

Sequenziamento ed analisi con il sistema ION TORRENT

Obiettivi

La piattaforma Ion Torrent di cui dispone Genomnia si compone di due sequenziatori, complementari per capacità e throughput:

1. Ion Personal Genome Machine (PGM)
2. Ion S5

A seconda della profondità di sequenziamento desiderata, dell'applicazione scelta e del tipo di sequenziatore, si possono utilizzare chip differenti che permettono di produrre in una sola corsa da un minimo di circa 500.000 reads fino ad un massimo di 60-80 milioni.

Le caratteristiche di questa piattaforma le conferiscono una versatilità adatta a garantire l'esecuzione di differenti applicazioni:

- SEQUENZIAMENTO MIRATO DI DNA, tramite tecnologia Ampliseq, per lo studio di alterazioni in singoli geni o in pannelli di geni legati ad una patologia;
- TRASCRITTOMICA (Total RNA-seq, Small RNA-seq, trascrittoma Ampliseq);
- SEQUENZIAMENTO MIRATO DI RNA, tramite tecnologia Ampliseq, per lo studio dell'espressione genica differenziale focalizzata su specifici pathways o gruppi di geni;
- EPIGENOMICA, per l'analisi su scala genomica dei siti di legame per fattori trascrizionali (Chip-seq) o delle regioni arricchite in CpG metilate (MBD-seq);
- METAGENOMICA;
- SEQUENZIAMENTO *DE NOVO* O RI-SEQUENZIAMENTO DI INTERI GENOMI virali o batterici.

Il SEQUENZIAMENTO MIRATO DI DNA si esegue quando è nota a priori la sequenza di specifici geni di interesse, siano essi pochi o alcune migliaia come nel caso dell'analisi dell'esoma umano. Il sequenziamento mirato di specifiche regioni può essere utilizzato per valutare la presenza di snv/indel in sequenze genomiche potenzialmente connesse all'insorgenza di malattie ereditarie a carattere mendeliano o di patologie somatiche, quali il cancro. Le regioni genomiche di interesse vengono dapprima amplificate mediante PCR e quindi usate come template nel processo di preparazione delle librerie.

Questa metodica si basa sulla tecnologia **Ion AmpliSeq™** che, utilizzando in multiplex diverse centinaia di coppie di primers, permette di selezionare, amplificandole contemporaneamente, tutte le regioni di DNA bersaglio; i prodotti di amplificazione così ottenuti sono successivamente sequenziati su piattaforma Ion Torrent. Il sistema **Ion AmpliSeq™** richiede come input quantitativi di DNA piuttosto bassi (da 1 a 100 ng a seconda dell'applicazione), permettendo quindi il sequenziamento di campioni disponibili in esigue quantità di partenza (es. biopsie). Le ridotte dimensioni degli ampliconi, inoltre, rendono il sistema adatto anche per l'amplificazione di campioni di DNA parzialmente degradato, come quello estratto da materiale biologico conservato in paraffina (FFPE).

Il sequenziamento dell'esoma, disponibile solo per il DNA umano, è un sequenziamento mirato delle sole regioni codificanti del genoma, che pur rappresentandone solo l'1% contengono circa l'85% delle mutazioni causative di malattie. Lo studio dell'esoma permette quindi di identificare singole varianti nucleotidiche (SNVs), variazioni del numero di copie (CNVs) piccole inserzioni e delezioni (indels), ma anche mutazioni nuove e non ancora caratterizzate alla base di malattie ereditarie a carattere mendeliano o complesse.

Esistono inoltre pannelli a DNA specifici per determinati set di geni, sia pre-disegnati (*ready to use*) da Life Technologies che realizzati su misura da Genomnia. I pannelli disponibili si suddividono in "Hot-Spot", che permettono di esaminare regioni non contigue del genoma alla ricerca di SNP già noti, isolati o organizzati in gruppi, e pannelli per il risequenziamento di intere regioni geniche contigue, alla ricerca di mutazioni sia già note che ancora sconosciute. Per la costruzione di pannelli personalizzati (*custom*) si utilizza il software Ion AmpliSeq™ Designer e le informazioni contenute in database di riferimento quali COSMIC (*Catalogue of Somatic Mutations in Cancer*) per i geni legati al cancro, NCBI ClinVar, per le mutazioni associate a malattie ereditarie, e dbSNP, per una collezione di tutte le alterazioni geniche note. I pannelli personalizzati a DNA permettono il sequenziamento mirato di regioni genomiche di interesse dei genomi di diversi organismi, quali uomo, topo, maiale, pecora, mucca, pollo, mais e altri ancora. È inoltre possibile eseguire l'upload della sequenza di riferimento dell'organismo d'interesse, se non già presente. I primers sono raggruppati in uno o più pool di amplificazione multiplexed; e il numero dei pool è funzione della complessità e delle dimensioni della regione genomica da analizzare.



Per dettagli sulla TRASCRITTOMICA, che comprende le analisi di Total RNA-seq e Small RNA-seq, si rimanda alla scheda specifica.

E' inoltre possibile realizzare l'analisi del trascrittoma utilizzando la tecnologia Ampliseq, attualmente disponibile sia per sequenze umane che per sequenze murine. Il trascrittoma Ampliseq permette di valutare i livelli di espressione genica grazie ad un singolo pannello di primers progettati per consentire l'interrogazione in multiplexing di più di 20.000 geni (oltre il 95% dei geni RefSeq di *Homo sapiens* e oltre il 90% per *Mus musculus*). Il disegno del pannello, composto da piccoli ampliconi (uno per gene), permette di ottenere ottimi risultati anche utilizzando RNA estratti da tessuti fissati, quali campioni fissati in formalina e inclusi in paraffina (FFPE). Il protocollo sviluppato per questa applicazione permette di evitare la selezione dell'RNA poly(A) o l'eliminazione dell'RNA ribosomiale prevedendo l'utilizzo come input di campioni di RNA totale, inizialmente retrotrascritto con primers random. La quantità di RNA totale normalmente richiesta è di soli 10 ng, ma, in caso di particolari necessità, è possibile eseguire la procedura anche a partire da quantità inferiori.

Il SEQUENZIAMENTO MIRATO DI RNA tramite tecnologia Ampliseq si basa sullo stesso principio descritto per il trascrittoma Ampliseq al paragrafo precedente. Pannelli di primers amplificano in multiplex un solo amplicone specifico per ognuno dei geni bersaglio, permettendo l'analisi contemporanea dei livelli di espressione di centinaia o migliaia di geni selezionati. Anche in questo caso sono disponibili pannelli *ready-to-use* e pannelli personalizzati che Genomnia può creare *ad hoc* utilizzando il software Ion AmpliSeq™ Designer. Dove possibile, l'amplicone disegnato su ognuno dei geni bersaglio copre una giunzione esone-esone, così da minimizzare l'amplificazione di eventuali tracce di DNA genomico. Inoltre, la regione target è scelta tra quelle presenti nei trascritti più rappresentati per ogni singolo gene RefSeq. Con la stessa tecnologia di sequenziamento mirato di RNA è possibile analizzare anche la presenza nel trascrittoma di fusioni geniche specifiche.

La lista completa ed aggiornata dei pannelli a DNA e a RNA attualmente disponibili in commercio è consultabile registrandosi al sito www.ampliseq.com

Le analisi di EPIGENOMICA offerte da Genomnia comprendono ChIP-seq, per lo studio su scala genomica di specifiche interazioni DNA-proteina, e MBD-seq, per lo studio delle regioni genomiche arricchite in CpG metilate (riferirsi alla scheda specifica). Con l'analisi ChIP-seq è possibile determinare la distribuzione nel genoma dei siti di legame di fattori trascrizionali o delle modificazioni epigenetiche a carico degli istoni. La cromatina dei campioni, nativa o fissata con formaldeide, viene frammentata per digestione enzimatica o per sonicazione, e quindi immunoprecipitata con anticorpi specifici per la proteina o la modifica epigenetica di interesse. I frammenti di DNA contenuti nei complessi cromatinici immunoprecipitati vengono poi estratti e ligati ad adattatori specifici per la successiva preparazione della libreria. Il mappaggio delle reads ottenute dal sequenziamento della libreria fornisce le coordinate genomiche delle regioni a cui si trovavano associate le molecole del fattore trascrizionale di interesse o della modificazione istonica in esame.

L'analisi di MBD-seq permette invece di determinare la distribuzione genomica delle regioni più ricche in CpG metilate. I campioni di DNA vengono frammentati tramite sonicazione e quindi sottoposti ad un passaggio di selezione mediante *pull-down* con biglie magnetiche funzionalizzate con il *methyl binding domain* della proteina umana MBD. Tale funzionalizzazione permette il legame delle regioni CpG metilate, consentendo la conseguente selezione dei frammenti genomici con livello più alto di metilazione. I frammenti ottenuti vengono usati per costruire librerie il cui sequenziamento consente di mappare sul genoma le regioni con livello di metilazione CpG elevato. Sia ChIP-seq che MBD-seq forniscono informazioni quantitative e permettono quindi di confrontare campioni diversi, eseguendo analisi differenziali.

L'analisi METAGENOMICA di popolazioni miste di microorganismi può essere eseguita mediante sequenziamento delle regioni ipervariabili del gene codificante per l'RNA ribosomiale 16S e permette di identificare le differenti specie batteriche presenti nel campione. Due set di primers disegnati per coprire, rispettivamente, le regioni ipervariabili V2, V4 e V8, e V3, V6-7 e V9 del 16S rDNA, permettono l'identificazione di un ampio spettro di batteri, a livello di famiglia, di genere e in certi casi anche di specie.

Il SEQUENZIAMENTO DI INTERI GENOMI virali e batterici e il sequenziamento *de novo* di microorganismi, può essere eseguito al fine di identificare nuove specie o di monitorare le condizioni di salute e di malattia o di valutare la presenza di specifiche contaminazioni ambientali, tipizzando genomi di virus e batteri noti.

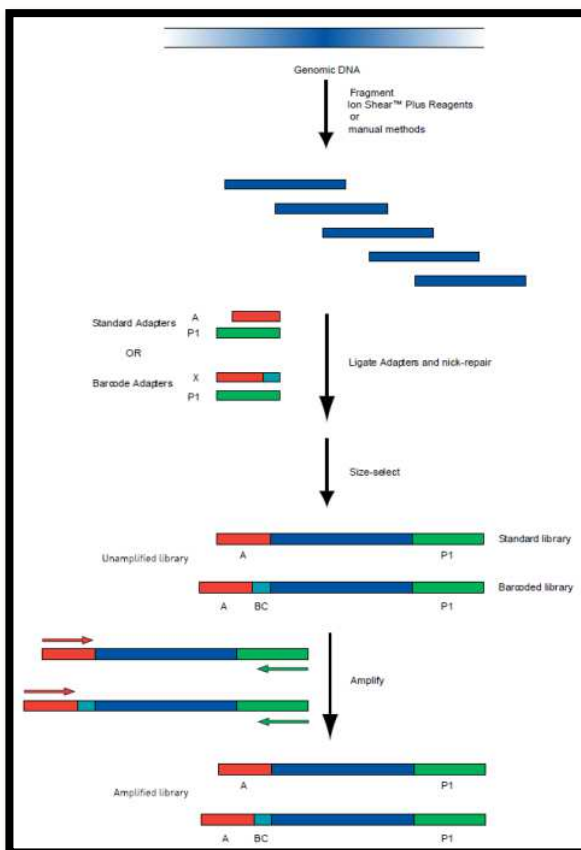
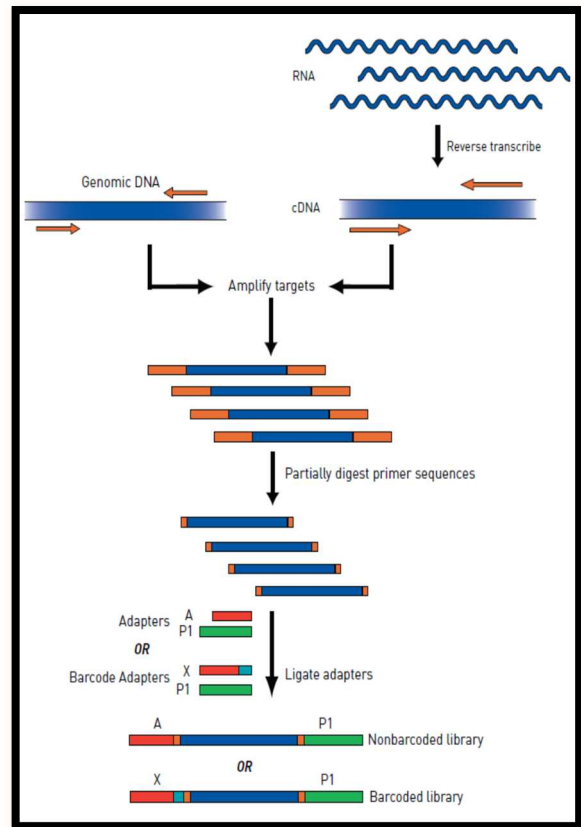


Costruzione delle librerie

Il procedimento per la costruzione delle librerie AmpliSeq è comune per campioni di DNA e di RNA. Come mostrato in figura, se il materiale di partenza è costituito da campioni di RNA viene eseguito un iniziale passaggio di retrotrascrizione in cDNA.

Le regioni bersaglio del DNA genomico o del cDNA sono amplificate in multiplex utilizzando pools di coppie di primers specifici. Gli ampliconi ottenuti vengono quindi trattati enzimaticamente in modo da digerire parzialmente i primers di amplificazione e fosforilare le estremità degli ampliconi stessi. Quindi, alle due estremità di ciascun amplicone vengono ligati covalentemente adattatori specifici (A/P1), richiesti per il sequenziamento con piattaforma Ion Torrent. In questa fase è possibile sostituire l'adattatore A con adattatori dotati di brevi sequenze di DNA differenti che fungono da codici a barre per l'identificazione univoca dei diversi campioni in caso di sequenziamento in multiplex.

I prodotti della ligazione vengono purificati con biglie magnetiche Agencourt AMPure XP (SPRI technology) e sottoposti ad amplificazione tramite PCR, allestita in modo da selezionare solo le molecole dotate degli adattatori ad entrambe le estremità. Le librerie così ottenute vengono sottoposte ad analisi quali-quantitativa mediante fluorimetria (Qubit) o microelettroforesi capillare (Agilent Bioanalyzer). Le librerie sono quindi pronte ad essere sequenziate, singolarmente o in multiplex.



La procedura eseguita per la costruzione di librerie di frammenti prevede l'iniziale frammentazione del DNA genomico mediante digestione enzimatica o per sonicazione, per ottenere frammenti omogenei della grandezza desiderata. In alcuni casi (es. ChIP-seq), il campione di DNA frammentato può essere il prodotto di specifici passaggi di lavorazione precedenti. Le estremità dei frammenti vengono riparate per essere poi unite covalentemente agli adattatori specifici mediante trattamento con ligasi. Le molecole della libreria così ottenute sono quindi sottoposte a selezione dimensionale e amplificate tramite PCR per ottenere la quantità di DNA necessaria a procedere con la preparazione del template sulle particelle IonSphere e con il successivo sequenziamento.

Come già accennato, librerie di frammenti possono essere preparate anche a partire dal DNA estratto da frammenti di cromatina immunoprecipitata con anticorpi specifici (ChIP-seq) o da campioni di DNA arricchito per regioni CpG metilate tramite *MBD pull-down*.

Appartengono a questa categoria di librerie anche quelle utilizzate negli studi di metagenomica. In particolare, le regioni ipervariabili del 16S rDNA sono amplificate con primer specifici e i prodotti di PCR così ottenuti sono usati per la costruzione di libreria di frammenti secondo la procedura sopra descritta.

Caratteristiche quali-quantitativi dei campioni

La quantità e qualità richieste per i campioni di partenza variano a seconda dell'applicazione e del tipo di controllo eseguito sul campione stesso prima di procedere con la costruzione della libreria. Le quantità richieste per la costruzione della libreria con i sistemi AmpliSeq sono comprese tra 1 e 100 ng di acidi nucleici. È tuttavia preferibile disporre di quantità di partenza maggiori, in modo da poter eseguire anche un controllo di qualità iniziale.

Per le applicazioni che comportano la preparazione di una libreria di frammenti sono richiesti da 10 ng ad 1 µg di DNA. Maggiori dettagli saranno specificati in fase di progettazione del servizio.

Analisi bioinformatica

Le analisi bioinformatiche sono differenziate per applicazione, ma, in generale, per i progetti riferiti all'uomo o ad organismi modello, partono tutte dal mappaggio delle read di sequenza sul genoma di riferimento in maniera non ambigua, ad alta precisione e quantificabile con sicurezza.

Il 'segnale' di sequenza (che sia mutazionale, epigenetico o trascrittomico) è poi quantificato e filtrato. Sono utilizzati approcci statistici per la valutazione degli aspetti differenziali (analisi della trascrizione, del controllo trascritturale epigenetico) o per la descrizione del panorama delle variazioni genomiche (SNP, inserzioni/delezioni).

Infine, i risultati sono annotati con riferimento all'annotazione della struttura genica nei suoi aspetti di base (esoni, introni, inizio della trascrizione, promotori, ecc) e in relazione ad elementi di annotazione avanzata specifici per il tipo di applicazione.

I risultati dell'analisi delle mutazioni (da pannelli di geni o esomi interi) vengono analizzati con algoritmi integrati di predizione funzionale e confrontati con liste di putativi geni malattia derivati da letteratura, da database o forniti dal cliente. Sempre per l'analisi di mutazione, disponiamo di una procedura 'ad hoc' per l'analisi di mutazione 'in trio', particolarmente potente per l'identificazione di putativi geni malattia (maggiori dettagli nella relativa scheda servizio).

Per quanto riguarda le analisi epigenetiche (ChIP-seq, MBD-seq), le regioni di interesse vengono annotate in relazione alle isole CpG, alle sequenze ripetute e a tracce UCSC relative a modifiche epigenetiche, come la metilazione e l'acetilazione di istoni. All'analisi MBD-seq è dedicata una scheda servizio specifica.

L'analisi del trascrittoma con tecnologia Ampliseq consente di quantificare il livello di espressione genica e identificare geni differenzialmente espressi tra due gruppi biologici. All'analisi RNA-seq è dedicata una scheda specifica.

Per quanto riguarda l'analisi degli smallRNA, siamo in grado di identificare e quantificare, oltre ai microRNA noti, anche le varianti di sequenza e lunghezza (isomiR) e i putativi nuovi microRNA. Per tutte queste specie molecolari, la procedura analitica si conclude con l'analisi differenziale.

Infine, per gli approcci legati al mondo della microbiologia proponiamo analisi metagenomiche che permettono non solo di identificare la composizione e quantificare la diversità di campioni singoli, ma anche di confrontare statisticamente le dinamiche temporali o dovute a trattamenti farmacologici e/o biologici in diversi campioni.

Informazioni per gli ordini

Richiedere ad admin@genomnia.it un'offerta per l'applicazione di interesse.



Trattamenti per DNA e RNA

Prodotto	N. Catalogo
QC: Controllo di qualità e dimensioni di DNA MBD/CHIP arricchito	DNA05
QC: controllo di qualità di preparazioni di DNA	DNA06
QC: controllo di qualità di preparazioni di RNA totale	RNA03
Purificazione di RNA Poly(A) ⁺ da RNA totale	RNA06
Arricchimento nella frazione degli Small RNA	RNA07
Deplezione di rRNA da RNA totale	RNA10
Arricchimento delle regioni ipervaribili del 16S	16S
Arricchimento per MBD da DNA genomico	MBD
Arricchimento per CHIP da DNA genomico	CHIP

Preparazioni di librerie

Prodotto	N. Catalogo
Disegno e sintesi pannello custom Ampliseq	PAN
Preparazione di libreria DNA con codice a barre	LDb
Preparazione di libreria RNA con codice a barre	LRb

Sequenziamento

Prodotto	N. Catalogo
Sequenziamento forward 200 bp tags con codice a barre	SEQI200B
Sequenziamento forward 400 bp tags con codice a barre	SEQI400B
Sequenziamento forward 600 bp tags con codice a barre	SEQI600B

Analisi Bioinformatica

Prodotto	N. Catalogo
Analisi Bioinformatica I: DNA (pannelli)	DNA-BF01
Analisi Bioinformatica II: DNA (esoma intero)	DNA-BF02
Analisi Bioinformatica III: DNA (esoma intero in trio)	DNA-BF03
Analisi Bioinformatica I: RNA (analisi del trascrittoma)	RNA-BF01
Analisi Bioinformatica II: RNA (isoforme e network)	RNA-BF02
Analisi Bioinformatica I: smallRNA (microRNA noti)	Small-BF01
Analisi Bioinformatica II: smallRNA (nuovi microRNA, isomiR e targets)	Small-BF02
Analisi Bioinformatica I: Metagenomica	METAGEN-BF01
Analisi Bioinformatica I: De Novo	ASSEMBLY-BF01
Analisi Bioinformatica I: Metilazione (identificazione delle regioni metilate)	MBD-BF01
Analisi Bioinformatica II: Metilazione (analisi di metilazione differenziale)	MBD-BF02
Analisi Bioinformatica III: Metilazione (analisi differenziale delle LINE/SINE)	MBD-BF03
Analisi Bioinformatica CHIP-Seq I	CHIP-BF01