



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Università degli Studi di Padova
Dipartimento di Scienze Oncologiche e Chirurgiche

SCUOLA DI DOTTORATO DI RICERCA IN:
ONCOLOGIA ED ONCOLOGIA CHIRURGICA
-CICLO XXII-

STUDIO DEL RUOLO DEI MICRORNA NELLA MATURAZIONE E NELLA LEUCEMOGENESI T

Direttore della Scuola: Ch.ma Prof.ssa Paola Zanovello

Supervisore: Ch.mo Prof. Alberto Amadori

Dottoranda: Margherita Ghisi

RIASSUNTO **-1-**

SUMMARY **-4-**

1. INTRODUZIONE **-6-**

- 1.1 Linfopoiesi T nel timo **-6-**
- 1.2 Leucemia linfoblastica acuta a cellule T (T-ALL) **-10-**
- 1.3 Linfoma linfoblastico T (T-LBL) **-13-**
- 1.4 Biosintesi e biologia dei microRNA **-14-**
- 1.5 Ruolo dei microRNA nella differenziazione e nella fisiologia dei linfociti **-19-**
- 1.6 Ruolo dei microRNA nella leucemogenesi **-24-**
 - 1.6.1 Leucemia Linfatica Cronica (CLL) **-25-**
 - 1.6.2 Linfoma **-26-**
 - 1.6.3 Leucemia Linfatica Acuta (ALL) **-27-**

2. SCOPO DELLA TESI **-29-**

3. MATERIALI E METODI **-30-**

- 3.1. Campioni primari e colture cellulari **-30-**
- 3.2. Isolamento delle sottopopolazioni timiche e linfoidi da timo e da sangue periferico **-31-**
- 3.3. Estrazione dell'RNA **-34-**
- 3.4. Quantificazione dell'RNA estratto e verifica della sua qualità **-35-**
 - 3.4.1 Analisi spettrofotometrica **-35-**
 - 3.4.2 Analisi elettroforetica **-35-**

3.5.	<u>Analisi del profilo d'espressione di microRNA mediante microarray</u>	-36-
3.6.	<u>Analisi dell'espressione di microRNA tramite TaqMan qRT-PCR</u>	-37-
3.7.	<u>Generazione di librerie di small-RNA</u>	-41-
3.8.	<u>Predizione computazionale degli mRNA bersaglio dei microRNA</u>	-45-
3.9.	<u>Analisi del profilo d'espressione genica (GEP) mediante microarray</u>	-46-
3.10.	<u>Produzione di vettori lentivirali</u>	-47-
3.11.	<u>Trasduzione cellulare <i>in vitro</i> mediante vettori lentivirali</u>	-49-
3.12.	<u>Controllo citofluorimetrico dell'efficienza di trasduzione</u>	-50-
3.13.	<u>Analisi della proliferazione cellulare mediante test dell'ATP</u>	-50-
3.14.	<u>Analisi del ciclo cellulare con Ioduro di Propidio</u>	-51-
3.15.	<u>Analisi statistiche</u>	-52-
	3.15.1 MiR microarray	-52-
	3.15.2 Librerie di small-RNA	-52-
	3.15.3 GEP array	-52-
4.	<u>RISULTATI</u>	-53-
4.1.	<u>Profilo d'espressione dei microRNA in sottopopolazioni timocitarie umane a diversi stadi di maturazione</u>	-53-

4.2. <u>Generazione di librerie di small-RNA da sottopopolazioni linfoidi T</u>	-58-
4.2.1. Generazione della pipeline bioinformatica	-58-
4.2.2. Analisi delle librerie di small-RNA ottenute da popolazioni linfoidi T	-61-
4.3. <u>Confronto fra i dati dei miR microarray e delle librerie di small-RNA: identificazione di miR comunemente regolati nel corso della maturazione dei linfociti T</u>	-65-
4.4. <u>Validazione tramite qRT-PCR dei miR differenzialmente regolati nella differenziazione T</u>	-67-
4.5. <u>Predizione computazionale dei potenziali bersagli di miR d'interesse</u>	-68-
4.6. <u>Combinazione dei profili d'espressione genica e dei miR di popolazioni timocitare per la predizione di bersagli</u>	-69-
4.7. <u>Allineamento di miR-150 all'UTR-3' di <i>Candidate Target 1</i></u>	-72-
4.8. <u>Livelli di espressione di miR-150 in diverse linee di T-ALL</u>	-73-
4.9. <u>Studi funzionali dell'over-espressione del miR-150 in una linea di T-ALL</u>	-75-
4.10. <u>Profilo d'espressione dei miR in campioni di T-LBL e linfonodi reattivi umani</u>	-80-
5. <u>DISCUSSIONE</u>	-83-
6. <u>BIBLIOGRAFIA</u>	-93-

RIASSUNTO

La maturazione delle cellule T avviene attraverso una serie di complesse modificazioni fenotipiche e genotipiche ed è guidata da fattori e meccanismi in parte ancora poco compresi. Recenti lavori hanno mostrato che l'espressione di alcuni microRNA (miR) è dinamicamente regolata nel corso dello sviluppo ematopoietico, della risposta immunitaria e della leucemogenesi. Tuttavia, attualmente, è ancora in gran parte sconosciuto il ruolo dei miR nello sviluppo fisiologico delle cellule T, nonché il significato della loro alterata espressione nella leucemogenesi T. Allo scopo di identificare miR potenzialmente coinvolti nella differenziazione dei linfociti T, abbiamo analizzato il profilo d'espressione dei miR in timociti umani a diversi stadi di maturazione: Doppi Positivi (DP; CD4⁺CD8⁺), Singoli Positivi CD4⁺ (SP CD4; CD4⁺CD8⁻) e Singoli Positivi CD8⁺ (SP CD8, CD4⁻CD8⁺). Parallelamente, al fine di approfondire la nostra conoscenza sull'espressione dei miR nelle popolazioni linfoidi T, abbiamo generato delle librerie di small-RNA a partire dall'RNA totale di timociti non frazionati, timociti DP, linfociti maturi CD4⁺ e CD8⁺ del sangue periferico.

L'analisi dei dati degli array ha mostrato che ogni popolazione timica presenta un profilo d'espressione dei miR caratteristico e distinto, che riflette le relazioni fra gli stadi di sviluppo dei precursori T. Nel corso della maturazione dei precursori T dallo stadio DP a quello SP si osserva una generale up-regolazione dell'espressione dei miR.

La generazione delle librerie di small-RNA ci ha permesso di studiare l'espressione sia dei miR noti, che di nuovi candidati miR nelle diverse popolazioni linfoidi T. Al fine di identificare i miR noti e nuovi potenziali fra le sequenze delle librerie di small-RNA, è stata sviluppata una *pipeline* bioinformatica. L'analisi computazionale delle 29.744 sequenze di small-RNA ricavate dalle nostre librerie ha portato all'identificazione di 139 sequenze corrispondenti a miR noti e 98 sequenze di candidati nuovi miR. Mediante un'analisi *bootstrap*, è stato calcolato che, per tutte e 4 le librerie, il *set* di miR maturi sequenziati rappresenta più dell'80% della totalità dei miR che si stima siano espressi nei campioni. L'analisi delle librerie ha confermato la generale up-regolazione dell'espressione dei miR nel corso della maturazione delle cellule T. Comparando i dati degli array e del sequenziamento delle librerie, è stato individuato un gruppo di miR noti che sono consistentemente regolati durante la differenziazione T. Il pattern di espressione nei diversi stadi di sviluppo T di alcuni di questi miR, tra cui miR-150, è

stato validato mediante qRT-PCR. In seguito, abbiamo fatto una ricerca dei target potenziali di questi miR integrando i risultati di 3 diversi software di predizione di target (Miranda, TargetScan and PicTar). Fra i candidati target del gruppo di miR d'interesse sono stati identificati molti geni coinvolti in processi biologici rilevanti, come la regolazione del ciclo cellulare, l'apoptosi, il differenziamento e la tumorigenesi.

Inoltre, abbiamo confrontato i profili di espressione genica delle popolazioni timocitarie con la lista di target predetti computazionalmente per i miR regolati con maggiore *fold-change* nel corso della differenziazione dei timociti da DP a SP. Combinando quest'ultimo approccio alla ricerca bioinformatica integrata di target, abbiamo identificato un gene della famiglia dei recettori Notch (Notch3), definito *Candidate Target 1*, che è noto giocare un ruolo importante nella differenziazione T e nella trasformazione leucemica e che viene predetto, in modo concorde da tre diversi software di predizione di target, come bersaglio di miR-150, uno dei miR maggiormente up-regolati nelle popolazioni timocitarie mature rispetto ai DP. Inoltre, il trascritto di *Candidate Target 1* risulta regolato in modo opposto al miR-150 nel passaggio dei timociti dallo stadio DP a quello di SP CD4. In particolare, abbiamo identificato un'elevata complementarità fra il miR-150 e l'UTR-3' del gene *Candidate Target 1*. Attualmente stiamo lavorando per validare sperimentalmente l'associazione fra miR-150 e *Candidate Target 1*.

In parallelo, abbiamo deciso di studiare gli effetti funzionali dell'over-espressione di miR-150 in linee cellulari di T-ALL. MiR-150 è espresso a livelli molto bassi in tutte le linee cellulari di T-ALL analizzate. Inducendo l'espressione forzata di miR-150 nella linea di T-ALL Jurkat, abbiamo osservato un significativo rallentamento del tasso di proliferazione cellulare associato ad un accumulo delle cellule in fase G2.

Infine, allo scopo di identificare pattern di espressione dei miR associati alla trasformazione neoplastica T, abbiamo confrontato i profili d'espressione dei miR delle sottopopolazioni timocitarie umane con il profilo di un gruppo di linfomi linfoblastici T (T-LBL) pediatrici ed un gruppo di linfonodi reattivi non neoplastici (LN) (forniti dal laboratorio del Dott. Rosolen, Dipartimento di Pediatria, Università di Padova). Il clustering gerarchico dei campioni ha mostrato che i T-LBL hanno un profilo d'espressione dei miR distinto sia da quello delle sottopopolazioni timocitarie, sia da quello dei linfonodi reattivi. Abbiamo inoltre osservato che tutti i 25 miR maggiormente regolati nel passaggio dei timociti dallo stadio DP a SP (a parte miR-128) risultano

espressi in modo differenziale nei T-LBL rispetto ad almeno una delle popolazioni timiche. Nel futuro, ci proponiamo di investigare il significato biologico di alcuni dei miR regolati nella maturazione e la trasformazione neoplastica delle cellule T, ponendo particolare attenzione al ruolo di miR-150 in questi processi.

SUMMARY

Maturation of T cells occurs through a series of steps and is driven by factors and mechanisms which remain poorly defined. Recent work shows that some microRNAs (miRs) are dynamically regulated in their expression during hematopoietic development, inflammation and leukemogenesis. However, little is known regarding the possible role of miRs during the physiologic development of human T cells or their significance in T cell-derived malignancies. In order to identify miRs which could be involved in the differentiation of T lymphocytes, we analyzed the expression profile of miRs in human T cell progenitors at different stages of maturation: Double Positive (DP; CD4⁺CD8⁺), Single Positive CD4⁺ (SP CD4; CD4⁺CD8⁻) and Single Positive CD8⁺ (SP CD8, CD4⁻CD8⁺). In parallel, in order to deepen our knowledge about microRNA expression in T lymphoid cells, smallRNA libraries were generated from total RNA of unsorted thymocytes, DP thymocytes, CD4⁺ and CD8⁺ mature peripheral blood lymphocytes. The array-based analysis showed that each thymic population displays a distinct miR expression profile, which reflected the developmental relationships between maturation stages in thymocytes. A general up-regulation of miR expression was observed in T precursors during the maturation from DP to SP stage.

The generation of small-RNA libraries enabled us to study the expression of both known and new candidate miRs in different T lymphoid populations. In order to identify known and new candidate miRNAs among the short-RNA sequences, a computational pipeline was developed. Computational analysis of the 29,744 small-RNA sequences obtained from our libraries led to the identification of 139 sequences corresponding to known miRs and 98 sequences of new candidate miRs. A bootstrap analysis estimated that the sequence data covered more than 80% of the total content of miRs in the 4 libraries. The analysis of the libraries confirmed the general up-regulation of miR expression during T cell maturation. By comparing the array and sequence data, we identified a group of known miRs which were consistently regulated during normal T cell maturation. The modulation of the expression during T cell differentiation of some of these miR, such as miR-150, was further validated by qRT-PCR. Subsequently, for this subset of miR, we searched for potential target genes by crossing the results of 3 different target prediction softwares (Miranda, TargetScan and PicTar). Among the predicted targets for this group of miRs we found many genes involved in biologically

relevant processes, including cell cycle regulation, apoptosis, differentiation and tumorigenesis.

Moreover, we compared the gene expression profiles from thymic subpopulations with the list of computationally predicted targets for the most significantly regulated miRs between mature SP and DP thymocytes. By combining this approach to the pure bioinformatic approach, we identified a gene of the family of Notch receptors (Notch3), referred to as *Candidate Target 1*, which is known to be important in T cell differentiation and in leukemic transformation and which is concordantly predicted by 3 different target prediction softwares as a target of miR-150, one of the top up-regulated miR during T cell maturation from DP to SP stage. Moreover, *Candidate Target 1* is regulated in thymocytes maturation in the opposite direction compared to miR-150. In particular, we identified a high complementarity between miR-150 and the 3'UTR of *Candidate Target 1*. We are now working in order to validate experimentally the association between miR-150 and *Candidate Target 1*.

In parallel, we decided to study the functional effects of miR-150 over-expression in T-ALL cell lines. MiR-150 is expressed at very low levels in all the T-ALL cell lines tested. Interestingly, when we forced the expression of miR-150 in Jurkat T-ALL cells, we observed a significant reduction of the proliferation rate associated with the accumulation of cells in the G2 phase of cell cycle.

Eventually, in order to identify patterns of miR expression that can be associated with malignant transformation, we compared the miR expression profiles of human thymocyte subpopulations to the profile of a group of T-cell lymphoblastic lymphomas (T-LBL) of childhood and to that of a group of reactive non-neoplastic lymph nodes (LN) (provided by Dr. Rosolen laboratory, Department of Pediatrics, University of Padova). The hierarchical clustering of the samples indicated that TLBLs have a group of miRs which are differentially expressed both in respect to reactive lymph nodes and to thymic healthy subsets. Interestingly, we observed that all the top 25 differentially regulated miRs during T differentiation from DP to SP stage (excluding miR-128) show a significant differential expression in T-LBL compared to at least one of the thymus populations. In the future, we plan to investigate the biological significance of some of the miRs regulated in the maturation and neoplastic transformation of T cells, playing particular attention to miR-150 role in these processes.

1. INTRODUZIONE

1.1 Linfopoiesi T nel timo

I linfociti T sono importanti mediatori cellulari della risposta immunitaria adattativa. Essi mediano la cosiddetta immunità cellulare, svolgendo una duplice funzione effettrice e regolatoria. Infatti, da un lato i linfociti T provvedono direttamente al riconoscimento e all'eliminazione di cellule infettate da patogeni intracellulari (quali virus ed alcuni tipi di batteri) e tumorali; dall'altro essi regolano la funzione dei linfociti B, secernono citochine, reclutano cellule coinvolte nella fase effettrice della risposta immunitaria e ne modulano la funzione. I linfociti T, a differenza dei linfociti B, non riconoscono antigeni isolati e solubili, ma soltanto antigeni peptidici presentati sulla superficie delle cellule dell'ospite in stretta associazione con le molecole del complesso MHC (Complesso Maggiore d'Istocompatibilità). Il recettore per l'antigene delle cellule T è costituito dal complesso molecolare TCR/CD3 composto da una parte variabile (l'etero-dimero noto come T cell receptor o TCR) e una parte monomorfa (costituita dal CD3 e dalla catena ζ). I geni che codificano per le due catene polipeptidiche che compongono il TCR sono generati mediante ricombinazione somatica di segmenti di DNA durante la maturazione cellulare. La ricombinazione somatica genera milioni di combinazioni genetiche codificanti per recettori diversi, dando così luogo ad un repertorio linfocitario altamente diversificato e distribuito clonalmente, cioè costituito da molti cloni cellulari con specificità antigenica diversa.

Fra le cellule appartenenti al lineage T sono comprese le cellule T-helper $CD4^+$ e le cellule T-citotossiche $CD8^+$ (entrambe facenti parte dei linfociti T $\alpha\beta$), i linfociti T $\gamma\delta$, le cellule T regolatorie $FoxP3^+$ e le cellule NK (Natural Killer) [1]. Qui di seguito, mi concentrerò in particolare sulla descrizione dello sviluppo dei linfociti T $\alpha\beta$.

I linfociti T, sia nell'uomo che nel topo, derivano da precursori ematopoietici multipotenti, detti TSP (*Thymus Seeding Progenitors*), che originano nel midollo osseo e nel fegato fetale per poi migrare, attraverso i vasi sanguigni, fino al timo. Il timo è un organo linfoide primario che non contiene in sé cellule staminali in grado di auto-rigenerarsi, ma che fornisce un microambiente specializzato in grado di indurre la maturazione dei progenitori midollari in cellule T mature [2]. Le cellule T in via di sviluppo nel timo sono dette timociti. Gli stadi dello sviluppo T nel timo hanno luogo in diverse nicchie stromali (fig. 1.1): infatti i timociti nel corso della loro maturazione

migrano attraverso il tessuto timico, guidati da fattori solubili e stimoli provenienti da diverse cellule stromali presenti nel microambiente, quali le cellule epiteliali timiche (TEC), le cellule dendritiche (DC) ed i macrofagi. Le TEC formano un reticolo di prolungamenti citoplasmatici all'interno della zona corticale del timo, le cellule DC sono presenti a livello della giunzione cortico-midollare e nella zona midollare del timo, mentre i macrofagi sono presenti soprattutto a livello della midollare. I TSP, una volta entrati nel timo a livello della giunzione cortico-midollare, migrano verso la zona corticale esterna sottocapsulare del timo; da qui, i precursori T attraversano la corticale, dove si verifica la maggior parte dei successivi eventi maturativi, ed infine penetrano nella zona midollare del timo, dove terminano il loro sviluppo, prima di fuoriuscire nel torrente circolatorio e migrare in periferia, dove svolgeranno le loro funzioni effettrici [3,4]. In questo modo, il passaggio dei timociti attraverso questo ordinamento tissutale permette che si verifichino le interazioni fisiche tra le cellule, necessarie per la maturazione dei linfociti T.

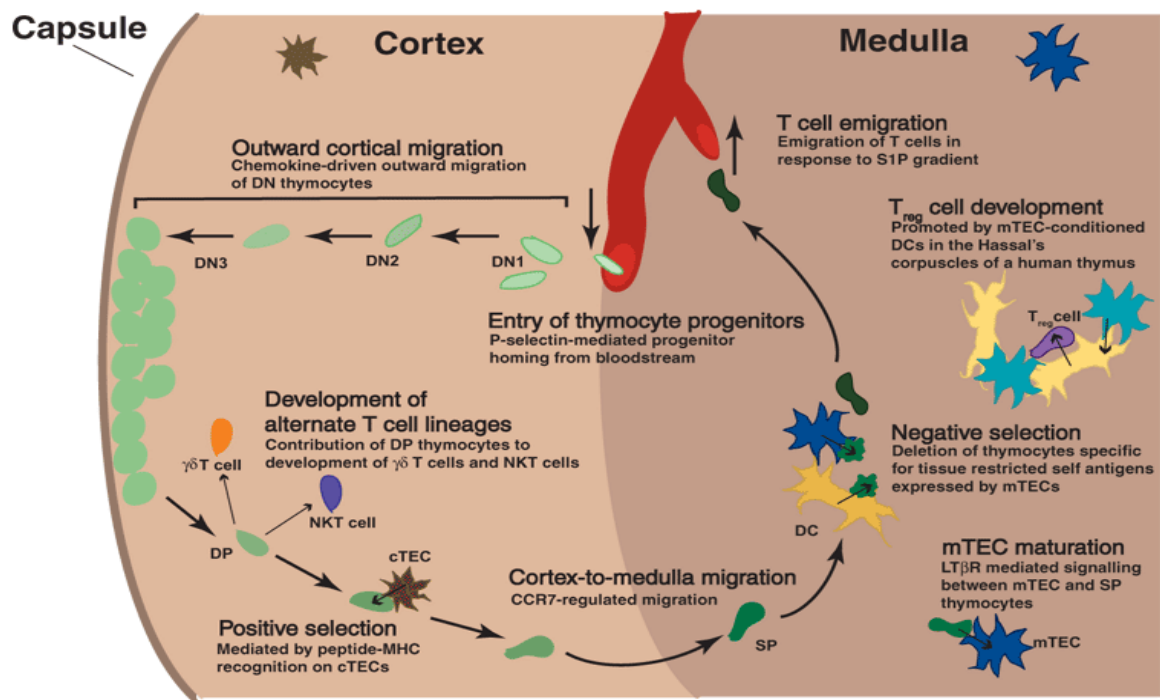


Figura 1.1. Lo sviluppo T nel microambiente timico. Il timo fornisce ai precursori T una varietà di nicchie specializzate che supportano e guidano la differenziazione e la selezione T. In questa figura sono illustrate le diverse fasi dello sviluppo dei precursori T nelle varie nicchie del microambiente timico [3].

In seguito alla colonizzazione del timo, la maturazione dei linfociti T procede attraverso una serie di tappe sequenziali che prevedono il riarrangiamento dei loci genici delle catene del TCR (TCRD, TCRG, TCRB, TCRA) ed eventi di selezione positiva e negativa che portano infine all'espressione di recettori TCR $\alpha\beta$ -TCR $\gamma\delta$ maturi ed in grado di riconoscere peptidi "non-self" presentati dal complesso maggiore di istocompatibilità (MHC). Questo processo alterna fasi di selezione a fasi di intensa proliferazione e passa attraverso una serie discreta di stadi fenotipici che sono distinti per l'espressione sulla superficie cellulare di specifici marcatori di differenziazione, fra cui in particolare i co-recettori CD4 e CD8. Questi marcatori vengono comunemente utilizzati per distinguere i diversi stadi di sviluppo timico.

Sia nell'uomo che nel topo, i timociti vengono suddivisi per stadio di maturazione in DN (Double Negative; CD4⁻/CD8⁻), ISP (Immature Single Positive; CD4⁺/CD3⁻ nell'uomo o CD8⁺/CD3⁻ nel topo), DP (Double Positive; CD4⁺/CD8⁺) e SP CD4 (Single Positive CD4; CD4⁺/CD8⁻/CD3⁺) o SP CD8 (Single Positive CD8; CD4⁻/CD8⁺/CD3⁺).

I DN sono i precursori T più immaturi presenti nel timo: essi sono una popolazione rara ed altamente eterogenea, in parte non ancora determinata in senso T, che deriva direttamente dai TSP e non esprime in superficie né un TCR maturo, né i co-recettori CD4 e CD8. Questa fase di maturazione è anche nota come stadio **pro-T**. Nel topo, la popolazione DN può essere a sua volta suddivisa in 4 sotto-stadi sulla base dell'espressione differenziale degli antigeni CD44 e CD25: DN1 (CD44⁺/CD25⁻), DN2 (CD44⁺/CD25⁺), DN3 (CD44⁻/CD25⁺) e DN4 (CD44⁻/CD25⁺). Nell'uomo invece all'interno dei DN sono state individuate solo 3 diverse sottopopolazioni sulla base dell'espressione di altri due antigeni, il CD34 ed il CD1a: DN1 (CD34⁺/CD1a⁻), DN2 (CD34⁺/CD1a⁺) e DN3 (CD34⁻/CD1a⁺) [5,6,7]. Lo stadio DN4 nell'uomo è assimilato al successivo stadio di ISP. I DN1, anche detti ETP (*Early Thymic Progenitors*), sono localizzati nel seno sotto-capsulare e nella corticale esterna del timo. Si tratta di cellule non ancora determinate in senso T, che mantengono la capacità di differenziarsi in cellule NK, DC, mieloidi e, in rari casi, anche in cellule B. In questo stadio, cominciano ad essere espresse le proteine RAG1 e RAG2 ed inizia la ricombinazione dei segmenti genici V (Variable), D (Diversity) e J (Joining) del locus TCRD. Nel successivo stadio DN2 si ha il riarrangiamento dei geni del TCRG. In questa fase i timociti sono altamente proliferanti e vanno incontro ad una rilevante espansione cellulare. I progenitori in cui il riarrangiamento delle catene δ e γ del TCR è funzionale in genere (anche se non in tutti i casi) rimangono DN e diventano linfociti $\gamma\delta$. I linfociti $\gamma\delta$ sono

una minoranza nell'uomo (0,5-5% dei linfociti T periferici). Essi sono CD3⁺ e generalmente CD4⁻ e CD8⁻. Negli animali questo tipo di linfociti T sono soprattutto localizzati negli epiteli e non ricircolano: si pensa che essi abbiano un ruolo nella difesa precoce contro un numero limitato di microorganismi incontrati a livello delle barriere epiteliali. In fase DN3 la proliferazione dei progenitori T termina e comincia la ricombinazione dei geni del TCR β , che si protrarrà fino allo stadio di ISP. I timociti che riarrangiano con successo almeno uno degli alleli del locus TCRB cominciano ad esprimere sulla superficie cellulare il complesso pre-TCR, cioè una forma immatura del TCR composta dalla catena β del TCR in associazione alla catena invariabile pre-T α , al CD3 ed alla catena ζ . I timociti esprimenti il pre-TCR sulla loro superficie cellulare sono anche detti **pre-T**. Segnali emanati in modo ligando-indipendente dai pre-TCR funzionali promuovono la sopravvivenza e la proliferazione dei timociti, l'inibizione di ulteriori riarrangiamenti dei loci TCRB, l'inizio dei riarrangiamenti dei loci TCRA ed il differenziamento delle cellule in doppie positive DP (CD4⁺ CD8⁺).

I timociti DP, localizzati nella zona corticale interna del timo, co-esprimono sulla loro superficie i co-recettori CD4 e CD8 e sono ormai irreversibilmente determinati in senso $\alpha\beta$. Nella popolazione DP il riarrangiamento produttivo del locus TCRA porta all'espressione di un complesso TCR $\alpha\beta$ maturo sulla superficie cellulare in associazione al CD3 [8].

In seguito all'espressione del recettore TCR maturo sulla superficie, le cellule DP sono sottoposte ad una selezione positiva per la capacità di riconoscere efficacemente gli antigeni presentati nel contesto delle molecole MHC-self e ad una selezione negativa per la reattività contro antigeni self dell'ospite. In questo modo, viene fatta la selezione del repertorio immunologico T e vengono eliminate le cellule auto-reattive, che, se immesse in circolo, potrebbero scatenare reazioni autoimmuni nell'organismo [9,10,11]. E' stato stimato che circa il 95% dei timociti va incontro ad apoptosi durante le diverse fasi di selezione nel timo. Le cellule DP che superano con successo questi processi selettivi proseguono migrando nella zona midollare del timo e maturano in timociti SP CD4 o CD8, che successivamente escono dal timo per migrare alla periferia come linfociti T naive. I timociti SP CD4 sono gli immediati precursori dei linfociti T periferici CD4⁺ ad attività helper; mentre i timociti SP CD8 sono i precursori dei linfociti T periferici CD8⁺ ad attività citotossica. I diversi stadi di sviluppo T nell'uomo e nel topo sono illustrati nella figura 1.2. Nel timo umano i DP costituiscono circa il 60-

80% dei timociti totali, i DN rappresentano solo l'1-5%, mentre i SP CD4 ed i SP CD8 costituiscono rispettivamente il 10-15% ed il 5-10% [1].

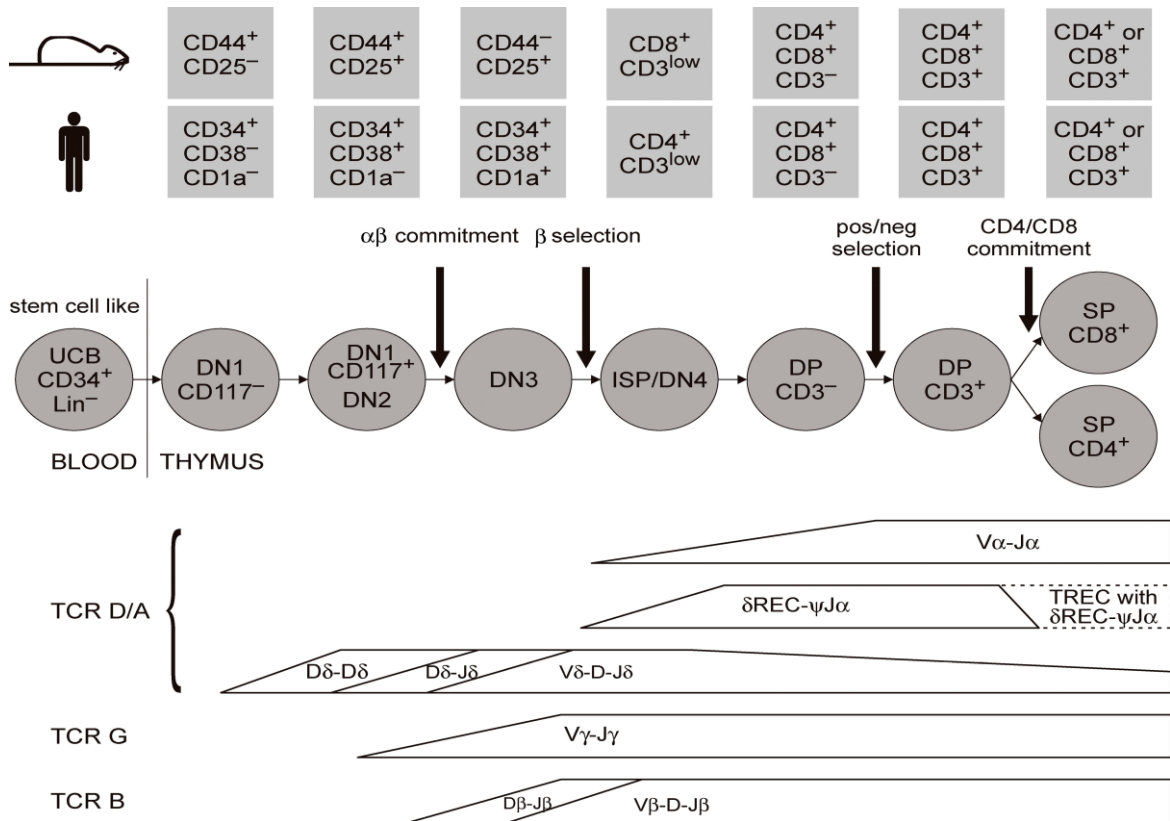


Figura 1.2. Stadi dello sviluppo delle cellule T umane e murine nel timo. Sono illustrati i diversi stadi di maturazione dei timociti con i marcatori di differenziazione che li caratterizzano nell'uomo e nel topo [8].

1.2 Leucemia linfoblastica acuta a cellule T (T-ALL)

La leucemia linfoblastica acuta a cellule T (T-ALL) è un'emopatia maligna dei timociti caratterizzata da un alto numero di linfoblasti T circolanti e nel midollo. Rappresenta il 10-15% dei casi di leucemia linfoblastica acuta in età pediatrica ed il 25% dei casi in età adulta ed è solitamente associata a prognosi più sfavorevole rispetto alle più comuni leucemie linfoblastiche acute a cellule B (B-ALL). Spesso questa forma di leucemia acuta si presenta clinicamente con conte elevate di globuli bianchi nel sangue, adenopatia e coinvolgimento del sistema nervoso centrale.

L'insorgenza della T-ALL è legata ad eventi di trasformazione che avvengono in punti cruciali del processo di differenziamento intratimico delle cellule T; è stata infatti

evidenziata una correlazione tra l'espressione di determinati oncogeni e l'arresto in definiti stadi dello sviluppo normale dei timociti.

Contrariamente alle leucemie mieloidi e alle leucemie linfoblastiche acute a cellule B (B-ALL), le T-ALL presentano un cariotipo normale in oltre il 50% dei casi. Nel 35% dei casi le traslocazioni cromosomiche riguardano loci del TCR (T-cell receptor) e comportano la giustapposizione dei forti elementi promotoriali ed enhancer dei geni del TCR a geni di fattori di trascrizione [12]. Tra i geni principalmente interessati da queste alterazioni ricordiamo HOX11, HOXA-HOXD, HOX11L2, TAL1, TAL2, LYL1, BHLHB1, LMO1 e LMO2. Alcuni di questi geni non sono fisiologicamente espressi nel timo, ma vengono attivati ectopicamente nelle cellule leucemiche; altri sono normalmente espressi dai timociti non-trasformati. Questi ultimi geni sono spesso regolatori essenziali dell'ontogenesi T, implicati in diverse tappe del differenziamento intratimico dei linfociti T; quindi la loro espressione aberrante comporta anomalie nel controllo del ciclo cellulare, della proliferazione e del differenziamento dei timociti. Altre anomalie citogenetiche riscontrate nelle T-ALL riguardano la generazione di proteine chimeriche con proprietà oncogeniche, delezioni cromosomali che comportano la perdita di geni oncosoppressori, duplicazioni geniche e mutazioni attivanti (vedi tabella nella figura 1.3A). In aggiunta, nella T-ALL si osserva spesso l'attivazione aberrante di alcuni fattori di trascrizione in assenza di evidenti alterazioni cromosomiche o genetiche [13,14].

In questi anni sta emergendo sempre più, in particolare, il ruolo del *pathway* di Notch nel processo della leucemogenesi T, in quanto la sua attivazione costitutiva nei progenitori ematopoietici, pur inibendo il corretto sviluppo sia dei linfociti T che B, porta esclusivamente a neoplasie a cellule T. La traslocazione t(7;9)(q34;q34.3), che pone una forma tronca del gene Notch1 sotto il controllo del locus TCR β si riscontra in meno dell'1% dei casi di T-ALL. Tuttavia, in più del 50% dei pazienti affetti da T-ALL è presente una mutazione attivante nel gene Notch1 o a carico di proteine coinvolte nella regolazione della stabilità di questo recettore (come l'ubiquitina ligasi Fbw7) [15]. Oltre a Notch1, altri membri della famiglia di Notch sono stati implicati nel processo di leucemogenesi T. In particolare, nonostante non siano attualmente note mutazioni attivanti, Notch3 risulta over-espresso nel 90% delle T-ALL [16].

I recettori Notch (nei mammiferi esistono 4 diversi membri della famiglia Notch: Notch1 e Notch2, Notch3 e Notch4) controllano la progressione nel ciclo cellulare e

dell'apoptosi in diversi tipi cellulari e sono regolatori essenziali del differenziamento dei progenitori ematopoietici della linea T [17,18].

A

Cytogenetic abnormality	Approximate frequency	Involved oncogenes or fusion genes
Translocations involving TCR genes on chromosomes 7q34 (TCRB and TCRG) and 14q11 (TCRA and TCRD)	35%	HOX11, HOX11L2, TAL1, TAL2, LYL1, BHLHB1, LMO1, LMO2, LCK, NOTCH1, cyclin D2
Abnormal expression	9–30%	SIL–TAL1 fusion
Fusion gene formation	10%	CALM–AF10 fusion
	4–8%	MLL fusions
	6%	ABL1 fusions
	Rare	NUP98 fusions
Chromosomal deletions of 9p21 and 6q	Up to 65% (9p21)	P15, P16
	20–30% (6q)	Unknown
Activating gene mutations	50–60%	NOTCH1, FLT3, NRAS
Gene duplications	~30% in T-ALL cell lines	MYB

ABL1, Abelson 1; BHLHB1, basic helix–loop–helix B1; CALM, clathrin assembly lymphoid–myeloid leukaemia; FLT3, FMS-related tyrosine kinase 3; HOX11, homeobox-11; LMO, LIM-only; LYL1, lymphoblastic leukaemia-derived sequence 1; MLL, mixed-lineage leukaemia; SIL, SCL-Interrupting locus; TAL, T-cell acute lymphocytic leukaemia; T-ALL, T-cell acute lymphoblastic leukaemia; TCR, T-cell receptor.

B

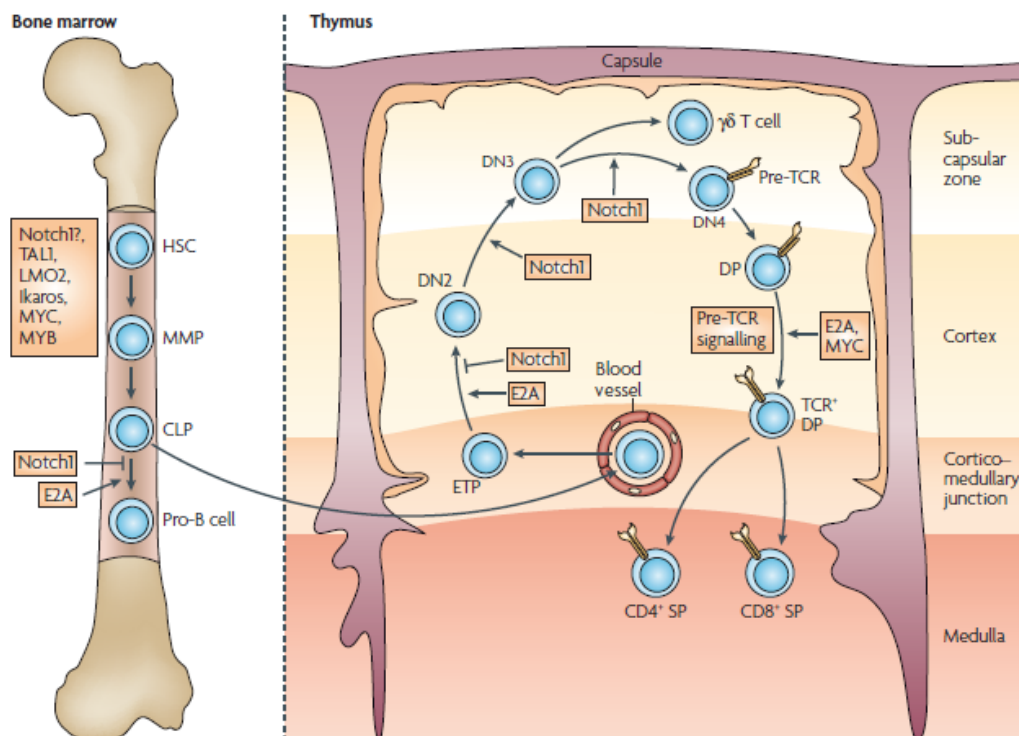


Figura 1.3. A) Aberrazioni citogenetiche comuni nella T-ALL. La tabella riporta le principali anomalie citogenetiche riscontrate nelle T-ALL. **B) Stadi dell’ematopoiesi e dello sviluppo delle cellule T.** Rappresentazione degli stadi dello sviluppo delle cellule T partendo dalla cellula staminale ematopoietica fino ai linfociti T maturi. Sono indicati gli oncogeni che sono noti essere associati con lo sviluppo di T-ALL. (HSC: haematopoietic stem cells, MPPs: multipotent progenitors, CLPs: common lymphoid progenitors, ETPs: early T-cell-lineage progenitors, DN: double negative, DP: double positive, SP: single positive) [13].

La via di trasduzione di Notch è coinvolta infatti in diverse fasi del differenziamento dei precursori ematopoietici in linfociti T maturi: Notch1 controlla il differenziamento delle cellule linfoidei progenitrici comuni in cellule T piuttosto che cellule B, agisce allo stadio di biforcazione $\alpha\beta$ - $\gamma\delta$ favorendo lo sviluppo del TCR $\alpha\beta$ rispetto al TCR $\gamma\delta$, è coinvolto nel passaggio da cellule doppie positive a cellule T CD4⁺ helper o CD8⁺ citotossiche ed infine controlla la suscettibilità delle cellule T all'apoptosi. Notch3 regola l'espressione del pre-TCR α nello stadio precedente a quello di timociti DP ed ha un ruolo importante nel passaggio dei precursori T dallo stadio DN3/ISP allo stadio DP. L'over-espressione della forma attiva sia di Notch1, che di Notch3 porta allo sviluppo di una forma aggressiva di T-ALL nei topi; mentre la delezione di Notch1 nelle HSC causa atrofia timica associata ad una totale inibizione della differenziazione T [16,19-24]. Questi dati sottolineano l'importante relazione tra il ruolo di questo recettore nello sviluppo delle cellule T e la patogenesi delle T-ALL.

1.3 Linfoma linfoblastico T (T-LBL)

Come la T-ALL, anche i linfomi linfoblastici T (T-LBL) originano dai precursori T del timo. I T-LBL rappresentano circa un terzo dei Linfomi Non-Hodgkin dell'età pediatrica. Il termine "*linfoma linfoblastico*" è stato coniato da Barcos e Luke per la similarità riscontrata tra le cellule tumorali di linfoma con i blasti della leucemia linfoblastica acuta (ALL). Di fatto, sono molte le analogie tra T-ALL e T-LBL: oltre all'origine comune, nonché alla morfologia e all'immunofenotipo indistinguibili, le due malattie presentano anche le stesse caratteristiche genetiche (traslocazioni cromosomiche di fattori trascrizionali). La principale discriminante clinica tra T-LBL e T-ALL riguarda la sede di presentazione della malattia, che è classicamente il mediastino per i T-LBL ed il midollo osseo per la T-ALL. Il T-LBL si presenta spesso con massa mediastinica, talora accompagnata da adenopatia subdiaframmatica, scarso coinvolgimento del midollo osseo e, raramente, del sistema nervoso centrale. Operativamente, la percentuale di cellule tumorali nel midollo osseo distingue i T-LBL (< 25%) dalle T-ALL (\geq 25%). Rispetto alla T-ALL, il T-LBL si manifesta in pazienti di età media maggiore, con un picco di incidenza nella seconda decade di vita; inoltre mostra una prevalenza nei maschi rispetto alle femmine (rapporto 2:1). Le elevate analogie cliniche suggeriscono che il T-LBL e la T-ALL possano rappresentare diverse

presentazioni (o evoluzioni) di un'unica entità biologica e che possano beneficiare della stessa terapia [13,25].

1.4 Biosintesi e biologia dei microRNA

I microRNA (miRs) sono una classe di piccoli RNA non codificanti (19-24 nt) in grado di modulare l'espressione genica a livello post-trascrizionale interferendo con la traduzione in proteina e/o favorendo la degradazione dell'mRNA dei loro bersagli. I miR sono generati mediante un elaborato meccanismo di espressione e processamento che ha inizio nel nucleo e termina nel citoplasma della cellula. Studi recenti hanno mostrato come la biogenesi dei miR sia molto più complessa, regolata e miR-specifica di quanto in precedenza si riteneva. Infatti, la trascrizione dei geni dei miR, come d'altro canto quella dei geni codificanti proteine, è controllata da elementi enhancer e promotoriali, dal legame di fattori di trascrizione e dalle modificazioni epigenetiche della struttura cromatinica. Inoltre, ulteriori livelli di regolazione dell'espressione dei miR sono ottenuti mediante processamento post-trascrizionale delle forme precursori, modificazioni/addizioni nucleotidiche (RNA *editing*) che possono alterare la stabilità e la specificità di *targeting* dei miR, l'esportazione selettiva dal nucleo al citoplasma e la localizzazione sub-cellulare dei microRNA stessi (fig. 1.4) [26,27].

I geni dei miR possono essere localizzati sia in regioni intergeniche, che negli introni (o più raramente in sequenze esoniche e in UTR) di geni codificanti. E' stato osservato che i miR derivanti da introni (circa il 25% dei miR noti) di geni codificanti sono trascritti e processati da un precursore comune insieme all'mRNA dei rispettivi geni ospitanti. In molti casi questi miR ed i loro geni ospitanti sono co-regolati nella loro espressione e sembrano svolgere funzioni correlate nella cellula [28]. Al contrario, i miR intergenici sono trascritti da un loro promotore indipendente. Molti miR (circa il 40% dei miR noti) sono raggruppati in *cluster* di due, tre o, in alcuni casi, anche più miR, che danno origine a trascritti primari policistronici. I miR che fanno parte di uno stesso *cluster* hanno spesso sequenze simili fra loro. Questa organizzazione genomica dei miR sembra indicare l'esistenza di meccanismi interni che ne determinano l'azione coordinata.

I miR sono trascritti ad opera delle RNA-polimerasi II e III in lunghi trascritti primari (100-1000 bp) detti **pri-miR**. Questi trascritti primari generalmente sono dotati di CAP al 5' e poly-A al 3', possono codificare per le sequenze di uno o più miR e sono

caratterizzati da una struttura a forcina composta da un braccio a doppio filamento di circa 33 bp che termina in un'ansa a singolo filamento.

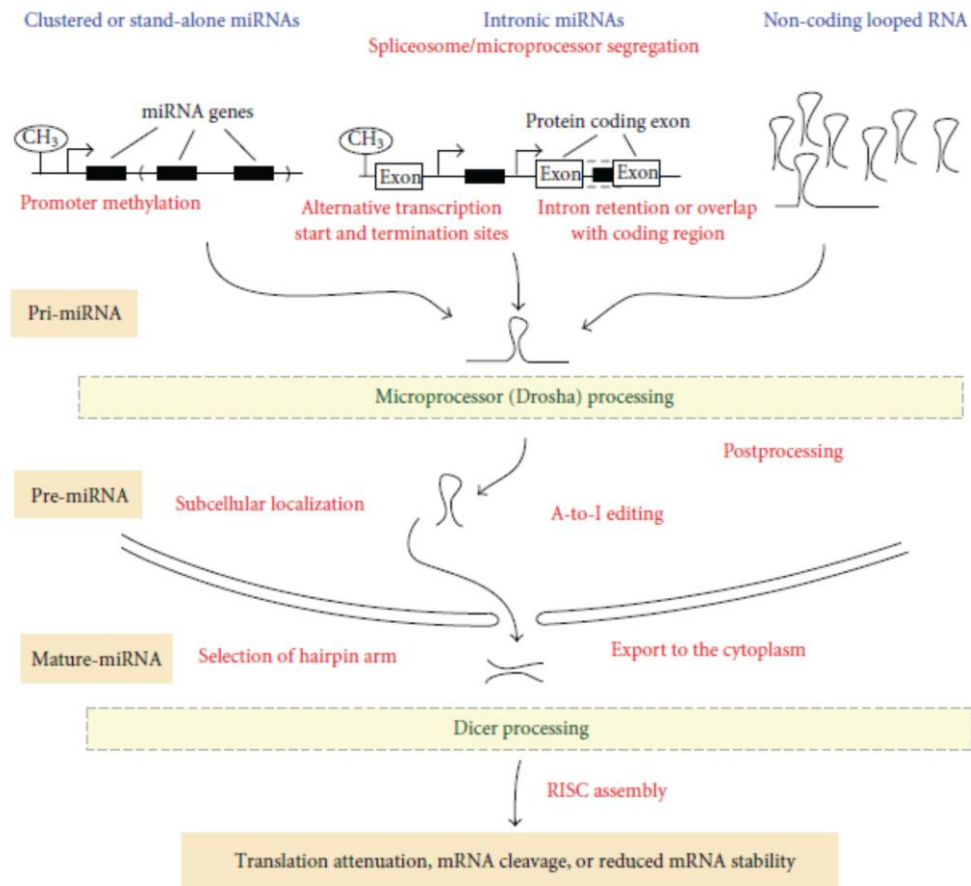


Figura 1.4. Regolazione della biogenesi dei miR. Diversi livelli di regolazione controllano la biogenesi dei miR, fra cui l'attivazione trascrizionale, lo splicing alternativo, il riconoscimento da parte di Drosha, il processamento post-trascrizionale, l'*editing* dell'RNA, la localizzazione sub-cellulare, l'esportazione nucleare, la selezione del *guide strand* [27].

Il *duplex* della struttura a forcina e le sequenze immediatamente fiancheggianti sono fondamentali per il riconoscimento ed il clivaggio del pri-miR da parte del microprocessore, un complesso multi-proteico di cui fanno parte l'RNasi di tipo III Drosha ed il suo co-fattore DGCR8. Il taglio endonucleasico mediato dal complesso Drosha-DGCR8 avviene nel nucleo durante la trascrizione dell'RNA.

Il clivaggio del microprocessore genera una struttura a forcina di circa 70 bp, detta **pre-miR**, che viene successivamente esportata dal nucleo al citoplasma cellulare tramite il

complesso Exportina5/ Ran-GTP. Nel caso di alcuni miR intronici (detti mirtroni), lo stesso processo di splicing permette di by-passare il taglio di Drosha, generando direttamente la forma pre-miR. Una volta arrivato nel citoplasma, il pre-miR viene processato da un secondo complesso proteico contenente l'RNasi III Dicer, che ne cliva la porzione ad ansa, generando un *miRNA-miRNA* duplex* di circa 22 nt. Questo *duplex* di RNA contiene il miR maturo (miR) e il suo filamento antisenso (miR*). Solo uno dei due filamenti del *miR-duplex*, detto “*guide strand*” o miR, è incorporato nel complesso RISC (*RNA-Induced Silencing Complex*) e ne guida il legame specifico agli mRNA bersaglio. In genere, è il filamento del *duplex* più termodinamicamente instabile a livello dell'estremità 5' ad essere caricato preferenzialmente nel RISC ed a fungere da *guide strand*, mentre l'altro filamento viene degradato. Il filamento che non si associa al RISC è chiamato “*passenger strand*” ed è indicato con il simbolo miRNA*. In alcuni casi è stato visto che entrambi i filamenti del *miRNA-miRNA* duplex* possono legarsi al RISC, mediando il silenziamento di bersagli diversi. La via canonica della biogenesi dei miR è descritta nella figura 1.5.

Il taglio di Dicer è accoppiato al caricamento del miR maturo a singolo filamento nel complesso RISC (*RNA-Induced Silencing Complex*). Quest'ultimo è il complesso citoplasmatico effettore dell'attività di silenziamento genico propria dei miR. Infatti, i miR, caricati nel RISC, reprimono la traduzione di specifici RNA messaggeri, appaiandosi con complementarietà imperfetta a livello del UTR-3' del trascritto bersaglio. Per il legame fra miR e target è di particolare importanza l'appaiamento perfetto dell'UTR-3' dell'mRNA bersaglio con l'ettamero in 5' del miR (nt 2-8) detto “*seed region*” [26,29,30]. Tuttavia, si è visto che in molti casi la presenza di un *seed match* minimo (6 nt) a livello dell'estremità in 5' del miR non è sufficiente a provocare una rilevabile down-regolazione dell'espressione dell'mRNA target e che la capacità effettiva di *targeting* dei bersagli da parte dei miR è influenzata anche da altri determinanti. Alcune caratteristiche del contesto dell'UTR-3' dell'mRNA che contribuiscono ad aumentare l'efficacia e la specificità di *targeting* dei miR sono:

- La presenza di complementarietà, oltre che con l'estremità in 5' del miR (dove si trova il *seed*), anche con l'estremità 3', in particolare a livello dei nt 12-17 del miR .
- La presenza di una A e di una A/U rispettivamente in posizione opposta alla base 1 e 9 del miR.
- Un'aumentata conservazione di sequenza in generale nelle 50 bp circa a valle e a monte del *seed match*.

- Un'elevato contenuto di AU al 3' del *seed match*.
- La prossimità del sito di legame a siti riconosciuti da altri miR co-espressi, che possono reprimere l'espressione dello stesso target in modo coordinato.
- Posizione del *seed match* lontana dal centro dell'UTR-3' e distante almeno 15 nt dal codone di stop [31,32].

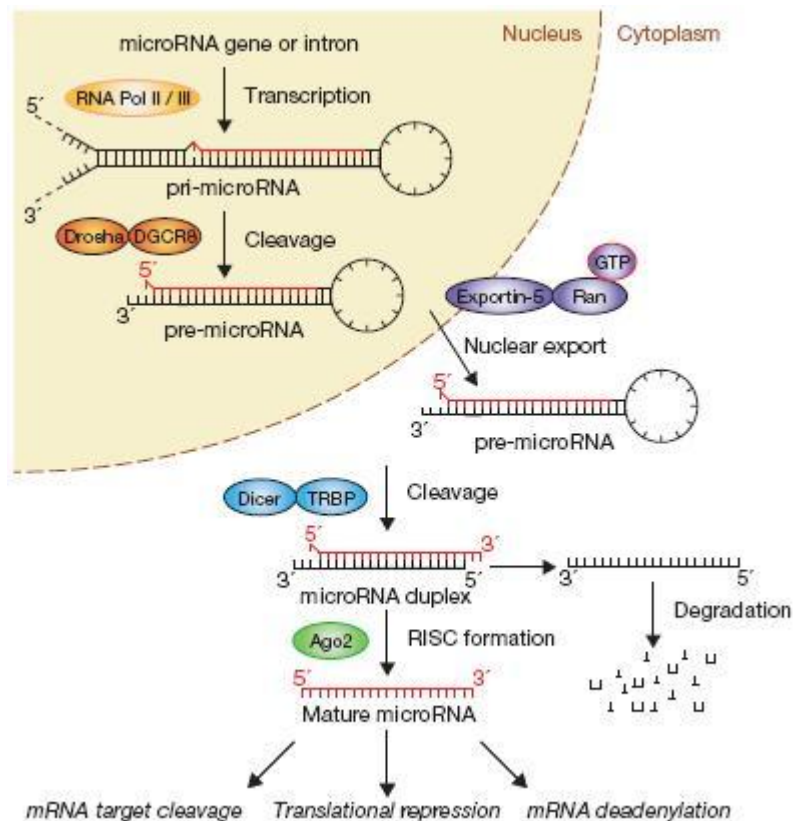


Figura 1.5. Il pathway canonico della biogenesi dei miR. Il meccanismo di processamento canonico dei miR include la produzione di un miR trascritto primario (pri-miR) da parte della RNA polimerasi II o III, seguita dal clivaggio da parte del complesso microprocessore Droscha-DGCR8 nel nucleo. La risultante struttura a forcina di circa 70 nt (pre-miR) è esportata nel citoplasma, dove è clivata dal complesso Dicer-TRBP, dando origine ad un *duplex miR:miR** di circa 20 nt. Il filamento funzionale del miR maturo è caricato quindi nel complesso RISC, guidandone l'azione di silenziamento post-trascrizionale nei confronti di specifici mRNA bersaglio [29].

Inoltre, recenti lavori hanno rivelato che i miR possono associarsi con complementarità imperfetta anche a siti diversi dall'UTR-3', localizzati nelle sequenze codificanti o nell'UTR-5' degli mRNA target [30]. In particolare, è stata individuata una classe di miR che interagisce con il promotore e l'UTR-5' dei propri target, stimolandone, piuttosto che inibendone la trascrizione genica con un fenomeno definito *RNA*

activation [33]. Dati emergenti sostengono l'ipotesi che alcuni miR oscillino tra la repressione e l'attivazione dell'espressione dei loro target in modo ciclo-dipendente: l'attivazione trascrizionale si verificherebbe solo in cellule quiescenti, bloccate in fase G0 del ciclo cellulare; mentre nei periodi di normale proliferazione l'attività repressoria sarebbe quella preferenziale [34].

In questi anni sono stati proposti diversi meccanismi per spiegare l'attività di silenziamento post-trascrizionale dei miR. Si è visto che essi agiscono sia bloccando la traduzione in proteina del trascritto, sia direttamente destabilizzando o portando alla degradazione dell'mRNA dei loro bersagli. Infatti, il complesso RISC-miR può inibire l'inizio della traduzione del target grazie alla multipla attività della proteina del RISC Argonaute-2 (Ago2), che da un lato compete con il fattore d'iniziazione eIF4E per il legame al CAP in 5' dell'mRNA e dall'altro interagisce con la proteina eIF6, inibendone l'associazione con le subunità ribosomiali 40S e 60S. Inoltre, mediante un meccanismo ancora sconosciuto, la presenza del RISC rallenta l'elongazione proteica e porta al distacco del ribosoma dal polipeptide nascente, causandone la successiva degradazione. La destabilizzazione dei trascritti target mediata dai miR comprende la de-adenilazione, il taglio endonucleasico e/o il *decapping* seguito da degradazione in senso 5'-3' dell'mRNA. Gli mRNA associati al complesso miR-RISC possono localizzarsi in compartimenti sub-cellulari chiamati P-body, in cui gli mRNA non tradotti si accumulano per essere successivamente riciclati o degradati.

Da cosa dipenda la scelta fra i diversi possibili meccanismi di silenziamento non è ancora stato ben compreso. Si pensa che l'azione dei miR dipenda in parte dal tipo di co-fattori che si associano al complesso RISC al momento della sua formazione, in parte dalle caratteristiche dell'appaiamento miR-target. Tuttavia, è stato visto che quando la complementarità con l'mRNA bersaglio è quasi perfetta, i miR tendono ad indurre il taglio endonucleasico mediato da Ago2. Questo tipo di *targeting* è predominante per i miR delle piante, mentre è più raro per i miR animali, che generalmente agiscono preferenzialmente inibendo la traduzione in proteina dei loro bersagli [29,30].

Recenti esperimenti di proteomica condotti su cellule di mammifero hanno dimostrato che l'up-regolazione di singoli miR può reprimere l'espressione di centinaia di target, portando in genere alla down-regolazione della proteina accompagnata da una lieve riduzione dei livelli del relativo trascritto [35,36]. Quindi, un singolo miR può regolare diversi mRNA bersaglio ed un singolo gene può essere a sua volta controllato nella sua

espressione da molti diversi miR. Al momento più di 700 miR sono stati identificati nell'uomo (miRBase Registry, release 13.0): alcuni di essi mostrano un *pattern* di espressione ubiquitario e sono altamente conservati in diverse specie, suggerendo un loro ruolo in processi coinvolti nell'omeostasi e nella regolazione della crescita cellulare; altri sembrano avere un'espressione ristretta a singoli organi o stadi maturativi, facendo ipotizzare un loro ruolo importante nella determinazione dello stadio differenziativo cellulare o della tessuto-specificità. Di conseguenza, la scoperta dei miR ha rivelato un nuovo livello di complessità e di regolazione che deve essere considerato nello studio di ogni processo biologico. A questo proposito, diversi lavori hanno mostrato il ruolo chiave svolto dai miR in diversi processi cellulari fondamentali sia fisiologici, come la differenziazione cellulare, la proliferazione e l'apoptosi, sia patologici, come la tumorigenesi [37].

1.5 Ruolo dei microRNA nella differenziazione e nella fisiologia dei linfociti

Studi come quello di Chen [38] e di Ramkissoon [39] hanno dimostrato l'esistenza di miR espressi preferenzialmente in cellule del sistema ematopoietico, quali miR-142, miR-146, miR-150, miR-155, miR-181 e miR-223. Inoltre, è stato osservato che l'espressione di questi ed altri miR varia nel corso dell'ematopoiesi e di conseguenza si è ipotizzato che questa modulazione potesse avere un significato funzionale. In questi anni diversi gruppi di ricerca hanno dimostrato sperimentalmente il ruolo fondamentale che i miR giocano nell'immunità sia innata che acquisita, controllando la differenziazione di diversi tipi di cellule ematopoietiche e regolandone la funzione immunitaria (fig. 1.6). E' interessante osservare che, nella maggior parte dei casi riportati, i miR non agiscono inducendo la differenziazione cellulare, bensì prevenendo la de-differenziazione o mantenendo sottopopolazioni cellulari in un determinato stato differenziativo [40]. Inoltre, molti miR hanno effetti differenti in cellule ed in contesti spazio-temporali diversi.

Dato che in questa tesi si tratterà dell'espressione dei miR in cellule T, si è ritenuto opportuno riassumere qui di seguito le principali scoperte riguardanti la funzione dei miR nella fisiologia dei linfociti (con particolare interesse per i linfociti T), tralasciando in questa sede i dati riguardanti le altre popolazioni ematopoietiche.

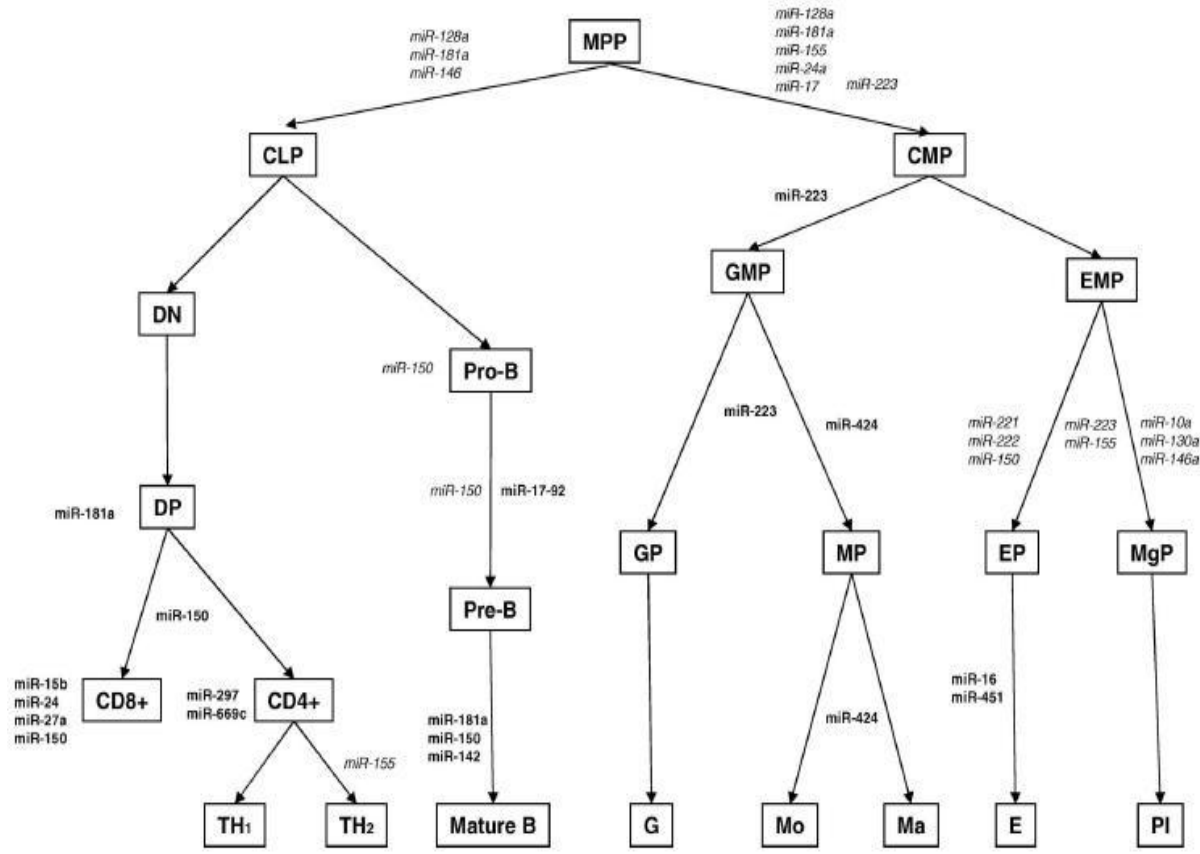


Figura 1.6. Ruolo dei miR nella normale ematopoiesi. La figura mostra i principali miR coinvolti nella determinazione differenziativa e nei vari passaggi di maturazione degli MPP (Progenitori Multipotenti) verso uno specifico lineage ematopoietico. I miR in grassetto sono up-regolati; quelli in corsivo sono down-regolati. CLP: progenitori comuni linfoidi; DN: doppi negativi; DP: doppi positivi; TH₁: T-helper 1; TH₂: T-helper 2; CMP: progenitori comuni mieloidi; GMP: precursori granulomonocitici; GP: precursori granulocitici; G: granulociti; MP: precursori monocito-macrofagici; Mo: monociti; Ma: macrofagi; EMP: precursori eritro-megacariocitici; EP: precursori eritrocitici; E: eritrociti; MgP: precursori megacariocitici; Pl: piastrine [59].

Il silenziamento genico di componenti del sistema di processamento dei miR (Drosha, Dicer, etc.) o la sregolata espressione di alcuni specifici miR compromette gravemente diverse fasi del processo ematopoietico e della risposta immunitaria e può portare allo sviluppo di disordini immunitari e di neoplasie ematopoietiche [41].

In particolare, l'inattivazione condizionale di Dicer-1 (Dcr1), enzima chiave della biogenesi dei miR, in linfociti T e B murini ha mostrato quanto i miR siano importanti nello sviluppo e nella differenziazione linfocitaria. Nel caso dello sviluppo T, la delezione di Dicer-1 (mediata dal transgene LcK-cre) in timociti immaturi (allo stadio

di DN) porta alla riduzione di 10 volte del numero di timociti totali, associato ad un drastico calo di cellule a stadi più avanzati di maturazione, quali i timociti DP e SP (CD4⁺ e CD8⁺), nonché i linfociti T periferici. Ciò nonostante, nei topi deficitari per Dicer-1 le percentuali delle diverse sottopopolazioni timiche rimangono inalterate rispetto ai controlli normali [42].

Quando invece Dicer-1 viene deleto allo stadio DP (mediante il transgene CD4-cre) si assiste ad una riduzione più modesta solo nel numero dei timociti maturi SP, con un dimezzamento del numero dei linfociti CD4⁺ periferici. In tutti i casi, le cellule T con la mutazione di Dicer mostrano una ridotta capacità proliferativa ed un'augmentata tendenza all'apoptosi, dopo attivazione *in vitro*. Inoltre, se coltivate in condizioni non polarizzanti, le cellule T CD4⁺ Dicer-delete mostrano una tendenza a differenziarsi in senso T helper 1 (Th1), associata all'incapacità di reprimere l'espressione di IFN- γ [43]. In ogni caso, la sottopopolazione T più colpita dalla delezione di Dicer-1 allo stadio DP sembra essere quella delle cellule T regolatorie (T reg), che diminuiscono di circa 6 volte sia a livello del timo che in periferia, portando allo sviluppo di una grave patologia autoimmune nei topi transgenici [44]. Questo dato è confermato dal fatto che anche la delezione di Drosha (mediata dal transgene CD4-cre) risulta in una patologia autoimmune del tutto simile a quella causata dal deficit di FoxP3, il principale fattore di trascrizione che controlla la differenziazione e la funzione di queste cellule. Inoltre, l'omeostasi e la capacità soppressoria delle T reg deficienti per Dicer-1 risulta marcatamente ridotta in condizioni non infiammatorie ed è addirittura annullata in condizioni infiammatorie.

Per quanto riguarda i linfociti B, la delezione di Dicer-1 in stadio precoce di sviluppo porta al blocco quasi totale della transizione dallo stadio pro-B allo stadio pre-B, dovuto almeno in parte all'apoptosi indotta nelle cellule pre-B Dicer^{-/-}.

La delezione condizionale di Dicer-1 solleva indubbiamente dei problemi sperimentali ed interpretativi dovuti all'esistenza di diversi membri della famiglia di Dicer, alla lunga emivita di questa proteina, al tempo di decadimento variabile di miR diversi, nonché alle funzioni alternative al processamento dei miR che sono state recentemente associate a questa proteina. Tuttavia, esperimenti di questo genere possono fornire un'idea generale dell'impatto dei miR sullo sviluppo T e B [41,45].

Parallelamente, sono stati condotti diversi studi modulando l'espressione di specifici miR o valutando le variazioni del profilo d'espressione dei miR in popolazioni cellulari ematopoietiche a diversi stadi di sviluppo o sottoposte a stimoli di diversa natura.

Questo ha portato all'identificazione di singoli miR con ruoli importanti nello sviluppo linfocitario.

La regolazione dinamica dell'espressione dei miR nel corso dello sviluppo T in un sistema murino è stata descritta dal gruppo di ricerca di Neilson [46]. In questo lavoro si mostra che la maggior parte dei miR è espressa in tutte le fasi dello sviluppo timocitario a livelli basali, mentre solo pochi miR (21, nello specifico) presentano transitori arricchimenti in corrispondenza di specifici stadi maturativi. In particolare, il **miR-181a** è riportato come fortemente e specificamente over-espresso nei timociti DP. Lo stesso miR-181, se espresso in cellule staminali ematopoietiche (HSC), ne induce la differenziazione a cellule B. Inoltre, sempre mir-181 insieme a miR-146 e miR-128, secondo lo studio di Georgantas [40], è in grado di inibire la differenziazione degli MPP (Progenitori Multipotenti) in CLP (Progenitori Comuni Linfoidi). Tuttavia, allo stesso tempo, questo miR è in grado di controllare la sensibilità intrinseca all'antigene delle cellule T, modulando la forza del segnale trasmesso dal TCR dopo stimolazione [47]. I livelli di sensibilità all'antigene sono regolati nel corso della maturazione di queste cellule al fine di permettere lo sviluppo del repertorio immunitario e della tolleranza nei linfociti T. Il miR-181 (come riportato per primo da Neilson) è quindi altamente espresso nella popolazione timocitaria immatura DP, che riconosce antigeni self a bassa affinità, mentre è down-regolato nelle popolazioni T più mature (quali i timociti SP ed i linfociti T periferici), che devono essere reattive solo nei confronti di antigeni non-self ad alta affinità. Di fatto il silenziamento del miR-181 nei timociti DP, diminuendone la sensibilità all'antigene, ne inibisce i processi di selezione positiva e negativa; al contrario, la sua over-espressione nei T maturi ne aumenta la sensibilità del TCR, portandoli a reagire anche ad antigeni peptidici normalmente identificati come antagonisti. Questi effetti del miR-181 sono mediati dal *targeting* coordinato di questo miR di diversi bersagli fra cui alcune fosfatasi associate alla via di ERK ed alcuni geni coinvolti nella selezione positiva dei timociti, quali Bcl-2, CD69 e TCR.

Oltre a miR-181, anche **miR-155** sembra svolgere un ruolo importante nella differenziazione e nell'attivazione linfocitaria. Alti livelli di miR-155 sono presenti nelle cellule B e T attivate, nonché nei monociti attivati. Da una parte, nei linfociti B, miR-155 contribuisce alla formazione del centro germinativo, all'induzione della risposta anticorpale T-dipendente, allo *switch* isotipico degli anticorpi ed alla formazione di cellule B memoria. D'altro canto, nei linfociti T, miR-155 partecipa al controllo della differenziazione delle cellule T helper. Nei topi 155^{-/-} le cellule T CD4⁺

producono aumentati livelli di citochine IL-4, IL-5 e IL-6 e sono intrinsecamente prone al differenziamento in senso Th2. L'alterata sintesi di citochine sembra essere dovuta all'up-regolazione di c-Maf, fattore di trascrizione attivante l'espressione di citochine associate al fenotipo Th2, nonchè target di miR-155. In aggiunta, FoxP3 regola direttamente la trascrizione del miR-155, che a sua volta contribuisce a mantenere le cellule T regolatorie altamente responsive a IL-2, fattore di crescita e di sopravvivenza fondamentale per le cellule T.

Altri due miR considerati importanti per il mantenimento delle cellule T regolatorie sono i **miR-146** e **miR-223**. Questi miR, già noti per i loro ruoli nel sistema immunitario innato, sono stati individuati come over-espressi specificamente nelle cellule T regolatorie rispetto alle cellule T in stato di attivazione, nonostante l'ampia sovrapposizione dei profili d'espressione dei miR in queste due popolazioni cellulari [48,49,50,51].

Infine, un altro miR molto importante nella fisiologia e nello sviluppo linfocitario è **miR-150**. Mir-150 è espresso specificamente nei linfociti B e T maturi non attivati, ma non nei loro precursori. L'over-espressione di miR-150 in cellule HSC causa il blocco della linfopoiesi B con una significativa riduzione dei livelli di cellule B in circolo, nella milza e nei linfonodi. Successivamente è stato dimostrato che l'espressione del miR-150 contrasta la transizione dei precursori linfocitari dallo stadio pro-B a quello pre-B, inibendo la generazione di cellule B mature [52]. Il gruppo di ricerca di Xiao ha dimostrato che miR-150 controlla l'espressione di c-Myb, fattore di trascrizione fondamentale in diverse fasi dello sviluppo linfocitario e necessario per la generazione della sottopopolazione B matura denominata B1. Il *pattern* di espressione di c-Myb nel corso dello sviluppo linfocitario è complementare a quello di miR-150. Combinando approcci di targeting genetico di tipo *loss-* e *gain-of-function*, si è visto che topi transgenici con deficit di c-Myb hanno un fenotipo molto simile ai topi over-esprimenti ectopicamente miR-150. Quest'ultimo tipo di topi transgenici risulta vitale, fertile e morfologicamente normale, ma mostra un aberrante blocco nello sviluppo dei linfociti B con perdita della sottopopolazione B1 in periferia. L'espressione ectopica di mir-150 provoca un'alterazione anche dello sviluppo dei linfociti T, in cui si osserva un'inibizione della transizione dei timociti dallo stadio DN3 allo stadio DN4. Facendo l'esperimento opposto, cioè silenziando l'espressione di miR-150 nei topi si osserva un fenotipo meno marcato rispetto al fenotipo *gain-of-function*: anche in questo caso i topi sono vitali, fertili e morfologicamente normali; lo sviluppo delle cellule T e delle cellule

B risulta in generale normale, a parte per l'espansione della sottopopolazione B1 in periferia [53].

Nel complesso, questi ed altri studi recenti mostrano il ruolo determinante che i miR svolgono in diverse fasi dello sviluppo e dell'attivazione dei linfociti T.

1.6 Ruolo dei microRNA nella leucemogenesi

Dato l'esteso impatto dell'azione dei miR sulla regolazione della normale ematopoiesi e di molti processi biologici fondamentali, non è sorprendente che l'alterazione dell'espressione di queste molecole sia coinvolta nella patogenesi dei tumori in generale e delle neoplasie ematopoietiche in particolare. I miR, per loro natura, possono comportarsi da oncosoppressori, andando a silenziare l'espressione di target oncogeni, o viceversa possono agire loro stessi come oncogeni, reprimendo l'espressione di geni oncosoppressori. Dato che uno stesso miR ha molti target potenziali, che in genere diventano effettivi in modo cellulo-specifico e solo in determinati contesti ambientali, paradossalmente un miR può avere in un certo contesto azione oncogenica ed in un altro oncosoppressiva [51].

L'analisi del profilo d'espressione di diversi tipi di neoplasie umane ha mostrato che spesso nei tumori si ha una globale down-regolazione dei miR maturi, accompagnata dall'up-regolazione di specifici gruppi di miR [54,55]. Inoltre i geni dei miR sono spesso localizzati in siti cromosomici fragili e in regioni genomiche frequentemente alterate nella carcinogenesi [56]. Dati recenti mostrano che i miR possono contribuire direttamente a diversi aspetti della tumorigenesi, come la disseminazione metastatica ed il mantenimento delle cellule staminali del cancro [57,58].

La sregolata espressione dei miR è una caratteristica comune a diverse neoplasie ematologiche e si è visto che il profilo d'espressione dei miR permette di classificare più accuratamente rispetto ai profili di espressione genica tumori poco differenziati. In questi anni alcuni specifici miR sono stati associati a diversi tipi di neoplasie ematologiche. E' interessante notare che molti dei miR che sono stati associati alla leucemogenesi, erano già stati dimostrati avere una funzione nello sviluppo fisiologico e nella differenziazione delle cellule ematopoietiche [51].

Qui di seguito si andrà a riassumere i principali dati riguardanti in particolare il coinvolgimento dei miR nella patogenesi delle leucemie linfatiche e dei linfomi.

1.6.1 Leucemia Linfatica Cronica (CLL)

La CLL è la forma di leucemia più comune nei paesi occidentali. Essa è caratterizzata dal progressivo accumulo di linfociti B CD5⁺ per lo più non proliferanti nel sangue, nel midollo osseo e nei tessuti linfatici. La prima prova del coinvolgimento dei miR nella tumorigenesi umana è stata scoperta proprio studiando la patogenesi di questa forma di leucemia. Nella CLL la più frequente aberrazione cromosomica consiste nella delezione in emizigosi o in omozigosi del locus 13q14, presente in più del 50% dei casi ed associata a malattia ad andamento indolente [59]. Per anni si è ricercata in questa regione la presenza di un gene oncosoppressore. Qualche anno fa il gruppo di Calin et al. ha mostrato che il *cluster* dei miR-15a e 16-1 è localizzato esattamente in quella regione e che entrambi i miR risultano down-regolati o deleti nel 68% dei campioni di CLL. In seguito, Cimmino et al. hanno scoperto che questi due miR controllano l'espressione di Bcl-2, una proteina anti-apoptotica over-espressa nella maggior parte delle CLL, apparentemente senza un'aberrazione genetica causale (con l'eccezione della giustapposizione di Bcl-2 al locus dell'immunoglobulina nel 5% dei casi). Il controllo di Bcl-2 da parte dei miR-15a e -16-1 è stato proposto come meccanismo patogenetico alla base della sua sregolata espressione nella CLL. Tuttavia, l'impatto di questo miR *cluster* nella leucemogenesi sembra essere molto più complesso ed ampio di quanto in precedenza ritenuto: infatti, da un'analisi del trascrittoma è risultato che circa il 14% dei geni umani viene influenzato dall'over-espressione di questo *cluster* [60,61,62].

Altri due miR che giocano un ruolo importante nella patogenesi della CLL sono miR-181b e miR-29b. Essi svolgono una funzione oncosoppressoria nella CLL, reprimendo l'espressione dell'oncogene TCL-1 (*T-cell Leukemia/Lymphoma-1*). I livelli di entrambi i miR, infatti, correlano inversamente con quelli di TCL-1 nei pazienti affetti da CLL. TCL-1 è un co-attivatore di Akt, un oncogene che inibisce i processi apoptotici e che gioca un ruolo chiave nella regolazione di diverse vie molecolari coinvolte nella sopravvivenza, nella proliferazione e nella morte cellulare. Elevati livelli di TCL-1 sono associati ad alti livelli di ZAP-70 ed allo stato non-mutato delle IgV_H, entrambi marcatori di leucemia ad andamento aggressivo [59]. Inoltre, studiando un gruppo di 94 CLL, Calin et al. hanno individuato una *signature* di 13 miR con valore prognostico [63]. Infine, nelle CLL è stata riscontrata una frequente up-regolazione di mir-155, similmente a quanto avviene nei linfomi B [64].

1.6.2 Linfoma

A miR-155 è stato attribuito un chiaro ruolo oncogenico in diversi tipi di neoplasie, fra cui in particolare alcuni tipi di linfoma B. In un modello di topo transgenico l'over-espressione di miR-155, in modo cellula B-specifico, induce, tramite un meccanismo ancora sconosciuto, lo sviluppo di una proliferazione clonale pre-leucemia pre-B che sfocia in una leucemia B manifesta nel corso del tempo. L'up-regolazione di questo miR è stata ben descritta nei casi pediatrici di linfoma di Burkitt (BL). Inoltre, due diversi studi hanno rilevato l'over-espressione di questo miR nei DLBCL (*Diffuse Large B-cell Lymphoma*): in questo tipo di linfoma, alti livelli di miR-155 caratterizzano il fenotipo B attivato (ABC-DLBCL o Activated B-Cell DLBCL) rispetto al fenotipo B da centro germinativo (GC-DLBCL o Germinal Center B-Cell-like DLBCL). Poichè questi due fenotipi di DLBCL hanno prognosi diverse, miR-155 è stato proposto come marcatore prognostico in questo tipo di neoplasia. Sempre miR-155 è stato descritto come up-regolato anche nei linfomi di Hodgkin (HL). Alti livelli di questo miR sono stati decritti a livello del centro germinativo durante la normale linfopoiesi e, dato che sia i DLBCL, che i BL ed i HL hanno origine dal centro germinativo, è stato postulato che l'iper-espressione di miR-155 causi un blocco della differenziazione linfocitaria proprio a livello del centro germinativo, mantenendo l'attivazione di meccanismi anti-apoptotici che favoriscono la trasformazione neoplastica [59,65,66].

Nel suo studio, Roehle et al. hanno comparato il profilo d'espressione di miR di 58 DLBCL, 46 FL (Linfomi Follicolari) e 7 linfonodi non-neoplastici, individuando un gruppo di 4 miR (miR-330, miR-17-5, miR-106a e miR-210) in grado di discriminare questi due tipi di linfoma fra loro e rispetto alla controparte normale con un'accuratezza del 98% [67].

Il miR *cluster* 17-92 è localizzato in un sito cromosomico (13q31) frequentemente amplificato nei linfomi B. Questo *cluster* è over-espresso nel 65% dei pazienti affetti da linfoma B. Inoltre, topi over-esprimenti miR-17-92 nei linfociti sviluppano disordini linfoproliferativi ed autoimmuni e muoiono prematuramente. Questi effetti sembrano essere secondari all'azione di alcuni membri del *cluster* sui target oncosoppressori PTEN e Bim, gene che controlla l'apoptosi nelle cellule B. E' stato dimostrato che l'up-regolazione del *cluster* 17-92, insieme a quella dell'oncogene Myc, accelera sinergicamente lo sviluppo tumorale. Successivamente, è stato scoperto che lo stesso Myc attiva l'espressione del *cluster* 17-92, il quale ha fra i suoi target il fattore di trascrizione E2F-1, che normalmente promuove la progressione del ciclo cellulare, ma

ad alti livelli funge da fattore pro-apoptotico [59,68]. Fra i miR con supposto ruolo oncosoppressore in diverse neoplasie ematologiche ci sono miR-143 e mir-145, la cui espressione è ridotta in diverse neoplasie B, come la B-CLL (Leucemia Linfatica Cronica) ed i linfomi B.

L'analisi di un sito fragile a livello del cromosoma 12 nei linfomi a cellule T (indotti da radiazione) ha mostrato che in questa regione sono localizzati ben 52 miR, di cui almeno uno, miR-203, sembra essere implicato nella patogenesi di questi tumori. Infatti, miR-203 ha fra i suoi target il noto oncogene ABL1 [69].

Tuttavia, il ruolo dei miR nei linfomi T rimane poco chiaro. Lum et al. hanno scoperto un sito ad elevata frequenza di inserzione retrovirale a monte del miR *cluster* 106a-363 nei linfomi T indotti nei topi mediante infezione con RadLV (Radiation Leukemia Virus). Nei tumori contenenti tale integrazione, il trascritto primario policistronico del miR *cluster* 106a-363 era over-espresso 5-20 volte rispetto ai tumori di controllo. Successivamente il gruppo di Landais ha rilevato l'over-espressione di questo stesso *cluster* di miR nel 46% delle T-ALL umane e ne ha confermato *in vitro* il potenziale oncogeno [70,71].

1.6.3 Leucemia Linfatica Acuta (ALL)

La ALL è una delle più comuni neoplasie dell'età infantile. Questo tipo di leucemia deriva dalla proliferazione clonale di progenitori linfoidei a livello del midollo osseo, che porta all'insorgenza di vari tipi di citopenia associati alla presenza di blasti a livello del sangue periferico. La ALL rappresenta l'80% dei casi di leucemia acuta nei bambini e ed il 20% dei casi negli adulti.

Al contrario della CLL e dei linfomi e nonostante la rilevanza epidemiologica di questo tipo di neoplasie (soprattutto nei bambini), pochissimi studi hanno investigato il profilo d'espressione dei miR nelle ALL e di conseguenza il ruolo dei miR nella patogenesi di questo tipo di leucemia è ancora in gran parte sconosciuto [72,73]. La prima evidenza del coinvolgimento di un miR nella ALL è stata la scoperta di un riarrangiamento cromosomico fra il gene della catena pesante dell'immunoglobulina e il *locus* del miR-125-1 in un paziente affetto da B-ALL [74]. In seguito, Mi et al. hanno condotto uno studio *genome-wide* sul profilo d'espressione dei miR nelle ALL e nelle AML (Leucemie Mieloidi Acute), identificando 27 miR differenzialmente espressi nei due tipi di leucemie acute. Fra questi miR, i miR-128a e b sono stati visti significativamente over-espressi, mentre i miR-223 e let-7b significativamente down-regolati nelle ALL rispetto alle AML [75]. Il gruppo di ricerca di Zanette ha comparato il profilo

d'espressione dei miR in un gruppo di 7 ALL miste (T, B e bi-fenotipiche) con quello di normali cellule B CD19⁺ prelevate dal sangue periferico di 6 donatori sani, stilando una lista di miR de-regolati nelle ALL. In questo modo, Zanette et al. hanno identificato fra i 5 miR maggiormente up-regolati nelle ALL rispetto alla controparte normale miR-128b, miR-204, miR-218, miR-331 e miR-181b. Al contrario, i 5 miR con i livelli più bassi di espressione nelle ALL rispetto alle cellule normali sono risultati essere miR-135b, miR-132, miR-199s, miR-139 e miR-150. Inoltre, sempre in questo lavoro, il miR *cluster* 17-92 è stato trovato up-regolato in tutti i campioni di ALL, similmente a quanto riportato per i linfomi B [72]. Questo *cluster* ha una nota funzione oncogena in diversi tipi di tumori umani sia solidi che ematopoietici. Esso è composto da 7 miR (miR-17-5p, miR-17-3p, miR-18a, miR-19a, miR-20a, miR-19b-1, miR-92-1) ed è localizzato nel cromosoma 13 in posizione 13q31. Li et al. hanno recentemente mostrato che il *cluster* 17-92 è particolarmente over-espresso nelle leucemie acute con riarrangiamenti del locus MLL [76]. Nella normale ematopoiesi, i miR-17-92 sono down-regolati durante la monocitopoiesi, la megacariocitopoiesi e la linfopoiesi B. Nei modelli murini l'over-espressione di questo *cluster* è associata a iperproliferazione ed inibizione dell'apoptosi e porta allo sviluppo di malattie linfoproliferative ad andamento aggressivo. C-Myc controlla trascrizionalmente l'espressione del *cluster* 17-92, il quale, a sua volta, regola l'espressione di un gruppo di geni target con ruoli cruciali nella leucemogenesi, come PTEN, BIM, E2F e RB1 [59,65,66,68].

In conclusione, rispetto ad altri tipi di tumori solidi ed ematologici, il ruolo dei miR nelle leucemie acute e nei linfomi linfoblastici a cellule T è stato oggetto di pochissimi studi. Inoltre, i pochi studi condotti generalmente mirano a descrivere in modo *genome-wide* il profilo d'espressione dei miR in queste neoplasie, senza andare ad approfondire la relazione fra modulazione dell'espressione e funzione dei miR nelle cellule, nonché le conseguenze di questa modulazione sull'espressione genica globale. Di conseguenza, l'importanza ed il significato della modulazione di queste piccole molecole nella patogenesi di questo tipo di leucemie è ancora in gran parte ignoto.

2. SCOPO DELLA TESI

Lo sviluppo delle cellule T avviene attraverso una serie di fasi maturative ed è guidata da fattori e meccanismi che rimangono ancora non del tutto noti. I microRNA (miR) sono una classe di piccoli RNA non codificanti (19-24 nt) che svolgono un ruolo importante nella modulazione post-trascrizionale dell'espressione genica. Recenti lavori hanno dimostrato che alcuni miR sono modulati dinamicamente nella loro espressione e svolgono funzioni essenziali all'interno del processo ematopoietico, della risposta infiammatoria e della leucemogenesi. Inoltre, molte evidenze suggeriscono che la sregolata espressione di miR sia una caratteristica comune a diverse neoplasie ematologiche. Tuttavia, attualmente, il ruolo dei miR nello sviluppo fisiologico delle cellule T, nonché il significato della loro alterata espressione nella leucemogenesi T è ancora in gran parte sconosciuto. Inoltre, finora la maggior parte degli studi sui miR sono stati condotti su organismi inferiori o su topo; pertanto si sa ancora molto poco della modulazione di queste molecole nel corso dei diversi processi fisiologici e patologici nell'uomo. Allo scopo di affrontare queste questioni, nel mio progetto di dottorato ho analizzato il profilo d'espressione dei microRNA in popolazioni T a diversi stadi di sviluppo al fine di individuare miR implicati nel processo maturativo dei linfociti T e potenzialmente associati alla trasformazione in senso neoplastico di queste cellule. Per perseguire più specificamente questo secondo obiettivo, nella fase conclusiva del mio dottorato, abbiamo comparato (in collaborazione con il laboratorio del Dott. Rosolen, Dipartimento di Pediatria, Università di Padova) il profilo d'espressione di popolazioni timocitarie sane con quello di un gruppo di linfomi linfoblastici T pediatrici (T-LBL) e di linfonodi reattivi non-neoplastici umani (LN).

3. MATERIALI E METODI

3.1 Campioni primari e colture cellulari

Le popolazioni timiche sono state ottenute da frammenti operatori di timo umano sano prelevati da 35 pazienti pediatriche sottoposti ad interventi chirurgici per difetti cardiaci congeniti presso l'Ospedale Policlinico Universitario di Padova. I bambini avevano un'età compresa fra i 2 giorni e i 5 anni e nessuno di essi era affetto da disordini ematologici o immunologici. Tutti i frammenti operatori sono stati prelevati con il consenso informato dei genitori.

Le popolazioni linfocitarie mature sono state ottenute dal sangue periferico di donatori sani.

I campioni di linfoma linfoblastico T (T-LBL) derivano dalle biopsie di tessuto linfonodale tumorale ottenute da 9 pazienti pediatriche affetti da T-LBL; mentre i linfonodi reattivi non-neoplastici (LN) sono stati ottenuti dalle biopsie linfonodali di 6 pazienti pediatriche non affetti da tumore. Sia i campioni di T-LBL, che i LN sono stati forniti dal laboratorio del Dott. Rosolen del Dipartimento di Pediatria dell'Università degli studi di Padova. Prima di essere inviato al laboratorio, il tessuto linfonodale in eccesso derivante da biopsie diagnostiche dei pazienti veniva esaminato e selezionato dal patologo locale. Una volta arrivate in laboratorio, le biopsie venivano immediatamente processate, o congelate e conservate a -80°C fino al momento dell'utilizzo. Il tessuto linfonodale veniva quindi sminuzzato ed omogenizzato meccanicamente per l'estrazione dell'RNA totale.

Inoltre, in questo lavoro sono state utilizzate le seguenti linee cellulari:

- le linee di T-ALL Jurkat, MOLT-3 e CCRF-HSB-2 sono state acquistate dalla ATCC (American Tissue Culture Collection), mentre le TALL-1 e le DND4.1 sono state ottenute dal Dr. A. Ferrando (Institute for Cancer Genetics-Columbia University, Columbia University Medical Center, New York, USA). Tutte le linee cellulari di T-ALL usate in questo studio sono state coltivate in terreno RPMI 1640 (EuroClone) completo, ovvero addizionato con 10% di FBS (Fetal Bovine Serum, Life Technologies), 1% di Sodio Piruvato, 10 mmol/L di HEPES (Cambrex Bio Science), 2 mmol/L di L-Glutammina e 1% di un mix di antibiotici ed antimicotici (Gibco-BRL).

- la linea 293T è stata ottenuta dalla ATCC ed è stata coltivata in terreno DMEM (Sigma) completo, ossia supplementato con 10% di FBS (Fetal Bovine Serum, Gibco-BRL), 1% di Sodio Piruvato, 10 mmol/L di HEPES (Cambrex Bio Science) e 1% di un mix di antibiotici ed antimicotici (Gibco-BRL). Le 293T, sono una linea cellulare derivata da cellule embrionali renali umane, altamente trasfettabile. Questa linea è stata utilizzata come linea *packaging* per la produzione di vettori virali.

3.2 Isolamento delle sottopopolazioni timiche e linfoidi da timo e da sangue periferico

I timociti sono stati isolati mediante dissezione meccanica del tessuto timico, filtraggio per la rimozione degli aggregati di cellule epiteliali e stromali timiche e successiva lisi osmotica degli eritrociti con soluzione di lisi contenente cloruro di ammonio [77]. Dopo un'incubazione di 5 minuti a 37°C con la soluzione di lisi, le cellule venivano lavate per due volte in PBS 1X e contaminate. La sospensione cellulare ottenuta in questo modo risultava costituita per più del 90% da cellule timocitarie, definite per la positività all'antigene CD2, un marcatore precoce dei timociti umani.

Le popolazioni timocitarie d'interesse per lo studio (timociti CD2⁺/CD4⁺/CD8⁻ singoli positivi, **SP CD4**; CD2⁺/CD8⁺/CD4⁻ singoli positivi, **SP CD8**; CD2⁺/CD4⁺/CD8⁺ doppi positivi, **DP**) sono state individuate mediante marcatura per CD4 (CD4-FITC; Beckman Coulter), CD8 (CD8-PE; Beckman Coulter) e CD2 (CD2-PC5; Beckman Coulter) e sortate tramite *FACSVantage fluorescence-activated cell sorter* (Becton Dickinson, San Diego, CA) ad un livello di purezza superiore al 95% (Fig. 3.1). In breve, dai 40 ai 60 milioni di cellule venivano incubate per 30 minuti in ghiaccio ed al buio in presenza degli anticorpi specifici. Dopo l'incubazione, le cellule, lavate una volta in PBS 1X, venivano risospese alla concentrazione di 2x10⁷ cellule/ml in terreno RPMI 1640 (Euroclone) completo e separate tramite *FACS sorting*. Per la separazione, le sottopopolazioni timiche d'interesse erano selezionate all'interno del *gate* delle cellule positive per CD2, in base alla positività per i marcatori CD4 e/o CD8.

I linfociti maturi CD4⁺ e CD8⁺ sono stati ottenuti da PBMC (*Peripheral Blood Mononuclear Cells*) mediante separazione immunomagnetica in negativo, utilizzando rispettivamente il CD8 T Cell Isolation Kit II human ed il CD4 T Cell Isolation Kit II human della ditta Miltenyi Biotec. Questi kit Miltenyi sfruttano un sistema di marcatura magnetica indiretta per isolare in negativo da PBMC umani rispettivamente le cellule T

citotossiche CD8⁺ e le cellule T helper CD4⁺. I PBMC utilizzati a tale scopo erano stati precedentemente isolati tramite centrifugazione a gradiente di densità con Ficoll-Paque da *buffy coat* di sangue periferico umano di volontari sani. Per la separazione dei linfociti, i PBMC venivano centrifugati a 300 g per 10 minuti; quindi il pellet veniva risospeso in 40 µl/10⁷ cellule di MACS buffer (PBS, 0,5% BSA, 2mM EDTA; pH 7,2) addizionato con 10 µl/10⁷ cellule di Biotin-Antibody Cocktail e le cellule venivano incubate per 10 minuti in ghiaccio. Il Biotin-Antibody Cocktail è un mix di anticorpi che marcano i diversi tipi di cellule non-CD4 (linfociti T γ/δ, linfociti T CD8⁺, linfociti B, cellule NK, cellule dendritiche, monociti, granulociti ed eritrociti) o non-CD8 (linfociti T γ/δ, linfociti T CD4⁺, linfociti B, cellule NK, cellule dendritiche, monociti, granulociti ed eritrociti) presenti nel sangue periferico. Trascorso il tempo di incubazione, alle cellule venivano aggiunti 30 µl/10⁷ cellule di MACS buffer e 20µl/10⁷ cellule di Anti-Biotin MicroBeads. Le cellule venivano incubate con le MicroBead per 15 minuti in ghiaccio, poi venivano lavate in MACS buffer, centrifugate a 300 g per 10 minuti ed infine risospese ad una concentrazione di 10⁸ cellule/ml in 500 µl di MACS buffer. A questo punto si procedeva alla separazione magnetica delle cellule marcate con VarioMACS Separator (Miltenyi Biotec). A tal fine, le cellule venivano caricate su colonnine LS (attivate dal passaggio di 3 ml di MACS buffer) in presenza del campo magnetico generato dal VarioMacs. La frazione delle cellule non marcate, contenente, a seconda del Kit usato, i linfociti T CD8⁺ o CD4⁺, veniva così eluita, mentre le cellule marcate magneticamente, legate alle MicroBead, venivano trattenute nella matrice delle colonna LS. Le popolazioni linfocitarie ottenute tramite questo procedimento presentavano livelli di purezza in tutti i casi superiori all'85%.

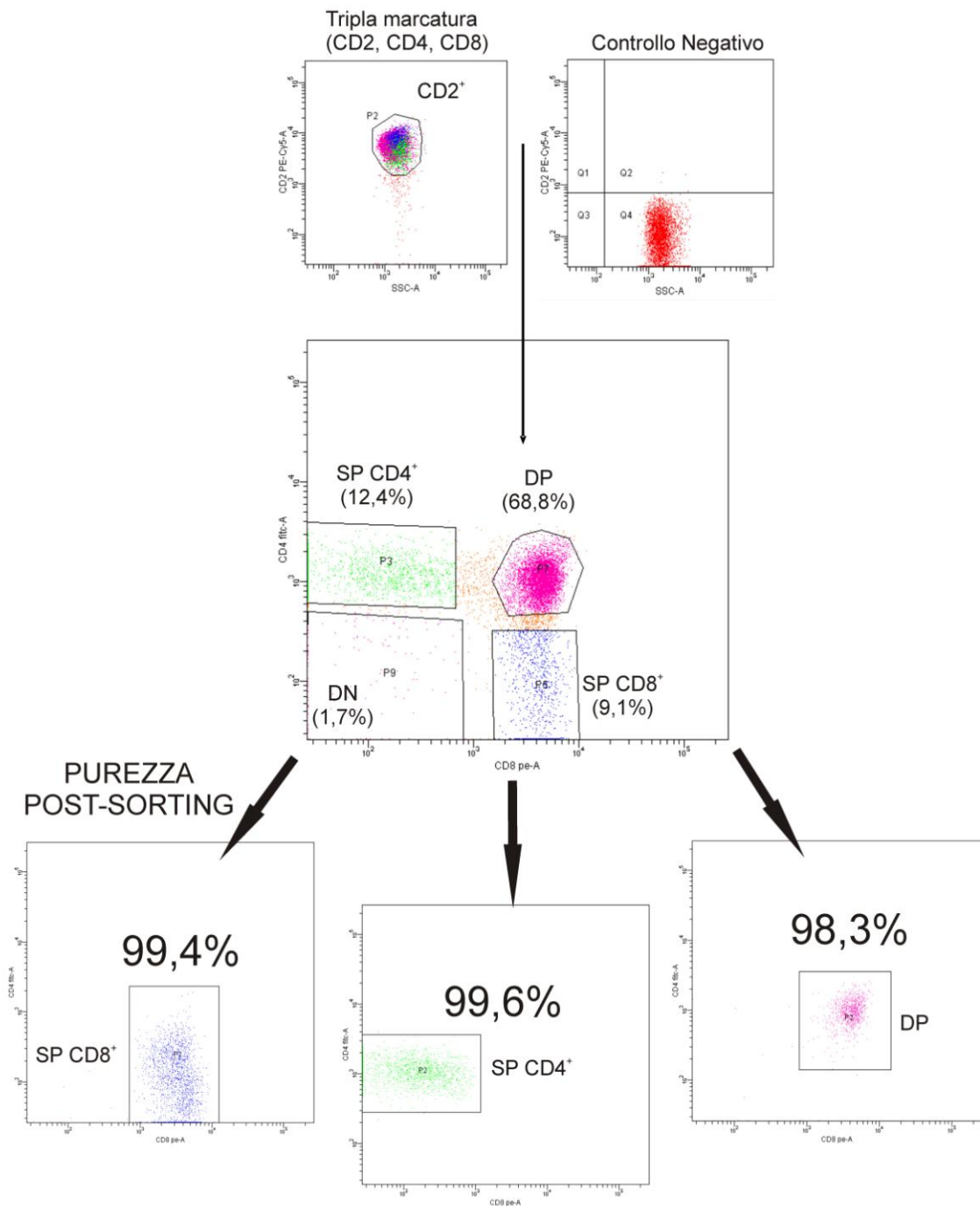


Figura 3.1. Purificazione mediante FACS sorting delle sottopopolazioni timiche. Il *gate* P2 è utilizzato per selezionare le cellule timiche, positive per CD2. All'interno del *gate* P2, la regione P3 individua i timociti CD4 SP (positivi per CD4 e negativi per CD8), la regione P7 i timociti DP (positivi per entrambi i marcatori CD4 e CD8), la regione P5 i timociti CD8 SP (positivi per CD8 e negativi per CD4), mentre la regione P9 comprende tutte le restanti cellule timocitarie DN, non esprimenti né CD4, né CD8 (precursori timocitari immaturi TN, CD4/CD8/CD3⁻; linfociti T $\gamma\delta$; precursori NK). I 3 citogrammi sottostanti mostrano la purezza media delle popolazioni timiche ottenute dopo *FACS sorting*.

3.3 Estrazione dell'RNA

L'RNA totale è stato purificato dai campioni con il metodo della separazione di fasi con Trizol Reagent (Invitrogen) e cloroformio. Operativamente, le cellule venivano lavate in PBS 1X e quindi fatte precipitare tramite centrifugazione a 1200 rpm per circa 8-10 minuti. I pellet ottenuti venivano lisati in 500 µl - 1 ml di Trizol, incubati per 5 minuti a temperatura ambiente, per permettere la dissociazione dei complessi nucleo-proteici, e quindi conservati a -80°C fino al momento dell'estrazione. L'omogenizzazione dei campioni con il Trizol causa la lisi cellulare, preservando l'integrità dell'RNA; inoltre, l'estrazione dell'RNA totale tramite il metodo della separazione di fase con il Trizol permette di isolare efficientemente diverse specie di RNA, comprese quelle a basso peso molecolare, quali i microRNA, che normalmente vengono perse con i classici metodi di estrazione in colonna.

Al momento dell'estrazione dell'RNA, ai campioni, scongelati a temperatura ambiente, veniva aggiunto il cloroformio (200 µl/ml di Trizol). Quindi i campioni venivano mescolati per inversione, incubati per 5 minuti a temperatura ambiente e centrifugati a 12000 g per 15 minuti a 4°C. In seguito alla centrifugazione la miscela si separava in una fase inferiore cloroformio-fenolica contenente le proteine e le membrane cellulari, un'interfase contenente il DNA ed una fase superiore acquosa contenente l'RNA. Dopo il trasferimento della fase acquosa in una nuova provetta, l'RNA era recuperato mediante precipitazione con alcol isopropilico (700 µl/ml di Trizol) e centrifugazione a 12000 g per 15 minuti a 4°C. Il pellet di RNA totale così ottenuto veniva lavato una volta in etanolo 75% (1 ml/ml Trizol), ri-centrifugato a 12000 g per 10 minuti a 4 °C, asciugato all'aria e quindi risospeso in un adeguato volume di acqua iniettabile *nuclease-free*.

Per eliminare eventuali contaminanti fenoliche, che possono interferire con l'efficienza del *labeling* dell'RNA e dell'ibridazione su microarray, l'RNA era fatto precipitare o.n. (*over-night*) a -20°C, aggiungendo 2,5 volumi di etanolo assoluto e 0,1 volume di Sodio Acetato 3 M alla soluzione acquosa di RNA. Quindi la soluzione veniva centrifugata a 12000 g per 30 minuti a 4°C, veniva fatto un ulteriore lavaggio in etanolo 75% ed infine l'RNA purificato era risospeso in acqua iniettabile *nuclease-free*.

3.4 Quantificazione dell'RNA estratto e verifica della sua qualità

L'RNA totale dei campioni è stato quantificato allo spettrofotometro (NanoDrop) e controllato per la qualità mediante elettroforesi capillare (RNA 6000 Nano Assay, Agilent Technologies).

3.4.1 Analisi spettrofotometrica

La quantificazione dell'RNA è stata effettuata misurando l'assorbanza del campione a 260 nm utilizzando lo spettrofotometro Nanodrop ND-1000 della Nanodrop Technology, utilizzando 1 µl di campione per la lettura. Questo spettrofotometro UV/visibile è in grado di lavorare con micro-volumi di campione (fino ad 1 µl) e di quantificare accuratamente e linearmente in un intervallo di concentrazione compreso tra 2 e 3700 ng, senza diluizioni. Il Nanodrop permette quindi di utilizzare meno materiale per l'esperimento; inoltre, non necessita dell'uso di cuvette o capillari, perché il campione viene pipettato direttamente sulla superficie di misura. Al fine della quantificazione, una goccia di soluzione acquosa di RNA veniva posizionata direttamente sull'apposita piastra di lettura ed il braccio dello strumento veniva chiuso, avvicinando le due fibre ottiche. In questo modo la tensione superficiale mantiene sulla colonna il campione liquido, mentre in meno di 10 secondi viene effettuata la misurazione tramite le due fibre ottiche. Lo spettro di assorbanza, la concentrazione (ng/µl) dell'RNA totale nel campione ed il rapporto fra l'assorbanza a 260 nm/280 nm e a 260 nm/230 nm (indici del grado di purezza del campione analizzato) venivano in seguito visualizzati sullo schermo del PC ed annotati.

La concentrazione dell'RNA estratto veniva ricavata automaticamente dallo strumento mediante la seguente formula matematica:

$$[\text{RNA}] \text{ ng}/\mu\text{l} = A_{260\text{nm}} \times \text{fattore di diluizione} \times \text{fattore di conversione}$$

(pari a 40 per l'RNA)

3.4.2 Analisi elettroforetica

L'integrità dell'RNA estratto è stata valutata mediante elettroforesi capillare con lo strumento Agilent Bioanalyzer 2100, secondo la procedura indicata dal kit *RNA 6000 Nano Assay* (Agilent Technologies), specifico per l'analisi dell'RNA totale eucariotico. Ogni RNA LabChip è costituito da un set di microcapillari nei quali viene polimerizzato un gel per la separazione dei frammenti di acido nucleico. Il software del Bioanalyzer

traccia per ciascun campione un profilo elettroforetico da cui ricavare la concentrazione e la qualità dell'acido nucleico, misurando il rapporto delle aree dei picchi corrispondenti all'RNA ribosomale 28S e 18S.

3.5 Analisi del profilo d'espressione di microRNA mediante microarray

L'analisi del profilo d'espressione dei miR mediante microarray è stata condotta sui campioni di timociti umani non frazionati, sulle cellule delle sottopopolazioni timiche separate mediante *FACS sorting*, nonché sui campioni di linfonodi reattivi umani (LN; 6 donatori) e di linfonodi neoplastici ottenuti da 9 pazienti pediatriche affetti da linfoma linfoblastico T (T-LBL). I profili d'espressione dei miR sono stati generati utilizzando l'*Agilent Human miRNA Microarray (v2.0) platform* (Agilent Technologies), contenente sonde in grado di riconoscere le sequenze di 723 microRNA umani e 76 microRNA virali umani, derivate dalla versione 10.1 del *Sanger database*.

Ai fini di tale analisi, l'RNA totale dei campioni è stato estratto in Trizol, purificato da eventuali contaminanti fenoliche e controllato per qualità e quantità come descritto ai punti 3.3 e 3.4 dei Materiali e Metodi. Per ogni popolazione timica sono stati ibridati su array 6 diversi replicati biologici. Tutti i campioni sono stati marcati utilizzando l'*Agilent's miRNA Complete Labeling and Hyb Kit* ed ibridati sugli array secondo le indicazioni del produttore. Operativamente, per ogni campione, 100 ng di RNA totale venivano defosforilati all'estremità 5' mediante un'incubazione di 30 minuti a 37°C con CIP (Calf Intestinal Phosphatase). Per la reazione, ai 100 ng di RNA del campione erano aggiunti 3µl del seguente CIP Master Mix:

- 0,7 µl 10X CIP Buffer
- 1,6 µl RNase-free H₂O
- 0,7 µl TaKaRa CIP (25 U/ul)

(Volume finale = 7 µl)

I campioni, dopo l'aggiunta di 5 µl di DMSO, venivano quindi denaturati al calore, mediante un'incubazione di 8 minuti a 100°C, e poi trasferiti immediatamente in ghiaccio. L'RNA denaturato veniva direttamente marcato in fluorescenza mediante ligazione di una molecola di *cyanine3-pCp* all'estremità 3'. Per la reazione di ligazione, al campione erano aggiunti 8 µl di Ligation Master Mix preparato nel modo seguente:

- 2 µl T4 RNA Ligase Buffer
- 2 µl RNase-free H₂O
- 3 µl pCp-Cy3
- 1 µl NEB T4 RNA Ligase (20 U/µl)

(Volume finale = 20 µl)

La reazione era incubata per 2 ore a 16°C.

Quindi l'RNA marcato veniva purificato dall'eccesso di *cyanine3-pCp* e da altri contaminati mediante passaggio in colonnine *Micro Bio-Spin 6* (Bio-Rad). L'RNA marcato e purificato veniva poi asciugato in un essiccatore a vuoto e risospeso in 18 µl di acqua *nuclease-free* a cui venivano aggiunti 4,5 µl di *10X GE Blocking Agent* e 22,5 µl di *2X Hi-RPM Hybridization Buffer* per arrivare ad un volume finale di 45 µl. I campioni erano quindi incubati a 100°C per 5 minuti e trasferiti velocemente in un bagno di acqua e ghiaccio, prima di essere ibridati sugli array. I 45 µl di campione erano caricati nello spazio apposito del *gasket slide*, successivamente il lato attivo dell'array veniva sovrapposto alla goccia del campione a formare la camera d'ibridazione. Gli array così preparati erano ibridati a 55°C per 20 ore in rotazione (20 rpm).

Terminato il tempo di ibridazione, gli array venivano lavati una volta in Gene Expression Wash Buffer 1 a temperatura ambiente ed una volta in Gene Expression Wash Buffer 2 a 37°C, quindi venivano velocemente asciugati e scannerizzati con l'Agilent Microarray Scanner. I vetrini sono stati scannerizzati in modalità Extended Dynamic Range Scan (che prevede 2 letture successive, rispettivamente a bassa ed alta intensità) e con una risoluzione di 5 µm.

L'estrazione e l'analisi dei dati grezzi ottenuti dagli array è stata effettuata con il software Feature Extraction 9.5.3.1 (Agilent Technologies).

3.6 Analisi dell'espressione di microRNA tramite TaqMan qRT-PCR

La validazione dei dati di miR microarray, nonché lo studio dell'espressione di hsa-miR-150, hsa-miR-196b, hsa-miR-128, hsa-miR-17 e hsa-miR-92 nelle linee di T-ALL è stata effettuata mediante TaqMan qRT-PCR.

L'analisi mediante qRT-PCR è stata condotta sul cDNA ottenuto per retro-trascrizione dell'RNA totale estratto dalle cellule dei campioni d'interesse (sottopopolazioni timiche

separate mediante *FACS sorting*, linee di T-ALL). Prima della retro-trascrizione, l'RNA è stato purificato da contaminanti fenolici e controllato per qualità e quantità come descritto ai punti 3.3 e 3.4 di Materiali e Metodi.

Le reazioni di retro-trascrizione dall'RNA e di RT-PCR sono state realizzate utilizzando il Taqman MicroRNA Assay (Applied Biosystem). Il protocollo di questo *assay* comprende una prima fase di retro-trascrizione dell'RNA totale a cDNA con *looped-primer* miRNA-specifici, seguita da una fase di real-time PCR con *primer* e sonde *Taq-Man* sempre miR-specifici. La real-time PCR è una tecnologia che permette la quantificazione in tempo reale dei prodotti di amplificazione. La rilevazione dei prodotti di PCR è resa possibile misurando la fluorescenza di una molecola *reporter*, la quale aumenta con l'accumularsi del prodotto di reazione. Nella real-time PCR di tipo TaqMan la molecola *reporter* è contenuta all'interno della sonda oligonucleotidica. Infatti, le sonde Taq-Man contengono una molecola *reporter* fluorescente (FAM) legata all'estremità 5' ed una molecola *quencher* non fluorescente legata all'estremità 3'. Nel corso della reazione di PCR le sonde TaqMan si appaiano in maniera specifica ad una sequenza complementare presente nel template e compresa fra il sito di legame del primer *forward* e di quello *reverse*. Quando la sonda è intatta, la vicinanza del gruppo *quencher* alla molecola *reporter* sopprime la fluorescenza di quest'ultima per un fenomeno di FRET (*Fluorescence Resonance Energy Transfer*). Nel corso della sua attività polimerasica, la DNA polimerasi taglia le sonde appaiate al loro bersaglio, separando in tal modo il *quencher* dal *reporter* e permettendo l'emissione di fluorescenza da parte di quest'ultimo. L'aumento del segnale di fluorescenza avviene quindi nel corso dell'amplificazione specifica della sequenza bersaglio grazie al clivaggio progressivo delle sonde.

Grazie all'uso di *looped-primer* nel processo di retro-trascrizione, questo tipo di test permette di valutare i livelli di espressione specificamente delle forme mature dei miR, escludendo dall'analisi i loro precursori.

Operativamente, per la retro-trascrizione dell'RNA estratto dai campioni è stato utilizzato il TaqMan MicroRNA Reverse Transcription Kit (Applied Biosystems). Per ogni campione veniva preparato un mix di reazione contenente i seguenti componenti:

- 1,5 µl di Reverse Transcription Buffer 10X
- 0,15 µl di dNTPs (100 mM, con dTTP)
- 1 µl di MultiScribe Reverse Transcriptase (50 U/µl)

- 0,19 µl RNase Inhibitor (20 U/µl)
 - 4,16 µl di acqua sterile *nuclease-free*
- (Volume finale: 7 µl/reazione)

Ai 7 µl di mix erano aggiunti 3 µl di RT-primer miR-specifici e 10 ng di RNA totale per un volume finale di 15 µl per reazione. Una volta mescolati i componenti della reazione, questi venivano incubati in ghiaccio per 5 minuti e quindi posti nel termociclatore (Thermo Hybayd PCR Thermalcycler). Il programma di retro-trascrizione veniva impostato con i seguenti parametri:

[16°C 30' - 42°C 30' - 85°C 5']

Il cDNA così ottenuto è stato usato per lo studio tramite TaqMan qPCR dei livelli di espressione di alcuni microRNA d'interesse (hsa-miR-146b e a, hsa-miR-151-5p, hsa-miR-150, hsa-miR-23a, hsa-miR-342-3p, hsa-miR-155, hsa-miR-29b, hsa-miR-196b, hsa-miR-128, hsa-miR-17, hsa-miR-92).

Per ovviare ad eventuali differenze tra campioni nella quantità di RNA utilizzato e nell'efficienza di retro-trascrizione, il valore d'espressione di ogni microRNA è stato normalizzato rispetto al valore di RNU44, uno *small nucleolar RNA* espresso allo stesso livello in tutti i campioni analizzati e pertanto utilizzato come controllo *housekeeping*.

In breve, per ciascun campione veniva preparato un mix di reazione costituito dai seguenti reagenti (tutti Applied Biosystem):

- 10 µl di TaqMan 2X Universal PCR Master Mix, No AmpErase UNG
 - 7,5 µl di acqua sterile *nuclease-free*
 - 1 µl di TaqMan MicroRNA Assay 20X
 - 1,5 µl del prodotto della reazione di RT
- (Volume finale: 20 µl/campione)

Ogni singola reazione di PCR era eseguita in duplicato in un volume di 20 µl, usando piastre ottiche da 96 pozzetti (Applied Biosystems).

La piastra da PCR, una volta distribuito il mix di reazione, veniva sigillata e caricata nel termociclatore ABI Prism 7900 Sequence Detection System (Perkin Elmer Applied Biosystems, Foster City, California, USA).

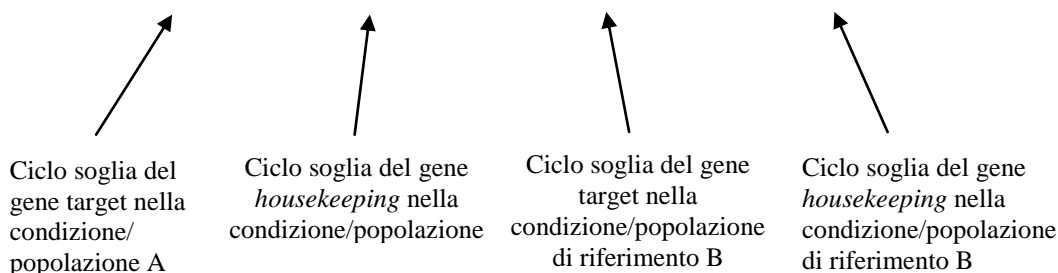
Il programma di amplificazione utilizzato era il seguente:

50°C 2' / 95°C 10'

[95°C 15" / 60°C 1'] CICLO (40 cicli)

L'analisi dei risultati della real-time PCR fornisce, per ogni campione, un valore di "Ciclo soglia" (Ct) che viene determinato nella fase esponenziale della reazione di amplificazione. Il Ct è il ciclo in cui si ha un segnale di fluorescenza 10 volte superiore al rumore di fondo. Esso è dipendente dal numero di copie iniziali di template ed è tanto più alto quanto minore è la concentrazione iniziale del template: la quantità iniziale di cDNA è quindi inversamente proporzionale al numero di cicli necessario per la rilevazione. Sebbene la metodica permetta una quantificazione assoluta, ai fini dello studio è stata sufficiente una quantificazione relativa, mettendo a confronto l'espressione dei singoli miR nella condizione/popolazione d'interesse (A) con quella nella condizione/popolazione di riferimento (B). Ciò è stato fatto con il metodo comparativo dei Ct (metodo $\Delta\Delta Ct$) basato sulla formula:

$$\Delta\Delta Ct = (Ct_{\text{target/trattato}} - Ct_{\text{housek./trattato}}) - (Ct_{\text{target/non trattato}} - Ct_{\text{housek./non trattato}})$$



Il rapporto numerico tra l'espressione del miR d'interesse nella condizione/popolazione A e l'espressione della stesso nella condizione di riferimento B è ricavabile dalla formula:

$$\text{(espressione del gene target nella condizione/popolazione A)} = 2^{-\Delta\Delta Ct} \times \text{(espressione del gene target nella condizione/popolazione B)}$$

Il valore di efficienza di reazione per tutte le coppie di *primer* utilizzate è compresa nell'intervallo 90-110%, condizione necessaria per applicare il metodo di analisi $\Delta\Delta Ct$.

3.7 Generazione di librerie di small-RNA

Le librerie di small-RNA sono state generate a partire dall'RNA totale estratto da campioni di timociti umani non frazionati, timociti DP e linfociti T maturi CD4⁺ e CD8⁺, ottenuti rispettivamente da frammenti operatori di timo umano o isolati dal sangue periferico di donatori sani, come descritto al paragrafo 3.1. Per la generazione di ciascuna libreria si è partiti da un *pool* di RNA totale derivato da almeno 3 diversi donatori.

Dall'RNA totale dei campioni, mediante corsa su un gel di poliacrilamide, si è separata la frazione degli small-RNA contenente RNA con peso molecolare compreso fra i 18 e i 26 nt. Gli small-RNA purificati sono stati modificati mediante l'aggiunta di oligonucleotidi adattatori all'estremità 3' e 5', retro-trascritti a cDNA ed infine sottoposti a diversi cicli di PCR con coppie di primer contenenti sequenze funzionali al successivo sequenziamento con la tecnologia del 454 Sequencing System (Roche).

Operativamente, inizialmente il *marker* (Decade Marker, Ambion) ed il *carrier* sono stati marcati radioattivamente con un'incubazione di 1 ora a 37°C in presenza di [γ -³²P] ATP e di T4 Polynucleotide Kinase (NEB). Per *carrier* si intende uno small-RNA sintetico di 23 nt che viene prima marcato radioattivamente e quindi unito al campione, permettendo in tal modo l'individuazione della banda degli small-RNA nel corso dei vari passaggi di corsa e purificazione su gel. Dopo la marcatura, *marker* e *carrier* sono stati purificati tramite passaggio attraverso ProbeQuant G-50 MicroColumns (GE Healthcare Life Sciences) e la loro qualità è stata controllata mediante corsa su un gel di poliacrilamide al 15% denaturante e autoradiografia.

Dieci-venti μ g di RNA totale di ciascun campione sono stati mescolati a 1 pmol di *carrier* marcato radioattivamente e ad un uguale volume di 2X RNA Loading Buffer (Ambion). L'RNA è stato poi caricato su gel di poliacrilamide al 15% denaturante e la frazione degli small-RNA è stata separata mediante corsa elettroforetica. Quindi è stata fatta una autoradiografia del gel che ha permesso di individuare e tagliare la banda di gel contenente gli RNA di peso molecolare compreso fra i 18 e i 26 nt. Gli small-RNA sono stati eluiti dal gel o.n. in acqua e NaCl 0,3 M; l'RNA è stato fatto precipitare incubando la soluzione per almeno 2 ore a -20°C con 2 volumi di etanolo assoluto e 10 μ g di glicogeno, quindi è stato lavato con etanolo 75% e risospeso in acqua sterile *nuclease-free*. La frazione degli small-RNA purificata è stata poi sottoposta a 2 successive reazioni di ligazione dei due oligonucleotidi ADATTATORI A e B (Roche)

rispettivamente all'estremità 3' e 5'. Al termine di ciascuna reazione, la frazione di small-RNA veniva fatta correre su gel di poliacrilamide al 12% denaturante. Dopo la corsa, il gel veniva sottoposto ad autoradiografia e venivano tagliate le bande corrispondenti agli small-RNA che avevano legato l'adattatore, prima all'estremità 3' (pesi molecolari dai 35 ai 45 nt) e poi all'estremità 5' (pesi molecolari dai 48 ai 62 nt). Dopo ogni reazione di ligazione e corsa elettroforetica la frazione degli small-RNA veniva recuperata come descritto sopra.

Successivamente, l'RNA è stato retro-trascritto con SuperScript II (Invitrogen) e quindi amplificato con Taq polymerase High Fidelity (Invitrogen) utilizzando una coppia di primer di fusione (#913 e #914, 10 µM, Roche) contenenti circa 20 bp complementari alla sequenza dell'ADATTATORE e 19 bp di una sequenza nota (PRIMER A o PRIMER B), che funge da sito di attacco dei primer durante l'amplificazione clonale ed il sequenziamento degli small-RNA.

Il prodotto di amplificazione (peso molecolare dai 57 ai 77 nt) è stato separato mediante corsa elettroforetica su gel di poliacrilamide al 12% non denaturante e purificato da gel in modo analogo a quanto descritto in precedenza. Dopo la purificazione, la frazione di small-RNA è stata sottoposta a due reazioni di digestione con l'enzima di restrizione Pac1 (NEB) per eliminare il *carrier* dal campione, intervallate da un'ulteriore reazione di PCR con una coppia di primers 454 (Roche) contenenti una sequenza nota e variabile da campione a campione che permette di distinguere i campioni che verranno poi amplificati e sequenziati insieme. Dopo le varie digestioni enzimatiche e dopo la PCR, l'RNA è stato purificato rispettivamente mediante corsa su gel mediante e precipitazione con etanolo assoluto e sodio acetato 3M. La frazione di small-RNA così purificata e modificata è stata utilizzata per il processo di amplificazione in emulsione e *massive parallel sequencing-by-synthesis* del sistema 454 [78]. Queste ultime due fasi della generazione della libreria sono state condotte dal laboratorio del Dott. DeBellis presso l'Istituto di tecnologie biomediche al CNR di Milano. La procedura impiegata è stata rappresentata nella figura 3.2. In breve, gli small-RNA sono denaturati a singolo filamento ed immobilizzati su specifiche DNA Capture Beads, che riconoscono le sequenze degli ADATTATORI introdotti alle estremità dell'RNA. Ogni bead lega un'unica sequenza della libreria. Le bead sono emulsificate con i reagenti di amplificazione in una miscela di acqua e olio e questo permette la formazione di micro-reattori contenenti ciascuno una bead legata ad un'unica sequenza, che viene amplificata clonalmente. In seguito, l'emulsione viene rotta e le biglie associate alle loro specifiche

sequenze clonalmente amplificate vengono caricate in una PicoTiterPlate per il *pyrosequencing*. Il diametro dei pozzetti della PicoTiterPlate permette l'entrata di una sola *bead* per pozzetto. Ai pozzetti sono aggiunti gli enzimi necessari alla reazione di *pyrosequencing*; successivamente il sistema fluidico del Genome Sequencer FLX introduce i singoli nucleotidi in ordine fisso nei diversi pozzetti. L'aggiunta al filamento in formazione di uno (o più) nucleotide(i) complementari alla sequenza della libreria risulta in un segnale chemiluminescente che è registrato dalla *CCD camera* dello strumento. La combinazione dell'intensità del segnale e della sua posizione nella PicoTiterPlate permette al software di registrare simultaneamente più di 1 milione di sequenze in 10 ore.

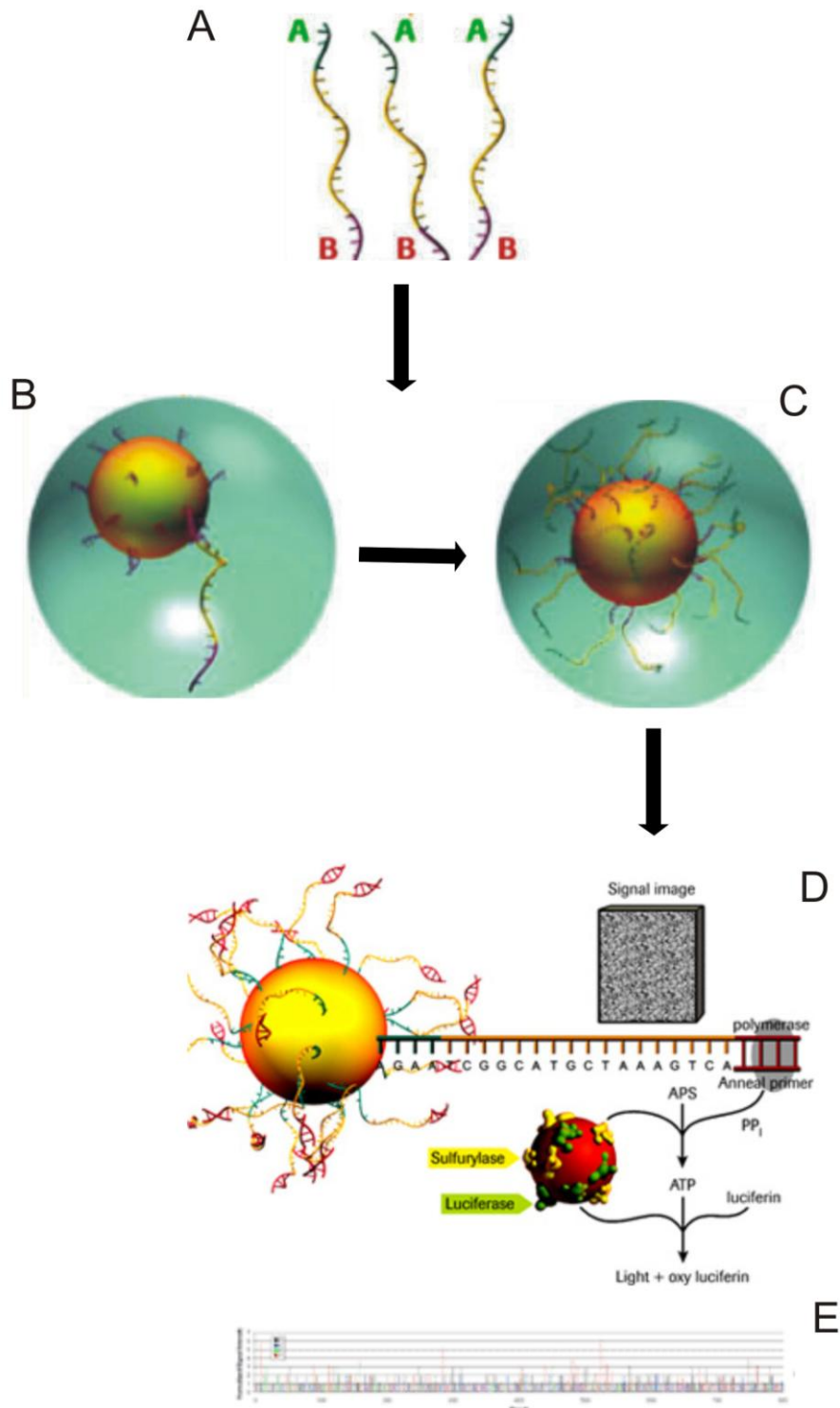


Figura 3.2. Preparazione delle librerie di small-RNA e sequenziamento con il sistema 454 (Roche).

La preparazione della libreria prevede la separazione della frazione degli small-RNA dall'RNA totale e l'aggiunta degli adattatori A e B alle estremità 3' e 5' (A). I singoli small-RNA sono legati a bead (B) ed amplificati clonalmente in micro-reattori mediante PCR in emulsione (C). I frammenti clonalmente amplificati, legati alle bead, sono caricati in una PicoTiter Plate e sequenziati mediante un procedimento di *massive pyrosequencing* (D). La combinazione di intensità e posizione del segnale, generate nella PicoTiter Plate, permette al software di determinare la sequenza di più di 100.000 frammenti all'ora.

3.8 Predizione computazionale degli mRNA bersaglio dei microRNA

La predizione computazionale dei potenziali geni target dei microRNA d'interesse è stata condotta integrando i risultati di tre diversi software di predizione computazionale: MiRanda (<http://www.microrna.org>), TargetScan (<http://www.targetscan.org>) e PicTar (<http://pictar.mdc-berlin.de>). Per elaborare la loro predizione, questi tre programmi sfruttano criteri e parametri di selezione in parte simili ed in parte distinti, quali:

- complementarità di sequenza fra microRNA e l'UTR-3' del mRNA bersaglio: in particolare, è di fondamentale importanza il perfetto appaiamento della *seed region* in 5' del microRNA (i.e. posizioni 2-8) con un sito complementare dell'UTR-3';
- conservazione inter-specie del sito di legame miR-UTR-3';
- energia libera di formazione del *duplex* miR:mRNA favorevole;
- struttura secondaria accessibile del mRNA bersaglio (in corrispondenza della regione contenente il sito di legame del miR).

Di conseguenza questi approcci computazionali possono generare liste di predizione solo parzialmente sovrapponibili, poichè un candidato identificato come potenziale bersaglio secondo un criterio di predizione può essere eliminato da un altro filtro che utilizza diversi valori di soglia o che considera parametri differenti.

Per ovviare questo problema, i risultati ottenuti dai tre algoritmi sono stati quindi integrati tra loro, individuando gli mRNA bersaglio comuni a tutte e tre le predizioni.

Dall'analisi dei *chip* di espressione genica è stata ottenuta una lista di mRNA differenzialmente regolati (FDR=0,001) nelle sottopopolazioni timiche considerate.

In seguito, per i primi 17 miR up-regolati ed i primi 2 miR down-regolati con maggiore *fold-change* nel corso della differenziazione dei timociti da DP a SP, è stata generata una lista di target predetti computazionalmente utilizzando Microcosm, *web resource* di miRbase che utilizza per la predizione di target il software MiRanda. I target predetti sono stati analizzati con R e Bioconductors per l'identificazione di mRNA differenzialmente regolati nei GEP (*Gene Expression Profile*) in senso opposto rispetto ai rispettivi miR (a cui erano stati precedentemente associati dall'analisi bioinformatica) nelle corrispondenti popolazioni.

Per confermare il risultato ottenuto per il bersaglio d'interesse Notch3 (o *Candidate Target 1*), questo particolare gene è stato inserito in ciascuno dei tre programmi di predizione sopracitati allo scopo di predire miRNA con cui esso potesse interagire. Una volta riconfermata in questo modo l'associazione fra Notch3 ed miR-150, l'mRNA *full-length* di Notch3 umano (uc002nan.1) e murino (uc008bvx.1) è stato ottenuto dal database UCSC, mentre la sequenza di miR-150 maturo è stata scaricata da miRBase (<http://www.mirbase.org>). La regione UTR-3' della forma umana e murina di Notch3 è stata estratta e convertita in formato FASTA. Quindi è stata studiata la sovrapposizione e la conservazione inter-specie dei siti di appaiamento con l'UTR-3' di Notch3 predetti per il miR d'interesse da MiRanda, PicTar e TargetScan.

3.9 Analisi del profilo d'espressione genica (GEP) mediante microarray

La possibilità di quantificare contemporaneamente l'espressione di numerosi geni è la caratteristica fondamentale della tecnologia dei microarray. L'mRNA estratto dalle cellule viene marcato e quindi ibridato a sequenze di DNA immobilizzate su una superficie solida. I sistemi principalmente utilizzati si dividono in due gruppi: microarray a cDNA e microarray ad oligonucleotidi. Nei microarray a cDNA le sequenze che vengono immobilizzate su un vetrino o su una membrana di nylon sono prodotti di PCR ottenuti da librerie di cDNA. Nei microarray a oligonucleotidi sequenze di 20-25 nucleotidi vengono sintetizzate direttamente sul supporto solido tramite una tecnica di fotolitografia, sviluppata da Affymetrix, che permette di sintetizzare contemporaneamente centinaia di migliaia di diversi oligonucleotidi. La sintesi *in situ* degli oligonucleotidi è caratterizzata da un'elevata riproducibilità che permette di confrontare segnali ottenuti da campioni ibridati su distinti microarray. Per questo motivo, si è scelto di utilizzare la tecnologia Affimetrix per lo studio dell'espressione genica dei timociti umani non frazionati e delle due sottopopolazioni timocitarie DP e SP CD4. Ai fini di tale analisi, l'RNA totale dei campioni è stato estratto in Trizol Reagent (Invitrogen), purificato da eventuali contaminanti fenoliche e controllato per qualità e quantità come descritto ai punti 3.3 e 3.4. Per ogni popolazione timica sono stati ibridati su array da 3 a 6 diversi replicati biologici. Secondo il protocollo Affimetrix, il cDNA a doppio filamento è stato sintetizzato a partire da almeno 1 µg di RNA totale usando il one-Cycle cDNA Synthesis Kit (Affymetrix, Santa Clara, CA) e

oligonucleotidi poly-dT contenenti sequenze promotrici riconosciute dalla T7 RNA polimerasi (GensetOligos). Il cDNA è stato poi usato come template per sintetizzare cRNA biotinilato tramite IVT Labeling Kit Kit (Affymetrix, Santa Clara, CA). Il cRNA così ottenuto è stato poi purificato con Rneasy Kit (Qiagen) e frammentato secondo le indicazioni del protocollo Affymetrix. Il cRNA biotinilato e frammentato è stato quindi ibridato su microarray Human genome U-133 PLUS (Affymetrix). Ibridazione, lavaggi e marcatura con streptoavidina coniugata a phycoeritrina sono stati condotti automaticamente nella gene-Chip Fluidic Station (Affymetrix). L'intensità di fluorescenza è stata rilevata da uno scanner ed elaborata tramite GeneChip software (Affymetrix). Questa analisi è stata condotta presso il laboratorio del Dott. Rosolen (Dipartimento di Pediatria, Università di Padova).

3.10 Produzione di vettori lentivirali

La produzione dei vettori lentivirali per l'over-espressione del miR-150 (LV-pre-miR-150) e dei relativi vettori di controllo (LV-control) è stata condotta utilizzando come linea di *packaging* la linea cellulare 293T, in virtù della sua elevata trasfettabilità. I vettori lentivirali d'interesse sono stati prodotti mediante trasfezione transiente di questa linea cellulare con tre plasmidi, utilizzando il metodo della precipitazione con $\text{Ca}_3(\text{PO}_4)_2$. In questa tecnica il DNA plasmidico viene mescolato con una soluzione 2M di CaCl_2 e con un *buffer* fosfato in modo da formare un fine precipitato che viene disperso sopra la coltura cellulare. I vettori di trasferimento prodotti appartengono alla generazione dei vettori lentivirali basati sul virus dell'immunodeficienza umana (HIV), detti SIN (*self inactivating*). Per la produzione di tali vettori sono stati utilizzati diversi plasmidi:

- il plasmide di trasferimento che costituisce il genoma del vettore stesso e che codifica il transgene di interesse;
- il costrutto di *packaging* gag-polgpt (p8.74), che ha la funzione di "impacchettare" l'RNA genomico virale;
- per la pseudotipizzazione delle particelle virali è stato utilizzato il plasmide pHCMV-G che esprime la proteina G del virus della stomatite vescicolare (VSV-G) sotto il controllo trascrizionale del promotore del citomegalovirus (CMV). Il tropismo del vettore è strettamente dipendente dalle proteine dell'envelope che interagiscono con i recettori cellulari; VSV-G permette di ampliare lo spettro

d'ospite del virus, infatti i suoi recettori sono fosfolipidi di membrana ubiquitari, presenti nella maggioranza delle cellule eucariotiche.

In breve, 24 ore prima della trasfezione, le cellule 293T sono state piastrate in fiasche da 75 cm² alla concentrazione di 3,5-4 x 10⁶ cellule/fiasca in 10 ml di terreno DMEM completo. Il giorno seguente le cellule sono state trasfettate con le seguenti quantità di DNA: 12 µg del costrutto d'espressione d'interesse, 6 µg del plasmide che esprime i geni *gag* e *pol* (p8.74) e 2 µg del plasmide codificante per la proteina G del virus della stomatite vescicolare (pHCMV-G). Al mix di DNA plasmidico sono stati aggiunti 62 µl di CaCl₂ (Fluka), portando ad un volume finale di 500 µl con H₂O sterile. La soluzione è stata aggiunta, goccia a goccia, a 500 µl di tampone HBS 2x (274 mM NaCl, 10mM KCl, 1,5 mM Na₂HPO₄*7H₂O, 12mM Destrosio, 42 mM HEPES, pH 7,1). Il mix di trasfezione è stato incubato per 20 minuti a temperatura ambiente; quindi la soluzione è stata aggiunta, goccia a goccia, al terreno di coltura delle cellule, che sono poi state riposte in incubatore a 37°C. Dopo 12-18 ore il terreno di coltura è stato rimosso per eliminare il calcio fosfato, ed è stato sostituito con 7 ml di terreno fresco di coltura. Il giorno dopo il surnatante virale è stato raccolto, centrifugato per 10 minuti a 3000 rpm a 4°C e filtrato con filtri dai pori di 0,45 µm di diametro, mentre alle 293T sono stati aggiunti altri 7 ml di terreno fresco alla coltura. Il giorno seguente si è proceduto con la seconda raccolta del surnatante virale, analogamente a quanto descritto sopra. I surnatanti virali raccolti sono stati concentrati 35 volte mediante centrifugazione a 24000 rpm per 2 ore a 4°C. Il pellet è stato poi risospeso in terreno DMEM semplice e conservato a -80°C.

Costrutti virali d'espressione:

- **Human pre-miR Expression Construct miR-150** (SBI), chiamato successivamente nel testo LV-pre-miR-150;
- **pCDH-CMV-MCS-EF1-copGFP** (SBI), chiamato successivamente nel testo LV-control.

La struttura dei due costrutti d'espressione è mostrata nella figura 3.3.

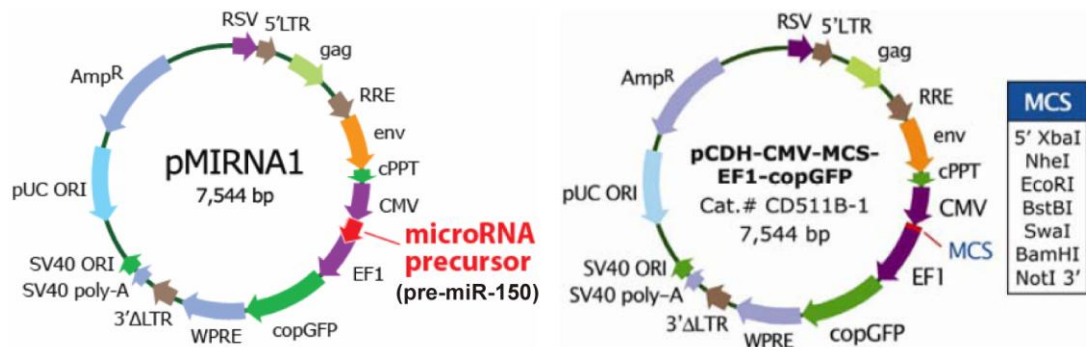


Figura 3.3. Mappa dei plasmidi d'espressione *HIV-based* Human pre-miR Expression Construct miR-150 (sinistra) e CDH-CMV-MCS-EF1-copGFP (destra).

3.11 Trasduzione cellulare *in vitro* mediante vettori lentivirali

Le cellule di T-ALL Jurkat sono state trasdotte *in vitro* con un vettore lentivirale codificante per il pre-miR-150 (LV-pre-miR-150) e con il relativo vettore di controllo (LV-control) (vedi punto 3.10). In breve, venivano piastrate 10^6 cellule di T-ALL per pozzetto in un volume finale di 1 ml di terreno RPMI completo in piastre di coltura da 6 pozzetti. In ogni pozzetto, al terreno completo erano aggiunti 100 μ l (LV-pre-miR-150) o 300 μ l (LV-control) di preparazione virale in presenza di solfato di protamina ad una concentrazione di 8 μ g/ml per facilitare l'infezione [79]. Dopo 14-18 ore, le cellule venivano lavate con terreno semplice e risospese in 3 ml di terreno RPMI completo. Dopo 48 ore, le cellule trasdotte con i vettori codificanti GFP erano osservate al microscopio a fluorescenza o al citofluorimetro per valutare l'efficienza di trasduzione in base alla positività per GFP. Inoltre, a 48-72 ore dalla trasduzione circa 10^6 cellule trasdotte venivano utilizzate per l'estrazione dell'RNA e lo studio tramite TaqMan real-time PCR dei livelli di espressione di miR-150 (come descritto al punto 3.6). Nei casi in cui la trasduzione con i vettori codificanti il pre-miR-150 ed i vettori di controllo era stata efficace (> 80% delle cellule era positiva per GFP), le cellule venivano utilizzate per vari saggi funzionali (saggi di proliferazione e analisi del ciclo cellulare), descritti più avanti nei Materiali e Metodi. Un pozzetto di cellule trasdotte veniva utilizzato per la conta cellulare a 72 ore dalla trasduzione: un milione di cellule era seminato per la trasduzione (analogamente a quanto detto sopra), quindi dopo 72 ore le cellule LV-pre-

miR-150 e LV-control venivano raccolte e contate al microscopio, al fine di valutare l'aumento del numero di cellule in coltura rispetto al tempo zero.

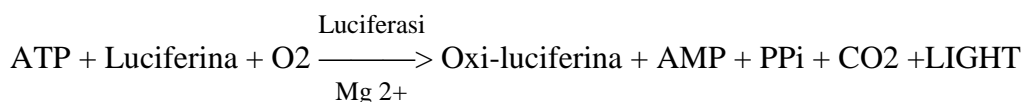
3.12 Controllo citofluorimetrico dell'efficienza di trasduzione

L'efficienza di trasduzione con i vettori lentivirali LV-pre-miR-150 e LV-control nella linea cellulare Jurkat è stata misurata, a 48 ore dalla trasduzione, mediante la lettura della percentuale di cellule positive per GFP rispetto ad un campione di cellule Jurkat *wild-type*. Questo è stato possibile grazie al fatto che entrambi i plasmidi d'espressione della SBI utilizzati per la produzione dei vettori virali (Human pre-miR Expression Construct miR-150 e pCDH-CMV-MCS-EF1-copGFP) contengono il gene reporter della GFP (*Green Fluorescent Protein*). Di conseguenza, tutte le cellule che avevano incorporato le particelle virali risultavano positive per GFP.

3.13 Analisi della proliferazione cellulare mediante test dell'ATP.

La proliferazione delle cellule Jurkat trasdotte con il vettore lentivirale esprime il pre-miR-150 o con il vettore di controllo è stata valutata utilizzando il kit ViaLight HS Kit (Cambrex Bioscience), in accordo con le istruzioni del produttore. Il kit si basa sulla misura, mediante bioluminescenza, dell'ATP (Adenosina Trifosfato) presente nelle cellule metabolicamente attive. Tutte le cellule devono produrre ATP per restare vitali e svolgere le loro funzioni fisiologiche; quindi i livelli di ATP cellulare sono un indice dello stato metabolico e proliferativo delle cellule.

Il metodo bioluminescente si basa sull'emissione di luce, prodotta dalla seguente reazione catalizzata dall'enzima luciferasi:



L'intensità della luce emessa è linearmente correlata alla concentrazione di ATP, la quale a sua volta è correlata al numero di cellule metabolicamente attive presenti in coltura. Infatti, dato che ogni cellula contiene una certa quantità di ATP, l'aumento dei livelli di ATP in una coltura cellulare nel corso del tempo indica l'aumento del numero di cellule vitali e dunque è un indice di proliferazione cellulare. La quantificazione

bioluminescente dei livelli di ATP cellulari è un saggio frequentemente impiegato per la valutazione del tasso proliferativo delle cellule anche grazie alla sua elevata sensibilità e al suo ampio range dinamico.

Per il test, le cellule Jurkat, a distanza di 48 ore dalla trasduzione con i vettori virali, sono state seminate in una piastra da 96 pozzetti ad una concentrazione di 5.000 cellule/pozzetto in 100 µl di terreno RPMI 1640 completo. La misurazione dei livelli di ATP nella coltura è stata fatta a 3, 24, 48 e 72 ore dalla semina. La prima misurazione, è stata effettuata a 3 ore dalla semina in modo da valutare i livelli basali di ATP ed eventuali differenze nel numero di cellule piastrate. Il saggio è stato condotto a temperatura ambiente (20-25°C), la temperatura ottimale per l'attività enzimatica della luciferasi. Al momento della misurazione, ad ogni pozzetto sono stati aggiunti 100 µl di NRR (*Nucleotide Releasing Reagent*). La soluzione è stata lasciata agire per 5 minuti in modo da permettere la lisi cellulare ed il rilascio dell'ATP. Quindi a ciascun pozzetto si sono aggiunti 20 µl di AMR (*ATP Monitoring Reagent*) e subito dopo la piastra è stata letta. I livelli di luminescenza emessi da ogni pozzetto al termine della reazione sono stati rilevati da un Top Scintillation Counter (Beckman Coulter, Fullerton, CA). Il tasso di proliferazione cellulare dei diversi campioni studiati è stato espresso in termini di intensità assoluta di luminescenza misurata in cpm (i.e. colpi per minuto).

3.14 Analisi del ciclo cellulare con Ioduro di Propidio

Per analizzare il profilo del ciclo cellulare, 10⁶ cellule Jurkat trasdotte (come descritto al punto 3.11) con il vettore LV-pre-miR-150 o LV-control sono state tenute in coltura per 3 giorni in terreno RPMI completo in una piastra da 6 pozzetti. Successivamente, le cellule sono state raccolte, lavate con PBS, contate, risospese in 1 ml di soluzione GM (glucosio 1.1 mM, NaCl 0.14 M, KCl 5 mM, Na₂HPO₄ 1.5 mM, KH₂PO₄ 1.1 mM, EDTA 0.5 mM) e fissate con l'aggiunta goccia a goccia di 3 ml di etanolo al 100% ogni 1-2 milioni di cellule. Le cellule sono state quindi centrifugate per 10 minuti a 1200 rpm, lavate con PBS 1X, ed infine risospese in 500 µl di PI (100µg/ml), contenente RNasi priva di DNasi (12 µg/ml) (Sigma). Dopo un'ora di incubazione a temperatura ambiente le cellule sono state analizzate al citofluorimetro usando un laser Argon di lunghezza d'onda di 488 nm e una linea di rivelazione FL2-A. Gli istogrammi relativi al contenuto di DNA sono stati ottenuti usando il software ModFit LT3.0 (Verity Software House).

3.15 Analisi statistiche

I risultati della real-time PCR e dei test funzionali su miR-150 sono stati espressi come valore medio \pm DS (Deviazione Standard). Le analisi statistiche sono state effettuate usando il test t di Student. Le differenze sono state considerate statisticamente significative con $P < 0.05$.

3.15.1 MiR microarray

Per la generazione dei dendrogrammi a partire dal profilo d'espressione dei miR nei campioni, è stato utilizzato un algoritmo di *clustering* gerarchico basato sul metodo della correlazione di Pearson e dell'*average-linkage*, fornito dal *software Genesis* (usato per il *clustering* delle sole popolazioni timiche) e dalla piattaforma *geWorkbench* (usata per il *clustering* dei campioni di T-LBL insieme ai LN e alle popolazioni timiche d'interesse). I miR differenzialmente espressi nei confronti fra le diverse popolazioni cellulari ibridate su miR microarray sono stati individuati mediante analisi con l'algoritmo SAM (*Significant Analysis of Microarrays*). Le differenze di espressione sono state considerate statisticamente significative con $FDR=0.001$.

3.15.2 Librerie di small-RNA

Per l'individuazione dei miR nuovi potenziali e noti fra le sequenze delle librerie di small-RNA, è stata sviluppata una *pipeline* bioinformatica illustrata nei Risultati. La completezza delle librerie di small-RNA nel rappresentare la totalità dei miR espressi dai campioni è stata stimata mediante un'analisi statistica *bootstrap* [80,81]. I miR differenzialmente espressi nelle diverse librerie sono stati individuati applicando il Fisher Exact test alle conte dei singoli miR noti presenti nelle librerie. Le differenze di espressione sono state considerate statisticamente significative con $P < 0.05$.

3.15.3 GEP array

L'elaborazione dei dati dei chip d'espressione genica è stata fatta con i software Genesis (<http://genome.tugraz.at/>) e BRB Array Tools (<http://linus.nci.nih.gov/pilot/index.htm>). Per determinare la regolazione differenziale dei geni si è utilizzata una $FDR=0,001$.

4. RISULTATI

4.1 Profilo d'espressione dei microRNA in sottopopolazioni timocitarie umane a diversi stadi di maturazione

Il profilo d'espressione dei microRNA (miRNA) nel corso dello sviluppo fisiologico delle cellule T nell'uomo è un argomento ancora in gran parte sconosciuto. Al fine appunto di identificare miR coinvolti nella differenziazione delle cellule T nell'uomo, abbiamo analizzato, mediante ibridazione su microarray, il profilo di espressione dei miR in timociti umani non frazionati ed in sottopopolazioni timocitarie umane a 3 diversi stadi di maturazione (Doppi Positivi o DP, Singoli Positivi CD4⁺ o SP CD4, Singoli Positivi CD8⁺ o SP CD8). La sottopopolazione timocitaria Doppia Negativa (DN; CD4⁻/CD8⁻), che rappresenta la fase più immatura dello sviluppo timico, è stata esclusa dall'analisi, poiché essa costituisce una percentuale molto bassa delle cellule del timo umano (2-4%) ed è perciò difficilmente isolabile in quantità sufficienti ed a livelli accettabili di purezza ai fini dell'esperimento.

Le popolazioni timiche d'interesse sono state ottenute da frammenti operatori di timo umano sano prelevati da pazienti pediatrici sottoposti ad interventi chirurgici per difetti cardiaci congeniti. I timociti sono stati isolati mediante dissezione meccanica del tessuto timico e definiti per la positività all'antigene CD2, un marcatore precoce delle cellule T umane. In ogni caso, più del 90% delle cellule estratte dal timo (con la procedura descritta al punto 3.2 dei Materiali e Metodi) è risultata positiva alla marcatura per l'antigene CD2.

Le popolazioni timocitarie d'interesse per lo studio (timociti SP CD4, CD2⁺/CD4⁺/CD8⁻; SP CD8, CD2⁺/CD4⁻/CD8⁺; DP, CD2⁺/CD4⁺/CD8⁺) sono state individuate mediante tripla marcatura per i co-recettori CD4 e CD8 e per l'antigene CD2 e separate tramite *FACS sorting*. La procedura di raccolta e separazione è stata ottimizzata in modo da ottenere delle popolazioni cellulari ad alti livelli di purezza ($\geq 95\%$).

Per l'ibridazione su *microarray*, l'RNA totale dei campioni è stato estratto in *Trizol*, purificato da eventuali contaminanti fenoliche, quantificato allo spettrofotometro e controllato mediante elettroforesi capillare. Per alcune popolazioni cellulari poco rappresentate è stato necessario fare *pool* dell'RNA ottenuto dai timociti di diversi pazienti.

I profili d'espressione dei miR per i campioni timocitari sono stati infine generati utilizzando l'*Agilent Human miRNA Microarray (v2.0) platform* (Agilent Technologies), contenente sonde in grado di riconoscere le sequenze di 723 microRNA umani e 76 microRNA virali umani, derivate dalla versione 10.1 del *Sanger database*. Per ogni popolazione sono stati ibridati su array sei replicati biologici. Dei 723 miR umani monitorati, 237 sono stati identificati come espressi nei timociti, secondo un processamento dei dati che impone la presenza di segnale d'ibridazione per i singoli miR almeno nel 50% dei *chip*.

Il *clustering* gerarchico del *pattern* d'espressione dei miR nei campioni (realizzato mediante un algoritmo di *clustering* gerarchico fornito dal software Genesis e basato sul metodo della correlazione di Pearson) ha mostrato che le diverse popolazioni timocitarie presentano distinti profili d'espressione di miR. Inoltre, tale differenziale espressione dei miR riflette le relazioni fra i diversi stadi maturativi dei timociti (fig. 4.1).

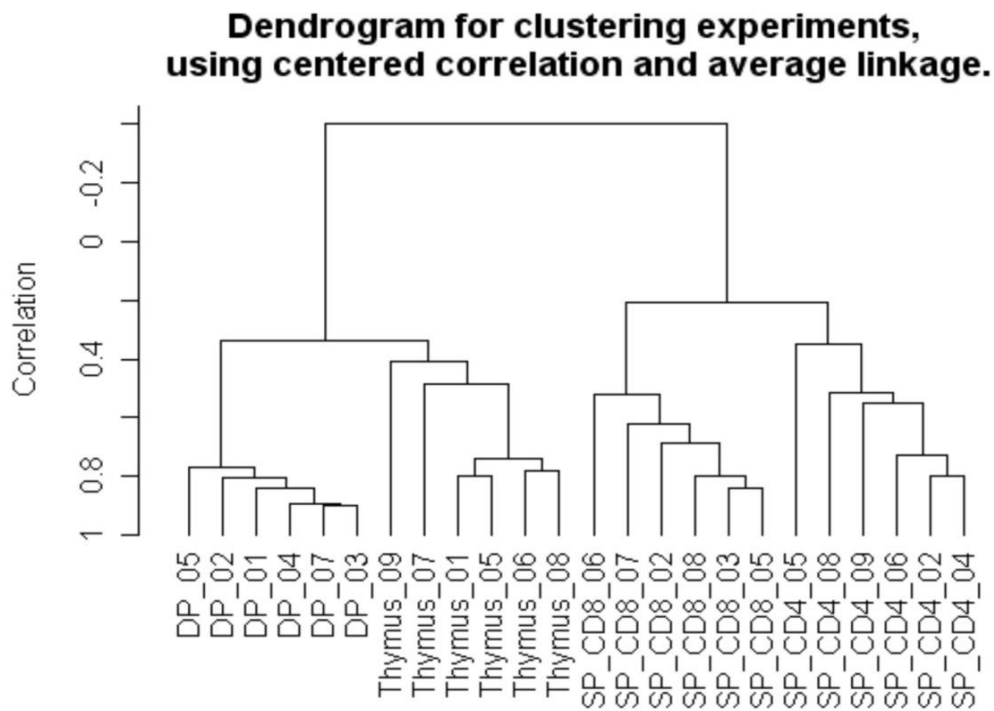


Figura 4.1. Dendrogramma derivato dal *clustering* gerarchico dei profili d'espressione dei miR nelle popolazioni timiche. L'analisi dei dati degli array indica che ogni popolazione timica presenta un profilo d'espressione dei miR caratteristico e distinto, che riflette le relazioni fra gli stadi di sviluppo dei precursori T.

Infatti, i DP presentano un profilo d'espressione dei miR che si discosta molto da quello di entrambe le popolazioni mature SP, le quali invece presentano *pattern* d'espressione più simili fra loro, anche se comunque distinti. Il timo totale si trova in una posizione intermedia, presentando un profilo distinto sia dalla popolazione DP che dalle più mature popolazioni SP, anche se più vicino a quello delle cellule DP. Questo riflette la composizione della popolazione timica umana che è costituita per una percentuale pari a circa 60-80% da cellule DP.

Mediante analisi SAM (*Significant Analysis of Microarrays*) sono stati individuati complessivamente 70 miR differenzialmente regolati (FDR=0,001) nel corso del passaggio dei precursori T dallo stadio di DP a quello di SP (fig. 4.2).

In particolare, 60 miR risultano up-regolati in entrambe le popolazioni SP rispetto alla popolazione DP, mentre solo 10 miR vengono down-regolati durante questa fase della maturazione. Questo dato sembra indicare che nel corso della maturazione T vi sia un generale aumento dell'espressione dei miR. La differenziazione cellulare avviene tramite l'attivazione di un programma genetico che prevede, da una parte il silenziamento di geni legati alla staminalità, dall'altra l'espressione di un gruppo di geni che determinano la specializzazione cellulare. Poiché i miR sono molecole con una funzione soppressoria, in questo contesto la loro up-regolazione potrebbe avere il ruolo di spegnere l'espressione di geni legati alla staminalità, favorendo la differenziazione cellulare.

Oltre ad essere in numero maggiore, i miR up-regolati nei timociti durante la transizione dallo stadio di DP a quello di SP presentano un intervallo di *fold-change* (FC) in generale più ampio rispetto al gruppo di miR che vengono down-regolati in questo stesso passaggio. Infatti, mentre i miR differenzialmente up-regolati nei SP rispetto ai DP hanno FC variabili fra le 61 e le 1,4 volte, nei miR down-regolati si osservano variazioni di FC più modeste, comprese fra le 0,3 e le 0,7 volte. In particolare, 16 miR risultano up-regolati con FC maggiore di 5, mentre solo 7 miR sono down-regolati con FC minore di 0,5 nei timociti SP rispetto ai DP. Nell'istogramma della figura 4.3 sono rappresentati i 10 miR up-regolati e i 5 miR down-regolati con FC più elevato nel passaggio dei precursori T da DP a SP.

Confrontando le due popolazioni SP CD4 e CD8 fra loro, sono stati osservati 15 miR differenzialmente regolati (FDR=0,001), di cui 8 risultano up-regolati e 7 down-regolati nelle cellule SP CD4 rispetto alle SP CD8 (fig. 4.4).

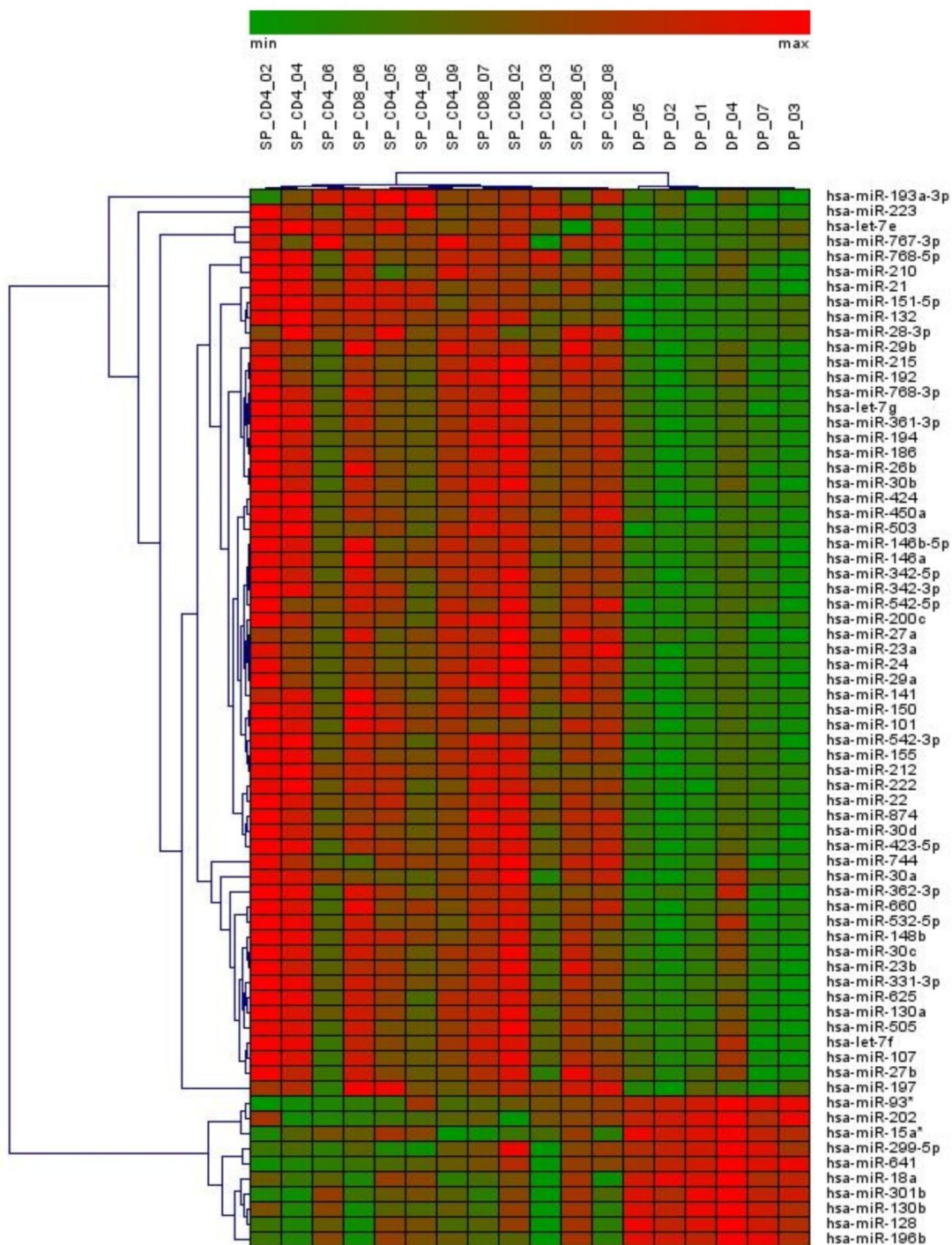


Figura 4.2. *Heatmap* rappresentante l'espressione differenziale dei miR in timociti umani SP rispetto a timociti DP. Il profilo d'espressione dei miR è stato ottenuto utilizzando l'Human miRNA microarray kit v.2.0 (Agilent Technologies) ed i dati grezzi sono stati analizzati con il Feature Extraction Software 9.5.3.1 (Agilent Technologies). Settanta miR sono stati identificati mediante analisi SAM come differenzialmente espressi in entrambe le popolazioni timocitarie mature SP CD4 e SP CD8 rispetto ai più immaturi timociti DP (FDR=0.001).

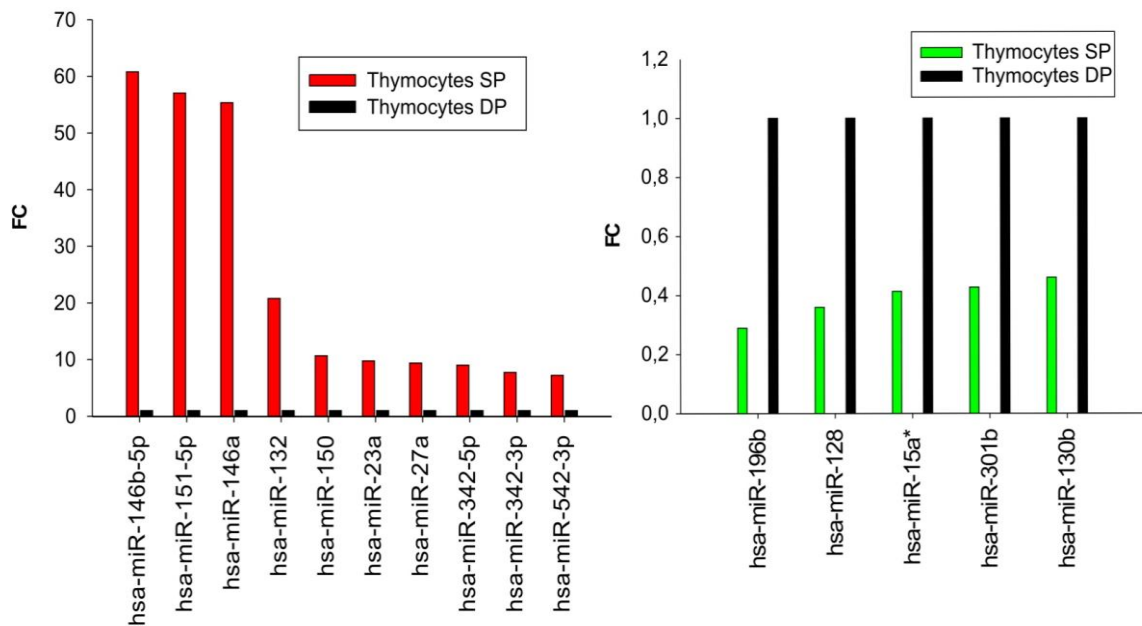


Figura 4.3. MiR up-regolati (sinistra) e down-regolati (destra) con FC più elevati risultati dal confronto fra timociti SP e DP nell'analisi dei microarray. Gli istogrammi mostrano, per i diversi miR, il *fold-change*, calcolato come rapporto fra il valore medio di intensità del segnale di ibridazione di entrambe le popolazioni SP CD4 e CD8 (trattate come un'unica popolazione) ed il valore medio di intensità della popolazione DP. I miR scelti per la rappresentazione in questa figura sono stati selezionati dalla lista dei miR differenzialmente espressi nel confronto fra SP e DP con FDR=0,001.

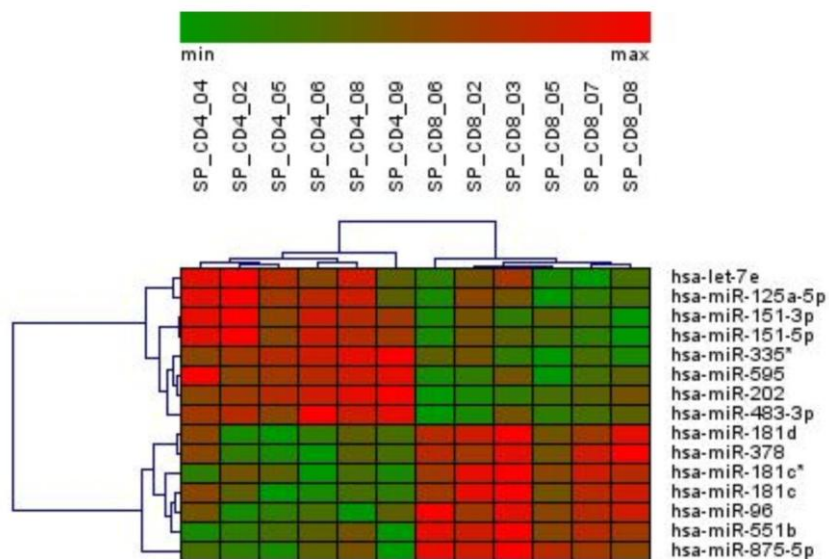


Figura 4.4. Heatmap rappresentante l'espressione differenziale dei miR in timociti umani SP CD4 e SP CD8. Quindici miR sono stati identificati mediante analisi SAM come differenzialmente espressi nel confronto fra le due popolazioni timocitarie mature SP CD4 e SP CD8 (FDR=0.001).

4.2 Generazione di librerie di small-RNA da sottopopolazioni linfoidi T

In parallelo, allo scopo di approfondire la nostra conoscenza sull'espressione dei microRNA nella linea linfoide T, sono state generate delle librerie di small-RNA a partire dall'RNA totale estratto da campioni di timociti umani non frazionati, timociti immaturi DP e linfociti T maturi CD4⁺ (LT CD4) e CD8⁺ (LT CD8), ottenuti rispettivamente da frammenti operatori di timo umano o isolati dal sangue periferico di donatori sani. Per la generazione di ciascuna libreria si è partiti da un *pool* di RNA totale derivato da almeno 3 diversi donatori. Dall'RNA totale dei campioni è stata purificata la frazione degli small-RNA contenente RNA con peso molecolare compreso fra i 18 e i 26 nt. Gli small-RNA purificati sono stati modificati mediante l'aggiunta di oligonucleotidi adattatori all'estremità 3' e 5', retro-trascritti a cDNA ed infine sottoposti a diversi cicli di PCR con coppie di *primer* contenenti sequenze funzionali al successivo sequenziamento con la tecnologia del *454 Massive Parallel Sequencing System (Roche)*. La procedura è stata spiegata in dettaglio al punto 3.7 di Materiali e Metodi. La generazione di librerie di small-RNA ci ha permesso non solo di studiare l'espressione di miR noti, ma anche di identificare candidati nuovi miR arricchiti specificamente nelle diverse popolazioni di cellule T.

4.2.1 Generazione della pipeline bioinformatica

Al fine di identificare i miR noti e nuovi potenziali fra le sequenze delle librerie di small-RNA, è stata sviluppata una *pipeline* bioinformatica (illustrata nella figura 4.5) costituita da una serie di filtri computazionali molto stringenti, in modo da ridurre al minimo il numero di falsi positivi.

Questa procedura bioinformatica si basa su protocolli analitici sviluppati in precedenza per l'interpretazione dei dati di sequenziamento ottenuti da librerie di small-RNA [82] e su una serie di assunzioni derivanti da articoli recentemente pubblicati in letteratura.

In breve, l'analisi bioinformatica delle librerie comprende una prima parte di "pulizia dei dati", durante la quale sono rimossi dalle sequenze *tag*, *primer* e *carrier* (introdotti per la generazione della libreria); le sequenze degli small-RNA sono selezionate sulla base delle dimensioni, che devono essere comprese fra i 18 ed i 25 nt; sono identificate le sequenze ridondanti con il software RepeatMasker.

Le sequenze non ridondanti di small-RNA così ottenute sono state allineate, sfruttando l'algoritmo BLAST, con diversi *database* di microRNA umani e di altre specie (miRBase hairpins e miRBase mature). In questo modo è possibile individuare fra le sequenze delle librerie la presenza di miR noti e conservati in diverse specie. In seguito, le sequenze degli small-RNA sono state allineate con l'algoritmo BLAST sul genoma umano ed i *loci* così individuati sono stati comparati con la posizione genomica di miR noti. Si è deciso di ammettere al massimo 1 nt di *mismatch* e 0 nt di *gap* nell'allineamento delle sequenze con il genoma umano per ammortizzare eventuali errori tecnici nel sequenziamento, o modificazioni/addizioni nucleotidiche a livello delle estremità dei miR. Infatti, in diversi lavori recenti [82,83], è stato visto che *in vivo* i miR possono subire delle modificazioni a livello delle loro estremità (in particolare a livello dell'estremità 3'), in seguito al processamento di Drosha o Dicer, con l'aggiunta in ultima posizione comunemente di un'adenosina o di un'uracile.

Per le sequenze di small-RNA non corrispondenti a miR già annotati nei database *online*, è stata condotta una ricerca sul genoma per individuare la presenza di putative sequenze precursori fiancheggiando la posizione genomica dello small-RNA. A questo scopo, siamo andati a vedere se nelle 80 basi a valle e a monte del *match* genomico si trovassero sequenze in grado di dare origine a strutture secondarie a forcina. Le strutture secondarie dei precursori predetti sono state studiate per la loro energia libera, considerando accettabili valori inferiori alla massima energia libera di *folding* osservata per i miR annotati in miRBase.

Gli small-RNA la cui posizione genomica cadeva al livello dell'ansa della struttura a forcina del precursore predetto, sono stati scartati.

In parallelo, è stata condotta un'analisi per individuare ed assimilare fra loro sequenze di small-RNA (con 0 o 1 *mismatch*) che mappavano nelle stesse regioni del genoma ed avevano un precursore predetto comune.

In letteratura è noto che spesso i miR sono raggruppati in *cluster* di due o più miR che vengono trascritti da un promotore comune, generando un trascritto primario policistronico. Partendo da questo assunto, i candidati nuovi miR sono stati testati per l'appartenenza ad un *cluster* contenente due o più miR (noti o nuovi putativi).

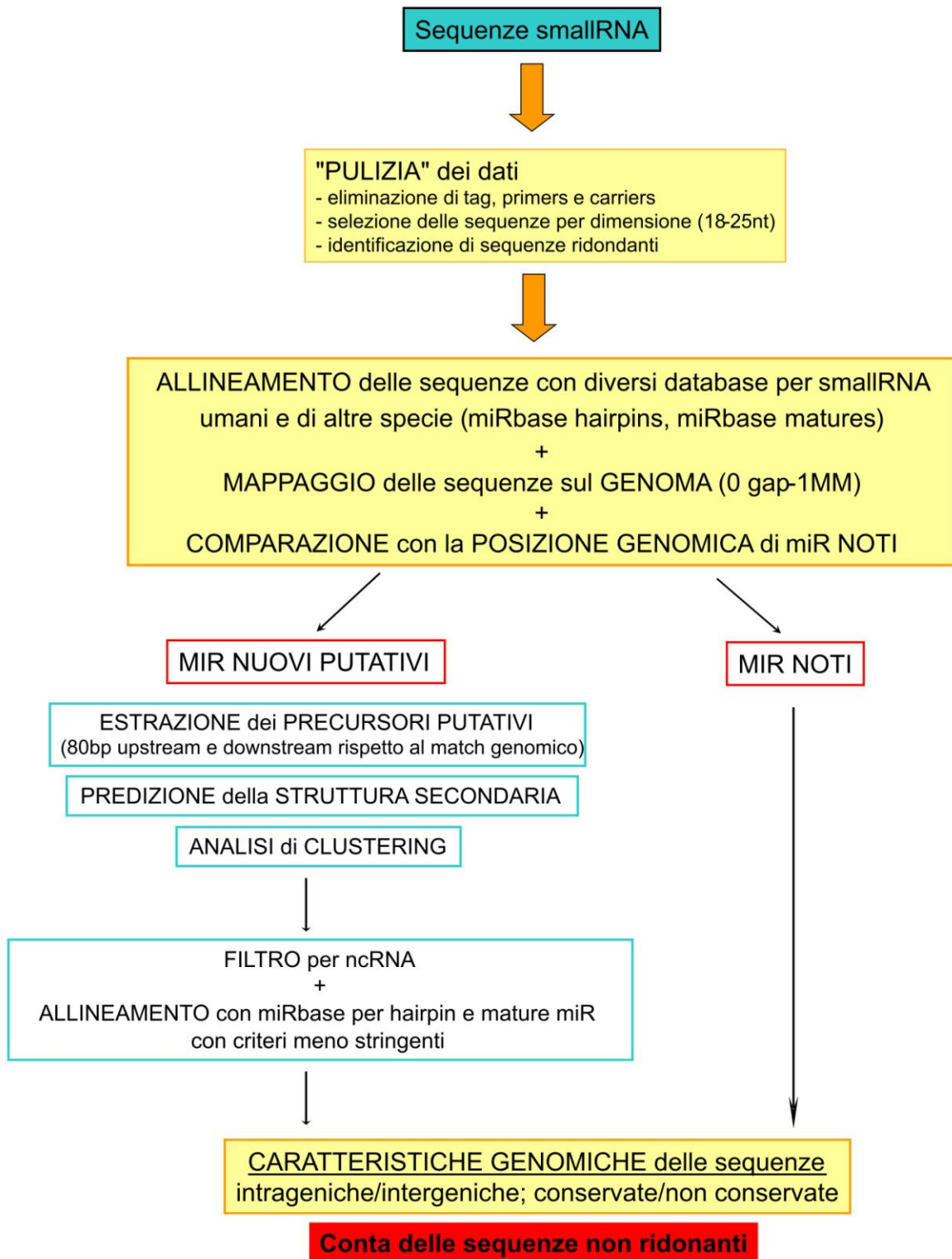


Figura 4.5. Pipeline bioinformatica utilizzata per la definizione dei miR noti e nuovi presenti nelle librerie di small-RNA. Nella figura sono illustrati i principali filtri computazionali che sono stati applicati ai prodotti del sequenziamento della frazione di small-RNA delle 4 popolazioni linfoidi T studiate.

Dalla lista dei nuovi candidati miR che avevano superato i precedenti filtri sono stati eliminati gli small-RNA le cui sequenze si allineavano con altri non-coding-RNA presenti nei *database on-line* (snoRNA, piRNA, tRNA, RNA virali eucariotici, frammenti di UTR, rRNA). I nuovi candidati miR sono quindi stati allineati con criteri meno stringenti ai miR presenti in miRBase (hairpin e mature). Infine, sia per i miR noti che per i nuovi candidati miR sono state studiate le caratteristiche genomiche di posizione intra-genica o extra-genica e di conservazione inter-specie.

4.2.2 Analisi delle librerie di small-RNA ottenute da popolazioni linfoidi T

Complessivamente, dal sequenziamento delle librerie dei nostri campioni sono state ottenute 29.744 sequenze di small-RNA. L'analisi bioinformatica delle nostre librerie ha identificato 139 sequenze non ridondanti corrispondenti a miR noti (già presenti nei *database* di miR umani) e 98 sequenze non ridondanti identificate come candidati nuovi miR. Per essere considerati candidati nuovi miR le sequenze dovevano essere presenti con più di una conta in almeno una delle 4 librerie analizzate. Il numero di sequenze non ridondanti di miR maturi noti, nuovi putativi e totali per le diverse popolazioni T è riportato nella tabella 4.1. In tutte e 4 le librerie sequenziate il numero di miR noti è superiore a quello dei nuovi miR putativi individuati (vedere tab. 4.1).

N° sequenze	LT CD4	LT CD8	DP	TIMOCITI NON FRAZIONATI
miR noti (FREQUENZA > 0 copie)	99	88	86	89
miR nuovi putativi (FREQUENZA > 1 copia)	45	44	61	31
miR maturi tot sequenziati	144	132	147	120

Tabella 4.1. La tabella riporta i numeri di miR noti e nuovi putativi identificati rispettivamente nei timociti non frazionati, nei timociti DP e nei linfociti T maturi CD4 e CD8 del sangue periferico (LT CD4 e LT CD8).

Circa il 42% dei miR noti è presente nelle librerie con più di 10 conte rispetto al 35% dei miR nuovi. Inoltre, il 40% dei miR noti rilevati nelle librerie contro il 30% dei candidati nuovi miR risulta espresso in modo condiviso da DP, LT CD4 e LT CD8 (fig. 4.6). La maggiore percentuale di miR nuovi con espressione specifica in una delle tre popolazioni cellulari considerate è in accordo con il concetto che tutt'oggi molti miR specificamente espressi in particolari sottopopolazioni cellulari siano ancora poco rappresentati nei database *on-line*.

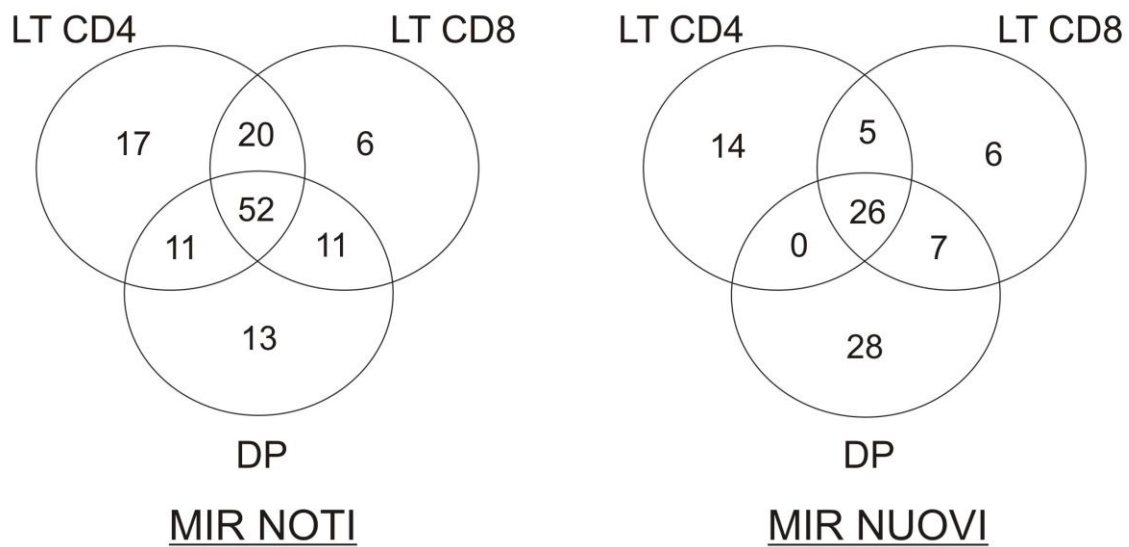


Figura 4.6. Diagrammi di Venn rappresentanti il numero di miR noti (sinistra) e nuovi putativi (destra) comuni o specificamente espressi dalle 3 diverse popolazioni T linfoidi considerate: timociti DP, linfociti maturi CD4 e CD8.

Abbiamo voluto stimare la completezza delle nostre librerie nel rappresentare la totalità dei miR espressi dai campioni in studio. A tale scopo è stata utilizzata una tecnica *bootstrap* [80,81] per estrapolare il numero totale di miR espressi da ogni campione. Questo valore è stato poi confrontato con il numero di miR maturi (noti e nuovi putativi) identificati mediante il sequenziamento delle nostre librerie. Sulla base di questa analisi è stato stimato che le librerie sequenziate coprono rispettivamente il 90% (LT CD4), l'85% (LT CD8), il 90% (DP) e l'84% (TIMOCITI NON FRAZIONATI) della totalità dei miR espressi nei campioni studiati (tab. 4.2).

	LT CD4	LT CD8	DP	TIMOCITI NON FRAZIONATI
COMPLETEZZA	90%	85%	90%	84%

Tabella 4.2. Completezza delle librerie. Mediante un'analisi *bootstrap*, è stato calcolato che, per tutte e 4 le librerie, il *set* di miR maturi identificati dal sequenziamento rappresenta più dell'80% del *set* di miR che si stima sia espresso da ogni campione.

Abbiamo calcolato le frequenze relative dei singoli miR nelle 4 librerie analizzate e, così facendo, abbiamo osservato che un particolare miR, il miR-142-3p, costituisce il 41%, il 58%, il 63% ed il 61% della totalità dei miR espressi rispettivamente dai LT CD4, LT CD8, timociti DP e timociti non frazionati. Il livello di espressione eccezionalmente alto ed ubiquitario nelle popolazioni considerate, suggerisce che questo miR sia un marcatore delle popolazioni linfoidi T, come già evidenziato in altri lavori [38,89]. Applicando il Fisher Exact test alle conte dei singoli miR noti presenti nelle librerie, abbiamo identificato 50 e 48 miR che risultano differenzialmente espressi ($p < 0,05$) rispettivamente nel confronto fra LT CD4 e timociti DP ed in quello fra LT CD8 e timociti DP. Analogamente a quanto fatto per i dati dei miR microarray, è stata individuata una lista di 41 miR noti differenzialmente regolati ($p < 0,05$) nel passaggio dai precursori DP ai linfociti T maturi, in modo concorde per i LT CD8 ed i LT CD4. Di questi 41 miR, 30 sono up-regolati nei LT maturi del sangue periferico rispetto ai precursori T DP, mentre 11 sono down-regolati.

Ci siamo voluti concentrare sullo studio dei miR regolati nel passaggio da precursori T DP a linfociti T maturi allo scopo di individuare miR implicati essenzialmente nel processo differenziativo e non nelle specifiche funzionalità dei LT CD4 rispetto ai CD8. Abbiamo quindi calcolato il rapporto (*fold-change*, FC) fra la frequenza relativa dei singoli miR noti nelle due popolazioni linfocitarie mature e nei timociti DP. In questo modo, è stata generata una lista di miR modulati in modo concorde in entrambe queste popolazioni rispetto ai DP con FC decrescente. Nella figura 4.7 sono rappresentati i 15 miR noti up-regolati (A) e i 10 down-regolati (B) con FC maggiore nel confronto fra LT CD4 e CD8 e timociti DP.

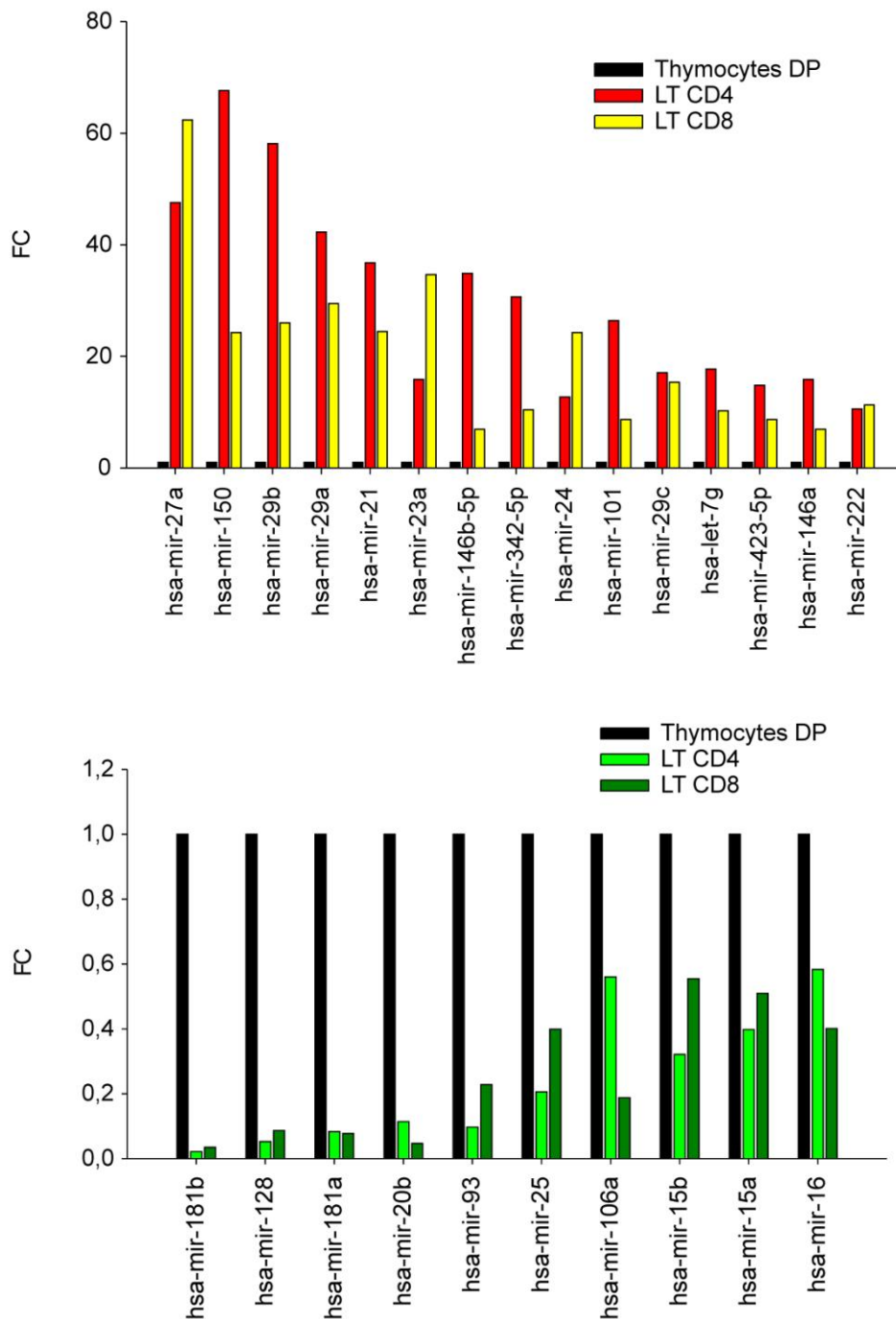


Figura 4.7. MiR up-regolati e down-regolati con maggiore FC risultati dall'analisi delle librerie. In questa figura sono rappresentati i 15 miR up-regolati (A) ed i 10 down regolati (B) con fold change (FC) maggiore, risultati dal confronto fra linfociti maturi CD4 e CD8 e timociti DP. Gli istogrammi mostrano, per i diversi miR, l'FC, calcolato come rapporto fra la frequenza relativa dei singoli miR rispettivamente nei LT CD4 e nei LT CD8 e la frequenza relativa degli stessi miR calcolata nella popolazione timocitaria DP. La frequenza relativa è calcolata come il rapporto percentuale della conta di un miR rispetto alla conta totale di tutti i miR noti presenti nella libreria. I miR scelti per la rappresentazione in questa figura sono stati selezionati dalla lista dei miR, differenzialmente regolati ($p < 0,05$), risultati con FC maggiore nel confronto fra LT (considerati come un'unica popolazione senza distinguere fra LT CD4 e CD8) e DP.

4.3 Confronto fra i dati dei miR microarray e delle librerie di small-RNA: identificazione di miR comunemente regolati nel corso della maturazione dei linfociti T

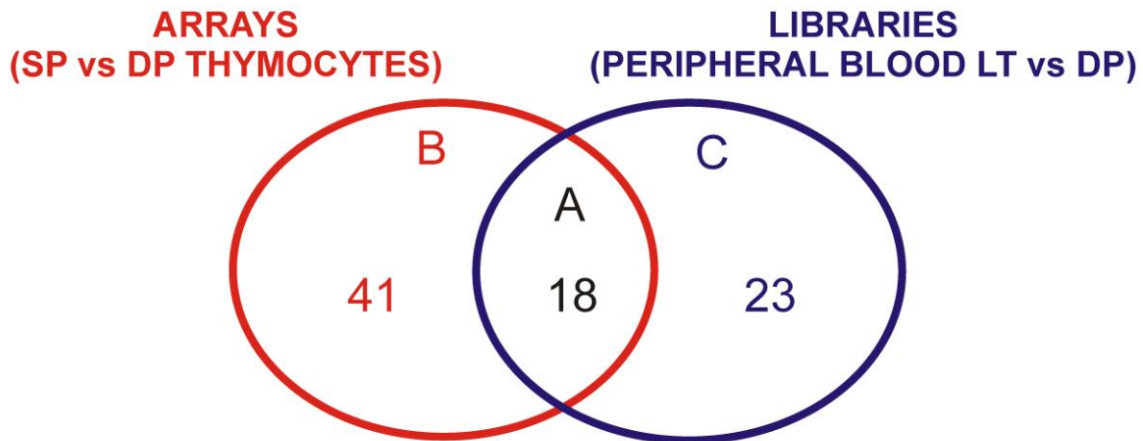
Dopo aver analizzato separatamente i dati derivanti dall'esperimento di miR microarray sulle sottopopolazioni timocitarie e quelli derivanti dalle librerie di small-RNA delle popolazioni linfoidi T, si è ricercata una correlazione tra i miR differenzialmente regolati nei LT rispetto ai DP (rilevati nel sequenziamento delle librerie di small-RNA) ed i miR differenzialmente regolati nei timociti SP rispetto ai DP (rilevati mediante l'analisi dei microarray) (fig. 4.8). La correlazione è stata determinata arbitrariamente, basandosi sulla regolazione concorde di uno specifico miR, rilevata con entrambi i metodi.

E' stato così individuato un gruppo di 18 miR che risultano regolati nel passaggio dei timociti dallo stadio di DP a quello di SP e che vengono mantenuti modulati nello stesso senso anche quando le cellule T migrano dal timo alla periferia. Questo gruppo di miR nella figura 4.8 è rappresentato dall'insieme intersezione "A" e corrisponde al 26% (per i microarray) ed al 44% (per le librerie di small-RNA) dei miR identificati mediante le due tecniche come differenzialmente regolati nel confronto fra le popolazioni considerate. In base a queste osservazioni, abbiamo ipotizzato che questo gruppo di miR possa svolgere un ruolo importante nel processo maturativo dei precursori T.

Riguardo ai restanti miR trovati diversamente regolati solo in uno dei due test, questi potrebbero essere attivati o soppressi in modo specifico nei timociti SP o nei linfociti T in associazione a fenomeni biologici particolari che si verificano nel microambiente timico o nella periferia. Un altro elemento di variabilità è introdotto dal fatto che si stanno confrontando dati di espressione ottenuti con tecniche diverse e normalizzati in modo indipendente.

Per quanto riguarda il confronto fra i profili d'espressione ed i dati di sequenziamento delle popolazioni mature T, abbiamo invece osservato solo una limitata sovrapposizione fra i miR differenzialmente regolati nelle due popolazioni timiche SP e le due popolazioni linfocitarie T CD4 e CD8. Infatti solo 2 miR risultano regolati in modo concorde nei due confronti. Questo sorprendente risultato, aldilà del problema legato all'uso di tecniche differenti, sembrerebbe indicare che l'espressione dei miR che differenziano la popolazione CD4 dalla CD8 vari moltissimo nel passaggio di queste cellule dall'ambiente timico, dove i timociti SP CD4 e CD8 ancora naive sono

sottoposti alla selezione negativa per l'eliminazione delle cellule auto-reattive, alla periferia, dove le cellule T naive, in seguito all'incontro con specifici antigeni, possono assumere funzioni effettrici o *memory*.



A			
Up-regolati:		Down-regolati:	
hsa-miR-146b-5p	hsa-miR-24	hsa-miR-26b	hsa-miR-128
hsa-miR-146a	hsa-miR-29a	hsa-miR-30c	
hsa-miR-150	hsa-miR-29b	hsa-miR-423-5p	
hsa-miR-23a	hsa-let-7g		
hsa-miR-27a	hsa-miR-101		
hsa-miR-342-5p	hsa-miR-21		
hsa-miR-342-3p	hsa-miR-222		

Figura 4.8. Correlazione tra i dati di miR microarray e quelli delle librerie di small-RNA. I Diagrammi di Venn mostrano la correlazione fra i miR differenzialmente regolati nel confronto timociti SP vs DP (miR microarray) e nel confronto LT maturi vs timociti DP (librerie di small-RNA). La correlazione è stata determinata arbitrariamente basandosi sulla regolazione concorde di uno specifico miR in entrambi i confronti. L'insieme intersezione "A" mostra un gruppo di 18 miR che risultano consistentemente regolati durante la maturazione T.

4.4 Validazione tramite qRT-PCR dei miR differenzialmente regolati nella differenziazione T

Il profilo d'espressione di alcuni miR, identificati mediante analisi dei microarrays e/o delle librerie di small-RNA come consistentemente modulati nel corso del differenziamento dei precursori T dallo stadio DP agli stadi successivi di maturazione, è stato validato tramite TaqMan qRT-PCR. Al fine della validazione, sono stati testati in PCR campioni di timociti DP, SP CD4 e SP CD8. In questo modo sono stati confermati i dati di espressione per 10 miR ($p < 0,05$), precedentemente identificati come differenzialmente regolati nelle 3 sottopopolazioni timiche d'interesse sulla base dei risultati dell'esperimento di miR microarray. I risultati della PCR mostrano una buona correlazione anche quantitativa con i dati degli array. Nella figura 4.9 sono mostrati i risultati delle PCR effettuate su almeno 3 repliche biologiche delle sottopopolazioni timiche umane d'interesse (DP, SP CD4 e SP CD8) per i 10 miR validati con questa tecnica.

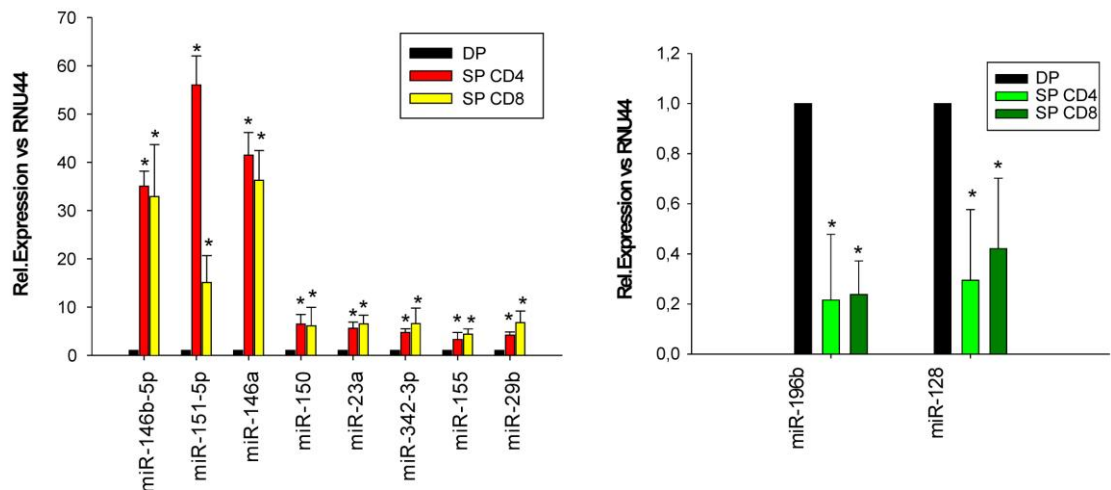


Figura 4.9. Validazione tramite TaqMan qRT-PCR dei profili d'espressione dei miR differenzialmente regolati nelle sottopopolazioni timiche d'interesse ($p < 0,05$). I livelli di espressione relativa dei diversi miR sono stati normalizzati rispetto all'espressione dello small-non-coding RNA RNU44. Nella figura è mostrata la media \pm SD di almeno 3 esperimenti per ogni miR.

4.5 Predizione computazionale dei potenziali bersagli di miR d'interesse

Per la predizione computazionale dei potenziali mRNA bersaglio, sono stati scelti 9 miR fra quelli differenzialmente regolati con *fold-change* maggiore nel passaggio da timociti DP a SP CD4 e CD8. Oltre al *fold-change*, un altro criterio impiegato per la selezione dei miR è stata la presenza di dati in letteratura che vedessero il miR coinvolto in processi biologici di potenziale interesse, quali la differenziazione di cellule ematopoietiche, la regolazione della funzione immunitaria, il controllo della proliferazione o dell'apoptosi e la tumorigenesi. Infine, si è cercato di includere in quest'analisi i miR che risultavano consistentemente modulati nel corso del differenziamento dei precursori T dallo stadio DP agli stadi successivi di maturazione, cioè quelli che venivano mantenuti regolati in modo concorde nel confronto fra timociti DP e SP e in quello fra timociti DP e linfociti T periferici (vedi punto 4.3 dei Risultati). Al fine dell'analisi, le sequenze di questi miR sono state introdotte in tre diversi software di predizione computazionale di target: TargetScan (<http://www.targetscan.org>), PicTar (<http://pictar.mdc-berlin.de>) e MiRanda (<http://www.microrna.org>). In un precedente studio [84] questi tre software sono stati identificati (confrontando cinque diversi software di predizione di target) come i tre metodi migliori di predizione di target di miR con una sensibilità di riconoscimento compresa fra il 65% e il 68%. MiRanda identifica i siti potenziali di legame miR-mRNA ricercando regioni altamente complementari nell'UTR-3' dei trascritti. Lo *Score* o punteggio di allineamento è assegnato mediante un algoritmo che dà più peso agli appaiamenti a livello dell'estremità 5' del miR. I siti di legame potenziali così ricavati sono successivamente valutati per la stabilità termodinamica dell'appaiamento miR-target. TargetScan usa un approccio un pò diverso, imponendo al principio il vincolo della perfetta complementarietà a livello della *seed region* del miR ed estendendo successivamente la ricerca di complementarietà a regioni al di fuori del *seed*. Inoltre TargetScan usa come ulteriore criterio di selezione la conservazione interspecie del sito di legame, usando per l'allineamento con il miR un gruppo di UTR-3' ortologhi di diverse specie. Come in MiRanda, anche in TargetScan gli appaiamenti predetti sono testati per la loro stabilità termodinamica. PicTar usa un metodo di ricerca molto simile a TargetScan, introducendo però in aggiunta anche una valutazione della co-espressione

nello spazio e nel tempo dei miR e dei target, attingendo dai *database on-line* di profili d'espressione genica [84].

Poichè questi tre diversi algoritmi sfruttano criteri e parametri di selezione in parte simili ed in parte distinti, le liste di potenziali target che generano sono solo parzialmente sovrapponibili. Al fine di ridurre il numero di falsi positivi, i risultati dei diversi software sono quindi stati integrati ed è stata individuata una lista di candidati target comuni a tutte e tre le predizioni (vedi tabella 4.3).

Fra i candidati target del gruppo di miR d'interesse sono stati identificati molti geni coinvolti in processi biologici rilevanti nel differenziamento cellulare e nella risposta immunitaria T, fra cui diversi membri della famiglia dei geni HOX (HOXA5, HOXC8, HOXB7, HOXB6, HOXA13 e HOXD10) ed alcuni componenti del pathway di Notch (NOTCH3, DTX1).

Incrociando i risultati di tre diversi software di predizione di target si riduce il numero di bersagli predetti, limitando il numero di falsi positivi; tuttavia in questo modo si aumenta molto il numero di falsi negativi, che vengono esclusi dalle successive analisi.

4.6 Combinazione dei profili d'espressione genica e dei miR di popolazioni timocitarie per la predizione di bersagli

E' stata condotta un'analisi del profilo d'espressione genica (GEP) delle sottopopolazioni timiche umane DP e SP CD4, nonché dei timociti totali. Da questa analisi sono per ora state escluse le popolazioni timiche SP CD8 e DN per insufficienza di materiale disponibile. Lo studio è stato effettuato utilizzando i microarray Human genome U-133 PLUS della ditta Affymetrix. Per ogni popolazione studiata, sono state ibridate sugli array da un minimo di 3 ad un massimo di 6 repliche biologiche.

L'elaborazione dei dati dei *chip* d'espressione genica è stata fatta con i software Genesis (<http://genome.tugraz.at/>) e BRB Array Tools (<http://linus.nci.nih.gov/pilot/index.htm>).

Complessivamente sono stati individuati 1665 geni differenzialmente regolati (FDR=0,001) nel corso del passaggio dei timociti dallo stadio di DP a quello di SP CD4. In particolare, 552 trascritti sono risultati up-regolati nei DP rispetto ai SP CD4 (con *fold-change* compreso fra 1,8 e 8,5); mentre 1113 trascritti sono risultati down-regolati (con *fold-change* variabile da 0,02 a 0,55).

mRNA BERSAGLIO PREDETTI
(MirBase-PicTar-TargetScan)

miR-196b	miR-128	miR-146	miR-150	miR-23a	miR-27a	miR-342	miR-155	miR-29b
MGAT4A	CSF1	TRAF6	ELOVL3	SEMA6D	PLK2	FHL3	ACTA1	COL11A1
OTX1	NEK2	IRAK1	PLP2	AUH	EYA4	RSBN1	H3F3A	COL16A1
TCF7	TROVE2	KLF7	ETF1	INTU	VIP	RASAL2	KCNN3	MFAP2
HAND1	SV2A	STRBP	PKP4	MRC1	TXN2	PPP3R1	IKBKE	NASP
RGL2	MFSD2	BIVM	CACNA1G	POU4F2	ZNF329	RASSF1	GOLPH3L	CLK2
CDYL	NRP2	PPP1R11	STX5	PKP4	UNKL	PDGFRA	CDC73	NFIA
TRERF1	ADCY3	KCTD15	EGR2	MYH4	AQP11	MRFAP1	ZNF642	RLF
COL1A2	TMSB10	DLGAP1	BSN	SOCS6	PPARG	RASA1	LRP1B	SETDB1
TSPAN12	FBLN2		EPHB2	MAP3K7IP3	C10orf137	ARID1B	FBXO11	GPX7
HOXA5	PDIA5		GDI1	CUL3	CDS1	NRG1	COL7A1	PER3
ABCB9	SEC61A1		RAD23B	PPP1CB	MMD	ZNF462	RNF123	MFAP2
HOXC8	ECE2		PRRT2	TADA1L	HOXA13	COL2A1	BOC	HMCN1
RBM26	DCP1A		ELK1	ANKHD1	ITSN2	MSI1	CSNK1A1	TRIM63
DOC2A	CCDC71		PLXNC1	CCL7	ADORA2B	FUT8	CSF1R	COL11A1
NME4	PLK2		GTF3C2	MYH1	PLCL2	MYLK2	MYO10	COL3A1
DDX19B	PAIP2		PA2G4	SNRPC	GALNT5		HIVEP2	COL3A1
DDX19A	RGL2		EDA	CAMTA1	PSMA1		DYNC111	SLC30A3
HOXB7	MDFI		SLC38A3	KIAA0922	B4GALT3		SYPL1	PRKRA
HOXB6	CYP39A1		GLRA1	NCOA6	SS18L1		HBP1	COL6A3
SMCR8	FOXQ1		NOTCH3	CCK	HOXB8		SDCBP	DUSP2
EFCAB5	HOXA13			PKIA	NR2F6		SGK3	EIF4E2
GATA6	CCM2			MAB21L2	RARA		CHD7	RND3
C20orf160	MOSPD3			LRP5	SPRY2		TP53INP1	MYCN
SAPS2	ORC5L			HOXD10	HORMAD2		SGK3	ABCB6
RGL2	TNPO3			UBE2O	NGFRAP1		ASTN2	SMEK2
	RNF38			GPRC5B	PDIA5		ADD3	DNMT3A
	C9orf97			CTNNBIP1	TAPT1		WEE1	RMND5A
	KLF4			STX17	GFPT2		AICDA	COL6A3
	SH2D3C			CPSF4	BAG2		PCDH9	CCNYL1
	NXT2			DUSP5	TMUB1		RCN2	SLC16A14
	NRBF2			SYS1	PHB		SALL1	DNMT3A
	SLC39A13			HSP90B1	NDUFS4		SOCS1	SESTD1
	PDE3A			TRIM63	INSR		CARHSP1	COL7A1
	LMBR1L			LHX4	CYP39A1		AKAP10	COMMD2
	DTX1			KCNK3	ISL1		KIAA1267	DNAJB11
	CABP1			PLAU	BCL3		CSNK1G2	MLF1
	ABCB9			EPS15	NXT2		MAP3K10	MED12L
	CCDC92			DLX1	ST14		SMARCA4	FAM116A
	CDH24			UBE2D1	BRSK1		CEBPB	CHIC2
	RGS6			DHX15	NEUROD6		BRD1	ELF2
	KCNK10			ZBTB26	NRBF2		DNAJB7	GAB1
	NEO1			PTK2B	RGS6			OTUD4
	C15orf27			ZNF423	FRS3			PDGFC

Tabella 4.3. Lista dei target comuni predetti dai tre software TargetScan, PicTar e MiRanda per i 9 miR selezionati.

Come già detto nell'Introduzione, i miR possono regolare negativamente l'espressione genica sia inibendo la traduzione in proteina, che destabilizzando il trascritto dei loro bersagli. La maggior parte dei miR nei mammiferi usa entrambe le modalità di regolazione, alterando, di conseguenza, sia i livelli di proteina, che i livelli di mRNA dei propri geni bersaglio [85,86].

A conferma di quanto detto, diversi studi [87,88] hanno riportato che, in seguito all'over-espressione o al silenziamento di specifici miR, si ha una down-regolazione di un gruppo di trascritti arricchito per mRNA target predetti computazionalmente. Inoltre, mRNA contenenti nel loro UTR-3' *seed match* riconosciuti da specifici miR tendono ad essere espressi a livelli ridotti in cellule o tessuti in cui quei miR sono naturalmente espressi a livelli elevati. Per questo lo studio del profilo d'espressione genica correlato al profilo d'espressione dei miR può essere un utile strumento per affinare e restringere il gruppo di bersagli potenziali predetti da software bioinformatici. Abbiamo così deciso di utilizzare anche questo secondo sistema per individuare candidati mRNA bersaglio per alcuni dei miR differenzialmente regolati nel corso della maturazione T.

Sono quindi state recuperate da miRBase (www.mirbase.org) le sequenze delle forme mature dei primi 17 miR up-regolati e dei primi 2 miR down-regolati con maggiore *fold-change* nel corso della differenziazione dei timociti da DP a SP. Per questi miR, sono state generate delle liste di target predetti computazionalmente mediante Microcosm, che utilizza per la predizione di target il software Miranda. Per quest'analisi abbiamo deciso di utilizzare la predizione di un solo software poichè, incrociando le predizioni di tre software, come già accennato al punto 4.5, si perdono molti bersagli potenziali identificati come falsi negativi. Abbiamo quindi utilizzato i dati di GEP per studiare l'andamento nei timociti DP e SP CD4 degli mRNA bersaglio predetti dall'analisi computazionale per i 19 miR scelti. Complessivamente, 252 target su 15155 predetti sono risultati differenzialmente regolati nel passaggio da DP a SP CD4 in senso opposto ai miR di cui sono stati predetti bersagli. Questo numero è sorprendentemente basso ed è un indice di quanto le predizioni bioinformatiche spesso possano essere incomplete. D'altro canto, è possibile che alcuni miR agiscano sui loro target prevalentemente a livello di proteina e, ovviamente, per questi casi, il confronto con i dati di espressione genica è poco informativo.

Fra i trascritti candidati target differenzialmente regolati nei GEP nel passaggio da timociti DP a SP CD4 abbiamo individuato in particolare un gene che viene predetto come bersaglio potenziale di miR-150 da tutti e tre i software di predizione di target:

TargetScan, PicTar e MiRanda. Questo candidato target è Notch3, un membro della famiglia dei recettori Notch, e verrà da qui in poi denominato *Candidate Target 1*.

4.7 Allineamento di miR-150 all'UTR-3' di Candidate Target 1

Il gene *Candidate Target 1* è presente nella lista degli mRNA target comunemente predetti per miR-150 da tre software bioinformatici diversi. Inoltre, il trascritto di *Candidate Target 1* risulta regolato in modo opposto al miR-150 nel passaggio dei timociti dallo stadio DP a quello di SP CD4. Notch3, ovvero *Candidate Target 1*, è un membro della famiglia di Notch che svolge un ruolo fondamentale nello sviluppo dei linfociti T e nella loro trasformazione neoplastica. Sulla base di questi dati abbiamo deciso di approfondire le analisi sulla relazione fra miR-150 e *Candidate Target 1*. Studiando in particolare l'appaiamento fra la sequenza target e quella del miR, abbiamo identificato un'elevata complementarietà fra il miR-150 e l'UTR-3' di *Candidate Target 1*. Come si può vedere nella figura 4.10, l'allineamento fra la regione del *seed* in 5' (nt 2-8) del miR-150 e l'UTR-3' di *Candidate Target 1* è perfetto ed è conservato nell'uomo e nel topo. Inoltre, la regione di complementarietà è allungata dalla presenza di un match conservato (in uomo e topo) anche a livello della posizione 9 del miR. Una caratteristica che è stata frequentemente riscontrata nei target validati dei miR in precedenti lavori è la presenza di complementarietà anche a livello della porzione 3' del miR, in particolare a livello delle basi 12-17 del miR [31]. Per quanto riguarda l'UTR-3' di *Candidate Target 1*, nell'uomo si osserva un allineamento a livello dei nt 11-12-13-15-17 del miR-150; mentre nel topo, subito dopo la posizione 9 del miR, si ha un *bulge* seguito da un appaiamento di 5 basi consecutive alle posizione 10-15 del miR.

Tutti questi elementi fanno di *Candidate Target 1* un buon candidato, appunto, come bersaglio dell'azione del miR-150.

Per validare l'associazione fra miR-150 e *Candidate Target 1* abbiamo clonato in un costrutto plasmidico l'UTR-3' di *Candidate Target 1* a valle del gene reporter della luciferasi. Nei prossimi mesi andremo a trasfettare la linea cellulare 293T con questo plasmide reporter Luc-UTR-3' e diverse dosi di miR-mimic-150 o di un miR di controllo e valuteremo le variazioni dell'attività della luciferasi in presenza ed in assenza dello specifico miR d'interesse. Inoltre, allo scopo di valutare l'effetto repressivo di miR-150 sull'espressione di *Candidate Target 1*, andremo ad over-esprimere miR-150 in linee

cellulari esprimenti *Candidate Target 1* e valuteremo tramite W.B. (Western Blot) e real-time PCR le variazioni a livello di proteina e di mRNA del candidato bersaglio.

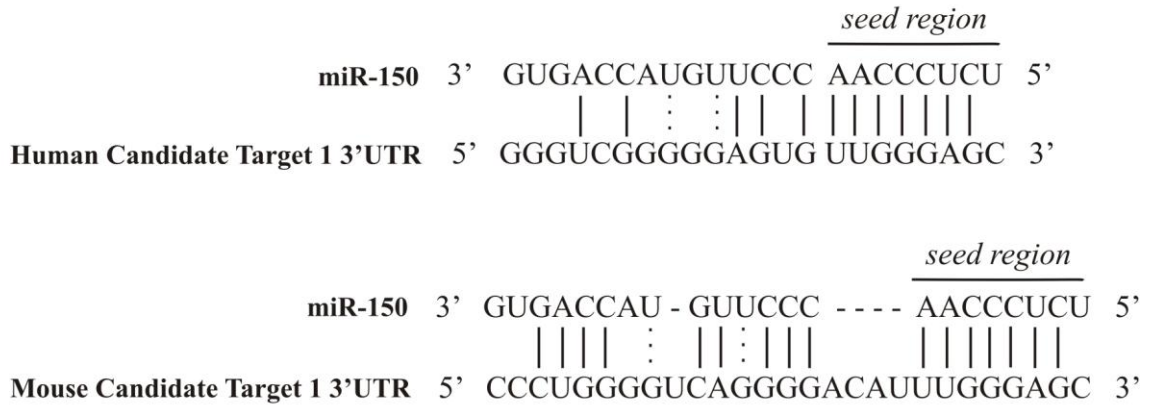


Figura 4.10. Allineamento di sequenza fra miR-150 e l'UTR-3' del *Candidate Target 1* nell'uomo e nel topo (calcolato dal software MiRanda). La linea continua indica la complementarietà fra UTR-3' e miR in corrispondenza della *seed region*; la linea tratteggiata indica le regioni delete.

4.8 Livelli di espressione di miR-150 in diverse linee di T-ALL

MiR-150 (come già detto nell'Introduzione) svolge un ruolo molto importante nello sviluppo sia dei linfociti B, che dei linfociti T. Questo miR è espresso in modo selettivo in cellule B e T mature e non attivate, ma non nei loro progenitori. Uno *screening* dei miR de-regolati nella ALL (sia B, che T) recentemente pubblicato [72] ha individuato miR-150 fra i quattro miR maggiormente down-regolati nelle leucemie acute rispetto alla controparte normale. Siamo quindi andati a misurare l'espressione di miR-150 in 5 diverse linee di T-ALL (CCRF-HSB2, TALL-1, DND4.1, MOLT-3, Jurkat). La quantificazione relativa dei livelli di espressione del miR-150 è stata condotta tramite TaqMan real-time PCR. Le linee di T-ALL sono state confrontate per i loro livelli di espressione con timociti DP e SP CD4.

Dato che in diversi lavori è stato riportata una generale down-regolazione nell'espressione dei miR nei tumori ed in particolare nelle linee cellulari neoplastiche, siamo andati a valutare nelle linee di T-ALL i livelli di espressione di altri 4 miR: 2 miR fortemente regolati (miR-146b, up-regolato nei timociti SP CD4 e CD8 rispetto ai DP; miR-128, up-regolato nei timociti DP rispetto ai SP CD4 e CD8) e 2 miR che non sono

modulati (miR17 e miR-92) nel corso della differenziazione T. Come si può vedere dalla figura 4.11, miR-150 in tutte le linee di T-ALL considerate è espresso a livelli bassi rispetto ai timociti DP, i quali a loro volta esprimono questo miR circa 6 volte meno dei timociti maturi SP CD4. Le diverse linee di T-ALL, pur presentando tutti livelli piuttosto bassi del miR-150, presentano fra loro una certa variabilità nei valori d'espressione di questo miR (con una differenza in termini di *fold-change* di circa 49 volte fra le TALL-1, la linea cellulare con la più alta espressione, e le DND4.1, quella con la più bassa espressione). Queste differenze in termini di *fold-change* non correlano con il fenotipo delle linee cellulari. Infatti, la linea Jurkat, che presenta un fenotipo maturo SP CD4⁺/CD3⁺, ha livelli di espressione di miR-150 molto simili alla linea MOLT-3, che invece presenta un fenotipo immaturo ISP (Immature Single Positive).

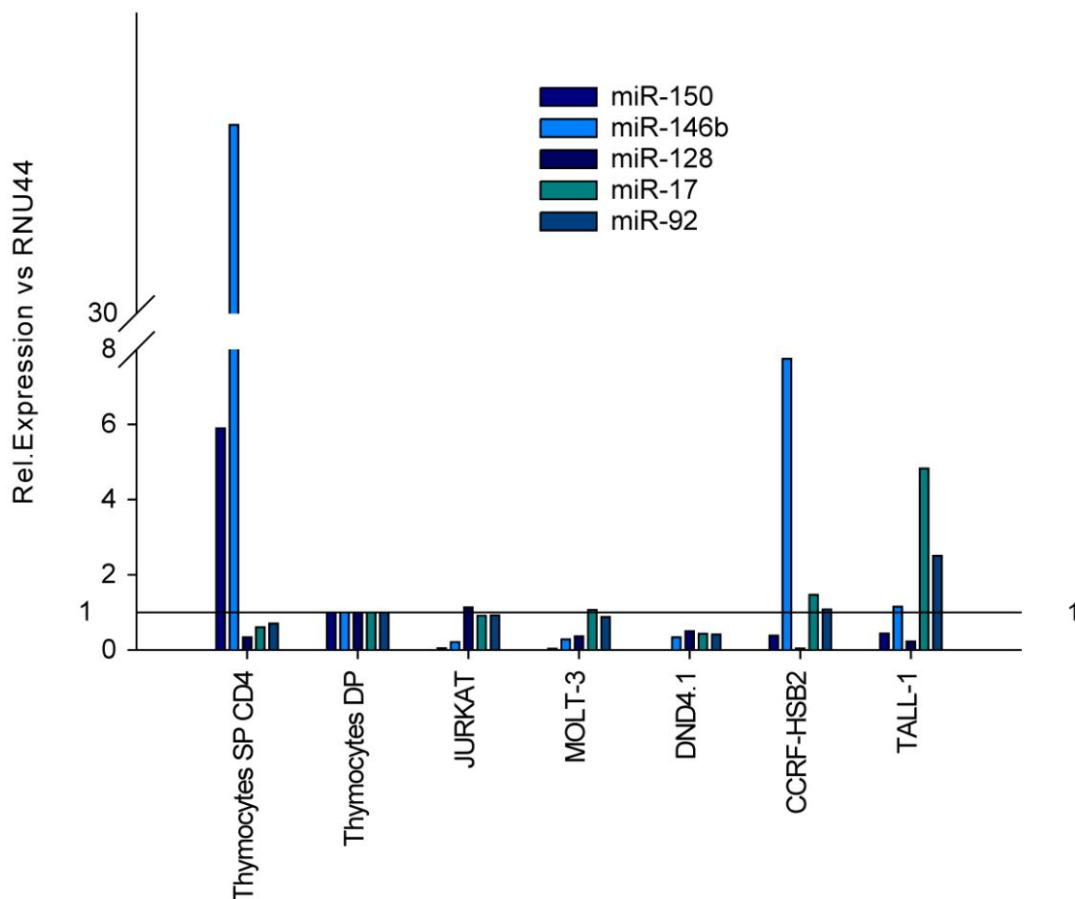


Figura 4.11. Misurazione tramite TaqMan qRT-PCR dell'espressione relativa di miR-150, miR-146b, miR-128, miR-17 e miR-92 nelle linee di T-ALL e nelle popolazioni timocitarie. I livelli di espressione relativa di miR-150, miR-146b, miR-128, miR-17 e miR-92 sono stati normalizzati rispetto all'espressione dello *small-non-coding* RNA RNU44.

Anche gli altri miR-146b e miR-128 presentano in generale livelli di espressione più bassi nelle linee di T-ALL rispetto alla popolazione timocitaria DP, presa come riferimento (con l'eccezione di miR-146b nelle linee CCRF-HSB2 e TALL-1 e miR-128 nella linea Jurkat). Tuttavia, a differenza del miR-146b e 128, l'espressione di miR-150 in tutte le cinque linee analizzate risulta costantemente down-regolata rispetto ai timociti DP con un *fold-change* medio di circa 35 volte (contro il *fold-change* medio di circa 2 del miR-146b e quello di circa 7 del miR-128). Al contrario, il livello d'espressione di miR-17 e miR-92 in tutte le linee di T-ALL analizzate (con l'eccezione delle DND4.1) è simile o superiore rispetto a quello dei timociti DP.

4.9 Studi funzionali dell'over-espressione del miR-150 in una linea di T-ALL

Considerata la bassa espressione del miR-150 nelle linee di T-ALL analizzate e la presenza fra i bersagli predetti di miR-150 di diversi geni che sono noti essere coinvolti nella regolazione della proliferazione, della sopravvivenza e della differenziazione cellulare, ci siamo chiesti se l'over-espressione di questo miR avesse degli effetti funzionali sulle linee di T-ALL. A questo proposito, la linea cellulare di T-ALL Jurkat è stata trasdotta con un vettore lentivirale esprimente il pre-miR-150 (LV-pre-miR-150) o con il relativo vettore lentivirale di controllo (LV-control). La struttura di questi costrutti virali (forniti entrambi dalla ditta SBI) differisce esclusivamente per la presenza o meno della sequenza del pre-miR a livello del sito MCS (*Multi-Cloning Site*) ed è mostrata nei Materiali e Metodi. Poichè entrambi i costrutti virali contengono al loro interno il gene *reporter* della GFP (*Green Fluorescent Protein*) posto sotto il controllo del promotore del gene EF1 α umano, l'efficienza di trasduzione è stata valutata a 48 ore mediante osservazione delle cellule al microscopio UV e misurazione al citofluorimetro della percentuale di cellule positive per GFP. Per la linea Jurkat l'efficienza di trasduzione è stata in ogni caso superiore al 99%, sia per il vettore LV-pre-miR-150, che per il vettore LV-control (vedi fig. 4.12).

I livelli di espressione di mir-150 nelle cellule Jurkat trasdotte con il vettore per l'over-espressione di questo miR ed il relativo vettore di controllo sono stati valutati tramite TaqMan real-time PCR. In questo modo si è riscontrato che nelle cellule trasdotte con

LV-pre-miR-150 il livello di miR-150 è aumentato in media circa 16 volte rispetto alle cellule trattate con il vettore di controllo (vedi fig. 4.13).

La trasduzione con i vettori sopracitati è stata testata anche su altre linee cellulari di T-ALL (MOLT-3, CCRF-HSB2, TALL-1, DND4.1), che però hanno mostrato tutti livelli più bassi di trasducibilità rispetto alle cellule Jurkat. Di conseguenza, le prove funzionali inerenti l'over-espressione del miR-150, per il momento, si sono concentrate su quest'ultima linea cellulare.

Le cellule Jurkat trattate con i vettori LV-pre-miR-150 o LV-control sono state quindi utilizzate per l'analisi della proliferazione cellulare mediante test dell'ATP. L'aumento dei livelli di ATP in una coltura cellulare nel corso del tempo indica l'aumento del numero di cellule metabolicamente attive e dunque è un indice di proliferazione cellulare. In breve, dopo 48 ore dalla trasduzione e dopo averne controllato al citofluorimetro la positività per GFP, le Jurkat sono state piastrate in quintuplicato in piastre da 96 pozzetti per il test dell'ATP e sono state analizzate a 3, 24, 48 e 72 ore dal momento della semina. Come si può osservare dagli istogrammi della figura 4.14, le cellule over-esprimenti miR-150 a distanza di 72 ore dalla semina (corrispondenti a 5 giorni dal momento della trasduzione) mostrano un dimezzamento del tasso proliferativo rispetto alle cellule trasdotte con il vettore di controllo. Questa differenza è già visibile a partire da 48 ore dal momento della semina delle cellule.

I risultati del test dell'ATP sono confermati dalla conta cellulare a 72 ore dalla trasduzione delle Jurkat trasdotte con i due costrutti lentivirali (vedi fig. 4.15). In breve, un milione di cellule Jurkat sono state piastrate alla concentrazione di 10^6 cellule/ml in una piastra da 6 pozzetti per la trasduzione con i vettori virali; quindi a 72 ore dalla trasduzione le cellule sono state contate. Come mostrano gli istogrammi della figura 4.15, a distanza di 72 ore dalla trasduzione, il numero di cellule trattate con il vettore di controllo è circa il doppio rispetto alle cellule trattate con il vettore per l'over-espressione del pre-miR-150.

La diminuzione del tasso di proliferazione cellulare può essere dovuta ad un rallentamento del ciclo cellulare e/o all'aumento di mortalità delle cellule. Il fatto che le Jurkat trasdotte con il vettore LV-pre-miR-150 mostrino nel tempo un aumento del numero di cellule in coltura e dei livelli di ATP (seppur minore rispetto a quello delle cellule di controllo) ci ha portato a pensare che l'over-espressione di miR-150 su questa linea cellulare abbia principalmente un effetto di rallentamento del ciclo cellulare, più che di induzione di apoptosi.

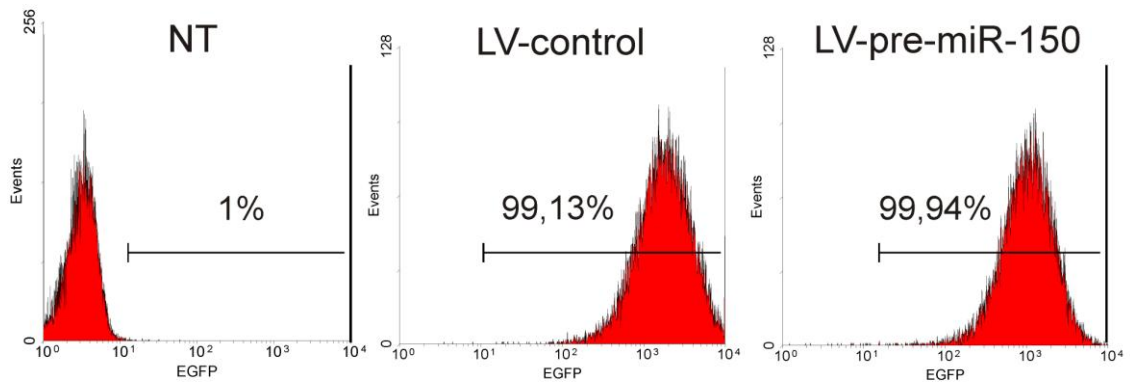


Figura 4.12. Valutazione citofluorimetrica della positività delle cellule alla GFP. Le cellule Jurkat, 48 ore dopo la trasduzione con i vettori virali LV-pre-miR-150 e LV-control, sono state analizzate al citofluorimetro e valutate per la percentuale di cellule GFP+.

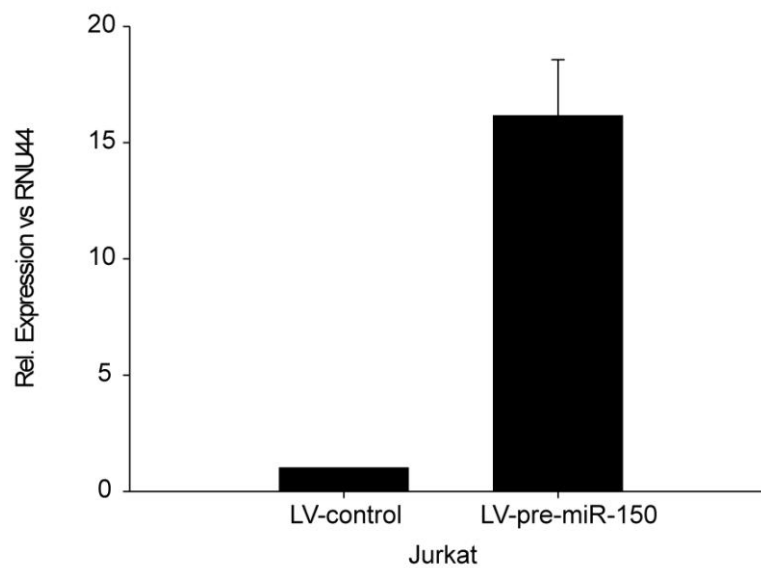


Figura 4.13. Misurazione tramite TaqMan qRT-PCR dell'espressione relativa di miR-150 nella linea cellulare Jurkat trasdotta con il vettore LV-pre-miR-150 o con il vettore LV-control. I livelli di espressione relativa di miR-150 sono stati normalizzati rispetto all'espressione dello *small-non-coding* RNA RNU44. Gli istogrammi mostrano la media di tre diversi esperimenti.

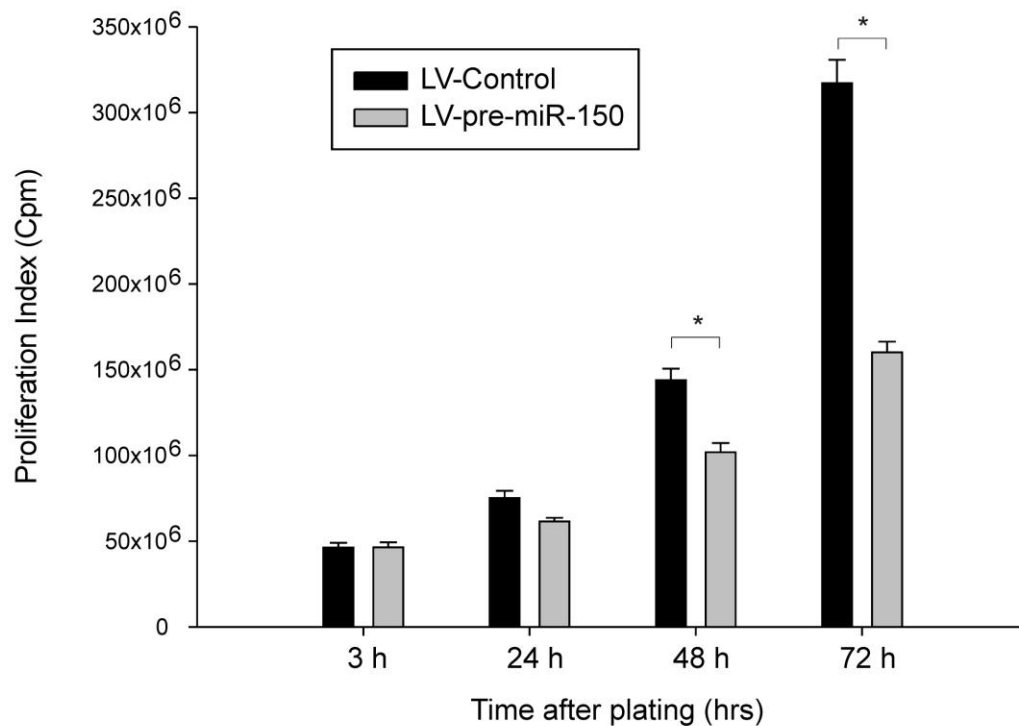


Figura 4.14. Analisi della proliferazione cellulare mediante test dell'ATP. Gli istogrammi rappresentano il tasso proliferativo misurato in termini di cpm (colpi per minuto) delle cellule Jurkat trasdotte con il vettore over-esprimente il pre-miR-150 (LV-pre-miR-150; barra grigia) ed il vettore di controllo (LV-control; barra nera). Nella figura è mostrato un esperimento rappresentativo di tre.

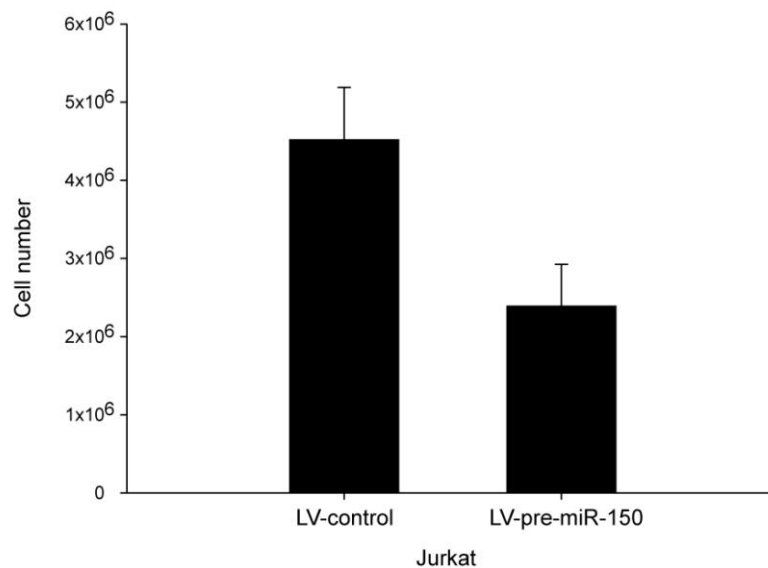


Figura 4.15. Conta cellulare a 72 ore dalla trasduzione. Negli istogrammi sono riportate le conte cellulari, a 72 ore dalla trasduzione, per le cellule Jurkat trattate rispettivamente con il vettore over-esprimente miR-150 o il vettore di controllo.

Al fine di verificare la nostra ipotesi, abbiamo fatto un'analisi della distribuzione nelle varie fasi del ciclo cellulare delle cellule a 72 ore dalla trasduzione con il vettore LV-pre-miR-150 o LV-control. Da un'analisi preliminare è risultato che, rispetto alle cellule di controllo, le cellule Jurkat in cui è indotta l'over-espressione di miR-150 vanno incontro ad un rallentamento del ciclo cellulare associato ad un accumulo delle cellule nella fase G2, precedente alla mitosi. Infatti, il 39% delle cellule di controllo sono in fase G2-M, contro il 57% delle cellule trattate con LV-pre-miR-150. Al contrario, nelle Jurkat LV-control è presente una frazione più elevata di cellule in fase S rispetto alle cellule LV-pre-miR-150 (fig. 4.16).

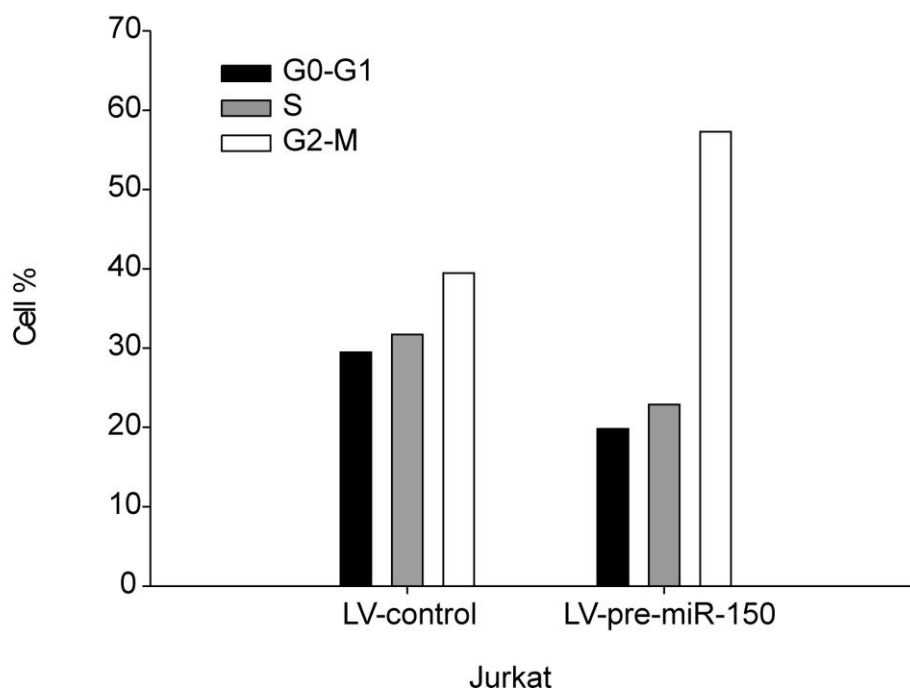


Figura 4.16. Analisi del ciclo cellulare mediante marcatura con Ioduro di Propidio. L'analisi del ciclo cellulare della linea Jurkat infettata con il vettore LV-pre-miR-150 o LV-control è stata condotta a distanza di 72 ore dalla trasduzione virale delle cellule. Gli istogrammi mostrano la distribuzione percentuale delle cellule nelle diverse fasi del ciclo cellulare (G0-G1, S, G2-M) in un esperimento rappresentativo.

Inoltre, abbiamo notato un aumento della percentuale di detriti nelle cellule trattate con il vettore LV-pre-miR-150 rispetto al controllo (49% vs 19%). Quest'ultima osservazione suggerisce che, in definitiva, la forzata espressione di questo miR nelle cellule di T-ALL induca una frazione delle cellule ad andare in apoptosi. In futuro, ci

proponiamo di replicare questa analisi estendendola anche ad altre linee cellulari di T-ALL. Inoltre, andremo a misurare direttamente con test mirati (valutazione del clivaggio di PARP mediante Western Blot e marcatura con Annexina-V) i livelli di apoptosi nelle cellule di T-ALL over-esprimenti o meno miR-150, al fine di valutare il contributo dell'apoptosi all'effetto osservato di rallentamento della proliferazione.

4.10 Profilo d'espressione dei miR in campioni di T-LBL e linfonodi reattivi umani

I linfomi linfoblastici T (T-LBL) rappresentano circa un terzo dei linfomi Non-Hodgkin dell'età pediatrica. Come già detto nell'Introduzione, ci sono significative analogie tra il T-LBL e la leucemia linfoblastica acuta T (T-ALL), tanto che frequentemente le due entità vengono considerate come una stessa patologia con diversa sede di presentazione e trattate nello stesso modo. In particolare, entrambi i tipi di neoplasia sembrano derivare dalla trasformazione di precursori T nel timo.

Al fine di identificare *pattern* di espressione dei miR associati alla trasformazione neoplastica delle cellule T, abbiamo comparato il profilo d'espressione delle sottopopolazioni timiche (DP, SP CD4, SP CD8 e timociti non frazionati) con il profilo d'espressione di un gruppo di 9 Linfomi Linfoblastici T (T-LBL) pediatrici e quello di un gruppo di 6 linfonodi reattivi non neoplastici (LN) (forniti dal laboratorio del Dr. Rosolen, Dipartimento di Pediatria, Università di Padova). I LN sono stati inclusi nell'analisi in modo da identificare *pattern* d'espressione differenziale dei miR nelle T-LBL rispetto ai timociti riconducibili alla composizione eterogenea dei campioni tumorali e non alla trasformazione neoplastica in sé. Infatti, i linfomi sono derivati da biopsie linfonodali che contengono al loro interno, oltre ai blasti T, anche tutta una serie di altri tipi cellulari (cellule B, cellule dendritiche, macrofagi, cellule endoteliali, etc). L'analisi dell'espressione dei miR nei campioni sopraindicati è stata condotta mediante ibridazione sullo *Human miRNA Microarray (v2.0)* della Agilent Technologies, utilizzato anche per l'analisi del profilo d'espressione dei miR nelle popolazioni timocitarie umane.

L'analisi non supervisionata dei dati di espressione dei miR ha confermato che le sottopopolazioni timocitarie hanno profili distinti che riflettono il processo di differenziazione, con maggiori affinità tra le popolazioni SP CD4 e SP CD8 ed un profilo nettamente distinto per le cellule DP. Inoltre, l'analisi ha evidenziato che i

linfomi linfoblastici T si comportano in modo aberrante rispetto alle cellule normali, mostrando un profilo d'espressione dei miR distinto da quello delle sottopopolazioni timocitarie. Le differenze tra T-LBL e popolazioni timocitarie non sono dovute solo al diverso contesto cellulare (biopsie verso cellule purificate), dato che i linfomi si separano nettamente dal tessuto linfonodale reattivo (LN). Il dendrogramma riportato in figura 4.17 rappresenta i risultati dell'analisi non supervisionata, condotta utilizzando un algoritmo di *clustering* gerarchico (basato sul metodo dell'*average linkage* e sulla correlazione di Pearson) fornito dalla piattaforma geWorkbench (<http://www.geworkbench.org>) (fig. 4.16). Nell'analisi sono stati inclusi solo i miR con valore di segnale d'espressione >100 in almeno uno dei fenotipi (116).

La maggior parte dei miR che in precedenza avevamo visto essere consistentemente regolati nel corso della maturazione dei precursori T, appaiono de-regolati nei linfomi. In particolare, i 20 miR maggiormente up-regolati ed i 5 miR maggiormente down-regolati nei timociti nel passaggio da DP a SP risultano tutti (a parte miR-128) differenzialmente espressi (t-test, $p < 0,05$) nei campioni di T-LBL rispetto ad almeno una delle popolazioni timiche normali. Alcuni di questi miR sono chiaramente espressi in altri tipi cellulari (LN) e quindi ulteriori validazioni sono indispensabili per verificare quali cellule sono responsabili dell'espressione nella biopsia tumorale.

Molti dei geni che svolgono funzioni importanti nell'embriogenesi e nella differenziazione delle cellule staminali, se mutati o espressi in modo aberrante, possono indurre trasformazione neoplastica. Allo stesso modo, anche quei miR che normalmente regolano finemente la differenziazione delle cellule e lo sviluppo delle loro specifiche funzionalità potrebbero, qualora espressi in modo aberrante, mediarne la trasformazione neoplastica. Sulla base di quest'ipotesi, in futuro, ci proponiamo di investigare il significato biologico di quei miR che risultano differenzialmente regolati nella maturazione e trasformazione neoplastica delle cellule T.

Quest'ultima parte del lavoro è stata portata avanti in collaborazione con il laboratorio del Dott. Rosolen (Dipartimento di Pediatria, Università di Padova, Padova) e della Dott.ssa Basso (Institute of Cancer Genetics, Columbia University, New York, USA).

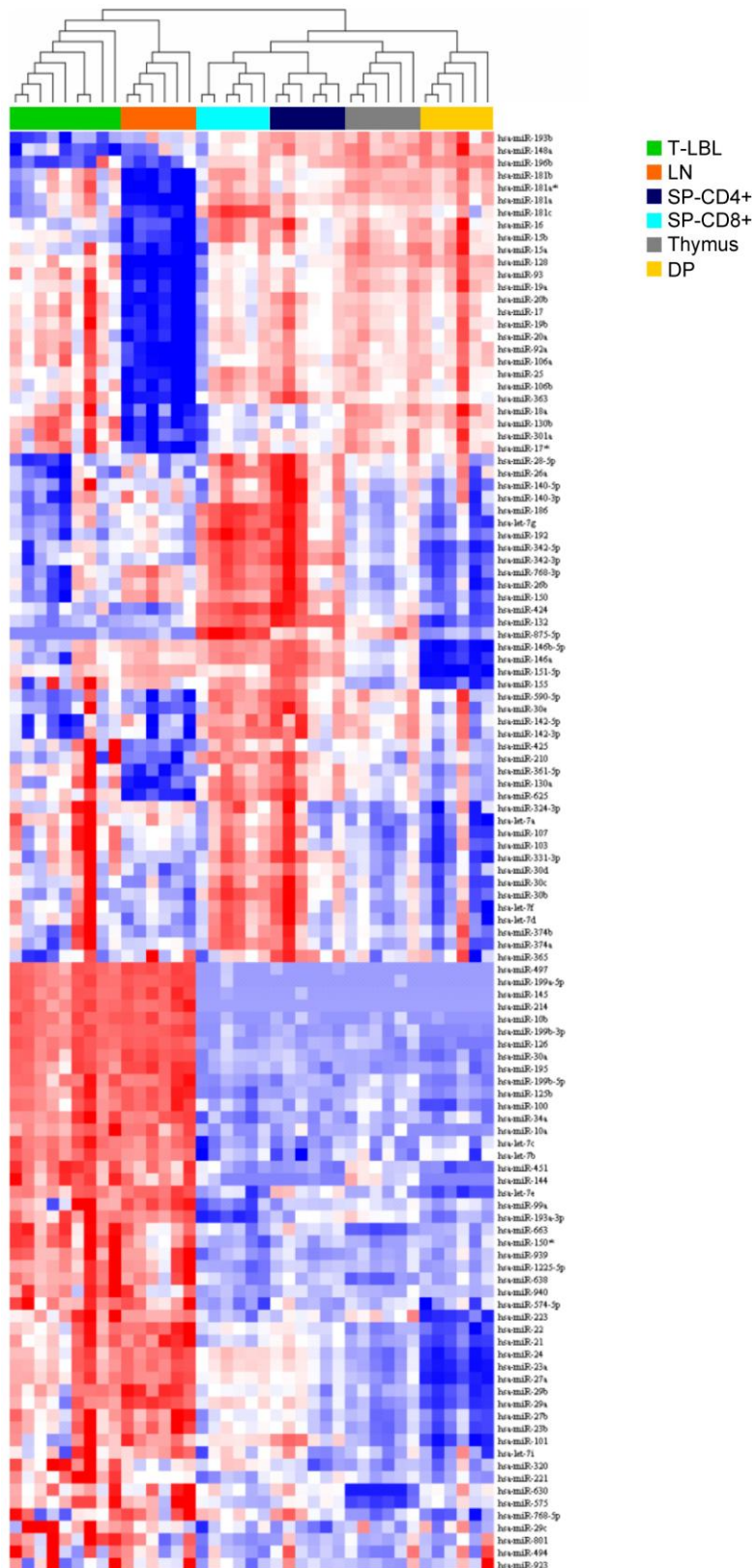


Figura 4.17. Heatmap rappresentante l'espressione differenziale dei miR in campioni di T-LBL (9), LN (6), in tessuto timico (6) ed in sottopopolazioni timocitarie umane (6 DP, 6 SP CD4, 6 SP CD8). Il dendrogramma riportato nella figura è stato generato utilizzando un algoritmo di *clustering* gerarchico fornito dalla piattaforma geWorkbench (<http://www.geworkbench.org>). L'analisi è stata condotta includendo solo i miR con valore di segnale d'espressione >100 in almeno uno dei fenotipi (116).

5. DISCUSSIONE

In questi anni sta emergendo sempre di più l'importanza dei miR nella regolazione di diversi processi fisiologici e patologici. I miR agiscono reprimendo a livello post-trascrizionale l'espressione genica dei propri target mediante la de-stabilizzazione dell'mRNA ed il blocco della traduzione in proteina. Ogni singolo miR agisce su molti mRNA target diversi (in alcuni casi centinaia) e, a sua volta, un mRNA può essere bersaglio di diversi miR, che agendo in modo concertato possono indurre diversi livelli di repressione [29]. Proprio grazie al loro complesso meccanismo d'azione i miR possono regolare finemente le funzioni cellulari e guidare processi che prevedono la modulazione coordinata di gruppi di geni, come la differenziazione cellulare.

Diversi studi di espressione dei miR hanno mostrato che il *pattern* d'espressione di queste molecole varia nel corso dello sviluppo ematopoietico, della risposta immunitaria e della leucemogenesi, suggerendo quindi che i miR possano giocare un ruolo importante in questi processi. A questo proposito, studi recenti hanno dimostrato che specifici miR svolgono funzioni essenziali in diverse fasi dell'ematopoiesi e della risposta immunitaria [37,41,48-51].

Il processo di differenziazione dei linfociti T è un processo complesso che prevede il passaggio dei precursori T attraverso una serie di cambiamenti genotipici e fenotipici che terminano nell'acquisizione del fenotipo maturo e delle capacità funzionali specifiche dei diversi tipi di cellule T. Nel corso di questo processo i precursori T, detti timociti, subiscono delle importanti modificazioni della loro espressione genica che sono state in passato e sono tutt'oggi oggetto di intenso studio. Al contrario, il ruolo dei miR nello sviluppo delle cellule T è ancora in gran parte sconosciuto ed i pochi studi condotti finora sull'argomento sono stati tutti realizzati in modelli murini. Questo probabilmente dipende dalla difficoltà nel reperire tessuto timico umano, rispetto ai linfociti maturi, che si possono facilmente ottenere dal sangue periferico di donatori. Infatti i timociti umani sono ottenuti da frammenti operatori prelevati da bambini che vengono sottoposti a timectomia parziale nel corso di interventi chirurgici per difetti cardiaci congeniti. Dato che sono note delle differenze nella maturazione timocitaria fra uomo e topo, sarebbe importante studiare anche nell'uomo come viene regolata l'espressione dei miR in questo processo.

Allo scopo quindi di identificare miR coinvolti nella differenziazione delle cellule T nell'uomo, in questo lavoro di tesi abbiamo analizzato tramite microarray il profilo

d'espressione dei miR in timociti umani non frazionati ed in sottopopolazioni timiche a tre diversi stadi di differenziazione: DP, SP CD4 e SP CD8. La sottopopolazione DN ($CD2^+/CD4^+/CD8^-$), che rappresenta una percentuale molto bassa dei timociti totali (circa 2-4%), è stata esclusa dall'analisi per insufficienza di materiale disponibile. Inoltre, allo scopo di approfondire la nostra conoscenza sull'espressione dei microRNA nella linea linfoide T, sono state generate delle librerie di small-RNA da timociti umani non frazionati, timociti DP e linfociti T $CD4^+$ e $CD8^+$ maturi da sangue periferico.

Dall'analisi dei microarray è risultato che le sottopopolazioni timocitarie considerate presentano profili d'espressione dei miR distinti che riflettono le relazioni maturative fra gli stadi di sviluppo dei precursori T. In particolare, la popolazione più immatura dei DP presenta un *pattern* di miR che si discosta nettamente da quello delle due popolazioni mature SP, le quali hanno profili d'espressione più simili fra loro. E' interessante osservare una generale up-regolazione dei miR nel passaggio dei timociti dallo stadio DP a quello SP, come se l'aumento dell'espressione dei miR fosse associato al processo maturativo. Generalmente, il *fold-change* dei miR down-regolati è molto più modesto di quello dei miR up-regolati. Anche i dati di sequenziamento delle librerie mostrano, in accordo con i risultati dell'esperimento di microarray, un tendenza all'up-regolazione dell'espressione dei miR (30 miR up-regolati contro 11 down-regolati) nella transizione dallo stadio DP dei timociti ai linfociti maturi del sangue periferico. Poiché i miR agiscono silenziando a livello post-trascrizionale i loro geni bersaglio, in questo contesto la loro up-regolazione potrebbe avere la funzione di spegnere geni legati alla staminalità, inducendo la differenziazione cellulare. Da questo punto di vista risulta interessante l'osservazione di diversi gruppi di ricerca [54,55] che hanno mostrato una generale down-regolazione dei miR maturi nei tumori rispetto alle controparti normali, spesso accompagnata dall'up-regolazione di specifici gruppi di miR. Questa osservazione suggerisce che i bassi livelli di miR riflettano la regressione dello stato differenziativo delle cellule trasformate o, nel caso delle leucemie e dei linfomi, il blocco maturativo dei precursori linfocitari.

Anche in altri lavori che hanno studiato la modulazione dei miR nel corso di processi differenziativi [46] sono stati osservati prevalentemente arricchimenti transitori di specifici miR, piuttosto che deplezioni. D'altro canto il meccanismo stesso d'azione dei miR suggerisce che l'up-regolazione di un miR sia in grado di provocare un effetto funzionale più potente rispetto alla sua down-regolazione. Infatti, in genere un singolo mRNA è contemporaneamente bersaglio di differenti miR. Di conseguenza, la down-

modulazione di un miR può essere mitigata dall'azione di altri miR, che ne vicariano la funzione silenziante nei confronti dei target. Al contrario, l'up-regolazione di un miR va a perturbare tutto il sistema, intensificando l'effetto di repressione su tutti i suoi target.

Fra i miR differenzialmente regolati nel passaggio da DP a SP sono presenti alcuni miR che sono noti avere un ruolo importante nella differenziazione ematopoietica del lineage linfoide (come miR-150 e miR-155) e mieloide (come miR-223 e miR-146), nonché nella risposta immunitaria (come miR-146). Sorprendentemente fra i miR differenzialmente regolati nella transizione DP-SP non appare nessun membro della famiglia del miR-181. In precedenza, studiando il *pattern* d'espressione dei miR nel corso della differenziazione T nel topo, Neilson et al. avevano identificato proprio miR-181a come altamente espresso nel timo e specificamente up-regolato in corrispondenza dello stadio DP dello sviluppo. L'up-regolazione di miR-181a in questo stadio correlava inversamente con l'espressione di Bcl-2, CD69 e TCR α , geni coinvolti nella selezione positiva dei timociti, successivamente dimostrati essere target di miR-181 [46]. In un altro lavoro, Li et al. hanno dimostrato che miR-181 influenza nelle cellule T la sensibilità del TCR all'antigene. In questo studio, l'espressione di miR-181 veniva trovata up-regolata nei timociti immaturi DP rispetto ai linfociti T maturi del sangue periferico, in cui i livelli di miR-181 si abbassano per limitare la reattività linfocitaria solo agli antigeni non-self ad alta affinità [47]. Nonostante i nostri dati di microarray non confermino la differenza d'espressione di miR-181 fra timociti DP e SP, il sequenziamento delle librerie di small-RNA ottenute dalle popolazioni linfoidi ha identificato due membri della famiglia di miR-181 fra i miR differenzialmente espressi nel confronto fra i timociti DP ed i linfociti maturi CD4⁺ e CD8⁺. Nello specifico, miR-181b e miR-181a sono rispettivamente il primo ed il terzo miR maggiormente up-regolati nei timociti DP rispetto ad entrambe le popolazioni linfocitarie mature. Tali miR risultano essere espressi a livelli elevati anche nei timociti non frazionati. Secondo i nostri dati, quindi sembra che miR-181 nell'uomo venga up-regolato nei timociti DP e mantenuto a livelli elevati nei timociti SP CD4 e CD8, per poi essere fortemente down-regolato quando questi escono dal timo per migrare in periferia come linfociti T maturi.

Un altro miR con un *pattern* d'espressione particolarmente interessante è miR-151-5p. Questo miR viene fortemente up-regolato nella maturazione timocitaria dallo stadio DP allo stadio SP. Tuttavia, l'entità di questa up-regolazione è diversa a seconda che i timociti DP si differenzino in senso CD4 (FC=108) o CD8 (FC=30). Nelle librerie di small-RNA, questo miR non viene trovato differenzialmente espresso nel passaggio da

DP a linfociti T maturi periferici. Questo suggerisce che miR-151-5p sia implicato in processi specifici del passaggio DP/SP intra-timico e che abbia un ruolo importante nella scelta differenziativa in senso CD4 *versus* CD8. In futuro, sarebbe interessante approfondire il ruolo di questo miR nel processo di selezione positiva e negativa dei timociti e della differenziazione in senso CD4 o CD8.

La generazione di librerie di small-RNA ci ha permesso non solo di studiare nelle popolazioni linfoidi d'interesse l'espressione dei miR noti, ma anche di identificare potenziali nuovi miR. Complessivamente, il sequenziamento delle nostre librerie ha portato all'identificazione di 139 miR noti e di 98 sequenze di candidati nuovi miR. Mediante un'analisi *bootstrap* [80,81] è stato stimato che per tutte e quattro le librerie generate il *set* di miR maturi identificati dal sequenziamento rappresenta più dell'80% della totalità dei miR espressi da ogni campione. Il numero di miR noti in tutte e quattro le librerie è superiore a quello dei nuovi miR. Questo è indice del fatto che buona parte dei miR espressi nei tessuti umani è già stata identificata ed annotata nei *database on-line*. I miR attualmente ancora sconosciuti sembrano essere principalmente quelli specificamente espressi da particolari sottopopolazioni cellulari o in determinate condizioni fisiologiche. Questa ultima affermazione è supportata dal fatto che nei timociti totali i miR nuovi rappresentano solo il 26 %, mentre nella sottopopolazione timocitaria DP rappresentano il 41,5 % della totalità dei miR maturi sequenziati. In questa direzione va anche il fatto che il 40% dei miR noti rilevati nelle librerie contro il 30% dei candidati nuovi miR sia espresso in modo condiviso da DP, LT CD4 e LT CD8.

Un altro dato interessante emerso dall'analisi delle librerie di small-RNA delle popolazioni linfoidi è l'esistenza di miR "dominanti" espressi a livelli alti ed ubiquitari da tutte e quattro le popolazioni T considerate, quale miR-142-3p e miR-16. In particolare, miR-142-3p, già identificato come marcatore delle popolazioni linfocitarie in altri studi, costituisce da solo più del 40% della totalità dei miR espressi dai linfociti T CD4⁺ e CD8⁺, nonché dai timociti DP e non frazionati. Altri tre miR (let-7a, miR-150 e miR-21) hanno invece conte altissime solo nei linfociti CD4⁺ e CD8⁺ maturi, mentre sono espressi a livelli molto più bassi nelle popolazioni timiche. Questo risultato è in accordo con un altro lavoro, condotto però nel modello murino, che ha analizzato con tre diverse metodiche (generazione di librerie di small-RNA, miR microarray e real-time PCR) la modulazione dell'espressione dei miR in cellule T CD8 naive, effettrici e

memoria. Anche in questo studio era stato individuato un gruppo ristretto di miR espresso in modo dominante nelle cellule T CD8: questi 7 miR dominanti (miR-16, miR-21, miR-142-3p, miR-142-5p, miR-150, miR-15b, let-7f) da soli costituivano per frequenza il 60% dei miR espressi dalle tre popolazioni T considerate. L'espressione di questi 7 miR era dinamicamente regolata nel corso della differenziazione delle cellule T stimulate dall'antigene. In generale, il livello d'espressione dei miR appariva inversamente correlato allo stato di attivazione delle cellule. Infatti, le cellule CD8 effettrici mostravano una globale down-regolazione nell'espressione dei miR (compresi 6 dei 7 miR dominanti sopracitati) rispetto alle cellule naive, mentre nelle cellule memoria i livelli di espressione dei miR tendevano a riaumentare, pur rimanendo più bassi in confronto a quelli delle cellule naive [89].

Allo scopo di identificare miR implicati essenzialmente nel processo differenziativo e non nelle specifiche funzionalità dei linfociti T CD4⁺ o CD8⁺, analizzando le librerie di small-RNA, abbiamo individuato una lista di miR noti differenzialmente regolati nel passaggio dai timociti immaturi DP ai linfociti T maturi, in modo concorde per i LT CD8 ed i LT CD4. Successivamente abbiamo confrontato questo gruppo di miR (ottenuti dall'analisi delle librerie di small-RNA) con la lista di miR risultati differenzialmente regolati nel passaggio dei timociti dallo stadio DP a quello SP (ottenuta dall'analisi dei miR microarray). In questo modo, è stato individuato un gruppo di 18 miR che risultano regolati nel passaggio dei timociti dallo stadio DP a quello SP e che vengono mantenuti modulati nello stesso senso anche quando le cellule T migrano dal timo alla periferia. All'interno di questo gruppo di miR si ritrovano molti miR variabili nel corso della maturazione T con *fold-change* particolarmente elevati, come per esempio miR-146a e b, miR-27 e miR-150. La consistente modulazione di questi 18 miR nel corso della differenziazione delle cellule T, ci ha portato ad ipotizzare che essi possano svolgere un ruolo importante in questo processo. Inoltre, alcuni miR presenti in questi gruppo sono già stati riportati in letteratura avere un'espressione sregolata in alcuni tipi di neoplasie ematologiche [51,59,65,72,73]. Di fatto, dal confronto del profilo di espressione dei microRNA nelle sottopopolazioni timiche umane, nei linfonodi reattivi e nei campioni di linfoma linfoblastico T è emerso che non solo i linfomi mostrano un *pattern* d'espressione dei miR distinto sia da quello dei timociti, sia da quello dei linfonodi reattivi non-neoplastici, ma anche che i 25 miR maggiormente regolati nei timociti nel passaggio da DP a SP risultano tutti (a parte

miR-128) differenzialmente espressi nei campioni di T-LBL rispetto ad almeno una delle popolazioni timiche normali. Di questi 25 miR, 11 appartengono al sopraccitato gruppo di 18 miR consistentemente regolati nel corso della maturazione dei precursori timici a linfociti T. Naturalmente queste osservazioni sono preliminari e necessitano di ulteriori approfondimenti e validazioni; tuttavia, sembrerebbero suggerire che, similmente a quanto è stato osservato per i geni codificanti proteine, alcuni dei miR che normalmente regolano finemente la differenziazione delle cellule e lo sviluppo delle loro specifiche funzionalità possano, qualora espressi in modo aberrante, mediarne la trasformazione neoplastica.

Dopo averne validato tramite real-time PCR il profilo d'espressione nelle sottopopolazioni timocitarie d'interesse, abbiamo deciso di investigare la funzione di alcuni dei miR (n=9) identificati, mediante analisi dei microarray e/o delle librerie di small-RNA, come consistentemente regolati nel corso dello sviluppo delle cellule T.

La scelta dei miR è stata dettata principalmente, oltre che dalla loro consistente regolazione differenziale nel corso della maturazione T, anche dalla presenza in letteratura di dati riguardanti il loro coinvolgimento in processi quali la differenziazione cellulare e la tumorigenesi. Allo scopo quindi di comprendere meglio la funzione nelle cellule T di questi 9 miR selezionati, abbiamo fatto una ricerca computazionale dei loro potenziali mRNA bersaglio.

L'identificazione dei geni target dei miR è attualmente affidata ad una scienza ancora molto problematica ed incerta. Infatti, il meccanismo di riconoscimento e silenziamento dei bersagli da parte dei miR è molto complesso e ancora poco compreso: in questi anni sono state identificate alcune caratteristiche dell'appaiamento del miR con l'UTR-3' del mRNA bersaglio che sono essenziali (anche se spesso non sufficienti) per l'efficacia del *targeting*. Si è capito che per il riconoscimento del target è fondamentale che ci sia perfetta complementarità fra il *seed* all'estremità 5' del miR e l'UTR-3' del mRNA bersaglio. In aggiunta, alcuni lavori hanno recentemente identificato una serie di caratteristiche del contesto dell'UTR-3' dell'mRNA che contribuiscono ad aumentare l'efficacia e la specificità di *targeting* dei miR [31,84].

I software di predizione di target disponibili al momento *on-line* utilizzano per elaborare la predizione criteri basati sulle caratteristiche di complementarità di sequenza fra microRNA e l'UTR-3' degli mRNA, sull'energia libera di formazione del *duplex* miRNA:mRNA, sull'accessibilità della struttura secondaria dell'mRNA a livello del

supposto sito di legame e sulla conservazione inter-specie del sito di legame stesso nell'UTR-3'. Le liste di potenziali target generate da diversi software di predizione sono solo parzialmente sovrapponibili. Questo dipende dal fatto che gli algoritmi utilizzati da questi software sono basati su criteri e parametri di selezione in parte simili ed in parte distinti. Inoltre, la predizione fatta da un singolo software spesso contiene un numero molto elevato di falsi positivi. Di conseguenza, in genere per la scelta del candidato target da validare sperimentalmente si tende ad incrociare le predizioni di più software diversi, considerando solo i target comuni per le successive analisi. Anche nel nostro caso, per identificare i potenziali bersagli dei 9 miR selezionati, abbiamo quindi integrato i risultati di tre diversi software di predizione di target (MiRanda, TargetScan e PicTar), generando una lista di candidati target comuni a tutte e tre le predizioni. Questo ha ridotto moltissimo il numero di target predetti, limitando molto il numero di falsi positivi, ma inevitabilmente eliminando anche un certo numero di falsi negativi. Fra i candidati target presenti nella lista sono stati identificati molti geni coinvolti nella differenziazione e nella proliferazione cellulare, nonché nella risposta immunitaria. In particolare, è stata notata la presenza di diversi membri della famiglia dei geni HOX (HOXA5, HOXC8, HOXB7, HOXB6, HOXA13 e HOXD10) e di alcuni componenti del pathway di Notch (NOTCH3, DTX1).

In parallelo, si è deciso di utilizzare un'altra metodica per l'identificazione di potenziali target dei miR d'interesse. La maggior parte dei miR nei mammiferi regola negativamente l'espressione genica dei propri target agendo sia a livello di proteina, che di trascritto. Sulla base di questo dato, diversi gruppi di ricerca hanno fatto delle analisi di correlazione tra il profilo d'espressione genica ed il profilo d'espressione dei miR, allo scopo di individuare trascritti differenzialmente regolati contenenti nel loro UTR-3' *seed match* per i miR modulati in senso opposto negli stessi campioni [83-86]. Allo scopo di fare questo tipo di confronto, abbiamo analizzato il profilo d'espressione genica dei timociti DP e SP CD4, generando una lista di 1665 trascritti differenzialmente regolati nel corso del passaggio fra questi due stadi maturativi. In seguito, abbiamo comparato questa lista di trascritti con la lista di bersagli potenziali generata dal software di predizione miRanda per i primi 17 miR up-regolati ed i primi 2 miR down-regolati con maggiore *fold-change* nel corso della differenziazione timocitaria da DP a SP. Abbiamo scelto di utilizzare per questa analisi la predizione di un solo software poichè incrociando le predizioni di più software, si tende a perdere molti candidati falsi negativi e si ottengono liste di target troppo piccole ai fini del

confronto con i profili d'espressione genica. In questo modo abbiamo individuato un ristretto gruppo di candidati target computazionali (252 mRNA target su 15155 predetti) che risultano differenzialmente regolati nel passaggio da DP a SP CD4 in senso opposto ai miR di cui sono predetti bersagli. Il risultato di questa seconda analisi sembra confermare l'elevato numero di falsi positivi presenti nelle liste di geni prodotte dai software di predizione di target attualmente disponibili *on-line*. D'altra parte è possibile che alcuni dei target predetti non siano espressi nei campioni considerati o siano regolati principalmente a livello di proteina da parte dei miR. Per questi casi naturalmente l'analisi del profilo d'espressione genica dei campioni è poco informativa.

Il basso numero di target individuati con questa metodica si riduce ulteriormente nel momento in cui si incrocia questa lista con quella dei target comuni predetti dai 3 diversi software bioinformatici.

Combinando queste due metodiche di ricerca dei target, abbiamo individuato un gene presente nella lista degli mRNA target comunemente predetti per miR-150 da tre software bioinformatici diversi e regolato in modo opposto al miR-150 nel passaggio dei timociti dallo stadio DP a quello di SP CD4. Questo gene (che abbiamo chiamato *Candidate Target 1*) è Notch3, un membro della famiglia dei recettori Notch, che svolge un ruolo fondamentale nello sviluppo dei linfociti T e nella loro trasformazione neoplastica [16].

L'allineamento fra la regione del *seed* in 5' (nt 2-8) del miR-150 e l'UTR-3' di Notch3 è perfetto, conservato nell'uomo e nel topo e presenta una serie di caratteristiche che fanno di Notch3 un buon candidato, appunto, come bersaglio dell'azione del miR-150. Basandoci su questi dati, abbiamo deciso di approfondire sperimentalmente la relazione fra miR-150 ed il nostro candidato bersaglio d'interesse.

Al fine di validare l'associazione fra miR-150 e *Candidate Target 1*, al momento, abbiamo clonato all'interno di un plasmide reporter l'UTR-3' di Notch3 umano a valle del gene della luciferasi. Nei prossimi mesi co-trasferiremo questo plasmide reporter Luc-UTR-3' insieme a diverse dosi di miR-mimic-150 o di un miR di controllo e valuteremo le variazioni dell'attività della luciferasi in presenza ed in assenza dello specifico miR d'interesse. Questa prova ci permetterà di validare direttamente il *targeting* di Notch3 da parte di miR-150.

Inoltre, dopo aver up-regolato l'espressione di miR-150 in linee cellulari esprimenti Notch3, valuteremo tramite W.B. e real-time PCR le variazioni a livello di proteina e di mRNA del candidato bersaglio.

MiR-150 fa parte del gruppo di 18 miR consistentemente regolati nel corso dello sviluppo delle cellule T. L'up-regolazione di questo miR nei timociti durante la transizione dallo stadio DP a quello di SP è stata validata mediante real-time PCR. Dalla letteratura è noto che miR-150 svolge un ruolo molto importante nello sviluppo sia dei linfociti B, che dei linfociti T. Questo miR è espresso in modo selettivo in cellule B e T mature e non attivate, ma non nei loro progenitori. L'over-espressione di miR-150 porta da una parte ad un blocco dello sviluppo dei linfociti B allo stadio pro-B, associata ad una ridotta vitalità cellulare *in vitro*; dall'altra ad un'inibizione della transizione dei timociti dallo stadio DN3 allo stadio DN4 [52-53]. Uno *screening* dei miR de-regolati nella ALL (sia B, che T) recentemente pubblicato [72] ha individuato miR-150 fra i quattro miR maggiormente down-regolati nelle leucemie acute rispetto alla controparte normale. Bersaglio validato di miR-150 è c-Myb, fattore trascrizionale essenziale nello sviluppo ematopoietico ed in particolare nella differenziazione T. E' stato dimostrato che l'over-espressione di v-Myb nel timo porta allo sviluppo di T-ALL nei topi [90]. Inoltre, la duplicazione del gene c-Myb è presente in circa l'8,4% dei casi di T-ALL umana [91]. Considerando i dati in nostro possesso e quelli generati in precedenza da altri gruppi di ricerca, abbiamo deciso di studiare gli effetti funzionali dell'espressione di miR-150 in linee di T-ALL. Abbiamo così osservato che miR-150 è espresso a livelli molto bassi in tutte le linee di T-ALL considerate rispetto sia ai timociti DP, sia ai timociti maturi SP CD4. Le diverse linee di T-ALL, pur presentando tutti livelli piuttosto bassi del miR-150, presentano fra loro una certa variabilità nei valori d'espressione di questo miR. Viene spontaneo chiedersi se questo abbassamento generalizzato dei livelli di miR-150 dia un vantaggio selettivo a queste cellule e se la variabilità osservata fra le linee di T-ALL nella sua espressione possa avere un significato funzionale. L'over-espressione di miR-150 nella linea di T-ALL Jurkat induce una significativa riduzione del tasso di proliferazione cellulare rispetto ai controlli. Inoltre, da una prima analisi del ciclo cellulare sembra che le cellule trattate con il vettore over-esprimente pre-miR-150 vadano incontro ad un rallentamento del ciclo associato ad un accumulo delle cellule nella fase G2, precedente alla mitosi. Le cellule trattate con il vettore over-esprimente miR-150 mostrano un aumento della frazione dei detriti cellulari, suggerendo che in definitiva la forzata espressione di questo miR nelle cellule di T-ALL le induca ad andare in apoptosi. Nei prossimi mesi, ci proponiamo di approfondire l'effetto dell'espressione ectopica di mir-150 sull'induzione di apoptosi e sul blocco del ciclo cellulare nelle Jurkat, nonché in altre

linee di T-ALL. Inoltre, sarebbe interessante valutare in futuro se negli effetti dell'over-espressione di miR-150 sulle linee di T-ALL sia implicato il silenziamento di Notch3 (*Candidate Target 1*), del target validato c-Myb, o piuttosto di altri target di questo miR ancora sconosciuti, magari implicati nel punto di controllo della transizione G2-M.

6. **BIBLIOGRAFIA**

1. Abbas AK. *Maturazione linfocitaria ed espressione dei geni del recettore per l'antigene*. In: Immunologia Cellulare e Molecolare. AK Abbas – AH Lichtman, V ed., Elsevier Italia Srl, Milano, 2006; 7:131-163.
2. Zlotoff DA, Schwarz BA, Bhandoola A: *The long road to the thymus: the generation, mobilization, and circulation of T-cell progenitors in mouse and man*. Semin Immunopathol. 2008; 30(4):371-82.
3. Ladi E, Yin X, Chtanova T, Robey EA. *Thymic microenvironments for T cell differentiation and selection*. Nat Immunol. 2006; 7(4):338-43.
4. Iwasaki H, Akashi K. *Thymus exclusivity: all the right conditions for T cells*. Immunity. 2006; 25(5):697-700.
5. Wu L. *T lineage progenitors: the earliest steps en route to T lymphocytes*. Curr Opin Immunol. 2006; 18(2):121-6.
6. Chi AW, Bell JJ, Zlotoff DA, Bhandoola A. *Untangling the T branch of the hematopoiesis tree*. Curr Opin Immunol. 2009; 21(2):121-6.
7. Weerkamp F, Pike-Overzet K, Staal FJ. *T-sing progenitors to commit*. Trends Immunol. 2006; 27(3):125-31.
8. Dik WA, Pike-Overzet K, Weerkamp F et al. *New insights on human T cell development by quantitative T cell receptor gene rearrangement studies and gene expression profiling*. J Exp Med. 2005; 201(11):1715-23.
9. Mueller DL. *Mechanisms maintaining peripheral tolerance*. Nat Immunol. 2010; 11(1):21-7.
10. Klein L, Hinterberger M, Wirnsberger G, Kyewski B. *Antigen presentation in the thymus for positive selection and central tolerance induction*. Nat Rev Immunol. 2009; 9(12):833-44.
11. Schnell S, Demolliere C, van den Berk P, Kirberg J, Jacobs H. *Constitutive expression of the pre-TCR enables development of mature T cells*. Int Immunol 2006; 18: 911-20.
12. Graux C, Cools J, Michaux L, Vandenberghe P, Hagemeijer A. *Cytogenetics and molecular genetics of T-cell acute lymphoblastic leukemia: from thymocyte to lymphoblast*. Leukemia 2006; 20:1496–1510.
13. Aifantis I, Raetz E, Buonamici S. *Molecular pathogenesis of T-cell leukaemia and lymphoma*. Nat Rev Immunol 2008; 8:380-90.
14. O'Neil J and Look AT. *Mechanisms of transcription factor deregulation in lymphoid cell transformation*. Oncogene 2007; 26:6838-6849.

15. Weng AP, Ferrando AA et al. *Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia*. Science 2004 306:269-71.
16. Bellavia D, Campese AF, Vacca A, Gulino A, Screpanti I. *Notch3, another Notch in T cell development*. Semin Immunol. 2003;15(2):107-12.
17. Koch U, Radtke F. *Notch and cancer: a double-edged sword*. Cell Mol Life Sci 2007; 64:2746-62.
18. Weerkamp F, van Dongen JJM and Staal FJT. *Notch and Wnt signaling in T-lymphocyte development and acute lymphoblastic leukemia*. Leukemia 2006; 20:1197-1205.
19. Radtke F, Wilson A, Mancini SJ, MacDonald HR. *Notch regulation of lymphocyte development and function*. Nat Immunol 2004; 5:247-53.
20. Sambandam A, Maillard I et al. *Notch signaling controls the generation and differentiation of early T lineage progenitors*. Nat Immunol 2005; 6:663-70.
21. Rothenberg EV. *Notchless T cell maturation?* Nat Immunol 2001; 2:189-90.
22. Spits H. *Development of alphabeta T cells in the human thymus*. Nat Rev Immunol 2002; 2:760-72.
23. Aster JC. *Deregulated Notch signalling in acute T-cell lymphoblastic leukemia/lymphoma: new insights, questions, and opportunities*. Int J Hematol 2005; 82(4):295-301.
24. Bellavia D, Campese AF et al. *Constitutive activation of NF-kappaB and T-cell leukemia/lymphoma in Notch3 transgenic mice*. EMBO J. 2000; 19(13):3337-48.
25. Thomas DA and Kantarjian HM. *Lymphoblastic lymphoma*. In: Advances in the treatment of adult acute lymphocytic leukemia-Part II. Hematology/Oncology Clinics of North America, 2001; 15(1):51-95.
26. Shomron N, Levy C. *MicroRNA-biogenesis and Pre-mRNA splicing crosstalk*. J Biomed Biotechnol. 2009; 2009:594678.
27. Shomron N, Golan D, Hornstein E. *An evolutionary perspective of animal microRNAs and their targets*. J Biomed Biotechnol. 2009; 2009:594738.
28. Kim YK, Kim VN. *Processing of intronic microRNAs*. EMBO J. 2007; 26(3):775-83.
29. Winter J, Jung S, Keller S, Gregory RI, Diederichs S. *Many roads to maturity: microRNA biogenesis pathways and their regulation*. Nat Cell Biol. 2009; 11(3):228-34.
30. Breving K, Esquela-Kerscher A. *The complexities of microRNA regulation: mirandering around the rules*. Int J Biochem Cell Biol. 2009. [Epub ahead of print]

31. Grimson A, Farh KK et al. *MicroRNA targeting specificity in mammals: determinants beyond seed pairing*. Mol Cell. 2007; 27(1):91-105.
32. Nielsen CB, Shomron N et al. *Determinants of targeting by endogenous and exogenous microRNAs and siRNAs*. RNA. 2007; 13(11):1894-910.
33. Janowski BA, Younger ST et al. *Activating gene expression in mammalian cells with promoter-targeted duplex RNAs*. Nat Chem Biol. 2007; 3(3):166-73.
34. Li LC, Okino ST et al. *Small dsRNAs induce transcriptional activation in human cells*. Proc Natl Acad Sci U S A. 2006; 103(46):17337-42.
35. Baek D, Villén J et al. *The impact of microRNAs on protein output*. Nature. 2008; 455(7209):64-71.
36. Selbach M, Schwanhäusser B et al. *Widespread changes in protein synthesis induced by microRNAs*. Nature. 2008; 455(7209):58-63.
37. Lawrie CH. *MicroRNA and hematology: small molecules, big function*. Br.J.Haematol. 2007, 137(6):503-12.
38. Chen CZ, Li L, Lodish HF, Bartel DP. *MicroRNAs modulate hematopoietic lineage differentiation*. Science. 2004; 303(5654):83-6.
39. Ramkissoon SH, Mainwaring LA et al. *Hematopoietic-specific microRNA expression in human cells*. Leuk Res. 2006; 30(5):643-7.
40. Georgantas RW 3rd, Hildreth R et al. *CD34+ hematopoietic stem-progenitor cell microRNA expression and function: a circuit diagram of differentiation control*. Proc Natl Acad Sci U S A. 2007; 104(8):2750-5.
41. Lu LF, Liston A. *MicroRNA in the immune system, microRNA as an immune system*. Immunology. 2009; 127(3):291-8.
42. Muljo SA, Ansel KM et al. *Aberrant T cell differentiation in the absence of Dicer*. J Exp Med. 2005; 202(2):261-9.
43. Cobb BS, Nesterova TB et al. *T cell lineage choice and differentiation in the absence of the RNase III enzyme Dicer*. J Exp Med. 2005; 201(9):1367-73.
44. Cobb BS, Hertweck A et al. *A role for Dicer in immune regulation*. J Exp Med. 2006; 203(11):2519-27.
45. Xiao C, Rajewsky K. *MicroRNA control in the immune system: basic principles*. Cell. 2009; 136(1):26-36.

46. Neilson JR, Zheng GX, Burge CB, Sharp PA. *Dynamic regulation of miRNA expression in ordered stages of cellular development.* Genes Dev. 2007; 21(5):578-89.
47. Li QJ, Chau J et al. *miR-181a is an intrinsic modulator of T cell sensitivity and selection.* Cell. 2007; 129(1):147-61.
48. Liang TJ, Qin CY. *The emerging role of microRNAs in immune cell development and differentiation.* APMIS. 2009; 117(9):635-43.
49. Bi Y, Liu G, Yang R. *MicroRNAs: novel regulators during the immune response.* J Cell Physiol.; 218(3):467-72.
50. Schickel R, Boyerinas B, Park SM, Peter ME. *MicroRNAs: key players in the immune system, differentiation, tumorigenesis and cell death.* Oncogene. 2008; 27(45):5959-74.
51. Vasilatou D, Papageorgiou S et al. *The role of microRNAs in normal and malignant hematopoiesis.* Eur J Haematol. 2009. [Epub ahead of print]
52. Zhou B, Wang S et al. *miR-150, a microRNA expressed in mature B and T cells, blocks early B cell development when expressed prematurely.* Proc Natl Acad Sci U S A. 2007; 104(17):7080-5.
53. Xiao C, Calado DP et al. *MiR-150 controls B cell differentiation by targeting the transcription factor c-Myb.* Cell. 2007; 131(1):146-59.
54. Lu J, Getz G et al. *MicroRNA expression profiles classify human cancers.* Nature. 2005; 435(7043):834-8.
55. Calin GA, Croce CM. *MicroRNA signatures in human cancers.* Nat Rev Cancer. 2006; 6(11):857-66.
56. Calin GA, Sevignani C et al. *Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers.* Proc Natl Acad Sci U S A. 2004; 101(9):2999-3004.
57. Hurst DR, Edmonds MD, Welch DR. *Metastamir: the field of metastasis-regulatory microRNA is spreading.* Cancer Res. 2009; 69(19):7495-8.
58. Nimmo RA, Slack FJ. *An elegant miRror: microRNAs in stem cells, developmental timing and cancer.* Chromosoma. 2009; 118(4):405-18.
59. Fabbri M, Croce CM, Calin GA. *MicroRNAs in the ontogeny of leukemias and lymphomas.* Leuk Lymphoma. 2009; 50(2):160-70.
60. Calin GA, Dumitru CD et al. *Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia.* Proc Natl Acad Sci U S A. 2002; 99(24):15524-9.

61. Cimmino A, Calin GA et al. *MiR-15 and miR-16 induce apoptosis by targeting BCL2*. Proc Natl Acad Sci U S A. 2005; 102(39):13944-9.
62. Calin GA, Cimmino A et al. *MiR-15a and miR-16-1 cluster functions in human leukemia*. Proc Natl Acad Sci U S A. 2008; 105(13):5166-71.
63. Calin GA, Ferracin M et al. *A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia*. N Engl J Med. 2005; 353(17):1793-801.
64. Marton S, Garcia MR et al. *Small RNAs analysis in CLL reveals a deregulation of miRNA expression and novel miRNA candidates of putative relevance in CLL pathogenesis*. Leukemia. 2008; 22(2):330-8.
65. Yendamuri S, Calin GA. *The role of microRNA in human leukemia: a review*. Leukemia. 2009; 23(7):1257-63.
66. Garzon R, Croce CM. *MicroRNAs in normal and malignant hematopoiesis*. Curr Opin Hematol. 2008; 15(4):352-8.
67. Roehle A, Hoefig KP et al. *MicroRNA signatures characterize diffuse large B-cell lymphomas and follicular lymphomas*. Br J Haematol. 2008; 142(5):732-44.
68. Chen J, Odenike O, Rowley JD. *Leukaemogenesis: more than mutant genes*. Nat Rev Cancer. 2010; 10(1):23-36.
69. Bueno MJ, Pérez de Castro I et al. *Genetic and epigenetic silencing of microRNA-203 enhances ABL1 and BCR-ABL1 oncogene expression*. Cancer Cell. 2008; 13(6):496-506.
70. Lum AM, Wang BB et al. *Retroviral activation of the mir-106a microRNA cistron in T lymphoma*. Retrovirology. 2007; 4:5.
71. Landais S, Landry S, Legault P, Rassart E. *Oncogenic potential of the miR-106-363 cluster and its implication in human T-cell leukemia*. Cancer Res. 2007; 67(12):5699-707.
72. Zanette DL, Rivadavia F et al. *MiRNA expression profiles in chronic lymphocytic and acute lymphocytic leukemia*. Braz J Med Biol Res. 2007; 40(11):1435-40.
73. Schotte D, Chau JC et al. *Identification of new microRNA genes and aberrant microRNA profiles in childhood acute lymphoblastic leukemia*. Leukemia. 2009; 23(2):313-22.
74. Sonoki T, Iwanaga E, Mitsuya H, Asou N. *Insertion of microRNA-125b-1, a human homologue of lin-4, into a rearranged immunoglobulin heavy chain gene locus in a patient with precursor B-cell acute lymphoblastic leukemia*. Leukemia. 2005; 19(11):2009-10.
75. Mi S, Lu J et al. *MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia*. Proc Natl Acad Sci U S A. 2007; 104(50):19971-6.

76. Li Z, Luo RT et al. *Consistent deregulation of gene expression between human and murine MLL rearrangement leukemias*. *Cancer Res*. 2009; 69(3):1109-16.
77. Indraccolo S, Tisato V et al. *Establishment and characterization of xenografts and cancer cell cultures derived from BRCA1 -/- epithelial ovarian cancers*. *Eur J Cancer*. 2006; 42(10):1475-83.
78. Leamon J.H., Braverman M.S and Rothberg J.M. *High throughput, massively parallel DNA sequencing technology for the era of personalized medicine*. *Gene therapy and regulation* 2007; 3(1):15-31.
79. Indraccolo S, Habeler W et al. *Gene transfer in ovarian cancer cells: a comparison between retroviral and lentiviral vectors*. *Cancer Res*. 2002; 62(21):6099-107.
80. Harrell, F. E. *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis*. Springer, New York, 2001.
81. Hinkley, A. C. D. a. D. V. *Bootstrap Methods and their Applications*. Cambridge University Press, New York, 1997.
82. Landgraf P, Rusu M et al. *A mammalian microRNA expression atlas based on small-RNA library sequencing*. *Cell* 2007; 129(7): 1401-14.
83. Basso K, Sumazin P et al. *Identification of the human B cell miRNome*. *Immunity* 2009; 30(5):744-52.
84. Mazière P, Enright AJ. et al. *Prediction of microRNA targets*. *Drug Discov Today* 2007; 12(11-12):452-8.
85. Farh KK, Grimson A et al. *The widespread impact of mammalian microRNAs on mRNA repression and evolution*. *Science* 2005; 310(5755): 1817-21.
86. Lim LP, Lau NC et al. *Microarray analysis shows that some microRNAs down-regulate large numbers of target mRNAs*. *Nature* 2005; 433(7027): 769-73.
87. Nicolas FE, Pais H et al. *Experimental identification of microRNA-140 targets by silencing and over-expressing miR-140*. *RNA* 2008; 14(12): 2513-20.
88. Li SS, Yu SL et al. *Target identification of microRNAs expressed highly in human embryonic stem cells*. *J Cell Biochem* 2009; 106(6): 1020-30.
89. Wu H, Neilson JR et al. *MiRNA profiling of naive, effector and memory CD8 T cells*. *PLoS One* 2007; 2(10):e1020.
90. Badiani PA, Kioussis D et al. *T-cell lymphomas in v-Myb transgenic mice*. *Oncogene*. 1996; 13(10):2205-12.

91. Lahortiga I, De Keersmaecker K et al. *Duplication of the MYB oncogene in T cell acute lymphoblastic leukemia*. Nat Genet 2007; 39(5):593-5.