



**SARDEGNA
RICERCHE**

**Sardegna FESR 2014/2020 - ASSE PRIORITARIO I
“RICERCA SCIENTIFICA, SVILUPPO TECNOLOGICO E INNOVAZIONE”**

**Azione 1.1.4 Sostegno alle attività collaborative di R&S per lo sviluppo di nuove
tecnologie sostenibili, di nuovi prodotti e servizi**

**PROGETTO CLUSTER TOP DOWN
“NIASMIC - Not Invasive Analysis of Somatic
Mutations In Cancer”
RAPPORTO TECNICO
Descrizione del protocollo sperimentale**



**SARDEGNA
RICERCHE**

Indice

Premessa	3
Trattamento dei campioni biologici e purificazione degli acidi nucleici	5
Preparazione delle librerie genomiche	6
Selezione delle regioni target	7
Sequenziamento	8
Analisi bioinformatica di primo livello	8
Analisi bioinformatica di secondo livello	9



**SARDEGNA
RICERCHE**

Premessa

Il protocollo descritto in questo documento è il risultato di un'attività di ricerca bibliografica e di sperimentazione eseguita ad oggi, per ragioni di causa di forza maggiore dovute allo stato di emergenza sanitaria a causa dell'insorgere della pandemia da COVID-19 ad inizio 2020 e dal perdurare della situazioni di criticità nel periodo Luglio 2020 - Marzo 2021, su un limitato numero di campioni. Considerando la variabilità dei campioni biologici, legata principalmente alla natura dei tessuti di provenienza, le differenze strumentali e di manualità che si possono presentare in diversi laboratori, non possiamo considerare il protocollo che viene descritto un protocollo riproducibile ad ampio spettro.

L'estrazione del DNA da tessuto tumorale (biopsia solida) presenta spesso problemi sia in termini quantitativi (bassa efficienza) che qualitativi (DNA a singola elica, non utilizzabile per le reazioni descritte). Per questo motivo il numero di cicli di PCR indicato nella preparazione delle librerie deve essere modificato a seconda della quantità di dsDNA effettivamente rilevata con saggio fluorimetrico.

I volumi di alcune reazioni sono particolarmente ridotti e necessitano di elevata precisione, ottenibile sia con un'estrema attenzione da parte dell'operatore che utilizzando plastiche di laboratorio (puntali, provette, ecc) di buona fattura e strumenti (pipette, termociclatori) perfettamente calibrati. L'utilizzo delle biglie magnetiche per le numerose purificazioni che si eseguono nelle diverse fasi del protocollo deve essere applicato con la massima precisione al fine di non avere perdite di materiale che, se sommate, portano a un risultato insufficiente in termini quantitativi che si riflette in un eccesso di duplicati di PCR, in particolare per le librerie da biopsia liquida e da biopsia solida.

Tutte le incubazioni a temperatura controllata sono state eseguite su termociclatore Applied Biosystems VeritiPro seguendo le indicazioni riportate sui manuali dei kit utilizzati, sia per le temperature della reazione che per quelle del coperchio della macchina. Le pipette utilizzate sono della marca RAININ e i rispettivi puntali a basso adsorbimento sono con filtro per evitare ogni possibile inquinamento.

Il protocollo può essere considerato aperto, nel senso che esistono sul mercato diversi kit per l'esecuzione delle fasi di preparazione dei campioni che possono essere intercambiabili perché danno risultati sovrapponibili. La continua evoluzione tecnologica sull'argomento non esclude che possano esserci miglioramenti futuri cambiando parte del protocollo o dei reagenti, come dimostrato dalla sempre maggiore disponibilità di reagenti spesso analoghi tra loro ma



**SARDEGNA
RICERCHE**

commercializzati da aziende diverse e dai miglioramenti dei protocolli che si trovano consultando la recente letteratura.

Il protocollo sperimentale messo a punto nel laboratorio NGSC del CRS4 nell'ambito del progetto NIASMIC si articola in diverse fasi:

- Raccolta dei campioni biologici e preparazione dei campioni per la purificazione degli acidi nucleici;
- Preparazione delle librerie genomiche
- Selezione dei frammenti genici di interesse
- Sequenziamento
- Analisi bioinformatica di primo livello
- Analisi bioinformatica di secondo livello



**SARDEGNA
RICERCHE**

PROTOCOLLO

Trattamento dei campioni biologici e purificazione degli acidi nucleici

Per la raccolta, la conservazione e la purificazione degli acidi nucleici dei campioni provenienti dalla chirurgia del reparto di Otorinolaringoiatria del Policlinico Universitario di Monserrato-Cagliari (tumori del distretto testa-collo) sono stati applicate le seguenti procedure.

I campioni di sangue, 2 provette con 6-10ml raccolti in Cell-free DNA Collection Tube (Roche Diagnostics), sono stati processati entro un tempo massimo di 12 ore dal prelievo per minimizzare la perdita di materiale utile alle analisi, in quanto il ctDNA è soggetto ad un elevato turn-over (15 minuti di emivita).

In seguito a una prima centrifugazione a 2500 x g per 10 minuti a 4°C, sono stati trasferiti per ogni prelievo due-tre aliquote da 1,4ml di plasma in provette Eppendorf da 1,5ml. Una seconda centrifugazione del plasma a 16000 x g per 15 minuti a 4°C viene eseguita per eliminare la parte corpuscolata residua e vengono trasferiti 1,3ml di surnatante in provette da 2ml. Le aliquote sono state conservare a -80°C. Inoltre, dopo la prima centrifugazione, dalla provetta di sangue si recupera l'anello di linfociti (buffy coat) in due provette da 2ml con un volume di 200 µl ciascuna per l'estrazione del DNA costituzionale.

Il campione di biopsia tissutale della neoplasia è stato posto nel liquido di trasporto, composto da formaldeide 4% e tampone fosfato 0,05M (Bio-Optica), e inviato al laboratorio. Per la conservazione a -80°C è stato trasferito in una soluzione di etanolo al 70% in acqua pura. L'estrazione del DNA libero circolante (cfDNA) è stata eseguita utilizzando il QIAamp MinElute ccfDNA Mini Kit (QIAGEN) a partire da aliquote di 2,6ml di plasma seguendo le indicazioni del protocollo.

Il DNA costituzionale (gDNA) dei pazienti è stato estratto dai buffy coat mediante l'utilizzo di QIAamp Blood DNA MiniKit (QIAGEN) seguendo il protocollo indicato.

I campioni di tessuto tumorale (circa 25-30mg di tessuto) sono stati sminuzzati meccanicamente e sottoposti a lisi in buffer ATL (QIAGEN) con l'aggiunta di proteinasi K per 3 ore in agitazione a 56°C. La successiva estrazione è stata eseguita seguendo il protocollo del QIAamp Blood DNA Mini Kit (QIAGEN).

Tutti i campioni di DNA sono stati eluiti in acqua e dosati allo spettrofotometro (Nanodrop) e al fluorimetro (Qubit DNA BroadRange) per la valutazione qualitativa e quantitativa degli acidi nucleici.



**SARDEGNA
RICERCHE**

Per i campioni di DNA già purificato provenienti dalle collaborazioni con Università di Sassari e Ospedale Gaslini di Genova, le procedure di raccolta, conservazione e purificazione degli acidi nucleici sono in gran parte sovrapponibili a quelle appena descritte. L'unica eccezione rilevante è rappresentata dal metodo di conservazione utilizzato per il materiale tissutale e dal rispettivo protocollo di purificazione del DNA. Tutti i collaboratori hanno estratto il DNA tumorale da sezioni di tessuto fissato e conservato in formalina (FFPE) utilizzando il GeneRead DNA FFPE Kit (QIAGEN).

Preparazione delle librerie genomiche

Le librerie di DNA sono state preparate con il kit KAPA HyperPlus (Roche Diagnostics) che consiste in diverse reazioni enzimatiche.

Per i campioni di DNA germinale (gDNA) e di DNA tumorale da biopsia solida (tDNA), 150ng di DNA in 17,5ul sono stati sottoposti a frammentazione enzimatica per 25 min a 37°C allo scopo di ottenere frammenti di circa 180-250bp. Per il DNA estratto da plasma (cfDNA) sono stati utilizzati 20ul di volume dell'eluato, a prescindere dalla concentrazione rilevata, in quanto spesso non rilevabile o non replicabile e non correlata ai risultati ottenuti. Data la naturale frammentazione del cfDNA non è necessario eseguire la reazione di frammentazione appena descritta.

Per tutti i campioni (gDNA, tDNA e cfDNA) è stata eseguita una reazione di riparazione delle estremità dei frammenti e l'adenilazione al 3' mediante reazione enzimatica per 30 min a 65°C con il fine di preparare il dsDNA alla successiva reazione di ligazione. Per la ligazione sono stati utilizzati gli adattatori xGen Dual Index UMI Adapters (Integrated DNA Technologies - IDT), alla concentrazione 0,7mM per i campioni di gDNA e tDNA e diluiti 1:5 (0,14 mM) per i cfDNA e la reazione è stata eseguita a 4°C per 16 ore in un volume di 55ml. In questo modo a ogni campione è stata assegnata una coppia di adattatori indicizzati univoci per poter permettere il multiplexing nelle reazioni successive. I prodotti della ligazione sono stati purificati mediante il metodo delle biglie magnetiche (AMPure XP beads, Beckman Coulter), in rapporto biglie/campione = 0,8 per eliminare i dimeri di adattatori. Successivamente, le librerie di gDNA e tDNA sono state sottoposte alla selezione dei frammenti nel range 300-450bp, ottimali per la reazione di cattura delle regioni target. Tale selezione è stata ottenuta utilizzando due diverse concentrazioni di biglie magnetiche: l'esclusione dei frammenti più grandi viene eseguito con un rapporto biglie/campione = 0,7 (35ul di biglie + 50ul di campione) per permettere solo ai frammenti più grandi di legare le biglie che vengono quindi scartate; dopo aver recuperato 80ul di surnatante, la selezione dei frammenti rimasti in soluzioni e con grandezza maggiore di 300bp si ottiene aggiungendo 10ul di biglie, portando quindi il rapporto biglie/campione a 0,8.



SARDEGNA RICERCHE

Le librerie, risospese in 20ul di H₂O, sono state amplificate con un diverso numero di cicli di PCR per ottenere le quantità in ng necessarie alla successiva reazione di ibridazione per la cattura delle regioni target. Per il gDNA e il tDNA sono stati sufficienti 5 cicli di PCR, mentre per ottenere un minimo di 250ng di prodotto dalle librerie di cfDNA sono necessari 9 cicli di amplificazione. Il protocollo di amplificazione prevede una prima denaturazione a 98°C per 45 secondi e l'appropriato numero di cicli di: denaturazione a 98°C per 15 secondi, l'annealing a 60°C per 30 secondi e l'estensione a 72°C per 30 secondi, seguiti da un'estensione finale a 72°C per 1 minuto. I prodotti di PCR, ottenuti con una mix composta da 25ul di 2X KAPA HighFidelity HotStart Ready Mix, 5ul di p5+p7 10X primer mix (concentrazione finale 2uM) e 20ul di DNA ligato, sono stati purificati con biglie magnetiche in rapporto 1:1 e dosati mediante fluorimetro (TECAN Infinite F200Pro + reagenti Qubit Broad Range Kit della Invitrogen). Il controllo di qualità mediante Agilent Bioanalyzer e kit DNA1000 ha evidenziato frammenti di 360bp per i prodotti da gDNA e tDNA. Le librerie ottenute da cfDNA presentano tre picchi: uno principale di 320bp, corrispondente alla grandezza dei frammenti avvolti intorno al nucleosoma più gli adattatori (143bp), uno di 500bp e uno di 700bp, corrispondenti a multipli dei prodotti della frammentazione nucleosomica.

Selezione delle regioni target

Per le reazioni di ibridazione sono stati prodotti dei pool di librerie con quantità equimolari per ciascun campione. I campioni vengono raggruppati in pool di librerie da cfDNA e pool di librerie da gDNA e da tDNA.

In un pool di 12 campioni vengono raccolti 250ng per ciascuna libreria per un totale di 3ug, al quale vengono poi addizionati 5ul di DNA Cot-1 (IDT) per mascherare le sequenze ripetute e 2ul di XGen TS-Universal Blockers (IDT) per mascherare le sequenze degli adattatori. Il volume risultante viene portato a secco in SpeedVac (Savant) e risospeso in 13ul di buffer di ibridazione, composto da 1x Hybridization Buffer e Hybridization Enhancer (IDT). Dopo una denaturazione a 95°C per 10 minuti, la temperatura viene portata a 65°C per 5 minuti prima dell'aggiunta di 4ul di una seconda miscela di ibridazione contenente un inibitore delle RNasi (RNaseOUT, Invitrogen) e il pool di esche biotinilate (bait) a RNA per il pannello di frammenti genici (Agilent). Dopo 1 minuto a 65°C la temperatura di ibridazione viene portata a 62°C per 16-20 ore.

Trascorso il tempo per l'ibridazione delle sonde al DNA target, si aggiungono alla soluzione le biglie magnetiche con streptavidina opportunamente preparate secondo il protocollo del kit xGen Hybridization and Wash (IDT). In seguito al pull-down dei frammenti ibridati mediante l'utilizzo di un rack magnetico, si eseguono una serie di lavaggi allo scopo di eliminare il DNA non ibridato o blandamente legato alle sonde. I lavaggi sono stati eseguiti utilizzando i reagenti e il protocollo che fanno parte del kit IDT, con una modifica che riguarda la temperatura di lavaggio, portata dai 65°C



**SARDEGNA
RICERCHE**

indicati a 62°C. Al termine dei lavaggi le biglie legate ai duplex RNAbait-DNA target vengono risospese in acqua e sottoposte ad amplificazione mediante 9 cicli di PCR utilizzando la miscela di Library Amplification Primer e 25ul di 2X KAPA HighFidelity HotStart Ready Mix fornita con il kit HyperPlus in un volume finale di 50ul. Terminata la reazione i prodotti di amplificazione vengono purificati con 75ul di AMPure XP beads (rapporto biglie/campione = 1,5) e dopo due lavaggi con 200ul di etanolo all'80%, il DNA viene risospeso con 22ul di Resuspension Buffer (Illumina) per recuperare 20ul di prodotto purificato. Si dosa la concentrazione con Qubit BR reagent e si controllano la qualità e la grandezza dei frammenti mediante elettroforesi in microfluidica con il kit DNA1000 e lo strumento Agilent Bioanalyzer.

Sequenziamento

Per il sequenziamento su Illumina HiSeq3000 le librerie sono state diluite alla concentrazione di 2nM, denaturate con NaOH 0,1N per 8 minuti e, dopo neutralizzazione del pH con Tris-HCl 200mM, ai 15ul risultanti sono stati addizionati 35ul della miscela ExAmp preparata seguendo il protocollo del HiSeq3000/400 PE Cluster Kit (Illumina). Per ottenere circa 90 milioni di paired end read per ciascun campione di cfDNA, corrispondente a un coverage teorico di 5000X, un pool di 12 campioni viene caricato su 3 lane di una flow cell. I pool di librerie da gDNA e da tDNA vengono caricati su un'unica lane in proporzione al coverage che si vuole ottenere, rispettivamente 100X e 1000X, corrispondenti a 2,5 milioni di cluster per gDNA e 25 milioni per tDNA.

La configurazione di corsa della flow cell di sequenziamento tiene conto del numero di cicli permessi dall'utilizzo del kit SBS 300 cycle (Illumina), della grandezza degli inserti delle librerie prodotte e della strategia UMI-UD Index utilizzata. Il sequenziatore viene quindi impostato per ottenere la lettura dei frammenti in paired end read da 146 basi, la lettura dell'index 1 e delle sequenze UMI con 17 cicli di sequenziamento e la lettura dell'index 2 con 10 cicli.

Analisi bioinformatica di primo livello

I dati di sequenziamento sono stati sottoposti a demultiplexing con il software Illumina bcl2fastq al fine di attribuire le reads a ciascun campione ed identificare gli UMIs, localizzati nella read di index1. I file in formato fastq sono poi stati allineati contro la sequenza di riferimento del genoma umano e analizzati con i tools FastQC¹ e Picard² per il controllo di qualità.

¹ FastQC: A Quality Control Tool for High Throughput Sequence Data. 2018 Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

² Picard Toolkit, 2019 Broad Institute, GitHub Repository. <http://broadinstitute.github.io/picard/>



**SARDEGNA
RICERCHE**

Analisi bioinformatica di secondo livello

L'analisi di secondo livello prevede l'utilizzo di 6 differenti tool per l'identificazione delle varianti rispetto alla sequenza di riferimento mediante una chiamata combinata del DNA tumorale o del cfDNA con il corrispettivo DNA germinale, in modo da ottenere un incremento in sensibilità nella detection di alleli a bassa frequenza.

Le varianti identificate dai 6 tool vengono poi combinate tra loro e annotate con l'ausilio del database COSMIC³, al fine di selezionare un set di mutazioni somatiche per la validazione finale. L'analisi si conclude con la comparazione delle mutazioni somatiche identificate nel ctDNA e quelle presenti nel DNA proveniente da tessuto biotico. In questo modo è possibile da un lato verificare e quantificare la sovrapposibilità tra le due fonti di DNA, dall'altro identificare mutazioni ad esclusiva rappresentazione nel ctDNA che potrebbero suggerire la presenza di ulteriori masse tumorali o metastasi.

³ COSMIC, (Catalogue Of Somatic Mutations In Cancer) v94, released 28-MAY-21 - <https://cancer.sanger.ac.uk/cosmic>