenuto percentuale (in peso) di matrice organica nella conchiglia, i dati ottenuti tramite analisi termogravimetrica indicano una percentuale minore del 2% ed omogenea nei campioni considerati nello studio (Gizzi, 2016).

1.2.2 Gli aspetti precoci durante lo sviluppo larvale

Come riassunto da Marin e colleghi (2012), nei bivalvi il primo evento della formazione della conchiglia inizia alla fine della gastrulazione, quando un gruppo di cellule epiteliali si ispessiscono definendo il cosiddetto “shell field” (Kniprath, 1981). Queste cellule, durante lo stadio di *trocofora*, si invaginano temporaneamente formando così la ghiandola del guscio; le cellule periferiche non invaginate producono una lamella extracellulare (il futuro periostraco) che fornirà il supporto organico laddove potrà essere deposta la prima conchiglia. Dopo la formazione di questa lamella, la ghiandola evagina e lo shell field si estende tramite l'appiattimento delle cellule e la loro mitosi, per poi diventare il mantello calcificante (Kniprath, 1981). La prima mineralizzazione tra periostraco e shell field porterà alla formazione della conchiglia primaria o primo *prodissoconch* che si verifica durante lo stadio di *trocofora*, al quale seguirà il secondo *prodissoconch* o conchiglia secondaria, durante lo stadio di *veliger*. Dopo la metamorfosi, il giovanile produce il *dissoconch*, separato dal

secondo *prodissoconch* da una cresta sulla superficie esterna della conchiglia (Jablonski & Lutz, 1980).

1.2.3 Le tipologie cellulari partecipanti al processo

Per la formazione della conchiglia, è stata tradizionalmente riconosciuta l'ipotesi della mediazione della matrice, in cui si sostiene che la matrice organica controlli il processo di formazione della conchiglia fornendo l'impalcatura, inducendo la nucleazione, regolando la crescita dei cristalli extracellulari e formando in questo modo le morfologie cristalline specifiche per i vari strati della conchiglia (Addadi et al., 1987; Addadi et al., 2006; Lowenstam & Weiner, 1989). Un'alternativa a questa ipotesi è quella della mediazione cellulare, che propone che la nucleazione avvenga negli emociti e/o nelle cellule epiteliali del mantello esterno e che cellule portanti cristalli trasportino questi ultimi al fronte di mineralizzazione (Mount et al., 2004; Gong et al., 2008; Xiang et al., 2014). Ad oggi, ad essere coinvolte nella formazione della conchiglia sono state osservate due differenti tipologie cellulari: da una parte abbiamo cellule provenienti dal mantello esterno, coinvolte nella nucleazione, nei processi di rimodellamento dei minerali di CaCO_3 (Kong et al., 2015) e nella secrezione della matrice organica, nella quale le proteine della matrice regolano la struttura delle conchiglia assemblando i cristalli di CaCO_3 in maniera specifica (Lowenstam & Weiner, 1989; Zhang & Zhang, 2006;

Marin et al., 2008; Ivanina et al., 2017); dall'altra troviamo gli emociti che partecipano alla deposizione dei cristalli intracellulari di CaCO_3 e li portano al sito di mineralizzazione (Mount et al., 2004; Fleury et al., 2008; Kádár, 2008; Mount & Pickering, 2009; Cho & Jeong, 2011; Johnstone et al., 2015; Li et al., 2016; Ivanina et al., 2017).

Per quanto riguarda il mantello, esso può essere diviso in diverse regioni specializzate: inferiore, interno ed esterno. L'epitelio esterno è stato collegato ai processi di formazione della conchiglia ed è inoltre in prossimità del fronte di mineralizzazione (Nudelman et al., 2006). Tra le cellule sulla superficie dell'epitelio esterno c'è una ulteriore zonazione (del bordo e palliali) e ciò sembra essere strettamente associato a diverse microstrutture nella formazione della conchiglia (Marin et al., 2008). Generalmente le proteine di matrice secrete dal bordo sono coinvolte nella formazione dello strato prismatico, mentre quelle palliali partecipano alla madreperla (Miyamoto et al., 1996). Come descritto in Marin et al., 2012, il bordo del mantello presenta una serie di pieghe (generalmente 3 nei bivalvi), e tra la cresta della piega esterna e quella centrale si crea una scanalatura nella quale cellule specializzate secernono il periostraco.

Il periostraco ha diversi ruoli: delimita uno spazio tra il mantello e la conchiglia separato dall'ambiente esterno e chiamato spazio extrapalliale che può essere

sovrasaturo di CaCO_3 , protegge dalla dissoluzione della conchiglia (soprattutto in ambienti acidi; Isaji, 1993), costituisce una barriera contro il fouling (Bers et al., 2006), ed essendo spesso pigmentato costituisce un mezzo per contrastare la predazione, inoltre nella larva (Kniprath, 1981) fornisce lo stampo primario per la mineralizzazione extracellulare. Il periostraco viene secreto come un film di precursori solubili instabili che sclerotizzano e diventano insolubili appena rilasciati in ambiente extracellulare (Waite, 1983). Inoltre, esso può poi mantenersi o essere abraso e perdersi.

Per quanto riguarda gli emociti, essi sono essenziali durante la risposta immunitaria innata, che si è anche mostrata essere collegata alla biomineralizzazione; infatti, ad esempio, geni immuno-associati mostrano un' espressione elevata durante i processi di rigenerazione della conchiglia (Sleight et al., 2015).

Recentemente, sempre più evidenze dimostrano la partecipazione diretta degli emociti (perlopiù granulociti) nella sintesi e trasporto dei cristalli di CaCO_3 così come per le proteine durante la rigenerazione (Mount et al., 2004; Mount & Pickering, 2009; Cho & Jeong, 2011; Johnstone et al., 2015; Li et al., 2016). Le inclusioni granulari degli emociti consistono in ioni Ca^{2+} e ciò è stato visto in molte specie di molluschi (Fleury et al., 2008; Ivanina et al., 2017; Kadar et al., 2009; Li et al., 2016), suggerendo che i granulociti possono agire come

riserve e trasportatori di calcio durante la formazione della conchiglia (Kadar et al., 2009; Cho & Jeong, 2011). Inoltre, cristalli intracellulari di CaCO_3 si osservano in emociti di vari molluschi (Mount et al., 2004). Gli emociti sono presenti sulla superficie dello strato prismatico allineati con i materiali fibrosi e i cristalli (Mount et al., 2004), ciò presuppone che gli emociti siano presenti al fronte di mineralizzazione. Oltre alla sintesi e al trasporto dei cristalli, gli emociti sono stati associati anche alla formazione della struttura organica, infatti dati multi-omici indicano che varie proteine della matrice sono fortemente espresse negli emociti di vari molluschi, soprattutto durante i processi di rigenerazione (Liao et al., 2015; Li et al., 2016; Ivanina et al., 2017; Shitalbahen, 2004; Johnstone et al., 2008).

1.2.4 Geni e prodotti genici coinvolti

Dalle informazioni raccolte da Song e colleghi (2019), i costituenti organici della conchiglia sintetizzati prevalentemente dalle cellule del mantello esterno possono essere macro-classificati in due categorie: solubili o insolubili. Per le insolubili abbiamo ad esempio la fibroina e la chitina, quest'ultima presente sia nello strato prismatico che in quello madreperlaceo, giocando probabilmente un ruolo importante nella biomineralizzazione (Furuhashi et al., 2009); inoltre le chitina-sintasi sono espresse nel bordo del mantello, contribuendo alla formazione della struttura per la calcificazione (Suzuki, 2007). Le insolubili

agiscono quindi come impalcatura e rafforzano le proprietà meccaniche del guscio, mentre le solubili determinano le proprietà minerali e cristallografiche (Addadi et al., 2006; Morse et al., 2007; Cusack & Freer, 2008; Marin et al., 2008; Marie et al., 2012). Per quanto riguarda lo strato madreperlaceo, costituito principalmente da tavolette di aragonite (Marie et al., 2015; Marie et al., 2017), tra le proteine di matrice identificate ci sono: la proteina acida di matrice Pif che regola la formazione della madreperla (Suzuki et al., 2009), MSI60 (Sudo et al., 1997), lustrina A (Shen et al., 1997), N16/perlina (Samata et al., 1999; Miyashita et al., 2000), perlucina, perlustrina (Weiss et al., 2000), N14, N66 (Kono et al., 2000), mucoperlina (Marin et al., 2000), AP7, AP24 (Michenfelder et al., 2003), P10 (Zhang et al., 2006b), perlwapina (Treccani et al., 2006), perlinibina (Mann et al., 2007), N19 (Yano et al., 2007), N40 (Yan et al., 2007) e BMSP (Suzuki et al., 2011). Invece, per quanto riguarda lo strato prismatico, costituito principalmente da cristalli di calcite cilindrici (Carter, 1985; Chateigner et al., 2000), tra le proteine di matrice identificate ci sono: KRMPs (Zhang & Zhang, 2006), MSI31 (Sudo et al., 1997), prismaticina-14 (Suzuki et al., 2004), aspeina (Tsukamoto et al., 2004), prisilkina-39 (Kong et al., 2009), P43 (Gong et al., 2008), Alv (Kong et al., 2018), e asprich proteins (Gotliv et al., 2005; Politi et al., 2007). Per quanto riguarda le proteine di matrice coinvolte nella formazione di entrambi gli strati si hanno: nacreina,

espressa in tutto l'epitelio del mantello e con importanti funzioni nella produzione di entrambi gli strati (Miyamoto, 2005; Takeuchi & Endo, 2006; Miyamoto et al., 1996), MSI7 (Zhang et al., 2003; Feng et al., 2009), PfY2 (Yi et al., 2017), shemitrine (Yano et al., 2006) e N25 (Yang et al., 2019). Nel fluido extrapalliale (EPF), ovvero nel microambiente acquoso che si trova tra mantello esterno e la faccia inferiore della conchiglia, e che serve da mezzo finale per la calcificazione (Saha et al., 1988; Lowenstam & Weiner, 1989), ci sono diversi tipi di ioni la cui concentrazione relativa è diversa da quella dell'acqua marina e dell'emolinfa (Saha et al., 1988; Wilt, 2002). Inoltre, sono presenti macromolecole secrete dalle cellule epiteliali del mantello esterno o da altri organi come proteine, polisaccaridi e lipidi che vengono trasportate nell'EPF (Kylmaja et al., 2016). Tra le proteine dell'EPF abbiamo EF fluid protein identificata in *Mytilus edulis* (Hattan et al., 2001; Yin et al., 2005), ACCBP ovvero una proteina di legame a carbonati di calcio amorfi, che sono i precursori di diversi polimorfi di CaCO_3 (Ma et al., 2007), SPARC identificata in *P. fucata* (Xie, 2016). In aggiunta, nello strato organico tra la madreperla di aragonite e lo strato prismatico di calcite in *P. fucata* (Yan et al., 2007) si trova una proteina simile alla calmodulina (CaLP), contenente domini di legame al calcio e questo suggerisce che CaLP potrebbe essere coinvolta nella crescita di entrambi gli strati. Per quanto riguarda invece i domini funzionali nel lavoro di

Arivalagan e colleghi (2016) sono stati identificati domini funzionali (kit di base condiviso per la costruzione della conchiglia) comuni alle 4 specie considerate ovvero *Crassostrea gigas*, *M. edulis*, *Mya truncata*, *Pecten maximus*. Una parte del kit di base è rappresentata dai seguenti domini: tirosinasi, anidraasi carbonica, dominio di legame alla chitina-2 (ChBt-2), fattore A di Von Willenbrand (VWA). La funzione principale del dominio della tirosinasi negli invertebrati è la sclerotizzazione, in questo processo polimeri proteici insolubili (Nagai et al. 2007) sono formati trasformando la tirosina in L-DOPA, lo stesso meccanismo impiegato dai bivalvi per la produzione del periostraco (Zhang et al. 2006a). Le anidraasi carboniche appartengono alla superfamiglia delle metallo proteine che catalizzano l'idratazione reversibile della CO₂ durante la formazione del CaCO₃, proteine contenenti questo dominio hanno ruoli sia nella calcificazione (Marie et al. 2008) sia nell'inibizione del processo di calcificazione (Miyamoto et al. 1996). Il ruolo principale di ChBt-2 nelle proteine di matrice è quello di interagire con la chitina in quanto componente principale della matrice (Mann, 1988). Il dominio VWA fornisce funzionalità adesive proteina-proteina (Whittaker & Hynes, 2002). I domini comuni alle conchiglie con uno strato di calcite sono: fattore di crescita epidermica (EGF), fibronectina 3 (FN3), proteina del siero del latte (WAP). Il dominio EGF è un motivo legante calcio (Stenflo et al., 2000).

Hanein e colleghi (1993), hanno dimostrato l'affinità della fibronectina per la calcite in presenza di acqua, quindi il dominio della fibronectina potrebbe aiutare nell'organizzazione specifica dei cristalli di calcite nell'idrogel della matrice. I whey-acidic-protein domains (WAP) sono domini inibitori delle proteasi molto conservati ed è stato dimostrato che hanno un ruolo nella formazione della conchiglia (Shen et al., 1997; Treccani et al., 2006; Marie et al., 2010). Per quanto riguarda lo strato di aragonite, il dominio Lam-G esibisce una funzione simile all'EGF (Panayotou et al., 1989). La presenza di un gran numero di proteine con domini associati a funzioni immunitarie sembra rafforzare l'ipotesi di una funzione protettiva della conchiglia (Arivalagan et al., 2016).

1.2.5 Un potenziale modello per la biomineralizzazione

Tenendo conto delle varie scoperte che si sono susseguite negli anni, Song e colleghi (2019) hanno proposto un potenziale modello e meccanismo per la biomineralizzazione della conchiglia (Fig. 4).

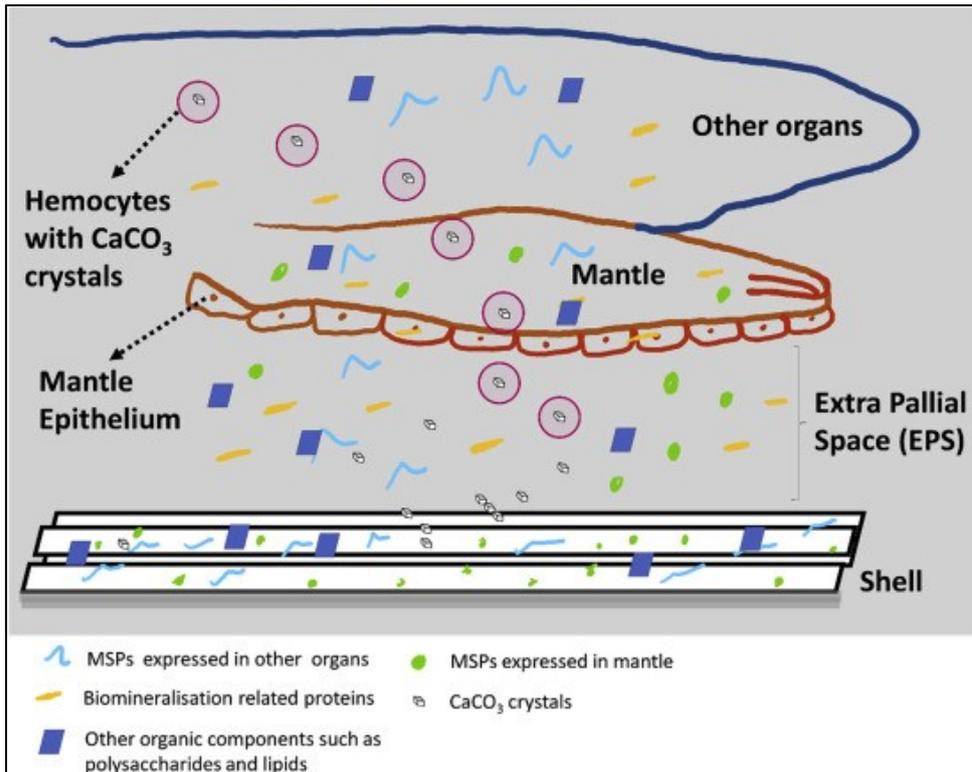


Fig. 4. Elementi coinvolti nella formazione del guscio. Da: www.researchgate.net.

Calcite e aragonite sono due polimorfi comuni del CaCO₃ presenti nella conchiglia e differiscono nell'organizzazione e orientamento delle molecole di carbonato (Stenzel, 1963). Hare (1963) studia le differenze nella composizione della matrice degli strati prismatici e madreperlacei e propone perciò che le proteine della matrice possano essere responsabili per la formazione di varie strutture della conchiglia e delle fasi minerali. È stato riportato che un'alta concentrazione di ioni Mg²⁺ promuove l'accumulo di aragonite in soluzione di CaCO₃ (Tokuyama, 1969), suggerendo quindi che la concentrazione degli ioni in soluzioni possa regolare il polimorfismo del carbonato nella conchiglia

(Blackwelder et al., 1976; Wilbur & Bernhardt, 1984). Al fine di determinare il significato delle matrici organiche, specialmente quello delle proteine nella formazione della conchiglia, Falini e colleghi (1996) assemblarono un substrato composto da chitina e fibroina per la nucleazione dei cristalli, scoprendo che proteine estratte dagli strati di aragonite e calcite inducevano maggiormente la formazione di aragonite e calcite rispettivamente. Diversi studi successivi hanno quindi supportato il modello della matrice, proponendo cioè che la calcificazione avvenga nella struttura di chitina, fibroina e proteine di matrice secreta dal mantello e auto-assemblante in strati successivi (Watabe, 1965; Addadi et al., 2006; Furuhashi et al., 2009). La prima osservazione di emociti al fronte di mineralizzazione è stata fatta da Mount e colleghi (2004). I dati sul secretoma e trascrittoma rivelano che alcune proteine di matrice sono non-secretorie (cioè non secrete dalle cellule epiteliali del mantello esterno), mentre altre sono abbondanti in altri tessuti oltre al mantello (Marie et al., 2010; Marie et al., 2012; Kinoshita et al., 2011; Mann et al., 2012; Wang et al., 2013); inoltre biominerali e proteine di matrice sono state osservate negli emociti (Johnstone et al., 2008; Calvoiglesias et al., 2016). Si può quindi supporre che la maggior parte delle proteine di matrice venga prodotta e secreta dalle cellule epiteliali del mantello esterno, mentre una minoranza da emociti o altri organi (Miyamoto et al., 2013; Wang et al., 2013). Le proteine di matrice prodotte dal

mantello esterno sono direttamente trasportate alla superficie di cristallizzazione tramite esosomi (Zhang et al., 2012) o pathway secretori classici, mentre quelle prodotte da altri organi sono prima trasferite all'emolinfa con pathway secretori classici o non (Gardella et al., 2002) e poi fagocitate dagli emociti (perlopiù granulociti). I cristalli di CaCO_3 sono prodotti sia dal mantello che sintetizzati dagli emociti e sono portati al fronte di biomineralizzazione attraverso il fluido extra palliale (Mount et al., 2004; Weiner & Addadi, 2011; Li et al., 2016).

1.3 Influenza dei parametri ambientali sul processo di biomineralizzazione

1.3.1 La plasticità fenotipica degli organismi

L'espressione genica può essere modificata in risposta a parametri sia biotici (come predazione, conspecifici, disponibilità cibo, ecc) che abiotici (come temperatura, fotoperiodo, ecc) (Gilbert, 2001). Questi parametri influenzano l'organismo (la sua espressione genica), producendo un fenotipo adattato a quell'ambiente. La variazione fenotipica intraspecifica, in genere, può essere prodotta da: pressione diversa da parte di ambienti differenti, variazione fenotipica indipendente dall'ambiente, effetti non adattativi dell'ambiente sul fenotipo (Schluter, 2000). L'abilità di produrre diversi fenotipi sotto i vari parametri ambientali è una questione critica per capire come le specie possano

affrontare cambiamenti futuri. La plasticità fenotipica è l'abilità di un organismo di produrre un range di fenotipi relativamente adattati, modificando morfologia, comportamento e attività biologica in risposta a fluttuazioni ambientali (DeWitt & Scheiner, 2004). Considerando l'esistenza dell'attività antropogenica sull'ambiente, gli organismi esibenti alta plasticità fenotipica potrebbero affrontare meglio disturbi su grande scala, come ad esempio il cambiamento climatico (Pigliucci et al., 2006).

1.3.2 I parametri principali

Gli organismi calcificanti, come i molluschi, sono probabilmente tra i più suscettibili al cambiamento dei parametri ambientali, mostrando variazioni morfologiche della conchiglia in base alle caratteristiche del substrato (topografia, tipologia, ecc), processi idrodinamici, e soprattutto al pH e alla temperatura (DeWitt & Scheiner, 2004; Vogel, 1996; Watson et al., 2012). I molluschi utilizzano il CaCO_3 , uno dei minerali più abbondanti in natura, come materiale strutturale/protettivo attraverso il processo di biomineralizzazione per la costruzione della conchiglia. Morfologia, mineralogia e chimica delle strutture di CaCO_3 sono largamente dipendenti dalla specie considerata e dall'ambiente circostante, infatti le proteine strutturali e gli enzimi coinvolti controllano le condizioni interne e rispondono ai parametri ambientali esterni (Lowenstam & Weiner 1989; Falini et al., 1996). La conchiglia dei molluschi

marini registra i fattori ambientali sotto forma di proprietà geochimiche, microstrutturali, nel tasso di crescita (Purroy et al., 2018) e, in base alla specie e all'habitat, le variabili ambientali possono essere anche più rilevanti di quelle fisiologiche nella crescita della conchiglia (Purroy et al., 2018, Nishida et al., 2012). I molluschi sono abili ad esercitare un fine controllo sul processo di biomineralizzazione, determinando quale tipo di polimorfo di CaCO_3 precipitare attraverso il controllo delle macromolecole di matrice (Falini et al., 1996). Fattori ambientali come radiazione solare, disponibilità di nutrienti, ossigeno, salinità e temperatura, influenzano tutti la spesa energetica degli organismi marini.

Gizzi e colleghi (2016) hanno valutato gli effetti della radiazione solare (SR) e della temperatura superficiale delle acque (SST), parametri ampiamente utilizzati per studi ecologici e che generalmente influenzano le proprietà scheletriche degli organismi marini calcificanti. Per studiare l'effetto di SR e SST sugli organismi marini, i gradienti latitudinali sono utili laboratori naturali, in cui SR e SST variano, permettendo di esaminare gli effetti a lungo termine su popolazioni diverse della stessa specie, adattate a differenti condizioni ambientali (Watson et al., 2012; Jansen et al., 2007). Come già detto, la morfologia della conchiglia è particolarmente sensibile ai parametri ambientali, variando in relazione a profondità (Claxton et al., 1998), corrente (Fuiman et

al., 1999), esposizione alle onde (Denny & Blanchette, 2000), tipologia di sedimento (Claxton et al., 1998; Newell & Hidu, 1982), temperatura e pH (Watson et al., 2012). In tempi recenti l'attenzione è stata particolarmente rivolta agli effetti dell'acidificazione degli oceani sui molluschi e sulle loro strutture calcificate (Clark et al., 2020). Infatti, la continua riduzione del pH marino e i relativi cambiamenti nella chimica del carbonato producono un costo maggiore per i processi di biomineralizzazione (Wittmann & Pörtner, 2013). La temperatura delle acque ha un ruolo importante nella crescita della conchiglia di *C. gallina*, temperature sotto i 10 °C rallentano fortemente la crescita, e temperature sopra i 28°C riducono l'assorbimento di energia e incrementano la spesa energetica tramite respirazione, quindi anche quest'ultime riducono la crescita della conchiglia (Ramón & Richardson, 1992; Moschino & Marin, 2006). La calcificazione di *C. gallina* sembra essere correlata alla temperatura e alla disponibilità di cibo, mostrando ampie bande di crescita durante inverno-primavera e bande strette durante il periodo estate-autunno (Ramón & Richardson, 1992). In generale c'è quindi un effetto complessivo derivante dalla combinazione dei diversi parametri.

1.3.3 Un confronto in letteratura

Per fare un confronto dei risultati ottenuti in letteratura, che saranno anche ripresi e integrati in relazione alle nostre osservazioni, secondo Gizzi e colleghi

(2016) e, con conclusioni pressoché analoghe, secondo Mancuso e colleghi (2019), la conchiglia degli individui di *C. gallina* delle aree più fredde dell'Adriatico, in cui è stato effettuato il campionamento e che sono caratterizzate da una minore salinità e una maggiore disponibilità di cibo, sono più allungate e strette nelle forma e con margini dorsoventrali più paralleli, mentre conchiglie di aree più calde sono più leggere, sottili, ovali, porose e meno resistenti alle fratture. Malgrado queste differenze nella forma non ci sono variazioni nel polimorfo del CaCO_3 o nei parametri composizionali. Gli autori attribuiscono la maggiore fragilità osservata nelle aree più calde a cambiamenti nel bilancio energetico dei processi fisiologici, che implicherebbero di poter allocare una minore energia per la formazione della conchiglia, inoltre un incremento della radiazione solare potrebbe influenzare densità e distribuzione del fitoplancton, ovvero della disponibilità dei nutrienti per i bivalvi (Häder et al., 2007). Ci sono però altri studi che mostrano un trend contrario a quello ritrovato da Gizzi e colleghi (2016) e da Mancuso e colleghi (2019), in cui ad una bassa SST è associata una maggiore solubilità di CaCO_3 nell'acqua (Watson et al., 2012; Vermeji, 1993) e ciò incrementerebbe il costo della formazione della conchiglia (Clarke, 1993). Inoltre, è stato dimostrato che una bassa SST riduce crescita (Heilmayer et al., 2004) e sviluppo (Peck et al., 2007). Invece, per quanto riguarda l'influenza della salinità, una salinità sub-

ottimale è uno stressore che porta sia ad un aumento dei costi per le difese da stress ossidativo che a problemi nel feeding e quindi minori tassi di crescita. Inoltre, una bassa salinità riduce lo stato di saturazione dell'aragonite (unico polimorfo ritrovato nella conchiglia di *C. gallina*) con un aumento dei costi della calcificazione (Carregosa, 2014; Monari et al., 2007; Siedlecki et al., 2017; Sanders et al., 2018; Gruffydd et al., 1984; Navarro, 1988; Kristensen, 1956; Maeda-Martinez, 1987; Hiebenthal et al., 2011; Wanamaker et al., 2007; Marsden & Pilkington, 1995; Almada-Villela, 1984). Questi trend generali devono essere contestualizzati in una situazione in cui è presente l'influenza dei fiumi, che apportano modifiche alla salinità, disponibilità di nutrienti, presenza di metalli pesanti ed altri contaminanti. (Vernberg, 1976; Widdows, 1985; Marcus et al., 1989; Widdows et al. 1990; Roper et al., 1991).

1.4 Applicazioni del sequenziamento NGS nel campo della biologia della conservazione

Molte specie marine sono minacciate dal cambiamento ambientale, inquinamento, malattie microbiche, pesca eccessiva, ecc, che possono ridurre dimensioni e diversità genetica. Lo sviluppo di tecniche come la trascrittomica e la genomica basate sul sequenziamento di nuova generazione (NGS) forniscono un notevole contributo alla biologia della conservazione delle specie, incrementando significativamente l'abilità dei ricercatori di ottenere

informazioni sui meccanismi molecolari adottati dalle specie per fronteggiare i cambiamenti ambientali (Ghiselli et al., 2018). Le tecniche di sequenziamento ad alta resa (high-throughput), come l'RNA-Sequencing (RNA-Seq), stanno acquisendo un crescente riconoscimento nella biologia della conservazione, per il monitoraggio di specie sensibili ed economicamente importanti (Carducci et al., 2020; Eizaguirre & Baltazar-Soares, 2014; Mable, 2019; He et al., 2016). In particolare, l'RNA-Seq è la tecnica di sequenziamento favorita per ottenere informazioni sulla risposta funzionale degli organismi alle condizioni ambientali. Infatti, in un contesto con diversi stressori questa tecnica permette di ottenere un quadro simultaneo delle alterazioni di diversi geni e pathway molecolari, anche in organismi non modello, in risposta a condizioni di stress (Bertucci et al., 2017; Milan et al., 2018).

Gli approcci NGS permettono inoltre di ottenere una panoramica sulla variabilità genetica, che rappresenta una misura dell'abilità delle specie di adattarsi alle perturbazioni ambientali. Infatti, specie mostranti bassi livelli di diversità genetica sono maggiormente a rischio di estinzione.

Capitolo secondo

2. MATERIALI E METODI

Il campionamento degli esemplari di *C. gallina* non ha richiesto permessi specifici in quanto la specie oggetto di studio non rientra tra quelle protette o a rischio e i siti di campionamento non sono in aree protette o di proprietà privata.

2.1 Campionamento

I campionamenti sono avvenuti nella tarda primavera del 2020 in due siti localizzati lungo la costa adriatica del Centro Italia ed effettuati tramite draga idraulica sul fondale sabbioso della zona subtidale a 3-7 m di profondità. I due siti sono stati scelti a causa dell'importanza economica che rivestono per due specifici distretti di pesca marittima. Inoltre, sono localizzati a nord e a sud rispetto al Monte Conero, e ciò causa la loro appartenenza a due differenti settori biogeografici (Bianchi & Morri, 2000), che comportano esposizioni a condizioni ambientali diverse che potrebbero influenzare il processo di calcificazione. La presenza del promontorio, infatti, influenza gli schemi di circolazione delle acque e di conseguenza anche la diffusione delle acque di deflusso provenienti dal maggior fiume della costa adriatica settentrionale, il Po.

Il campionamento è stato effettuato a: Silvi Marina coordinate: latitudine 42° 48' 514'' N; e longitudine 13° 57' 419'' E.; Senigallia coordinate: latitudine 43° 37' 38'' N e longitudine 13° 26' 28'' E.

In entrambi i siti sono stati campionati trenta individui, di dimensioni comparabili riguardo la larghezza (da una estremità all'altra delle valve) e l'altezza (direzione umbo-ventrale). Per Silvi Marina (Tab. 1) la larghezza e l'altezza media sono rispettivamente di 27,5 mm ($\sigma=0,96$ mm) e 24,7 mm ($\sigma=0,63$ mm); per Senigallia (Tab. 2) sono di 26,3 mm ($\sigma=0,89$) e 23,6 mm ($\sigma=0,64$ mm). Dimensioni simili garantiscono un'età degli individui confrontabile, nel nostro caso intorno ai 2-3 anni (Mancuso et al., 2019; Richardson, 2001; Gaspar et al., 2004; Austad, 1996), ciò permette di escludere questa variabile dalle eventuali differenze tra gli individui dei due siti. Gli individui sono stati successivamente suddivisi in tre pool da dieci esemplari per ogni sito, ovvero SM1, SM2, SM3 per Silvi Marina e S1, S2, S3 per Senigallia.

		larghezza	umbo-ventrale
pool 1	1	27,00	24,20
	2	27,20	24,90
	3	26,00	24,20
	4	27,00	24,00
	5	28,20	25,20
	6	28,00	23,90
	7	28,60	24,20
	8	27,20	24,20
	9	26,60	24,00
	10	26,20	24,50
pool 2	11	27,60	24,90
	12	26,60	25,10
	13	27,90	24,90
	14	27,20	24,20
	15	28,90	24,00
	16	28,20	24,90
	17	25,90	23,20
	18	28,90	25,20
	19	28,00	26,20
	20	26,50	25,20
pool 3	21	27,20	25,20
	22	26,90	24,90
	23	28,90	25,20
	24	25,90	24,50
	25	25,90	24,20
	26	27,90	24,20
	27	29,00	24,90
	28	27,90	25,20
	29	28,30	25,90
	30	27,90	25,00

Tab. 1. Dimensioni delle valve delle vongole campionate a Silvi Marina riportate in mm.

		larghezza	umbo-ventrale
pool 1	1	25,30	23,20
	2	25,90	23,50
	3	25,90	23,90
	4	25,90	23,00
	5	26,20	24,00
	6	26,50	24,00
	7	28,50	24,50
	8	26,00	24,00
	9	26,50	24,20
	10	26,00	23,00
pool 2	11	24,20	22,70
	12	25,20	23,20
	13	26,50	24,00
	14	26,90	24,00
	15	26,50	24,20
	16	26,20	24,00
	17	27,00	23,50
	18	26,50	23,50
	19	26,00	23,50
	20	26,80	23,80
pool 3	21	24,90	22,90
	22	26,70	23,00
	23	27,20	24,20
	24	26,50	22,00
	25	28,00	22,50
	26	24,80	23,00
	27	26,20	23,20
	28	26,00	24,50
	29	26,90	24,50
	30	27,50	24,50

Tab. 2. Dimensioni delle valve delle vongole campionate a Senigallia riportate in mm.

2.2 La conservazione dei campioni

Ogni individuo (in totale sessanta, cioè un gruppo da trenta per ogni sito e poi diviso in tre pool da dieci) è stato conservato in RNAlater® (Invitrogen). Questo è un reagente acquoso, non tossico, che permette di inattivare le RNasi e di stabilizzare l'RNA cellulare nei campioni di tessuto (integri e non congelati), che per questo studio sono rappresentati dal mantello di entrambe le valve di ogni individuo. La soluzione evita la necessità di processare immediatamente i campioni o di congelarli in azoto liquido per una successiva elaborazione; pezzi di tessuto possono quindi essere raccolti ed immersi in soluzione RNAlater per lo stoccaggio, senza compromettere la qualità o la quantità di RNA che sarà ottenuta in seguito a un processo di purificazione. Le provette così allestite (mantello in soluzione di RNAlater) sono state tenute per 2 ore a 4 °C, al fine di equilibrare i tessuti nella soluzione attraverso la permeazione di quest'ultima nelle cellule, successivamente sono state trasferite a -80 °C per la conservazione. Di ciascun individuo processato sono stati conservati il corpo (privato come già detto del mantello) e le relative valve. Inoltre, ulteriori cinque pool da dieci individui, da tenere come riserva in caso di eventuali di problemi, sono stati congelati a -80 °C per la conservazione.

2.3 Il disegno sperimentale

In questo paragrafo saranno elencati, in maniera riassuntiva, i vari passaggi effettuati durante lo studio. Ciò permetterà di avere una panoramica generale del disegno sperimentale nel suo complesso; successivamente, i vari passaggi avranno dei paragrafi di approfondimento dedicati.

2.3.1 Le metodiche sperimentali

L'ottenimento dei dati necessari all'analisi dell'espressione genica del mantello ha inizialmente richiesto delle metodiche sperimentali, dalle quali si è ottenuto il materiale per le successive elaborazioni bioinformatiche dei dati. Per quanto riguarda le metodiche sperimentali, dal materiale conservato descritto nel paragrafo precedente si è proseguito con la purificazione dell'RNA totale dei singoli individui. Un'aliquota del volume finale risultata dal processo di purificazione è stata fatta correre su gel di agarosio per una prima conferma visiva della riuscita del processo di purificazione. Un'altra aliquota è stata invece usata per la quantizzazione al Qubit 4.0 fluorometer (ThermoFisher Scientific) al fine di ottenere la concentrazione di RNA (ng/ μ l) e tramite NanoDrop (ThermoFisher Scientific) per valutare degli indici di purezza.

2.3.2 Il sequenziamento

Per ogni pool è stata costituita una provetta contenente la stessa quantità di RNA proveniente da ogni esemplare. Da questa provetta rappresentatrice di

ogni pool sono poi state eseguite la quantificazione tramite Qubit 4.0 fluorometer (ThermoFisher Scientific) e la valutazione degli indici di purezza tramite NanoDrop (ThermoFisher Scientific), tutto ciò per assicurarsi di inviare dei campioni aventi i requisiti minimi richiesti dalla ditta di sequenziamento. Successivamente un'aliquota delle provette rappresentatrici è stata inviata alla CDgenomics (New York, USA; <https://www.cd-genomics.com/>) per il sequenziamento. Per poter procedere con il sequenziamento richiesto (Stranded mRNA-Seq), la CDgenomics ha nuovamente controllato la quantità e l'integrità dell'RNA, quest'ultima attraverso l'indice RIN e successivamente è stata costruita la libreria di frammenti idonei al sequenziamento su una piattaforma Illumina. È stato quindi eseguito il sequenziamento da cui si sono ottenute le reads grezze, questo materiale ci è stato poi inviato accompagnato da altre informazioni sul controllo qualità e sui risultati del sequenziamento.

2.3.3 Le analisi bioinformatiche

Per quanto riguarda le analisi bioinformatiche dei dati, i file contenenti le reads grezze per ogni pool sono stati processati tramite il software CLC Genomics Workbench (QIAGEN, <https://digitalinsights.qiagen.com/>), nello specifico è stato fatto un accoppiamento delle reads paired-end, un successivo controllo qualità, una rifinitura tramite il processo di trimming ed infine, per quanto riguarda l'elaborazione con il software CLC Genomics Workbench, un

assemblaggio *de novo* con le reads ottenute dal sequenziamento dei pool di Senigallia, uno con le reads ottenute dal sequenziamento dei pool di Silvi Marina ed uno totale considerando tutte le reads. Gli assemblaggi prodotti sono stati successivamente analizzati con il software BUSCO (<https://busco.ezlab.org/>), che ha permesso una valutazione degli assemblaggi in termini di completezza del contenuto genico. Per poter proseguire con la determinazione dei geni differenzialmente espressi tra gli individui appartenenti ai due siti di campionamento, i file con le reads rifinite (cioè sottoposte al processo di trimming) sono state importate sulla piattaforma Galaxy (<https://usegalaxy.org/>), insieme al trascrittoma di riferimento pubblicato da Coppe e colleghi (2012). Questo trascrittoma di riferimento è stato utilizzato per le successive analisi nonostante abbia una minore qualità rispetto all'assemblaggio totale ottenuto dai nostri dati (confronto tramite BUSCO), il motivo di questa scelta risiede nel fatto che i dati pubblicati da Coppe e colleghi (2012) sono corredati da un file di annotazione, indispensabile al proseguimento delle analisi. È stato quindi fatto un allineamento/mappaggio delle reads sul trascrittoma di riferimento con il software HISAT2 disponibile sulla piattaforma Galaxy. Gli output generati sono stati elaborati fuori dalla piattaforma tramite il programma HTSeq-count (<https://htseq.readthedocs.io>), che ha permesso di contare il numero di reads mappanti in un certo intervallo

di riferimento. I file delle conte sono stati importati in Galaxy per poter ottenere, tramite il software DESeq2 della piattaforma, i dati sui geni differenzialmente espressi; questi sono stati riordinati e selezionati in base all'indice p-adjusted (un indicatore di significatività) e poi divisi in up- e down-regolati in base all'indice \log_2 fold change che esprime il rapporto tra l'espressione di un gene tra i due set di dati, in questo caso tra quelli dei due siti di campionamento. Utilizzando l'annotazione dei geni differenzialmente espressi è stato possibile risalire alle sequenze nucleotidiche, e da queste alle sequenze amminoacidiche per i sei frame. Di queste sequenze amminoacidiche è stata fatta l'annotazione funzionale tramite la piattaforma OmicsBox (BioBam, <https://www.biobam.com/omicsbox/>), utilizzando il modulo Blast2GO. Ciò ha permesso di ottenere una panoramica sulla funzione biologica e sulle categorie funzionali dei geni differenzialmente espressi. Infine, i dati ottenuti sono stati discussi alla luce dei parametri ambientali e delle osservazioni fatte sulle valve, alla ricerca di una eventuale relazione tra le differenze nell'espressione genica e quelle fenotipiche riguardanti la biomineralizzazione degli individui dei due siti.

2.3.4 SEM, microanalisi e diffrazione a raggi X

Parallelamente, sono state eseguite la diffrazione a raggi X e la microanalisi sulle polveri ottenute dalla frammentazione di una delle due valve di un

individuo per entrambi i siti, per rilevare eventuali differenze nel polimorfo del CaCO_3 e negli elementi presenti nel guscio. Inoltre, utilizzando un microscopio elettronico a scansione (SEM), sono state acquisite immagini da sezioni umbono-ventrali di valve appartenenti ad individui di entrambi i siti, per poter fare un confronto visivo della struttura e dell'organizzazione degli strati del guscio. Infine, durante le acquisizioni, in alcune aree specifiche delle sezioni è stata effettuata la microanalisi.

2.4 Purificazione dell'RNA totale

Il protocollo per l'estrazione dell'RNA ha previsto le seguenti fasi, eseguite per ogni individuo:

1. Omogeneizzazione

A 25-50 mg di tessuto vengono aggiunti 500 μl di TRIzol® (Invitrogen), cioè una soluzione monofasica di fenolo, guanidina isotiocianato e altri componenti registrati che facilitano l'isolamento di vari tipi di RNA di dimensioni più o meno elevate. Il reagente TRIzol® preserva l'integrità dell'RNA grazie all'elevata efficacia di inibizione nei confronti delle RNasi, inoltre disgrega le cellule e dissolve i componenti cellulari durante l'omogeneizzazione del campione, che viene effettuata impiegando dei potter di plastica trattati con acqua DEPC (dietilpirocarbonato), in ghiaccio.

2. Separazione di fase

Aggiunta di 0.1 ml di cloroformio alla soluzione precedente e, dopo essersi assicurati di aver chiuso saldamente la provetta, si procede agitando vigorosamente per 15 secondi. Segue un'incubazione di tre minuti a 15-30 °C e una fase di centrifugazione a 12,000 x g per 15 minuti ad una temperatura compresa fra i 2-8 °C. L'RNA andrà a localizzarsi nella fase acquosa.

3. Precipitazione

Trasferita la fase acquosa in una nuova provetta trattata con acqua DEPC si ha l'aggiunta di 250µl di alcool isopropilico, che provoca la precipitazione dell'acido nucleico. Segue un'incubazione di 10 minuti a 15-30 °C e una centrifugazione a 12,000 x g per 15 minuti ad una temperatura di 2-8 °C. L'RNA precipiterà andando a costituire il pellet che ha una consistenza gelatinosa separandosi dalla fase acquosa. A questo punto è possibile rimuovere il sovrantante, mantenendo in provetta solo l'acido nucleico.

4. Lavaggio

La fase di lavaggio avviene con l'aggiunta di 500 µl di etanolo al 75% trattato con acqua DEPC. La soluzione viene mescolata con l'aiuto di un vortex; segue uno step di centrifugazione a 7,500 x g per cinque minuti a 2-8 °C.

5. Risospensione

Si lascia asciugare l'acido nucleico all'aria per 30 minuti sotto cappa. Si procede con l'aggiunta di 30 μ l di acqua trattata con DEPC. La conservazione avviene a -80 °C.

2.5 Elettroforesi su gel di agarosio

Per una conferma visiva dell'integrità del materiale purificato è stata eseguita una corsa elettroforetica su gel di agarosio (1,5%), su cui vengono caricati 2 μ l di RNA estratto (esempio in Fig. 5). Un campione di RNA di buona qualità deve porre in risalto uno striscio in cui siano ben visibili le bande appartenenti ai due RNA ribosomiali 28S e 18S. Nel caso dei bivalvi, il pattern di migrazione dell'RNA ha delle differenze rispetto a quello degli altri organismi, e potrebbe non mostrare un picco in corrispondenza del 28S, come si vedrà meglio dai report sulla valutazione dell'indice RIN eseguita dalla ditta di sequenziamento.

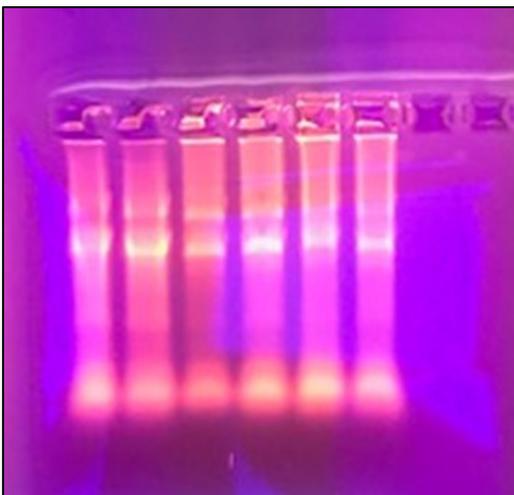


Fig. 5. Risultato corsa con parte dei campioni.

2.6 Quantificazione dell'acido nucleico e valutazione degli indici di purezza

2.6.1 La quantificazione dell'RNA

La quantificazione dell'RNA totale è stata effettuata per ognuna delle sessanta provette, ed in seguito anche sulle sei provette rappresentative di ogni pool (la costituzione delle sei provette ha un paragrafo dedicato), ottenendo la concentrazione di RNA espressa in ng/ μ l (Fig. 6, 7). Per fare questo è stato utilizzato il Qubit 4.0 fluorometer (ThermoFisher Scientific), uno strumento che utilizza molecole fluorescenti emettenti un segnale solo quando sono legate al target, ciò minimizza l'effetto dei contaminanti sul risultato. Infatti, rispetto ad una valutazione tramite UV, l'assorbanza a 260 nm di DNA, proteine, nucleotidi liberi e sali non influenza il risultato, essendo le molecole utilizzate specifiche per il target selezionato. Il Qubit permette di utilizzare microvolumi di campione, fino a 1 μ l, garantendo alti livelli di accuratezza. La quantificazione avviene attraverso la lettura preventiva di due soluzioni standard, che permette di ottenere una retta di taratura sulla quale, in base alla fluorescenza del nostro campione, si potrà risalire alla sua concentrazione. Per le due soluzioni standard, che vengono preparate prima del saggio con i reagenti fornite dalla ditta, si uniscono 190 μ l di buffer a 10 μ l di soluzione standard 1 e si fa lo stesso per la soluzione standard 2. Per il campione da quantificare, si

uniscono 198 μ l di buffer a 2 μ l di campione. Le tre provette sono miscelate tramite vortex e lasciate a temperatura ambiente per 3 minuti, avviene poi la lettura dei due standard e infine del campione seguendo le indicazioni dello strumento. Le concentrazioni permettono di capire quanti microlitri di campione si devono prelevare per ottenere una precisa quantità di RNA in microgrammi, necessaria allo step di sequenziamento, richiesta della CDgenomics.

2.6.2 La valutazione degli indici di purezza

La valutazione degli indici di purezza è stata effettuata sulle sei provette rappresentative dei pool. Il calcolo è stato eseguito tramite NanoDrop (ThermoFisher Scientific), uno spettrofotometro UV-Visibile capace di lavorare con microvolumi di campione, nel nostro caso sono stati utilizzati 2 μ l. Lo strumento sfrutta una tecnologia basata sulla tensione superficiale che piccoli volumi di liquidi esercitano quando si trovano collocati tra due superfici vicine. In tal modo una goccia di campione posizionata sull'apposita piastra di lettura crea una colonna di liquido a diretto contatto con due fibre ottiche, e può essere analizzata in modo semplice e veloce. Il NanoDrop permette di determinare automaticamente la lettura del campione a lunghezze d'onda di 260-280-230 nm, la sua concentrazione ed i relativi rapporti per determinare il grado di purezza del campione analizzato. Per quanto riguarda gli indici di

purezza, l'assorbanza viene registrata nei rapporti 260/280 e 260/230, i valori ottenuti permettono di valutare la purezza dell'RNA estratto. L'intervallo ottimale è compreso tra 1,8 e 2. Gli indici forniscono informazioni sulla purezza dell'acido nucleico rispetto ad eventuali residui proteici, fenolici, guanidinici o di altri reagenti rimasti nel campione dopo il processo di purificazione, che assorbono a 280 nm e 230 nm e quindi abbassano il rispettivo rapporto. Per quanto riguarda i nostri campioni (Fig. 6, 7), tutti gli indici 260/280 sono risultati superiori a 1,89; invece gli indici 260/230 si sono posizionati nell'intervallo 0,75-1,11 indicando possibili contaminazioni da solventi organici che però non hanno richiesto ulteriori passaggi. Infatti, come sarà indicato in seguito, la CDgenomics ha nuovamente valutato la qualità dei campioni prima del sequenziamento e non ci sono stati avvisi riguardo la non idoneità dei campioni.

		ng/ul	260/280	260/230	ul per 4ug	Vtot_pool
pool 1	1	1540,0	1,94	1,11	2,60	27,12
	2	1190,0			3,36	
	3	1340,0			2,99	
	4	1220,0			3,28	
	5	1610,0			2,48	
	6	1550,0			2,58	
	7	1580,0			2,53	
	8	1750,0			2,29	
	9	1420,0			2,82	
	10	1820,0			2,20	
pool 2	11	1340,0	1,93	0,94	2,99	27,81
	12	1640,0			2,44	
	13	1540,0			2,60	
	14	1340,0			2,99	
	15	1840,0			2,17	
	16	1660,0			2,41	
	17	1240,0			3,23	
	18	1410,0			2,84	
	19	1220,0			3,28	
	20	1390,0			2,88	
pool 3	21	1550,0	1,92	0,84	2,58	29,09
	22	2760,0			1,45	
	23	1780,0			2,25	
	24	1260,0			3,17	
	25	1260,0			3,17	
	26	1070,0			3,74	
	27	1360,0			2,94	
	28	1410,0			2,84	
	29	1570,0			2,55	
	30	909,0			4,40	
Silvi Marina						
	pool1	pool2	pool3	per ogni pool sono stati inviati 20 ul		
nanodrop	606,7 ng/ul	714,6 ng/ul	796,6 ng/ul			
qubit 4.0 fluorometer	1360 ng/ul	1380 ng/ul	1430 ng/ul			

Fig. 6. Concentrazioni e indici per i campioni di Silvi Marina.

		ng/ul	260/280	260/230	ul per 4ug	Vtot_pool
pool 1	1	1130,0	1,91	0,9	3,54	33,56
	2	1120,0			3,57	
	3	1200,0			3,33	
	4	989,0			4,04	
	5	1200,0			3,33	
	6	1300,0			3,08	
	7	1470,0			2,72	
	8	1320,0			3,03	
	9	1290,0			3,10	
	10	1050,0			3,81	
pool 2	11	1220,0	1,92	1,00	3,28	28,03
	12	1460,0			2,74	
	13	1730,0			2,31	
	14	1480,0			2,70	
	15	1410,0			2,84	
	16	1490,0			2,68	
	17	1280,0			3,13	
	18	1290,0			3,10	
	19	1370,0			2,92	
	20	1720,0			2,33	
pool 3	21	1500,0	1,89	0,75	2,67	25,05
	22	1860,0			2,15	
	23	1390,0			2,88	
	24	1820,0			2,20	
	25	1740,0			2,30	
	26	1580,0			2,53	
	27	1220,0			3,28	
	28	1870,0			2,14	
	29	1760,0			2,27	
	30	1520,0			2,63	
Senigallia						
	pool1	pool2	pool3	per ogni pool sono stati inviati 20 ul		
nanodrop	650,8 ng/ul	736,3 ng/ul	738 ng/ul			
qubit 4.0 fluorometer	1270 ng/ul	1530 ng/ul	1370 ng/ul			

Fig. 7. Concentrazioni e indici per i campioni di Senigallia.

2.7 Costituzione dei pool

Come già accennato, per l'invio del materiale alla ditta di sequenziamento sono state allestite sei provette, una per ogni pool, prelevando dalle dieci provette di ogni pool una quantità di RNA pari a 4 µg (Fig. 6, 7). Le sei provette sono state

marcate con le sigle con cui sono stati indicati i pool, ovvero SM1, SM2, SM3 e S1, S2, S3. Il motivo di tutto ciò risiede nei requisiti che deve avere il materiale per poter essere sequenziato, indicati dalla ditta. Nello specifico, per quanto riguarda il protocollo Stranded mRNA-Seq, per un'applicazione *de novo* viene consigliato l'invio di minimo 4 µg di RNA, ad una concentrazione minima di 40 ng/µl e in un volume minimo di 20 µl. Successivamente, come spiegato nel paragrafo precedente, di queste sei provette è stata calcolata la concentrazione tramite Qubit e gli indici di purezza tramite NanoDrop (Fig. 6, 7). Infine, sono stati prelevati 20 µl e inviati per il sequenziamento.

2.8 Sequenziamento NGS

2.8.1 Le valutazioni preliminari

Il sequenziamento del materiale è stato affidato ad una ditta esterna, la CDgenomics (New York, USA; <https://www.cd-genomics.com/>). All'arrivo dei campioni, la ditta ha provveduto alla quantizzazione del materiale tramite metodo Ribogreen** Victor X2 fluorometry, e alla valutazione dell'integrità dell'RNA tramite Agilent RNA Screen Tape. La quantizzazione effettuata dalla ditta ha indicato concentrazioni minori delle nostre, dovute all'utilizzo di un diverso metodo di quantizzazione, ma comunque ampiamente sufficienti; anche i volumi sono risultati leggermente diversi da quelli prelevati con i nostri strumenti, probabilmente per motivi di taratura. L'integrità del materiale è stata

espressa attraverso l'indice RIN (RNA integrity number). Questo numero è calcolato da un software che valuta il risultato di una corsa elettroforetica di una aliquota del materiale; il software costruisce un grafico associando l'intensità della fluorescenza emessa dagli RNA a fine corsa (asse Y) con la posizione relativa dei punti di fluorescenza lungo il gel (asse X), poi calcola l'area sotto la curva. I grafici che corrispondono a numeri RIN alti, cioè a materiali di ottima qualità, mostrano dei picchi in corrispondenza delle bande 18S e 28S, mentre sono molto piatti nel resto della curva, essendoci poche copie di RNA per tipo. Nel caso dei bivalvi, come già riportato, il pattern di migrazione dell'RNA è diverso rispetto agli altri organismi e non mostra un picco in corrispondenza del 28S (Fig. 8). Per un sequenziamento efficace è richiesto un RIN > 7, tutti i nostri campioni sono risultati idonei e con ottimi punteggi, infatti il minimo è stato 8,1 (SM3) mentre gli altri hanno valori maggiori o uguali a 9,2 (Fig. 9).

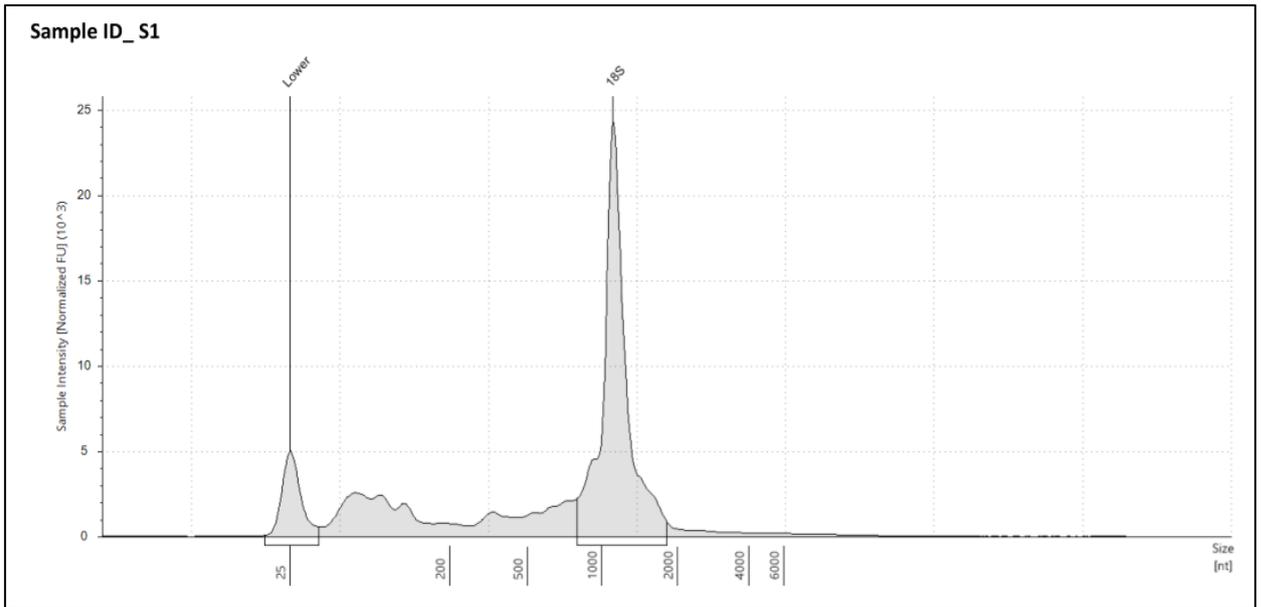


Fig. 8. Grafico RIN di uno dei campioni inviati per il sequenziamento.

3. QC Results summary						
Sample Name	Concentration (ng/ul)	Volume (ul)	Amount(ug)	RIN	Results	Reason
S1	333.34	23	7.67	9.5	Hold	
S2	464.52	20	9.29	9.2	Hold	
S3	391.9	21	8.23	9.5	Hold	
SM1	330.55	25	8.26	9.4	Hold	
SM2	394.49	24	9.47	9.2	Hold	
SM3	452.33	23	10.4	8.1	Hold	

Fig. 9. Punteggi RIN dei campioni inviati per il sequenziamento.

2.8.2 La preparazione della libreria

Lo step successivo ha visto la preparazione della libreria di frammenti per un sequenziamento su piattaforma Illumina. Il kit usato dalla CDgenomics è il TruSeq Stranded mRNA Sample Prep Kit (Illumina, San Diego, CA), per il

quale sono state seguite le istruzioni del fornitore. Il kit è specifico per mRNA con code di poly-A, infatti utilizza particelle sulle quali sono attaccati oligo-dT che catturano i trascritti ibridandosi con le loro code di poly-A, questo kit permette inoltre di mantenere l'informazione sulla direzionalità dei trascritti. Ai frammenti costituenti la libreria vengono aggiunti esternamente degli adattatori, divisi a livello funzionale in una regione (indicata come A1 e A2) necessaria al legame con gli oligonucleotidi presenti sulla superficie della cella a flusso, indispensabile per la generazione dei cluster; e in una regione (indicata come SP1 e SP2) adiacente alla sequenza da leggere e che serve al sequenziamento vero e proprio, è infatti qui che si appaieranno i primer. Ai frammenti vengono inoltre aggiunte sequenze specifiche per attribuire ogni frammento al campione di appartenenza, necessarie nei sequenziamenti multiplex.

2.8.3 Il sequenziamento

Preparata la libreria di frammenti, si procede al sequenziamento su piattaforma Illumina - NovaSeq - S1 - PE 150 Cycle. Il sequenziamento richiesto è paired-end, con una lunghezza delle reads di 151 bp, ovvero ogni frammento verrà letto nelle due direzioni e per una lunghezza di 151 bp. La tecnologia Illumina SBS (Sequencing By Synthesis) utilizza un metodo basato su terminatore fluorescente reversibile che permette di rilevare le singole basi nel momento in

cui sono incorporare nel filamento di DNA. Il risultato è un accurato sequenziamento base per base che minimizza gli errori di sequenza contesto specifici, soprattutto in regioni ripetute e omopolimeri. Il processo chimico del sequenziamento avviene in un supporto denominato cella a flusso, ovvero un vetrino dotato di canali o corsie; la tipologia di cella utilizzata dalla Novaseq è detta preconfigurata, ovvero ha dimensioni e spaziature definite per i cluster. La prima fase del sequenziamento è la generazione dei cluster, processo in cui ciascun frammento della libreria viene clonato in migliaia di copie identiche: inizialmente i frammenti si ibridano sulla cella a flusso grazie agli oligonucleotidi presenti su di essa e complementari alle regioni esterne degli adattatori dei frammenti, poi si eseguono vari cicli di Bridge PCR in cui gli oligonucleotidi presenti sulla cella vengono utilizzati come primer per la sintesi dei filamenti complementari a quelli ibridati, i quali serviranno a loro volta da stampo per una nuova sintesi e così via, ottenendo infine un insieme di filamenti identici raggruppati tra loro e pronti al sequenziamento; tutto ciò è funzionale all'amplificazione del segnale che verrà emesso dai nucleotidi terminatori. I reagenti per il sequenziamento, inclusi i nucleotidi marcati tramite fluorescenza, vengono aggiunti alla miscela di reazione sulla cella e, quando quest'ultima sarà sottoposta a imaging durante ogni step, ciascuno dei marcatori emetterà una lunghezza d'onda univoca che lo renderà identificabile.

2.8.4 I risultati del sequenziamento

Dal report sul sequenziamento fornito dalla ditta (Fig. 10) si può osservare, oltre agli altri dati che verranno presentati nel capitolo dedicato, una delle informazioni più importanti, ovvero il Phred Score, che rappresenta un indice della qualità e accuratezza delle basi rilevate (processo di “base calling” chiarito a breve) e quindi della qualità del sequenziamento; dai nostri dati si osserva che più del 97% delle reads di ogni campione ha uno score maggiore o uguale a Q20 cioè statisticamente contengono 1 base errata (o meno) ogni 100 lette, tra queste più del 93% ha uno score maggiore o uguale a Q30 ovvero 1 base errata (o meno) ogni 1000 lette. I segnali sulla fluorescenza ottenuti dalla piattaforma di sequenziamento sono convertiti in sequenze attraverso il base calling, in cui si associano le singole basi ai rispettivi segnali registrati (in questo caso alla lunghezza d’onda del marcatore). Le sequenze ottenute sono registrate in un file FASTQ, dove le informazioni di ogni read sono suddivise in 4 righe: Linea 1 contiene le informazioni sullo strumento, corsa, cella di flusso, corsia e pozzetto che permettono l’identificazione univoca del cluster da cui si è ottenuta la read, insieme al suo numero nella coppia paired e alla sequenza indice; Linea 2 ha l’effettiva sequenza in basi azotate della read; Linea 3 non contiene informazioni ed è spesso occupata da un “+”; Linea 4 contiene le informazioni qualitative (Phred Score) base per base codificate in

codice ASCII, ogni carattere di questa riga è da associare alla base che si trova nella posizione corrispondente nella Linea 2.

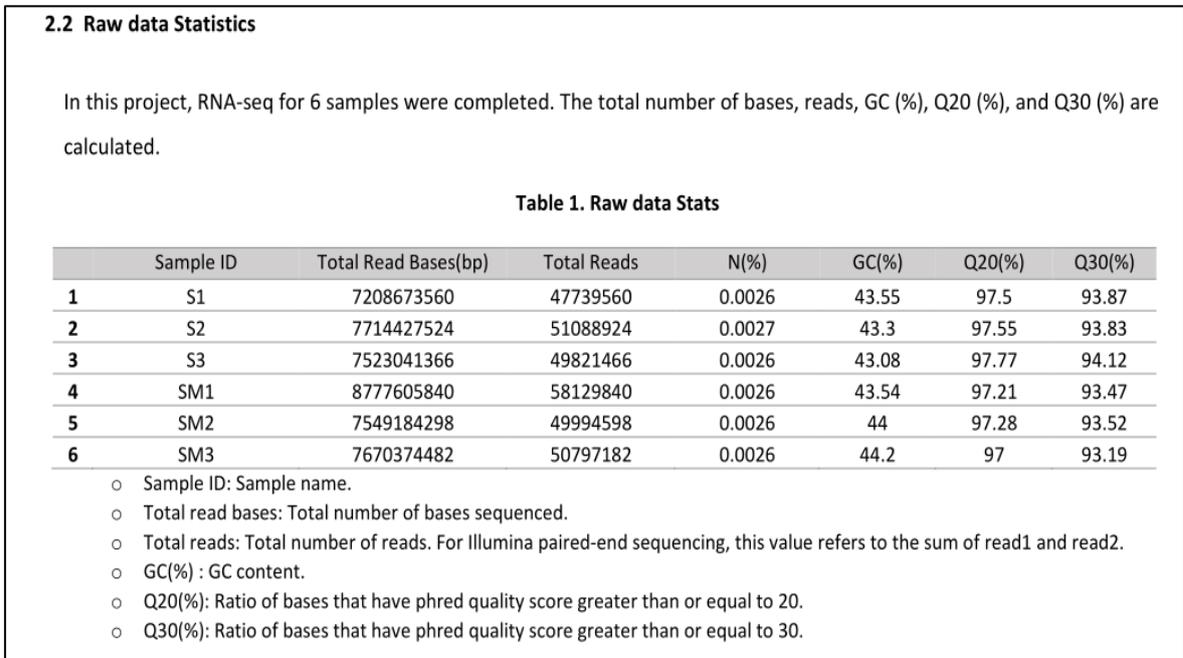


Fig. 10. Risultati del sequenziamento.

2.9 Le analisi bioinformatiche dei dati

In questi paragrafi saranno descritti in maniera approfondita le metodiche bioinformatiche utilizzate in questa tesi. Per i risultati di tali analisi si rimanda al capitolo dedicato.

2.9.1 Valutazione della qualità delle reads

I dati ottenuti dal sequenziamento sono stati inizialmente processati tramite il software CLC Genomics Workbench (QIAGEN, <https://digitalinsights.qiagen.com/>). I dodici file FASTQ (una coppia paired-

end per ogni pool) sono stati caricati nel software, il quale ha automaticamente provveduto all'accoppiamento dei due file appartenenti alla stessa coppia paired-end. L'accoppiamento ha prodotto quindi sei file, tre per ogni sito, che mantengono l'indicazione S1, S2, S3 per Senigallia e SM1, SM2, SM3 per Silvi Marina. Ottenuti questi sei file si è potuto procedere al controllo qualità tramite lo strumento "*Sequencing data quality control*" messo a disposizione dalla piattaforma e specifico per dati trascrittomici. Gli output del controllo qualità consistono in tre report: uno grafico, uno sulle sequenze duplicate ed uno con informazioni supplementari. Nel report grafico si trova una piccola tabella riepilogativa con le sequenze totali ed i nucleotidi complessivi del file sottoposto al controllo. La prime statistiche presenti nel report grafico hanno come ordinata la percentuale delle sequenze, associata ad una specifica caratteristica in ascissa, diversa per ogni grafico prodotto; In particolare, un grafico riguarda la distribuzione delle lunghezze, un grafico il contenuto percentuale di GC, che in assenza di problemi di contaminazione deve presentare una distribuzione (% sequenze su % contenuto GC) gaussiana, un grafico tiene conto della percentuale delle basi ambigue presenti (indicate nelle reads con "N"), ed, infine, un grafico riguarda la distribuzione della qualità in base al Phred Score. La seconda parte delle statistiche è invece incentrata sulla posizione di determinate caratteristiche lungo le sequenze ed in particolare

venie prodotto un grafico della coverage, indicante la percentuale di sequenze che coprono le singole posizioni da 1 a 151, un grafico del contributo nucleotidico, che indica la partecipazione di ogni base azotata lungo le varie posizioni delle reads (da queste informazioni si possono individuare eventuali contaminazioni, infatti in una situazione normale i contributi delle basi dovrebbero essere abbastanza simili e mostrare un andamento costante), un grafico con il contenuto GC nelle varie posizioni, anche qui in condizioni normali si dovrebbe avere un andamento costante, un grafico sulle basi ambigue lungo le posizioni, in cui si possono vedere eventuali errori di sequenziamento che ricadono nelle stesse posizioni, cioè avvenuti negli stessi cicli; un grafico sulla distribuzione della qualità (Phred Score) lungo le posizioni, per vedere le porzioni più affidabili delle reads, le porzioni di bassa qualità ed esterne verranno probabilmente tagliate nel processo di trimming. La terza parte delle statistiche riguarda i livelli di duplicazione presenti nelle reads ed è presente un grafico sull'arricchimento dei 5mers lungo le posizioni, ovvero delle sequenze di 5 nt over-rappresentate che a causa della loro piccola dimensioni sono trattate a parte rispetto a quelle duplicate, un grafico sui livelli di duplicazione delle sequenze, per sequenze duplicate si intendono sequenze di una certa lunghezza che si ripetono varie volte, nel grafico vediamo quante sequenze (asse Y) si ripetono un certo numero di volte (asse X), questo grafico

è collegato sia al file di report sulle sequenze duplicate, in cui vengono elencate tutte quelle trovate (con la loro sequenza in basi azotate), sia a quello con le informazioni supplementari, in cui sono evidenziate le 25 sequenze maggiormente ripetute.

2.9.2 Trimming dei dati

I sei file con le reads grezze, dopo la valutazione qualitativa, sono stati sottoposti al processo di *trimming*, ovvero una procedura che scarta le reads con punteggi di bassa qualità, rimuove porzioni duplicate appartenenti agli adattatori ed elimina porzioni delle reads con alta concentrazione di basi ambigue o a bassa qualità. Questa procedura è fondamentale per proseguire con l'assemblaggio e le successive analisi, poiché permette di ottenere un materiale di qualità superiore e più facile da assemblare, avendo eliminato le porzioni duplicate e ambigue. Per il trimming è stato usato lo strumento “*Trimming*“, messo a disposizione dalla piattaforma CLC Genomics Workbench.

L'output, oltre ai file contenenti le reads rifinite, consiste in un singolo file di report (per ogni file sottoposto al processo) riportante una tabella riepilogativa che mostra il numero e la lunghezza media delle reads prima e dopo il processo, per fornire rispettivamente un'idea delle reads completamente scartate e della lunghezza media delle porzioni rimosse. Si ha poi una tabella in cui le informazioni sui tagli effettuati sono divise in base alla tipologia dei tagli,

ovvero quelli sulle porzioni di bassa qualità e quelli sulle porzioni ambigue; a questa segue una tabella sulle statistiche dei tagli che hanno riguardato le sequenze degli adattatori, in cui si può vedere l'effettiva sequenza individuata e rimossa. Infine, il report contiene un istogramma che mette in relazione il numero di nucleotidi rimossi (asse X, da non confondere con la posizione delle basi rimosse) con il numero di sequenze (asse Y) a cui è stata rimossa quella quantità di basi.

Per avere un confronto diretto della qualità delle reads prima e dopo il processo di trimming, è stata fatta una prova ripetendo il controllo qualità per uno dei sei file, questa volta trimmed, affiancando poi i risultati nelle due condizioni (grezza e trimmed) per valutare l'impatto del trimming sulla qualità generale del nostro materiale.

2.9.3 Assemblaggio delle reads filtrate

Con i sei file rifiniti sono stati eseguiti tre assemblaggi *de novo*: uno per ogni sito di campionamento, utilizzando il materiale corrispondente come input (ad esempio, nell'assemblaggio per i campioni del sito di Senigallia sono stati utilizzati i file trimmed S1, S2, S3) e uno totale in cui sono stati utilizzati tutti e sei i file. Anche questa operazione è stata eseguita grazie allo strumento “*De novo assembly*” disponibile sulla piattaforma CLC Genomics Workbench. Nell'assemblaggio, le reads sono usate per formare sequenze più lunghe dette

contig (che in questo caso dovrebbero equivalere ai trascritti completi), ciò è possibile sfruttando le varie sovrapposizioni di sequenza dei frammenti che fanno parte dello stesso trascritto. L'output, oltre ai file contenenti gli assemblaggi, consiste in un report (per ogni assemblaggio) in cui troviamo una tabella con i dati assoluti e percentuali sul contributo nucleotidico e sulle basi ambigue e due tabelle per le misurazioni dei contig, con o senza l'inclusione degli scaffold nelle misure (ovvero di regioni gap stimate tra due contig e riempite con basi ambigue). In queste due tabelle troviamo informazioni specifiche degli assemblaggi come ad esempio le misure dei contig N75, N50 e N25; ad esempio l'N50 (di cui è riportata la lunghezza) è il contig più corto tra i contig più lunghi che sommati insieme coprono il 50% della lunghezza totale in basi dell'assemblaggio, ovvero i contig vengono sommati in ordine decrescente di lunghezza (cioè partendo dal più lungo) fino a coprire il 50% delle basi totali, arrivando all'N50 che sarà l'ultimo di questa lista. Dalla lunghezza dell'N50 si può capire che almeno la metà dei nucleotidi dell'assemblaggio appartengono a contig di lunghezza pari o superiore all'N50, perché questi ultimi sono stati sommati in ordine decrescente, e quindi ci si può avere un'idea della compattezza dell'assemblaggio. Abbiamo poi le lunghezze del contig più corto, più lungo, la lunghezza media, il numero dei contig e la

lunghezza totale dell'assemblaggio. Infine, un grafico dell'output mostra l'accumulo di lunghezza totale in base al numero dei contig considerati.

2.9.4 Valutazione della qualità dell'assemblaggio

I tre assemblaggi prodotti (Senigallia, Silvi Marina e il totale) sono stati esportati dalla piattaforma CLC Genomics Workbench per poter essere valutati tramite il software BUSCO (<https://busco.ezlab.org/>), la cui sigla sta per Benchmarking Universal Single-Copy Orthologs ovvero analisi comparativa di geni ortologi universali in singola copia. È usato per valutare gli assemblaggi in base al loro contenuto genico, i geni ortologi sono geni omologhi codificanti per la stessa proteina in specie diverse, che si sono originati da un comune evento di speciazione. Questa analisi è stata eseguita anche per il trascrittoma di riferimento di Coppe e colleghi (2012) per poter fare un confronto tra i nostri dati e quelli attualmente disponibili in letteratura. Il software contiene database di specie diverse e permette di selezionare il taxon di interesse all'interno del quale trovare dei match con i contig da analizzare e nel nostro caso è stato selezionato il phylum Mollusca. In base ai match viene stabilita la proporzione tra contig completi, frammentati, duplicati o mancanti rispetto ai gruppi di ortologi presenti nel database del taxon scelto. Un match completo, ad esempio, è un contig contenente buona parte dell'informazione genica di uno dei gruppi ortologi del database. Nel nostro caso, i gruppi

presenti nel database del phylum al momento dell'analisi erano 5295, e le percentuali dei vari tipi di match del nostro assemblaggio sono state calcolate in base a questo numero. Infatti, la porzione "missing" non si riferisce ai contig del nostro assemblaggio non matchati, ma ai gruppi ortologi presenti nel database e che non hanno avuto riscontro nell'assemblaggio, e dato che questi ortologi sono universali (condivisi da più del 90% delle specie del taxon selezionato) dalla loro assenza si possono ottenere informazioni sulla completezza degli assemblaggi. Non sono quindi presenti nei database i geni specie-specifici, cioè appartenenti ad una sola specie e neanche quelli che non rispettano la soglia del 90%. I risultati della valutazione sono stati convertiti in formato grafico, per una considerazione più immediata. Il risultato grafico mostra le percentuali dei vari tipi di match, che possono essere: completi, se i match hanno ottenuto un punteggio idoneo riguardo la qualità e la lunghezza dell'allineamento; frammentati, se hanno ottenuto un punteggio idoneo di qualità ma non di lunghezza, cioè se sono presenti trascritti incompleti; missing, per i gruppi di ortologi di cui non è stato trovato riscontro nei nostri dati, o se è stato trovato non ha ottenuto un punteggio di qualità idoneo.

2.9.5 Determinazione dei geni differenzialmente espressi (DEGs)

Partendo dai file trimmed, che sono stati esportati dal CLC Genomics Workbench con le coppie paired non più accorpate, e dal trascrittoma di

riferimento di Coppe e colleghi (2012), che nonostante avesse una minore qualità rispetto al nostro assemblaggio totale (confronto tramite BUSCO) è stato utilizzato perché comprensivo di annotazione, si è proseguito con la determinazione dei geni differenzialmente espressi tra gli individui appartenenti ai due siti di campionamento. I file sopra elencati (trimmed e riferimento) sono stati quindi importati nella piattaforma Galaxy (<https://usegalaxy.org/>), che permette l'utilizzo di software di analisi direttamente online. Nei file trimmed è contenuta l'informazione sui livelli di espressione, infatti considerando un gene altamente espresso si ritroveranno varie copie della sua sequenza tra le reads. Il primo step prevede un allineamento/mappaggio delle trimmed reads sul trascrittoma di riferimento con il software HISAT2 disponibile sulla piattaforma Galaxy; per ogni ciclo di analisi, HISAT2 ha preso come input i due file paired appartenenti allo stesso campione e il riferimento sul quale devono essere allineati e ha quindi allineato le singole reads lungo i contig del trascrittoma salvando per ogni coppia paired, in un solo file, le coordinate di inizio e fine rispetto al riferimento. Gli output generati (sei file in totale), contenenti le coordinate delle reads sul trascrittoma, permettono di valutare la presenza di determinate reads nelle stesse posizioni, e sono quindi stati elaborati all'esterno di Galaxy (mediante l'esecuzione di comandi da terminale linux) tramite il programma HTSeq-count

(<https://htseq.readthedocs.io>), che ha permesso la conversione delle coordinate di allineamento in un file di conte. Come input è stato inserito anche il file di annotazione del trascrittoma di riferimento, cioè con i codici identificativi delle porzioni del trascrittoma in cui c'è una ORF, e ciò è necessario per il mantenimento del corretto *gene_id* e ciò permetterà di proseguire con le analisi senza ambiguità. Questi output (conte S1, S2, S3, SM1, SM2, SM3) sono stati importati in Galaxy per poter sfruttare il software DESeq2 della piattaforma, in modo da ottenere i dati sui geni differenzialmente espressi. Forniti in input al software i file delle conte relativi ai due gruppi di dati, ovvero selezionati i tre i file di conte per Silvi Marina, e i tre per Senigallia, il programma DESeq2 confronterà tra gli identificativi in comune le rispettive conte, per poi salvare i dati sulla differenza di espressione corredata di diversi indici. Tra gli output si ottiene un MA-plot con ordinata \log_2 fold change (ratio tra i due livelli di espressione) e in ascissa la media delle conte normalizzate in entrambi i siti; in base alla distanza del \log_2 fold change dalla linea di threshold, si può capire la quantità geni up- o down-regolati, di cui quelli colorati in rosso ritenuti statisticamente significativi in base al p-adjusted. Un grafico delle componenti principali (PCA) mostra le sei conte distribuite in un piano 2D, dove gli assi sono le due componenti che contengono la maggior parte della variabilità dei dati, ogni componente principale infatti deriva dalla combinazione lineare delle

variabili (in questo caso i geni) e quindi racchiude il peso che ogni gene ha sulla differenziazione e distribuzione dei file di conte tra loro; per l'ordine con cui sono calcolate le componenti principali, la PC1 è la più importante e quella che rappresenta meglio la variabilità dei dati. La matrice di distanza presente è un'altra analisi di cluster, che rielabora i dati ottenuti dalle conte per creare una heatmap in cui la scala cromatica fornisce un'idea della somiglianza tra tutte le possibili coppie di conte, e un dendrogramma che raggruppa quest'ultime in ordine di somiglianza. Un grafico di dispersione distribuisce le stime sulla differenza di espressione dei geni sulla base delle conte normalizzate, in questa nuvola di dati è poi tracciata una curva di regressione, e in base a quest'ultima sono selezionati i DEGs, ovvero i punti limitrofi alla curva, e verranno esclusi gli outlier, ovvero i geni troppo dispersi all'interno dello stesso set di conte e che influenzerebbero il \log_2fc in maniera errata. Infine, il grafico della distribuzione dei p-value contiene informazioni riguardo l'appartenenza dei geni a determinate distribuzioni statistiche, ad esempio potrebbero far tutti parte della stessa distribuzione e quindi trovarsi in una situazione in cui non ci sono differenze di espressione statisticamente significative (al netto dei falsi positivi), e in cui si avrà una distribuzione dei p-value omogenea; all'opposto, potrebbero essere tutti statisticamente significativi (situazione praticamente impossibile visto il grande numero di geni) cioè appartenere a due diverse

distribuzioni, in questo caso si avrebbe un alto picco nei valori inferiori allo 0,05 che declinerebbe velocemente andando verso valori più alti di p-value (che saranno falsi negativi secondo la premessa fatta); il caso più probabile è quello di avere entrambe le distribuzioni sovrapposte, in cui una parte dei geni risente effettivamente dei diversi parametri ambientali mentre un'altra, probabilmente la maggior parte, no. Il risultato principale è quello tabulare, in cui per ogni DEG sono riportati l'identificativo, il base mean (ascissa dell'MA-plot, cioè la media normalizzata delle conte), il \log_2 fold change (ordinata dell'MA-plot, esprime il rapporto dell'espressione di un gene tra i due set di dati), lfcSE (acronimo per errore standard del \log_2 fold change), e tre indici di significatività statistica ovvero stat (Wald test), p-value, padj (p-adjusted). Per la successiva elaborazione il file tabulare è stato importato in Excel e i dati sono stati riordinati in base all'indice p-adjusted selezionando poi quelli statisticamente significativi in base alla definizione di un valore soglia, nel nostro caso di 0.01. Questo sottoinsieme di DEGs è stato portato su un nuovo foglio Excel e riordinato a sua volta per il \log_2 fold change, questa volta per dividerli in up- e down-regolati (con riferimento al confronto fatto nel DESeq2, cioè SM vs S), questi due sottoinsiemi sono presi e divisi in due fogli Excel. I geni up- e down-regolati sono poi stati copiati in due file testo, per poterli utilizzare come input, insieme al trascrittoma di riferimento, nel software Extractseq.bat, ottenendo le

sequenze nucleotidiche dei geni di interesse richiamate dal trascrittoma grazie al loro identificativo. Ottenute le sequenze nucleotidiche, è stato utilizzato il software Gotranseq per tradurre le sequenze nucleotidiche in amminoacidiche, in tutti e sei i frame; per standardizzare la numerazione dei frame è stato poi usato il software Findorf, che l'ha resa idonea alla successiva procedura di annotazione.

2.9.6 Annotazione dei DEGs

L'annotazione funzionale dei DEGs è stata eseguita attraverso la piattaforma OmicsBox (BioBam, <https://www.biobam.com/omicsbox/>), di cui si è utilizzato il modulo per l'analisi funzionale. Questo modulo, attraverso vari passaggi a cascata nei quali gli output di un passaggio diventano gli input per quello successivo, utilizza il metodo Blast2GO per restituire un'annotazione funzionale GO (Gene Ontology) delle sequenze sottoposte all'analisi. Gene Ontology è un vocabolario unificante che raccoglie e standardizza la nomenclatura biologica, i termini GO fanno riferimento a tre diverse categorie: *Funzione Molecolare*, cioè termini che si focalizzano sull'attività del prodotto genico piuttosto che sulla entità che la svolge; *Componente Cellulare*, termini indicanti i distretti cellulari nei quali il prodotto genico svolge la sua attività, così come i compartimenti cellulari o i complessi macromolecolari; *Processo Biologico*, i cui termini indicano il processo gerarchicamente superiore o il

programma biologico in cui è coinvolta l'attività del prodotto genico. Il metodo Blast2GO genera annotazioni GO attraverso tre step principali: 1) Blast che permette di trovare sequenze ortologhe a quelle fornite come set query al programma, nel nostro caso i DEGs, ottenendone i codici identificativi; 2) Mapping per il recupero di tutti i termini GO associati alle sequenze omologhe ottenute con Blast, utilizzando i loro codici identificativi per ottenere i termini dai database Uniprot, GO e dalla tabella DBXRef del database GO; 3) Annotation per la valutazione, selezione ed assegnazione alle sequenze query dei termini GO raccolti nel mapping. I risultati dell'annotazione, ottenuti per entrambi i gruppi di DEGs (up- e down-regolati), consistono in diversi file contenenti informazioni specifiche: Direct GO Count MF/CC/BP, cioè tre istogrammi che mostrano i termini GO trovati più frequentemente all'interno dei nostri set e quindi comuni a molte sequenze, divisi rispettivamente nelle categorie Funzione Molecolare (MF), Componente Cellulare (CC), Processo Biologico (BP); GO Distribution by Level, un istogramma riepilogativo dei termini più frequenti delle tre categorie già descritte, in questo caso i termini di ogni categoria che possono essere associati tra loro vengono fatti confluire in termini di grado più elevato rispetto alla gerarchia GO; Combined Graph, nel quale termini GO di gradi diversi tra loro, che rappresentano sottoinsiemi più o meno grandi delle nostre sequenze, sono collegati in base alla loro associazione

funzionale e posizionati in base al loro grado, andando a formare una mappa del significato biologico dei vari gruppi di sequenze; WordCloud, una rappresentazione visiva dei principali termini GO, l'importanza del termine nel set è proporzionale alla grandezza del carattere con cui è scritto.

2.10 SEM, microanalisi e diffrazione a raggi X

2.10.1 SEM

La preparazione del materiale (5 campioni per ogni sito) ha richiesto l'inclusione delle singole valve in resina, ciò ha permesso il successivo taglio in direzione umbo-ventrale per l'ottenimento delle sezioni- Una faccia delle sezioni è stata levigata con carta vetrata fine e trattata per 1 minuto con una soluzione di acido acetico all'1% v/v, al fine di rimuovere i detriti e gli artefatti prodotti dal taglio. Le sezioni sono state fissate dal lato non trattato su un dischetto portacampione con nastro biadesivo, è stato poi apposto un ponte di alluminio collegante la faccia superiore della sezione al portacampione. Infine, è seguita la metallizzazione del campione con vapori di oro per rivestire il lato di interesse con uno strato nanometrico (5 nm) carico positivamente che, all'arrivo degli elettroni, li possa diffondere lungo tutta la superficie e scaricarli attraverso il portacampione. Le immagini sono state acquisite (utilizzando un microscopio elettronico a scansione TESCAN VEGA3) lungo tutto lo spessore della sezione e a diversi ingrandimenti (da 170x a 5000x), ottenendo

panoramiche complete e focalizzazioni sui vari strati costituenti la conchiglia, sia in modalità elettroni secondari (SE) per l'analisi topografica della superficie del campione, sia in modalità elettroni retrodiffusi (BSE) per informazioni su fasi, composizione e organizzazione.

2.10.2 Microanalisi

In diverse aree appartenenti ai tre strati (prismatico, di transizione, madreperlaceo) è stata eseguita la microanalisi per ottenere informazioni sugli elementi chimici presenti e le loro proporzioni relative all'interno dell'area.

2.10.3 Diffrazione a raggi X

Al fine di procedere con la diffrazione a raggi X è stato necessario ottenere un materiale sotto forma di polvere da sottoporre all'analisi. Per questo step, è stata presa una singola valva da un individuo per ogni sito che sono poi state lavate e pulite con acqua e lasciate ad asciugare all'aria. Una volta asciutte si è proceduto tramite mortaio con la frantumazione separata di entrambe le valve in polvere fine, facendo attenzione a non far entrare in contatto il materiale di una con quello dell'altra. Le due polveri sono state infine inserite in due contenitori denominati secondo il sito di provenienza. La diffrazione a raggi X permette di ottenere informazioni riguardanti la struttura cristallina del guscio, come ad esempio la presenza di determinati polimorfi di CaCO_3 , che in questo caso potrebbe essere associata ai diversi parametri ambientali (soprattutto la

salinità) a cui sono sottoposti gli individui nei due siti. Inoltre, è stata anche eseguita una microanalisi per ottenere la stechiometria dei componenti del guscio ed osservare la possibile presenza di elementi disciolti.

Capitolo terzo

3. RISULTATI

3.1 Valutazione della qualità delle reads

Il controllo qualità delle reads (file S1), ottenute dal sequenziamento dell'RNA estratto dai 3 pool di Senigallia e 3 pool di Silvi Marina, ha evidenziato risultati sovrapponibili nella prima parte (così come nelle altre) del report grafico (% sequenze associate ad una specifica caratteristica). In particolare, il numero di sequenze e nucleotidi presenti è rispettivamente di 47.739.560 e 7.208.673.560; la distribuzione delle lunghezze è uniforme e pari a 151 bp, come indicato per il sequenziamento. Il grafico del contenuto percentuale di GC ha un andamento gaussiano (% di sequenze rispetto alla % GC). Questo è molto importante perché indica che non ci sono state contaminazioni di sequenze durante i vari passaggi che con le loro copie avrebbero creato dei picchi all'interno della distribuzione. Il grafico sul contenuto di basi ambigue mostra un ottimo risultato in quanto il totale delle sequenze si posiziona su valori molto vicini allo 0%. Questo risultato è coerente con quanto visto nel data report dell'azienda che ha effettuato il sequenziamento e che indicava un contenuto medio di basi ambigue dello 0,0026%. Anche il grafico sulla distribuzione della qualità (Phred Score) ha mostrato ottimi risultati dato che la quasi totalità delle

reads ha ottenuto punteggi superiori a 30, ovvero contengono meno di 1 base errata ogni 1000 assegnate, e tra queste la forte maggioranza delle reads ha ottenuto punteggi tra 35 e 40. Dalla seconda parte del report grafico (valore di specifiche caratteristiche lungo le posizioni) si ottengono i seguenti risultati: il grafico della coverage indicante che tutte le reads coprono tutte le 151 posizioni essendo, come già detto, tutte della stessa lunghezza; il grafico del contributo nucleotidico (Fig. 11) mostrante nelle prime posizioni delle oscillazioni, ciò indica che nella porzione iniziale le reads sono accomunate dalla stessa sequenza e ciò potrebbe essere dovuto all'adattatore utilizzato nel sequenziamento oppure alla bassa diversità dei primi cicli. Dopo le iniziali oscillazioni il grafico assume un andamento omogeneo caratteristico dei campioni senza contaminazioni, per poi deviare gradualmente nelle posizioni finali.

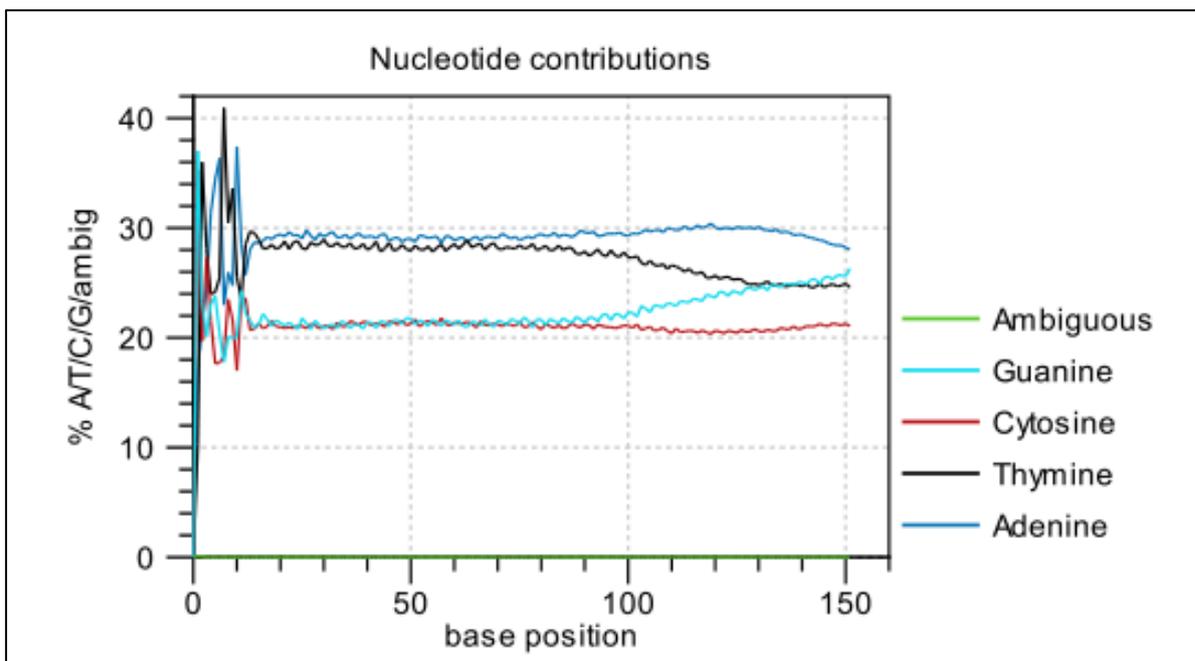


Fig. 11. Grafico del contributo nucleotidico.

Per il grafico del contenuto GC lungo le posizioni valgono le riflessioni del punto precedente in quanto anche qui dalle oscillazioni iniziali si passa ad un andamento omogeneo. Dal grafico sulle basi ambigue lungo le posizioni si può notare esattamente dove cadono le ambiguità, in questo caso la maggioranza degli errori (il cui numero è comunque molto basso) è avvenuto durante i primi cicli di sequenziamento, ovvero si trova principalmente nelle prime posizioni e con un picco dello 0,07%, in linea con i normali processi di sequenziamento. Dal grafico sulla distribuzione della qualità indicata in percentili (Fig. 12), dove i percentili esprimono la percentuale di sequenze che cadono al di sotto di un determinato valore di qualità (in ogni posizione), coerentemente con i risultati visti precedentemente, nelle prime basi si ha una bassa qualità, che

immediatamente aumenta per stabilizzarsi su ottimi punteggi come 37. Infatti nelle prime 60 bp i percentili 95, 75, 50 (mediana), 25 e 5 sono sovrapposti sul punteggio 37, ciò significa che il 95% delle sequenze possiede quel punteggio in quelle posizioni, mentre il 5% è distribuito al di sotto; da queste posizioni in poi il percentile 5 si sposta al punteggio 25, ovvero il 75% delle sequenze ha punteggio di 37, il 20% è distribuito tra il punteggio 37 e il 25 e il 5% è sotto il 25.

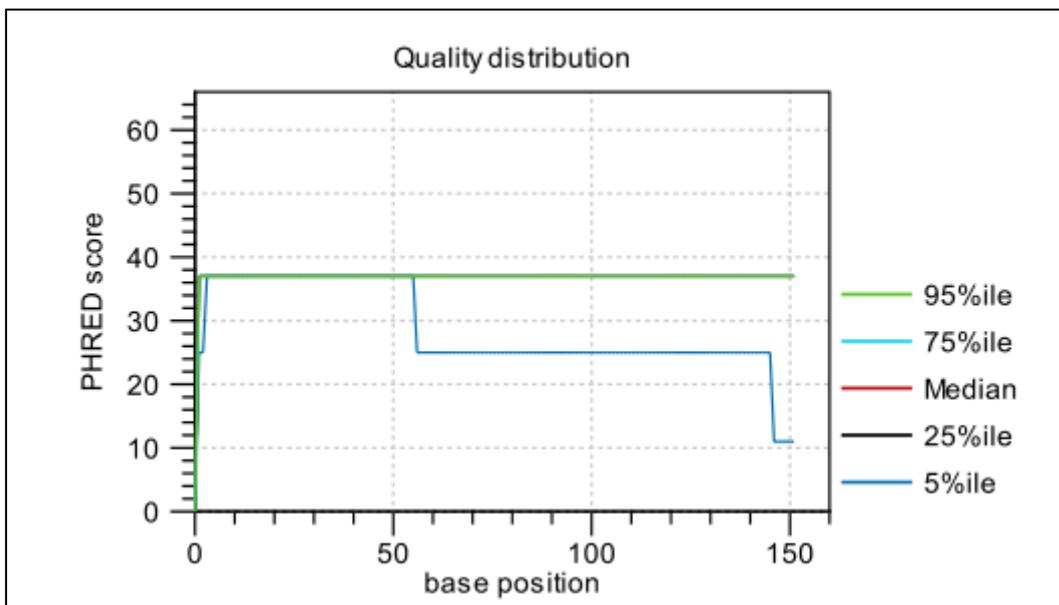


Fig. 12. Grafico sulla distribuzione della qualità.

La terza parte del report riguarda i livelli di duplicazione. Dal grafico sui 5mers (Fig. 13) si può notare che i quattro maggiormente presenti mostrano un aumento delle ripetizioni (fino all'1%) nelle posizioni finali, forse dovuto agli adattatori; infine, il grafico sulle sequenze che si ripetono per un certo numero

di volte mostra che la maggioranza ha livelli di ripetizione molto bassi (inferiori a 10) che in parte sono attesi quando si analizzano dei dati trascrittomici, a causa dei livelli di espressione.

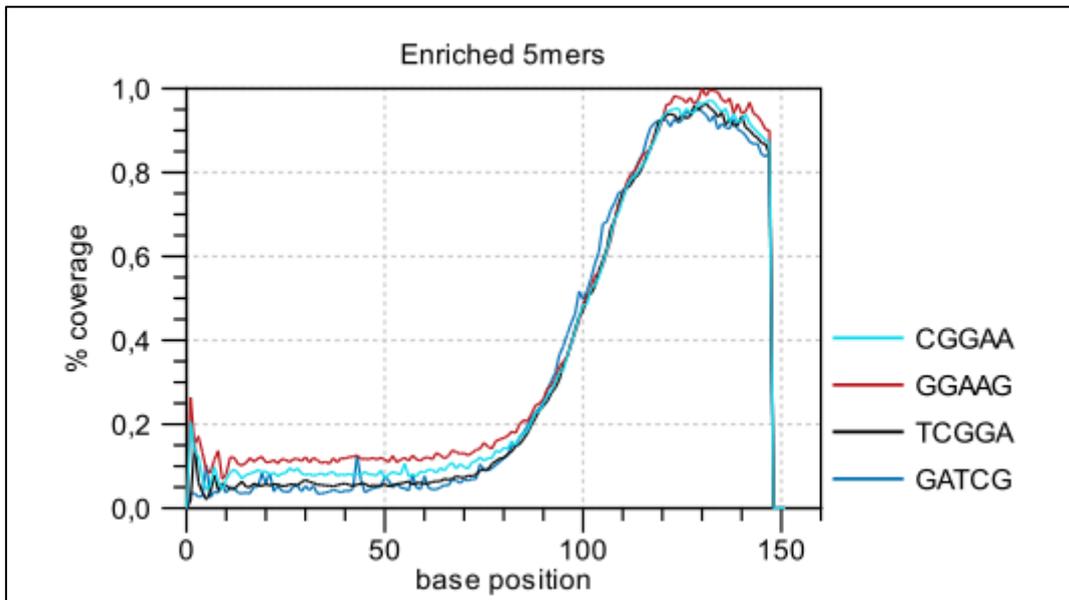


Fig. 13. Grafico di arricchimento dei 5mers.

Nel file di report supplementare troviamo le 25 sequenze maggiormente ripetute, delle quali la prima è una sequenza mononucleotidica di guanina con una ripetizione nelle reads dello 0,20%.

3.2 Trimming dei dati

Proseguendo con il file S1 come modello, dal report del trimming si osserva che partendo da 47.739.560 reads di 151 bp esatte prima del processo si è passati a 47.739.141 reads di 136 bp medie, con lunghezze spazianti tra 85 e 151 bp. Il taglio riguardante gli adattatori ha coinvolto il 46,94% delle reads.

Dall'istogramma (Fig. 14), indicante quante sequenze hanno subito il taglio di un certo numero di nucleotidi, si osserva che la maggior parte dei tagli ha rimosso porzioni di 1-40 bp, il restante numero di tagli diminuisce velocemente nella fascia 40-80 bp.

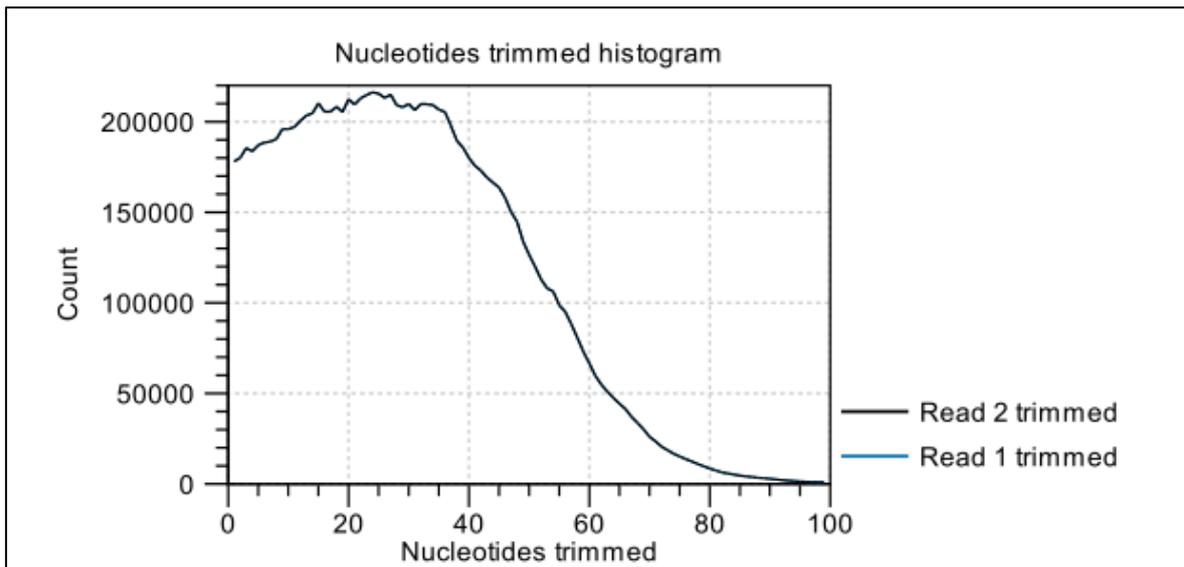


Fig. 14. Iistogramma del numero di nucleotidi tagliati.

3.2.1 Confronto qualità tra reads grezze e trimmed

Il controllo qualità effettuato con le reads grezze è stato ripetuto con le trimmed dello stesso file e i risultati sono stati poi confrontati. Il grafico sulla distribuzione della qualità mostra un aumento delle reads con punteggio 37, ovvero il migliore ottenuto in questo set, che passano dal 58% al 64%, questa situazione si ritrova nel grafico dei percentili (Fig. 15) in cui si osserva un miglioramento della qualità nelle posizioni finali.

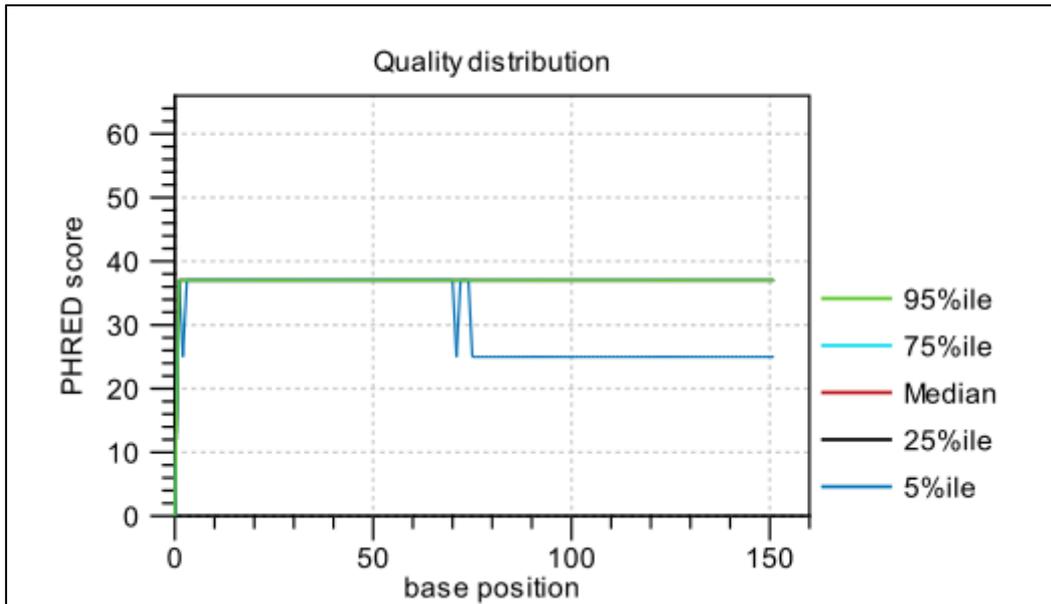


Fig. 15. Distribuzione qualità post-trimming.

Il grafico della coverage conferma che dopo il trimming non tutte le reads coprono tutte le posizioni fino alla 151. La contribuzione nucleotidica (Fig. 16) resta sbilanciata nelle posizioni iniziali ma migliora notevolmente in omogeneità nelle posizioni finali.

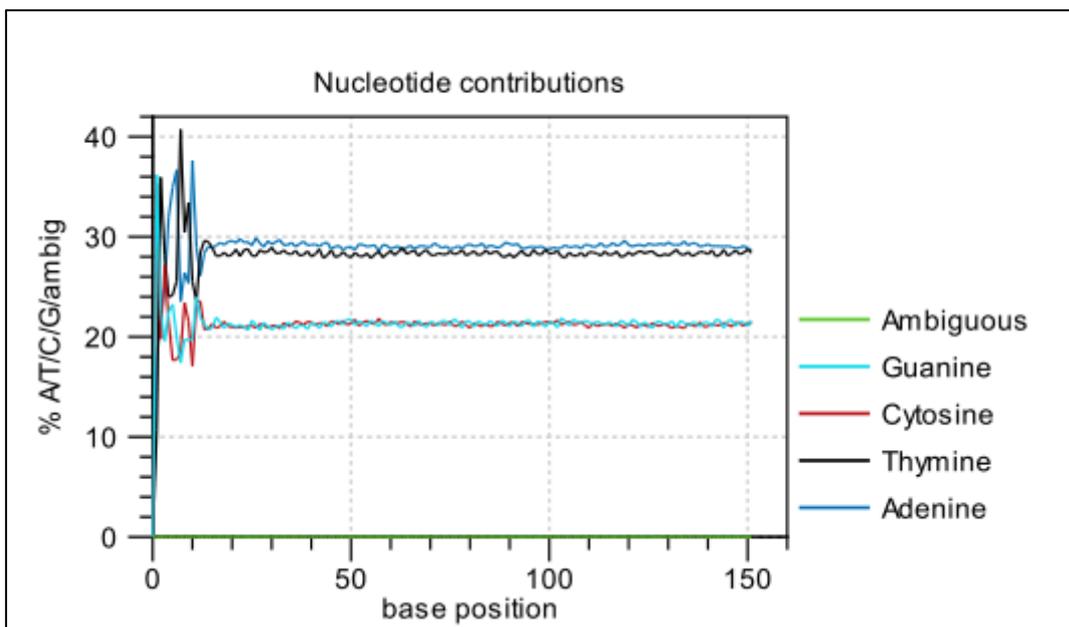


Fig. 16. Contributo nucleotidico post-trimming.

Il contenuto di basi ambigue lungo le posizioni cala drasticamente, soprattutto nelle posizioni iniziali che essendo esterne sono state rimosse più facilmente, infatti aumenta in proporzione il peso delle ambiguità nelle posizioni centrali che non sono state rimosse. Per quanto riguarda i 5mers (Fig. 17) c'è stato un cambio nelle sequenze più rappresentate, che prima lo erano nelle posizioni finali e proprio per questo sono state probabilmente eliminate durante il trimming (insieme a quelle adattatrici), lasciando il posto a nuove sequenze meno rappresentate di cui una però (mononucleotidica, costituita dalla base adenina) si mantiene allo 0,8% circa in tutte le posizioni; questa sequenza si ritrova anche nel file di report supplementare, ovvero è in realtà una sequenza duplicata della quale il programma ha interpretato una porzione come un 5mer.

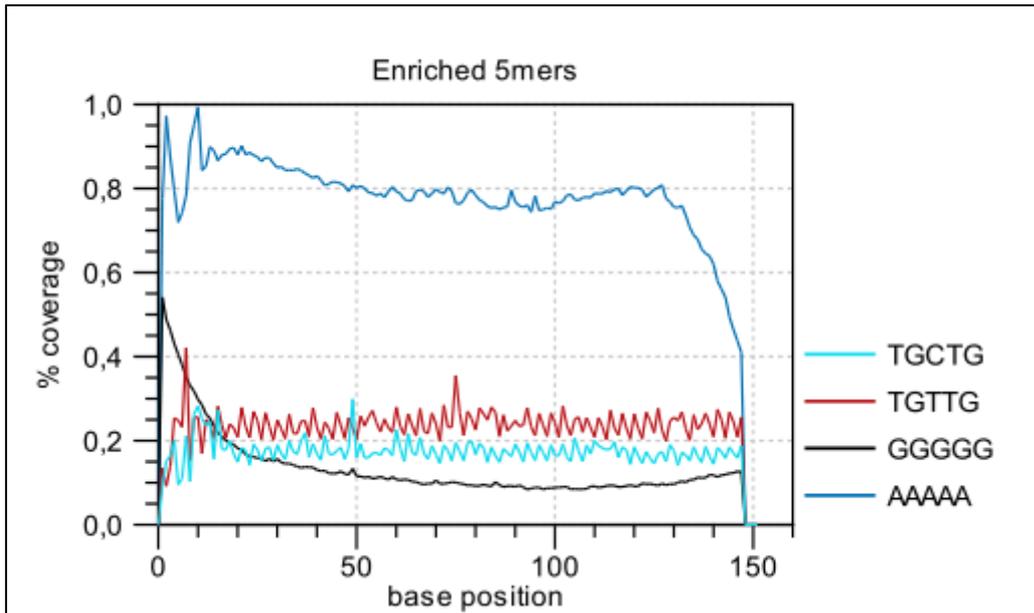


Fig. 17. Stato di arricchimento dei 5mers post-trimming.

3.3 Assemblaggio delle reads filtrate

Considerando l'assemblaggio di Senigallia, nel report si trovano inizialmente i dati percentuali del contributo nucleotidico e delle basi ambigue, nello specifico si ha: adenina 29,8% - citosina 20,0% - guanina 18,3% - timina 31,8% - N 0,1%; il contenuto in GC è del 38,3% (somma delle due percentuali); la percentuale di basi ambigue risulta maggiore rispetto a quella del singolo file. I risultati principali dell'assemblaggio si trovano nelle due tabelle di misurazioni, contenenti informazioni che differiscono leggermente tra loro a causa dell'inclusione o meno delle regioni gap. Considerando il primo caso, osserviamo un N75 di 294 bp ovvero il 75% dei nucleotidi appartiene a contig di lunghezza uguale o maggiore a 294 bp, discorso analogo per N50 (422 bp) e

N25 (757 bp). Si osservano inoltre la lunghezza minima di 200 bp, massima di 16.412 bp e media di 413 bp. I contig assemblati sono in totale 223.827. La tabella di distribuzione delle lunghezze conferma quanto detto precedentemente e osservato nell’N25, ovvero la maggioranza dei contig ha una lunghezza inferiore a 1 kb; infine, dal grafico sull’accumulo della lunghezza totale in base al numero di contig considerati (partendo da quelli più lunghi) possiamo estrapolare quanti contig servono per coprire una determinata porzione di trascrittoma.

3.4 Valutazione della qualità dell’assemblaggio

I risultati della valutazione BUSCO (<https://busco.ezlab.org/>) dei tre assemblaggi (Silvi Marina, Senigallia, Totale) consistono in un istogramma in cui per ognuno è riportata la percentuale di tutte le tipologie di match (completi, duplicati, frammentati e mancanti). Queste percentuali fanno riferimento ai gruppi di geni ortologi presenti nel database Mollusca selezionato. I tre assemblaggi (Fig. 18) hanno ottenuto percentuali molto simili tra loro, approssimativamente il 22% di match completi, l’8% di frammentati e il restante 70% di mancanti.

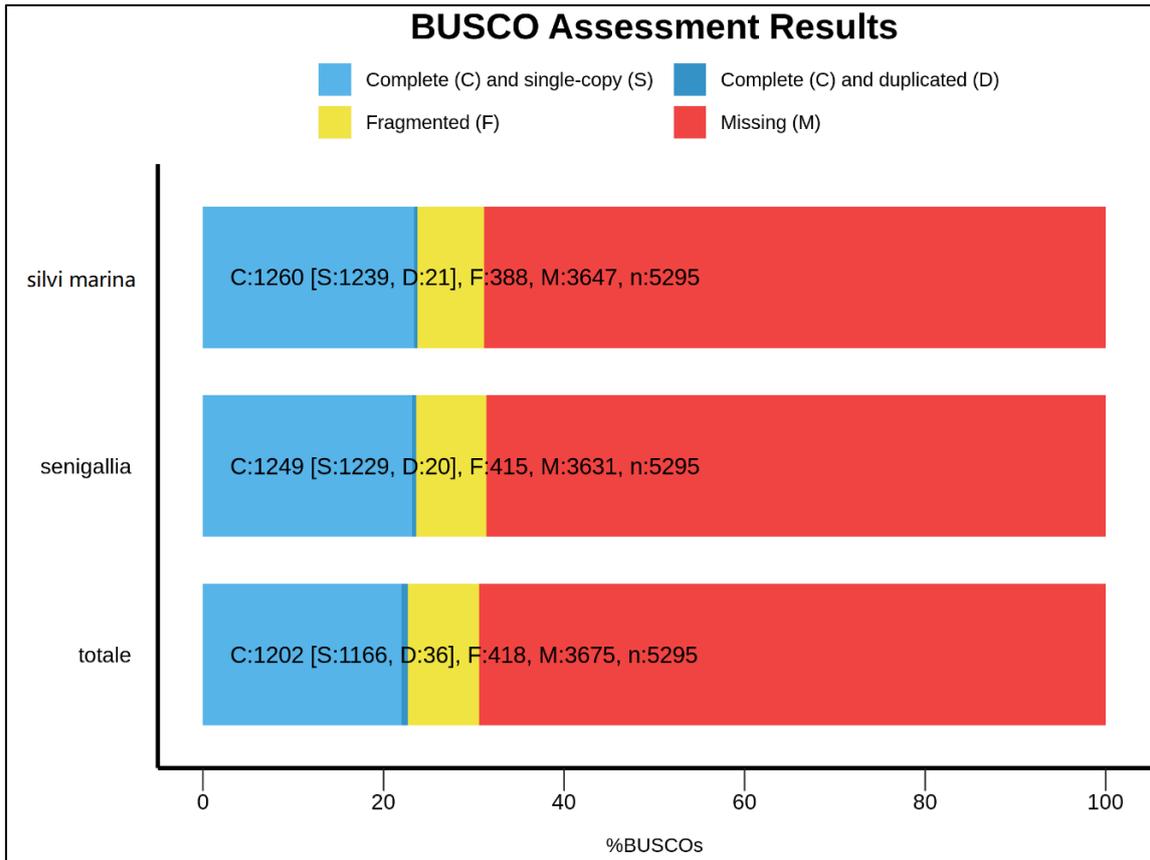


Fig. 18. Valutazione assemblaggi Silvi Marina, Senigallia, Totale.

3.5 Determinazione dei geni differenzialmente espressi (DEGs)

I risultati finali della determinazione dei geni differenzialmente espressi sono stati ottenuti con il software DESeq2, punto di arrivo dei precedenti step che hanno reso idonei i dati alla valutazione e all'individuazione dei DEGs. Il grafico delle componenti principali (PCA, Fig. 19) mostra un'effettiva divisione in due cluster basati sul sito di appartenenza, ovvero i dati dei campioni di Senigallia sono raggruppati tra loro così come quelli di Silvi Marina. Questa divisione avviene lungo la PC1 (rappresentante il 38% della

variabilità totale tra i dati), ma non lungo la PC2 (24%). Essendo però gli assi del grafico classificati in ordine di importanza, una divisione lungo la PC1 ha un peso molto maggiore di una non divisione lungo la PC2.

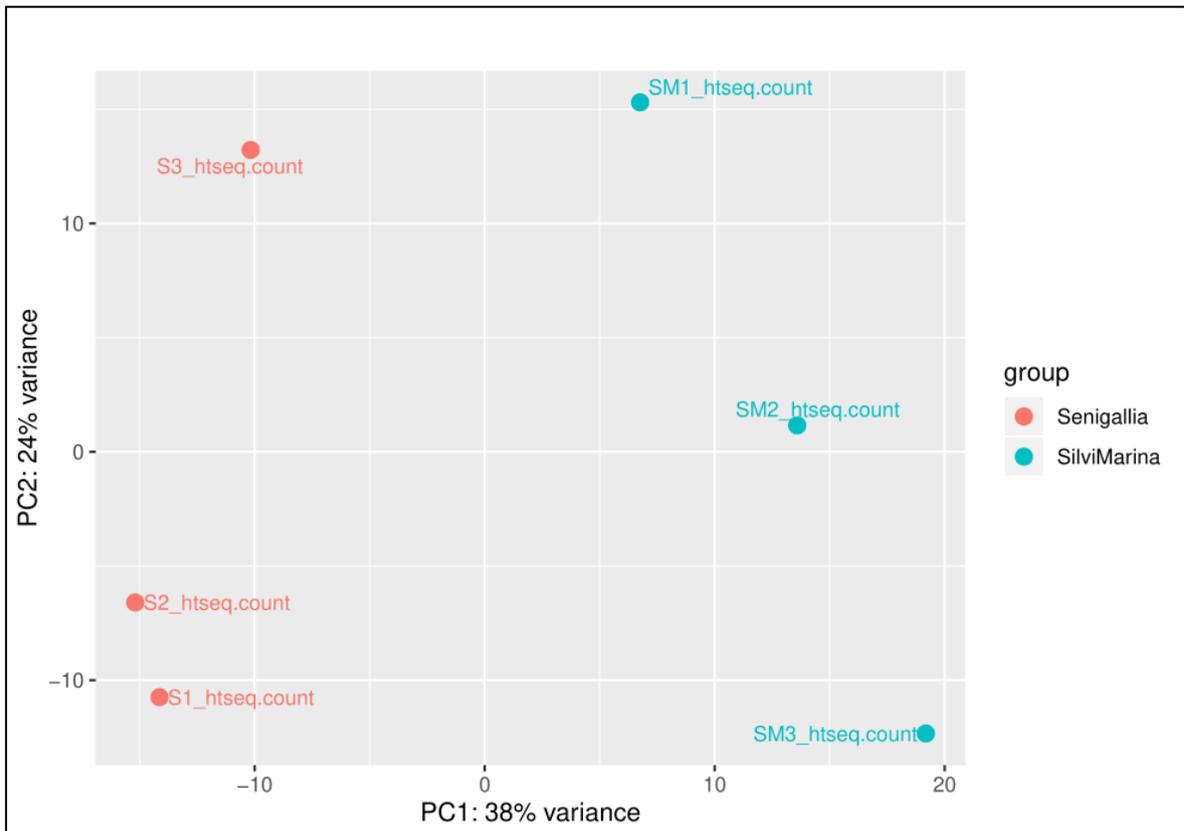


Fig. 19. Grafico delle componenti principali PCA.

La matrice di distanza (Fig. 20) fornisce una panoramica della diversità e della somiglianza tra le conte, e soprattutto nei dendrogrammi laterali segue la divisione in base ai siti di campionamento, confermando il risultato precedente.

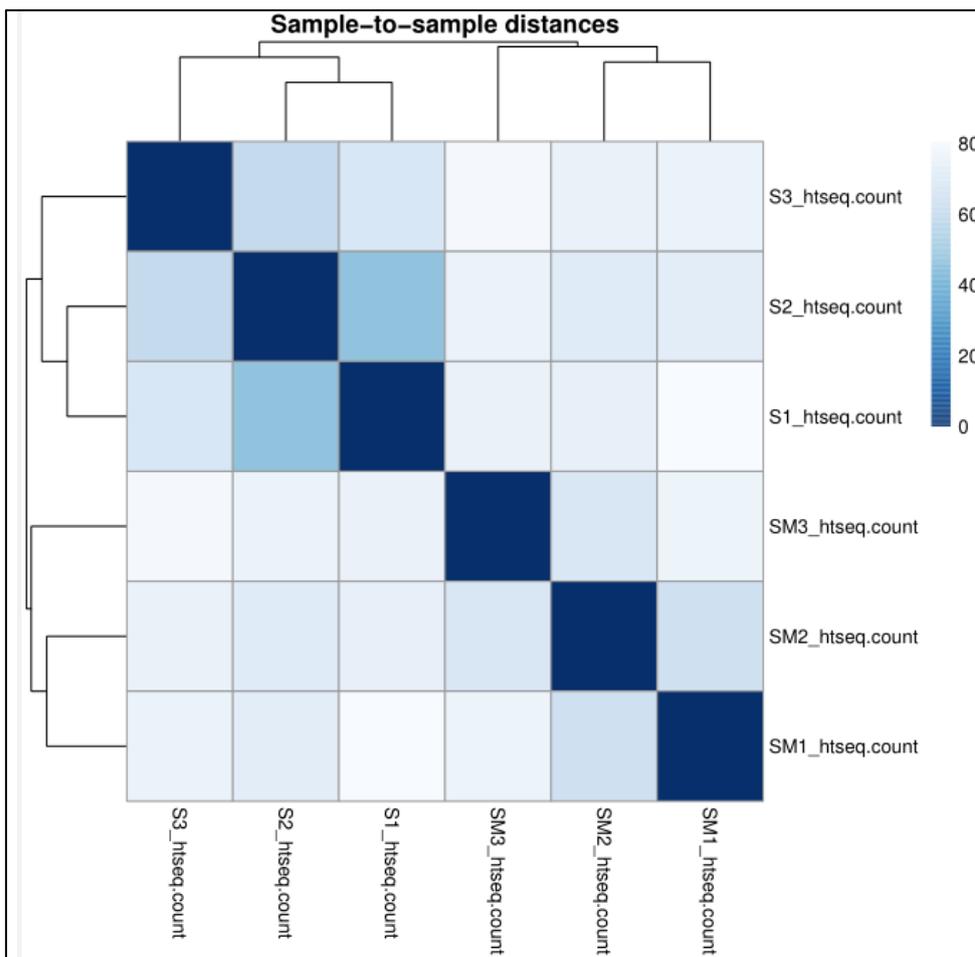


Fig. 20. Heatmap e dendrogrammi ottenuti dalla matrice di distanza.

Dal grafico di dispersione (Fig. 21) si può osservare come il programma abbia valutato e selezionato i geni differenzialmente espressi, tramite il calcolo delle stime di dispersione basate sulle conte normalizzate. In questo caso i geni selezionati sono gli spot blu limitrofi alla curva di regressione in rosso, mentre i punti neri cerchiati in blu sono gli outlier scartati, ovvero quei geni le cui conte all'interno dello stesso set hanno valori troppo diversi tra loro che quindi inficerebbero in modo anomalo l'indice $\log_2 fc$.

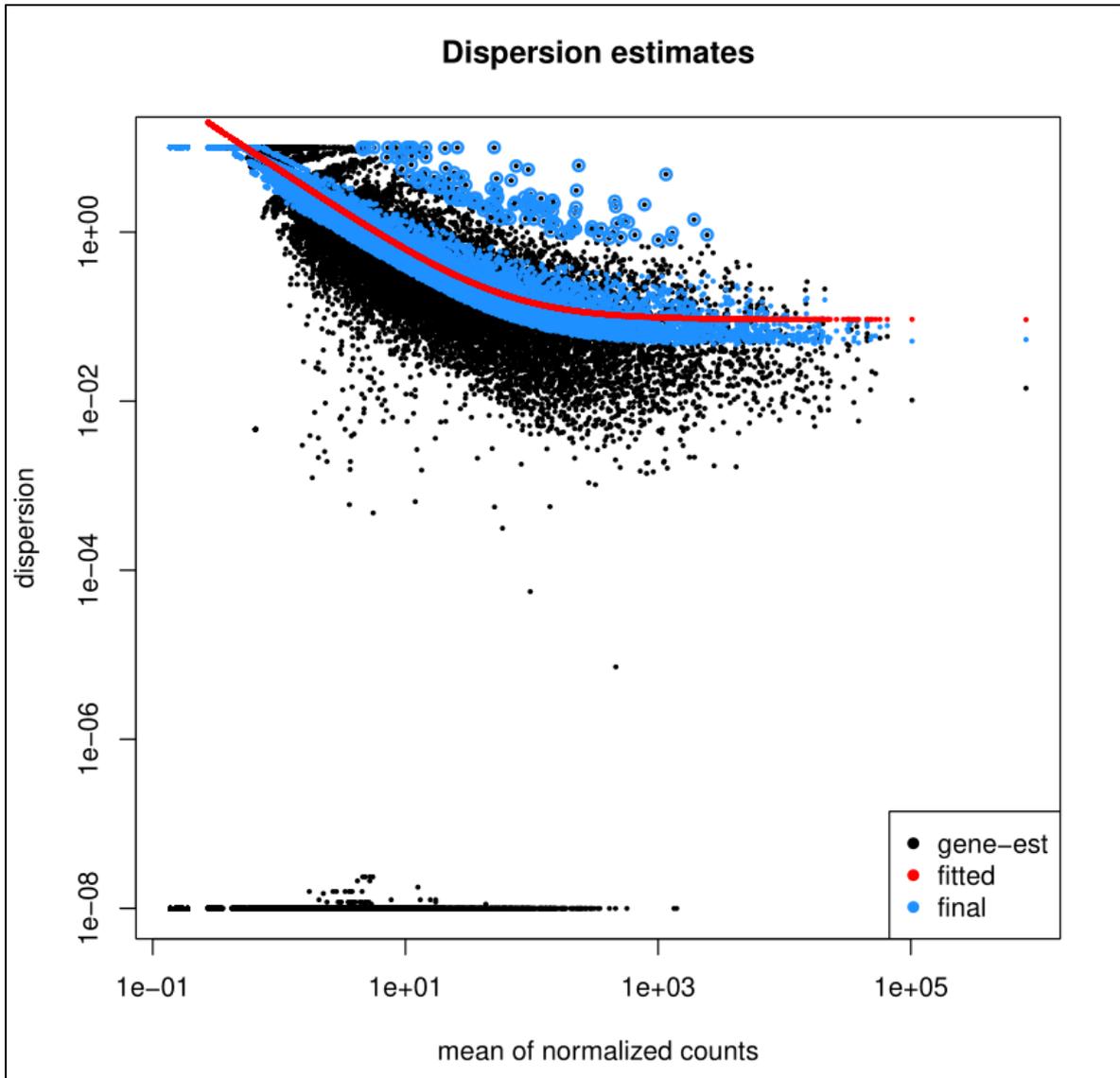


Fig. 21. Grafico di dispersione.

Dall'istogramma dei p-value (Fig. 22) si può riconoscere (ad eccezione dei picchi presenti nella seconda metà del grafico) una distribuzione che proviene dalla somma di due differenti realtà presenti nei dati, infatti alcuni dei geni appartengono a due diverse distribuzioni statistiche (in questo caso presumibilmente le due distribuzioni sono dovute ai due siti di campionamento,

in cui si hanno parametri ambientali diversi) e la loro distribuzione specifica mostrerebbe alti picchi su valori di p-value inferiori a 0,05. A questa distribuzione si aggiunge quella dei geni che fanno invece parte della stessa distribuzione ovvero la cui espressione non è influenzata dai siti di campionamento. Questa distribuzione specifica sarebbe uniforme in tutti i valori di p-value. Dalla somma di queste due distribuzioni se ne ottiene una simile a quella riportata.

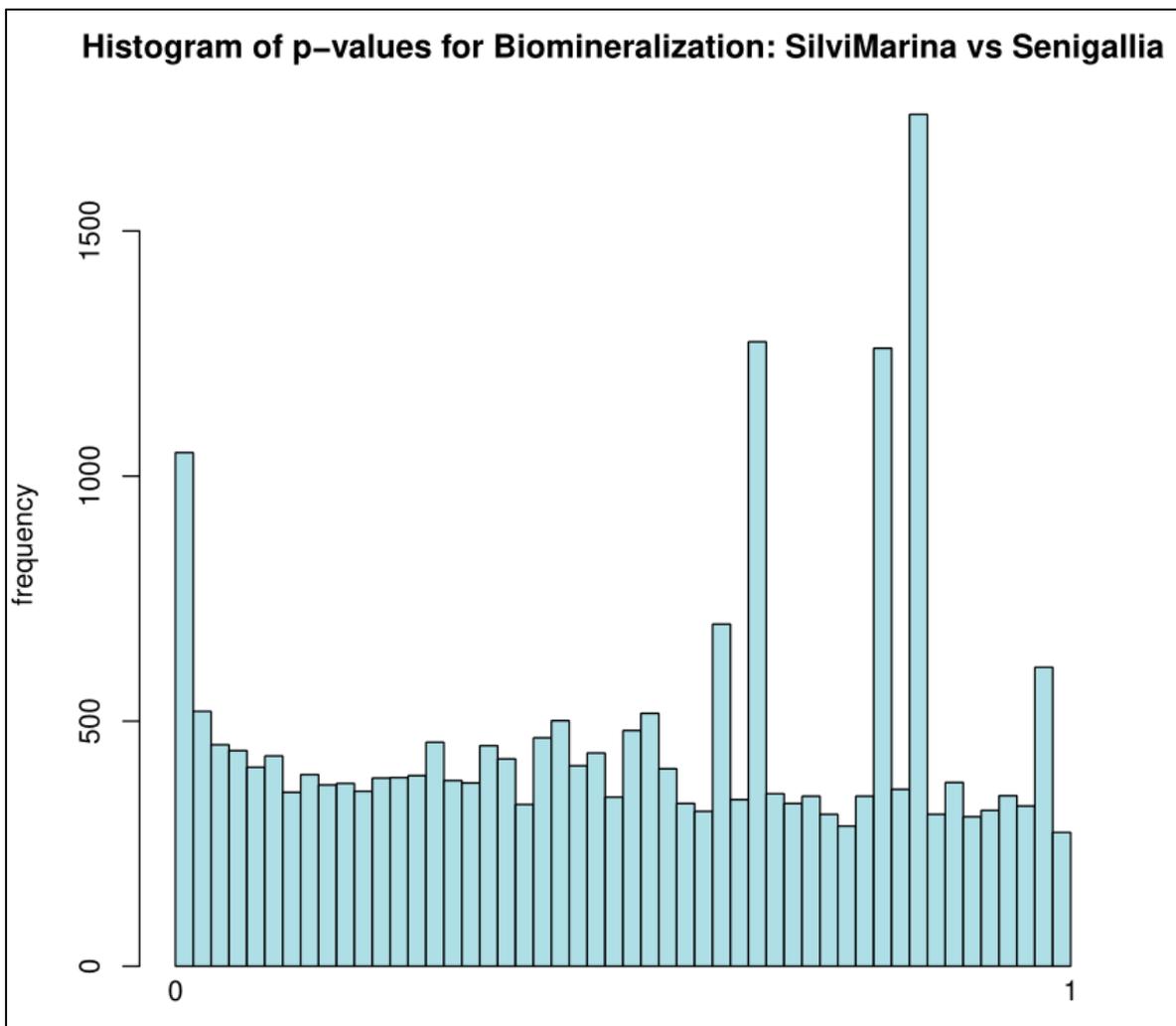


Fig. 22. Istogramma dei p-value.

Tutti i geni presenti nelle prime posizioni, quindi, non sono dei veri positivi, e ciò ha richiesto delle metodiche di correzione che sono risultate nel calcolo del p-adjusted, usato poi per la selezione dei DEGs.

L'MA-plot (Fig. 23), avendo come ordinata il \log_2fc , fornisce una panoramica della quantità di geni up- e down-regolati, a seconda della loro posizione rispetto alla retta orizzontale. Gli spot colorati in rosso sono i DEGs statisticamente significativi, ovvero con un p-adjusted sotto la soglia (0.01).

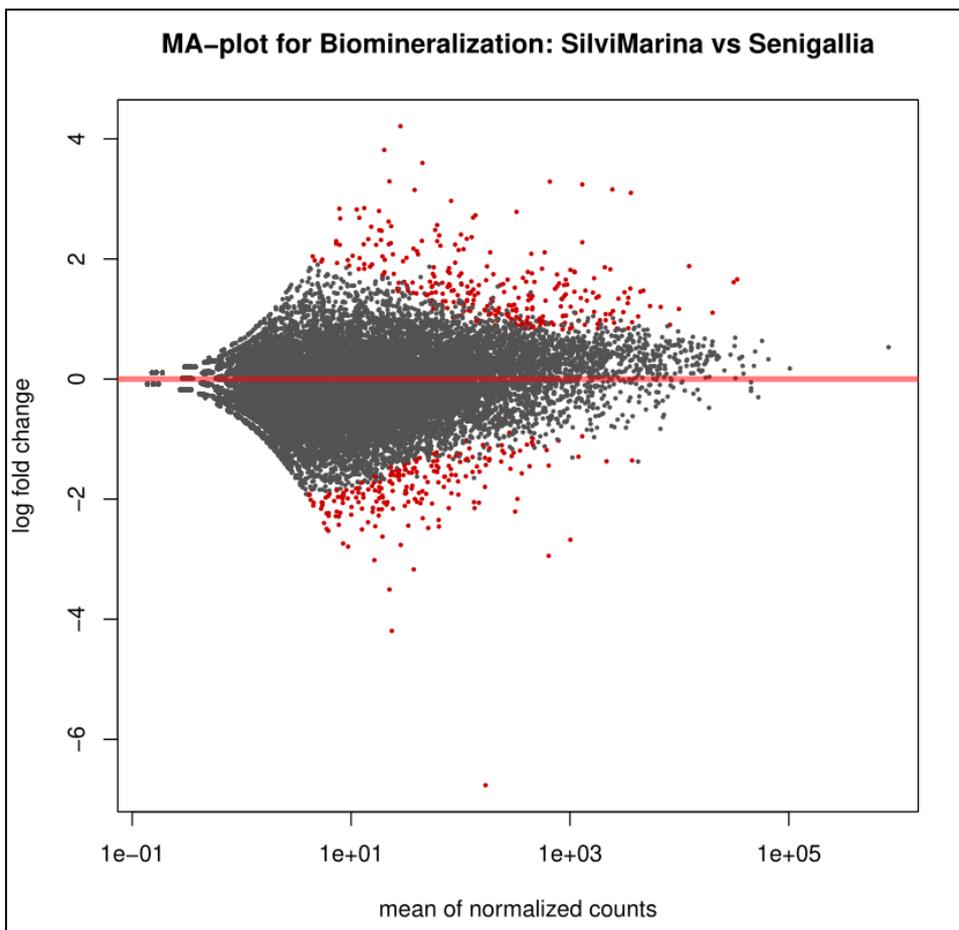


Fig. 23. MA-plot.

Tutti i geni così trovati sono raccolti nel file di report tabulare, contenente le informazioni necessarie alla selezione e divisione manuale dei dati, come il p-adjusted usato per il filtraggio e la selezione dei dati e il \log_2fc usato per la divisione in up- e down-regolati; dopo questi passaggi il risultato finale della determinazione dei DEGs (Silvi Marina vs Senigallia) è rappresentato da 50 geni down-regolati e 108 geni up-regolati.

3.6 Annotazione dei DEGs

3.6.1 Up-regolati Silvi Marina (o down-regolati Senigallia)

Per quanto riguarda gli istogrammi con i termini GO trovati più frequentemente e appartenenti alle tre categorie *Funzione Molecolare*, MF; *Componente Cellulare*, CC; *Processo Biologico*, BP, i principali termini MF sono: attività GTPasica, legame al GTP, costituente strutturale del citoscheletro, legame all'ATP, legame all'RNA, componente strutturale dei ribosomi; per i termini CC si hanno: citoplasma, microtubuli, membrana, componenti integrali di membrana, piccola subunità ribosomiale; infine i principali termini BP sono: processi basati sui microtubuli, processi ossido-riduttivi, organizzazione citoscheletro, citocinesi mitotica, regolazione del segnale della fosfatidilinositolo-3-chinasi, traduzione. Nell'istogramma GO Distribution by Level (Fig. 24), contenente termini maggiormente inclusivi rispetto ai precedenti, in BP abbiamo principalmente processi cellulari, processi

metabolici e risposta a stimoli; in MF attività catalitica, attività di legame, attività di molecole strutturali; in CC entità anatomiche cellulari, complessi contenenti proteine.

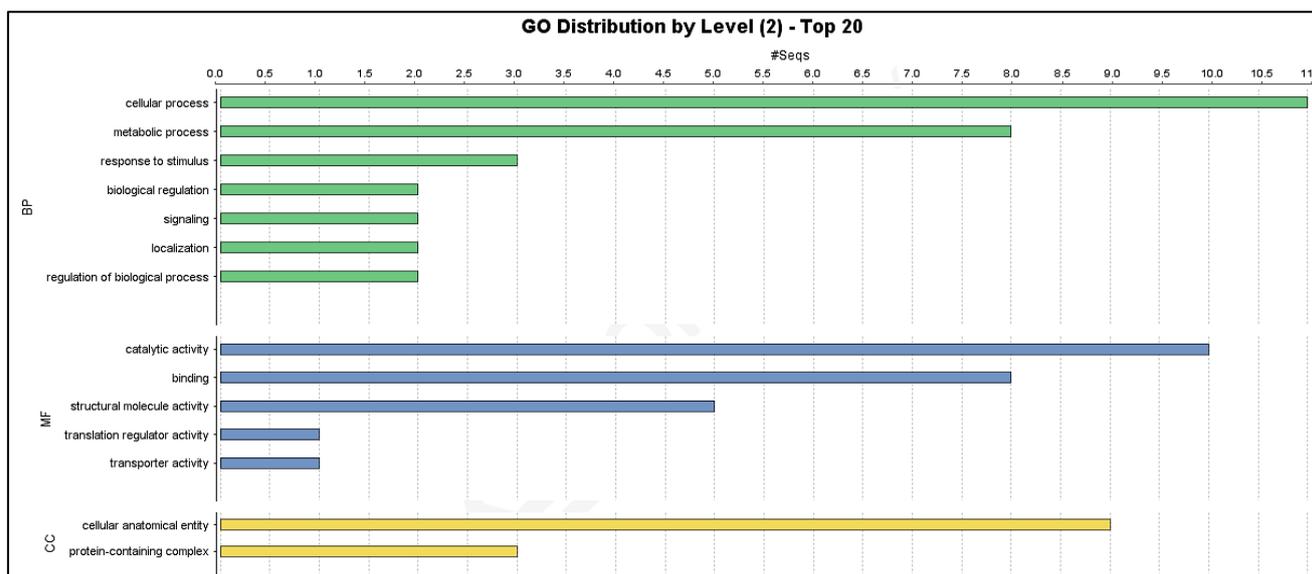


Fig. 24. Distribuzione dei termini GO per le tre categorie BP, MF, CC.

Nel Wordcloud i termini in maggior rilievo sono: entità anatomica cellulare, struttura anatomica intracellulare, membrana, citoplasma. I risultati maggiormente interessanti sono probabilmente quelli derivanti dalle ricerche effettuate per ogni DEG del gruppo. Per quanto riguarda il focus di questo studio, ovvero la biomineralizzazione, i geni up-regolati in Silvi Marina con funzioni correlate a tale processo sono: Neuromedin-U receptor 2-like, recettore di membrana coinvolto nei processi di trasporto e segnalazione del calcio (Raddatz et al., 2000), e implicato nella crescita ossea (Rucinski et al.,

2008); Sarcoplasmic calcium-binding protein, proteina abbondante nel tessuto muscolare dei molluschi che presenta siti di legame per il calcio ed è coinvolta nel suo storage; Zinc transporter ZIP14-like, proteina trasportatrice di zinco implicata nei processi immunitari; Rho GTPase, coinvolta in vari processi inerenti alla polarità cellulare, traffico vescicolare, ciclo cellulare; Serum amyloid A-5 protein-like, proteina della fase acuta; Alkaline phosphatase, ruolo di defosforilazione e regolatrice della mineralizzazione ossea attraverso il controllo dei livelli di difosfato (Hessle et al., 2002); Retinal dehydrogenase 1, proteina ossidante le retinaldeidi in acido retinoico.

3.6.2 Down-regolati Silvi Marina (o up-regolati Senigallia)

Osservando gli istogrammi con i termini GO trovati più frequentemente e appartenenti alle tre categorie *Funzione Molecolare*, *Componente Cellulare*, *Processo Biologico*, i principali termini MF sono: attività di legame al DNA dei fattori di trascrizione, legame al DNA, legame agli ioni metallici, costituente strutturale dei ribosomi; per i termini CC abbiamo: complesso VI della miosina, raddomero, anello di actina, complesso V della miosina, respirasomi, centrosomi, centrioli; infine i principali termini BP sono: regolazione della trascrizione dell'RNA polimerasi II, traduzione. Nell'istogramma GO Distribution by Level (Fig. 25), in BP abbiamo principalmente processi metabolici, processi cellulari regolazione biologica,

regolazione dei processi biologici, localizzazione; in MF legame, regolazione funzioni molecolari, attività molecole strutturali; in CC entità anatomica cellulare, complesso contenente proteine.

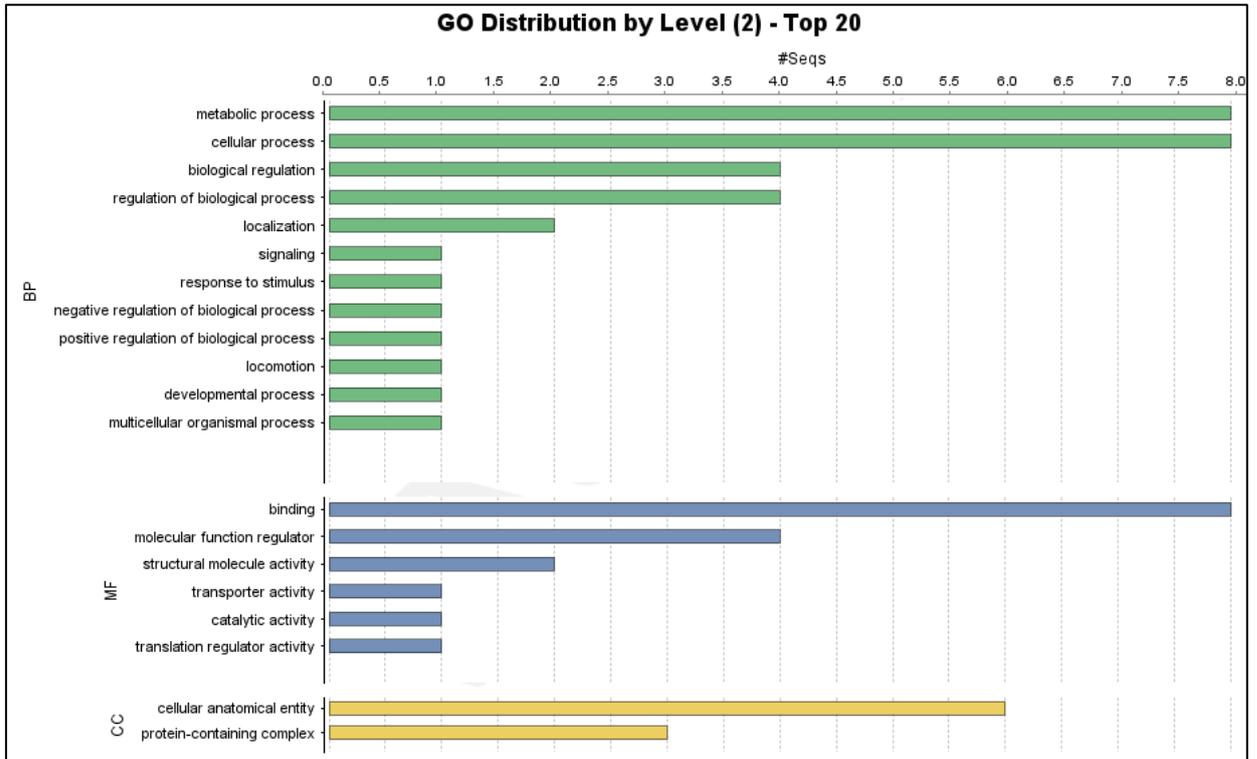


Fig. 25. Distribuzione dei termini GO per le tre categorie BP, MF, CC.

Il Wordcloud evidenzia maggiormente: entità anatomica cellulare, organello intracellulare non legato alla membrana, ribosoma. Dalle ricerche effettuate per i singoli DEGs, i geni up-regolati di Senigallia con funzioni correlate alla biomineralizzazione sono: transcription factor AP-1-like, fattore di trascrizione fortemente connesso alla biomineralizzazione; 60S ribosomal protein L4-like, costituente della subunità ribosomiale maggiore e coinvolta nel legame

all'RNA; Calmodulin, componente della matrice organica del guscio; Fos-related antigen 2, appartiene alla famiglia dei fattori di trascrizione AP1; Transgelin-3-like isoform X2, proteina di matrice del guscio della famiglia delle calponine; Cytochrome c oxidase subunit 1, componente della catena di trasporto degli elettroni; Early growth response protein 1-B-like, attivatore trascrizionale coinvolto nella differenziazione cellulare, nella mitosi e nello sviluppo osseo dei vertebrati (Toan et al., 2019).

3.7 SEM, microanalisi e diffrazione a raggi X

3.7.1 SEM

La scelta delle aree da acquisire, durante le varie prove effettuate al microscopio, è stata fatta focalizzandosi sulla ricerca di eventuali differenze o analogie riguardo la stratificazione della conchiglia lungo la sezione e la tipologia, composizione e organizzazione degli strati. Particolare attenzione è stata rivolta all'individuazione di differenze possibilmente prodotte dalle condizioni presenti nelle diverse zone di campionamento. Per quanto riguarda la stratificazione della conchiglia lungo la sezione, le acquisizioni più esplicative sono quelle a basso ingrandimento che ne offrono una panoramica completa, corredate da altre a maggiore ingrandimento nel punto di transizione e nelle varie zone, per approfondirne le differenze strutturali che le distinguono nelle diverse tipologie. Dalla Fig. 26 si possono osservare i due strati presenti,

separati da uno strato di transizione, lungo l'intera sezione e le loro proporzioni relative. Nella Fig. 27 è mostrato un ingrandimento maggiore della zona di transizione, punto di incontro dello strato prismatico superiore e granulare inferiore.

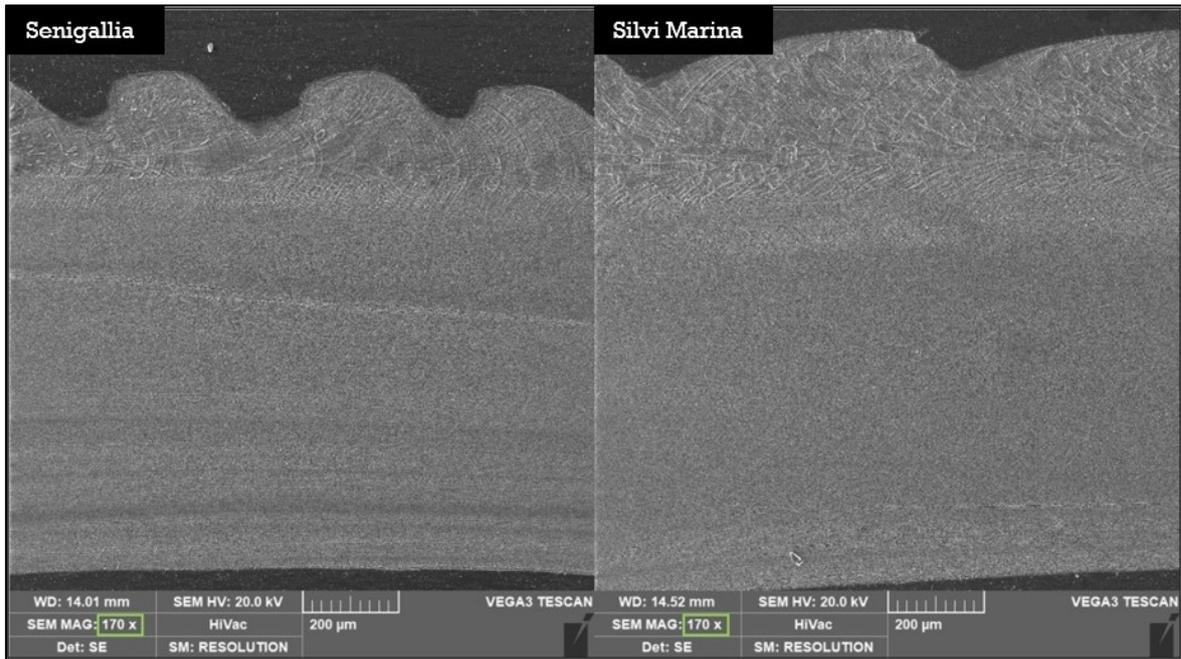


Fig. 26. Organizzazione degli strati.

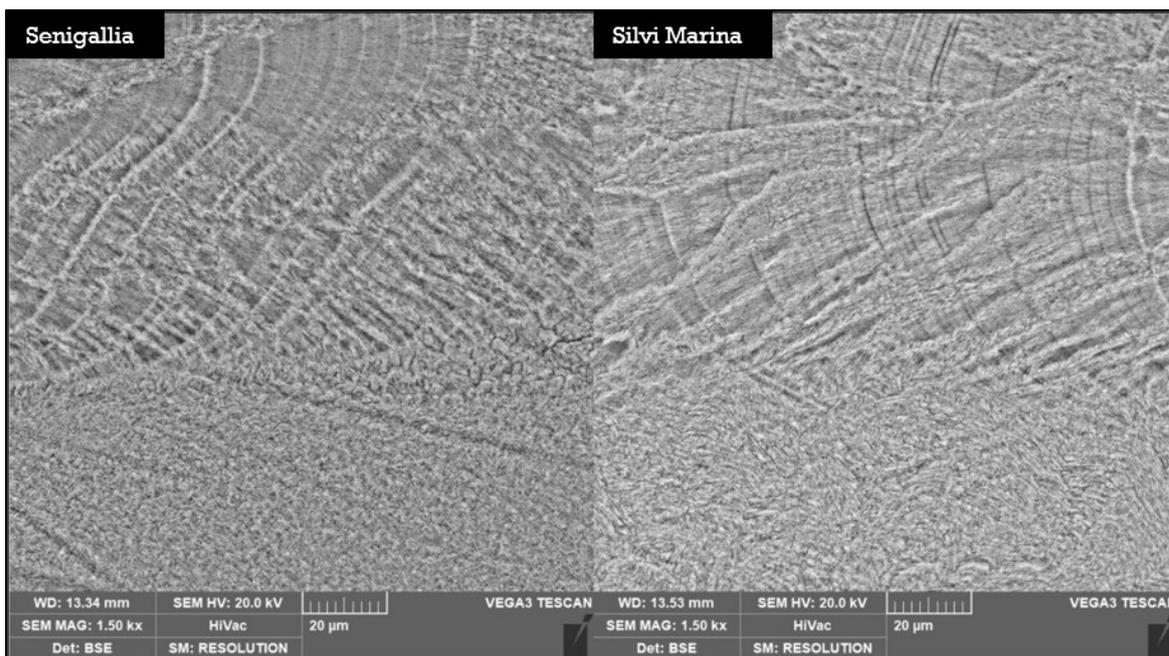


Fig. 27. Zona di transizione.

Seguono poi le acquisizioni a maggior ingrandimento dello strato prismatico superiore e di quello madreperlaceo granulare inferiore (Fig. 28, 29). La composizione delle varie sezioni può essere osservata nella Fig. 30 in cui si notano inoltre chiaramente i fronti di deposizione della conchiglia lungo le creste.

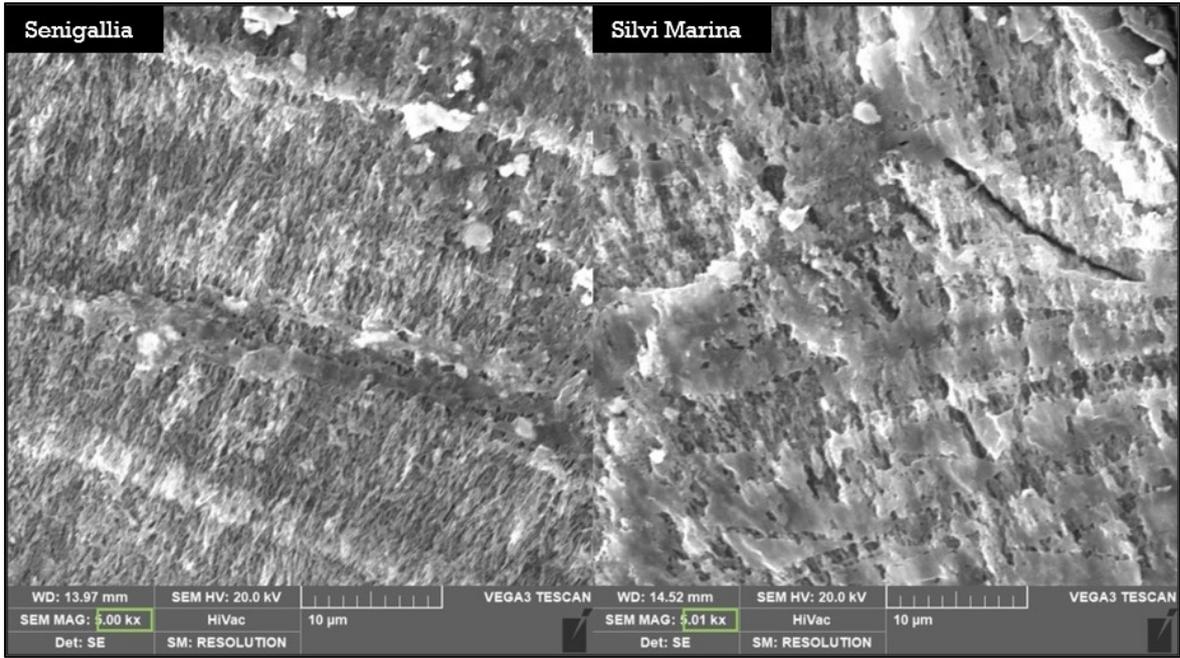


Fig. 28. Strato prismatico.

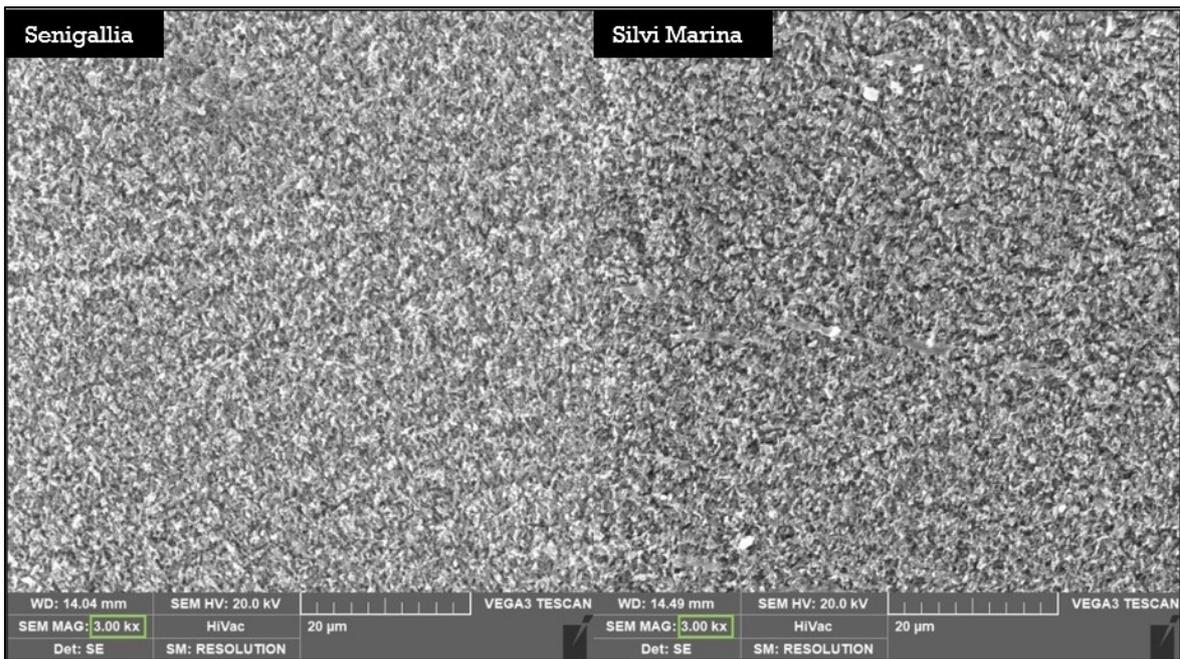


Fig. 29. Strato madreperlaceo.

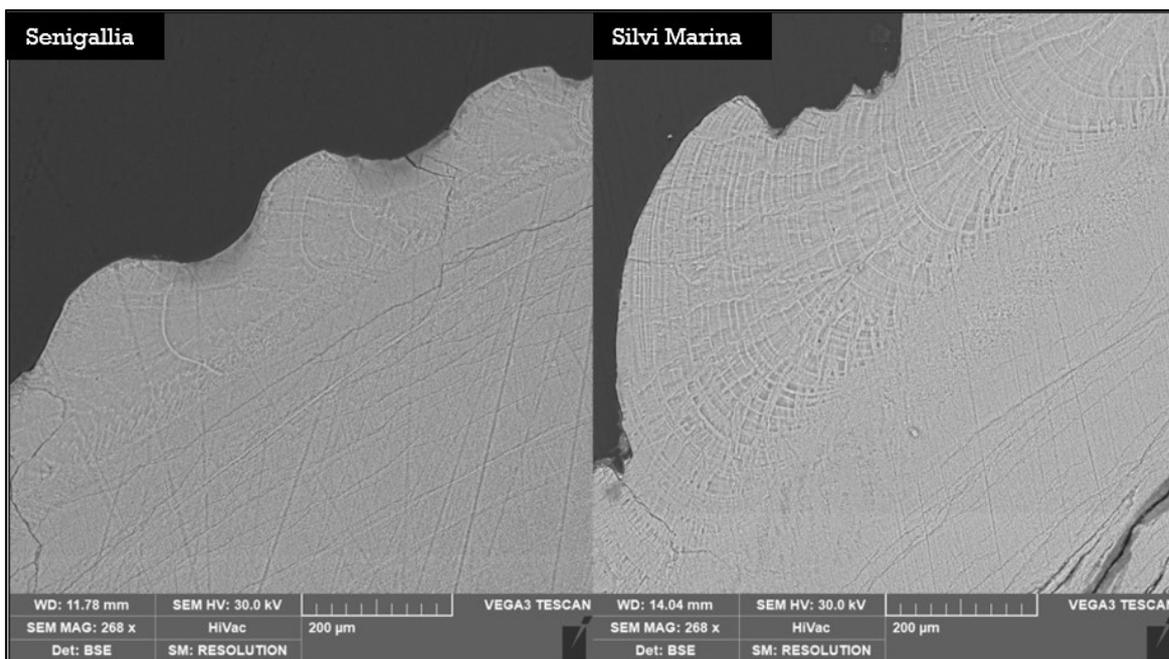


Fig. 30. Particolare delle creste.

Infine, le differenze maggiormente visibili che potrebbero essere state prodotte dalle diverse condizioni ambientali (che saranno discusse nel capitolo dedicato), si trovano nelle Fig. 31, 32, 33, 34, dove si può notare come a livello delle creste, cioè nello strato prismatico superiore, nelle vongole di Silvi Marina siano presenti delle zone scure, non presenti in quelle di Senigallia, che indicano una struttura caratterizzata da maggiore fragilità e porosità.

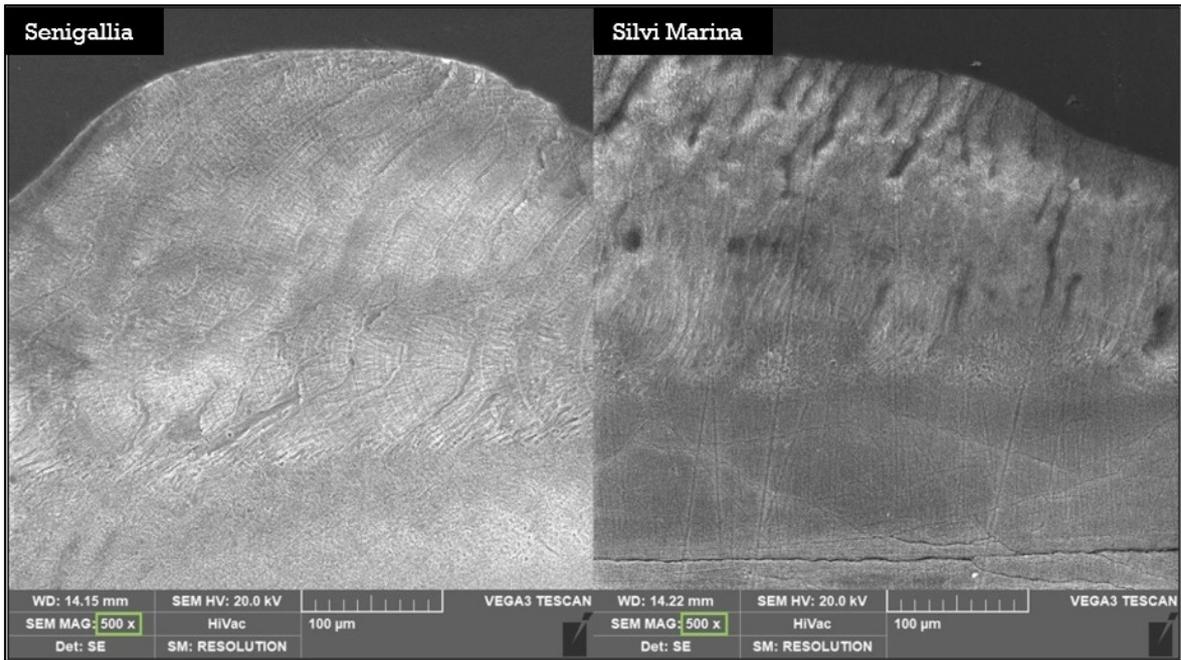


Fig. 31. Differenze nella struttura della conchiglia tra le vongole di Senigallia e Silvi Marina.

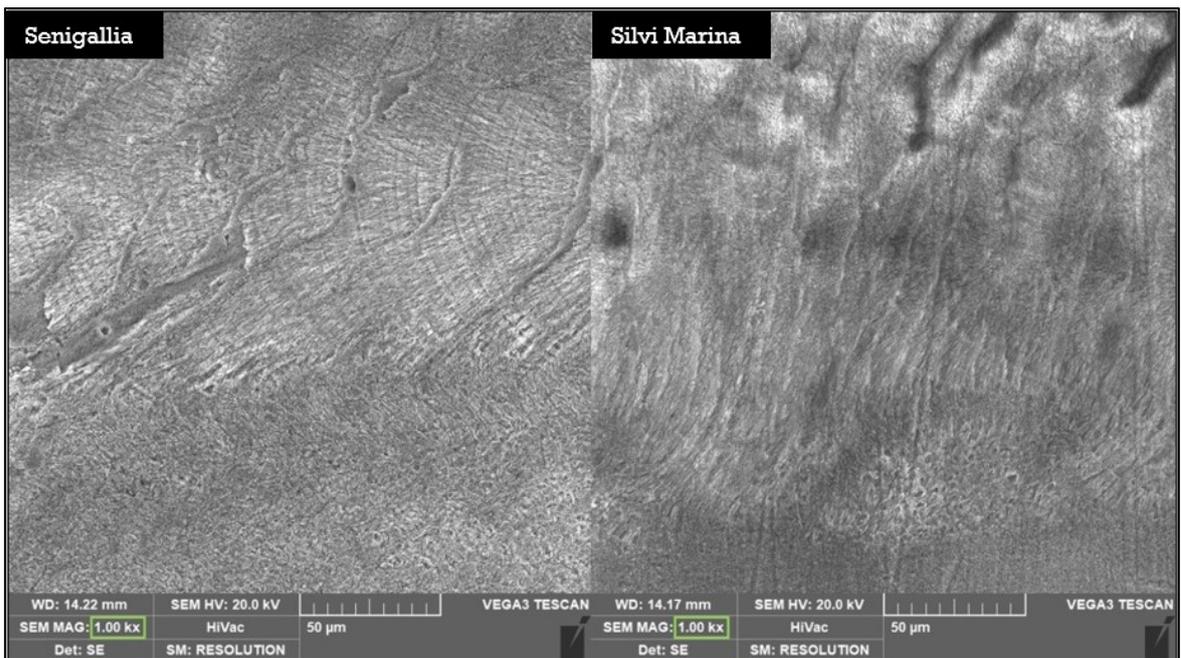


Fig. 32. Differenze nella struttura della conchiglia tra le vongole di Senigallia e Silvi Marina.

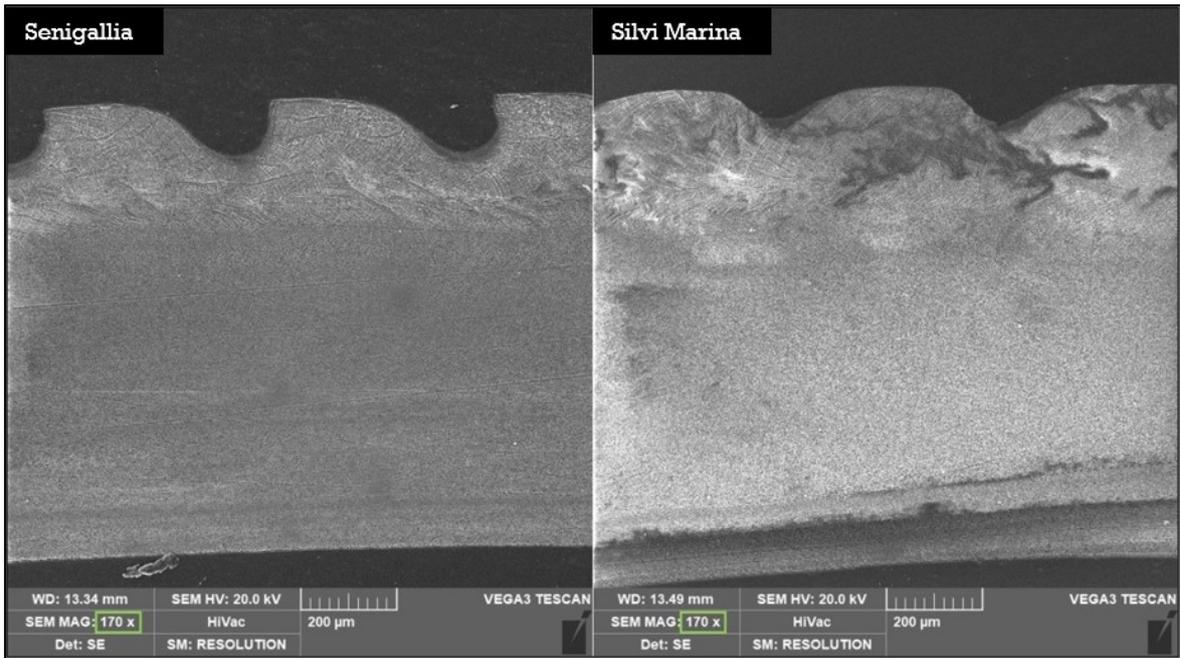


Fig. 33. Differenze nella struttura della conchiglia tra le vongole di Senigallia e Silvi Marina.

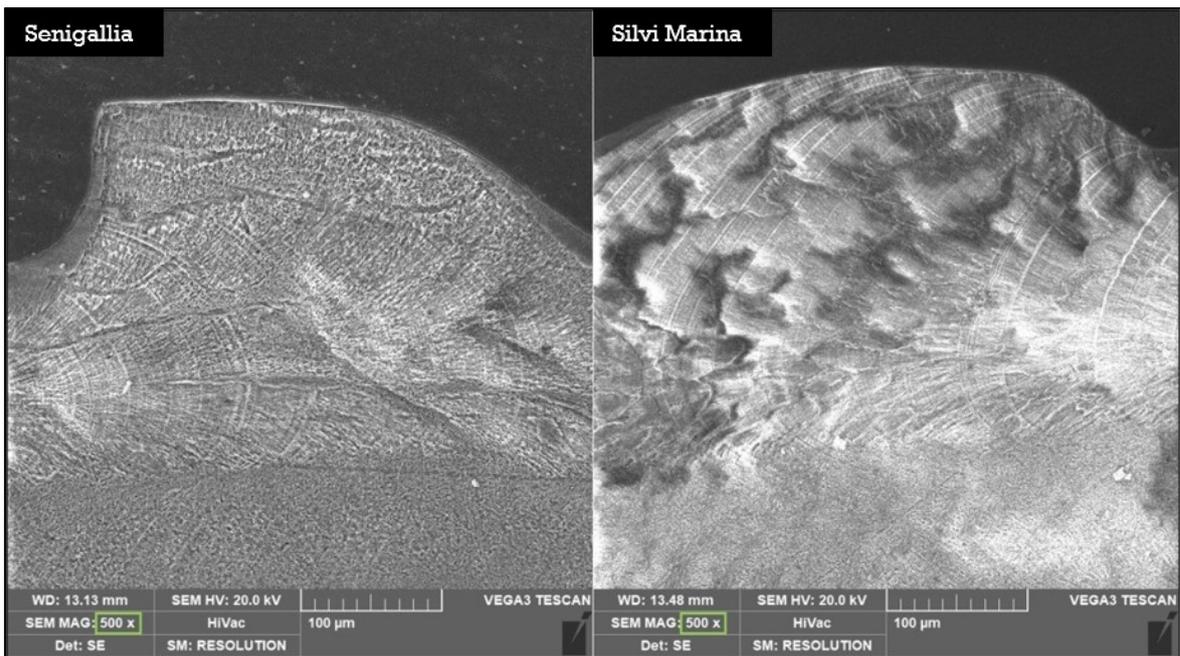


Fig. 34. Differenze nella struttura della conchiglia tra le vongole di Senigallia e Silvi Marina.

3.7.2 Microanalisi

I risultati delle microanalisi, effettuate per ottenere la composizione chimica del guscio e le proporzioni relative degli elementi presenti, sono coerenti tra loro e mostrano per tutte le aree la presenza esclusiva di carbonio, ossigeno, sodio e calcio. Le proporzioni sono pressoché analoghe (considerando la percentuale di errore nella rilevazione), e per i nostri scopi questa omogeneità composizionale si ritrova soprattutto confrontando le stesse aree in conchiglie provenienti dalle due zone di campionamento (Fig. 35, 36).

Smart Quant Results			
Element	Weight %	Atomic %	Error %
C K	7.09	12.62	6.87
O K	46.77	62.5	9.93
NaK	0.65	0.6	12.91
CaK	45.5	24.27	0.93

Fig. 35. Composizione delle valve delle vongole di Senigallia.

Smart Quant Results			
Element	Weight %	Atomic %	Error %
C K	7.42	13.02	6.85
O K	48	63.25	9.9
NaK	0.7	0.64	12.86
CaK	43.89	23.09	0.94

Fig. 36. Composizione delle valve delle vongole di Silvi Marina.

3.7.3 Diffrazione a raggi X

La diffrazione a raggi X non ha rilevato la presenza di polimorfi diversi del CaCO_3 , ovvero, gli spettri ottenuti da entrambe le valve (ognuna da un sito) sono prettamente aragonitici e quindi sovrapponibili tra loro.

Capitolo quarto

4. DISCUSSIONE

Lo scopo di questa tesi è osservare le eventuali differenze riguardanti la biomineralizzazione in individui di *Chamelea gallina* provenienti da due diversi siti lungo la costa adriatica. Per fare questo sono stati utilizzati gli strumenti dell'analisi trascrittomica e della microscopia elettronica a scansione per ottenere informazioni rispettivamente sul livello di espressione dei geni del mantello e sulla microstruttura e composizione interna della conchiglia.

Per poter giungere a conclusioni affidabili è essenziale partire da un materiale di più alta qualità possibile e che quest'ultima si mantenga nei vari passaggi di elaborazione dei campioni e dei dati. Abbiamo infatti già visto, per quanto riguarda il materiale di partenza, cioè l'RNA, l'ottenimento di elevati valori dell'indice di integrità dell'RNA (RIN), che hanno attestato un'ottima riuscita della metodica di estrazione e soprattutto assicurato materiale biologico qualitativamente affidabile da cui partire con le analisi. Nella fase di sequenziamento è stata mantenuta un'ottima qualità, infatti tutte le reads hanno registrato sia un alto Phred Score che una bassissima percentuale di basi ambigue. È stato inoltre ottenuto un miglioramento grazie al *trimming*, che ha rimosso le sequenze adattatrici, le duplicazioni terminali e le basi ambigue

terminali, favorendo un miglioramento del livello generale delle sequenze, ad esempio bilanciando il contributo nucleotidico, alzando il Phred Score nella porzione finale delle reads e diminuendo la quantità di basi ambigue. Questo, in fase di assemblaggio, è stato fondamentale per l'ottenimento di un trascrittoma di buona qualità. Infatti, il contenuto in GC di 38,3% (somma delle due percentuali delle singole basi) è in linea con quanto ritrovato in altri assemblaggi fatti sui bivalvi (Philipp et al., 2012; Takeuchi et al., 2012; Bjanmark et al., 2016). Tuttavia, la percentuale di basi ambigue, rispetto a quella del singolo file, risulta leggermente maggiore (ma comunque molto bassa) e questo potrebbe essere dovuto sia all'unione di porzioni con basi ambigue durante la formazione dei contig sia al riempimento delle regioni gap. Osservando le lunghezze dei contig, e considerandoli come trascritti (essendo un trascrittoma), un N50 di 422 bp appare basso se confrontato con quelli ritrovati da Bjanmark e colleghi (2016) e da Philipp e colleghi (2012) in cui si osservavano rispettivamente valori di 1.086 bp e 767 bp. Questa differenza potrebbe essere dovuta all'alto polimorfismo allelico, che ostacola la formazione di contig di lunghezza maggiore (Takeuchi et al., 2012). La validazione dell'assemblaggio fatta con BUSCO ha evidenziato una percentuale alta (70%) di gruppi ortologi mancanti rispetto a quelli presenti nel database che potrebbe far supporre una bassa completezza dei dati ottenuti.

Questo risultato dovrebbe essere letto alla luce di un elemento importante, ovvero la limitatezza delle risorse molecolari disponibili in letteratura per i molluschi ed in particolare per *C. gallina*. Questo potrebbe avere delle ripercussioni nella valutazione BUSCO. Infatti, utilizzando gruppi di geni ortologhi, la mancanza nel database di sequenze di specie filogeneticamente vicine alla nostra (o della stessa *C. gallina*) rischierebbe di non far riconoscere uno specifico contig come match valido, alzando di conseguenza la percentuale dei gruppi ortologhi mancanti, nonostante in realtà ce ne siano di più rispetto a quelli indicati. La valutazione BUSCO effettuata sul trascrittoma di *C. gallina* ottenuto da Coppe e colleghi (2012), ritenuto finora la risorsa molecolare di riferimento per questa specie, ha ottenuto approssimativamente il 4% di match completi, l'1% di frammentati e il restante 95% di mancanti, indici di dati qualitativamente peggiori rispetto a quelli ottenuti in questo lavoro di tesi.

Per avere un quadro di riferimento in cui inserire e discutere i risultati qui ottenuti, è necessario considerare i dati finora presenti in letteratura, che forniscono un contesto dentro cui orientarsi e da cui trarre analogie e differenze. Riprendendo ed ampliando quanto detto nel capitolo introduttivo, Gizzi e colleghi (2016) riportavano che le conchiglie degli individui di *C. gallina* delle aree più calde dell'Adriatico, dove c'è anche una maggiore salinità e una minore disponibilità di cibo, sono più leggere, sottili, porose e meno resistenti

alle fratture rispetto a quelle presenti a maggiori latitudini. Malgrado queste differenze, però, lungo il gradiente latitudinale non si trovano variazioni nel polimorfo del CaCO_3 o nei parametri composizionali. La maggiore fragilità della conchiglia, osservata nelle aree più calde, è stata attribuita a cambiamenti nel bilancio energetico dei processi fisiologici, con la conseguenza di poter allocare una minore energia per la formazione della conchiglia. Inoltre, un incremento della radiazione solare potrebbe influenzare la densità e la distribuzione del fitoplancton, diminuendo quindi la disponibilità dei nutrienti per i bivalvi (Häder et al., 2007). Un trend contrario a quello appena descritto è però stato riportato in altri studi, nei quali ad una bassa temperatura superficiale delle acque è associata una maggiore solubilità di CaCO_3 nell'acqua (Watson et al., 2012; Vermeji, 1993) e ciò incrementerebbe il costo della formazione della conchiglia, diminuendo lo stato di saturazione del CaCO_3 (Clarke, 1993). Per quanto riguarda invece l'influenza della salinità, una salinità sub-ottimale è un fattore di stress che porta sia ad un aumento dei costi per le difese da stress ossidativo che a problemi nel feeding, e di conseguenza a minori tassi di crescita. Inoltre, una bassa salinità riduce lo stato di saturazione dell'aragonite (unico polimorfo ritrovato nella conchiglia di *C. gallina*), con un aumento dei costi della calcificazione (Carregosa, 2014; Monari et al., 2007; Siedlecki et al., 2017; Sanders et al., 2018; Gruffydd et al.,

1984; Navarro, 1988; Kristensen, 1956; Maeda-Martinez, 1987; Hiebenthal et al., 2011; Wanamaker et al., 2007; Marsden & Pilkington, 1995; Almada-Villela, 1984). Nell'articolo pubblicato da Mancuso e colleghi (2019), se si osservano i parametri scheletrici delle conchiglie di taglia commerciale, si notano trend molto simili a quelli ritrovati per la stessa categoria di taglia da Gizzi e colleghi (2016), dove la porosità aumenta all'aumentare della temperatura superficiale delle acque, contrariamente quindi alla densità della conchiglia e ciò riporterebbe ad associare una maggiore porosità e fragilità ai siti a bassa latitudine nell'Adriatico. Questi trend generali, apparentemente in contraddizione tra loro, devono essere contestualizzati considerando anche l'influenza dei fiumi, che apportano modifiche alla salinità e alla disponibilità dei nutrienti (Vernberg, 1976, Widdows, 1985, Marcus et al., 1989, Widdows et al. 1990, Roper et al., 1991). Le discrepanze osservate nei vari studi citati, riguardanti la correlazione tra la densità della conchiglia (a cui sono collegati altri parametri scheletrici) e determinati parametri ambientali, suggeriscono che la formazione di questa importante struttura della vongola non dipende semplicemente dalle caratteristiche fisiche dell'ambiente, ma più probabilmente da una complessa rete in cui fattori fisici e biologici interagiscono tra loro nella definizione dei parametri scheletrici finali della conchiglia stessa.

Le analisi sui geni differenzialmente espressi (DEGs) condotte sui dati trascrittomici ottenuti dalle vongole di Silvi Marina e Senigallia ci permettono di raggruppare i DEGs di ogni sito all'interno di categorie basate sulle loro caratteristiche comuni. Per quanto riguarda i geni up-regolati di Silvi Marina, Neuromedin-U receptor 2-like, Zinc transporter ZIP14-like, Rho GTPase, Serum amyloid A-5 protein-like, Retinal dehydrogenase 1 sono geni che nei vertebrati sono stati comunemente visti essere coinvolti nel processo di osteogenesi (Aydemir & Cousins, 2018; Strzelecka-Kiliszek & Bożycki, 2017; Ebert et al., 2015; Nallamshetti et al., 2013). È da notare che l'osteogenesi nei vertebrati e la biomineralizzazione nei molluschi presentano diversi punti in comune riguardanti la composizione della matrice, le interazioni al suo interno e varie proteine effettrici. Infatti, nel lavoro di Takami e colleghi (2013) la proteina morfogenetica ossea BMP-2 dimostra di conservare le sue proprietà di induzione della formazione di tessuti duri sia nei vertebrati che negli invertebrati come i molluschi. Questo gruppo è quindi associato in maniera indiretta alla biomineralizzazione, ma è comunque importante da considerare alla luce della letteratura disponibile. Invece, più direttamente associati alla biomineralizzazione dei molluschi, sono i geni Sarcoplasmic calcium-binding protein e Alkaline phosphatase, rispettivamente con funzioni di storage del calcio (Shi et al., 2013; Miyamoto et al., 1996; Huang et al., 2007) e di

interazione con le proteine di matrice del guscio (Zhang et al., 2006b). In base a quanto visto, è lecito supporre una attività di biomineralizzazione in atto negli individui di *Silvi Marina*.

Prendendo ora in considerazione i geni up-regolati di *Senigallia*, abbiamo anche qui una coppia di geni coinvolti nello sviluppo osseo dei vertebrati (Toan et al., 2020), ovvero il Fos-related antigen 2 e l'Early growth response protein 1-B-like. Inoltre, sono stati trovati geni altamente espressi nei tessuti dei molluschi coinvolti nella biomineralizzazione (Sleight et al., 2015; Nemoto et al., 2019), come la 60S ribosomal protein L4-like e il Cytochrome c oxidase subunit 1. Infine, ritroviamo due geni fortemente connessi al processo di biomineralizzazione, ovvero la Calmodulina, componente di matrice del guscio che induce la nucleazione dell'aragonite e che regola la crescita della calcite nello strato prismatico (Yan et al., 2007), e il Transcription factor AP-1-like, fattore di trascrizione il cui omologo in *Pinctada fucata*, nel lavoro di Zheng e colleghi (2016), è stato visto essere necessario per l'attivazione dei promotori di molte proteine di matrice del guscio come N19, KRMP, perlina, prisilkina-39 ed è inoltre coinvolto nei processi di rigenerazione/riparazione del guscio. Anche per *Senigallia*, considerando quanto riportato, si può sostenere che siano in atto processi associati alla biomineralizzazione ed in particolare alla rigenerazione/riparazione della conchiglia.

I risultati genetici sono stati messi in relazione anche con quelli ottenuti dalle prove effettuate al SEM, che forniscono un'immagine abbastanza chiara delle differenze interne alla conchiglia degli individui dei due siti. Infatti, osservando le sezioni di conchiglia si può facilmente notare come negli individui di Silvi Marina la zona superiore della conchiglia, cioè a livello delle creste, presenti molte zone scure allungate indicanti una maggiore porosità nella conchiglia di questi individui. Inoltre, sembra che le vongole di Silvi Marina presentino una maggiore estensione lineare che fa ipotizzare che sia una conchiglia in crescita o che vengano sacrificate la durezza e la compattezza della conchiglia per conseguire una maggiore estensione lineare. Infine, ad un ingrandimento di 5000x la microstruttura degli individui di Silvi Marina sembra essere più sfogliata/trabecolare, sempre nella zona superiore della sezione. Il resto della sezione invece presenta caratteristiche perlopiù analoghe tra gli individui dei due siti. Dalle osservazioni al SEM mediante detector ad elettroni retrodiffusi (BSE), le caratteristiche tra i due siti sono simili, mostrando la direzione del fronte di deposizione e l'omogeneità della composizione interna dei vari strati. Questa omogeneità è anche supportata sia dalla diffrazione a raggi X, in cui non si ritrovano polimorfi del CaCO_3 diversi dall'aragonite (come ad esempio la vaterite), con un risultato analogo a quello evidenziato da Gizzi e colleghi (2016). Infatti, la microanalisi di diverse aree lungo i tre strati non hanno

rilevato differenze chimiche nella stechiometria degli elementi evidenziando che nelle conchiglie di entrambi i siti si ritrovano esclusivamente carbonio, ossigeno, calcio e sodio e all'incirca nelle stesse proporzioni. Dunque, l'unità fondamentale che viene prodotta durante la biomineralizzazione e con cui è costruita la conchiglia, cioè l'aragonite, sembra non essere influenzata dalle diverse condizioni ambientali che caratterizzano i due siti e a cui si trova sottoposta questa specie.

Considerando quanto detto finora, i risultati ottenuti sono in accordo con quelli ritrovati nei precedenti lavori di Gizzi e colleghi (2016) e Mancuso e colleghi (2019), che hanno tenuto conto dei parametri ambientali registrati lungo la costa adriatica; e sembrerebbe cioè che nel mare Adriatico la porosità della conchiglia aumenti con l'aumento della temperatura delle acque, della radiazione solare, della salinità e con la diminuzione della clorofilla. Non bisognerebbe però supporre che l'unico parametro indicante una mineralizzazione efficiente sia la densità, infatti, anche l'estensione lineare è un parametro scheletrico, e i due sembrano non essere associati tra loro (come si vede nel lavoro di Mancuso e colleghi, 2019), ovvero l'investimento nella biomineralizzazione sembrerebbe poter esser fatto principalmente verso uno dei due, sacrificando in parte l'altro. I due parametri scheletrici, infatti, si associano in modo opposto all'aumento o alla diminuzione dei diversi

parametri ambientali, ovvero le conchiglie prelevate a minore latitudine hanno una maggiore estensione lineare (che probabilmente beneficia della salinità e della temperatura), mentre quelle presenti a latitudini maggiori hanno una minore porosità e quindi una maggiore densità (che probabilmente beneficia della maggiore presenza di clorofilla).

Quindi per gli individui di Silvi Marina (sottoposti ad un ambiente a maggiore salinità, maggiore temperatura delle acque e minore clorofilla), che presentano una conchiglia più grande ma allo stesso tempo più porosa, si potrebbe pensare che il sacrificio di una parte della durezza per una maggiore estensione sia indirizzato alla creazione di uno spazio interno maggiore che possa contenere gonadi più grandi e quindi anche aumentare il numero di gameti emessi durante il periodo riproduttivo, con chiare ripercussioni sul successo riproduttivo. Questo comportamento è inoltre in accordo con quanto ritrovato nel lavoro di Carducci e colleghi (2020), in cui l'analisi trascrittomiche operata partendo dalle ghiandole digestive di *C. gallina* e quella istologica sulle gonadi ha mostrato che, nel periodo primaverile in cui è stato effettuato il campionamento, gli individui di Silvi Marina non presentano processi attivi riguardanti la produzione dei gameti permettendo di allocare risorse nella estensione lineare della conchiglia, ampliando di conseguenza il corpo dell'organismo e con esso le gonadi.

Al contrario, per le vongole di Senigallia, sottoposte a minore salinità, minore temperatura delle acque e maggiore clorofilla, si osservano conchiglie leggermente più piccole ma più dense, con una microstruttura più compatta e con meno depressioni/pori. Nel periodo campionato questi esemplari sono già in riproduzione come anche evidenziato da Carducci e colleghi (2020), suggerendo che queste vongole riescono ad investire meno energia per la formazione della conchiglia. È inoltre dimostrato che i bivalvi e gli altri organismi calcificanti rispondono alla pressione predatoria attraverso modifiche dei parametri scheletrici del guscio, che comprendono l'aumento dello spessore e quello della densità (Beadman et al., 2003; Leonard et al., 1999; Arneri et al., 1998). Quindi, nonostante si sappia poco dei predatori di *C. gallina* nell'Adriatico (Froggia, 1989), è lecito supporre che la radiazione solare e la temperatura delle acque influiscano sulla pressione predatoria, contribuendo a definire in un dato luogo la densità e la presenza di determinati tipi di predatori utilizzando tecniche specifiche (ad esempio rottura, foratura del guscio, ecc). In presenza di una determinata pressione predatoria, sarebbe quindi plausibile la creazione di un guscio più resistente per opporre resistenza. Questo sarebbe anche in accordo con il fatto di aver rivelato la trascrizione di geni coinvolti in processi destinati alla riparazione ed al rinforzo del guscio.

Queste diverse ipotesi, nonostante siano ragionevolmente possibili, non pretendono di fornire una soluzione esaustiva della complessa rete che vede i parametri ambientali e quelli biologici intrecciarsi tra loro per produrre il risultato finale osservato nelle conchiglie.

Sicuramente le due tipologie di conchiglie hanno un diverso impatto economico sulla pesca di questa specie. Infatti, considerando le zone a più bassa latitudine, la presenza di una struttura più porosa a cui è associata una minore durezza, come dimostrato dalle analisi sul carico di frattura di Gizzi e colleghi (2016), porterebbe ad una maggiore frazione di vongole scartata perché danneggiata dalla pesca tramite draga idraulica. D'altra parte, però, la maggiore estensione lineare della conchiglia permetterebbe il raggiungimento di una taglia commerciale in tempi più brevi.

Capitolo quinto

5. CONCLUSIONE

I risultati ottenuti in questa tesi evidenziano che le vongole dei due siti presentano differenze nell'espressione dei geni coinvolti nella biomineralizzazione e nella struttura della conchiglia che rimane invariata per quanto riguarda la composizione in quanto fatte di aragonite. Questo potrebbe essere dovuto ad un diverso investimento in termini di energia che gli organismi dei due siti fanno anche in relazione alla disponibilità di cibo e al periodo riproduttivo. Infatti, sembra evidente che le vongole del sito marchigiano investano in tempi più brevi nella gametogenesi a discapito della grandezza della conchiglia. Le vongole del sito abruzzese si riproducono tardivamente rispetto a quelle di Senigallia e pertanto investono di più nella crescita della conchiglia che essendo più porosa risulta però anche più fragile. Questo suggerisce due aspetti negativi per la pesca di questa risorsa: le vongole di Senigallia sembrano impiegare più tempo per arrivare alla taglia commerciale mentre quelle di Silvi Marina pur raggiungendo prima una taglia commerciale appropriata possono essere maggiormente danneggiate dall'attività di pesca.

I dati qui ottenuti aggiungono un tassello importante nella comprensione e dell'impatto che i vari fattori hanno sul fenotipo della conchiglia e in che modo

si legano tra loro andando a fornire informazioni importanti in termini di biologia della conservazione, per la tutela di questa e di altre specie affini. Attività volte alla gestione controllata della risorsa forniranno un sostegno al settore economico basato su questa specie, garantendone al contempo una maggiore protezione.

Capitolo sesto

6. BIBLIOGRAFIA

Addadi, L., et al. A chemical model for the cooperation of sulfates and carboxylates in calcite crystal nucleation: relevance to biomineralization. *Proc. Natl. Acad. Sci. U.S.A.* 84, 2732-2736 (1987).

Addadi, L., Joester, D., Nudelman, F. & Weiner, S. Mollusk shell formation: a source of new concepts for understanding biomineralization processes. *Chemistry* 12, 980-987 (2016).

Aleman, J. Comparaison morphologique de la structure de la conchille et de l'enroulement en spirale chez *Chamele gallina* et chez *Venus verrucosa*. *Arch. Anat. Micro. Morphol. Expèr.* 61-74 (1986).

Alhejoj, I., Bandel, K. & Salameh, E. Aquatic mollusks: occurrences, identification and their use as bioindicators of environmental conditions (salinity, trace elements and pollution parameters) in Jordan. Springer International Publishing, Cham, 295-318 (2017).

Almada-Villela, P. C. The effects of reduced salinity on the shell growth of small *Mytilus edulis*. *J. Mar. Biol. Assoc. United Kingdom* (1984).

Arivalagan, J., et al. Insights from the Shell Proteome: Biomineralization to Adaptation. *Mol. Biol. Evol.* 1, 66-77 (2017).

Arneri, E., Giannetti, G. & Antolini, B. Age determination and growth of *Venus verrucosa* L. (Bivalvia: Veneridae) in the southern Adriatic and the Aegean Sea. *Fish. Res.* 38, 193-198 (1998).

Austad, S. N. The uses of intraspecific variation in aging research. *Exp. Gerontol* (1996).

Aydemir, T. B. & Cousins, R. J. The Multiple Faces of the Metal Transporter ZIP14 (SLC39A14). *J. Nutr.* 148(2), 174-184 (2018).

Beadman, H., Caldow, R., Kaiser, M. & Willows, R. How to toughen up your mussels: using mussel shell morphological plasticity to reduce predation losses. *Mar. Biol.* 142, 487-494 (2003).

Bers, A. V., et al. Chemical defence in mussels: antifouling effect of crude extracts of the periostracum of the blue mussel *Mytilus edulis*. *Biofouling* 22, 251-259 (2006).

- Bertucci, A., et al. Transcriptomic responses of the endangered freshwater mussel *Margaritifera margaritifera* to trace metal contamination in the Dronne River France. *Environ. Sci. Pollut. R.* 24, 27145-27159 (2017).
- Bianchi, C. N. & Morri, C. Marine biodiversity of the Mediterranean Sea: situation, problems and prospects for future research. *Mar. Pollut. Bull.* 40, 367-376 (2000).
- Björnmark, N. A., et al. Transcriptomics provides insight into *Mytilus galloprovincialis* (Mollusca: Bivalvia) mantle function and its role in biomineralisation. *Mar. Genom.* 27, 37-45 (2016).
- Blackwelder, P. L., Weiss, R. E. & Wilbur, K. M. Effects of calcium, strontium, and magnesium on the coccolithophorid *Cricosphaera (Hymenomonas) carterae*. I. Calcification. *Mar. Biol.* 34, 11-16 (1976).
- Calvoiglesias, J., et al. Characterization of a monoclonal antibody directed against *Mytilus* spp larvae reveals an antigen involved in shell biomineralization. *PLoS One* 11:e0152210 (2016).
- Carducci, F., et al. Omics approaches for conservation biology research on the bivalve *Chamelea gallina*. *Sci. Rep.* 10, 19177 (2020).
- Carregosa, V., et al. Physiological and biochemical responses of three Veneridae clams exposed to salinity changes. *Comp. Biochem. Physiol. Part B Biochem. Mol. Biol.* 177-178, 1-9 (2014).
- Carter, J. Classification and phylogenetic significance of molluscan shell microstructure. *Mollusks* 13, 50-71 (1985).
- Casali, C. Résumé des paramètres biologiques sur *Venus gallina* L. en Adriatique (Synopsis of biological data on *Venus gallina* L. in the Adriatic Sea). *FAO Fish. Rep.* 290, 171-173 (1984).
- Cattaneo, P. & Bernardi, C. Molluschi bivalvi vivi ed echinodermi, tunicati e gasteropodi marini vivi. *Food In* N°1, 102 (2008).
- Chateigner, D., Hedegaard, C. & Wenk, H. R. Mollusc shell microstructures and crystallographic textures. *J. Struct. Geol.* 22, 1723-1735 (2000).
- Cho, S. M. & Jeong, W. G. Prismatic shell repairs by hemocytes in the extrapallial fluid of the Pacific Oyster, *Crassostrea gigas*. *Korean J. Malacol.* 27, 223-228 (2011).
- Clark, M. S., et al. Deciphering mollusc shell production: the roles of genetic mechanisms through to ecology, aquaculture and biomimetics. *Biol. Rev.* 95, 1812-1837 (2020).
- Clarke, A. Temperature and extinction in the sea: a physiologist's view. *Paleobiology.* 19, 499-518 (1993).

- Claxton, W. T., Wilson, A. B., Mackie, G. L. & Boulding, E. G. A genetic and morphological comparison of shallow and deepwater populations of the introduced dreissenid bivalve *Dreissena bugensis*. *Can. J. Zool.* 76, 1269-1276 (1998).
- Coppe, A., et al. Sequencing and characterization of striped venus transcriptome expand resources for clam fishery genetics. *PLoS One* 7(9):e44185 (2012).
- Cusack, M. & Freer, A. Biomineralization: elemental and organic influence in carbonate systems. *Chem. Rev.* 108, 4433-4454 (2008).
- Dalgiç, G., Okumuş, I. & Karayücel, S. The effect of fishing on growth of the clam *Chamelea gallina* (Bivalvia: Veneridae) for the Turkish Black Sea coast. *J. Mar. Biol. Assoc. UK* 90, 261-265 (2009).
- Delgado, M., Silva, L. & Juárez, A. Aspects of reproduction of striped venus *Chamelea gallina* in the Gulf of Cádiz (SW Spain): implications for fishery management. *Fish. Res.* 146, 86-95 (2013).
- Denny, M. W. & Blanchette, C. A. Hydrodynamics, shell shape, behavior and survivorship in the owl limpet *Lottia gigantea*. *J. Exp. Biol.* 203, 2623-2639 (2000).
- DeWitt, T. J. & Scheiner, S. M. Phenotypic Plasticity: Functional and Conceptual Approaches. Oxford Univ. Press, Oxford, UK (2004).
- Ebert, R., et al. Acute phase serum amyloid A induces proinflammatory cytokines and mineralization via toll-like receptor 4 in mesenchymal stem cells. *Stem Cell Res.* 15(1), 231-239 (2015).
- Eizaguirre, C. & Baltazar-Soares, M. Evolutionary conservation-evaluating the adaptive potential of species. *Evol. Appl.* 7, 963-967 (2014).
- Ezgeta-Balic, D. et al. An energy budget for the subtidal bivalve *Modiolus barbatus* (Mollusca) at different temperatures. *Mar. Environ. Res.* 71, 79-85 (2011).
- Falini, G., Albeck, S., Weiner, S. & Addadi, L. Control of aragonite or calcite polymorphism by mollusk shell macromolecules. *Science* 271, 67-69 (1996).
- Feng, Q., et al. The structure-function relationship of MSI7, a matrix protein from pearl oyster *Pinctada fucata*. *Acta Biochim. Biophys. Sin.* 41, 955-962 (2009).
- Fitzer, S. C., Phoenix, V. R., Cusack, M. & Kamenos, N. A. Ocean acidification impacts mussel control on biomineralisation. *Sci. Rep.* 4, 6218 (2014).
- Fleury, C., et al. Shell repair process in the green ormer *Haliotis tuberculata*: a histological and microstructural study. *Tissue Cell* 40, 207-218 (2008).

- Froggia, C. Aspetti biologici, tecnologici e statistici della pesca delle vongole (*Venus gallina*) (Biological, technological and statistical observations on the fishery targeting common clams, *Venus gallina*). Incontri Tecnici, Laboratorio di Tecnologia della Pesca, Consiglio Nazionale delle Ricerche. 9, 7-22 (1975).
- Froggia, C. Clam fisheries with hydraulic dredges in the Adriatic Sea in Marine Invertebrates Fisheries: Their Assessment and Management. (ed. Caddy, J. F.) 507-524 (Wiley, New York, 1989).
- Fuiman, L. A., Gage, J. D. & Lamont, P. A. Shell morphometry of the deep sea protobranch bivalve *Ledella pustulosa* in the Rockall Trough, North-East Atlantic. J. Mar. Biol. Ass. UK 79, 661-671 (1999).
- Furuhashi, T. et al. Pyrolysis GC/MS and IR spectroscopy in chitin analysis of molluscan shells. Bioscience, biotechnology, and biochemistry 73, 93-103 (2009).
- Furuhashi, T., et al. Molluscan shell evolution with review of shell calcification hypothesis. Comp. Biochem. Physiol. Biochem. Mol. Biol. 154, 351-371 (2009).
- Gardella, S., et al. The nuclear protein HMGB1 is secreted by monocytes via a non-classical, vesicle-mediated secretory pathway. EMBO Rep. 3, 995-1001 (2002).
- Gaspar, M. B. & Monteiro, C. C. Reproductive cycles of the razor clam *Ensis siliqua* and the clam *Venus striatula* off Vilamoura Southern Portugal. J. Mar. Biol. Assoc. U.K. 78, 1247-1258 (1998).
- Gaspar, M. B., Pereira, A. M., Vasconcelos, P. & Monteiro, C. C. Age and growth of *Chamelea gallina* from the Algarve coast (Southern Portugal): influence of seawater temperature and gametogenic cycle on growth rate. J. Molluscan Stud. 70, 371-377 (2004).
- Ghiselli, F., et al. Comparative transcriptomics in two bivalve species offers different perspectives on the evolution of sex-biased genes. Genome Biol. Evol. 10, 1389-1402 (2018).
- Gilbert, S. F. Ecological developmental biology: developmental biology meets the real world. Dev. Biol. 233, 1-12 (2001).
- Gizzi, F. Reproduction, growth and biomineralization of calcifying marine organisms and their relationships with environmental parameters. Alma Mater Studiorum Università di Bologna. Dottorato di ricerca in Biodiversità ed evoluzione, 28 Ciclo (2016).
- Gizzi, F., et al. Shell properties of commercial clam *Chamelea gallina* are influenced by temperature and solar radiation along a wide latitudinal gradient. Sci. Rep. 6, 36420 (2016).

- Gong, N., et al. Characterization of calcium deposition and shell matrix protein secretion in primary mantle tissue culture from the marine pearl oyster *Pinctada fucata*. *Mar. Biotechnol.* 10, 457-465 (2008).
- Gong, N., et al. Immunolocalization of matrix proteins in nacre lamellae and their in vivo effects on aragonitic tablet growth. *J. Struct. Biol.* 164, 33-40 (2008).
- González-Domínguez, et al. Combined proteomic and metallomic analyses in *Scrobicularia plana* clams to assess environmental pollution of estuarine ecosystems. *Mar. Pollut. Bull.* 113, 117-124 (2016).
- Gotliv, B. A., et al. Asprich: a novel aspartic acid-rich protein family from the prismatic shell matrix of the bivalve *Atrina rigida*. *Chembiochem.* 6, 304-314 (2005).
- Gruffydd, L. L. D., Huxley, R. & Crisp, D. J. The reduction in growth of *Mytilus edulis* in fluctuating salinity regimes measured using laser diffraction patterns and the exaggeration of this effect by using tap water as a diluting medium. *J. Mar. Biol. Assoc. UK* 64, 401-409 (1984).
- Häder, D. P., Kumar, H. D., Smith, R. C. & Worrest, R. C. Effects of solar UV radiation on aquatic ecosystems and interactions with climate change. *Photochem. Photobiol. Sci.* 6, 267-285 (2007).
- Hanein, D., Geiger, B. & Addadi, L. Fibronectin adsorption to surfaces of hydrated crystals. An analysis of the importance of bound water in protein-substrate interactions. *Langmuir.* 9, 1058-1065 (1993).
- Hare, P. E. Amino acids in the proteins from aragonite and calcite in the shells of *Mytilus californianus*. *Science* 139, 216-217 (1963).
- Hattan, S. J., Laue, T. M. & Chasteen, N. D. Purification and characterization of a novel calcium-binding protein from the extrapallial fluid of the mollusc, *Mytilus edulis*. *J. Biol. Chem.* 276, 4461-4468 (2001).
- He, X., Johansson, M. L. & Heath, D. D. Role of genomics and transcriptomics in selection of reintroduction source populations. *Conserv. Biol.* 30, 1010-1018 (2016).
- Heilmayer, O., Brey, T. & Pörtner, H. O. Growth efficiency and temperature in scallops: a comparative analysis of species adapted to different temperatures. *Funct. Ecol.* 18, 641-647 (2004).
- Hessle, L., et al. Tissue-nonspecific alkaline phosphatase and plasma cell membrane glycoprotein-1 are central antagonistic regulators of bone mineralization. *Proc. Natl. Acad. Sci. U.S.A.* 99, 9445-9449 (2002).

- Hiebenthal, C., Philipp, E., Eisenhauer, A. & Wahl, M. Interactive effects of temperature and salinity on shell formation and general condition in Baltic Sea *Mytilus edulis* and *Arctica islandica*. *Aquat. Biol.* 14, 289-298 (2011).
- Huang, J., et al. A novel extracellular EF-hand protein involved in the shell formation of pearl oyster. *Biochim. Biophys. Acta* 1770(7), 1037-1044 (2007).
- Isaji, S. J. Formation of organic sheets in the inner shell layer of *Geloina* (Bivalvia, *Corbiculidae*) - An adaptive response to shell dissolution. *The Veliger* 36, 166-173 (1993).
- Ivanina, A. V., et al. Biomineralization-related specialization of hemocytes and mantle tissues of the Pacific oysters *Crassostrea gigas*. *J. Exp. Biol.* 220(Pt 18), 3209-3221 (2017).
- Ivanina, A. V., Kurochkin, I. O., Leamy, L. & Sokolova, I. M. Effects of temperature and cadmium exposure on the mitochondria of oysters (*Crassostrea virginica*) exposed to hypoxia and subsequent reoxygenation. *J. Exp. Biol.* 215, 3142-3154 (2012).
- Jablonski, D. & Lutz, R. A. Molluscan larval shell morphology - Ecological and paleontological applications. *Skeletal Growth of Aquatic Organisms*. Eds: DC Rhoads, RA Lutz, Plenum Press, New-York (1980).
- Jansen, J. M., et al. Geographic and seasonal patterns and limits on the adaptive response to temperature of European *Mytilus* spp. and *Macoma balthica* populations. *Oecologia*. 154, 23-34 (2007).
- Jiang, Y., et al. Recombinant transgelin-like protein 1 from *Mytilus* shell induces formation of CaCO₃ polymorphic crystals in vitro. *FEBS Open Bio.* 10(10), 2216-2234 (2010).
- Johnstone, M. B., Ellis, S. & Mount, A. S. Visualization of shell matrix proteins in hemocytes and tissues of the Eastern oyster, *Crassostrea virginica*. *J. Exp. Zool. B Mol. Dev. Evol.* 310, 227-239 (2008).
- Johnstone, M. B., et al. Cellular orchestrated biomineralization of crystalline composites on implant surfaces by the eastern oyster, *Crassostrea virginica* (Gmelin, 1791). *J. Exp. Mar. Biol. Ecol.* 463, 8-16 (2015).
- Kádár, E. Haemocyte response associated with induction of shell regeneration in the deep-sea vent mussel *Bathymodiolus azoricus* (Bivalvia: Mytilidae). *J. Exp. Mar. Biol. Ecol.* 362, 71-78 (2008).
- Kadar, E., Lobo-Da-Cunha, A. & Azevedo, C. Mantle-to-shell CaCO₃ transfer during shell repair at different hydrostatic pressures in the deep-sea vent mussel *Bathymodiolus azoricus* (Bivalvia: *Mytilidae*). *Mar. Biol.* 156, 959-967 (2009).

- Kanazawa, T. & Shinichi, S. Environmental and physiological controls on shell microgrowth pattern of *Ruditapes philippinarum* (Bivalvia: *Veneridae*) from Japan. *J. Molluscan Stud.* 74, 89-95 (2007).
- Keller, N., Del Piero, D. & Longinelli, A. Isotopic composition, growth rates and biological behaviour of *Chamelea gallina* and *Callista chione* in the Gulf of Trieste. *Mar. Biol.* 140, 9-15 (2002).
- Kim, H., et al. The role of nacreous factors in preventing osteoporotic bone loss through both osteoblast activation and osteoclast inactivation. *Biomaterials* 33(30), 7489-7496 (2012).
- Kinoshita, S., et al. Deep sequencing of ESTs from nacreous and prismatic layer producing tissues and a screen for novel shell formation-related genes in the pearl oyster. *PLoS One* 6:e21238 (2011).
- Kniprath, E. Ontogeny of the molluscan shell field. *Zoologica Scripta* 10, 61-79 (1981).
- Kong, J., et al. Alv protein plays opposite roles in the transition of amorphous calcium carbonate to calcite and aragonite during shell formation. *Cryst. Growth Des.* 7, 3794-3804 (2018).
- Kong, W., et al. Calcium carbonate mineralization mediated by in vitro cultured mantle cells from *Pinctada fucata*. *Biochem. Biophys. Res. Commun.* 7, 1053-1058 (2015).
- Kong, Y., et al. Cloning and characterization of Prsilkin-39, a novel matrix protein serving a dual role in the prismatic layer formation from the oyster *Pinctada fucata*. *J. Biol. Chem.* 284, 10841-10854 (2009).
- Kono, M., Hayashi, N. & Samata, T. Molecular mechanism of the nacreous layer formation in *Pinctada maxima*. *Biochem. Biophys. Res. Commun.* 269, 213-218 (2000).
- Kouchinsky, A. Shell microstructures in Early Cambrian molluscs. *Acta Palaeontol. Pol.* 45, 119-150 (2000).
- Kristensen, I. Differences in density and growth in a cockle population in the Dutch Wadden Sea. *Archives Neerlandaises de Zoologie* 12, 351-453 (1956).
- Kylmaoja, E., Nakamura, M. & Tuukkanen, J. Osteoclasts and remodeling based bone formation. *Curr. Stem Cell Res. Ther.* 11, 626-633 (2016).
- La Guardia, M. J., et al. In situ accumulation of HBCD, PBDEs, and several alternative flame-retardants in the bivalve (*Corbicula fluminea*) and gastropod (*Elimia proxima*). *Environ. Sci. Technol.* 46, 5798-5805 (2012).
- Leonard, G. H., Bertness, M. D. & Yund, P. O. Crab predation, waterborne cues, and inducible defenses in the blue mussel, *Mytilus edulis*. *Ecology.* 80, 1-4 (1999).

- Li, Q., Zhao, X., Khong, L. & Yu, H. Transcriptomic response to stress in marine bivalves. *Invert. Surviv. J.* 10, 84-93 (2013).
- Li, S., et al. Hemocytes participate in calcium carbonate crystal formation, transportation and shell regeneration in the pearl oyster *Pinctada fucata*. *Fish Shellfish Immunol.* 51, 263-270 (2016).
- Liao, Z., et al. In-depth proteomic analysis of nacre, prism, and myostracum of *Mytilus* shell. *J. Proteomics* 122, 26-40 (2015).
- Lowenstam, H. A. & Weiner, S. *On Biomineralization*. Oxford University Press (1989).
- Luchmann, K. H. et al. Biochemical biomarkers and hydrocarbons concentrations in the mangrove oyster *Crassostrea brasiliana* following exposure to diesel fuel water-accommodated fraction. *Aquat. Toxicol.* 105, 652-660 (2011).
- Ma, Z., et al. A novel extrapallial fluid protein controls the morphology of nacre lamellae in the pearl oyster, *Pinctada fucata*. *J. Biol. Chem.* 282, 23253-23263 (2007).
- Mable, B. K. Conservation of adaptive potential and functional diversity: integrating old and new approaches. *Conserv. Genet.* 20, 89-100 (2019).
- Maeda-Martinez, A. N. The rates of calcium deposition in shells of molluscan larvae. *Comp. Biochem. Physiol. Part A Physiol.* 86, 21-28 (1987).
- Mancuso, A., et al. Environmental influence on calcification of the bivalve *Chamelea gallina* along a latitudinal gradient in the Adriatic Sea. *Sci. Rep.* 9, 11198 (2019).
- Mann, K., Edsinger-Gonzales, E. & Mann, M. In-depth proteomic analysis of a mollusc shell: acid-soluble and acid-insoluble matrix of the limpet *Lottia gigantea*. *Proteome Sci.* 10, 28 (2012).
- Mann, K., et al. Perlinhibin, a cysteine-, histidine-, and arginine-rich miniprotein from abalone (*Haliotis laevis*) nacre, inhibits in vitro calcium carbonate crystallization. *Biophys. J.* 93, 1246-1254 (2007).
- Mann, S. Mineralization in biological systems. *Struct. Bond.* 54, 125-174 (1983).
- Mann, S. Molecular recognition in biomineralization. *Nature* 332, 119-124 (1988).
- Marcus, J. M., Scott, G. L. & Heizer, D. H. The use of oyster shell thickness and condition index measurements as physiological indicators of no heavy metal pollution around three coastal marinas. *J. Shellfish Res.* 8, 87-94 (1989).

- Marie, B., et al. Deep conservation of bivalve nacre proteins highlighted by shell matrix proteomics of the Unionoida *Elliptio complanata* and *Villosa lienosa*. *J. R. Soc. Interface* 14, 20160846 (2017).
- Marie, B., et al. Different secretory repertoires control the biomineralization processes of prism and nacre deposition of the pearl oyster shell. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20986-20991 (2012).
- Marie, B., et al. Nacre calcification in the freshwater mussel *Unio pictorum*: carbonic anhydrase activity and purification of a 95kDa calcium-binding glycoprotein. *Chem. Bio. Chem.* 9, 2515-2523 (2008).
- Marie, B., et al. Proteomic analysis of the organic matrix of the abalone *Haliotis asinina* calcified shell. *Proteome Sci.* 8, 54 (2010).
- Marie, B., et al. Unveiling the evolution of bivalve nacre proteins by shell proteomics of unionoidae. *Key Eng. Mater.* 672, 158-167 (2015).
- Marin, F., et al. Mucins and molluscan calcification Molecular characterization of mucoperlin, a novel mucin-like protein from the nacreous shell layer of the fan mussel *Pinna nobilis* (Bivalvia, Pteriomorpha). *J. Biol. Chem.* 275, 20667-20675 (2000).
- Marin, F., Le Roy, N., and Marie, B. The formation and mineralization of mollusk shell. *Front. Biosci.* 4, 1099-1125 (2012).
- Marin, F., Luquet, G., Marie, B. & Medakovic, D. Molluscan shell proteins: primary structure, origin, and evolution. *Curr. Top. Dev. Biol.* 80, 209-276 (2008).
- Marsden, I. & Pilkington, R. Spatial and temporal variations in the condition of *Austrovenus stutchburyi* Finlay, 1927 (Bivalvia: Veneridae). Avon- Heathcote estuary, Christchurch. *New Zealand Nat. Sci.* 22, 57-67 (1995).
- Matozzo, A., et al. Exposure to anoxia of the clam *Chamelea gallina*. Effects on immune responses. *J. Exp. Mar. Biol. Ecol.* 325, 163-174 (2005).
- Matozzo, V., et al. First evidence of immunomodulation in bivalves under seawater acidification and increased temperature. *PLoS One* 7(3) (2012).
- Michenfelder, M., et al. Characterization of two molluscan crystal-modulating biomineralization proteins and identification of putative mineral binding domains. *Biopolymers* 70, 522-533 (2003).
- Milan, M., et al. Ecotoxicological effects of the herbicide glyphosate in non-target aquatic species: transcriptional responses in the mussel *Mytilus galloprovincialis*. *Environ. Pollut.* 237, 442-451 (2018).

- Milan, M., et al. Host-microbiota interactions shed light on mortality events in the striped venus clam *Chamelea gallina*. *Mol. Ecol.* 28, 4486-4499 (2019).
- Milan, M. et al. Transcriptomic profiling of *Chamelea gallina* from sites along the Abruzzo coast (Italy), subject to periodic localized mortality events. *Mar. Biol.* 163, 163-169 (2016).
- Ministerial decree n.27 del 17/6/2019, Ministry of Agricultural Food, forestry, and Tourism policies. Adozione del Piano di gestione nazionale per le attività di pesca con il sistema draghe idrauliche e rastrelli da natante così come identificati nella denominazione degli attrezzi di pesca in draghe meccaniche comprese le turbosoffianti (HMD) e draga meccanizzata (DRB). *Gazzetta ufficiale Italiana* (2019).
- Miyamoto, H., et al. A carbonic anhydrase from the nacreous layer in oyster pearls. *Proc. Natl. Acad. Sci. U.S.A.* 93, 9657-9660 (1996).
- Miyamoto, H., et al. The diversity of shell matrix proteins: genome-wide investigation of the pearl oyster, *Pinctada fucata*. *Zoolog. Sci.* 30, 801-816 (2013).
- Miyamoto, H., Miyoshi, F. & Kohno, J. The carbonic anhydrase domain protein nacrein is expressed in the epithelial cells of the mantle and acts as a negative regulator in calcification in the mollusc *Pinctada fucata*. *Zool. Sci.* 22, 311-315 (2005).
- Miyashita, T., et al. Complementary DNA cloning and characterization of pearlins, a new class of matrix protein in the nacreous layer of oyster pearls. *Mar. Biotechnol.* 2, 409-418 (2000).
- Monari, M., et al. Effects of salinity on the clam *Chamelea gallina* haemocytes. Part II: Superoxide dismutase response. *Mar. Biol.* 151, 1059-1068 (2007).
- Monari, M., et al. Heat shock protein 70 response to physical and chemical stress in *Chamelea gallina*. *J. Exp. Mar. Biol. Ecol.* 397, 71-78 (2011).
- Morse, J. W., Arvidson, R. S. & Lüttge, A. Calcium carbonate formation and dissolution. *Chem. Rev.* 107, 342-381 (2007).
- Morton, B. "Bivalve". *Encyclopedia Britannica*, 13 (2020).
- Moschino, V. & Marin, M. G. Seasonal changes in physiological responses and evaluation of "well-being" in the Venus clam *Chamelea gallina* from the Northern Adriatic Sea. *Comp. Bioch. Physiol.* 145A, 433-440 (2006).
- Moschino, V., Deppieri, M. & Marin, M. G. Evaluation of shell damage to the clam *Chamelea gallina* captured by hydraulic dredging in the Northern Adriatic Sea. *ICES J. Mar. Sci.* 60(2), 393-401 (2003).
- Mount, A. & Pickering, C. M. Testing the capacity of clothing to act as a vector for non-native seed in protected areas. *J. Environ. Manage.* 91, 168-179 (2009).

- Mount, A. S., Wheeler, A., Paradkar, R. P. & Snider, D. Hemocyte-mediated shell mineralization in the eastern oyster. *Science* 304, 297-300 (2004).
- Nagai, K., Yano, M., Morimoto, K. & Miyamoto, H. Tyrosinase localization in mollusc shells. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 146, 207-214 (2007).
- Nallamshetty, S., et al. Deficiency of Retinaldehyde Dehydrogenase 1 Induces BMP2 and Increases Bone Mass In Vivo. *PLoS One* 8(8):e71307 (2013).
- Navarro, J.M. The effects of salinity on the physiological ecology of *Choromytilus choromytilus* chorus (Molina, 1762) (Bivalvia: Mytillidae). *J. Exp. Mar. Biol. Ecol.* 122, 19-33 (1988).
- Nemoto, M., et al. Integrated transcriptomic and proteomic analyses of a molecular mechanism of radular teeth biomineralization in *Cryptochiton stelleri*. *Sci. Rep.* 9, 856 (2019).
- Newell, C. R. & Hidu, H. The effects of sediment type on growth rate and shell allometry in the soft-shelled clam *Mya arenaria*. *L. J. Exp. Mar. Biol. Ecol.* 65, 285-295 (1982).
- Nishida, K., Ishimura, T., Suzuki, A. & Sasaki, T. Seasonal changes in the shell microstructure of the bloody clam, *Scapharca broughtonii* (Mollusca: Bivalvia: Arcidae). *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 363-364, 99-108 (2012).
- Noh, J., et al. Bioaccumulation of polycyclic aromatic hydrocarbons (PAHs) by the marine clam, *Macra veneriformis*, chronically exposed to oil-suspended particulate matter aggregates. *Environ. Sci. Technol.* 52, 7910-7920 (2018).
- Nudelman, F., Gotliv, B. A., Addadi, L. & Weiner, S. Mollusk shell formation: mapping the distribution of organic matrix components underlying a single aragonitic tablet in nacre. *J. Struct. Biol.* 153, 176-187 (2006).
- Orban, E., et al. Nutritional and commercial quality of the striped venus clam, *Chamelea gallina*, from the Adriatic Sea. *Food Chem.* 101, 1063-1070 (2007).
- Panayotou, G., et al. Domains of laminin with growth-factor activity. *Cell* 56, 93-101 (1989).
- Peck, L. S., Powell, D. K. & Tyler, P. A. Very slow development in two Antarctic bivalve molluscs, the infaunal clam *Laternula elliptica* and the scallop *Adamussium colbecki*. *Mar. Biol.* 150, 1191-1197 (2007).
- Philipp, E. E. et al. Massively parallel RNA sequencing identifies a complex immune gene repertoire in the lophotrochozoan *Mytilus edulis*. *PLoS One* 7:e33091 (2012).
- Pigliucci, M., Murren, C. J. & Schlichting, C. D. Phenotypic plasticity and evolution by genetic assimilation. *J. Exp. Boil.* 209, 2362-2367 (2006).

- Politi, Y., et al. Asprich mollusk shell protein: in vitro experiments aimed at elucidating function in CaCO₃ crystallization. *Crystengcomm* 9, 1171-1177 (2007).
- Popov, S. V. Composite prismatic structure in bivalve shell. *Acta Palaeontol. Pol.* 31, 3-28 (1986).
- Poppe, G. T. & Goto, Y. *European Seashells. Vol II (Scaphopoda, Bivalvia, Cephalopoda)*, 1-221 (Verlag Christa Hemmen, Germany, 1993).
- Purroy, A., et al. Drivers of shell growth of the bivalve, *Callista chione* (L. 1758) - Combined environmental and biological factors. *Mar. Environ. Res.* 134, 138-149 (2018).
- Raddatz, R., et al. Identification and characterization of two neuromedin U receptors differentially expressed in peripheral tissues and the central nervous system. *J Biol. Chem.* 275(42), 32452-32459 (2000).
- Ramón, M. & Richardson, C. A. Age determination and shell growth of *Chamelea gallina* (Bivalvia: *Veneridae*) in the western Mediterranean. *Mar. Ecol. Prog. Ser.* 89, 15-23 (1992).
- Richardson, C. A. Molluscs as archives of environmental change. *Oceanography and Marine Biology, An Annual Review*, Volume 39 (2001).
- Rittschof, D. & McClellan-Green, P. Molluscs as multidisciplinary models in environment toxicology. *Mar. Pollut. Bull.* 50, 369-373 (2005).
- Romanelli, M., Cordisco, C. A. & Giovanardi, O. The long-term decline of the *Chamelea gallina* L. (Bivalvia: *Veneridae*) clam fishery in the Adriatic Sea: is a synthesis possible? *Acta Adriat.* 50, 171-205 (2009).
- Roper, D.S., Pridmore, R.D., Cummings, V.J. & Hewitt, J.E. Pollution related differences in the condition cycles of Pacific oyster *Crassostrea gigas*. *Mar. Environ. Res.* 31, 197-214 (1991).
- Rucinski, M., et al. Neuromedin U directly stimulates growth of cultured rat calvarial osteoblast-like cells acting via the NMU receptor 2 isoform. *Int. J. Mol. Med.* 22(3), 363-368 (2008).
- Saha, A., Jana, T. & Choudhury, A. The extrapallial fluid of *Macoma birmanica*: an environment for calcium carbonate deposition. *Philipp. J. Crop Sci.* 117, 395-399 (1988).
- Samata, T., et al. A new matrix protein family related to the nacreous layer formation of *Pinctada fucata*. *FEBS Lett.* 462, 225-229 (1999).
- Sanders, T., Schmittmann, L., Nascimento-Schulze, J. C. & Melzner, F. High calcification costs limit mussel growth at low salinity. *Front. Mar. Sci.* 5, 352 (2018).
- Saruwatari, K., et al. Nucleation and growth of aragonite crystals at the growth front of nacles in pearl oyster, *Pinctada fucata*. *Biomaterials* 30, 3028-3034 (2009).

- Schluter, D. The Ecology of Adaptive Radiation. Oxford Univ. Press, Oxford, UK (2000).
- Shen, X., et al. Molecular cloning and characterization of lustrin A, a matrix protein from shell and pearl nacre of *Haliotis rufescens*. J. Biol. Chem. 272, 32472-32481 (1997).
- Shi, Y., et al. Characterization of the pearl oyster (*Pinctada martensii*) mantle transcriptome unravels biomineralization genes. Mar. Biotechnol. (NY). 15(2), 175-187 (2013).
- Shitalbahen, V. P. A Novel Function of Invertebrate Collagen in the Biomineralization Process During the Shell Repair of Eastern Oyster, *Crassostrea virginica*. Clemson, SC: Clemson University (2004).
- Siedlecki, S. A., et al. The Importance of Freshwater to Spatial Variability of Aragonite Saturation State in the Gulf of Alaska. J. Geophys. Res. Ocean. 122, 8482-8502 (2017).
- Sleight, V. A., et al. Characterisation of the mantle transcriptome and biomineralisation genes in the blunt-gaper clam, *Mya truncata*. Mar. Gen. 27, 47-55 (2016).
- Sleight, V. A., Thorne, M. A., Peck, L. S. & Clark, M. S. Transcriptomic response to shell damage in the Antarctic clam, *Laternula elliptica*: time scales and spatial localisation. Mar. Genomics 20, 45-55 (2015).
- Sobral, P. & Widdows, J. Influence of hypoxia and anoxia on the physiological response of the clam *Ruditapes decussatus* from southern Portugal. Mar. Biol. 127, 455-461 (1997).
- Song, X., Liu, Z., Wang, L. & Song, L. Recent Advances of Shell Matrix Proteins and Cellular Orchestration in Marine Molluscan Shell Biomineralization. Front. Mar. Sci. 6, 41 (2019).
- Stenflo, J., Stenberg, Y. & Muranyi, A. Calcium-binding EGF-like modules in coagulation proteinases: function of the calcium ion in module interactions. Biochim. Biophys. Acta 1477, 51-63 (2000).
- Stenzel, H. Aragonite and calcite as constituents of adult oyster shells. Science 142, 232-233 (1963).
- Strzelecka-Kiliszek, A. & Bożycycki, Ł. Cross-talk between Src kinases and Rho small GTPases regulates biomineralization and simplify imaging of the mineralization process. Postepy Biochem. 63(2), 93-109 (2017).
- Sudo, S., et al. Structures of mollusc shell framework proteins. Nature 387, 563-564 (1997).
- Suzuki, M., et al. An acidic matrix protein, Pif, is a key macromolecule for nacre formation. Science 325, 1388-1390 (2009).
- Suzuki, M., et al. Characterization of Prismaticin-14, a novel matrix protein from the prismatic layer of the Japanese pearl oyster (*Pinctada fucata*). Biochem. J. 382(Pt 1), 205-213 (2004).

- Suzuki, M., et al. Identification and characterisation of a calcium carbonate-binding protein, blue mussel shell protein (BMSP), from the nacreous layer. *Chembiochem* 12, 2478-2487 (2011).
- Suzuki, M., Sakuda, S. & Nagasawa, H. Identification of chitin in the prismatic layer of the shell and a chitin synthase gene from the Japanese pearl oyster, *Pinctada fucata*. *Biosci. Biotechnol. Biochem.* 71, 1735-1744 (2007).
- Takami, A., Kato, H., Takagi, R. & Miyashita, T. Studies on the *Pinctada fucata* BMP-2 Gene: Structural Similarity and Functional Conservation of Its Osteogenic Potential within the Animal Kingdom. *Int. J. Zool.* 2013, 787323 (2013).
- Takeuchi, T. & Endo, K. Biphasic and dually coordinated expression of the genes encoding major shell matrix proteins in the pearl oyster *Pinctada fucata*. *Mar. Biotech.* 8, 52-61 (2006).
- Takeuchi, T., et al. Draft genome of the pearl oyster *Pinctada fucata*: a platform for understanding bivalve biology. *DNA Res.* 19(2), 117-130 (2012).
- Teixeira, M., et al. Toxic effects of the antihistamine cetirizine in mussel *Mytilus galloprovincialis*. *Water Res.* 114, 316-326 (2017).
- Toan, N. K., Tai, N. C., Kim, S. A. & Ahn, S. G. Soluble Klotho regulates bone differentiation by upregulating expression of the transcription factor EGR-1. *FEBS Lett.* 594(2), 290-300 (2020).
- Tokuyama, A. Effects of organic matter on solubilities and crystal form of carbonates. *Am. Zool.* 9, 681-688 (1969).
- Treccani, L., Mann, K., Heinemann, F. & Fritz, M. Perlwapin, an abalone nacre protein with three four-disulfide cores (Whey Acidic Protein) domains, inhibits the growth of calcium carbonate crystals. *Biophys. J.* 91, 2601-2608 (2006).
- Tsukamoto, D., Sarashina, I. & Endo, K. Structure and expression of an unusually acidic matrix protein of pearl oyster shells. *Biochem. Biophys. Res. Commun.* 320, 1175-1180 (2004).
- Valli, G., Zardini, D. & Nodari, P. Cycle reproductif et biométrie chez *Chamelea gallina* (L.) (Mollusca, Bivalvia) dans le Golfe de Trieste (Reproductive cycle and biometry of the *Chamelea gallina* stock in the Gulf of Trieste). *Rapp. Comm. Int. Mer Méditerran.* 29, 339-340 (1985).
- Vaughn, C. C. & Hoellein, T. J. Bivalve impacts in freshwater and marine ecosystems. *Annu. Rev. Ecol. Evol. Syst.* 49, 183-208 (2018).
- Vermeij, G. J. A Natural History of Shells. Princeton University Press, New Jersey (1993).
- Vernberg, J. F. Physiological responses of marine biota to pollutants. Academic Press, N.Y. (1976).

- Visciano, P., et al. Concentrations of contaminants with regulatory limits in samples of clam (*Chamelea gallina*) collected along the Abruzzi Region Coast in Central Italy. *J. Food Prot.* 78, 1719-1728 (2015).
- Vogel, S. *Life in Moving Fluids: The Physical Biology of Flow*. Princeton Univ. Press, Princeton, New Jersey (1996).
- Waite, J. H. Quinone-tanned scleroproteins. *The Mollusca*, vol. 4: Physiology. Academic Press, New York (1983).
- Wanamaker, A. D., et al. Experimental determination of salinity, temperature, growth, and metabolic effects on shell isotope chemistry of *Mytilus edulis* collected from Maine and Greenland. *Paleoceanography* 22, PA2217 (2007).
- Wang, X., et al. Oyster shell proteins originate from multiple organs and their probable transport pathway to the shell formation front. *PLoS One* 8:e66522 (2013).
- Watabe, N. Studies on shell formation. XI. Crystal-matrix relationships in the mollusk shells. *J. Ultrastruct. Res.* 12, 351-370 (1965).
- Watson, S. A., et al. Marine invertebrate skeleton size varies with latitude, temperature and carbonate saturation: implications for global change and ocean acidification. *Global Change Biol.* 18, 3026-3038 (2012).
- Weiner, S. & Addadi, L. Crystallization pathways in biomineralization. *Cells Tissues Organs* 41, 21-40 (2011).
- Weiss, I. M., Kaufmann, S., Mann, K. & Fritz, M. Purification and characterization of perlucin and perlustrin, two new proteins from the shell of the mollusc *Haliotis laevigata*. *Biochem. Biophys. Res. Commun.* 267, 17-21 (2000).
- Whittaker, C. A. & Hynes, R.O. Distribution and evolution of von Willebrand/integrin A domains: widely dispersed domains with roles in cell adhesion and elsewhere. *Mol. Biol. Cell* 13, 3369-3387 (2002).
- Widdows, J. Physiological measurements. In *Effects of Stress and Pollution on Marine Animals* (ed. Bayne, B.L.). Praeger Scientific Publishers, N.Y. (1985).
- Widdows, J., et al. Measurement of physiological energetics (scope for growth) and chemical contamination in mussels (*Area zebra*) transplanted along a contamination gradient in Bermuda. *J. Exp. Mar. Biol. Ecol.* 138, 99-118 (1990).
- Wilbur, K. M. & Bernhardt, A. M. Effects of amino acids, magnesium, and molluscan extrapallial fluid on crystallization of calcium carbonate: in vitro experiments. *Biol. Bull.* 166, 251-259 (1984).

- Wilt, F. H. Biomineralization of the spicules of sea urchin embryos. *Zoolog. Sci.* 19, 253-261 (2002).
- Wittmann, A. C. & Po Rtner, H. O. Sensitivities of extant animal taxa to ocean acidification. *Nat. Clim. Chang.* 3(11), 995-1001 (2013).
- Xiang, L., et al. Amorphous calcium carbonate precipitation by cellular biomineralization in mantle cell cultures of *Pinctada fucata*. *PLoS One* 9:e113150 (2014).
- Xie, J. The Mechanism Study of Extrapallial Fluid Protein in the Shell Formation of *Pinctada fucata*. Beijing Shi: Tsinghua University (2016).
- Yan, Z., et al. Biomineralization: functions of calmodulin-like protein in the shell formation of pearl oyster. *Biochim. Biophys. Acta.* 1770(9), 1338-44 (2007).
- Yan, Z., et al. N40, a novel nonacidic matrix protein from pearl oyster nacre, facilitates nucleation of aragonite in vitro. *Biomacromolecules* 8, 3597-3601 (2007).
- Yang, D., et al. A basic protein, N25, from a mollusk modifies calcium carbonate morphology and shell biomineralization. *J. Biol. Chem.* 294(21), 8371-8383 (2019).
- Yano, M., Nagai, K., Morimoto, K. & Miyamoto, H. A novel nacre protein N19 in the pearl oyster *Pinctada fucata*. *Biochem. Biophys. Res. Commun.* 362, 158-163 (2007).
- Yano, M., Nagai, K., Morimoto, K. & Miyamoto, H. Shematin: a family of glycine-rich structural proteins in the shell of the pearl oyster *Pinctada fucata*. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 144, 254-262 (2006).
- Yi, Y., et al. A novel matrix protein, PfY2, functions as a crucial macromolecule during shell formation. *Sci. Rep.* 7, 6021 (2017).
- Yin, Y., et al. Structural characterization of the major extrapallial fluid protein of the mollusc *Mytilus edulis*: implications for function. *Biochemistry* 44, 10720-10731 (2005).
- Zhang, C. & Zhang, R. Matrix proteins in the outer shells of molluscs. *Mar. Biotechnol.* 8, 572-586 (2006).
- Zhang, C., et al. A novel putative tyrosinase involved in periostracum formation from the pearl oyster (*Pinctada fucata*). *Biochem. Biophys. Res. Commun.* 342, 632-639 (2006a).
- Zhang, C., et al. A novel matrix protein p10 from the nacre of pearl oyster (*Pinctada fucata*) and its effects on both CaCO₃ crystal formation and mineralogenic cells. *Mar. Biotechnol.* 8, 624-633 (2006b).
- Zhang, G., et al. The oyster genome reveals stress adaptation and complexity of shell formation. *Nature* 490, 49-54 (2012).

Zhang, Y., et al. A novel matrix protein participating in the nacre framework formation of pearl oyster, *Pinctada fucata*. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 135, 565-573 (2003).

Zheng, X., et al. The AP-1 transcription factor homolog Pf-AP-1 activates transcription of multiple biomineral proteins and potentially participates in *Pinctada fucata* biomineralization. *Sci. Rep.* 5, 14408 (2015).