



MINISTERO DELLO SVILUPPO ECONOMICO
DIREZIONE GENERALE PER LA LOTTA ALLA CONTRAFFAZIONE
UFFICIO ITALIANO BREVETTI E MARCHI

DOMANDA DI INVENZIONE NUMERO	102010901844372
Data Deposito	31/05/2010
Data Pubblicazione	01/12/2011

Classifiche IPC

Titolo

METODO PER LA PREPARAZIONE E AMPLIFICAZIONE DI LIBRERIE RAPPRESENTATIVE DI CDNA PER IL SEQUenziAMENTO MASSIVO, LORO USO, KIT E CARTUCCE PER KIT DI AUTOMAZIONE

DESCRIZIONE

a corredo di una domanda di brevetto per invenzione dal titolo:

“METODO PER LA PREPARAZIONE E AMPLIFICAZIONE DI LIBRERIE
RAPPRESENTATIVE DI CDNA PER IL SEQUENZIAMENTO MASSIVO, LORO
5 USO, KIT E CARTUCCE PER KIT DI AUTOMAZIONE”

a nome : CONSIGLIO NAZIONALE DELLE RICERCHE

inventori: Apollonia Tullo, Marina Mangiulli, Elisabetta Sbisà, Graziano Pesole

10 SETTORE DELL'INVENZIONE

La presente invenzione concerne un metodo rapido, semplice, efficace ed economico per la preparazione e amplificazione di librerie rappresentative di cDNA per il sequenziamento massivo di nuova generazione, kit e cartucce (“cartridges”) per kit di automazione che utilizzano tale metodo manualmente o in maniera automatica.

15

TECNICA ANTERIORE

Nel corso degli ultimi anni sono state sviluppate nuove piattaforme di sequenziamento di nuova generazione (“next-generation sequencing technologies”, NGS), tra le quali la “454 GS FLX” della Roche, la “SOLiD” dell’Applied Biosystems e il “Genome Analyzer II”
20 dell’Illumina. Inizialmente utilizzate quasi esclusivamente per il sequenziamento e/o risequenziamento di interi genomi, e per la determinazione di polimorfismi a singolo nucleotide (SNP - Single Nucleotide Polymorphism), attualmente vengono sempre più impiegate in modo efficace per molte altre applicazioni, tra le quali l’identificazione di marcatori biologici e agenti patogeni in studi di biodiversità e metagenomica in ambito
25 biomedico, agroalimentare, ambientale e industriale.

Tra le più recenti e promettenti applicazioni, il sequenziamento dell’intero trascrittoma sta diventando fondamentale per lo studio della regolazione dell’espressione genica. Infatti, questo tipo di analisi permette l’identificazione e la quantificazione oltre che degli mRNA, anche di piccoli RNA, come micro-RNA e altri RNA non codificanti, il cui significato
30 biologico è tuttora largamente sconosciuto. Tra le piattaforme di sequenziamento di nuova generazione a oggi disponibili, il pirosequenziatore 454 GS FLX della Roche permette di ottenere per ogni corsa oltre un milione di “reads” e, a differenza delle altre piattaforme, ciascuna sequenza può raggiungere una lunghezza di circa 400 basi. Sequenze di questa

lunghezza permettono, oltre che l'individuazione e la quantificazione relativa degli mRNA, anche l'identificazione di nuovi siti e varianti di "splicing", espressi in diverse condizioni fisiopatologiche. La lunghezza inferiore delle "reads" delle piattaforme prodotte da Illumina (100-150 bp) e Applied Biosystems (35-50 bp), le rende invece particolarmente adatte al sequenziamento di piccoli RNA e alla quantificazione dei livelli di trascritti.

Indipendentemente dalla piattaforma di NGS usata, l'analisi del trascrittoma richiede come primo passaggio la costruzione di librerie di cDNA adatte per il sequenziamento massivo.

Uno dei principali problemi è rappresentato dal fatto che spesso la quantità di RNA di partenza è molto esigua e che alcuni trascritti, pur avendo un importante significato biologico, sono molto poco abbondanti. A questo si aggiunge che circa il 90-95% dell'RNA totale che viene estratto è costituito da RNA ribosomale (rRNA) che, per una buona riuscita dell'esperimento, è indispensabile rimuovere. Questo significa partire da un maggiore quantitativo di RNA totale non sempre disponibile.

I protocolli esistenti per la costruzione di librerie di cDNA prevedono le seguenti tappe: retrotrascrizione degli RNA, sintesi del secondo filamento e in alcuni casi amplificazione delle specie di RNA retrotrascritte. La retrotrascrizione degli mRNA si può fare utilizzando oligo-dT o piccoli primers a sequenza casuale (random primers). In genere si preferisce usare oligo-dT perché gli rRNA non sono poliadenilati e in questo modo vengono esclusi nella reazione di retrotrascrizione. L'uso di oligo-dT ha il grosso svantaggio di escludere dalla reazione di retrotrascrizione oltre agli rRNA, anche tutte le specie di RNA non poliadenilate come le specie di RNA con funzione di regolazione o le specie di RNA batterici o virali (Metatrascrittomica), che si vogliono identificare in cellule umane in alcune condizioni fisiologiche o patologiche, come per esempio nel caso della sclerosi multipla. Inoltre le lunghe code di omopolimeri di T o A possono dare origine a sequenze di scarsa qualità, in particolare per il sequenziatore 454 GS FLX della Roche. Per ovviare a questo problema tecnico possono essere utilizzati dei primers oligo-dT con una sequenza al 5' contenente un sito di restrizione, che taglia 14-16 nt a valle. In questo modo la coda di T può essere rimossa dopo la retrotrascrizione; oppure si possono utilizzare per la retrotrascrizione degli oligo-dT in cui ogni 4 o 5 T è inserito un altro nucleotide, in modo da interrompere la lunga coda di T.

Dopo la sintesi del secondo filamento complementare al filamento retrotrascritto, in alcuni casi si procede a una reazione di amplificazione utilizzando degli adattatori, che vengono legati alle estremità dei frammenti retrotrascritti. Questo consente di amplificare il

materiale di partenza e in particolare le specie di RNA meno rappresentate, altrimenti difficili da rilevare. Per la reazione di amplificazione in genere è usata una Taq polimerasi. Questo protocollo ha comunque l'inconveniente di produrre una popolazione di frammenti più piccoli di 2Kb; in questo modo, la nebulizzazione, che è il processo di frammentazione
5 usato nella procedura standard di preparazione di librerie "shotgun" per il sequenziamento, potrebbe non dare risultati ottimali, in quanto più adatto a molecole di alto peso molecolare.

Protocolli alternativi prevedono, invece, l'isolamento degli mRNA con oligo-dT, la frammentazione delle molecole di RNA, la purificazione dei bassi pesi molecolari da gel e
10 poi la retrotrascrizione. In questo caso si omette il passaggio di nebulizzazione. Non è prevista l'amplificazione dei frammenti retrotrascritti, con l'inconveniente di non riuscire a identificare le specie di RNA meno rappresentate.

Attualmente non esiste un unico protocollo di preparazione di librerie rappresentative di cDNA che tenga conto di tutte le problematiche della tecnica anteriore. Pertanto è
15 necessario sviluppare un metodo che risolva tutte le problematiche descritte e che sia rapido, semplice ed economico.

DESCRIZIONE DELL'INVENZIONE

La presente invenzione consente la preparazione di librerie di cDNA sequenziabili per
20 mezzo delle principali piattaforme di sequenziamento massivo oggi disponibili sul mercato.

Gli aspetti innovativi della procedura riguardano lo studio della successione delle diverse fasi per la preparazione del campione e l'ottimizzazione delle performance del sequenziamento massivo. In particolare il nuovo processo è in grado di operare con:

- 25 1. uso di esigue quantità di RNA totale di partenza (400-500 ng)
2. riduzione considerevole dell'rRNA (circa 90%)
3. retrotrascrizione a partire da 30-40 ng di trascritti poliadenilati e non poliadenilati con primers specifici disegnati dagli autori
4. ligazione dei retrotrascritti
- 30 5. amplificazione con la Polimerasi Phi29 ad altissima fedeltà e processività
6. elevate rese di DNA amplificato (40 µg).

Con la strategia messa a punto è possibile sottoporre a sequenziamento massivo anche campioni di RNA preziosi e disponibili in piccole quantità, altrimenti difficilmente

analizzabili, soprattutto per quanto riguarda i trascritti poco rappresentati. Estremamente rilevante è anche il fatto che le quantità di partenza di RNA (sia totale che privo di rRNA) sono dimezzate rispetto ad altri protocolli e che con le alte rese raggiunte attraverso la procedura dell'invenzione è possibile utilizzare la libreria di cDNA per ulteriori successivi esperimenti di validazione.

Gli aspetti innovativi della procedura riguardano la messa a punto della successione delle diverse fasi per la preparazione del campione e l'ottimizzazione delle performance del sequenziamento massivo.

In questo contesto gli autori hanno messo a punto una procedura rapida, semplice, efficace, economica, che consente di partire da quantitativi dimezzati di RNA totale (400-500 ng), di ridurre in maniera notevole l'rRNA (circa 90%) e di costruire una libreria di trascritti poliadenilati e non poliadenilati rappresentativa anche delle specie meno abbondanti in quanto prevede dopo la retrotrascrizione una fase di amplificazione.

Una libreria di cDNA rappresenta l'insieme dei trascritti espressi da uno specifico sistema cellulare. Geni altamente espressi saranno presenti nella libreria più volte, mentre geni espressi a bassi livelli saranno meno presenti o addirittura assenti. Una libreria di cDNA si dice rappresentativa quando contiene, in almeno una copia, tutti i possibili trascritti, compresi quelli meno abbondanti.

Le DNA polimerasi differiscono tra loro per diverse proprietà, come la processività, l'attività di correzione di bozze, l'attività di "strand displacement", ovvero la capacità di spiazzare il DNA incontrato a valle durante la sintesi, e le rese. Tra le DNA polimerasi disponibili, la seguente tabella indica quali hanno alti livelli di attività di correzione di bozze e attività di "strand displacement", e relativo tasso di errore.

	3'→5' Esonucleasi	Tasso di errore (x10 ⁻⁵)	Strand Displacement
<i>Bst</i> DNA Polymerase, Large Fragment	–		++++
<i>Bst</i> DNA Polymerase, Full Length	–		–
<i>Taq</i> DNA Polymerase	–	285	–
Vent _R ® DNA Polymerase	++	57	++
Vent _R ® (exo–) DNA Polymerase	–	190	+++
Deep Vent _R ®	+++		++

DNA Polymerase			
Deep Vent _R ® (exo-) DNA Polymerase	-		+++
9°N _m TM DNA Polymerase	+		+++
T7 DNA Polymerase	++++	15	-
<i>E. coli</i> DNA Polymerase I	++	9	-
T4 DNA Polymerase	++++	<1	-
phi29 DNA Polymerase	++++	9,5	+++++
Phusion® High-Fidelity DNA Polymerase	++++	0.44	
Phusion® High-Fidelity Hot Start DNA Polymerase	++++	0.44	
Phusion® High-Fidelity Hot Start II DNA Polymerase	++++	0.44	
Phire® Hot Start DNA Polymerase	+		-
Phire® Hot Start II DNA Polymerase	+		-
Phusion® Flash High-Fidelity DNA Polymerase	++++		
LongAmp TM Taq DNA Polymerase	++		-

Gli autori hanno utilizzato la Polimerasi Phi29, che si distingue per le seguenti caratteristiche: i) attività di “strand-displacement”, che consente un’efficace e uniforme amplificazione anche in presenza di strutture secondarie e di sequenze difficili da amplificare; ii) attività di “proof-reading” tra le più elevate oggi disponibili nell’ambito delle Taq Polimerasi ad alta fedeltà; iii) alta processività, che consente la generazione di prodotti di sintesi molto lunghi, ideali per lo step di nebulizzazione, senza dissociarsi dal template; iv) elevate rese di amplificato a partire da piccole quantità di template. Polimerasi analoghe possono essere ugualmente utilizzate.

10 Infatti il metodo della presente invenzione permette di ottenere, partendo da una quantità esigua di RNA totale (circa 400-500 ng), circa 30-40 ng di RNA (notevolmente ridotto della componente di rRNA) da retrotrascrivere e amplificare con una resa finale superiore

ai 40 µg di DNA ad alto peso molecolare. Pertanto il metodo è particolarmente adatto per conseguire ottimi risultati di frammentazione con la nebulizzazione (si veda Figura 2). Questo consente di avere materiale più che sufficiente non solo per le procedure di sequenziamento massivo, ma anche per ulteriori analisi e validazioni, come la
5 determinazione del profilo di espressione genica attraverso Real Time PCR. Infatti, il metodo dell'invenzione consente l'amplificazione uniforme di tutti i trascritti presenti nell'RNA di partenza, grazie a proprietà intrinseche alla Phi29 polimerasi e all'intera procedura di amplificazione. Nella presente invenzione è possibile sottoporre a sequenziamento massivo ed ulteriori validazioni anche campioni di RNA preziosi e
10 disponibili in piccole quantità, altrimenti difficilmente analizzabili, soprattutto per quanto riguarda i trascritti poco rappresentati.

Nella presente invenzione è stato sviluppato un metodo per la preparazione di librerie di cDNA per il sequenziamento massivo definendo la successione dei protocolli da utilizzare. Pertanto forma oggetto dell'invenzione un metodo per ottenere una libreria rappresentativa
15 di cDNA da un campione comprendente le fasi di:

- a) estrazione dell'RNA totale dal campione in cui detto RNA totale è presente in una quantità di almeno 400 ng;
- b) rimozione dell'RNA ribosomiale dall'RNA totale estratto per ottenere una frazione di RNA depleto dell'rRNA in cui detta frazione è presente in una quantità di almeno 30 ng;
- 20 c) retrotrascrizione della frazione di RNA depleto dell'rRNA in cDNA;
- d) purificazione dei cDNA sintetizzati;
- e) ligazione e amplificazione dei cDNA purificati ottenendo una libreria di cDNA di almeno 40 µg,

in cui la fase di retrotrascrizione avviene in presenza di primers fosforilati al 5' aventi la
25 sequenza:

5'-GCGGCCGCNNNNNN-3',

dove N è uno qualsiasi dei quattro nucleotidi A, T, G e C, e i primers sono presenti in quantità equimolari tra loro;

e in cui la fase di amplificazione avviene in presenza di una polimerasi con elevate attività
30 di "strand-displacement" e di "proof-reading", elevate processività e rese.

Preferibilmente la polimerasi è la Phi29 polimerasi.

In un aspetto preferito, i primers sono presenti in una quantità da 25 a 50 µM e l'RNA depleto dell'rRNA è presente in quantità da 30 ng a 300 ng.

E' ulteriore oggetto dell'invenzione una libreria rappresentativa di cDNA ottenuta con il metodo sopra descritto, e il suo uso per il sequenziamento massivo delle sequenze di cDNA ivi contenute.

5 E' ulteriore oggetto dell'invenzione un kit per la preparazione di una libreria rappresentativa di cDNA comprendente da un campione:

- mezzi per estrarre l'RNA totale dal campione;
- mezzi per rimuovere l'RNA ribosomiale dall'RNA totale estratto;
- mezzi per retrotrascrivere l'RNA in cDNA comprendenti primers fosforilati al 5' aventi la sequenza:

10 $5\text{'-GCGGCCGCGNNNNNN-3'}$,

dove N è uno qualsiasi dei quattro nucleotidi A,T, G e C, e i primers sono presenti in quantità equimolari tra loro;

- mezzi di purificazione dei cDNA sintetizzati;
 - mezzi per ligare e amplificare i cDNA sintetizzati purificati comprendenti polimerasi
- 15 Phi29.

E' ulteriore oggetto dell'invenzione una cartuccia ("cartridge") per kit di automazione per attuare il metodo sopra descritto, in maniere manuale automatizzata o semi-automatizzata.

LEGENDE DELLE FIGURE

20 La presente invenzione verrà descritta in esempi non limitativi, facendo riferimento alle seguenti figure:

Figura 1. Profilo Agilent di un RNA di prova sottoposto al "RiboMinus Eukaryote Kit for RNA Seq" (Invitrogen) (in blu) confrontato col profilo del relativo RNA totale di partenza (in rosso). Si noti come i picchi corrispondenti agli RNA ribosomiali 28S e 18S, evidenti
25 nel campione di partenza, risultano quasi assenti nel campione-post Ribominus, che viene così arricchito nella componente non ribosomiale dell'RNA.

Figura 2. Profilo Agilent di: A. una libreria dopo la nebulizzazione e B. dopo la rimozione dei bassi pesi molecolari. Entrambi i profili corrispondono alle condizioni raccomandate per la preparazione di librerie shotgun da sottoporre a sequenziamento con la 454 GS FLX.

30 Figura 3. Risultati del sequenziamento di due campioni, ottenuti con la piattaforma 454 GS FLX della Roche.

MATERIALI E METODI

Estrazione dell'RNA totale

L'RNA totale è stato estratto da linee cellulari e da tessuti utilizzando l'*RNeasy plus mini* kit (QIAGEN Cat. No. 74134), che garantisce un'efficiente rimozione del DNA genomico e un arricchimento in trascritti lunghi più di 200 nucleotidi (gran parte degli RNA al di sotto di queste dimensioni sono selettivamente rimossi).

Rimozione dell'RNA ribosomiale (rRNA)

La sottrazione della componente ribosomiale dall'RNA totale di partenza è stata effettuata utilizzando il *RiboMinus Eukaryote Kit for RNA Seq* (Invitrogen Cat. No. [A1083708](#)).

La rimozione dell'rRNA è stata condotta secondo la procedura indicata dall'azienda produttrice, qui di seguito riportata:

1. Portare a 75°C un bagnetto termostato
2. Ad un tubo sterile di 1,5ml aggiungere:

Componente	Volume per reazione
RNA totale	minimo 400 ng (*) in <20µl
Ribominus probe (15pmol/µl)	10µl
Hybridization Buffer	300µl

(*) Sebbene il kit indichi 2µg come quantitativo minimo del materiale di partenza, gli autori hanno verificato che, partendo da 400-500 ng di RNA totale, è possibile ottenere circa 30-40 ng di RNA privato della frazione ribosomiale, sufficienti per la costruzione e l'amplificazione di una libreria rappresentativa di cDNA, preparata secondo la procedura della presente invenzione.

3. Incubare a 75°C per 5 min per denaturare l'RNA

4. Far raffreddare lentamente il campione e incubare a 37°C per 30 min. Per consentire un'ibridazione sequenza-specifica tra rRNA e *Ribominus probe* (sonde specifiche per l'rRNA eucariotico e recanti biotina al 5') è importante che il raffreddamento avvenga lentamente.

Mentre il campione è a 37°C, preparare le biglie magnetiche rivestite di streptavidina,

5. Risospendere la soluzione contenente le biglie miscelando al vortex
6. Pipettare 750µl della sospensione di biglie in un tubo sterile da 1,5ml
7. Porre il tubo in un separatore magnetico per 1 min, aspettando che le biglie si concentrino lungo la parete del tubo più vicina al magnete, aspirare e gettare il

sovrantante

8. Aggiungere alle biglie 750µl di H₂O-DEPC e risospenderle miscelando delicatamente al vortex
9. Porre nuovamente il tubo nel separatore magnetico per 1 min, aspirare e gettare il sovrnatante
- 5 10. Ripetere i passaggi 8 e 9 una volta
11. Risospendere le biglie in 750µl di Hybridization Buffer e trasferire 250µl di biglie in un nuovo tubo
12. Porre il tubo con i restanti 500µl di biglie nel separatore magnetico per 1 min, aspirare e gettare il sovrnatante
- 10 13. Risospendere il pellet di biglie in 200µl di Hybridization Buffer e porre il tubo a 37°C sino a quando non saranno usate
14. Terminata l'incubazione del campione posto a ibridare a 37°C per 30 min, centrifugare brevemente per raccoglierlo sul fondo del tubo
15. Trasferire il campione (circa 330µl) alle biglie preparate in precedenza (step 13) e miscelare bene
- 15 16. Incubare nel bagnetto a 37°C per 15 min. Durante l'incubazione, miscelare delicatamente ogni 2 min circa. Alla fine, centrifugare brevemente per raccogliere il campione sul fondo del tubo
17. Porre il tubo nel separatore magnetico per 1 min per consentire al complesso rRNA-probe di aderire alle pareti. Recuperare il sovrnatante di circa 530µl contenente l'RNA privo di rRNA
- 20 18. Porre il tubo contenente i 250µl di biglie (step 11) nel separatore magnetico per 1 min, aspirare e gettare il sovrnatante
19. A questo tubo con le biglie attaccate alla parete, aggiungere il sovrnatante di circa 530µl contenente l'RNA di interesse (step 17). Miscelare bene. Quindi l'RNA già privo di rRNA, viene messo ad incubare una seconda volta con le biglie di streptavidina per garantire una migliore rimozione
- 25 20. Incubare il tubo a 37°C per 15 min. Durante l'incubazione, miscelare delicatamente ogni 2 min circa. Alla fine, centrifugare brevemente per raccogliere il campione sul fondo del tubo
- 30 21. Trasferire il sovrnatante contenente l'RNA privo della componente ribosomiale, circa 530µl, ad un nuovo tubo da 2 ml

Gli step successivi, necessari per purificare e concentrare l'RNA isolato, richiedono l'utilizzo del *Ribominus Concentration Module* (Invitrogen Cat. No. [K155005](#)), utilizzato secondo le indicazioni dell'azienda produttrice:

- 5 22. Aggiungere al tubo contenente l'RNA privo della componente ribosomiale, un volume di Binding Buffer e 1 volume di Etanolo 100%. Mescolare bene
23. Caricare 700µl di questa miscela nella colonnina fornita dal kit
24. Centrifugare a $\geq 12000 \times g$ per 1 min a temperatura ambiente. Eliminare l'eluato
25. Ripetere i passaggi 23 e 24 fino a che il campione (step 22) non è interamente caricato sulla colonnina
- 10 26. Aggiungere alla colonnina 200µl di Wash Buffer precedentemente complementato con etanolo 100% (6ml di etanolo 100% per 1,5ml di Wash Buffer)
27. Centrifugare a $\geq 12000 \times g$ per 1 min a temperatura ambiente. Eliminare l'eluato
28. Ripetere il lavaggio con altri 200µl di Wash Buffer complementato con etanolo 100%
- 15 29. Gettare il *Collection Tube* fornito dal kit e porre la colonnina nel *Wash Tube*
30. Centrifugare la colonnina a massima velocità per 2-3 min a temperatura ambiente per rimuovere ogni residuo di Wash Buffer. Gettare il *Wash Tube*
31. Porre la colonnina nel *Recovery Tube* fornito dal kit
32. Aggiungere 10-15µl (•) di H₂O RNasi-free al centro della colonnina. Incubare a
20 temperatura ambiente per 1 min
33. Centrifugare a massima velocità per 1 min a temperatura ambiente
34. Il *Recovery Tube* ora contiene l'RNA privo della componente ribosomiale, purificato e concentrato. Porre subito in ghiaccio per procedere con le applicazioni successive o congelare a -80°C sino all'uso. Questo è l'ultimo passaggio previsto
25 dal kit.

(•) nella presente invenzione il campione è stato eluito in 15µl di H₂O RNasi-free fornita dal kit.

La concentrazione dell'RNA eluito è stata determinata attraverso il Nanodrop 2000c, micro-volume spectrophotometer, Thermo Scientific.

- 30 L'efficienza di rimozione degli RNA ribosomiali dall'RNA totale di partenza è stata valutata per mezzo del bioanalyzer Agilent 2100 usando l'apposito chip per RNA. Come visibile in Figura 1, dove è riportato un caso esemplificativo dei risultati ottenuti, il profilo

relativo all'RNA sottoposto alla suddetta procedura mostra un'efficiente rimozione degli RNA ribosomiali 18S e 28S rispetto all'RNA totale di partenza.

Retrotrascrizione

5 L'RNA depleto della componente ribosomiale è stato convertito in cDNA a singolo filamento utilizzando il "SuperScript III (SSIII) First-strand synthesis system for RT-PCR" (Invitrogen Cat. No. 18080-051).

Il relativo protocollo è stato modificato relativamente ad alcuni componenti, tempi e condizioni di reazione. Estremamente rilevante è stato l'utilizzo di random primers opportunamente disegnati dagli autori per rendere applicabile la strategia di amplificazione del cDNA, prevista dal kit "Quantitect Whole Transcriptome", e inserita nella presente
10 invenzione come step 5' nella procedura qui descritta, alla preparazione del campione per le principali tecnologie di sequenziamento ultra-massivo oggi disponibili. Gli autori hanno anche messo a punto la concentrazione di tali primers nella reazione di retrotrascrizione.

La retrotrascrizione dell'RNA è stata effettuata secondo il seguente protocollo definito dagli autori, che ha sostituito il primo step di sintesi dei cDNA singolo filamento del kit
15 "Quantitect Whole Transcriptome" (QIAGEN Cat. No. 207043):

1'-Preparare la Mix 1 in un tubo eppendorf da 0,5ml, aggiungendo i componenti qui di seguito riportati:

Componente	Quantità/Volume per reazione
RNA post-Ribominus	minimo 30 ng (**)
<i>Tag-Lig-hexamers</i> (50µM)	1µl (***)
dNTPs (10mM)	1µl
H ₂ O-DEPC	sino a 10µl

20 (***) Si potrebbe partire da un quantitativo di RNA depleto della componente ribosomiale inferiore a 30 ng (addirittura anche 1pg), ma in tal modo aumenta il rischio che trascritti a basso numero di copie possano risultare parzialmente rappresentati o assenti nel campione all'inizio dello step di amplificazione del trascrittoma. Questo, a sua volta, non può garantire alta riproducibilità nell'amplificazione dei cDNA. Comunque, si può valutare se
25 partire da un quantitativo di RNA inferiore, in base alle proprie esigenze sperimentali e al numero di reads che lo sperimentatore intende ottenere per ogni trascritto del campione in esame, in particolare per quelli meno rappresentati.

(***) I primers, indicati come *Tag-Lig-hexamers*, sono fosforilati al 5' e hanno sequenza: 5'-GCGGCCGCNNNNNN-3', dove N sta per uno qualsiasi dei quattro nucleotidi. Sono

esameri random preceduti al 5' da una sequenza nota, lunga 8 bp, corrispondente al sito di restrizione dell'enzima "rare-cutter" NotI, presente nel genoma umano molto raramente, all'incirca una volta ogni 10 Mb. La necessità di inserire una sequenza nota al 5' è legata alla natura della procedura di amplificazione utilizzata, del kit "Quantitect Whole
5 Transcriptome" (QIAGEN Cat. No. 207043), che prevede uno step di ligazione di tutte le molecole di cDNA a singolo filamento, a dare concatenameri lunghi circa 20kb. La presenza della sequenza nota consente di identificare i punti in cui è avvenuta la ligazione. Il fatto che sia una sequenza particolarmente rara la rende più funzionale come etichetta dei cDNA sintetizzati.

10

2'-Incubare la Mix1 a 75°C per 5 min

3'-Porre subito in ghiaccio per 5 min.

4'-Durante questo intervallo, preparare la Mix2, aggiungendo ogni componente nell'ordine indicato con le modifiche da apportate dagli autori:

Componente	Volume per reazione
10X RT Buffer per SSIII	2µl
MgCl ₂ (25Mm)	4µl
DTT (0.1M)	2µl
RNaseOUT (40U/µl)	1µl
Trascrittasi inversa SuperScript III (200U/µl)	1µl
DMSO (SIGMA-Aldrich)	0,5µl
Volume totale	10,5µl

15

5'-Aggiungere la Mix2 alla Mix1, mescolare e raccogliere sul fondo per mezzo di una breve centrifugata.

6'-Incubare la mix finale così ottenuta, come di seguito riportato con le modifiche da noi apportate:

Temperatura	Durata
25°C	10 min
45°C	90 min
95°C	5 min

20

7'-Dopo aver bloccato la reazione in ghiaccio, aggiungere 1µl di RNasiH e incubare a 37°C per 20 min.

8'-I cDNA sintetizzati possono essere conservati a -20°C o immediatamente usati per le fasi successive della procedura.

Purificazione dei cDNA sintetizzati

La reazione di sintesi del cDNA è stata sottoposta a purificazione attraverso il kit *MinElute PCR Purification* (QIAGEN Cat. No. 28004), per eliminare primers, nucleotidi, enzimi e sali, seguendo le istruzioni fornite nel kit ed eluendo la miscela di cDNA in 10µl di Buffer EB (10mM Tris·Cl, pH 8,5).

Ligazione e amplificazione dei cDNA

I cDNA purificati e risospesi nel Buffer EB (§) sono stati ligati tra loro in lunghi concatenameri e successivamente amplificati, utilizzando gli step successivi al primo di retrotrascrizione del kit *QuantiTect Whole Transcriptome* (QIAGEN Cat. No. 207043) (vedi protocollo di retrotrascrizione).

(§) E' stato verificato che il Buffer EB non interferisce con le procedure successive di ligazione e amplificazione del cDNA previste dal kit *QuantiTect Whole Transcriptome* (QIAGEN Cat. No. 207043).

Per la reazione di ligazione dei cDNA è stato seguito il seguente protocollo previsto dal suddetto kit, eccetto che per la miscela dei retrotrascritti purificati (step 2'')

1''-Preparare la Mix di ligazione, aggiungendo i singoli componenti nel seguente ordine:

Componente	Volume per reazione
Ligation Buffer	6µl
Ligation Reagent	2µl
Ligation Enzyme1	1µl
Ligation Enzyme2	1µl
Volume totale	10µl

2''-Aggiungere i 10µl della Mix di ligazione ai 10µl della miscela di cDNA purificati, mescolare e centrifugare brevemente.

3''-Incubare a 22°C per 2h,

Terminata l'incubazione:

4''-Allestitire la Mix di amplificazione, che fa uso della Polimerasi Phi29 *REPLI-g* (dal kit *QuantiTect Whole Transcriptome*, QIAGEN Cat. No. 207043), secondo il seguente schema:

Componente	Volume per reazione
REPLI-g Midi Reaction Buffer	29µl
REPLI-g Midi DNA Polymerase	1µl
Volume totale	30µl

5''-Aggiungere i 30µl della Mix di amplificazione ai 20µl della reazione di ligazione, miscelare bene e raccogliere sul fondo tramite breve centrifugata

6''-Incubare a 30°C per 8 ore (o in alternativa per 2h, come indicato nel kit)

7''-Fermare la reazione di amplificazione incubando a 95°C per 5 min

5 8''-Conservare il cDNA amplificato a -20°C fino a successivo utilizzo.

Le librerie di cDNA così ottenute, nella forma di concatenameri lunghi circa 20 kb, sono state sequenziate con ottimi risultati seguendo i protocolli standard di preparazione di librerie shogun per il sequenziamento con la piattaforma 454 GS FLX della Roche.

10 RISULTATI

Esempi di profili Agilent relativi ad una libreria rappresentativa dopo la nebulizzazione del campione amplificato e dopo la rimozione dei frammenti a basso peso molecolare sono riportati in Figura 2A e 2B. I risultati del sequenziamento relativi a due campioni sono presentati in Figura 3. Seguendo il metodo dell'invenzione, è stato ottenuto un numero di
15 "reads" (vedi colonna *passed filter*) molto vicino ai valori corrispondenti alle più alte performance del pirosequenziatore 454 GS FLX, indicate dalla casa produttrice.

Il protocollo sviluppato nella presente invenzione può essere utilizzato per la realizzazione di un nuovo kit e/o cartuccia ("cartridge") per kit di automazione, in grado di garantire, in tempi rapidi e a costi contenuti, l'uniforme e fedele amplificazione di tutti i trascritti in
20 campioni di piccole quantità di RNA da retrotrascrivere (30-40 ng) - precedentemente depleto della componente ribosomiale - grazie alle singolari proprietà della Phi29 Polimerasi. Allo tempo stesso il kit e/o cartuccia ("cartridge") per kit di automazione consente il sequenziamento per studi di trascrittomica e metatracrittomica attraverso le
25 principali piattaforme di NGS (454 GS FLX della Roche, SOLiD dell'Applied Biosystems e Genome Analyzer II dell'Illumina). Grazie alle alte rese di DNA amplificato (~40 µg), inoltre, è possibile utilizzare la libreria di cDNA prodotta per ulteriori successivi esperimenti di validazione. Il kit può quindi rappresentare un'evoluzione del "Quantitect Whole Transcriptome" (QIAGEN), in quanto contiene gli stessi ultimi due step del relativo protocollo, ovvero la ligazione dei cDNA a singolo filamento e l'amplificazione dei
30 concatenameri ad opera della Phi29 polimerasi, mentre ne sostituisce il primo step di retrotrascrizione con quello descritto nei Materiali e Metodi, paragrafo Retrotrascrizione, condotto con i primers disegnati dagli autori e nelle condizioni messe a punto nella

presente invenzione, tali da rendere l'amplificato adatto al successivo sequenziamento ultra-massivo.

5 Considerando la complessità delle procedure previste dalle piattaforme di NGS ed il costo elevato di un singolo run, la disponibilità di un kit unico e standardizzato per la preparazione di campioni di trascrittoma e in grado di ottimizzare le performance del sequenziamento ed il numero di reads utili all'analisi è di estremo interesse per gli utilizzatori. E' da considerare infatti che il numero di piattaforme NGS presenti sul mercato è in continua e rapida crescita e ad oggi solo in Europa si contano: 200 unità per Roche, 210 per Illumina e 80 per Applied Biosystems, molte delle quali sono in service.

RIVENDICAZIONI

1. Un metodo per ottenere una libreria rappresentativa di cDNA da un campione comprendente le fasi di
- 5 a) estrazione dell'RNA totale dal campione in cui detto RNA totale è presente in una quantità di almeno 400 ng;
- b) rimozione dell'RNA ribosomiale dall'RNA totale estratto per ottenere una frazione di RNA depleto dell'rRNA in cui detta frazione è presente in una quantità di almeno 30 ng;
- c) retrotrascrizione della frazione di RNA depleto dell'rRNA in cDNA;
- 10 d) purificazione dei cDNA sintetizzati;
- e) ligazione e amplificazione dei cDNA purificati ottenendo una libreria di cDNA di almeno 40 µg,
- in cui la fase di retrotrascrizione avviene in presenza di primers fosforilati al 5' aventi la sequenza:
- 15 $5\text{'-GCGGCCCGCNNNNNN-3'}$,
- dove N è uno qualsiasi dei quattro nucleotidi A,T, G e C, e i primers sono presenti in quantità equimolari tra loro;
- e in cui la fase di amplificazione avviene in presenza di una polimerasi con elevate attività di "strand-displacement" e di "proof-reading", elevate processività e rese.
- 20 2. Metodo secondo la rivendicazione 1 in cui la polimerasi è la Phi29 polimerasi.
3. Metodo secondo una delle rivendicazioni precedenti in cui i primers sono presenti in una quantità da 25 a 50 µM e l'RNA depleto dell'rRNA è presente in quantità da 30 ng a 300 ng.
4. Metodo secondo la rivendicazione 3 in cui i primers sono presenti in una quantità di
- 25 circa 50 µM.
5. Libreria rappresentativa di cDNA ottenuta con il metodo della rivendicazione da 1 a 4.
6. Uso della libreria rappresentativa di cDNA secondo la rivendicazione 5 per il sequenziamento massivo.
7. Kit per la preparazione di una libreria rappresentativa di cDNA da un campione,
- 30 comprendente:
- mezzi per estrarre l'RNA totale dal campione;
 - mezzi per rimuovere l'RNA ribosomiale dall'RNA totale estratto;

- mezzi per retrotrascrivere l'RNA in cDNA comprendenti primers fosforilati al 5' aventi la sequenza:

5'-GCGGCCGCNNNNNN-3',

dove N è uno qualsiasi dei quattro nucleotidi A,T, G e C, e i primers sono presenti in

5 quantità equimolari tra loro;

- mezzi di purificazione dei cDNA sintetizzati;

- mezzi per ligare e amplificare i cDNA sintetizzati purificati comprendenti polimerasi Phi29.

8. Cartuccia per kit di automazione per attuare il metodo secondo le rivendicazioni da 1 a

10 4.

1/3

Sample 1

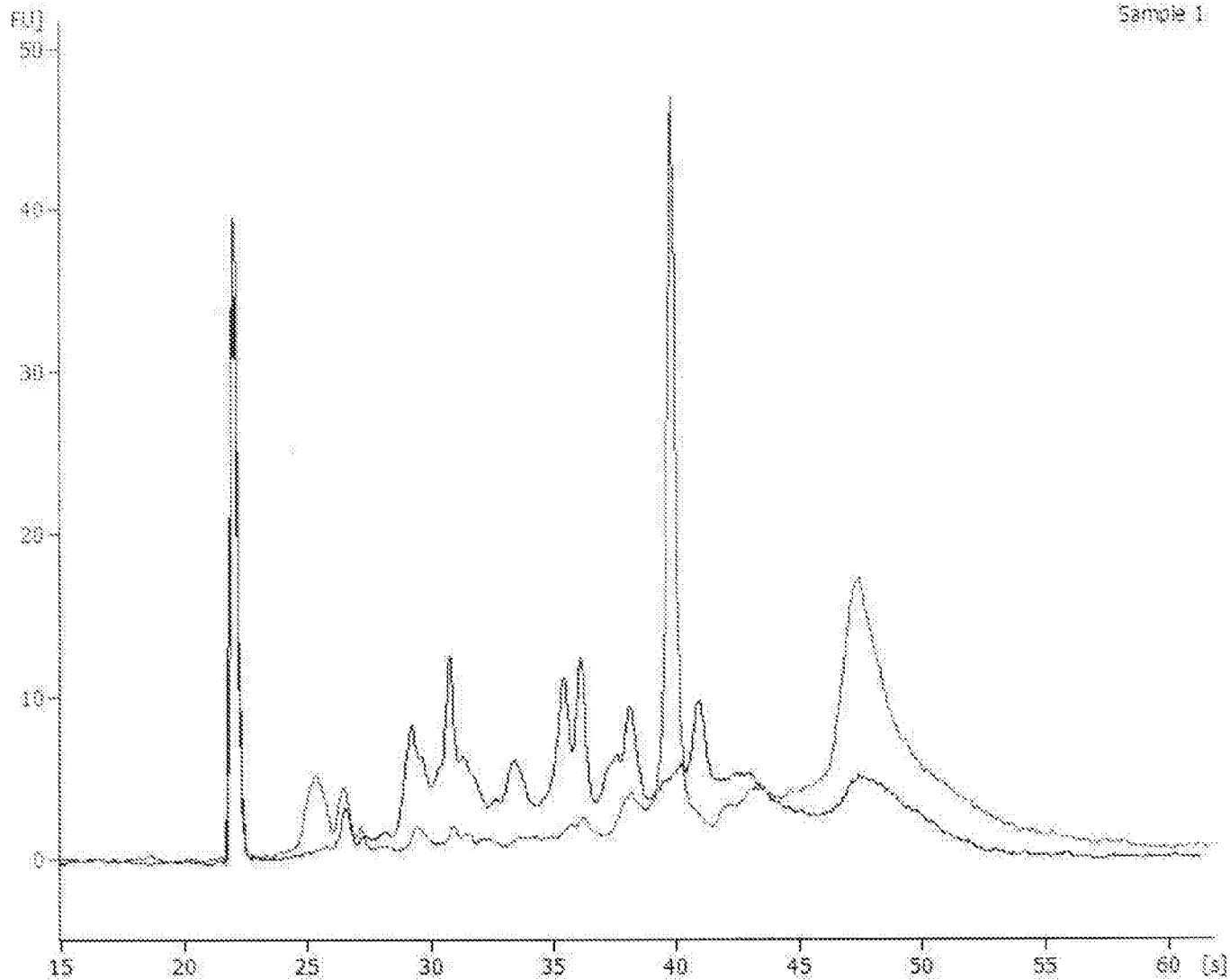


Fig. 1

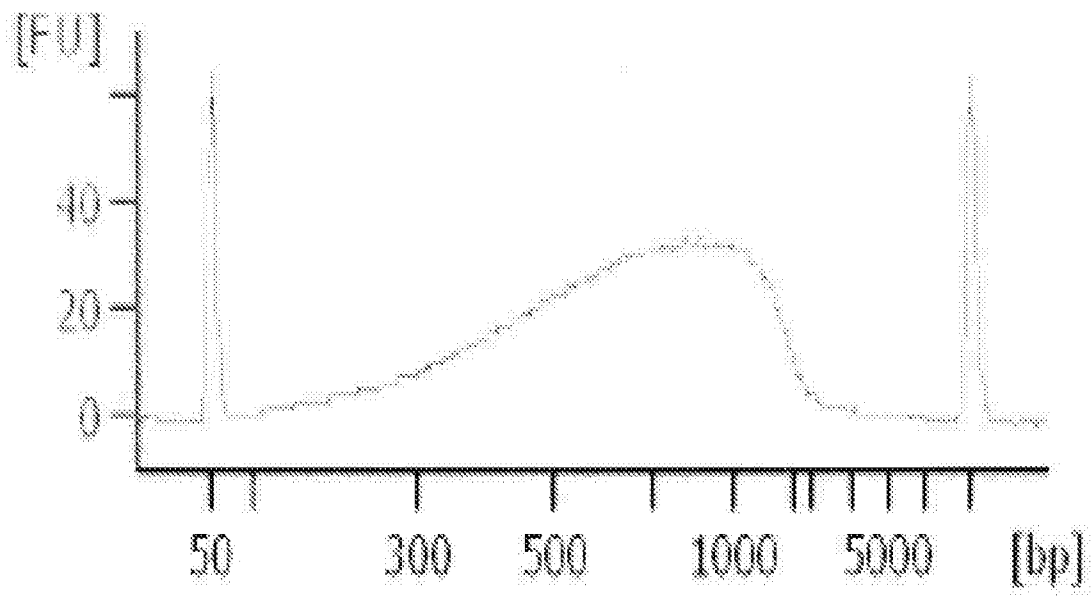
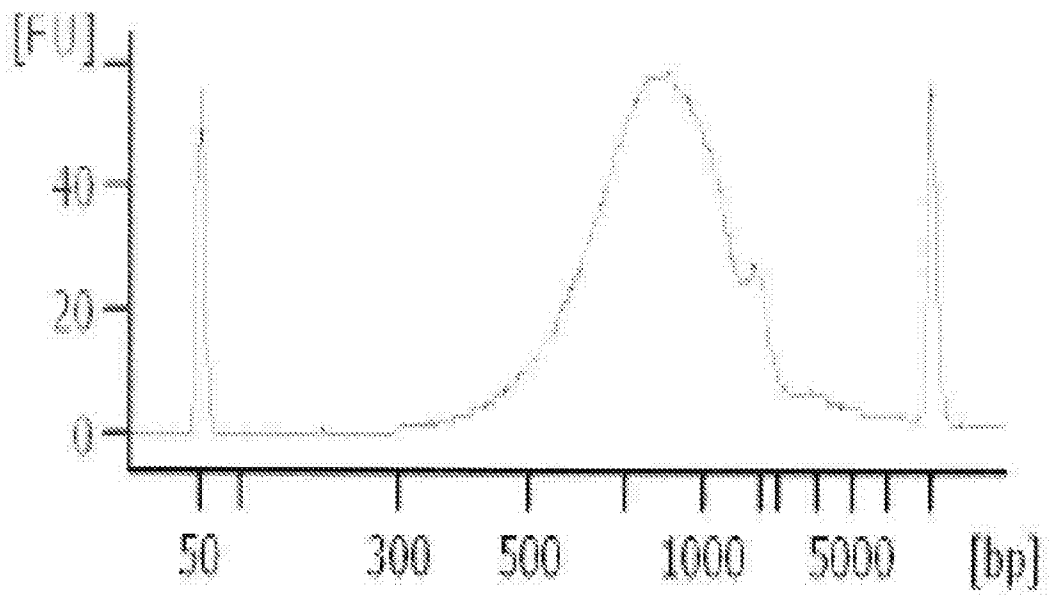
A.**B.**

Fig. 2

Region	Raw Wells	Keypass	Passed Filter (% Keypass)	Avg Length	Total Bases	Dot Failed (% Keypass)	Mixed Failed (% Keypass)	ShortQuality (% Keypass)	Primer Failed (% Keypass)
1	785,502	750,827	400,882 (53.4%)	242.0 ± 136.0	97,001,844 bp	111,773 (14.9%)	36,168 (4.8%)	167,676 (22.3%)	34,328 (4.6%)
2	1,001,160	964,643	468,391 (47.6%)	331.2 ± 146.6	155,119,877 bp	125,647 (12.8%)	86,982 (8.8%)	292,167 (29.7%)	11,456 (1.2%)
TOTAL	1,786,662	1,735,470	869,273		252,151,721 bp	237,420	123,150	459,843	45,784

Fig. 3