



UNIVERSITÀ DEGLI STUDI DI PADOVA

FACOLTÀ DI SCIENZE MM.FF.NN.

CORSO DI LAUREA IN BIOLOGIA MOLECOLARE

ELABORATO DI LAUREA

**PREPARAZIONE DI cDNA PER ANALISI
METATRASCRIPTOMICA**

TUTOR: DOTT. STEFANO CAMPANARO

Dipartimento di Biologia
Università degli Studi di Padova

LAUREANDA: MICHELA SERENA

ANNO ACCADEMICO: ANNO 2008-2009

INDICE

ABSTRACT	7
Capitolo 1: STATO DELL'ARTE	8
Introduzione alla metagenomica e alla metatrascrittomica	8
L'importanza dell'rRNA 16S	9
Capitolo 2: MATERIALI E METODI	11
2.1. Prelievo e filtraggio	11
2.2. Estrazione e precipitazione RNA	11
2.3. Rimozione rRNA 16S-23S	12
2.4. Trattamento con DNasi	14
2.5. Rimozione degli small RNA	14
2.6. Sintesi cDNA	15
2.7. Purificazione DNA	16
2.8. Amplificazione tramite <i>TempliPhi</i>	16
2.9. PCR: <i>polymerase chain reaction</i>	17
Capitolo 3: RISULTATI E DISCUSSIONE	18
3.1. Determinazione della concentrazione di RNA dopo l'estrazione	18
3.2. Determinazione della concentrazione di RNA dopo precipitazione	19
3.3. Determinazione della concentrazione di RNA dopo rimozione 16S e 23S	19
3.4. Determinazione della concentrazione di RNA dopo rimozione sRNA	21
3.5. Corsa in gel di Agarosio, dopo amplificazione con <i>TempliPhi</i>	21
3.6. Corsa in gel di Agarosio, per l'analisi dei prodotti di PCR	22
Capitolo 4: BIBLIOGRAFIA	25

ABSTRACT

La metatrascrittomica, parallelamente alla metagenomica, è una disciplina scientifica che si basa sul sequenziamento del trascrittoma di una comunità di microrganismi diversi che coesistono in uno stesso ambiente. L'unica differenza sta nel materiale di partenza: da un lato RNA (metatrascrittoma), dall'altro DNA (metagenoma).

Tale disciplina permette di superare i numerosi problemi spesso dettati dalla difficile coltivazione di moltissimi microrganismi, causati da particolari loro esigenze di crescita, quali temperature spesso molto elevate, pressioni pari a quelle dei fondali oceanici, concentrazioni saline notevoli. Non essendo infatti coltivabili nelle normali condizioni di laboratorio è necessario effettuare un campionamento direttamente dal loro habitat, dove crescono liberamente e, mediante sequenziamento dell'rRNA 16S, è poi possibile risalire alle diverse specie in esso presenti.

L'attività di ricerca di cui mi sono occupata riguarda lo studio di due campioni microbici distinti prelevati nella Laguna di Venezia: uno a Marghera e uno presso la Palude della Rose, di cui si vogliono analizzare sia il profilo di espressione genica, a partire dagli mRNA, sia la diversità microbica, a partire dagli rRNA 16S.

Inoltre, seppur in maniera secondaria, mi sono dedicata all'amplificazione del gene per l'rRNA 16S a partire da un campione metagenomico proveniente sempre dalla Laguna di Venezia.

Capitolo 1: STATO DELL'ARTE

Introduzione alla metagenomica e alla metatrascrittomica

I microrganismi sono in grado di influenzare la maggior parte dei cicli biogeochimici che caratterizzano l'ambiente terrestre e acquatico, e sono per questo molto studiati, soprattutto per comprenderne le attività metaboliche.

A questo spiccato interesse verso il mondo microbico si oppone però un problema di non poca importanza, ossia il fatto che circa il 99% delle specie microbiche esistenti sia difficilmente coltivabile in laboratorio con le modalità tuttora conosciute, e quindi identificabile mediante i tradizionali test biochimici e molecolari. Numerosi sono infatti i problemi legati a particolari loro esigenze di crescita (talvolta addirittura sconosciute), spesso non riproducibili in coltura, come temperatura, pH, pressione, concentrazione salina; inoltre molte volte in una comunità microbica si instaura una stretta interdipendenza tra le diverse specie, cosa che viene invece a mancare in una coltura pura. In queste condizioni risulta pertanto difficile poter effettuare una buona stima della biodiversità del campione in analisi ed è proprio per questi motivi che si sono sviluppate nuove metodiche che, nel loro insieme, costituiscono una vera e propria disciplina: la metagenomica, ovvero l'analisi contemporanea di tutti i genomi dei diversi microrganismi di un dato habitat.

Uno tra i più importanti studi di questo tipo è stato fatto da J. Craig Venter e colleghi: "Environmental Genome Shotgun Sequencing of the Sargasso Sea" (1), in cui l'analisi delle popolazioni microbiche del Mar dei Sargassi è stata condotta applicando un sequenziamento shotgun su tutto il metagenoma da cui poi, mediante analisi bioinformatica, si è potuta predire l'esistenza di più di 6 milioni di proteine, assegnandone più della metà a famiglie note, e allo stesso tempo scoprendo nuovi geni. Inoltre, servendosi di determinati marker filogenetici, relativi a proteine omologhe universalmente distribuite, è stato possibile classificare le specie in esso presenti.

Un lavoro analogo è stato quello svolto da Jack. A. Gilbert e colleghi: "Detection of Large Numbers of Novel Sequences in the Metatranscriptomes of Complex Marine Microbial Communities" (2), in cui ci si è invece concentrati sull'analisi metatrascrittomica di una comunità microbica marina. A partire dagli mRNA estratti dal campione, retrotrascritti in cDNA, amplificati ed infine sequenziati, si è potuto analizzare il profilo di espressione genica all'interno dell'intera comunità. Servendosi infatti del forte supporto bioinformatico è stato possibile individuare migliaia di geni, così come predirne di nuovi ed è inoltre stato possibile valutare come il profilo di espressione genica possa variare in risposta a diverse condizioni ambientali. Parallelamente a questo tipo di analisi, per poter caratterizzare le diverse specie presenti, sono state analizzate le sequenze di rRNA 16S, estratte anch'esse direttamente dal campione.

L'importanza dell'rRNA 16S

L'rRNA ribosomale è la tipologia di RNA più abbondante nella cellula, ed è un componente essenziale dei ribosomi.

Il ribosoma è una particella ribonucleoproteica compatta che si compone di due subunità, una più grande ed una più piccola, ciascuna delle quali è costituita da RNA e proteine. I ribosomi batterici vengono definiti 70S, in funzione della loro velocità di sedimentazione, espressa in Svedberg e possiamo distinguere una subunità maggiore (50S), caratterizzata per la presenza di rRNA 5S e 23S, ed una minore (30S), caratterizzata per la presenza di rRNA 16S (3).

Generalmente nei batteri i geni ribosomali sono organizzati in un operone e possono esservene più copie sparse nel genoma (figura 1.1).



Figura 1.1. Rappresentazione schematica: operone batterico codificante geni ribosomali.

L'rRNA 16S è oggi molto utilizzato per la determinazione delle relazioni evolutive tra microrganismi, in quanto si tratta di una molecola relativamente grande, dotata della medesima funzione in ogni organismo, in quanto coinvolta nella sintesi delle proteine, e per questo caratterizzata da regioni altamente conservate, che presentano la stessa sequenza in tutti i batteri. Allo stesso tempo sono presenti anche regioni altamente variabili, che presentano omologia solo tra batteri della stessa specie, e che possono quindi essere sfruttate per analisi filogenetiche. È stato dimostrato che un procariote, la cui sequenza di rRNA 16S differisce per più del 3% da quella di tutti gli altri microrganismi, appartiene ad una nuova specie (4).

Al giorno d'oggi le tecnologie della PCR e del sequenziamento costituiscono un potentissimo strumento di analisi in questo ambito, poiché permettono di ottenere una visione della biodiversità microbica nettamente superiore a quella riscontrabile mediante semplice coltivazione. Infatti, servendosi di *primers* specifici disegnati sulla base delle regioni altamente conservate dell'rRNA 16S, è possibile amplificare l'intero gene ribosomale per ottenere una classificazione tassonomica il più fedele possibile alla realtà. Con un approccio metatrascrittomico è peraltro possibile bypassare lo step della PCR poiché, estraendo direttamente gli rRNA 16S dal campione, basterebbe retrotrascriverli in cDNA e sequenziarli. Inoltre, grazie all'apporto di nuove tecnologie, finalizzate a consentire un sequenziamento su larga scala, a prezzi e tempi contenuti, risulta ora essere molto più pratico ed efficiente sequenziare geni.

Fino a pochi anni fa il metodo classico consisteva nell'inserzione dei frammenti di DNA genomico in plasmidi, amplificati in batteri, e nel successivo sequenziamento con il metodo Sanger, basato sull'impiego di dideoossinucleotidi marcati e nell'elettroforesi, su gel o capillare. Per far fronte ai numerosi limiti imposti da tale tecnica, quali il basso numero di basi lette per run, la clonazione

del DNA in batteri e l'elettroforesi dei singoli frammenti, sono ora disponibili sequenziatori di nuova generazione, molto più veloci ed efficienti.

In questo caso verrà utilizzato il sequenziatore *454 GS FLX (Roche)* basato sulle tecniche della PCR in emulsione e del pirosequenziamento. La PCR in emulsione è basata sull'utilizzo di adattatori specifici, che vengono legati ai frammenti genomici, complementari a determinati oligo, legati invece a particolari *beads*, che costituiranno i *primers* per l'amplificazione, grazie a cui è possibile ottenere fino a circa 10 milioni di copie di DNA per ogni *bead*, evitando il classico metodo del clonaggio in batteri. Il pirosequenziamento sfrutta invece la capacità della luciferasi di emettere luce, in maniera proporzionale al numero di nucleotidi incorporati durante la polimerizzazione del filamento complementare allo stampo, evitando la classica elettroforesi (5).

Attualmente sono stati delineati 52 *phyla* e la maggior parte di essi è proprio rappresentata da microrganismi non coltivabili (6) (figura 1.2).

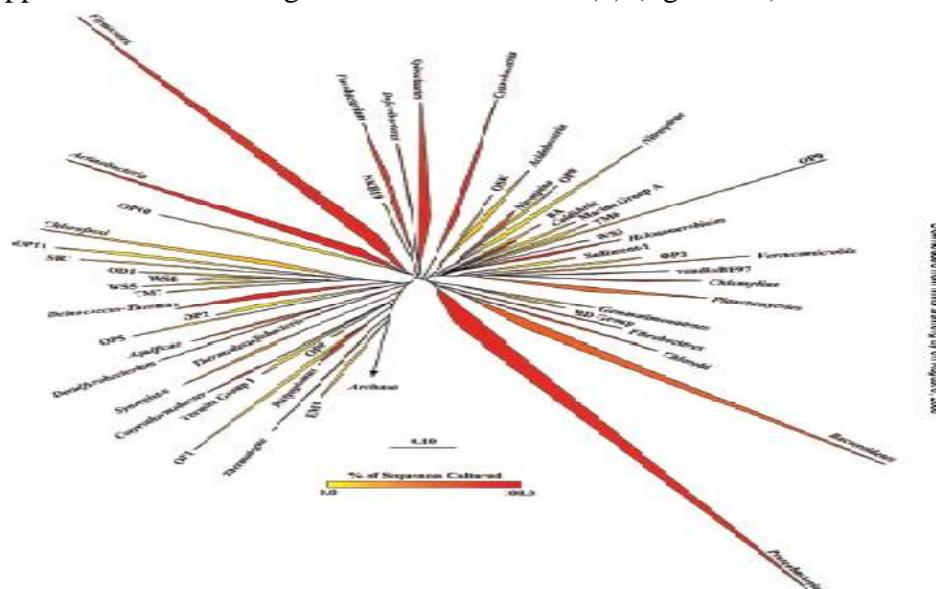


Figura 1.2. Albero filogenetico dei *Bacteria*, che mostra *phyla* conosciuti e nuovi *phyla* candidati, ottenuti dall'analisi di sequenze direttamente da campioni ambientali. Il colore rosso corrisponde alla percentuale di sequenze in quel *phylum* ottenuta da microrganismi coltivabili. Si nota come i *phyla* candidati non contengano sequenze da batteri ottenuti tramite coltura. (immagine ottenuta dal lavoro di Jo Handelsman *Metagenomics: Application of genomics to Uncultured Microorganisms*) (6).

Nel mio lavoro si è utilizzato un approccio metatrascrittomico, incentrato prevalentemente su due obiettivi: analisi dell'espressione genica e analisi della diversità microbica di due distinti campioni ambientali, prelevati in ambienti completamente diversi della Laguna di Venezia: l'uno presso la Palude delle Rose, in un contesto prevalentemente naturale e poco inquinato; l'altro a Marghera, in un ambiente molto più contaminato e ad alto tasso d'inquinamento.

Il punto cruciale del progetto consiste quindi nell'estrazione dell'mRNA da un lato e dell'rRNA dall'altro, seguita poi da retrotrascrizione, in modo da poter infine passare al sequenziamento.

Inoltre, seppur in maniera secondaria, mi sono occupata dell'amplificazione del gene codificante l'rRNA 16S per la preparazione di una libreria, a partire da un campione metagenomico, proveniente dalla Palude delle Rose.

Capitolo 2: MATERIALI E METODI

2.1. Prelievo e filtraggio

I siti di campionamento scelti per effettuare le analisi sono due: Marghera e Palude delle Rose, il primo molto più inquinato rispetto al secondo.

Sono stati prelevati circa 20 l di acqua, i quali sono subito stati processati attraverso tre diversi filtri: uno da 10 µm, seguito da uno da 3µm, per eliminare principalmente gli eucarioti, e uno di policarbonato da 0,22 µm, specifico per i batteri. Successivamente è stato conservato il tutto in ghiaccio secco, per minimizzare eventuali variazioni di espressione genica, causate da sbalzi di temperatura. Rientrati in laboratorio abbiamo congelato i campioni a -80°C.

2.2. Estrazione e precipitazione RNA

I passaggi descritti qui di seguito sono stati impiegati separatamente per entrambi i campioni: Marghera e Palude delle Rose.

mirVana¹ miRNA Isolation Kit

Dopo aver congelato i filtri in azoto liquido, e averli frantumati con un pestello, il primo passaggio per l'estrazione dell'RNA consiste nell'impiego di una soluzione di lisi denaturante, in grado di inattivare le RNasi cellulari: il *Lysis/Binding Buffer*. Aliquotiamo 2000 µl di tale soluzione e, dopo aver risospeso e agitato il tutto, aggiungiamo *miRNA Homogenate Additive*, in una quantità pari ad 1/10 del volume totale, per omogenare il lisato, e lasciamo il tutto in ghiaccio per 10'. Il principio fondamentale su cui si basa l'estrazione dell'RNA sta nel fenolo, che agisce con proprietà diverse, a seconda che il pH sia acido o basico: in questo caso utilizziamo una soluzione a pH acido, in modo tale da avere le proteine in fase fenolica, portare l'RNA in fase acquosa e lasciare invece il DNA all'interfaccia tra le due fasi. Dividiamo la soluzione in due eppendorf ed aggiungiamo quindi una soluzione di fenolo cloroformio a pH acido, ad un volume pari a 750 µl in ognuna, agitiamo per circa 1' utilizzando un vortex e centrifughiamo per 5' alla massima velocità (11,6x1000g). Trasferiamo la fase acquosa in una nuova eppendorf e riscaldiamo la *Elution Solution* (Nuclease-free, 0,1mM EDTA) a 95°C.

In seguito aggiungiamo etanolo 100% in quantità pari a 1,25 volte il volume prelevato e pipettiamo il tutto in un tubino provvisto di filtro; centrifughiamo per 15', in modo che la soluzione passi attraverso il filtro, lasciando l'RNA legato ad esso, ed eliminiamo tale soluzione. Ripetiamo gli stessi passaggi aliquotando 700 µl di *miRNA Wash Solution 1* e poi 500 µl di *miRNA Wash Solution 2/3*.

Infine, messo il filtro su un nuovo tubino, aggiungiamo 100 µl di *Elution Solution* (95°C), spinniamo il tutto e preleviamo l'eluato, contenente l'RNA.

A questo punto effettuiamo una prima lettura al Nanodrop, per quantificare l'RNA presente nei nostri campioni (due eppendorf per ognuno), seguita da corsa in Agilent, per verificarne l'integrità e la quantità (paragrafo 3.1).

Precipitazione RNA con Ammonio Acetato ed Etanolo

Partendo dai nostri campioni, portiamo a 100 µl con acqua, aggiungiamo 1/10 del volume totale di ammonio acetato, 3M, pH 5,2 e etanolo 100% in quantità pari a 2,5 volte il volume totale e mettiamo il tutto in precipitazione a -80°C *over night*.

Centrifughiamo il tutto alla massima velocità (11,6 x 1000g), per 30' a 6°C. Si vede un *pellet* di dimensioni notevoli, evidentemente per l'alta presenza di sali rimasti dopo la procedura di estrazione. Eliminato il surnatante, aggiungiamo 400 µl di etanolo 80% in ogni eppendorf, agitiamo il tutto per favorire il distacco del *pellet* e centrifughiamo a T ambiente, per 10-15'. Già dopo questo primo passaggio il *pellet* si è ridotto di parecchio, ad indicare l'avvenuto scioglimento dei sali in acqua; quindi ripetiamo un secondo lavaggio con etanolo 80%, eliminiamo il surnatante e, asciugato il *pellet* all'aria, riponiamo l'RNA a -80°C.

2.3. Rimozione rRNA 16S-23S

Per lo studio della diversità microbica sarà ora necessario tener da parte circa 4 µg di RNA totale per ogni campione, per i quali si andrà poi ad analizzare l'RNA 16S. Con l'RNA rimanente procederemo invece con la rimozione degli rRNA 16S e 23S, in maniera da isolare solo gli mRNA, per l'analisi metatrascrittomica.

Dall'estrazione dell'RNA precedente erano rimasti circa: 25,8 µg di RNA totale per Marghera e 28,3 µg per Palude (paragrafo 3.1); risospendiamo in 15 µl di *RNA Storage Solution* ciascuno e otteniamo:

É[RNA] Marghera: 1,72 µg/µl; É[RNA] Palude: 1,89 µg/µl

Prelevate le rispettive quantità pari a 4 µg: 2,32 µl per Marghera e 2,12 µl per Palude, le mettiamo a -80°C. Dai campioni rimasti preleviamo 0,2 µl per effettuare le misure al Nanodrop (paragrafo 3.2). Rimangono ora 12,48 µl per Marghera e 12,68 µl per Palude, con i quali procederemo utilizzando il kit *MICROExpress* .

MICROExpress Kit

Il principio su cui si basa la rimozione degli rRNA 16S e 23S è la complementarità di sequenza tra determinati oligo, forniti dal kit *MICROExpress* , e le regioni di 16S e 23S più conservate nelle diverse specie. Una volta avvenuto l'*annealing*, è possibile, mediante l'utilizzo di *beads*

magnetiche, andare a catturare tali oligo, ibridati alle sequenze target, in modo da rimuoverle dall'intero pool di RNA (figura 2.1).

Purtroppo il kit non è efficiente per tutti i tipi di batteri, in quanto gli oligo da esso forniti sono stati progettati basandosi su un consensus e di conseguenza non sempre sono in grado di ibridarsi in maniera efficiente al target, a causa del differente grado di conservazione di tali regioni nelle diverse specie microbiche.

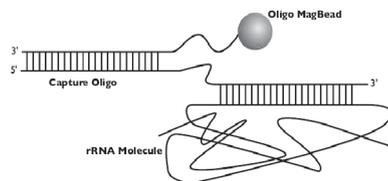


Figura 2.1. Rappresentazione schematica: ibridazione rRNA-Capture Oligo e cattura mediante Oligo MagBead. (immagine tratta da: *MICROBExpress¹*, *Instruction Manual*). Per avere maggiori informazioni riguardo alle specie compatibili:

www.ambion.com/catalog/CatNum.php?1905

Poichè questo kit non è efficiente per quantità di RNA superiori a 10 µg, dividiamo i campioni in due eppendorf ed eseguiamo due passaggi di rimozione. Marghera: 6,24 µl (15,76 µg) ciascuna; Palude: 6,34 µl (12 µg) ciascuna.

Pipettiamo 200 µl di *Binding Buffer* (favorisce l'*annealing*) in ogni eppendorf, aggiungiamo poi l'rRNA, agitiamo il tutto ed aggiungiamo 4 µl di *Capture Oligo Mix*, soluzione contenente gli oligonucleotidi. Centrifughiamo brevemente la soluzione e mettiamo per 10' a 70°C per eliminare possibili *loops* creati nel campione e poi a 37°C per 15' o più, per favorire l'*annealing* tra oligo e rRNA.

Nel frattempo procediamo con la preparazione delle *beads* magnetiche: aliquotiamo in una eppendorf 50 µl di *Oligo MegBeads* per campione (tot. 200 µl) e collochiamo poi il tubino in un campo magnetico (CM), generato da un piccolo magnete, per 3' per poi catturare le sferette metalliche e rimuovere il surnatante. Aggiungiamo 200 µl di *Nuclease-free Water*, rimuoviamo la eppendorf dal magnete e risospendiamo il tutto. Collochiamo di nuovo la soluzione nel CM ed eliminiamo il surnatante. Ripetiamo gli stessi passaggi con 200 µl di *Binding Buffer*. Dopo aver eliminato il surnatante, aliquotiamo altri 200 µl di *Binding Buffer*, rimuoviamo dal CM, risospendiamo, e portiamo a 37°C.

Aliquotiamo ora 50 µl di tale soluzione in ogni eppendorf, contenente quindi RNA + *Capture Oligo Mix*, agitiamo e mettiamo a 37°C per 15'. In questo passaggio le *beads* si legano agli oligo, in modo da poter poi essere catturate, mediante CM (per 3'), insieme agli rRNA 16S e 23S ad essi ibridatisi. Prelevato il surnatante (contenente mRNA e sRNA) è necessario aggiungere alle *beads* catturate 400 µl di *Wash Solution*, preriscaldata a 37°C, rimuovere dal CM e risospendere, per poi ricattare di nuovo le *beads*, prelevare il surnatante e conservarlo insieme al precedente. Quest'ultimo lavaggio è necessario per rimuovere eventuali mRNA legati in maniera aspecifica alle *beads*.

Infine è necessario precipitare di nuovo l'rRNA rimasto, arricchito in mRNA, mediante precipitazione con etanolo e sali, in modo da ridurre il volume e renderlo così più concentrato per i passaggi successivi.

A questo punto effettuiamo un'ulteriore analisi al Nanodrop, seguita da corsa in Agilent (paragrafo 3.3.a), per poi procedere con il secondo passaggio di rimozione.

2.4. Trattamento con DNasi

Il trattamento con la DNasi è necessario per rimuovere eventuali residui di DNA. È essenziale alla fine inattivare la DNasi con il calore, in quanto questa potrebbe poi risultare dannosa nel momento della retrotrascrizione, pertanto metteremo la soluzione a 65-70°C per 5-10'. Poiché la DNasi necessita però di ioni Mg^{2+} (presenti nel *buffer*) che, durante il trattamento termico, degraderebbero l'RNA, è conveniente aggiungere prima l'EDTA, in grado di chelare tali ioni.

Utilizziamo DNasi (5U/ μ l) e diluiamo 1:25 in *buffer Ambion 1X*. Considerando che 1U degrada 1 μ g di DNA plasmidico a 37°C per 10', ne utilizzeremo 0,2 U.

Preparazione DNasi diluita:

É 0,5 μ l di DNasi

É 10,3 μ l di acqua

É 1,2 μ l di *buffer 10X* (contiene Mg^{2+})

Risospendiamo ora l'RNA in 17,5 μ l di acqua e ne preleviamo 0,4 μ l per la lettura al Nanodrop e la corsa in Agilent (paragrafo 3.3.b). Rimangono 17,1 μ l di RNA diluito. Aggiungiamo:

É 1,9 μ l di *buffer 10X*

É 1 μ l di DNasi diluita come sopra

Lasciamo per 15' a 37°C ed aggiungiamo poi 0,83 μ l di EDTA (500 mM) in modo che sia ad una concentrazione di 20 mM; scaldiamo a 70°C per 5' per inattivare l'enzima.

2.5. Rimozione degli small RNA

Riconsideriamo ora anche i 4 μ g di RNA totale messi da parte prima dell'utilizzo del kit *MICROBExpress*, per entrambi i campioni.

Procediamo poi con il kit *MEGAClear*, per rimuovere gli rRNA 5S e i tRNA.

MEGAClear Kit

Tale kit è appropriato per la purificazione di ssRNA superiori a 100 nt, in modo da poter eliminare così gli sRNA (rRNA 5S e tRNA).

Portiamo tutti i campioni (mRNA Marghera/Palude; RNA tot Marghera/Palude) a 100 μ l con la *Elution Solution*, aggiungiamo 350 μ l di *Binding Solution* e 250 μ l di Etanolo 100%. Pipettiamo il tutto in quattro tubini distinti, provvisti di filtro, centrifughiamo per 15' ed eliminiamo il surnatante. Laviamo due volte con 500 μ l di *Wash Solution* e centrifughiamo ulteriormente a vuoto. L'RNA rimasto è ora legato al filtro. Aliquotiamo 50 μ l di *Elution Solution*, mettiamo nel termo blocco per 10' a 70°C, ed infine centrifughiamo per 1' in modo da favorire il distacco degli acidi nucleici dal filtro e conserviamo poi l'eluato. Precipitiamo con etanolo e sali, per purificare e concentrare i campioni; in seguito effettuiamo una lettura al Nanodrop, seguita da corsa in Agilent (paragrafo 3.4).

2.6. Sintesi cDNA

Per la retrotrascrizione ci serviamo della trascrittasi inversa *SuperScript II*, che ha bisogno di inneschi per agire, costituiti da random esameri, cioè sequenze generate in modo puramente casuale, in maniera tale da potersi ibridare per complementarità a numerosi siti sparsi lungo le diverse sequenze presenti nel campione. Una volta ibridatesi, la retrotrascrittasi può procedere con la sintesi del primo filamento di DNA, formando una molecola ibrida RNA/DNA. Per sintetizzare la seconda elica sono necessari 3 enzimi: una RNasi H, in grado di creare dei *nicks* nella molecola di RNA, che costituiranno gli inneschi per il secondo enzima: la Polimerasi I di *E.Coli*, che sintetizzerà la seconda elica, complementare alla prima (essa è dotata di un'attività esonucleasica 5'→3' in grado di rimuovere l'RNA man mano che procede nella polimerizzazione); è poi necessaria una Ligasi di *E. coli*, per congiungere tra loro i frammenti neo sintetizzati. Infine, per rendere le estremità piatte, per l'aggiunta degli adattatori ai fini del sequenziamento (figura 2.2), utilizziamo la T4 DNA Polimerasi.



Figura 2.2. Rappresentazione schematica: frammento genomico legante due adattatori distinti (biotinilato quello in 3').

SuperScript® One-Cycle cDNA Kit

Risospendiamo i campioni in 10 µl di acqua, aggiungiamo 2 µl di Random Esameri 50 µM (1 µg/µl) e centrifughiamo brevemente. Mettiamo il tutto a 70°C per 10' per consentire la denaturazione di strutture secondarie possibilmente create, e poi in ghiaccio per permettere agli esameri di legarsi alle sequenze:

In ogni campione aliquotiamo 7 µl così composti:

É4 µl di buffer per il primo filamento

É2 µl di DTT 0,1 M

É1 µl di dNTPs 10mM

Agitiamo ed incubiamo a 42°C per 2', aggiungiamo 1 µl dell'enzima *SuperScript II RT* (200 U/µl) e incubiamo a 42°C per 1 ora. Mettiamo tutto in ghiaccio e procediamo alla sintesi del secondo filamento di DNA, aliquotando in ogni campione, sempre in ghiaccio, 130 µl così composti::

É91 µl di *RNase-free water*;

É30 µl di buffer per il secondo filamento;

É3 µl di dNTPs 10 mM;

É1 µl di Ligasi 10 U/µl;

É4 µl di DNA Polimerasi I 10 U/µl;

É1 µl di RNasi H 2 U/µl

Agitiamo ed incubiamo per 2 ore a 16°C. Aggiungiamo 2 µl di T4 DNA Polimerasi (5U/µl) e rimettiamo a 16°C per 5'

Infine è sufficiente aliquotare 10 µl di EDTA 0,5 M per inattivare l'enzima.

2.7. Purificazione DNA

Per purificare i campioni è conveniente utilizzare delle particolari *beads* magnetiche, in grado di legare il DNA ed eliminare tutti i sali presenti in soluzione e qualsiasi altro residuo derivante dalle soluzioni introdotte in precedenza.

AGENCOURT® AMPure® PCR Purification Kit

Tale kit prevede l'utilizzo di un *buffer* ottimizzato per legare selettivamente sequenze superiori a 100 basi, mediante particolari *beads* magnetiche, in modo da poter poi eliminare tutti i residui di oligo, nucleotidi, sali ed enzimi con un semplice lavaggio.

Aliquotiamo 291,6 µl di *Agencourt AMPure*, contenente le *beads*, in ogni campione e, dopo aver agitato la soluzione, possiamo metterla nel campo magnetico, per 10'. Aspiriamo la soluzione, evitando di andare a contatto con le *beads*, leganti il DNA, e laviamo queste ultime con 200 µl di etanolo 70%, per eliminare eventuali residui; infine aspiriamo l'etanolo e ripetiamo una seconda volta. Togliamo il campione dal CM e lasciamo asciugare all'aria per 10-20' per poi aliquotarvi 40 µl di *Elution buffer* (H₂O distillata, TRIS Acetato) e agitare vigorosamente, per staccare il DNA dalle *beads*. Infine, collocando di nuovo la soluzione nel CM, ricatturiamo le *beads*, ora libere, e aspiriamo la soluzione rimanente, contenente le sequenze.

2.8. Amplificazione tramite *TempliPhi*

Poiché la quantità di template che abbiamo ottenuto è scarsa, è conveniente amplificare le nostre sequenze. Nel nostro lavoro utilizziamo la polimerasi del fago 29, particolarmente indicata per amplificare quantità davvero minime di DNA, come nel nostro caso, in cui è necessaria un'amplificazione cospicua, ai fini del sequenziamento. La reazione è in grado di avvenire a temperatura costante ed è inoltre molto efficace, grazie all'attività di *proofreading* di tale polimerasi.

Il principio su cui si basa questo tipo di amplificazione si discosta dalla classica PCR, in quanto i *primers* utilizzati sono sequenze randomiche di sei basi, in grado di legarsi casualmente lungo il genoma. Quando la polimerasi, a partire da un *primer*, procede lungo il filamento ed incontra il *primer* successivo, avanza con la polimerizzazione, provocando il distacco di quest'ultimo dal filamento stampo, insieme al filamento che da lì era stato sintetizzato. Quel che si ottiene alla fine di questa amplificazione è una sorta di *hyperbranched DNA* (figura 2.3).



Figura 2.3. Rappresentazione schematica: *Hyperbranched DNA* (immagine tratta da: *Sequencing genomes from single cells by polymerase cloning*) (7).

TempliPhi DNA amplification: short protocol

Diluiti i DNA in 20 µl di acqua per ogni campione (mRNA Palude, RNA totale Palude, mRNA Marghera, RNA totale Marghera), aliquotiamo in tre eppendorf:

É0,2 µl É0,5 µl É1 µl per tutti i quattro campioni (12 eppendorf in totale), in modo da poter poi valutare quello maggiormente amplificato.

Come controllo positivo utilizziamo 1 ng di plasmide di controllo (paragrafo 3.5.b). Aliquotiamo 20 µl di *buffer* di reazione (denaturante) in ogni eppendorf e scaldiamo a 95°C per 30" per denaturare il DNA e poi mettiamo a 4°C, per consentire ai random esameri di legarsi in posizioni casuali lungo il genoma. Aggiungiamo ora 20 µl di *TempliPhi premix* (contenente la polimerasi), precedentemente scongelato, agitiamo e mettiamo per 18 ore a 30°C. Alla fine del ciclo inattiviamo l'enzima incubando a 65°C per 10" e poi reffreddiamo a 4°C. Proseguiamo con una corsa in gel d'Agarosio 1% (paragrafo 3.5.a).

Infine i nostri campioni sono pronti per essere sequenziati.

2.9. PCR: *polymerase chain reaction*

Seppur in maniera secondaria, mi sono inoltre occupata di allestire una reazione di PCR per amplificare il gene codificante l'rRNA 16S, a partire da due campioni di DNA genomico estratti da due dei filtri provenienti dalla Laguna di Venezia (Palude delle Rose) (paragrafo 2.1), ai fini di creare una libreria di 16S.

Abbiamo pertanto utilizzato *primers* batterici costruiti in modo da amplificare il gene ribosomale dalla posizione 300 a 700.

La mix 1X da preparare è costituita da:

ÉTemplato: 1,5 µl

É10U/µl Polimerasi Phusion: 0,2 µl

É200 µM dNTPs: 0,4 µl

É0,5 µM Primer For: 0,7 µl

É0,5 µM Primer Rew: 0,7 µl

É1X Buffer (contenente Mg²⁺): 4 µl

ÉH₂O: 12,5 µl

Partendo da due aliquote di genomico (DNA1; DNA2) non quantificato, derivanti dalla procedura di estrazione dai due filtri, prepariamo tre distinte diluzioni: 1:1; 1:100; 1:1000, per valutare poi, in base all'amplificazione, quale possa essere sufficiente per la creazione della libreria. Come controllo positivo utilizziamo *E. coli* (1,5 µg/µl) e come controllo negativo, acqua. Prepariamo quindi una MasterMix 10X. Il ciclo di PCR utilizzato è:

Denaturazione iniziale: 98°C, 20"

35X { Denaturation: 98°C, 20"
Annealing: 60°C, 45"
Extension: 72°C, 1"

Final Extension: 72°C, 60"

Al termine proseguiamo con una corsa in gel d'Agarosio 1% (paragrafo 3.6).

Capitolo 3: RISULTATI E DISCUSSIONE

3.1. Determinazione della concentrazione di RNA dopo l'estrattazione

Lettura al Nanodrop a partire dai campioni di RNA estratti mediante utilizzo del kit *mirVana 1*, divisi in due aliquote, per Marghera (1; 2) e Palude (1; 2).

	230 nm	260 nm	280 nm	320 nm	C (ng/μl)	260/280	260/230
Marghera1	2,073	3,567	1,628	0,024	143	2,19	1,72
Marghera2	2,012	2,922	1,634	0,042	115	2,10	1,79
Palude1	2,031	3,22	1,497	0,042	128	2,14	1,58
Palude2	2,166	3,866	1,808	0,043	155	2,14	1,78

RNA TOT Marghera in 100 μl: 14,3 μg + 11,5 μg = 25,8 μg

RNA TOT Palude: in 100 μl: 12,8 μg + 15,5 μg = 28,3 μg

Osservando il valore di assorbanza a 260 nm, considerata la diluizione dell'RNA nei campioni, si deduce che esso sia presente in quantità abbondante; infatti anche il rapporto 260/280, corrispondente al rapporto acidi nucleici/proteine, essendo maggiore di 2, indica che la procedura di purificazione è stata efficace.

Per valutare l'integrità dell'RNA, viene effettuata, al MicroCribi (*Microarray Service & RNA Analysis Service*), una corsa in Agilent, mediante *RNA 6000 Nano Chip Kit*, o *RNA 6000 Pico Chip Kit*, a seconda della quantità di RNA da noi fornita. In questo caso circa 100 ng (figura 3.1; 3.2).

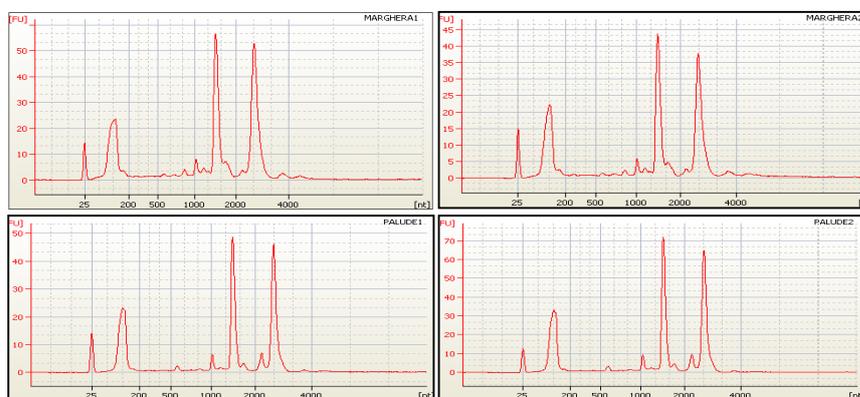


Figura 3.1. In alto (da sx a dx): Marghera1-Marghera2; In basso (da sx a dx): Palude1-Palude2

In tutti i quattro casi riscontriamo due picchi molto alti a circa 1500 nt e 2500 nt, rispettivamente corrispondenti agli rRNA 16S e 23S, e un terzo picco a 100 nt, relativo ai tRNA e agli rRNA 5S; i rimanenti segnali corrispondono invece agli mRNA. Si vede però, valutando il rapporto tra i segnali relativi ai 23S e 16S, come questo sia inferiore rispetto all'atteso, poichè teoricamente dovrebbe aggirarsi intorno a 2: ciò potrebbe essere dovuto ai numerosi processamenti cui i campioni sono stati sottoposti ed inoltre anche il fatto di averli trasportati dal sito di campionamento al laboratorio, con le conseguenti variazioni termiche, potrebbe aver influito sulla stabilità delle molecole. In ogni caso la quantità di rRNA risulta essere molto abbondante, a scapito degli mRNA, proprio come ci aspettiamo, in quanto quest'ultimo è normalmente presente nelle cellule ad una concentrazione molto bassa, che si aggira intorno al 4-6 %.

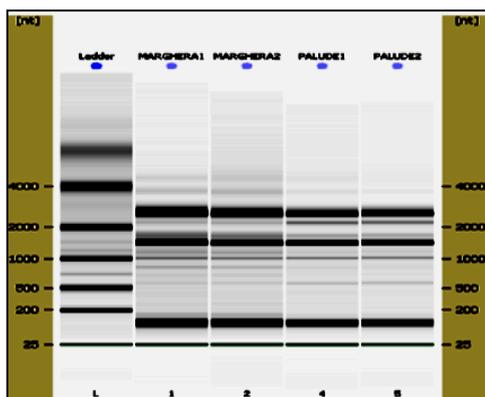


Figura 3.2. Immagine gel:
(da sx a dx): ĤLadder ĤMarghera1 ĤMarghera2
ĤPalude1 ĤPalude2.

Si vede come le bande piŹ intense siano proprio quelle corrispondenti agli rRNA 16S e 23S (1500 nt e 2500 nt), seguite da quella per gli rRNA 5S e i tRNA (a circa 100 nt).

3.2. Determinazione della concentrazione di RNA dopo precipitazione

Preleviamo 0,2 μ l dalla soluzione di RNA precipitato in 15 μ l *RNA Storage Solution*, aggiungiamo 1,3 μ l di *RNA Storage Solution*, per un totale di 1,5 μ l e determiniamo la concentrazione di RNA al Nanodrop, per entrambi i campioni.

	C (ng/ μ l)	260/280	260/230
Marghera	336,7	2,03	1,74
Palude	252,6	2,01	1,65

Marghera:

[RNA] in 0,2 μ l: 2,525 μ g/ μ l
RNA in 2,32 μ l (messi da parte): 5,86 μ g
RNA in 12,48 μ l (soluzione rimanente): 31,51 μ g

Palude:

[RNA] in 0,2 μ l: 1,895 μ g/ μ l
RNA in 2,12 μ l (messi da parte): 4,02 μ g
RNA in 12,68 μ l (soluzione rimanente): 24,03 μ g

In entrambi i casi la qualitŹ dell'RNA Ź ancora buona, poichŽ il rapporto 260/280 si aggira intorno a 2. E' da notare perŹ che stranamente la quantitŹ di RNA per Marghera Ź aumentata rispetto a prima. Questo fatto Ź alquanto inatteso, poichŽ le misure precedenti (paragrafo 3.1) dimostravano come la procedura di purificazione fosse stata efficace. Probabilmente si tratta di errori di misura.

3.3. Determinazione della concentrazione di RNA dopo rimozione 16S e 23S

a) Dopo aver risospeso i campioni in 15 μ l di *RNA Storage Solution*, dopo il primo passaggio di rimozione, ne preleviamo 0,4 μ l e li diluiamo in 2,6 μ l di *RNA Storage Solution* per poi leggerne le concentrazioni al Nanodrop:

	C (ng/ μ l)	260/280	260/230
Marghera	94,8	2,03	1,30
Palude	62,4	1,93	1,24

Marghera

[RNA] in 0,4 μ l: 0,71 μ g/ μ l
RNA in 14,6 μ l (soluzione rimanente): 10,37 μ g

Palude

[RNA] in 0,4 μ l: 0,468 μ g/ μ l
RNA in 14,6 μ l (soluzione rimanente): 6,83 μ g

Possiamo vedere come la quantitŹ di RNA si sia ridotta di circa 1/3 rispetto a prima e, anche se il rapporto 260/280 Ź sceso poco al di sotto di 2 per quanto riguarda il campione Palude, possiamo ritenerci abbastanza soddisfatti per questo

primo passaggio di rimozione 16S e 23S. Infatti questa notevole diminuzione di RNA è giustificabile proprio dal fatto che l'rRNA costituisce la maggior parte dei trascritti presenti nella cellula (circa 90%) e di conseguenza questo dato fornisce un buon parametro per poter affermare che l'impiego del kit è stato efficace.

b) Dopo il secondo passaggio di rimozione, a partire da 17,5 µl di RNA risospeso in acqua, ne diluiamo 0,4 µl in 2,6 µl di acqua ed effettuiamo la lettura.

	C (ng/µl)	260/280	260/230
Marghera	67,6	1,90	1,12
Palude	64,3	1,80	1,06

Marghera

[RNA] in 0,4 µl: 0,507 µg/µl

RNA in 17,1 µl (soluzione rimanente): 8,67 µg

Palude

[RNA] in 0,4 µl: 0,482 µg/µl

RNA in 17,1 µl (soluzione rimanente): 8,24 µg

Si può vedere come la quantità di RNA per Marghera sia leggermente diminuita, mentre quella per Palude è stranamente aumentata: questo potrebbe essere dovuto, oltre che a errori di misura, all'eventuale presenza di oligonucleotidi rimasti a seguito della procedura di rimozione degli rRNA 16S e 23S.

Al MicroCribi viene inoltre effettuata una corsa in Agilent, a partire da circa 100 ng di campione, indistintamente per Marghera o Palude (figura 3.3; 3.4).

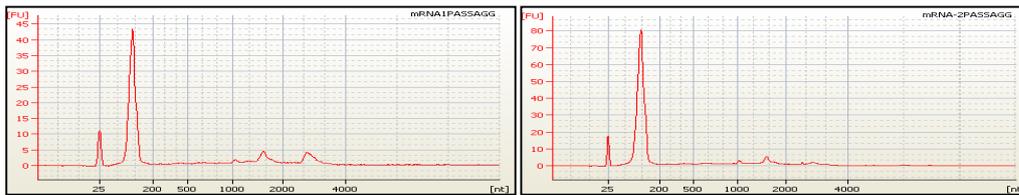


Figura 3.3. (da sx a dx): dopo primo passaggio di rimozione; dopo secondo passaggio di rimozione.

Si vede nettamente, già dopo il primo passaggio, come gli rRNA 16S e 23S siano drasticamente calati, nonostante sia rimasto un lieve picco, probabilmente dovuto agli rRNA di specie batteriche non compatibili con il kit. Ricordiamo infatti che gli oligo sono costruiti sulla base di un consensus stimato nelle regioni altamente conservate dell'rRNA 16S e 23S, ma questo non garantisce una completa efficacia nell'ibridazione con i targets per tutte le diverse specie microbiche presenti nel campione e non è quindi da escludere il fatto che alcuni di tali trascritti siano ancora presenti. In ogni caso, confrontando le altezze dei picchi, si vede come, avendo rimosso gli rRNA 16S e 23S, i tRNA e gli rRNA 5S siano più concentrati.

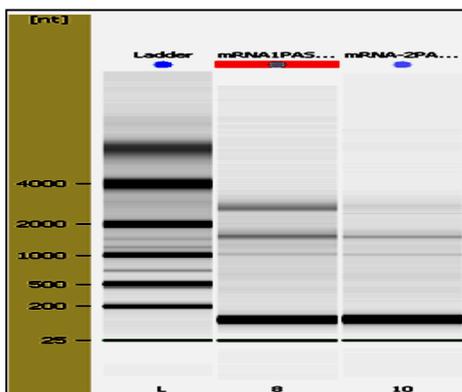


Figura 3.4. Immagine gel:

(da sx a dx): Ladder mRNA dopo primo passaggio mRNA dopo secondo passaggio.

Si vede chiaramente come la banda più marcata sia quella corrispondente a circa 100 nt (rRNA 5S; tRNA), mentre quelle relative agli rRNA 16S e 23S siano molto meno marcate rispetto a prima, soprattutto dopo il secondo passaggio.

3.4. Determinazione della concentrazione di RNA dopo rimozione sRNA

Dopo la precipitazione dei quattro campioni, risospendiamo in 9,7 μl di acqua ed effettuiamo le relative letture al Nanodrop, diluendone 0,75 μl in 0,75 μl di acqua.

	C(ng/ μl)	260/280	260/230
Marghera tot, 4 μg	29,1	2,13	0,88
Palude tot, 4 μg	26,2	2,10	0,82

RNA tot Marghera, da 4 μg : 0,52 μg in 8,95 μl
 RNA tot Palude, da 4 μg : 0,47 μg in 8,95 μl

Notiamo subito che i campioni sono di buona qualità (260/280 > 2).

Per quanto riguarda i campioni di mRNA, la rilevazione da parte dello strumento è pressoché impossibile o comunque poco affidabile, a causa delle bassissime concentrazioni: Marghera: 3,9 ng/ μl (0,070 μg in 8,95 μl); Palude: 5,9 ng/ μl (0,106 μg in 8,95 μl). Presso il MicroCribi viene effettuata una corsa in Agilent, a partire da circa 5 ng di campione (mRNA Marghera/Palude) (figura 3.5; 3.6).

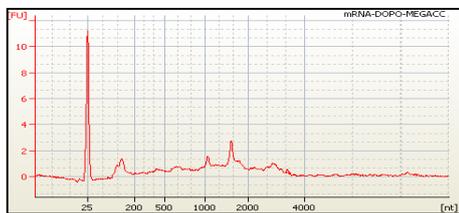


Figura 3.5. Dopo rimozione sRNA.

Si può notare come ora gli mRNA presentino un segnale maggiore e come invece il segnale relativo ai tRNA e rRNA 5S sia drasticamente diminuito.

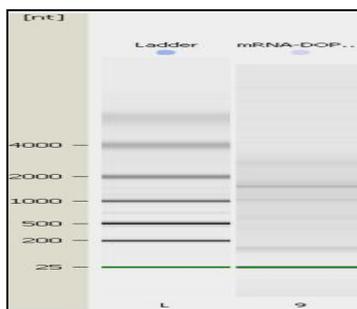


Figura 3.6. Immagine gel:

(da sx a dx): Ladder mRNA dopo MEGAclean.

Si osserva che anche la banda a livello dei 100 nt è molto meno marcata rispetto a prima. E ora infatti visibile solo lo smear dato dagli mRNA.

3.5. Corsa in gel d'Agaroso, dopo amplificazione con TempliPhi

a) Dopo l'amplificazione con *TempliPhi*, facciamo un elettroforesi su gel d'Agaroso 1%, caricando 0,4 μl di amplificato in ogni pozzetto (figura 3.7).

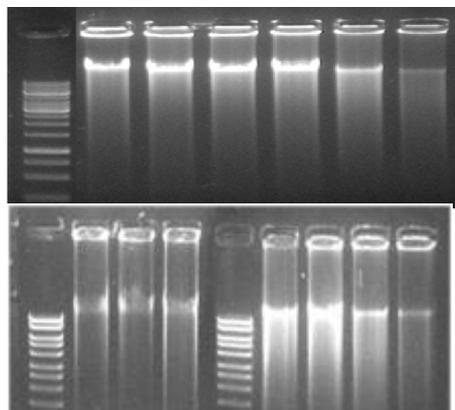


Figura 3.7. Immagine gel:

in alto (da sx a dx): Marker 1 Kb plasmide di controllo (a partire da 1 ng) RNA tot Marghera (a partire da 1 μl) RNA tot Marghera (a partire da 0,5 μl) RNA tot Marghera (a partire da 0,2 μl) mRNA Marghera (a partire da 0,5 μl) mRNA Marghera (a partire da 0,2 μl);

in basso (da sx a dx): Da sx a dx: Marker 1 Kb RNA tot Palude (a partire da 1 μl) RNA tot Palude (a partire da 0,5 μl) RNA tot Palude (a partire da 0,2 μl) marker 1 Kb mRNA Marghera (a partire da 1 μl) mRNA Palude (a partire da 1 μl) mRNA Palude (a partire da 0,5 μl) mRNA Palude (a partire da 0,2 μl).

È possibile vedere dalla figura 3.7, che gli RNA totali sia Marghera, che Palude, sono presenti in gran quantità, sia che il templato fosse 0,2 μl , che 0,5 μl , che 1 μl ,

indice di una buona amplificazione. Per quanto riguarda i rispettivi mRNA, si vede invece come sia stato maggiormente amplificato quello con 1 μ l di template, seguito da quello con 0,5 μ l, e 0,2 μ l. Questo tipo di amplificazione è purtroppo poco affidabile per la formazione di un elevato numero di chimere, cioè di artefatti, anche se si è però visto come questo inconveniente possa essere ridotto mediante un processo di deramificazione (7).

b) Per valutare che l'amplificazione non possa avvenire in assenza di template, abbiamo inoltre effettuato delle amplificazioni di prova, in modo da escludere il fatto che vi possano essere delle contaminazioni nella polimerasi.

Caricati sempre 0,4 μ l di amplificato (figura 3.8).

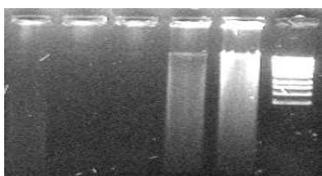


Figura 3.8. Immagine gel:

(da sx a dx): Φ plasmide di prova (a partire da 0,01 ng) Φ 1 μ l H₂O Φ controllo negativo (senza template, solo *TempliPhi*) Φ plasmide di prova (a partire da 0,1 ng) Φ plasmide di prova (a partire da 1 ng) Φ marker 1 Kb.

Si vede come 0,01 ng di plasmide siano insufficienti, in quanto l'amplificazione pare assente, a differenza di 0,1 ng e 1 ng; inoltre sembra non esserci alcuna amplificazione, né in acqua, né nel controllo negativo. Questo sembra in contraddizione con quanto riportato nel kit, dove si dice che anche in assenza di template è possibile individuare una banda alta, data da amplificazione aspecifica.

3.6. Corsa in gel d'Agarosio, per l'analisi dei prodotti di PCR

Per quanto riguarda le due aliquote di DNA genomico (DNA1; DNA2) estratte dai due filtri di Palude delle Rose, per la creazione della libreria di 16S (paragrafo 2.9), terminati i 35 cicli di PCR, facciamo un'elettroforesi su gel d'Agarosio 1%, caricando 0,4 μ l di amplificato (figura 3.9).

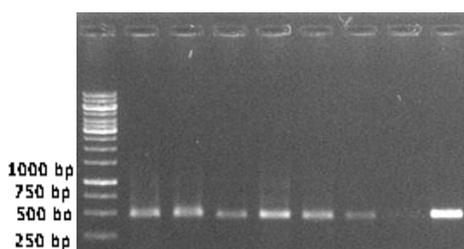


Figura 3.9. Immagine gel.

(da sx a dx): Φ marker 1 Kb Φ DNA1 Φ DNA1 (diluito 1:100) Φ DNA1 (diluito 1:1000) Φ DNA2 Φ DNA2 (diluito 1:100) Φ DNA2 (diluito 1:1000) Φ controllo negativo (acqua) Φ controllo positivo *E. coli* (a partire da 2,25 μ g).

Dalla figura 3.9 si può vedere innanzitutto come l'amplificazione sia avvenuta in tutti i casi in cui fosse presente il template, compreso il controllo positivo, mentre giustamente il controllo negativo è assente. Inoltre, proprio come ci aspettiamo, poiché i *primers* utilizzati consentono di amplificare il gene ribosomale batterico da 300 bp a 700 bp, osserviamo che le bande si trovano effettivamente ad un'altezza poco inferiore a 500 nt. Infine, confrontando le diverse bande, corrispondenti alle tre diluizioni per entrambi i campioni, possiamo riscontrare che anche quelli diluiti 1:1000 presentano una notevole quantità di amplificato e possiamo pertanto concludere dicendo che tale diluizione risulta essere più che sufficiente per la preparazione della libreria.

Capitolo 4: BIBLIOGRAFIA

1. **Venter, J. C., Remington, K., Heidelberg, J. F.** 2004. Environmental Genome Shotgun Sequencing of the Sargasso Sea. *Science* 304: 66-72.
2. **Gilbert, J. A., Field, D., Huang, Y.** 2008. Detection of Large Numbers of Novel Sequences in the Metatranscriptomes of Complex Marine Microbial Communities. *PLoS ONE* 3 (8): e3042: 1-13.
3. **Lewin, B.** 2006. Il Gene VIII (Zanichelli). Capitolo 6: La sintesi delle proteine.
4. **Madigan, M., T., Martinko, J., M.** 2007 Brock Biologia dei Microrganismi (Casa Editrice Ambrosiana). Capitolo 11: Sistematica molecolare ed evoluzione microbica.
5. **Margulies, M., Egholm, M., Altman, W., E.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376-380.
6. **Handelsman, J.** 2004. Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiol. Mol. Biol. Rev.*, 68: 669-685.
7. **Zhang, K., Martiny, A., C., Reppas, N., B.** 2006. Sequencing genomes from single cells by polymerase cloning. *Nature Biotechnology* 24: 680-686.

