



UNIVERSITÀ DEGLI STUDI DEL SANNIO

Dipartimento di Scienze e Tecnologie

Corso di Laurea Triennale in Biotecnologie

Tesi di Laurea in
BIOINFORMATICA

**STUDIO DEL CENTROSOMA E DELLO SPLICING
CITOPLASMATICO ATTRAVERSO DATI DISPONIBILI DI
TRASCRIPTOMICA**

Relatore:

Prof. Francesco Napolitano

Candidato:

Donatella Pierri

Matricola: 171003374

ANNO ACCADEMICO 2022/2023

Sommario

Sommario.....	3
INTRODUZIONE	5
CAPITOLO 1	6
1.1 CENTROSOMA.....	6
1.2 DISFUNZIONE DEL CENTROSOMA	7
1.3CENTRIOLI	8
1.4SPLICING	9
1.4.1. SPLICING ALTERNATIVO	10
1.4.2. SPLICING CITOPLASMATICO	12
CAPITOLO 2	15
2.1 TRASCRITTOMICA	15
2.2MICROARRAY	15
2.2.1 PRODUZIONE MICROARRAY.....	16
2.2.2 ARRAY PER FOTOLITOGRAFIA	17
2.2.3 SPOTTED MICROARRAY.....	18
2.2.4 MICROARRAY DI OLIGONUCLEOTIDI	19
2.3 RNA-Seq	21
2.3.1 SEQUENZIAMENTO COMPLEMENTARE DEL DNA (cDNA-Seq).....	23
2.3.2 SEQUENZIAMENTO DELL’RNA A SINGOLA CELLULA (scRNA-Seq) ..	24
2.3.3 APPLICAZIONI RNA-Seq.....	25

CAPITOLO 3	26
3.1 GSE62933.....	27
3.2 GSE51791.....	31
3.3 GSE199820.....	34
3.4 GSE43143.....	39
3.5 GSE152983.....	43
3.5 GSE56796.....	49
3.6 GSE179439.....	55
CONCLUSIONE	61
Bibliografia.....	63

INTRODUZIONE

Il centrosoma, noto come il centro organizzatore dei microtubuli, è una struttura fondamentale all'interno delle cellule eucariotiche, responsabile dell'organizzazione e del corretto funzionamento del citoscheletro cellulare. È coinvolto in una vasta gamma di processi cellulari importanti, tra cui la divisione cellulare, il trasporto intracellulare e la segnalazione cellulare. La sua complessità e il suo ruolo vitale lo rendono un obiettivo primario per la ricerca scientifica, specialmente quando si tratta di comprendere le dinamiche molecolari che regolano la sua funzione. Negli ultimi anni è anche stato studiato un possibile ruolo del centrosoma nel processo dello splicing dell'RNA, arrivando ad ipotizzare che tale processo possa avvenire anche nel citoplasma. Lo splicing è un meccanismo fondamentale che consente la produzione di varianti di RNA mature a partire da precursori di RNA (pre-mRNA). Questo processo regola la diversificazione delle proteine nelle cellule, permettendo loro di svolgere ruoli specifici in diversi contesti biologici. E qui entra in gioco la bioinformatica.

Attraverso la bioinformatica, e in particolare la trascrittomica, siamo in grado di analizzare enormi quantità di dati genomici e trascrittomici per comprendere meglio i processi cellulari, inclusi quelli che potrebbero coinvolgere il centrosoma e lo splicing citoplasmatico. La bioinformatica ci consente di identificare e analizzare i vari fattori coinvolti nello splicing, compresi gli RNA-binding proteins e i complessi spliceosomiali che regolano questo processo. Studi in corso evidenziano un legame interessante tra il centrosoma e l'ipotesi dello splicing citoplasmatico. Si è scoperto che alcuni geni coinvolti nella regolazione del centrosoma possono essere soggetti a splicing alternativo, il che significa che possono essere processati in modi diversi per produrre varianti di proteine con funzioni diverse. Queste varianti possono influenzare direttamente la funzione del centrosoma e, di conseguenza, il comportamento cellulare. Questa tesi si propone di approfondire la comprensione dei meccanismi legati al centrosoma attraverso l'impiego della bioinformatica, concentrandosi in particolare sui programmi trascrizionali coinvolti. Al fine di raggiungere questo obiettivo, verranno identificati data set, di cui saranno fatte analisi preliminari, per individuare correlazioni significative e tendenze nei dati. Tale approccio consentirà di ottenere una visione più completa e dettagliata dei processi molecolari che regolano il centrosoma e la sua interazione con i programmi trascrizionali. Attraverso l'analisi bioinformatica dei dati di sequenziamento dell'mRNA, si potrebbero identificare varianti di splicing associate al centrosoma e studiarne l'impatto sulla biologia cellulare.

L'unione tra il centrosoma e lo splicing citoplasmatico rappresenta un affascinante campo di studio che offre preziose intuizioni sulla biologia cellulare. Grazie alla bioinformatica, possiamo esplorare questo legame in profondità, aprendo la strada a nuove scoperte e applicazioni nella ricerca biomedica e nella medicina.

CAPITOLO 1

1.1 CENTROSOMA

In biologia molecolare, il centrosoma, noto anche come MTOC (Micro Tubele Organizing Center), è una struttura essenziale posizionata nelle vicinanze del nucleo, responsabile della formazione, della demolizione e dell'organizzazione dei microtubuli all'interno della cellula. Questa struttura, non è delimitata da una membrana, rendendo i suoi confini sostanzialmente indistinguibili. Tuttavia, la presenza di una coppia di strutture cilindriche all'interno del centrosoma, chiamata centrioli, consente di individuare facilmente la posizione. I centrioli sono coinvolti principalmente nella divisione cellulare e nell'organizzazione del citoscheletro nelle cellule animali. Sono cilindri costituiti da nove triplette di microtubuli disposti a raggiera, che insieme formano un cilindro cavo dello spessore di circa 170 nanometri e della lunghezza di 500 nanometri. I centrioli si trovano generalmente in coppia, orientati in modo perpendicolare l'uno rispetto all'altro e spesso sono chiamati centriolo anziano (madre) e centriolo giovane (figlio)

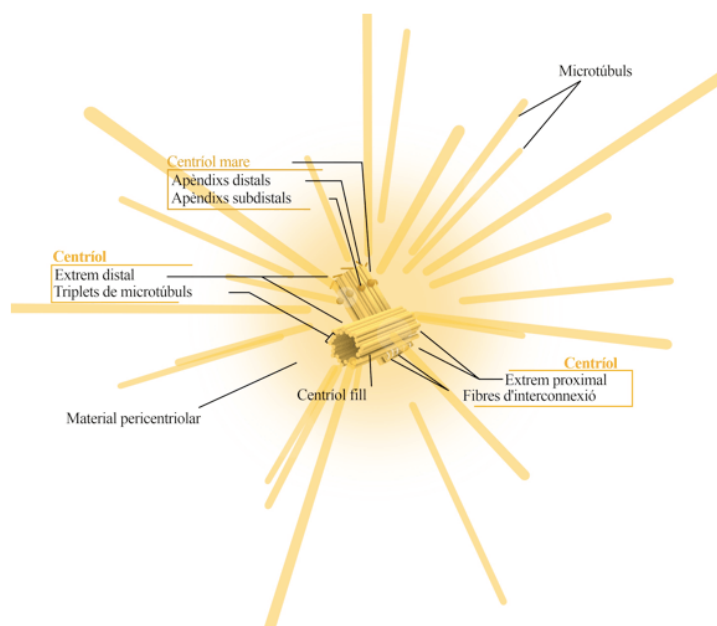


Figura 1: : il centrosoma formato da due centrioli e il citoplasma, dove sono dispersi i dimeri di tubulina.¹

Durante la fase S del ciclo cellulare, prima della separazione, i centrioli si dispongono ad angolo retto tra loro. Si forma quindi un centriolo figlio, perpendicolare a ciascuno dei due originali, garantendo che ogni nuova cellula abbia una coppia di centrioli. Durante la profase mitotica, i due centrosomi risultanti dalla duplicazione dei centrioli si separano e migrano verso i poli opposti della cellula, iniziando ad organizzare il fuso mitotico. Questa struttura costituita da microtubuli, si lega ai cromosomi, separando i cromatidi fratelli e guidandoli verso i poli opposti della cellula per assicurare una distribuzione equa del materiale genico alle cellule figlie. Oltre al ruolo nella formazione del fuso mitotico, i centrioli sono fondamentali per la creazione di protrusioni microtubulari nella cellula, come microvilli, ciglia e flagelli.

1.2 DISFUNZIONE DEL CENTROSOMA

Negli ultimi decenni, è emersa una maggiore comprensione della complessità funzionale e dell'importanza dei centrosomi in vari processi cellulari, tra cui la modellazione, la divisione e la migrazione cellulare. Inoltre, si è scoperto che le anomalie dei centrosomi sono correlate a una vasta gamma di malattie umane e condizioni patologiche, tra cui il cancro, i disturbi riproduttivi, le malattie cerebrali e le ciliopatie. Lo sviluppo del cervello è influenzato da anomalie nel numero e nella struttura dei cromosomi, come dimostrato da varie ricerche (Chavali, Putz, & Gergely, 2014). Malattie genetiche come la microcefalia ereditaria primaria (MCPH), la sindrome di Seckel (SCKL) e il nanismo primordiale osteodisplastico microcefalico di tipo II (MOPDII) sono attribuite a disfunzioni dei centrosomi (Woods, Bond, & Enard, 2005) (Bober & Jackson, 2017). Queste patologie causano riduzione delle dimensioni della corteccia cerebrale e disabilità intellettiva nei neonati affetti, sebbene l'organizzazione complessiva del cervello sia solitamente conservata (Nigger & Holland, 2017). Le mutazioni genetiche associate a MCPH coinvolgono geni che regolano l'assemblaggio e la maturazione dei centrosomi, come CEP135, CEP152, CEP63, CPAP, STIL, CDK5RAP2, SAS6 e PLK4 (Bond, Roberts, Springell, & al., 2005) (Guernsey, Jiang, & Hussin, 2010) (Kalay, Yigit, Aslan, & al, 2011). Le proteine legate ai centrosomi, come ANKA, svolgono un ruolo importante nella regolazione della neurogenesi attraverso l'organizzazione dei microtubuli (Camargo Ortega, Falk, Johansson, & al, 2019). Tuttavia, la relazione tra mutazione genetica e manifestazioni cliniche rimane complessa. Lo studio di disturbi neurologici è stato reso difficile dagli organismi modello incapaci di riprodurre

¹ <https://images.app.goo.gl/XnSTacsARqV2vxju8>

fedelmente le caratteristiche cerebrali umane. Gli organoidi cerebrali 3D in vitro sono emersi come una promettente alternativa, consentendo la ricreazione del cervello umano e della microcefalia (Lancaster, Renner, & al, 2013) (Setia & Muotri, 2019). Ciononostante, gli organoidi cerebrali presentano delle limitazioni nella replicazione completa della funzionalità celebrale umana (Pulvers, Bryk, Fish, & al, 2010). L'utilizzo di cellule staminali pluripotenti indotte (iPSC) da pazienti affetti da microcefalia, insieme all'editing genomico, può contribuire alla comprensione dei meccanismi della malattia. In definitiva, l'integrazione di organoidi cerebrali umani 3D e strumenti genomici avanza la ricerca sulla microcefalia e sullo sviluppo cerebrale umano. Le disfunzioni del centrosoma possono causare una vasta gamma di malattie e disturbi, tra cui l'infertilità dovuta ad anomalie dei centrosomi negli spermatozoi o negli ovociti. Nei processi di oncogenesi, i centrioli vengono eliminati per prevenire la partenogenesi, ma i meccanismi sottostanti rimangono poco chiari. Comprendendo meglio il comportamento dei centrosomi si può aprire la strada a nuove terapie per malattie e disturbi associati alla loro disfunzione.

1.3CENTRIOLI

Negli anni '70, fu osservato che i centrioli erano capaci di auto-replicarsi, questo ha portato all'ipotesi che i centrosomi contengono un proprio set di acidi nucleici (Alliegro, Henry, & Alliegro, 2010). Sebbene sia chiaro che il DNA non si trovi nei centrosomi, è stato osservato che l'RNA si lega ai centrosomi in diversi tipi di cellule, inclusi gli embrioni precoci di *Drosophila*, *Xenopus* e *Zebrafish*, così come nelle vongole e nelle cellule di mammifero coltivate in vitro (Allilegro, Allilegro, & Palazzo, 2006). Questi risultati suggeriscono che la localizzazione dell'RNA e altri meccanismi di regolazione post-trascrizionale potrebbero contribuire all'attività e/o alla funzionalità del centrosoma (Amato, Morleo, Giaquinto, Di Bernardo, & Franco, 2014). I centrosomi stanno assumendo sempre più importanza come siti citoplasmatici per la conservazione di informazioni, arricchiti con molecole di RNA e proteine di elaborazione dell'RNA. Inoltre, ci sono sempre più evidenze a supporto dell'ipotesi che i componenti dello spliceosoma nucleare assemblati al centrosoma svolgono un ruolo di coordinamento nei sottoprogrammi di splicing, nella pluripotenza e nel differenziamento cellulare. (Busselez, Pancione, & al, 2023)

1.4SPLICING

In biologia molecolare e genetica, lo splicing è un processo di modifica del pre-mRNA appena trascritto, durante o dopo la trascrizione, in cui gli introni vengono rimossi e gli esoni vengono uniti. Questo è un passaggio essenziale per la formazione dell'RNA messaggero maturo, il quale può essere poi utilizzato per la corretta traduzione e sintesi delle proteine.

Nei geni eucariotici, la struttura del gene prevede la presenza di regioni codificanti, chiamate esoni, alternate a regioni non codificanti, chiamate introni. Affinché l'mRNA possa svolgere il suo ruolo nella sintesi proteica, è necessario rimuovere gli introni dal trascritto primario, in modo che solo gli esoni rimangano e possano essere tradotti in proteine funzionali.

L'evento dello splicing avviene attraverso due successive reazioni di transesterificazione, dove i legami fosfodiesterici sul pre-mRNA vengono disfatti e riformati. Questo complesso processo è mediato da uno strumento molecolare noto come spliceosoma, composto da circa 150 proteine e 5 RNA. Questi RNA, denominati piccoli RNA nucleari (snRNA), riconoscono le sequenze ai confini tra introni ed esoni, con almeno uno di essi coinvolto direttamente nella catalisi della reazione, comportandosi come un ribozima.

Lo spliceosoma svolge le sue funzioni attraverso una serie di complessi passaggi: le snRNP (ribonucleoproteine nucleari) entrano ed escono dalla reazione in momenti specifici, ciascuna con ruoli distinti. Queste svolgono una varietà di funzioni, tra cui il riconoscimento dei siti di splicing e il coordinamento dei tagli e delle giunzioni dell'RNA. Queste interazioni, insieme a riarrangiamenti successivi, guidano e contribuiscono alla precisione della reazione di splicing.

Le interazioni iniziali coinvolgono l'accoppiamento di basi complementari tra l'snRNA U1 e il sito di splicing 5' del pre-mRNA, mentre nell'ulteriore progressione della reazione, l'snRNA U6 riconosce il sito di splicing. Altri esempi di appaiamenti includono il riconoscimento del punto di ramificazione da parte dell'snRNA U2 e l'interazione tra snRNA U2 e U6, che contribuisce a portare il sito di splicing 5' e il punto di ramificazione vicini. Proteine come il fattore ausiliario U2 (U2AF) e la branch-point binding protein (BBP) sono coinvolte in questi complessi passaggi iniziali. Successivamente, altre proteine facilitano ulteriori interazioni tra l'RNA e le snRNP, contribuendo al processo di splicing.

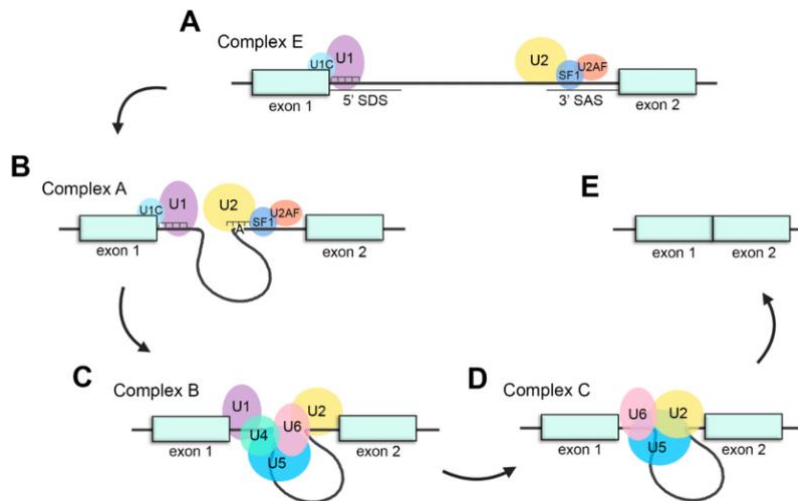


Figura 2: rappresentazione del processo di splicing²

1.4.1. SPLICING ALTERNATIVO

Lo splicing alternativo è un processo molecolare che consente di ottenere diverse varianti di mRNA mature a partire dallo stesso trascritto primario. Questa versatilità aumenta notevolmente la diversità dei trascritti rispetto al numero di geni presenti nel genoma umano, amplificando il repertorio proteico dell'organismo. Dopo la traduzione, queste diverse isoforme proteiche possono svolgere funzioni biologiche differenti all'interno delle cellule. Recenti stime basate su dati di RNA-Seq affermano che il 90%, 60%, e il 25% dei geni Homo Sapiens, Drosophila Melanogaster e Caenothabditis Elegans subiscono splicing alternativo (Auinash & al, 2011), producendo due o più mRNA alternativi che codificano per proteine differenti, ma derivano dallo stesso gene. Questo processo può essere costitutivo, avvenendo normalmente e costantemente in diversi tessuti dell'organismo, oppure può essere regolato, indotto da fattori interni o esterni in determinati momenti metabolici. Praticamente, lo splicing alternativo si riferisce alle diverse modalità con cui può avvenire il processo di splicing dell'mRNA, che è una delle modifiche post-trascrizionali che avvengono durante la trascrizione del DNA. Il prodotto della trascrizione è un pre-mRNA che, prima di partecipare al processo di traduzione, deve subire modifiche e maturazione nel nucleo della cellula. Oltre allo splicing, altre modifiche post-trascrizionali includono il capping 5', la poliadenilazione e l'editing. Lo splicing, sia alternativo che costitutivo, è regolato da proteine che riconoscono sequenze specifiche presenti sia negli introni che negli esoni del pre-mRNA. Queste sequenze possono influenzare la selezione dei siti di splicing,

² https://www.researchgate.net/figure/Schematic-representation-of-the-splicing-process-A-Assembly-of-the-spliceosome-U1_fig2_349280641

determinando quali esoni verranno inclusi nel mRNA maturo. Le sequenze che favoriscono il riconoscimento dei siti di splicing sono chiamate exonic o intronic splicing enhancers (ESE o ISE), a seconda che si trovino sugli esoni o negli introni. Al contrario, le sequenze che ostacolano il riconoscimento dei siti di splicing sono chiamate exonic o intronic splicing silencers (ESS o ISS). Nella maggior parte dei casi, le sequenze enhancer sono riconosciute da proteine della famiglia SR (serine-arginine rich), mentre le sequenze silencer sono riconosciute da proteine hnRNP (heterogeneous nuclear ribonucleoproteins). Gli eventi di splicing alternativo sono il risultato della variazione nella capacità delle proteine di legare queste sequenze regolatrici. La "forza" di riconoscimento di un sito di splicing, e quindi la probabilità di splicing alternativo, dipende non solo dalla presenza di ESE, ISE, ESS e ISS nella sequenza del pre-mRNA, ma anche dalla disponibilità delle proteine che le legano e dalla loro interazione con il complesso dello spliceosoma. Si stima che ogni gene umano dia origine, in media, a circa 4 proteine diverse. Questa diversificazione proteica è in parte spiegata dal meccanismo di splicing alternativo, il quale consente la produzione di diverse varianti di mRNA e, di conseguenza, di proteine, anche da un singolo gene. Questo meccanismo offre una spiegazione parziale del motivo per cui negli organismi superiori non vi sia un rapporto lineare tra il numero di geni e la complessità dell'organismo, in quanto la varietà delle proteine dipende anche dalla capacità di utilizzare più o meno diffusamente lo splicing alternativo. Un processo analogo è coinvolto nell'espressione dei geni delle immunoglobuline nei linfociti, permettendo la produzione di un'ampia gamma di anticorpi con diversi specifici siti di legame antigene. Inoltre, gli eventi di splicing alternativo sembrano giocare un ruolo significativo nella variabilità genetica, contribuendo alle differenze tra individui della stessa specie, e possono influenzare anche l'evoluzione introducendo nuove varianti genetiche e funzioni proteiche (Keren, 2010). Esistono diverse modalità di splicing alternativo, tra cui:

- Salto dell'esone: (exon skipping): un esone può essere eliminato dal trascritto primario. Questa è la modalità più comune di splicing nei mammiferi.
- Esone mutualmente esclusivo: solo uno dei due esoni alternativi viene mantenuto nel mRNA maturo, non entrambi
- Sito di taglio alternativo 5': viene utilizzato un sito di taglio al 5' alternativo, cambiando l'estremità 3' dell'esone a monte.
- Sito di taglio alternativo 3': viene utilizzato un sito di taglio al 3' alternativo, cambiando l'estremità 5' dell'esone a valle.

- **Introne trattenuto:** i siti di taglio di un introne possono non essere riconosciuti, quindi l'introne non viene eliminato dal trascritto primario. La differenza con il salto dell'esone risiede nel fatto che nella modalità dell'introne trattenuto la sequenza mantenuta non è circondata da introni. Nel caso in cui l'introne trattenuto si trovi all'interno della regione codificante, è essenziale che non alteri la cornice di lettura degli esoni. Qualora avvenisse tale alterazione, potrebbe risultare nella produzione di una proteina tronca o non funzionale.

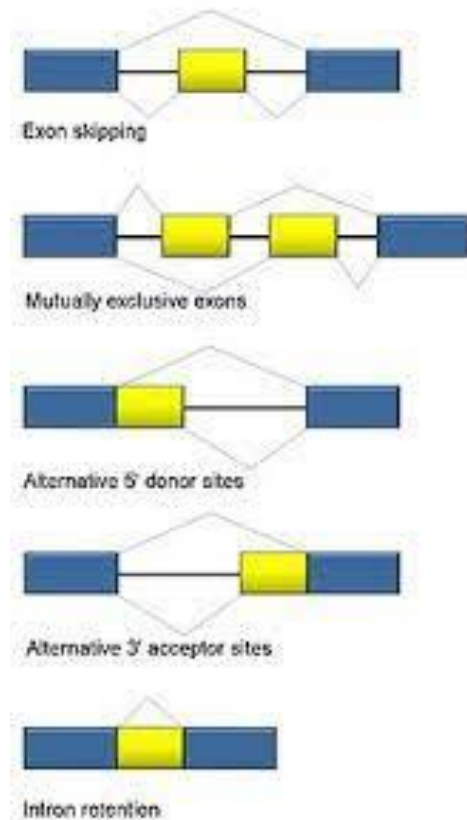


Figura 3: eventi di splicing alternativo³

1.4.2. SPLICING CITOPLASMATICO

Secondo l'ipotesi dello splicing citoplasmatico, questo potrebbe avvenire in presenza di introni e sequenze introniche presenti all'interno dei trascritti che lasciano il nucleo. La ritenzione degli introni, insieme al salto dell'esone, all'utilizzo di esoni mutuamente esclusi e ai siti di splicing alternativi donatori o accettori, rappresenta una delle modalità di splicing alternativo. Gli introni trattenuti fungono da substrato per lo splicing citoplasmatico e

³ <https://it.wikipedia.org/wiki/Splicing>

costituiscono una base per l'identificazione dei componenti dello spliceosoma con attività citoplasmatica (Wang, et al., 2008)

L'attività dello spliceosoma nel citoplasma è un argomento controverso e ampiamente discusso. Lo spliceosoma, è diviso in complesso maggiore e minore, con il complesso maggiore che funziona nel nucleo, dove si presume che si assembli su ogni nuova molecola di pre-mRNA trascritta. Tuttavia, lo spliceosoma minore opera su una sottoclasse di introni che si trovano a frequenze molto inferiori nei genomi eucariotici. Questi introni si distinguono per la presenza di siti di splicing 5' e sequenze di punto di ramificazione altamente conservate e differenti rispetto alla maggior parte degli altri introni. La localizzazione cellulare in cui opera lo spliceosoma minore è stata oggetto di intenso dibattito e studio (Turunen, Niemala, Verma, & Frilander, 2013). Uno studio (Konig, Matter, Bader, Thiele, & Muller, Splicing segregation: the minor spliceosome acts outside the nucleus and controls cell proliferation., 2007) riporta che lo spliceosoma minore funziona prevalentemente nel citoplasma e controlla la proliferazione delle cellule agendo sugli introni di tipo u12. Tutto questo mette in discussione e smentisce alcune pubblicazioni che mostravano il funzionamento dello spliceosoma minore nel nucleo (Konig, Matter, Bader, Thiele, & al, Splicing segregation: the minor spliceosome acts outside the nucleus and controls cell proliferation., 2007) (Pessa, Will, Meng, Schneider, & al, 2008). Effettivamente, la frequenza e l'intensità dello splicing citoplasmatico, controllato dallo spliceosoma, sono probabilmente molto inferiori rispetto a quanto osservato nel centro di elaborazione cellulare principale. Questo fenomeno sembra manifestarsi solo in circostanze estremamente specifiche e specializzate, come nei neuroni o nei megacariociti, dove le cellule sono altamente compartimentate.

La presenza di splicing citoplasmatico è inequivocabile poiché è stato osservato in piastrine enucleate e in dendriti neuronali isolati, privi di nucleo. Tuttavia, è importante notare che, sebbene lo splicing extranucleare possa verificarsi, il suo meccanismo specifico è ancora sconosciuto e potrebbe differire significativamente da quello osservato durante lo splicing nucleare. Un esempio pionieristico di splicing citoplasmatico nei mammiferi è stato presentato da Denis et al (Denis, Tolley, Bunting, Schwartz, & al, 2005) utilizzando le piastrine. Hanno dimostrato che i trascritti di IL1b che contengono un introne trattenuto si accumulano nelle proiezioni pro-piastrina e persistono dopo la formazione e la maturazione delle piastrine. Questi trascritti citoplasmatici con introni trattenuti non solo evitano lo splicing convenzionale ed escono dal nucleo, ma sfuggono anche ai fattori di decadimento nonsense mediato. I trascritti di IL1b con introni trattenuti nelle piastrine persistono fino

all'attivazione delle piastrine, quando vengono sottoposti a splicing nella piastrina enucleata e danno origine al normale RNA messaggero maturo di IL1b, che viene quindi tradotto in proteina IL1b funzionale. È importante notare che le piastrine mature sono cellule enucleate, quindi per definizione qualsiasi splicing che possa verificarsi deve avvenire al di fuori del nucleo. Questi risultati dimostrano che esiste un sistema endogeno di splicing extranucleare all'interno di una cellula di mammifero e hanno un'evidente rilevanza biologica. Durante lo stesso periodo, sia le proteine che i costituenti acidi nucleici dello spliceosoma sono stati osservati nei dendriti dei neuroni primari di ratto (Glanzer, Miyashiro, & Sul, 2005). Sebbene queste molecole possano svolgere una funzione secondaria al di fuori del nucleo cellulare, lo stesso studio ha rivelato che i dendriti isolati meccanicamente, separati fisicamente dal soma e dal nucleo, erano in grado di eseguire lo splicing di costrutti di RNA reporter. Questo risultato ha dimostrato la possibilità che il processo di splicing avvenga al di fuori del nucleo dei neuroni. L'individuazione delle molecole dello spliceosoma nel citoplasma neuronale suggerisce che alcuni trascritti potrebbero subire un'elaborazione simile allo splicing convenzionale. È rilevante notare che lo splicing dendritico dei costrutti reporter in questo studio non ha coinvolto siti di splicing canonici. Anche se alcune delle sequenze risultanti dallo splicing nei dendriti hanno mostrato aderenza alle sequenze donatrici e accettrici canoniche, nessun esempio ha rispettato entrambe le sequenze canoniche. È possibile che ciò sia dovuto all'introduzione di una sequenza di introne trattenuto artificiale o alla saturazione della macchina dello splicing; un'altra spiegazione potrebbe essere l'esistenza di un codice di splicing citoplasmatico.

Questi risultati hanno altresì suggerito la presenza di trascritti citoplasmatici endogeni che ospitano substrati intronici trattenuti nei dendriti e potrebbero essere soggetti a splicing extranucleare. Poiché queste proiezioni dendritiche si trovano a notevoli distanze dal corpo cellulare, un evento di splicing locale potrebbe evitare la necessità di trasportare molecole da e verso il nucleo per l'elaborazione. Inoltre, in un modo simile alla regolazione della traduzione dendritica in risposta alla stimolazione sinaptica, lo splicing dendritico potrebbe rappresentare un nuovo metodo di regolazione genica post-trascrizionale, analogamente a quanto osservato con le piastrine contenenti IL1b.

CAPITOLO 2

2.1 TRASCRITTOMICA

La trascrittomica è un campo della biologia molecolare che si occupa dello studio dei trascritti di RNA in uno specifico tipo di cellula o tessuto. Le tecniche di trascrittomica possono fornire preziose informazioni sulla regolazione genica, sulle vie metaboliche coinvolte in vari processi biologici e sulla risposta cellulare a stimoli esterni; esse includono l'uso di DNA microarray e la più recente RNA-Seq.

2.2 MICROARRAY

La scoperta dei microarray ha rivoluzionato la ricerca sul DNA e sull'RNA, consentendoci di analizzare parallelamente diverse decine di migliaia di trascritti per gene ad un costo notevolmente ridotto. Questo è stato possibile grazie alla disponibilità di sequenze genomiche completamente sequenziate. Prima del 2000, il numero di organismi sequenziati era piuttosto piccolo, tra cui 38 batteri, 1 fungo (*S.cerevisiae*), 2 invertebrati (*Caenorhabditis Elegans* e *Drosophila Melanogaster*) e 1 pianta (*Arabidopsis Thaliana*) (Volker & al, *Microarrays as Research Tools and Diagnostic Devices*, 2015). Dopo circa 10 anni, il catalogo degli organismi sequenziati comprendeva oltre 16.000 progetti (Prutt, Tatusova, Brown, & al, 2012). L'esplosione del sequenziamento del genoma è dovuta allo sviluppo di un'ulteriore tecnologia all'inizio del ventunesimo secolo, definita sequenziamento di nuova generazione (NGS) (Margulies, Egholm, Altman, & al, 2005). Dall'ultimo aggiornamento del NCBI (ottobre 2013), il numero è aumentato a 24.788 progetti di genoma procariotici registrati che rappresentano 4.528 specie diverse. Contemporaneamente al completamento delle prime sequenze genomiche, è stata sviluppata la tecnologia dei microarray di DNA. I primi esperimenti si sono concentrati sulla proliferazione dell'espressione utilizzando come organismo modello *S. cerevisiae* per monitorare i cambiamenti dell'attività trascrizionale di ogni gene noto. Per *S. cerevisiae*, dopo il completo sequenziamento del genoma sono state fatte le prime analisi trascrizionali (DeRisi, Iyer, & Brown, 1997); (Hauser, Vingron, Scheideler, & al, *Transcriptional profiling on all open reading frames of Saccharomyces cerevisiae*, 1998); (Wodicka, Dong, Mittman, & al, 1997), è stato il modello per lo sviluppo dei metodi biochimici e bioinformatici necessari per la proliferazione trascrizionale. Il processo di microarray comprende diversi passaggi:

- **PREPARAZIONE DEL CAMPIONE:** Bisogna estrarre l'RNA totale dal campione biologico di interesse, convertirlo in cDNA (DNA complementare) utilizzando un enzima chiamato trascrittasi inversa.
- **ETICHETTATURA DEL cDNA:** Il cDNA viene etichettato con un marcatore fluorescente o un altro agente rilevabile.
- **IBRIDAZIONE:** Il cDNA etichettato viene ibridato su un microarray, che è un supporto solido (vetrino o chip di vetro) su cui sono stampate delle sonde oligonucleotidiche complementari ai geni di interesse. Ogni sonda corrisponde ad un gene specifico.
- **SCANSIONE E ANALISI:** Dopo l'ibridazione, il microarray viene scansionato per rilevare l'intensità del segnale fluorescente su ciascuna sonda. L'intensità del segnale è proporzionale alla quantità del trascritto di RNA corrispondente al campione. I dati scansionati vengono quindi analizzati per identificare i geni che sono differenzialmente espressi tra i campioni o per altri scopi di studio.

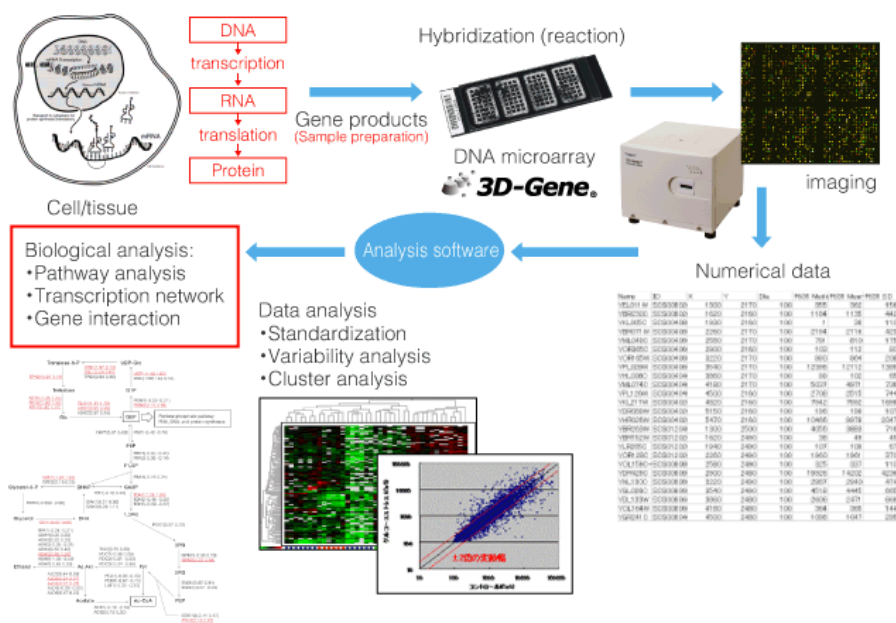


Figura 4: DNA microarray⁴

2.2.1 PRODUZIONE MICROARRAY

I microarray possono essere fabbricati usando diverse tecnologie, come la stampa di micro solchi, con un microspillo appuntito su una lastrina di vetro dove verrà attaccata covalentemente la sonda (probe) di materiale genetico ottenuta per clonazione sfruttando la

⁴ https://www.3d-gene.com/en/about/chip/chip_003.html

PCR (fotolitografia). Questa tecnica richiede apparecchi robotici molto sofisticati, il centro dell'apparecchio è costituito da pennini che prelevano uno o più campioni di cDNA e poi vengono trasferiti sui vetrini. Una volta che la sonda è su vetrino si effettua il processing, dove la sonda viene legata al supporto tramite una reazione innescata dalla luce ultravioletta. Infine il cDNA viene reso a singola catena attraverso una denaturazione termica o chimica. Con questa tecnica però era possibile creare microarray a bassa densità.

2.2.2 ARRAY PER FOTOLITOGRAFIA

In questo caso gli oligonucleotidi vengono sintetizzati in situ, questa tecnica è stata utilizzata per la prima volta dall'Affimetrix, che ne detiene il brevetto. Con la tecnica della fotolitografia è possibile sintetizzare molte migliaia di differenti oligonucleotidi sulla superficie di un vetrino. Anche se questa tecnica dà una sintesi molto accurata, la massima lunghezza degli oligonucleotidi che è possibile raggiungere è di 25 nucleotidi, ma ovviamente oligonucleotidi di queste dimensioni non sono sufficienti a dare specificità al microarray, per questo servono 3 oligonucleotidi che legano il gene, e altri 3 che presentano un mismatch per il controllo negativo. Sui microarray a bassa densità solitamente si usano marcatori radioattivi, per ottenere buoni microarray è essenziale isolarli dall'umidità e dalla polvere, per questo motivo esistono camere apposite per l'ibridazione dei microarray che vengono sigillate. Il DNA fluorescente dei campioni sperimentali viene mescolato con un DNA di un controllo marcato con un colorante fluorescente diverso, così che se la quantità di RNA espressa da un gene nella cellula di interesse è aumentata (up regolata) rispetto al campione, lo spot sarà del colore del primo fluorescente. Viceversa se l'espressione del gene è diminuita (down regolata) rispetto al campione di riferimento lo spot sarà colorato del secondo fluorescente. La fluorescenza viene rilevata grazie ad uno scanner a laser, il segnale rilevato dallo scanner viene sottoposto ad algoritmi di filtrazione e di pulizia e convertito in valori numerici. Oggi con questa tecnica è possibile analizzare i livelli di espressione di un singolo gene ottenendo degli ottimi risultati, invece la combinazione dello studio di molte migliaia di geni risulta molto complicato e può portare spesso a dei falsi positivi, questo accade a causa della mancanza di standardizzazione o dal fatto che alcuni cDNA possono cross-ibridare altre sonde, che avrebbero dovuto rivelare altri geni.

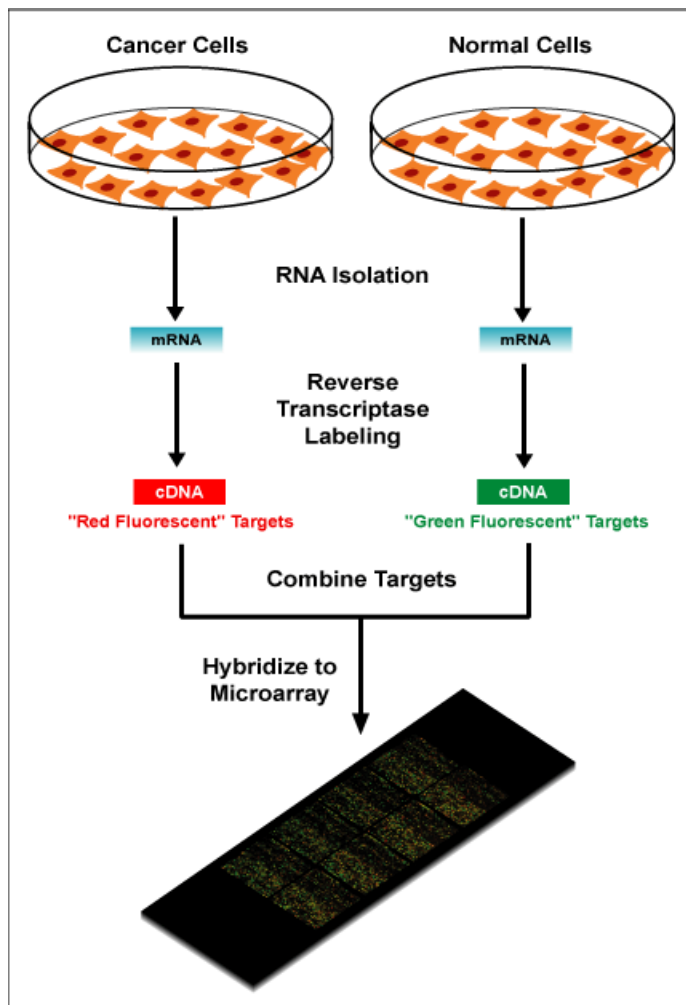


Figura 6: tecnica di microarray spottato dove si possono esaminare i profili di espressione di due diversi campioni cellulari
<http://it.wikipedia.org/wiki/Microarray#/media/File:Microarray-schema.gif>

2.2.4 MICROARRAY DI OLIGONUCLEOTIDI

Nei microarray di oligonucleotidi (o single-channel microarrays), i probes vengono costruiti per riconoscere parti di sequenze di mRNA. Vi sono matrici microarray di tal specie commercializzate da numerose ditte come GE Healthcare, Affimetrix, o Agilent, contengono oligonucleotidi importanti per alcune analisi di routine o grosse parti di genomi; inoltre possono essere prodotte matrici su richiesta per soddisfare bisogni per la ricerca o per la diagnostica. Gli Array oligonucleotidici possono essere prodotti o per deposizione piezoelettrica dell'intera lunghezza dell'oligonucleotide, o per sintesi in sito (fotolitografia). Array di oligonucleotidi lunghi sono composti da 60 basi e prodotte con la tecnologia ink-jet printing su substrati di silicio. Array di oligonucleotidi corti sono composti da 25\30 basi e sono prodotti per sintesi fotolitografica su substrati di silicio (Affimetrix) o per deposizione piezoelettrica su matrice di acrilamide (GE Healthcare).

Più recentemente la NimbleGem System ha sintetizzato nuove matrici dette Maskless Array, che possono essere utilizzate in numerosi oligonucleotidi test (probe): i nucleotidi che formeranno gli oligonucleotidi presentano un gruppo protettore fotolabile che, quando è presente, ne impedisce il legame dell'oligonucleotide in crescita. Questo gruppo può essere allontanato con una fonte luminosa che permette agli oligonucleotidi di reagire. Si usano delle maschere per determinare quali nucleotidi in quale posizione devono essere attivati dalla luce, in questo modo possono essere costruite sequenze oligonucleotidiche specifiche in determinate posizioni.

Oltre a queste classiche tecnologie, sono stati sviluppati anche sistemi basati su semiconduttori. Per esempio, CombiMatrix utilizza un sistema di questo tipo per produrre dei propri microarray. (Volker & al, *Microarrays as Research Tools and Diagnostic Devices*, 2015). Migliaia di microelettrodi di platino vengono attivati simultaneamente per sintetizzare singoli oligonucleotidi attraverso una precisa sintesi digitale controllata. L'attivazione di ciascun microelettrodo avviene attraverso una reazione elettrochimica che porta alla produzione di acido. Questo processo determina la deprotezione del filamento oligonucleotidico in crescita, preparandolo per la successiva fase di sintesi. Grazie alla capacità di indirizzare singolarmente migliaia di microelettrodi in parallelo, è possibile produrre comodamente singoli microarray (Ghindilis, Smith, Schwarzkopf, & al, 2007). Questa tecnica consente di sintetizzare simultaneamente oltre 12.000 oligonucleotidi, ciascuno lungo 50 basi. Oltre al DNA naturale, sono state impiegate varianti non naturali come L-DNA (Hauser, Vingron, Scheideler, & al, *Transcriptional profiling on all open reading frames of Saccharomyces cerevisiae*, 1998) e PNA (Busselez, Pancione, & al, 2023) (Sforza, Tedeschi, Bencivenni, & al, 2014), al fine di evitare la cross-reattività e aumentare l'interazione con i bersagli nel campione applicato. Le sonde di L-DNA offrono il vantaggio di interagire con la stessa cinetica del suo filamento antiparallelo, come accade nel d-DNA. Ciò significa che le conoscenze acquisite con i microarray di d-DNA possono essere applicate agli equivalenti l-DNA. Per legare bersagli come RNA, DNA, peptidi/epitopi o molecole chimiche, vengono utilizzati i così detti array di codici postali (Volker & al, *Microarrays as Research Tools and Diagnostic Devices*, 2015). Questi array utilizzano sonde di fusione, che legano i bersagli in soluzione utilizzando un oligomero naturale di d-DNA complementare al bersaglio DNA\RNA, fuso in un l-DNA. Questo oligomero di l-DNA si lega a un oligonucleotide l-DNA complementare immobilizzato sulla superficie del microarray, garantendo un'ibridazione priva di sfondo sugli array. Tuttavia, il costo elevato per la sintesi di oligomeri di l-DNA può rappresentare un ostacolo per ulteriori sviluppi.

D'altra parte, i P-DNA, analoghi sintetici dell'acido nucleico con una spina dorsale pseudopeptidica, mostrano eccellenti proprietà nel riconoscere specifiche sequenze e sono meno suscettibili alle variazioni della forza ionica. Dal momento che non sono carichi, manca la repulsione tra la spina dorsale caricata negativamente presente nei duplex di DNA, con conseguente maggiore affinità, come mostrato dalle curve di fusione più elevate dei duplex PNA/DNA o PNA/RNA rispetto alle loro controparti naturali (Brandt & Hoheisel, 2004) (Jacob, Brandt, Stephan, & al, 2004) (Sforza, Tedeschi, Bencivenni, & al, 2014). Questa tecnica presenta diversi limiti: innanzitutto, è necessario conoscere la sequenza dei geni di interesse per progettare le sonde da utilizzare sul microarray. Inoltre, durante l'ibridazione, possono verificarsi "rumori di fondo" dovuti a appaiamenti incompleti o imprecisi, riducendo così la specificità della tecnica. Infine, un punto cruciale che ha spinto verso l'adozione delle tecnologie NGS è la quantificazione imprecisa, che viene espressa con un valore continuo anziché con una misura precisa (Wang & Gerstein, 2009). I problemi con i microarray includono artefatti di ibridazione incrociata, scarsa quantificazione di geni bassi e altamente espressi e la necessità di conoscere la sequenza a priori (Kukurba & Montgomery, 2015)

A causa delle limitazioni tecniche riscontrate, la trascrittomica ha adottato metodi basati sul sequenziamento che hanno subito un'evoluzione significativa nel corso degli anni. Questi approcci sono passati dal sequenziamento Sanger delle librerie di tag di sequenza espressa, ai metodi che sfruttano tag chimici come nell'analisi seriale dell'espressione genica, per giungere infine alla tecnologia attuale, che impiega il sequenziamento di nuova generazione del DNA complementare (cDNA), in particolare RNA-Seq.

2.3 RNA-Seq

RNA-Seq, noto anche come sequenziamento dell'RNA, è una tecnica innovativa che sfrutta il sequenziamento di nuova generazione (NGS) per rivelare la presenza e la quantità delle molecole di RNA presenti in un campione biologico. Il processo di RNA-Seq coinvolge diversi fasi:

- **PREPARAZIONE DEL CAMPIONE:** L'RNA viene estratto dal campione biologico di interesse, che può essere tessuto, cellule o altro materiale biologico. L'RNA totale viene trattato per eliminare l'RNA ribosomiale e arricchire l'RNA messaggero (mRNA) e altri RNA di interesse.
- **SEQUENZIAMENTO:** L'RNA viene convertito in cDNA utilizzando la trascrittasi inversa. il cDNA viene amplificato e sequenziato utilizzando tecniche di

sequenziamento ad alta efficienza, come il sequenziamento ad alto rendimento (Next Generation Sequencing, NGS).

- **ANALISI BIOINFORMATICA:** I dati di sequenziamento vengono analizzati con metodi bioinformatici per mappare i frammenti di cDNA sul genoma di riferimento e quantificare l'abbondanza relativa dei trascritti. Questo permette di identificare i geni differenzialmente espressi, gli eventi di splicing alternativo, le modifiche post-trascrizionali e altro ancora.

L'RNA-Seq offre la possibilità di analizzare i trascritti alternativi derivati dallo splicing genico, le modifiche post-trascrizionali, le fusioni geniche, le mutazioni\SNP ⁶ e i cambiamenti dell'espressione genica nel corso del tempo, nonché le differenze nell'espressione genica tra gruppi o trattati diversi (Maher, et al., 2009).

Oltre all'mRNA, l'RNA-Seq consente l'analisi di diverse popolazioni di RNA, che comprendono l'RNA totale, piccoli RNA come miRNA e tRNA, nonché il profilo ribosomiale (Ingolia, Brar, Rouskin, McGreachy, & al, 2012). Questa tecnica può essere impiegata anche per definire i confini degli esoni\introni e per verificare o modificare i confini dei geni 5' e 3'. Altri sviluppi significativi includono l'analisi dell'alterazione del numero di copie, la rilevazione della contaminazione microbica, l'identificazione degli elementi trasponibili, la deconvoluzione⁷ dei diversi tipi cellulari e la scoperta dei neoantigeni⁸, tutti resi possibili grazie ai progressi negli algoritmi bioinformatici (Thind, et al., 2011)

⁶ Polimorfismo a singolo nucleotide, una variazione del materiale genico dovuta da un unico nucleotide.

⁷ Meccanismo di filtraggio

⁸ I neoantigeni sono marker presenti sulla superficie delle cellule tumorali ma assenti nel tessuto normale, possibili target per lo sviluppo di nuovi vaccini

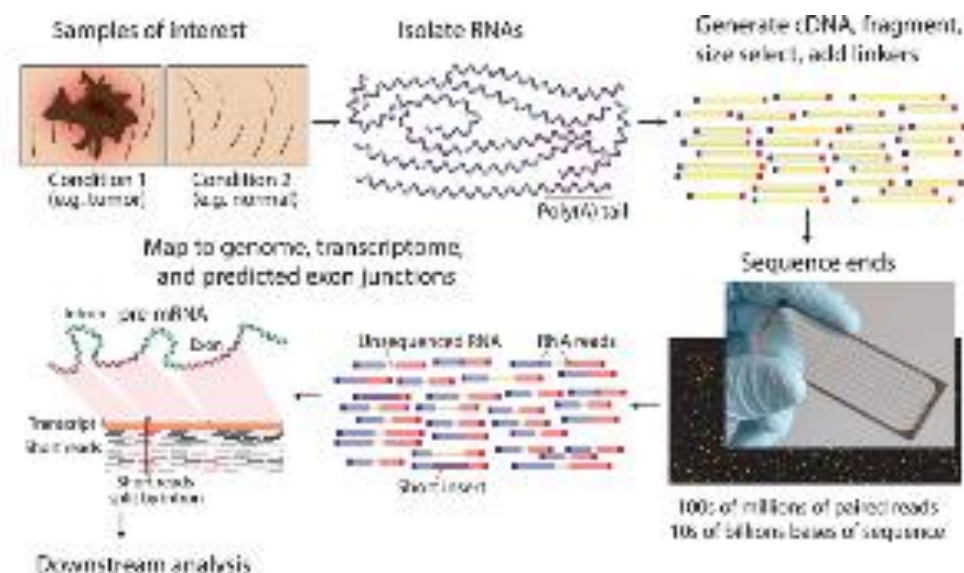


Figura 7: tipico esperimento di RNA-Seq

Griffith M, Walker JR, Spies NC, Ainscough BJ, Griffith OL (August 2015). "Informatics for RNA Sequencing: A Web Resource for Analysis on the Cloud". PLOS Computational Biology. 11 (8): e1004393. Bibcode:2015PLSCB..11E4393G. doi:10.1371/journal.pcbi.1004393. PMC 4527835. PMID 26248053.⁹

2.3.1 SEQUENZIAMENTO COMPLEMENTARE DEL DNA (cDNA-Seq)

La libreria di cDNA, ottenuta dai vari tipi di RNA, viene quindi sottoposta a sequenziamento per essere convertita in un formato leggibile dal computer. Esistono numerose tecnologie di sequenziamento ad alto rendimento per il cDNA, sviluppate da aziende come Illumina, Thermo Fisher, BGI/MGI, PacBio e Oxford Nanopore Technologies (Oikonomopoulos, Bayega, Fahiminiya, Djambazian, & al, 2020). Per il sequenziamento a breve lettura, Illumina è una delle tecnologie più diffuse. In questa tecnica, gli adattatori vengono legati al cDNA, il quale poi viene ancorato ad una cella di flusso. I Cluster sono generati attraverso cicli di amplificazione e denaturazione del ponte, mentre la sintesi della sequenza avviene tramite cicli di sintesi di filamenti complementari ed eccitazione laser di basi con terminatori reversibili. La scelta della piattaforma di sequenziamento e dei relativi parametri è guidata dal disegno sperimentale e dai costi. Tra le considerazioni più comuni del disegno sperimentale vi sono la determinazione della lunghezza del sequenziamento, la profondità del sequenziamento, l'opzione tra sequenziamento singolo o a due estremità, il numero di repliche, l'utilizzo del multiplexing, la randomizzazione e l'aggiunta di spike-in (Conesa, Madrigal, Tarazona, Gomez-Cabrero, & al, 2016).

⁹ Griffith M, Walker JR, Spies NC, Ainscough BJ, Griffith OL (August 2015). "Informatics for RNA Sequencing: A Web Resource for Analysis on the Cloud". PLOS Computational Biology. 11 (8): e1004393. Bibcode:2015PLSCB..11E4393G. doi:10.1371/journal.pcbi.1004393. PMC 4527835. PMID 26248053.

2.3.2 SEQUENZIAMENTO DELL'RNA A SINGOLA CELLULA (scRNA-Seq)

Tecnologie convenzionali come i microarray e l'analisi standard del RNA-Seq valutano l'espressione genica di RNA proveniente da ampie popolazioni cellulari. Tuttavia, in popolazioni cellulari eterogenee, tali approcci possono mascherare le rilevanti differenze esistenti tra le singole cellule all'interno di queste popolazioni (Shapiro, Biezuner, & Linnarsson, 2013) (Kolodziejczyk, Kim, Svensson, Marioni, & al, 2015)

Il sequenziamento dell'RNA a singola cellula fornisce un quadro dettagliato degli schemi di espressione genica nelle singole cellule. Anche se non è sempre possibile ottenere un profilo completo di tutti gli RNA espressi in ciascuna cellula, data la limitata quantità di materiale disponibile, è possibile identificare modelli di espressione genica tramite analisi di raggruppamento genico. Questo approccio consente di individuare tipi cellulari rari all'interno di una popolazione, talvolta sconosciuti fino a quel momento. Ad esempio, nel 2018 sono state scoperte rare cellule specializzate nel polmone, chiamate ionociti polmonari, che esprimono il regolatore della conduttanza transmembrana della fibrosi cistica, attraverso l'analisi di dati scRNA-Seq condotta su epitelii delle vie aeree polmonari (Montoro, Haber, Biton, Vinarsky, & al, 2018) (Plasschaert, Žilionis, Choo-Wing, & al, 2018)

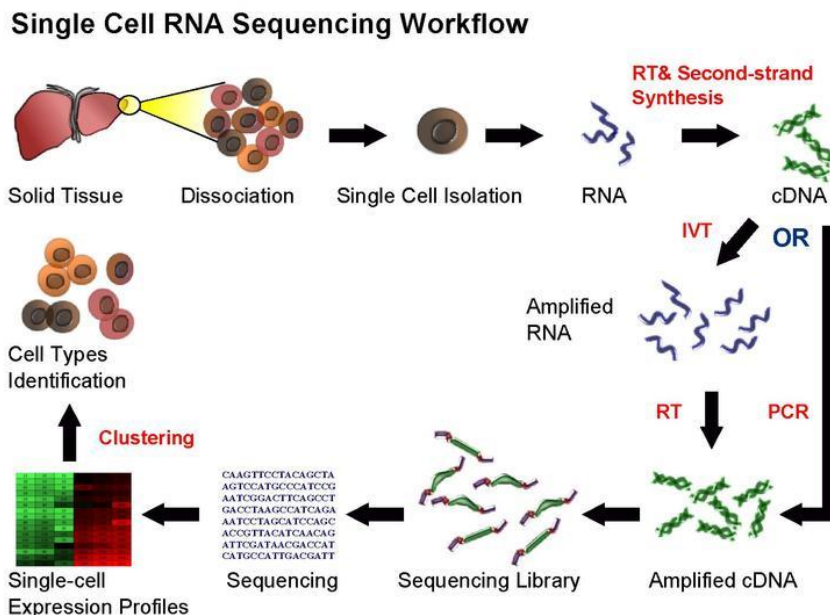


Figura 8: Single-cell RNA sequencing¹⁰

¹⁰ https://en.wikipedia.org/wiki/File:RNA-Seq_workflow-5.pdf

2.3.3 APPLICAZIONI RNA-Seq

Un'applicazione comune ma estremamente efficace dell'RNA-Seq è la ricerca delle differenze nell'espressione genica tra due o più condizioni, ad esempio tra campioni trattati e non trattati. Questo processo è noto come analisi dell'espressione differenziale e produce output sotto forma di geni differenzialmente espressi (DEG). Questi geni possono essere regolati in su o in giù, indicando un aumento o una diminuzione nell'espressione rispetto alla condizione di riferimento. Numerosi strumenti sono disponibili per l'esecuzione di analisi dell'espressione differenziale, la maggior parte dei quali sono implementati in ambienti come R, Python o attraverso la riga di comando Unix. Tra gli strumenti comunemente utilizzati vi sono DESeq, Edgar (Robinson, McCarthy, & Smyth, 2010) e voom+limma, tutti accessibili tramite R/Bioconductor ([Bioconductor - Open source software for bioinformatics](#))

Le analisi per un elenco di geni differenzialmente espressi sono di due versioni, che servono a convalidare le osservazioni e a fare interferenza biologica. A causa delle sfide legate all'espressione differenziale e all'RNA-Seq, le importanti scoperte vengono confermate mediante un metodo ortogonale applicato agli stessi campioni, come la PCR in tempo reale, oppure attraverso un altro esperimento, talvolta preregistrato, condotto su una nuova coorte di campioni. Il metodo più comune per ottenere una compressione biologica di livello superiore di risultati è l'analisi di arricchimento dei set di geni¹¹, determina se la sovrapposizione tra due set di geni è statisticamente significativa, in questo caso la sovrapposizione tra geni differenzialmente espressi e set di geni provenienti da pathway/database noti (*ad esempio*, Gene Ontology, KEGG, Human Phenotype Ontology) o da analisi complementari negli stessi dati (come le reti di co-espressione). Gli strumenti comuni per l'arricchimento del set genico includono interfacce web (*ad esempio*, ENRICHR, g: profiler, WEBGESTALT (Liao, Wang, Jaehnig, & al, 2019) e pacchetti software.

Il trascrittoma, ossia l'insieme delle molecole di RNA prodotte in una cellula in un dato momento, è estremamente complesso e dinamico, influenzato dall'ambiente circostante, dallo stadio del ciclo cellulare, e da fattori esterni come condizioni patologiche. Questa complessità riflette la diversità delle condizioni fisiologiche e patologiche che le cellule possono incontrare. Il confronto tra il trascrittoma di cellule sane e malate può rivelare differenze nell'espressione genica, consentendo di comprendere meglio le basi molecolari delle malattie e di sviluppare terapie mirate. Questo tipo di analisi ha notevoli implicazioni

¹¹ metodo per identificare classi di geni o proteine che sono sovrarappresentati in un ampio insieme di geni o proteine e possono avere un'associazione con fenotipi diversi

cliniche, permettendo una diagnosi precoce e una prognosi più accurata, nonché lo sviluppo di strategie terapeutiche personalizzate.

L'analisi dei trascrittomi può coinvolgere campioni biologici diversi, dai fluidi corporei ai tessuti. I cambiamenti nell'espressione genica sono associati a una vasta gamma di patologie, molte delle quali hanno una base genetica multifattoriale. Le nuove conoscenze omiche hanno sottolineato l'importanza delle interazioni complesse tra geni e molecole, portando a un concetto più ampio di malattia, che non si limita più a cause monogeniche.

La trascrittomica riveste un ruolo fondamentale nella diagnosi prenatale di patologie genetiche e neurologiche, consentendo una comprensione precoce delle malattie che si manifestano solo in fasi successive dello sviluppo. Nelle malattie come la sclerosi multipla, che possono presentare alterazioni fisiopatologiche prima dei sintomi clinici, l'analisi precoce dei trascrittomi può guidare sia la comprensione della malattia che l'anticipazione delle cure.

Nell'ambito oncologico, l'analisi dei trascrittomi è fondamentale per classificare i tipi di tumori, prevedere la risposta alla terapia e identificare nuovi bersagli terapeutici. Le modificazioni nel trascrittoma possono precedere i cambiamenti istopatologici, offrendo un'opportunità unica per intervenire precocemente nella progressione della malattia.

Infine, i profili di espressione genica possono essere utilizzati per valutare la risposta ai farmaci, identificando i geni coinvolti nella sensibilità ai trattamenti e personalizzando le terapie in base alle caratteristiche individuali dei pazienti (Collotta, 2018).

CAPITOLO 3

Sono stati presi in considerazione vari studi da cui sono derivati diversi data set (GSE62933, GSE51791, GSE19980, GSE43143, GSE152983, GSE56796, GSE179439) al fine di ottenere una comprensione più approfondita dell'espressione dei geni centrosomiali durante il processo di differenziamento cellulare. Per ottenere una comprensione più approfondita dei data set, è essenziale utilizzare strumenti avanzati come il database GRAIN per identificare i geni differenzialmente espressi. Successivamente, attraverso l'analisi dei geni identificati con il sito DAVID, è possibile condurre un'analisi di arricchimento per identificare le pathway biologiche coinvolte. Durante questa fase, si considera sia il p-value, che fornisce una misura della significatività statistica, sia il metodo di correzione di Benjamini-Hochberg, che tiene conto del numero di test effettuati per ridurre il rischio di falsi positivi. Per lo scopo della tesi è importante selezionare le pathway più rilevanti, valutando sia la significatività statistica che la rilevanza biologica delle pathway identificate.

Verranno quindi citate le pathway che sono più pertinenti al contesto della ricerca, tenendo conto dei risultati ottenuti dall'analisi di arricchimento e della loro relazione con l'oggetto di studio della tesi.

3.1 GSE62933

Lo studio (Zhang, Zheng, Li, & al, 2018), affronta la crescente carenza di donatori di fegato, ad oggi l'unico metodo disponibile ed efficace per i pazienti con insufficienza epatica, presentando il trapianto di epatociti come valida alternativa. Tuttavia, gli epatociti adulti mostrano limitata attività proliferativa in vivo e sono difficili da conservare, limitando le loro applicazioni. Le cellule simili agli epatociti (iHeps), ottenute dalla riprogrammazione diretta dei fibroblasti o dalla differenziazione delle cellule staminali pluripotenti indotte umane (iPSC), hanno mostrato limitazioni nella funzionalità epatocellulare a causa di una conversione incompleta. Allo stesso modo, una riprogrammazione incompleta delle iPSC può portare alla carcinogenesi in vivo, riducendo le loro applicazioni cliniche. Pertanto, le cellule staminali epatiche umane (HpSC) della linea epatica sono emerse come una promettente fonte alternativa agli epatociti primari.

Le HpSC, isolate dal fegato fetale e postnatale, possono differenziarsi in epatociti in vitro ed in epatociti più maturi in vivo, superando così le limitazioni delle iHeps. Le HpSC esprimono vari marcatori, tra cui CD29, CD49f, CD13, CD117, CD133, EpCAM e DLK, oltre a marcatori di cellule staminali e di linee epatocitarie e colangiocitarie come l'albumina (ALB), l'alfa-fetoproteina (AFP), la citocheratina 18 (CK18) e CK19) e, cosa più importante, marcatori di cellule staminali ematopoietiche e mesenchimali (cioè CD34, CD45, CD90 e CD109). Tuttavia, il profilo di espressione dei marcatori specifici per isolare le HpSC dalle cellule epatiche fetali primarie umane (FLC) non è ancora chiaro.

CD90 (Thy-1) è stato originariamente individuato nelle cellule staminali ematopoietiche¹² e mesenchimali¹³ ed è stato anche identificato come un marcatore delle cellule ovali. È stato osservato nelle cellule staminali/progenitrici epatiche durante lo sviluppo del fegato e nelle cellule staminali del cancro epatico, tuttavia, è raramente presente nelle cellule epatiche adulte. Ad ogni modo, l'uso esclusivo di CD90 come marcatore non è sufficientemente sensibile o specifico per l'isolamento delle cellule staminali epatiche umane (HpSC).

La proteina 1 contenente il dominio CUB (CDCP1) potrebbe essere coinvolta nell'adesione cellulare o nell'interazione con la matrice extracellulare, poiché la sua espressione è stata correlata alle

¹² Cellula non ancora differenziata, pluripotente, danno origine a tutte le cellule del sangue.

¹³ Cellule staminali adulte multipotenti in grado di replicarsi e trasformarsi in altri tipi di cellule

metastasi delle cellule di carcinoma in diversi organi, come il colon, il polmone, la mammella e il pancreas. Inoltre, vari studi hanno suggerito che CDCP1 potrebbe anche essere coinvolta nell'identificazione di fenotipi cellulari specifici, come le cellule staminali ematopoietiche (Conze, Lammers, Kuci, Scherl-Mostageer, & al, 2003), (Takeda, Fujimori, Ogawa, & al, 2010) e mesenchimali, oltre alle cellule staminali neurali (Buhning, Kuci, Conze, Rathke, & al, 2004). Tuttavia, non è ancora stato condotto uno studio specifico per investigare il ruolo di CDCP1 nell'isolamento delle cellule staminali epatiche umane (HpSC) dai fegati fetali primari umani.

Nel precedente studio, è stato prima confermato se le cellule epatiche fetali primarie umane (FLC) mostrassero caratteristiche delle cellule staminali epatiche (HpSC). Attraverso il test di formazione delle colonie clonali, le FLC primarie umane hanno dimostrato la capacità di crescere in colonie cellulari con una densità fino a 100 cellule/cm² su piastre rivestite di collagene per via endovenosa. Inoltre, è stata osservata la presenza di cellule coesprimenti marcatori specifici per gli epatociti e i colangiociti, confermando il loro potenziale di differenziamento. Utilizzando la citometria a flusso, sono stati poi identificati diversi marcatori di superficie cellulare presenti nelle FLC primarie umane, tra cui CD29, CD49f, CD24 e CD44, mentre risultavano negativi per altri marcatori come CD54, CD117, CD138, CD133 e CD140a. In particolare, l'espressione di CD90 è circa del 40,1% delle FLC primarie umane, ed è stato identificato un altro marcatore candidato, CDCP1, espresso nel 18,5% delle FLC primarie umane, suggerendo il suo potenziale come marcatore specifico per l'isolamento di HpSC.

In questa ricerca, è stato confermato che CD90 è espresso nelle cellule staminali/progenitrici epatiche dei fegati adulti o fetali. Per verificare se CDCP1 potesse essere un marcatore specifico per l'isolamento di HpSC, le FLC primarie umane sono state sottoposte a citometria a flusso, utilizzando CDCP1 combinato con CD90 e CD66, producendo quattro sottopopolazioni distinte. Le colture a bassa densità hanno dimostrato che la sottopopolazione CDCP1⁺ CD90⁺ CD66⁻ aveva la maggiore capacità di formare colonie rispetto alle altre. Inoltre, tramite un saggio di selezione cellulare clonale, è stata valutata l'efficienza di clonogenicità, rilevando una maggiore efficienza nelle frazioni CDCP1-positive rispetto a quelle CDCP1-negative. Analizzando il profilo dell'antigene dei marcatori HpSC su CDCP1⁺ CD90⁺ CD66⁻, è stato scoperto che CD13 era espresso in oltre l'80% delle cellule, insieme ad altri marcatori come CD24, CD49f, CD44, CD55 e CD166. Al contrario, altri marcatori di cellule staminali mostravano tassi positivi relativamente più bassi. Inoltre, tramite qPCR e dati di microarray, è stato osservato un arricchimento dei geni bersaglio della via del segnale Wnt/ β -catenina nelle HpSC, coerente con altri studi. È stata rilevata anche un'alta espressione di geni correlati alle cellule staminali, che potrebbero essere utili per tracciare e comprendere le HpSC nel fegato in via di sviluppo.

Per confermare la clonogenicità di una singola cellula staminale putativa, è stata eseguita un'analisi di tracciamento cellulare, suggerendo che le HpSC clonali e proliferative sono state isolate come sottopopolazione CDCP1⁺ CD90⁺ CD66⁻ da FLC primarie umane. Il trapianto di epatociti

emerge come una valida alternativa al trapianto di fegato per le malattie epatiche, date le limitate risorse di organi donatori. Tuttavia, le attuali strategie di conversione non riescono a specificare completamente il destino cellulare. Le cellule staminali epatiche umane (HpSC) potrebbero rappresentare una fonte promettente di epatociti trapiantabili, ma non è chiaro se esistano nel tessuto adulto. Sono state utilizzate varie tecniche per caratterizzare le HpSC presenti nelle cellule epatiche primarie umane, confermando la presenza di una sottopopolazione di cellule CDCP1 + CD90 + CD66 - con capacità autorinnovativa e bipotenziale in vitro, oltre a dimostrare il loro attecchimento e maturazione in vivo dopo il trapianto nel fegato di topo.

è stato osservato che, nonostante l'identificazione di vari marcatori di superficie cellulare, come EpCAM e DLK, CDCP1 emerge come un marcatore specifico e promettente per l'isolamento delle HpSC umane. Le cellule CDCP1 + CD90 + CD66 - hanno dimostrato capacità di autorinnovamento, espansione clonale e differenziazione in epatociti maturi, mantenendo una stabilità nel tempo e mostrando una longevità prolungata nei fegati riceventi dopo il trapianto.

Nel precedente studio si evidenzia la potenziale applicabilità delle HpSC nella terapia cellulare per le malattie del fegato, poiché i fegati umanizzati con HpSC hanno mostrato un'efficienza di ripopolamento elevata, una maturazione completa in epatociti funzionali e una normale capacità metabolica dei farmaci, offrendo così nuove prospettive nel trattamento delle malattie epatiche.

In uno studio (Zhang, Zheng, Li, & al, 2015), è stato impiegato un approccio innovativo utilizzando topi Alb-TRECK/SCID, Per superare i limiti degli attuali modelli murini per lo studio delle cellule epatiche umane. Questo modello murino presenta caratteristiche che lo rendono particolarmente adatto per la differenziazione in vivo di epatociti umani immaturi e la generazione di fegati umanizzati. La creazione di topi chimerici con fegati umanizzati offre un prezioso strumento per lo studio della biologia delle cellule epatiche umane, del metabolismo dei farmaci e per altre applicazioni precliniche. Tuttavia, il successo del trapianto di epatociti umani nei topi richiede un'attenta gestione dell'insufficienza epatica nei topi riceventi, garantendo un microambiente favorevole (nicchia di supporto) per favorire l'espansione e la differenziazione delle cellule umane trapiantate.

Le cellule epatiche fetali primarie umane e le cellule staminali epatiche sono state trapiantate con successo nei topi Alb-TRECK/SCID affetti da insufficienza epatica fulminante letale, indotta dalla tossina difterica (DT). In alcuni topi, il tasso di ripopolamento è stato quasi del 100%. Rispetto alle cellule epatiche embrionali primarie umane, il trapianto di cellule staminali epatiche umane ha dimostrato di salvare i topi Alb-TRECK/SCID affetti da insufficienza epatica fulminante letale.

Inoltre, i fegati umanizzati ottenuti dalle cellule staminali epatiche umane hanno mostrato una secrezione più elevata di albumina umana nei sieri dei topi e hanno dimostrato di funzionare come veri "fegati umani", in grado di metabolizzare i farmaci ketoprofene e debrisoquina. Il lavoro sperimentale riguarda l'utilizzo di topi Alb-TRECK/SCID, questi topi sono stati confermati come omozigoti attraverso incroci ripetuti per tre generazioni. È stato somministrato DT (Diphtheria Toxin) per via intraperitoneale a topi Alb-TRECK/SCID di 8 settimane per valutare il danno epatico tramite l'attività sierica dell'aspartato aminotransferasi (AST).

Le cellule epatiche progenitrici umane sono state ottenute da donatori di età embrionale e coltivate in vitro seguendo procedure standard. Le cellule sono state poi isolate utilizzando specifici marcatori cellulari e trapiantate nei topi Alb-TRECK/SCID.

Sono stati condotti esami biochimici epatici sui topi dopo il trapianto cellulare per valutare l'efficacia del trattamento. Inoltre, è stata somministrata BrdU per valutare la proliferazione cellulare dopo il trattamento. Sono state eseguite analisi istologiche e immunocitochimiche dei tessuti epatici per valutare i cambiamenti a livello cellulare e tissutale dopo il trattamento con le cellule progenitrici epatiche umane.

Gli RNA sono stati estratti da diversi tipi di tessuti epatici umanizzati e da cellule epatiche primarie utilizzando il reagente Isogen. Successivamente, è stato sintetizzato il cDNA utilizzando un kit di trascrizione inversa. La PCR quantitativa è stata eseguita utilizzando sonde specifiche per geni di interesse come ALB, AFP, CYP3A4, CYP2C9, CYP2C19 e hACTB, e i dati sono stati analizzati con un sistema di rilevamento di sequenze. Gli RNA totali sono stati estratti separatamente da FLC primari umani, HpSC e fegati di topo Alb-TRECK/SCID utilizzando un mini kit RNeasy. Questi campioni di RNA sono stati poi ibridati con un pool di "two commercial normal ovary RNA" su un microarray di oligonucleotidi del genoma umano intero. I dati di espressione genica sono stati confrontati tra specie utilizzando gli ID HomoloGene nel set di dati curato da Mouse Genome Informatics dell'ortologia uomo-topo. <http://www.informatics.jax.org>. I risultati dell'analisi dei microarray sono stati depositati nel database Gene Expression Omnibus con l'accesso GSE62933 ([GSE62933](#)).

Nel complesso, è stato dimostrato che i topi Alb-TRECK/SCID costituiscono un modello animale ideale per studiare l'insufficienza epatica fulminante letale e la rigenerazione epatica, oltre allo sviluppo di malattie epatiche. Questo modello può facilitare la valutazione della differenziazione degli epatociti immaturi umani in vivo e il test del metabolismo dei farmaci umani. Dopo il trapianto di epatociti umani immaturi, inclusi FLC primari umani e HpSC, nei topi Alb-TRECK/SCID trattati con DT, i fegati danneggiati sono stati ripopolati

con successo con cellule umane. In particolare, i fegati umanizzati derivati da trapianto di HpSC umani hanno mostrato funzioni epatiche umane più elevate, come la secrezione di albumina umana e una capacità di metabolismo dei farmaci più simile a quella umana. Pertanto, questo modello di fegati umanizzati nei topi Alb-TRECK/SCID offre possibilità di applicazioni funzionali, come studi sul metabolismo dei farmaci, interazioni tra farmaci e trasfezione virale epatica, e supporta ricerche sull'HpSC umano, inclusa la valutazione in vivo dei processi di differenziazione e sviluppo delle cellule staminali epatiche.

3.2 GSE51791

Nel seguente studio (Dianat N. , Steichen, Combettes, Raveux, & al, 2014), è stato investigato il potenziale delle cellule staminali embrionali umane (hESC) di differenziarsi in colangiociti, che sono cellule epiteliali biliari con origine simile agli epatociti durante lo sviluppo embrionale. Un nuovo approccio è stato adottato per guidare la differenziazione delle hESC verso il lignaggio colangiocitico utilizzando condizioni di coltura definite e prive di alimentatore. Dopo una prima fase di differenziazione in progenitori epatici, le hESC sono state ulteriormente differenziate in colangiociti mediante l'utilizzo di specifici fattori di crescita come l'ormone della crescita, il fattore di crescita epidermico e l'interleuchina-6, seguiti dal trattamento con taurocolato di sodio.

Queste condizioni di coltura hanno consentito anche la generazione di colangiociti da epatoblasti derivati da HepaRG. Le cellule simili ai colangiociti, ottenute sia dalle hESC che da HepaRG, hanno espresso marcatori specifici dei colangiociti come citocheratina 7, osteopontina, fattore di trascrizione SOX9 e fattore nucleare epatocitario 6. Inoltre, sono state identificate proteine importanti per le funzioni dei colangiociti, tra cui il regolatore della conduttanza transmembrana della fibrosi cistica, il recettore della secretina e recettori nucleari. Queste cellule hanno dimostrato di formare ciglia primarie e di rispondere alla stimolazione ormonale mediante l'aumento del Ca^{2+} intracellulare.

Attraverso un'analisi genomica integrativa, è stato confermato che le cellule simili ai colangiociti differiscono dalla linea epatocitaria a livello genico. Crescendo in una matrice tridimensionale, i colangiociti sviluppano una polarità epiteliale/apicobasale e formavano cisti funzionali e dotti biliari. Inoltre, è stato dimostrato che cellule simili ai colangiociti potevano essere generate anche da cellule staminali pluripotenti indotte dall'uomo, sottolineando l'efficacia e la flessibilità di questo approccio con cellule staminali/progenitrici di diverse origini.

In conclusione, è stato sviluppato un metodo robusto ed efficiente per differenziare le cellule staminali pluripotenti in cellule simili ai colangiociti, che presentano caratteristiche strutturali e funzionali simili alle cellule del dotto biliare nel fegato normale. Queste cellule rappresentano una risorsa preziosa per lo studio in vitro dei meccanismi molecolari dello sviluppo del dotto biliare e potrebbero avere un ruolo significativo in strategie terapeutiche, inclusi approcci bioingegnerizzati per il trattamento di patologie epatiche.

Nelle fasi iniziali dell'organogenesi epatica, gli epatoblasti, precursori delle cellule epatiche, possono differenziarsi in epatociti e colangiociti, le cellule epiteliali biliari che rivestono i dotti intra ed extraepatici dell'albero biliare. Intorno all'ottava settimana di gestazione umana, gli epatoblasti formano una placca duttale, che è considerata il compartimento delle cellule staminali epatiche fetali e neonatali. Si ritiene che la via NOTCH sia coinvolta nell'impegno biliare e nella formazione della placca duttale, con l'interazione tra cellule mesenchimali JAGGED1-positive e epatoblasti NOTCH2-positivi. Questa interazione induce l'espressione di SOX9, marcatore precoce delle cellule biliari che controlla la morfogenesi del dotto biliare (Antoniou, Raynaud, Cordi, Zong, & al, 2009). La proteina Sal-like 4 (SALL4) è anche coinvolta nel controllo del destino delle cellule epatiche, inibendo la differenziazione degli epatoblasti in epatociti e favorendo la loro differenziazione in colangiociti.

La funzione principale dei colangiociti è quella di regolare attivamente la composizione della bile, modificando i componenti biliari derivati dagli epatociti attraverso eventi secretori e di riassorbimento. Oltre a ciò, essi partecipano alla disintossicazione da xenobiotici. I colangiociti rappresentano bersagli primari di danno in diverse malattie epatiche colestatiche, che spaziano da disturbi ereditari come la sindrome di Alagille e la fibrosi cistica, alla colangite autoimmune e alla cirrosi biliare primitiva, che sono le principali indicazioni per il trapianto di fegato pediatrico. Tuttavia, il loro numero limitato (3% della massa epatica totale) e la loro localizzazione intraepatica hanno reso difficile lo sviluppo di modelli cellulari in vitro per approfondire la loro funzione. Attualmente, ci sono poche linee cellulari umane di colangiociti disponibili, derivanti principalmente da colangiocarcinoma o da cellule normali immortalate da SV40 (Humphreys, Williams, Adams, & Afford, 2010), e la maggior parte delle cellule studiate in vitro proviene da colangiociti di ratti (Vroman & LaRusso, 1996). Tuttavia, esistono differenze significative tra i modelli di roditori e quelli umani, con i colangiociti umani, soprattutto quelli più piccoli, che esprimono il regolatore della conduttanza transmembrana della fibrosi cistica (CFTR) (Alpini, Ulrich, Roberts, & Ueno, 1997), mentre quelli dei roditori non lo fanno.

Le cellule staminali pluripotenti, capaci di differenziarsi in una vasta gamma di tipi cellulari e di proliferare indefinitamente, rappresentano una fonte promettente di cellule differenziate con potenziale (Dianat N. , Steichen, Vallier, Weber., & al, 2013). Diversi studi si sono concentrati sullo sviluppo di protocolli efficaci per generare cellule simili agli epatociti da queste cellule, incluso il nostro lavoro che ha generato con successo cellule simili agli epatociti da cellule staminali embrionali umane (hESC) e cellule staminali pluripotenti indotte dall'uomo (hiPSC). L'approccio degli studiosi ha seguito le fasi dello sviluppo epatico, permettendo la differenziazione graduale delle hESC e delle hiPSC in cellule endodermiche definitive, progenitori epatici e cellule simili agli epatociti in un ambiente definito privo di cellule alimentatrici o siero (Dianat N. , Steichen, Vallier, Weber., & al, 2013). Tuttavia, è ancora oggetto di dibattito se i progenitori epatici generati attraverso diversi protocolli, specialmente nelle nostre condizioni definite, siano capaci di differenziarsi in colangiociti.

Nel principale studio, è stato esaminato il potenziale delle hESC, delle hiPSC e dei progenitori epatici derivati da HepaRG (epatoblasti) di differenziarsi ulteriormente lungo la linea colangiocitica. Sono state sviluppate per la prima volta condizioni robuste per guidare la differenziazione di questi diversi tipi di progenitori in cellule funzionali simili ai colangiociti. Hanno scoperto che queste cellule mostravano una varietà di funzionalità specifiche, inclusa la capacità di trasportare gli acidi biliari, la formazione di ciglia e la risposta alla stimolazione della via di segnalazione del calcio. Inoltre, è stato osservato che, in condizioni di maturazione 3D, i colangiociti formavano strutture tubulari con polarità apicobasale e cisti.

Questo studio rappresenta la prima dimostrazione della differenziazione di cellule funzionalmente simili ai colangiociti da progenitori epatici generati da cellule staminali pluripotenti umane e cellule HepaRG. L'approccio adottato si basa su condizioni di coltura completamente definite, prive di siero e di cellule alimentatrici.

Durante la maturazione in vivo, gli epatoblasti si differenziano in colangiociti sotto l'influenza di diversi fattori secreti presenti nel microambiente intorno alla vena porta. È stata osservata l'espressione di vari fattori di trascrizione nei nostri progenitori epatici in fase di differenziazione, tra cui HNF3 β ¹⁴ HNF6 e HNF1 β , importanti per lo sviluppo dei dotti biliari. L'analisi del trascrittoma ha indicato l'attivazione di vie di segnalazione coinvolte nella differenziazione biliare.

¹⁴ fattore nucleare epatocitario 3 β

Nonostante l'assenza di una via di segnale specifica per la differenziazione dei colangiociti, l'aggiunta di GH¹⁵ sugli epatoblasti seguita da IL-6 ha dimostrato di essere efficace nell'indurre la differenziazione verso la linea dei colangiociti e la loro proliferazione.

Le analisi genomiche hanno permesso di identificare un pannello di geni e percorsi caratteristici dei colangiociti, confermando l'impegno dei progenitori in questa linea cellulare. I colangiociti differenziati esprimevano marcatori tipici dei dotti biliari e mostravano una varietà di funzionalità, compresa la capacità di trasportare gli acidi biliari, rispondere a segnalazioni cellulari e formare strutture tubulari.

Complessivamente, i colangiociti generati con questo approccio potrebbero fornire un modello utile per lo studio dei meccanismi molecolari e patogenetici dello sviluppo del dotto biliare, così come delle malattie epatiche associate. Questo potrebbe aprire la strada allo sviluppo di nuove strategie terapeutiche, comprese quelle basate su fegati bioingegnerizzati.

3.3 GSE199820

Tra le centinaia di cellule umane, le cellule staminali pluripotenti indotte ed embrionali (iPSC/ESC) hanno meccanismi post-trascrizionali unici (Chen & Hu, 2017), come specifici eventi di elaborazione dell'mRNA, che stabiliscono e mantengono lo stato pluripotente (Iwasaki, Kawahara, Okubo, Yamakawa, & al, 2022). È notevole osservare che, per diversi geni, le discrepanze nei livelli proteici tra le cellule staminali embrionali (ESC) non sono accompagnate da variazioni corrispondenti nei livelli di mRNA (Lu, Markowitz, Unwin, Leek, & al, 2009), implicando una stretta regolazione post-trascrizionale a livello proteico. Ad esempio, il gene CLOCK mostra una stabilità nei livelli di mRNA prima e dopo la differenziazione, mentre la sua espressione proteica varia a seconda del tipo cellulare ed è assente nelle cellule staminali pluripotenti indotte (iPSC/ESC) (Umemura, Koike, Ohashi, Meng, & al, 2017). Considerando che nei primati i livelli proteici sono più conservativi rispetto a quelli di mRNA (Khan, Ford, Cusanovich, & al, 2013), questa stretta regolazione proteica con livelli costanti di mRNA potrebbe svolgere un ruolo significativo nella pluripotenza e nella differenziazione durante lo sviluppo.

Per identificare in modo sistematico i geni regolati diversamente a livello di mRNA e proteine, sono state condotte analisi del trascrittoma mediante sequenziamento dell'RNA o microarray e analisi del proteoma mediante spettrometria di massa (MS) su diversi tipi di cellule. Tipicamente, si osserva una forte correlazione nei livelli di geni correlati alle vie

¹⁵ ormone della crescita

metaboliche principali, ma una correlazione più bassa per geni ribosomiali e spliceosomiali. Tuttavia, resta aperta la questione se i geni regolati post-trascrizionalmente con livelli proteici strettamente controllati siano essenziali per il mantenimento della cellula.

In questo studio, sono stati confrontati i profili globali di mRNA e proteine tra le iPSC/ESC e le cellule differenziate per identificare geni essenziali regolati a livello post-trascrizionale. Sono stati individuati 228 geni regolati post-trascrizionalmente esclusivamente nelle iPSC/ESC, con funzioni biologiche specifiche legate all'RNA e agli acidi nucleici. Attraverso lo screening dei siRNA, è emerso che il 54% (84 geni) di questi geni era necessario per la sopravvivenza cellulare, con 20 di essi identificati come specifici per le iPSC. Inoltre, è stato evidenziato che la maggior parte dei 20 geni è regolata principalmente nella fase di localizzazione dell'mRNA, tre geni sono sottoregolati nella fase di degradazione proteica nei fibroblasti dermici umani (HDF) e sette geni sono sovraregolati nella fase di traduzione nelle iPSC. Complessivamente, questo studio suggerisce l'importanza della regolazione post-trascrizionale specifica del tipo cellulare nelle iPSC.

Nonostante gli effetti onnipresenti, la comprensione dei meccanismi di regolazione post-trascrizionale risulta molto meno approfondita rispetto a quelli di regolazione trascrizionale nelle iPSC/ESC. È emerso che 20 geni regolati a livello post-trascrizionale (noti come geni uPRA) sono cruciali per il mantenimento delle iPSC. Questi geni mostrano un'elevata espressione proteica senza una significativa sovraregolazione di espressione di mRNA nelle iPSC. Le proteine HSPA8, EIF3D e NCBP2 subiscono rapida degradazione nello stadio di degradazione del proteasoma nelle HDF. Diciotto mRNA iPSC/ESC-uPRA vengono trasportati dal nucleo al citoplasma in quantità maggiore nelle iPSC/ESC rispetto alle HDF, mentre si osserva una significativa localizzazione dell'RNA nucleare a livello globale nelle HDF. L'efficienza di traduzione di POLD2, CPSF6, RBM22, BUD31 e IMP4 risulta sovraregolata nei monosomi, mentre la traduzione di RSL1D1 e NIFK è sovraregolata nei polisomi nelle iPSC. Tuttavia, la maggior parte dei geni uPRA è arricchita nella frazione polisomica leggera nelle HDF, suggerendo differenti effetti inibitori dell'mRNA di uPRA sulla traduzione dopo il legame con i ribosomi nelle cellule differenziate. Per identificare tale inibizione, è stato scelto un gene uPRA, SRRT, per ulteriori analisi. La percentuale di mRNA citoplasmatico di SRRT risulta essere 3 volte inferiore nelle HDF, e l'mRNA stesso mostra una maggiore associazione con monosomi, polisomi leggeri e pesanti nelle HDF rispetto alle iPSC. L'mRNA esogeno citosolico di SRRT viene tradotto nelle HDF, suggerendo che il livello di mRNA citoplasmatico potrebbe influenzare l'espressione proteica corretta. Un altro possibile meccanismo di inibizione della traduzione nelle cellule

differenziate è la pausa ribosomiale, in cui l'mRNA si lega a ribosomi inattivi o seleziona codoni di inizio non canonici (Tresenrider, Morse, Jorgensen, Chia, & al, 2021). Una terza possibilità è che le proteine aberranti vengano degradate immediatamente dopo la loro produzione attraverso un percorso di controllo della qualità dell'mRNA e del ribosoma (D'orazio & Green, 2021). Ulteriori studi sono necessari per chiarire quale meccanismo si applichi a ciascun gene.

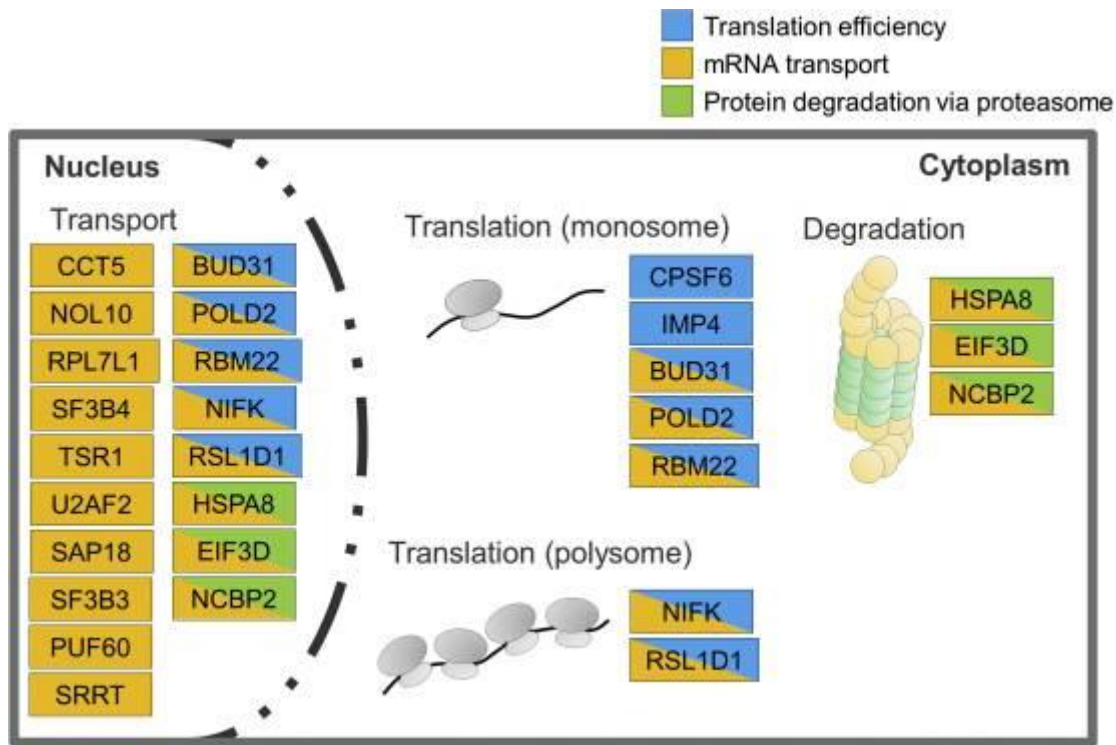


Figura 9: : illustrazione di come 20 geni essenziali iPSC/ESC-uPRA sono regolati nelle iPSC e nelle cellule differenziate¹⁶

Tutte le 20 proteine iPSC/ESC-uPRA identificate nello studio (Iwasaki, Kawahara, Okubo, Yamakawa, & al, 2022) sono coinvolte in un'ampia gamma di funzioni cellulari, dalla trascrizione alla post-traduzione. Per le funzioni legate alla trascrizione, 10 geni iPSC/ESC-uPRA hanno ruoli nella polimerasi del DNA, nello splicing e nella maturazione dell'mRNA: POLD2 fa parte del complesso della DNA polimerasi; SF3B3, SF3B4 e RBM22 sono componenti del fattore di splicing; e PUF60, BUD31, SAP18, CPSF6, U2AF2 e SRRT sono necessari per lo splicing del pre-mRNA, con particolare rilevanza di SRRT come "guardiano molecolare dello stato pluripotente cellulare" che facilita lo splicing corretto nelle iPSC (Kaivon & Makeyven, 2020). Per le funzioni legate alla traduzione, otto geni iPSC/ESC-uPRA sono coinvolti nelle funzioni del complesso ribosomiale: NCBP2 ed EIF3D legano il cap dell'mRNA; EIF3D è necessario per i fattori specializzati di avvio della traduzione;

¹⁶<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9097716/>

RPL7L1 è un presunto componente del complesso ribosomiale; mentre RSL1D1, NIFK, NOL10, IMP4 e TSR1 sono necessari per l'elaborazione dell'rRNA. Per le funzioni post-traduzionali, CCT5 e HSPA8 sono componenti del complesso chaperone, che assiste le proteine appena sintetizzate nel corretto ripiegamento.

L'osservazione che questi geni sono regolati a livello post-trascrizionale è particolarmente intrigante, poiché tutti agiscono come regolatori post-trascrizionali. È stato scoperto che i tipi di motivi dei geni uPRA sono piuttosto diversi attraverso l'analisi in silico, e diverse combinazioni di proteine che legano l'RNA potrebbero essere coinvolte nel mantenimento delle iPSC. In conclusione, lo studio analizzato ha rivelato che le espressioni proteiche di 228 geni sono regolate a livello post-trascrizionale nelle iPSC. Inoltre, i livelli proteici di 20 di questi geni sono specificamente aumentati nelle iPSC e le proteine stesse sono essenziali per il mantenimento delle iPSC. Infine, sebbene i meccanismi di repressione della traduzione dei geni uPRA nelle HDF richiedano ulteriori indagini, è stato dimostrato che diversi meccanismi regolano le espressioni geniche post-trascrizionali tra i tipi di cellule.

DATA SET	FACTOR OF INTEREST	EXPERIMENTAL GROUP	CONTROL GROUP
GSE199820	Fraction	Cytosol	Nucleus

CATEGORY	TERM	P-VALUE	BENJAMINI
INTERPRO	Nucleic acid-binding, OB-fold	3,8E-2	1,0E0
INTERPRO	Ubiquitin conserved site	4,1E-2	1,0E0

UP_SEQ_FEATURE	TOPO_DOM: Mitochondrial matrix	4,4E-2	1,0E0
GOTERM_BP_DIRECT	Regulation of translation	4,7E-2	1,0E0
INTERPRO	Ubiquitin subgroup	4,9E-2	1,0E0
UP_SEQ_FEATURE	DOMAIN:CCDC144C -like coiled-coil	5,1E-2	1,0E0
GOTERM_BP_DIRECT	Protein import into mitochondrial outer membrane	5,2E-2	1,0E0
GOTERM_BP_DIRECT	Response to virus	6,6E-2	1,0E0
GOTERM_BP_DIRECT	Maturation of SSU- Rrna	7,0E-2	1,0E0
GOTERM_BP_DIRECT	Translational elongation	7,3E-2	1,0E0
GOTERM_BP_DIRECT	<i>Stem cell proliferation</i>	1,5E-3	1,2E-1
UP_KM_BIOLOGICAL_PROCES S	<i>Differentiation</i>	3,8E-3	5,0E-2

3.4 GSE43143

La tecnologia delle cellule staminali pluripotenti indotte (iPSC) offre una preziosa opportunità per studiare i disturbi neuropsichiatrici, consentendo la crescita di neuroni specifici del paziente in vitro. Tradizionalmente, i fibroblasti cutanei da biopsia sono stati la principale fonte di cellule per la riprogrammazione. Tuttavia, l'impiego di altre cellule somatiche ottenute in modo meno invasivo sarebbe preferibile, soprattutto nei bambini con disturbi dello spettro autistico (ASD) e altre condizioni neurologiche.

Oltre ai fibroblasti, le iPSC sono state ottenute dal sangue cordonale, dai linfociti, dai cheratinociti¹⁷ dei capelli e dalla polpa dentale (Chen, Foxe, Pedrosa, & al, 2013). La polpa dentale, in particolare, si presenta come una buona fonte per i disturbi dello sviluppo neurologico nei bambini, in quanto il prelievo del materiale non è invasivo. Al fine di valutarne l'idoneità per la modellazione della malattia, è stata condotta un'analisi dell'espressione genica mediante RNA-seq su neuroni differenziati derivati da iPSC costituite da polpa dentale estratta da denti decidui (T-iPSC) e da fibroblasti (F-iPSC). Questa è stata la prima analisi RNA-seq che ha confrontato i profili di espressione genica nei neuroni derivati da iPSC costituite da diverse cellule somatiche. Globalmente, i profili di espressione genica erano piuttosto simili, con solo 329 geni che mostravano espressione differenziale significativa, di cui 63 rimanevano significativi dopo la correzione per l'analisi dell'intero genoma.

La principale differenza osservata è stata il livello inferiore di espressione di numerosi membri delle quattro famiglie di geni HOX nei neuroni derivanti da T-iPSC. Inoltre, è stato osservato un aumento dell'espressione di diversi fattori di trascrizione espressi nel prosencefalo in via di sviluppo, come FOXP2, OTX1 e LHX2.

L'analisi dei percorsi ha rivelato che i geni espressi in modo differenziale, mostrando livelli di espressione più elevati nei neuroni derivati da T-iPSC, erano arricchiti per i geni implicati nella schizofrenia (SZ). Questi risultati suggeriscono che i neuroni derivati da T-iPSC sono idonei per modellare la malattia del disturbo neuropsichiatrico e possono offrire vantaggi rispetto a quelli derivati da F-iPSC.

Sono state sfruttate le cellule staminali pluripotenti indotte (iPSC) per modellare malattie in vitro, concentrandoci su una vasta gamma di disturbi neuropsichiatrici, tra cui la schizofrenia (SZ) e i disturbi dello spettro autistico (ASD) (Lin, Pedrosa, Wang, Zheng, & al, 2012). Finora, i fibroblasti ottenuti da biopsie cutanee sono stati la fonte predominante di cellule

¹⁷ Cellule dell'epidermide che sintetizzano la cheratina

per la riprogrammazione delle iPSC, ma questo può costituire un ostacolo per modellare i disturbi neurologici infantili a causa dell'invasività della procedura. Inoltre, le iPSC derivate da bambini con ASD e altri disturbi sono state ottenute utilizzando fibroblasti o, più recentemente, sangue periferico, ma questo approccio può essere problematico a causa dell'avversione dei bambini alle procedure mediche. Di conseguenza, ci sarebbe bisogno di fonti alternative di cellule per la riprogrammazione delle iPSC ottenute in modo non invasivo.

La polpa dentale derivata dai denti decidui emerge come una potenziale fonte di cellule somatiche per la riprogrammazione delle iPSC. Le iPSC derivate dalla polpa dentale (T-iPSC) mostrano caratteristiche di pluripotenza e sono capaci di differenziarsi in neuroni. Tuttavia, è necessario un approfondito profilo molecolare per valutare la somiglianza dei neuroni derivati da T-iPSC con quelli derivati da altre fonti di cellule somatiche, come i fibroblasti.

Pertanto, è stata condotta un'ampia analisi del profilo di espressione genica dei neuroni derivati da T-iPSC utilizzando il profilo dell'intero trascrittoma (RNA-Seq) e confrontandolo con iPSC derivati da neuroni ottenuti da fibroblasti (F-iPSC). I risultati mostrano un alto grado di correlazione tra le due fonti di neuroni, ma anche sottili differenze che possono influenzare la scelta di utilizzare T-iPSC o F-iPSC per modellare specifici disturbi neuropsichiatrici.

Durante lo studio condotto su iPSC TIPS4 e TIPS4-C5, NPC e neuroni a 14 giorni dopo la differenziazione neuronale, è stata effettuata un'analisi del trascrittoma dell'intero genoma (RNA-Seq). Durante la transizione da iPSC a NPC, sono stati identificati 1368 geni con espressione differenziale (627 più alti negli NPC; 741 inferiori negli NPC, FDR <0,05). Un confronto tra iPSC e neuroni ha rivelato 2543 geni con espressione differenziale (1286 aumentati nei neuroni; 1257 inferiori nei neuroni, FDR <0,05).

Tra i geni che hanno mostrato diminuzioni significative nell'espressione negli NPC e nei neuroni vi erano POU5F1 (OCT4), LIN28A e TDGF1, noti per essere espressi ad alti livelli nelle cellule staminali pluripotenti e inattivati durante la differenziazione. Inoltre, diversi RNA lunghi non codificanti (lncRNA), tra cui ESRG, RP11-1144P22.1, LINC00678 e RP11-25619.2, sono emersi come sottoregolati durante la transizione da iPSC a NPC, suggerendo un ruolo nel mantenimento della pluripotenza.

Durante la differenziazione in NPC e neuroni, si è osservato un aumento dell'espressione di membri della famiglia dei geni HOX e di geni codificanti fattori di trascrizione neuronali

come POU3F3, MYT1L e DLX1. In particolare, MYT1L, insieme a POU3F2 e ASCL1, è noto per la capacità di riprogrammare i fibroblasti direttamente nei neuroni.

Infine, durante la transizione dagli NPC ai neuroni, sono emersi 289 geni con espressione differenziale, tra cui diversi lncRNA che mostravano un aumento significativo dell'espressione nei neuroni rispetto agli NPC, suggerendo un coinvolgimento nella maturazione neuronale.

L'analisi del percorso ha evidenziato un arricchimento per i geni coinvolti nella schizofrenia tra quelli che aumentavano nei neuroni, mentre i geni coinvolti nella regolazione del ciclo cellulare erano diminuiti. Questi risultati indicano che le iPSC derivate dalla polpa dentale possono essere utilizzate efficacemente per modellare i disturbi neuropsichiatrici.

Nel campo della modellizzazione delle malattie utilizzando iPSC, è essenziale sfruttare fonti di cellule somatiche facilmente accessibili per la riprogrammazione, come fibroblasti cutanei, cheratinociti dei capelli, leucociti CD34+, cellule epiteliali delle urine e polpa dentale. Tuttavia, mentre la biopsia cutanea può essere invasiva e causare disagio, la polpa dentale derivante dai denti decidui caduti naturalmente durante l'infanzia rappresenta una fonte di cellule per la riprogrammazione non invasiva e priva di ulteriori stress per i bambini. Sotto il profilo biologico, la polpa dentale potrebbe emergere come una fonte privilegiata di iPSC per modellare i disturbi neuropsichiatrici, considerando le sue origini durante lo sviluppo embrionale dei vertebrati. Le cellule della polpa dentale derivano dall'ectoderma, in particolare dalle cellule della cresta neurale, mentre i fibroblasti provengono dall'ectoderma programmato per diventare epidermide. Questa differenza di origine potrebbe influenzare l'espressione genica e, di conseguenza, la capacità delle iPSC di modellare la malattia in vitro.

In uno studio di confronto tra neuroni derivati da iPSC di polpa dentale (T-iPSC) e fibroblasti (F-iPSC), sono emerse differenze significative nell'espressione di geni chiave. Nei neuroni T-iPSC, sono stati osservati livelli più bassi di espressione per geni coinvolti nello sviluppo del rombencefalo, mentre diversi fattori di trascrizione associati allo sviluppo del prosencefalo erano aumentati. Questi risultati suggeriscono che i neuroni derivati da T-iPSC potrebbero offrire vantaggi in condizioni come la schizofrenia e i disturbi dello spettro autistico, associati a deterioramento cognitivo e linguistico.

Tuttavia, resta da chiarire se queste differenze nell'espressione genica persistano utilizzando altri metodi di differenziazione neuronale. In definitiva, l'analisi del percorso dei geni espressi in modo differenziale supporta l'idea che i neuroni derivati dalle T-iPSC possano essere più adatti per modellare specifici disturbi neuropsichiatrici.

DATA SET	FACTOR OF INTEREST	EXPERIMENTAL GROUP	CONTROL GROUP
GSE43143	Cell type	IPSC	NPC

CATEGORY	TERM	P-VALUE	BENJAMINI
GOTERM_BP_DIRECT	Positive regulation of tumor necrosis factor production	8,3E-2	1,0E0
GOTERM_BP_DIRECT	Central nervous system neuron development	8,5E-2	1,0E0
KEGG_PATHWAY	Leukocyte transendothelial migration	8,6E-2	1,0E0
GOTERM_BP_DIRECT	Calcium-independent cell-cell adhesion via plasma membrane cell-adhesion molecules	8,9E-2	1,0E0
GOTERM_BP_DIRECT	Female gonad development	9,3E-2	1,0E0
GOTERM_BP_DIRECT	Cell fate determination	9,3E-2	1,0E0

GOTERM_BP_DIRECT	Peptidyl-tyrosine autophosphorylation	9,3E-2	1,0E0
GOTERM_BP_DIRECT	Regulation of DNA methylation	9,3E-2	1,0E0
GOTERM_BP_DIRECT	Negative regulation of axon extension	9,7E-2	1,0E0
GOTERM_BP_DIRECT	Negative regulation of actin filament polymerization	9,7E-2	1,0E0
GOTERM_MF_DIRECT	<i>RNA polymeraseII transcription factor activity, sequence specific DNA- Binding</i>	3,3E-10	6,8E-8
GOTERM_BP_DIRECT	<i>Positive regulation of cell poliferation</i>	4,8E-6	9,7E-4
GOTERM_BP_DIRECT	<i>Positive regultion of stem cell proliferation</i>	2,6E-5	4,2E-3

3.5 GSE152983

Di recente è stato descritto un aumento aberrante dei trascritti citoplasmatici SFPQ che trattengono gli introni (IRT) e la concomitante errata localizzazione della proteina SFPQ come nuovi segni distintivi della sclerosi laterale amiotrofica (SLA). Tuttavia, la

generalizzabilità e i potenziali ruoli degli IRT citoplasmatici nella salute e nella malattia rimangono poco chiari. In questo studio (Tyzack, Neeves, Crerar, Klein, & al, 2021), utilizzando il sequenziamento profondo risolto nel tempo di frazioni nucleari e citoplasmatiche di cellule staminali pluripotenti indotte umane sottoposte a neurogenesi motoria, è stato rivelato che le mutazioni del gene VCP che causano la SLA portano ad un accumulo aberrante di IRT specifico per compartimento. Nello specifico, sono stati identificati più di 100 IRT con maggiore abbondanza citoplasmatica nei campioni di SLA. Inoltre, questi IRT citoplasmatici aberranti possiedono attributi sequenza-specifici e un'affinità di legame prevista differenziale con le proteine leganti l'RNA. Sorprendentemente, TDP-43, SFPQ e FUS - proteine leganti l'RNA note per l'errata localizzazione nucleare-citoplasmatica nella SLA - si legano in modo abbondante e specifico a questo pool citoplasmatico aberrante di IRT. I dati ottenuti sono quindi coerenti con un nuovo ruolo degli IRT citoplasmatici nella regolazione dell'abbondanza proteica specifica del compartimento. Questo studio fornisce nuove informazioni molecolari sui potenziali meccanismi patogenetici alla base della SLA e evidenzia gli IRT citoplasmatici aberranti come potenziali bersagli terapeutici.

Gli studi recenti hanno evidenziato che la ritenzione degli introni è più comune nei mammiferi di quanto inizialmente pensato, influenzando le trascrizioni della maggior parte dei geni. È stato osservato che le cellule neurali, con la loro particolare polarità e compartimentazione, mostrano livelli più elevati di ritenzione degli introni rispetto ad altri tipi cellulari. Questo fenomeno è particolarmente significativo durante la differenziazione neuronale precoce e svolge un ruolo funzionale nell'omeostasi neuronale. La ritenzione degli introni è stata precedentemente associata alla regolazione del trascrittoma attraverso l'interazione con le vie di degradazione dell'RNA. Anche se si pensava che i trascritti che trattengono gli introni risiedessero principalmente nel nucleo dove venivano degradati, crescono le prove della loro presenza nel citoplasma. Tuttavia, la loro frequenza e il loro ruolo rimangono poco chiari. Uno studio precedente ha evidenziato che un introne trattenuto nella trascrizione *Calm3* determina la sua localizzazione dendritica, suggerendo una funzione di indirizzamento per la ritenzione degli introni citoplasmatici. Questo suggerisce che le sequenze di RNA introniche potrebbero avere ruoli oltre la funzione nucleare e che i programmi di ritenzione degli introni citoplasmatici potrebbero essere rilevanti per la funzione neurologica umana e la patologia associata, come la sclerosi laterale amiotrofica (SLA).

La SLA è una malattia neurodegenerativa progressiva e incurabile che colpisce i motoneuroni. Il suo segno molecolare distintivo è l'errata localizzazione delle proteine chiave leganti l'RNA dal nucleo al citoplasma, anche se il meccanismo di questo fenomeno rimane poco chiaro. Le mutazioni genetiche coinvolgono regolatori cruciali dell'elaborazione dell'RNA, suggerendo che cambiamenti post-trascrizionali, inclusi quelli durante lo sviluppo neurologico, possano contribuire alla patogenesi della SLA. Le ricerche hanno confermato che la ritenzione degli introni è più comune nei mammiferi di quanto inizialmente stimato, esercitando un'influenza sulle trascrizioni di molti geni (Yap, Lim, Khandelia, Friedman, & al, 2012). È degno di nota che le cellule neurali, con la loro peculiare polarità e compartimentazione, presentano livelli più elevati di ritenzione intronica rispetto ad altri tipi cellulari (Mauger, Lemoine, & Scheiffele, 2016). Questo fenomeno riveste un ruolo significativo nello splicing durante le prime fasi della differenziazione neuronale e contribuisce all'omeostasi neuronale. Inoltre, la ritenzione degli introni è stata associata alla regolazione del trascrittoma attraverso l'interazione con le vie di degradazione dell'RNA. Sebbene sia noto che i trascritti che trattengono gli introni risiedono principalmente nel nucleo dove vengono degradati, crescono le prove che indicano la loro presenza nel citoplasma. Tuttavia, la frequenza e il ruolo degli introni citoplasmatici rimangono poco studiati. Uno degli studi focalizzati su un introne trattenuto nella trascrizione *Calm3* ha dimostrato il suo ruolo nell'indirizzare la localizzazione (Sharangdhar, Sugimoto, Heraud-Farlow, & al, 2017), suggerendo una funzione di "codificazione postale" per la ritenzione intronica citoplasmatica. Questa scoperta apre la strada a nuove interpretazioni delle sequenze di RNA introniche al di là del loro ruolo nucleare, suggerendo che i programmi di ritenzione intronica citoplasmatica siano rilevanti per la funzione neurologica umana e, pertanto, per lo sviluppo di malattie neurologiche.

La sclerosi laterale amiotrofica (SLA) rappresenta una condizione a esordio in età adulta caratterizzata da una progressione rapida e inarrestabile, che porta alla degenerazione relativamente selettiva dei motoneuroni. Il segno patologico molecolare distintivo della SLA è l'errata localizzazione delle proteine chiave leganti l'RNA (RBP) (Luisier, Tyzack, Hall, & al, 2018) dal nucleo al citoplasma, sebbene il meccanismo sottostante tale fenomeno rimanga poco chiaro. Le mutazioni genetiche coinvolte nella SLA riguardano regolatori cruciali dell'elaborazione dell'RNA, normalmente espressi durante lo sviluppo, suggerendo un possibile ruolo dei cambiamenti post-trascrizionali, inclusi quelli verificatisi durante lo sviluppo neurologico, nella patogenesi molecolare della malattia.

Recenti studi hanno individuato la ritenzione degli introni come l'evento di splicing predominante che caratterizza le prime fasi della restrizione del lignaggio dei motoneuroni da parte delle cellule staminali pluripotenti indotte dall'uomo (hiPSC), con variazioni causate da diverse mutazioni genetiche associate alla SLA. Tuttavia, rimane incerto se questo processo influenzi principalmente i compartimenti subcellulari nucleari o citoplasmatici. Pochi studi hanno esaminato specificamente la ritenzione intronica nei compartimenti cellulari differenziati (Price, Hwang, Tao, & al, 2020), e, fino ad oggi, nessuno ha analizzato in modo completo i programmi di ritenzione intronica citoplasmatica durante la neurogenesi motoria umana o caratterizzato sistematicamente gli effetti di mutazioni causali della SLA. In questa ricerca (Tyack, Neeves, Crerar, Klein, & al, 2021), è stato combinato il frazionamento cellulare delle cellule staminali pluripotenti indotte dall'uomo (hiPSC) durante la neurogenesi motoria con il sequenziamento profondo dell'RNA (RNA-seq), generando circa 100 milioni di letture finali accoppiate per campione. Questo approccio mirava a ottenere informazioni sulla "logica" molecolare che regola i programmi di ritenzione intronica in condizioni sia sane che patologiche. Tale metodologia rappresenta una risorsa preziosa per i ricercatori di tutte le discipline delle neuroscienze, poiché comprende sei punti temporali durante la neurogenesi motoria per quattro linee cellulari di controllo (derivate da quattro individui sani) e tre linee cellulari di SLA (derivate da due pazienti portatori di mutazioni nel gene VCP), con campionamento sia nucleare che citoplasmatico.

Questo approccio ha consentito di ottenere importanti informazioni sulla natura della ritenzione intronica aberrante in un modello di cellule staminali umane affette da SLA. In particolare, è stata fornita una classificazione per la ritenzione intronica aberrante basata sulla distribuzione nucleocitoplasmatica, sugli attributi cis e sulle affinità di legame degli introni con le principali proteine leganti l'RNA (RBP). In modo sorprendente, sono stati identificati più di 100 trascritti che trattengono gli introni (IRT) nel citoplasma delle colture di SLA, indicando che questo fenomeno è più diffuso di quanto precedentemente riconosciuto.

Inoltre, è stato confermato il legame diretto del fattore di splicing della prolina e della proteina ricca di glutammina (SFPQ) con gli IRT di SFPQ all'interno del citoplasma del modello di SLA hiPSC, con prove preliminari di un'interazione funzionale tra di loro. In sintesi, questa ricerca ha identificato una nuova classe di IRT citoplasmatici che possono predire l'errata localizzazione da nucleare a citoplasmatica delle chiave proteiche leganti l'RNA, il che rappresenta un segno molecolare distintivo riconosciuto della SLA.

Nel presente studio, gli autori hanno cercato di determinare se questa ritenzione intronica anomala, localizzata nel citoplasma, fosse un fenomeno diffuso nella SLA. Per farlo, hanno combinato la differenziazione diretta di specifiche cellule staminali pluripotenti indotte (hiPSC) in motoneuroni spinali con il frazionamento cellulare e il sequenziamento profondo dell'RNA [poli(A)]. Hanno dimostrato che la ritenzione anomala degli introni nel citoplasma è diffusa nella SLA, coinvolgendo almeno 112 trascrizioni, tra cui SFPQ. Inoltre, hanno osservato un aumento dell'abbondanza citoplasmatica delle sequenze introniche, suggerendo un'anomala esportazione nucleare e/o stabilizzazione citoplasmatica di un pool specifico di trascritti che trattengono gli introni nelle colture mutanti VCP. È interessante notare che un aumento simile di eventi di ritenzione intronica anomala è stato precedentemente osservato in set di dati RNA-Seq di motoneuroni portatori di mutazioni nei geni SOD1 o FUS.

Successivamente, gli autori hanno classificato gli eventi di ritenzione intronica anomala in tre categorie in base alla loro distribuzione nucleocitoplasmatica e hanno focalizzato la loro attenzione sul giorno 14 del protocollo di neurogenesi motoria, dove è stata osservata la maggior parte degli eventi anomali di ritenzione intronica. Hanno scoperto che gli introni trattenuti in modo anomalo prevalentemente nel nucleo mostrano caratteristiche simili a un "codice di ritenzione intronica" precedentemente identificato. Inoltre, gli autori hanno identificato una classe specifica di trascritti che trattengono gli introni nel citoplasma, con un'affinità notevolmente elevata per le proteine leganti l'RNA (RBP) mal localizzate nella SLA, come TDP-43, FUS e SFPQ. Questi risultati suggeriscono che un sottoinsieme di trascritti che trattengono gli introni nel citoplasma possa avere un ruolo distinto rispetto a quelli precedentemente descritti e potrebbe contribuire alla patogenesi della SLA.

DATA SET	FACTOR OF INTEREST	EXPERIMENTAL GROUP	CONTROL GROUP
GSE152983	Days of differentiation	35	0

CATEGORY	TERM	P-VALUE	BENJAMINI
----------	------	---------	-----------

INTERPRO	Transcriptional repressor NocA- like	9,5E-3	1,0E0
GOTERM_BP_DIRECT	Adult walking behavior	1,0E-2	1,0E0
UP_SEQ_FEATURE	REGION: Disordered	1,0E-2	1,0E0
GOTERM_BP_DIRECT	Nervous system development	1,3E-2	1,0E0
GOTERM_BP_DIRECT	Dendrite morphogenesis	1,3E-2	1,0E0
INTERPRO	Domain of unknown function DUF4074	1,4E-2	1,0E0
GOTERM_BP_DIRECT	Transcription, DNA-templated	1,5E-2	1,0E0
GOTERM_BP_DIRECT	Glossopharyngeal nerve morphogenesis	1,8E-2	1,0E0
GOTERM_BP_DIRECT	Cellular response to jasmonic acid stimulus	1,8E-2	1,0E0
GOTERM_BP_DIRECT	<i>Positive regulation of cell proliferation</i>	3,3E-3	5,5E-1

GOTERM_MF_DIRECT	<i>Microtubule binding</i>	1.0E-2	7,3E-1
INTERPRO	<i>Splicing factor ELAV\HuD</i>	1,09E-2	1,0E0
UP_KW_BIOLOGICAL_PROESS	<i>Differentiation</i>	3,3e-2	4,9e-1
GOTERM_BP_DIRECT	<i>RNA splicing</i>	7,1 E-2	1,0E0
GOTERM_BP_DIRECT	<i>mRNA splice site selection</i>	9,1E-1	1,0E0

3.5 GSE56796

La corteccia cerebrale, responsabile di una vasta gamma di compiti cognitivi e motori complessi, è strutturata attraverso un intricato processo di sviluppo noto come corticogenesi. Questo processo coinvolge l'espansione delle cellule progenitrici, la differenziazione neuronale, la migrazione delle cellule nervose e la formazione di circuiti neuronali. Sebbene la maggior parte delle informazioni sui modelli di espressione genica durante la corticogenesi derivi da studi condotti su modelli murini (Belgard, Oliver, Abaan, Sirey, & al, 2011), recenti avanzamenti nella tecnologia del sequenziamento dell'RNA (RNA-seq) hanno consentito di ottenere una comprensione più approfondita dello sviluppo corticale umano. Le differenze tra la corteccia cerebrale del topo e quella umana sono notevoli, con quest'ultima che presenta una complessità e una dimensione molto maggiori. L'elaborato processo di sviluppo corticale umano, che dura molto più a lungo rispetto a quello dei topi, presenta anche distinzioni molecolari significative. Molti marcatori corticali sono espressi in modo diverso tra topo e uomo, evidenziando le differenze nel profilo di espressione genica. Mentre molti studi sul trascrittoma della corteccia cerebrale umana si sono basati su dati ottenuti da tessuti post-morte utilizzando tecniche di microarray, la tecnologia del sequenziamento profondo è ora considerata essenziale per comprendere appieno i cambiamenti trascrittomici durante lo sviluppo corticale. Per questo studio (Boles, Kiehl, Corneo, Yao, & al, 2014), è stato utilizzato un protocollo per differenziare le cellule staminali

embrionali umane (hESC) in neuroni corticali, seguito da RNA-seq per generare un database completo di trascrittomi che comprende la corticogenesi umana in vitro. I risultati di questo studio sono resi disponibili come risorsa attraverso un database di dati grezzi e un'interfaccia web interattiva <http://cortecon.neuralsci.org/> che consente l'accesso alle analisi computazionali. Fino a poco tempo fa, le conoscenze sui determinanti molecolari della corticogenesi umana venivano principalmente dedotte da studi su modelli non umani. Tuttavia, negli ultimi anni, con gli sviluppi nella differenziazione delle cellule staminali embrionali umane (hESC) in neuroni corticali (Kirwan, Smith, Robinson, & Livesey, 2012), questo sistema è emerso come un prezioso modello di "sviluppo in un piatto" della corticogenesi umana. In questo contesto, è stata creata la risorsa "Cortecon" per assistere i ricercatori nell'esplorare questo argomento. Cortecon cataloga i cambiamenti nel trascrittoma durante la corticogenesi umana in vitro utilizzando la tecnologia RNA-seq, dimostrando una varietà di analisi successive per evidenziarne l'importanza e il valore. Inoltre, è stato sviluppato un sito web facilmente accessibile per rendere questa risorsa disponibile alla comunità scientifica.

Attraverso un confronto con i set di dati BrainSpan RNA-seq, i neuroni corticali generati utilizzando questo protocollo sono stati identificati come più strettamente associati alle regioni corticali prefrontali. Questa associazione è significativa data la rilevanza delle regioni prefrontali in varie condizioni patologiche, come l'autismo (Courchesne, Mouton, Calhoun, & Semendeferi, 2011), la schizofrenia (Pomarol-Clotet, Canales-Rodriguez, Salvador, & al, 2010) e la compromissione cognitiva legata all'età (Mander, Rao, Lu, Saletin, & Lindquist, 2013). Sebbene tale associazione sia indicativa di un'identità corticale prefrontale, è importante notare che questa regione richiede più tempo per maturare rispetto ad altre aree corticali, come quelle somatosensoriali o visive. Questo potrebbe essere attribuito alla persistenza prolungata di geni associati allo sviluppo corticale, un aspetto che merita ulteriori approfondimenti.

Utilizzando il set di dati Cortecon, è possibile esaminare i meccanismi di sviluppo corticale identificati in altri modelli animali per comprenderne il ruolo nell'uomo. Il marcatore gliale radiale NES è stato rilevato durante l'intero periodo di generazione delle cellule progenitrici, come previsto. Inoltre, l'ordine di nascita dei neuroni è risultato coerente con quanto osservato in altre specie, confermando i risultati umani. Ad esempio, la Relina, un marcatore delle cellule Cajal-Retzius nate precocemente, ha mostrato un aumento precoce seguito da un aumento più marcato durante una fase successiva, in linea con quanto riportato per il topo.

La segnalazione FGF e WNT emerge come cruciale anche nella corticogenesi umana, con molti geni correlati che mostrano espressione simile nei dati di Cortecon rispetto al topo. Ad esempio, l'FGF8, noto per il suo ruolo nel patterning dell'area corticale del topo e nella generazione dei neuroni dello strato profondo, ha mostrato un aumento durante le prime fasi della corticogenesi umana nel set di dati Cortecon. I recettori FGF sono stati dimostrati essere essenziali per questo processo, con un picco di espressione durante le fasi iniziali dello sviluppo corticale.

Infine, FGF10, che nel topo regola la differenziazione delle cellule neuroepiteliali in astrociti radiali, ha mostrato un profilo di espressione diverso durante lo sviluppo corticale umano, indicando un ruolo potenzialmente diverso in questo contesto. Questi risultati sottolineano l'importanza dei segnali FGF e WNT nella corticogenesi umana e la loro possibile diversità rispetto al modello murino.

Per esaminare le variazioni nei componenti del percorso WNT, sono stati analizzati i regolatori WNT nel set di dati Cortecon. WNT7B e WNT8B, noti per il loro ruolo chiave nella formazione della corteccia nel topo, hanno mostrato un aumento dell'espressione durante la specifica corticale nel set di dati Cortecon. Oltre alla stimolazione del percorso, è stato evidenziato il ruolo fondamentale dell'antagonismo WNT nello sviluppo corticale, con l'antagonista WNT DKK1 che ha mostrato un aumento durante la fase ND, un calo significativo durante la fase CS e un picco successivo nella fase UL. Allo stesso modo, FRZB ha seguito un andamento simile, ma con un picco maggiore durante la fase UL e un picco più basso durante la fase ND. L'analisi del set di dati Cortecon da diverse prospettive consente ai ricercatori di esplorare ipotesi riguardanti lo sviluppo corticale umano. Un'interessante scoperta emersa durante questa analisi riguarda il catalogo delle trascrizioni di splicing alternativo presenti in diversi stadi dello sviluppo corticale. È stato osservato che i set di geni delle forme di giunzione possono concentrare l'arricchimento delle malattie in categorie più specifiche. Inoltre, sono state identificate diverse isoforme espresse in diversi stadi della corticogenesi in vitro per TGIF1, associato all'oloprosencefalia, e MARK1, implicato nell'ASD. Questi dati suggeriscono che forme di giunzione specifiche potrebbero essere più rilevanti per alcune malattie in base al momento in cui vengono espresse nel programma di sviluppo. Ad esempio, le trascrizioni di TGIF1 espresse specificamente nelle prime fasi dello sviluppo corticale potrebbero essere cruciali per la progressione dell'oloprosencefalia. Pertanto, la risorsa Cortecon sarà preziosa per coloro interessati ai meccanismi di splicing e al loro contributo allo sviluppo, ai cambiamenti nelle espressioni genetiche specifiche e a una vasta gamma di malattie neurologiche. Attraverso l'utilizzo di

questa risorsa, i ricercatori hanno a disposizione dettagliate informazioni sui percorsi di sviluppo, grazie all'applicazione di strategie di analisi di rete. Ad esempio, l'analisi della rete per l'autismo evolutivo ha rivelato che i recettori e i ligandi associati all'ASD svolgono un ruolo chiave in tutte le fasi dello sviluppo corticale. I geni legati alla segnalazione della serotonina e alla segnalazione GABAergica sono particolarmente significativi, poiché sia la serotonina che il GABA sono noti per la loro importanza nello sviluppo corticale e nell'autismo. La rete ASD costruita utilizzando i dati mostra che vari recettori della serotonina e enzimi correlati al percorso della serotonina sono cruciali negli stadi PP e UL, suggerendo un ruolo vitale della serotonina nello sviluppo degli strati corticali superiori, che potrebbero contribuire alla patologia dell'ASD. Inoltre, diversi recettori GABA sono coinvolti in diversi stadi della corticogenesi in vitro, con sottotipi individuali che mostrano modelli di espressione temporale distinti. L'applicazione di questo approccio di rete consente di collegare tra loro diversi cluster, permettendo l'interrogazione delle interazioni tra di essi. Ad esempio, i geni nel cluster 18, composti principalmente da regolatori trascrizionali, possono influenzare l'espressione dei geni nel cluster 34, composto principalmente da molecole di segnalazione. Questo suggerisce l'importanza del cluster 18 nella regolazione dei geni nel cluster 34. L'analisi dei regolatori temporali di ciascun cluster fornisce informazioni su come essi sono connessi e regolati, permettendo di scoprire i meccanismi alla base dello sviluppo corticale umano in salute e malattia.

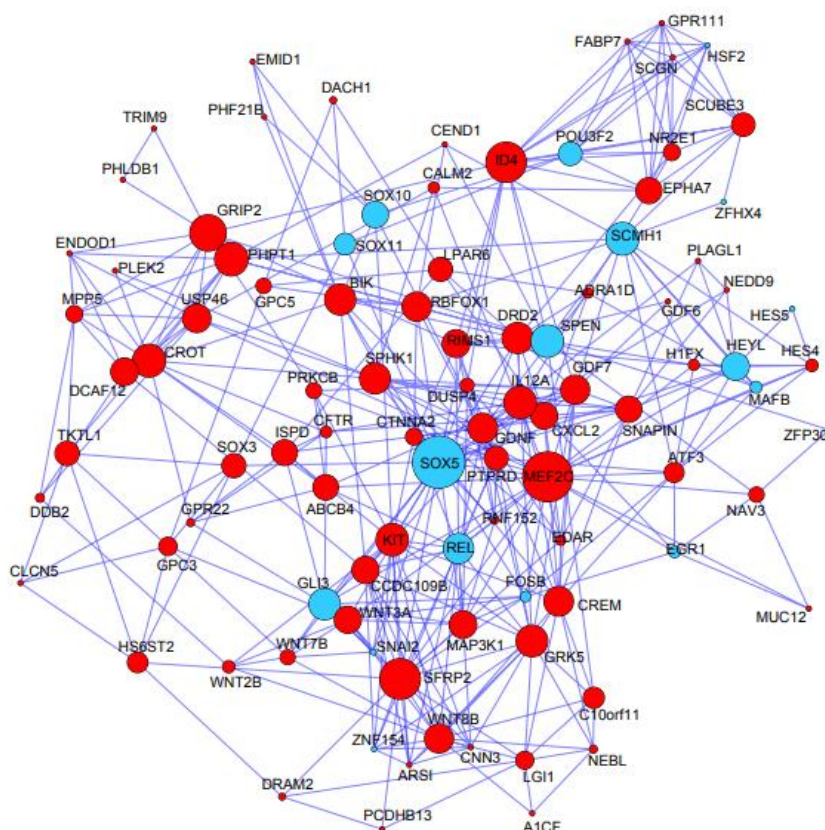


Figura 10: grafico dei fattori di trascrizione del cluster 18 (blu) e del cluster 34 (rosso)¹⁸

DATA SET	FACTOR OF INTEREST	EXPERIMENTAL GROUP	CONTROL GROUP
GSE5679	Characteristic	In_vitro_Hesc_cortecon_day7	In_vitro_Hesc_Test_day7
6	s	7	7

CATEGORY	TERM	P-VALUE	BENJAMINI

¹⁸ <https://pubmed.ncbi.nlm.nih.gov/24991954/>

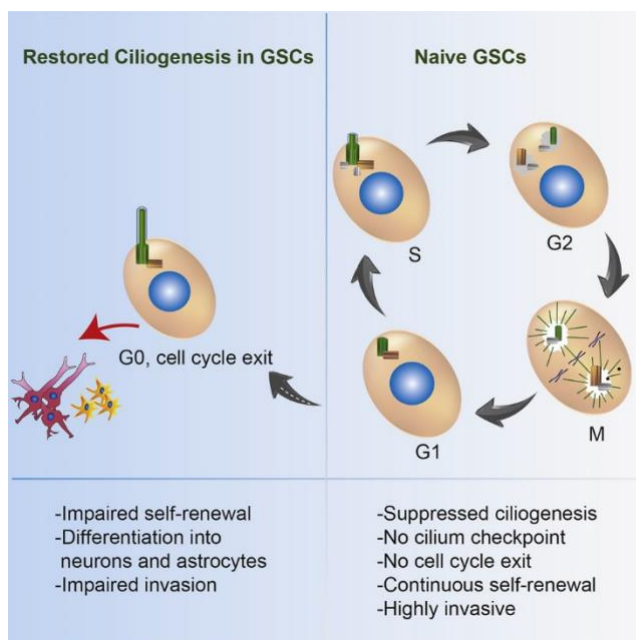
GOTERM_MF_DIRECT	Transmembrane transporter activity	2,7E-2	1,0E0
SMART	WD40	2,7E-2	1,0E0
GOTERM_MF_DIRECT	Thyroid hormone transmembrane transporter activity	3,0E-2	1,0E0
KEGG_PATHWAY	Protein digestion and absorption	3,5E-2	1,0E0
UP_SEQ_FEATURE	DOMAIN:RIIa	3,6E-2	1,0E0
GOTERM_BP_DIRECT	Hexose transport	4,3E-2	1,0E0
GOTERM_MF_DIRECT	D-glucose transmembrane transporter activity	4,5E-2	1,0E0
UP_SEQ_FEATURE	DOMAIN: EF- hand	5,0E-2	1,0E0
UP_SEQ_FEATURE	REGION: AA6	5,2E-2	1,0E0
GOTERM_BP_DIRECT	Canonical Wnt signaling pathway	5,4E-2	1,0E0

UP_KM_DISEASE	<i>Primary ciliary dyskinesia</i>	2,5E-3	5,1E-2
GOTERM_MF_DIRECT	<i>MiRNA binding</i>	1,1E-2	8,7E-1
UP_KM_CELLULAR_COMPONENT	<i>Cilium</i>	1,4E-2	8,8E-2

3.6 GSE179439

Le GSC¹⁹ derivate dal paziente reclutano livelli elevati di proteine che promuovono il disassemblaggio del ciglio temporale, impedendo la formazione dei cigli. L'esaurimento dei componenti del complesso di disassemblaggio delle ciglia induce la ciliogenesi in un gruppo di GSC tramite il trasferimento del recettore alfa del fattore di crescita derivato dalle piastrine (PDGFR- α) in un ciglio appena formato. Questo ripristino della ciliogenesi porta le GSC a passare dall'autorinnovamento alla differenziazione.

Attraverso test di invasione del glioma basati su organoidi e xenotrapianti cerebrali nei topi, è stato dimostrato che la differenziazione indotta dalla ciliogenesi può prevenire l'infiltrazione delle GSC nel cervello. Questi risultati (Goranci-Buzhala, et al., 2021), evidenziano un ruolo cruciale del ciglio come interruttore molecolare nel determinare il destino delle GSC e suggeriscono l'induzione del ciglio come potenziale strategia terapeutica per contrastare la proliferazione delle GSC.



¹⁹ Cellule staminali di glioma

Figura 11: L'induzione del ciglio innesca la differenziazione delle cellule staminali del glioma

<https://pubmed.ncbi.nlm.nih.gov/34496239/>

Il glioblastoma (GBM), il tumore cerebrale maligno più comune, continua a rappresentare una sfida terapeutica significativa, con un tempo di sopravvivenza mediano di soli 15 mesi nonostante le terapie attualmente disponibili. Le cellule staminali di glioma derivate dal paziente (GSC) sono particolarmente cruciali nell'iniziazione, mantenimento e invasione del tumore, rappresentando quindi un bersaglio primario per interventi terapeutici. Tuttavia, la comprensione della biologia delle GSC è ancora incompleta.

Recenti studi hanno evidenziato il ruolo del ciglio primario, una struttura cellulare coinvolta nella ricezione di segnali, nel controllo del ciclo cellulare e nel destino delle cellule staminali neurali. Il ciglio si assembla durante l'uscita dal ciclo cellulare e si disassembla al rientro nel ciclo, agendo come un "checkpoint del ciglio" che regola la progressione del ciclo cellulare. Componenti proteici come CPAP, Nde1, HDAC6, Aurora-A e OFD1 sono coinvolti nell'assemblaggio di un complesso di disassemblaggio delle ciglia (CDC) che promuove il disassemblaggio delle ciglia (Gabriel, et al., 2016). Il tempismo preciso di questo processo influenza la durata della transizione G1-S e, di conseguenza, l'autorinnovamento delle cellule staminali neurali. Il prolungamento della fase G1 a causa di un ritardato disassemblaggio delle ciglia può portare a una differenziazione precoce delle cellule staminali neurali e quindi a una diminuzione del loro pool.

In sintesi, comprendere i meccanismi che regolano il ciclo vitale del ciglio e il suo impatto sull'autorinnovamento delle cellule staminali neurali potrebbe aprire nuove strade per terapie mirate contro il GBM.

È stato osservato che la perdita delle ciglia è una caratteristica comune in diversi tipi di cancro, tra cui il cancro al seno, alla prostata, alla pelle, al melanoma e al pancreas. Tuttavia, solo pochi studi hanno indagato la perdita delle ciglia nelle cellule del glioblastoma (GBM), sollevando interrogativi sulla sua relazione con la ciliogenesi e la formazione normale delle ciglia (Yang, Roine, & Mäkelä, 2013).

Uno studio condotto su tumori di GBM ha rivelato che le interruzioni si verificano nelle prime fasi della ciliogenesi, suggerendo un coinvolgimento delle proteine CDC in questo processo critico per la soppressione della ciliogenesi (Moser, Fritzler, & Rattner, 2014). Si ipotizza che la soppressione della ciliogenesi nelle cellule staminali di glioma (GSC) possa conferire loro un vantaggio selettivo nell'autorinnovamento continuo, senza uscita e differenziazione dal ciclo cellulare.

In questo studio (Goranci-Buzhala, et al., 2021), è stato dimostrato che le GSC derivate dai pazienti e i tessuti di glioma clinico mostrano una ciliogenesi soppressa a causa del reclutamento elevato dei CDC nei corpi basali. Prevenire il reclutamento dei CDC e alterare il loro comportamento di localizzazione dinamica nel corpo basale è risultato sufficiente per indurre la formazione del ciglio in un sottoinsieme di GSC. Dopo l'induzione del ciglio, il recettore PDGFR- α viene sequestrato nel ciglio appena formato, causando una riduzione complessiva dei livelli di PDGFR- α e portando le GSC da uno stato di autorinnovamento a uno stato di differenziazione.

Infine, è stato dimostrato che le GSC indotte con ciglia non riescono ad infiltrarsi negli organoidi cerebrali umani derivati da cellule staminali pluripotenti indotte (iPSC) e nel cervello dei topi. Questi risultati suggeriscono che l'induzione del ciglio può svolgere un ruolo cruciale nel determinare il destino delle GSC, aprendo nuove prospettive per l'interferenza terapeutica nel GBM.

Il glioblastoma è caratterizzato da cellule staminali tumorali altamente autorinnovanti e tumorigeniche, che mostrano somiglianze sorprendenti con le cellule staminali neurali. Queste cellule tumorali sono in grado di invadere rapidamente gli organoidi cerebrali umani e il cervello dei topi, formando tumori che replicano le caratteristiche dei tumori riscontrati nei pazienti. Identificare le peculiarità delle cellule staminali tumorali e comprendere la loro regolazione molecolare può essere fondamentale per sviluppare strategie mirate per contrastare la loro proliferazione e invasione.

Le cellule staminali neurali (NPC) sono conosciute per la loro capacità di autorinnovarsi e differenziarsi, con le ciglia primarie che svolgono un ruolo cruciale nel regolare questo processo. Si è ipotizzato che nelle cellule staminali tumorali del glioblastoma (GSC) le ciglia siano soppressate nelle fasi iniziali della loro formazione, a causa di livelli elevati di componenti molecolari che promuovono il loro disassemblaggio. Questo potrebbe conferire alle GSC un vantaggio selettivo nell'autorinnovamento continuo, indipendentemente dal normale controllo cellulare.

In questo studio, è stato identificato che l'esaurimento dei componenti molecolari coinvolti nel disassemblaggio delle ciglia può indurre persistentemente la formazione di ciglia nelle GSC. L'induzione della ciliogenesi ha portato alla perdita delle proprietà staminali delle GSC e all'acquisizione di programmi di differenziazione. Di conseguenza, le GSC con ciglia indotte non sono state in grado di invadere gli organoidi cerebrali umani tridimensionali e il cervello dei topi.

Il reclutamento dei componenti del complesso di disassemblaggio delle ciglia (CDC) nella base ciliare è stato collegato al processo di disassemblaggio delle ciglia temporali all'inizio del rientro nel ciclo cellulare. Tuttavia, le interazioni molecolari sottese a questo reclutamento e al disassemblaggio delle ciglia sono state fino ad ora poco comprese. Si ipotizza che il CDC possa contenere un enzima in grado di depolimerizzare il citoscheletro dei microtubuli delle ciglia durante il loro disassemblaggio. L'identificazione di Kif24, una chinesina depolimerizzante dei microtubuli, e di Nek2 come componenti del CDC ha iniziato a chiarire i meccanismi alla base del disassemblaggio delle ciglia mediato da questo complesso. È interessante notare che i livelli di Nek2, una chinasi coinvolta nel disassemblaggio delle ciglia, sono aumentati nelle GSC native, che presentano una ciliogenesi soppressa.

Inducendo la ciliogenesi, le GSC tendono a comportarsi in modo simile alle cellule staminali neurali progenitrici (NPC) sottoposte a differenziazione, suggerendo che le ciglia appena indotte siano strutturalmente e funzionalmente normali. Questo supporta l'idea che le ciglia possano essere ripristinate funzionalmente in un sottoinsieme di GSC classificati come simili alle cellule progenitrici oligodendrogliali (OPC), che esprimono un livello elevato del recettore PDGFR- α . Il ripristino delle ciglia sembra innescare la perdita di staminalità delle GSC e l'acquisizione di programmi di differenziazione. L'induzione delle ciglia ha anche portato alla riduzione dei livelli totali di PDGFR- α e alla diminuzione della segnalazione aberrante del PDGFR- α , suggerendo un possibile ruolo delle ciglia nel controllo della quantità e della segnalazione a valle di questa proteina.

La segnalazione mediata dal recettore PDGFRA è nota per essere regolata dal ciglio primario. È interessante notare che l'espressione di Nek2-KD ha sensibilizzato solo le GSC simili alle OPC che sovraesprimono PDGFR- α , suggerendo una certa specificità nella risposta. Alcuni fattori che potrebbero contribuire a questa selettività includono i livelli più elevati di Nek2 nelle GSC simili alle OPC e la presenza di programmi di induzione delle ciglia che potrebbero essere attivi solo in questi sottotipi cellulari.

Al momento, non sono chiare le ragioni per cui l'induzione del ciglio non è stata osservata nelle linee cellulari non rispondenti. Potrebbe essere dovuto alla presenza di centrioli difettosi o a mutazioni genetiche nei geni coinvolti nella formazione delle ciglia. Ulteriori studi saranno necessari per comprendere appieno le basi molecolari di questa non risposta.

In sintesi, il checkpoint del ciglio emerge come un promettente interruttore molecolare che può essere manipolato per influenzare gli aspetti fondamentali delle GSC. L'identificazione dei componenti del CDC e la comprensione della loro interazione con il processo di

ciliogenesi stanno iniziando a spiegare perché le GSC tendono ad auto-rinnovarsi e a invadere il cervello. La differenziazione indotta dalle ciglia sembra avere un impatto significativo sulla capacità invasiva delle GSC, suggerendo che prendere di mira il CDC potrebbe rappresentare una strategia promettente per limitare l'invasione delle GSC e quindi la progressione del tumore cerebrale.

DATA SET	FACTOR OF INTEREST	EXPERIMENTAL GROUP	CONTROL GROUP
GSE179439	Genotype/variation	NEK2-KD	Control

CATEGORY	TERM	P-VALUE	BENJAMINI
GOTERM_BP_DIRECT	Negative regulation of canonical Wnt signaling pathway	5,0E-2	1,0E0
GOTERM_BP_DIRECT	Signal transduction	5,0E-2	1,0E0
GOTERM_BP_DIRECT	Limb bud formation	5,2E-2	1,0E0
GOTERM_BP_DIRECT	Notochord development	5,6E-2	1,0E0

GOTERM_BP_DIRECT	Positive regulation of insulin-like growth factor receptor signaling pathway	6,1E-2	1,0E0
GOTERM_BP_DIRECT	Negative regulation of cell growth involved in cardiac muscle cell development	6,1E-2	1,0E0
GOTERM_BP_DIRECT	Female pregnancy	6,5E-2	1,0E0
GOTERM_BP_DIRECT	Digestive tract morphogenesis	6,6E-2	1,0E0
GOTERM_BP_DIRECT	Extracellular matrix assembly	6,6E-2	1,0E0
GOTERM_BP_DIRECT	Central nervous system myelination	7,2E-2	1,0E0

GOTERM_BP_DIRECT	<i>Positive regulation of neuroblast proliferation</i>	3,0E-5	4,3E-3
GOTERM_BP_DIRECT	<i>Plasma cell differentiation</i>	2,4E-2	7,1E-1
GOTERM_BP_DIRECT	<i>Osteoblast differentiation</i>	2,4E-2	7,1E-1

CONCLUSIONE

Attraverso le analisi di trascrittomica, derivate dai precedenti data set, su campioni sia sani che patologici, è stato individuato il colangiocarcinoma come un contesto in cui si manifestano alterazioni spliceosomiche associate al centrosoma, suggerendo un potenziale coinvolgimento di tali meccanismi nella patogenesi di questa malattia. In questa tesi, è stato approfondito il ruolo cruciale dei condensati ribonucleoproteici (RNP) nel controllo del metabolismo e dello splicing dell'RNA nelle cellule animali. Ponendo l'attenzione sul centrosoma, il principale centro di organizzazione dei microtubuli nelle cellule animali, è stata impiegata la trascrittomica per mappare le reti di interazione proteica dei condensati RNP. Una limitazione della trascrittomica è che fornisce solo una valutazione statica della trascrizione genica in un dato momento temporale e non tiene conto di altri importanti livelli di regolazione dell'espressione genica, come la stabilità del trascritto, la traduzione proteica e la stabilità proteica. Questo significa che i cambiamenti nell'RNA trascritto possono non riflettere necessariamente i cambiamenti nella produzione proteica effettiva.

Inoltre, la trascrittomica può essere influenzata da vari fattori tecnici e biologici, come la qualità del campione, la sensibilità e la specificità dei metodi di rilevamento, nonché la complessità della regolazione genica nelle diverse condizioni biologiche. Queste limitazioni possono portare a una comprensione incompleta dei processi biologici studiati e possono richiedere l'integrazione di dati provenienti da altre tecniche omiche per ottenere una visione più completa e accurata dell'espressione genica e della sua regolazione.

In uno studio è emerso che numerose proteine dello spliceosoma nucleare sono strettamente associate alle sub-strutture del centrosoma coinvolte nella divisione nucleare e nella formazione delle ciglia.

In particolare, è stato individuato un componente dello spliceosoma nucleare, noto come BUD31, il quale interagisce sperimentale alla proteina satellite dei centrioli OFD1, confermando il ruolo funzionale di queste interazioni. Attualmente, l'idea che eventi di splicing possano verificarsi occasionalmente, o essere completati, nel citoplasma, rimane non dimostrata, anche se allo stato attuale è stata messa in dubbio l'idea secondo cui lo spliceosoma minore sarebbe localizzato in un compartimento diverso da quello dello spliceosoma maggiore.

Le tecniche di rilevamento, ad esempio, l'uso di analisi bioinformatiche, come microarray o RNA-Seq, per identificare varianti di splicing possono essere influenzate dalla qualità dei dati e dall'accuratezza delle previsioni. Nonostante i progressi nello studio di possibili meccanismi che indicherebbero eventi di splicing citoplasmatico, molte delle sue funzioni specifiche e i meccanismi regolatori coinvolti possono ancora essere sconosciuti o non completamente compresi. Ciò potrebbe impedire l'accesso a nuove applicazioni diagnostiche e terapeutiche basate sulla capacità di manipolare selettivamente lo splicing citoplasmatico.

In sintesi, sebbene lo splicing citoplasmatico potrebbe essere un processo fondamentale per la regolazione dell'espressione genica e la diversificazione proteica, la sua attuale conoscenza è molto limitata e richiede ulteriori studi e sviluppo di nuove tecnologie per essere completamente compreso e sfruttato. Questi sviluppi potrebbero avere un impatto significativo sulla medicina personalizzata e sulla terapia delle malattie genetiche e complesse.



Bibliografia

- Alliegro, M. A., Henry, J. J., & Alliegro, M. C. (2010). *Rediscovery of the nucleolinus, a dynamic RNA-rich organelle associated with the nucleolus, spindle, and centrosomes*. Proc. Natl. Acad. Sci. USA.
- Allilegro, M. C., Allilegro, M. A., & Palazzo, R. E. (2006). *Centrosome-associated RNA in surf clam oocytes*. Proc. Natl. Acad. Sci. USA.
- Alpini, G., Ulrich, C., Roberts, S., & Ueno, Y. (1997). *Molecular and functional heterogeneity of cholangiocytes from rat liver after bile duct ligation*. Am J Physiol.
- Amato, R., Morleo, M., Giaquinto, L., Di Bernardo, D., & Franco, B. (2014). *A network-based approach to dissect the cilia/centrosome complex interactome*. BMC Genom.
- Antoniou, A., Raynaud, P., Cordi, S., Zong, Y., & al, e. (2009). *Intrahepatic bile ducts develop according to a new mode of tubulogenesis regulated by the transcription factor SOX9*. Gastroenterology.
- Auinash, K., & al, e. (2011). *Functional consequences of developmentally regulated alternative splicing*. Nat. Rev Genet.

- Belgard, T., Oliver, P., Abaan, H., Sirey, T., & al, e. (2011). *A transcriptomic atlas of mouse neocortical layers*. Neuron.
- Bober, M. B., & Jackson, A. P. (2017). *Microcephalic osteodysplastic primordial dwarfism, type II: a clinical review*. Curr. Osteoporos. Rep. 15, 61–69.
- Boles, N., Kiehl, T., Corneo, B., Yao, S., & al, e. (2014). *CORTECON: a temporal transcriptome analysis of in vitro human cerebral cortex development from human embryonic stem cells*. Neuron.
- Bond, J., Roberts, E., Springell, K., & al., e. (2005). *A centrosomal mechanism involving CDK5RAP2 and CENPJ controls brain size*. Nat. Genet. 37, 353–355.
- Brandt, O., & Hoheisel, J. D. (2004). *Peptide nucleic acids on microarrays and other biosensors*. Trends Biotechnol.
- Buhring, H., Kuci, S., Conze, T., Rathke, G., & al, e. (2004). *CDCP1 identifies a broad spectrum of normal and malignant stem/progenitor cell subsets of hematopoietic and nonhematopoietic origin*. Stem Cell.
- Busselez, J., Pancione, M., & al, e. (2023). *New insights into the centrosome-associated spliceosome components as regulators of ciliogenesis and tissue identity*. Wiley Periodicals LLC.
- Camargo Ortega, G., Falk, S., Johansson, P. A., & al, e. (2019). *The centrosome protein AKNA regulates neurogenesis via microtubule organization*. Nature 567, 113–117.
- Chavali, P., Putz, M., & Gergely, F. (2014). *Small organelle, big responsibility: the role of centrosomes in development and disease*. Philos. Trans. R. Soc. Lond. B Biol. Sci. 369, 20130468.
- Chen, J., Foxe, J., Pedrosa, E., & al, e. (2013). *Transcriptome Comparison of Human Neurons Generated Using Induced Pluripotent Stem Cells Derived from Dental Pulp and Skin Fibroblasts*. PLoS One.

- Chen, Q., & Hu, G. (2017). *Post-transcriptional regulation of the pluripotent state*. Curr. Opin. Genet. Dev.
- Collotta, M. (2018). Trascrittomica: la differenza che conta. *Rivista Società Italiana di Medicina Generale*, N.4 VOL.25.
- Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., & al, e. (2016). *A survey of best practices for RNA-seq data analysis*. Genome Biology.
- Conze, T., Lammers, R., Kuci, S., Scherl-Mostageer, M., & al, e. (2003). *CDCP1 is a novel marker for hematopoietic stem cells*. Ann N Y Acad Sci.
- Courchesne, E., Mouton, M., Calhoun, M., & Semendeferi, K. (2011). *Neuron number and size in prefrontal cortex of children with autism*. JAMA.
- Denis, M. M., Tolley, N. D., Bunting, M., Schwertz, H., & al, e. (2005). *Escaping the nuclear confines: signaldependent pre-mRNA splicing in anucleate platelets*. Cell.
- DeRisi, J. L., Iyer, V. R., & Brown, P. O. (1997). *Exploring the metabolic and genetic control of gene expression on a genomic scale*. Science.
- Dianat, N., Steichen, C., Combettes, L., Raveux, A., & al, e. (2014). *Generation of functional cholangiocyte-like cells from human pluripotent stem cells and HepaRG cells*. Hepatology.
- Dianat, N., Steichen, C., Vallier, L., Weber., & al, e. (2013). *Human pluripotent stem cells for modelling human liver diseases and cell therapy*. Curr Gene Ther.
- D'orazio, K., & Green, R. (2021). *Ribosome states signal RNA quality control*. Mol. Cell.
- Gabriel, E., Wason, A., Ramani, A., Gooi, L., Keller, P., & al, e. (2016). *CPAP promotes timely cilium disassembly to maintain neural progenitor pool*. EMBO J.
- Ghindilis, A. L., Smith, M. W., Schwarzkopf, K. R., & al, e. (2007). *CombiMatrix oligonucleotide arrays: genotyping and gene expression assays employing electrochemical detection*. Biosens Bioelectron.

- Glanzer, J., Miyashiro, K. Y., & Sul, J. Y. (2005). *RNA splicing capability of live neuronal dendrites*. Proc Natl Acad Sci U S A.
- Goranci-Buzhala, G., Mariappan, A., Ricci-Vitiani, L., Josipovic, N., Pancioni, S., Gottardo, M., & al, e. (2021). *Cilium induction triggers differentiation of glioma stem cells*. Cell Rep.
- Guernsey, D. L., Jiang, H., & Hussin, J. (2010). *Mutations in centrosomal protein CEP152 in primary microcephaly families linked to MCPH4*. Am. J. Hum. Genet. 87, 40–51.
- Hauser, N. C., Vingron, M., Scheideler, M., & al, e. (1998). *Transcriptional profiling on all open reading frames of Saccharomyces cerevisiae*. Yeast.
- Hauser, N. C., Vingron, M., Scheideler, M., & al, e. (1998). *Transcriptional profiling on all open reading frames of Saccharomyces cerevisiae*. Yeast.
- Humphreys, E., Wlliams, K., Adams, D., & Afford, S. (2010). *Primary and malignant cholangiocytes undergo CD40 mediated Fas dependent apoptosis, but are insensitive to direct activation with exogenous Fas ligand*. Plos One.
- Ingolia, N. T., Brar, G. A., Rouskin, S., McGreachy, A. M., & al, e. (2012). *The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments*. Nature Protocols.
- Iwasaki, M., Kawahara, Y., Okubo, C., Yamakawa, T., & al, e. (2022). *Multi-omics approach reveals posttranscriptionally regulated genes are essential for human pluripotent stem cells*. iScience.
- Jacob, A., Brandt, O., Stephan, A., & al, e. (2004). *Peptide nucleic acid microarrays*. Methods Mol Bio.
- Kaivon, Y., & Makeyven, E. (2020). *A transcriptome-wide antitermination mechanism sustaining identity of embryonic stem cells*. Nat. Commun.
- Kalay, E., Yigit, G., Aslan, Y., & al, e. (2011). *CEP152 is a genome maintenance protein disrupted in Seckel syndrome*. Nat. Genet. 43, 23–26.

- Keren, H. (2010). *Alternative splicing and evolution: diversification, exon definition and function*. Nature Reviews Genetics.
- Khan, Z., Ford, M., Cusanovich, D., & al, e. (2013). *Primate Transcript and Protein Expression Levels Evolve under Compensatory Selection Pressures*. Science.
- Kirwan , P., Smith, J., Robinson, H., & Livesey, F. (2012). *Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses*. Nat. Neurosci.
- Kolodziejczyk, A. A., Kim, J. K., Svensson, V., Marioni, J. C., & al, e. (2015). *The technology and biology of single-cell RNA sequencing*. Molecular Cell.
- Konig, H., Matter, N., Bader, R., Thiele, W., & al, e. (2007). *Splicing segregation: the minor spliceosome acts outside the nucleus and controls cell proliferation*. Cell.
- Konig, H., Matter, N., Bader, R., Thiele, W., & Muller, F. (2007). *Splicing segregation: the minor spliceosome acts outside the nucleus and controls cell proliferation*. Cell.
- Kukurba, K. R., & Montgomery, S. B. (2015). *RNA Sequencing and Analysis*. Cold Spring Harbor Protocols.
- Lancaster, M. A., Renner, M. A., & al, e. (2013). *Cerebral organoids model human brain development and microcephaly*. Nature 501, 373–379.
- Liao, Y., Wang, J., Jaehnig, E. J., & al, e. (2019). *"WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs*. Nucleic Acids Research.
- Lin, H., Pedrosa, E., Wang, T., Zheng, D., & al, e. (2012). *Allele-biased expression in differentiating human neurons: Implications for neuropsychiatric disorders*. PLoS One.
- Lu, R., Markowitz, F., Unwin, R., Leek, J., & al, e. (2009). *Systems-level dynamic analyses of fate change in murine embryonic stem cells*. Nature.
- Luisier, R., Tyzack, G., Hall, C., & al, e. (2018). *Intron retention and nuclear loss of SFPQ are molecular hallmarks of ALS*. Nat Commun.

- Maier, C. A., Kumar-Sinha, C., Cao, X., Kalyana-Sundaram, S., Han, B., & al, e. (2009). *Transcriptome sequencing to detect gene fusions in cancer*. Nature.
- Mander, B., Rao, V., Lu, B., Saletin, J., & Lindquist, J. (2013). *Prefrontal atrophy, disrupted NREM slow waves and impaired hippocampal-dependent memory in aging*. Nat. Neurosci.
- Margulies, M., Egholm, M., Altman, W. E., & al, e. (2005). *Genome sequencing in microfabricated high-density picolitre reactors*. Nature.
- Mauger, O., Lemoine, F., & Scheiffele, P. (2016). *Targeted intron retention and excision for rapid gene regulation in response to neuronal activity*. Neuron.
- Montoro, D. T., Haber, A. L., Biton, M., Vinarsky, V., & al, e. (2018). *A revised airway epithelial hierarchy includes CFTR-expressing ionocytes*. Nature.
- Moser, J., Fritzler, M., & Rattner, J. (2014). *Ultrastructural characterization of primary cilia in pathologically characterized human glioblastoma multiforme (GBM) tumors*. BMC Clin. Pathol.
- Nigger, E. A., & Holland, A. J. (2017). *Once and only once: mechanisms of centriole duplication and their deregulation in disease*. Nat. Rev. Mol. Cell Biol. 19, 297–312.
- Oikonomopoulos, S., Bayega, A., Fahiminiya, S., Djambazian, H., & al, e. (2020). *Methodologies for Transcript Profiling Using Long-Read Technologies*. Frontiers in Genetics.
- Pessa, H. K., Will, C. L., Meng, X., Schneider, C., & al, e. (2008). *Minor spliceosome components are predominantly localized in the nucleus*. Proc Natl Acad Sci U S A.
- Plasschaert, L. W., Žilionis, R., Choo-Wing, R., & al, e. (2018). *A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte*. Nature.
- Pomarol-Clotet, E., Canales-Rodriguez, E., Salvador, R., & al, e. (2010). *Medial prefrontal cortex pathology in schizophrenia as revealed by convergent findings from multimodal imaging*. Mol. Psychiatry.

- Price, A., Hwang, T., Tao, R., & al, e. (2020). *Characterizing the nuclear and cytoplasmic transcriptomes in developing and mature human cortex uncovers new insight into psychiatric disease gene regulation*. *Genome Res*.
- Prutt, K. D., Tatusova, T., Brown, G. R., & al, e. (2012). *NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy*. *Nucleic Acids Res.* .
- Pulvers, J. N., Bryk, J., Fish, J. L., & al, e. (2010). *Mutations in mouse Aspm (abnormal spindle-like microcephaly associated) cause not only microcephaly but also major defects in the germline*. *Proc. Natl Acad. Sci. USA* 107, 16595–16600.
- Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data*. *Bioinformatics*.
- Setia, H., & Muotri, A. R. (2019). *Brain organoids as a model system for human neurodevelopment and disease*. *Semin. Cell Dev. Biol.* 95, 93–97.
- Sforza, S., Tedeschi, T., Bencivenni, M., & al, e. (2014). *Use of peptide nucleic acids (PNAs) for genotyping by solution and surface methods*. *Methods Mol Biol*.
- Shapiro, E., Biezuner, T., & Linnarsson, S. (2013). *Single-cell sequencing-based technologies will revolutionize whole-organism science*. *Nature Reviews. Genetics*.
- Sharangdhar, T., Sugimoto, Y., Heraud-Farlow, J., & al, e. (2017). *A retained intron in the 3'- UTR of Calm3 mRNA mediates its Staufen2- and activity-dependent localization to neuronal dendrites*. *EMBO Rep*.
- Takeda, H., Fujimori, Y., Ogawa, H., & al, e. (2010). *CD318/CUB-domain-containing protein 1 expression on cord blood hematopoietic progenitors*. *Exp Ther Med*.
- Thind, A. S., Monga, I., Thakur, P. K., Kumari, P., Dindhoria, K., & al, e. (2011). *Demystifying emerging bulk RNA-Seq applications: the application and utility of bioinformatic methodology*. *Briefings in Bioinformatics*.

- Tresenrider, A., Morse, K., Jorgensen, V., Chia, M., & al, e. (2021). *Integrated genomic analysis reveals key features of long undecoded transcript isoform-based gene repression*. Mol. Cell.
- Turunen, J. J., Niemala, E. H., Verma, B., & Frilander, M. J. (2013). *The significant other: splicing by the minor spliceosome*. WIREs RNA.
- Tyzack, G., Neeves, J., Crerar, H., Klein, O., & al, e. (2021). *Aberrant cytoplasmic intron retention is a blueprint for RNA binding protein mislocalization in VCP-related amyotrophic lateral sclerosis*. Brain.
- Umemura, Y., Koike, N., Ohashi, M., Meng, Q., & al, e. (2017). *Involvement of posttranscriptional regulation of Clock in the emergence of circadian clock oscillation during mouse development*. Proc. Natl. Accade. Sci.
- Volker, A., & al, e. (2015). *Microarrays as Research Tools and Diagnostic Devices*. RNA and DNA Diagnostics.
- Volker, A., & al, e. (2015). *Microarrays as Research Tools and Diagnostic Devices*. RNA and DNA Diagnostics.
- Vroman, B., & LaRusso, N. (1996). *Development and characterization of polarized primary cultures of rat intrahepatic bile duct epithelial cells*. Lab Investig J Tech Methods Pathol.
- Wang, E., Sandberg, R., Lou, S., Khrebtkova, I., Zhang, L., Mayr, C., . . . Burge, C. B. (2008). *Alternative isoform regulation in human tissue transcriptomes*. Nature.
- Wang, Z., & Gerstein, M. (2009). *RNA-Seq: a revolutionary tool for transcriptomics*. Nat Review Genet .
- Wodicka, L., Dong, H., Mittman, M., & al, e. (1997). *Genome-wide expression monitoring in Saccharomyces cerevisiae*. Nat Biotechnol.

- Woods, C. G., Bond, J., & Enard, W. (2005). *Autosomal recessive primary microcephaly (MCPH) a review of clinical, molecular, and evolutionary findings*. Am. J. Hum. Genet. 76, 717–728.
- Yang, Y., Roine, N., & Mäkelä, T. (2013). *CCRK depletion inhibits glioblastoma cell proliferation in a cilium-dependent manner*. EMBO Rep.
- Yap, K., Lim, Z., Khandelia, P., Friedman, B., & al, e. (2012). *Coordinated regulation of neuronal mRNA steady-state levels through developmentally controlled intron retention*. Genes.
- Zhang, R., Zheng, Y., Li, B., & al, e. (2018). *Hepatic stem cells with self-renewal and liver repopulation potential are harbored in CDCPI-positive subpopulations of human fetal liver cells*. Stem Cell Res Ther.
- Zhang., R., Zheng, Y., Li, B., & al, e. (2015). *Human hepatic stem cells transplanted into a fulminant hepatic failure Alb-TRECK/SCID mouse model exhibit liver reconstitution and drug metabolism capabilities*.