



Università degli Studi di Cagliari

DOTTORATO DI RICERCA  
Terapia Pediatrica e Farmacologia dello Sviluppo  
Ciclo XXVIII

Determinanti genetici dell'espressione  
dell'emoglobina HbF

Settore scientifico disciplinare di afferenza  
MED/03

Presentata da:	Laura Manunza
Coordinatore Dottorato	Prof. Paolo Moi
Tutor	Prof. Paolo Moi

Esame finale anno accademico 2014 – 2015



## **Indice**

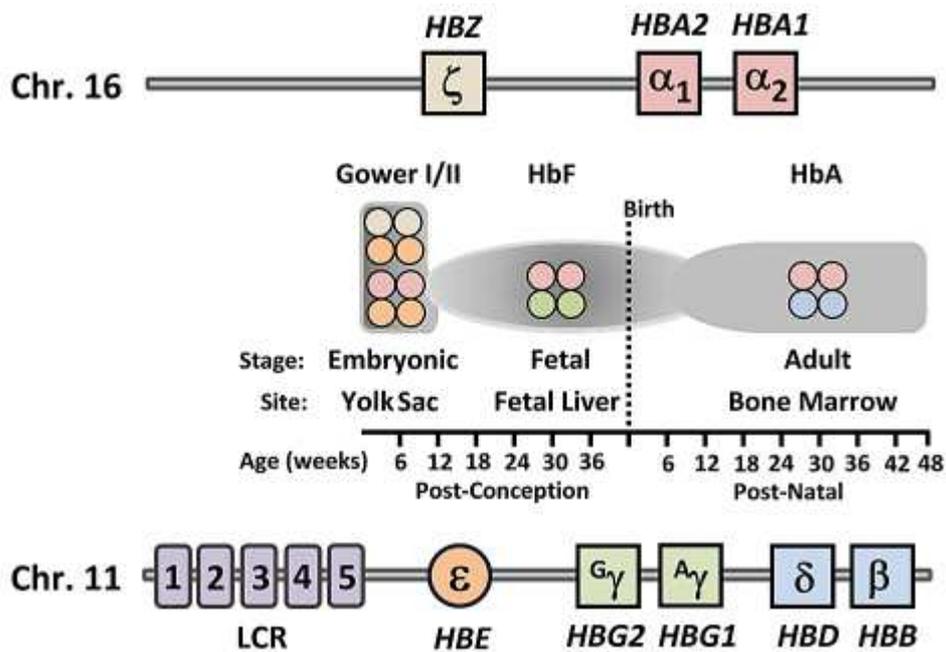
Introduzione.....	1
Scopo della tesi.....	12
Materiali e Metodi.....	13
Risultati.....	15
Discussione e Conclusioni.....	22
Bibliografia.....	25

## Introduzione

La molecola emoglobinica (Hb) è un tetramero composto da due eterodimeri formati da una catena polipeptidica del tipo  $\alpha$ -globinico e da una catena polipeptidica appartenente al gruppo  $\beta$ -globinico. Durante l'ontogenesi la composizione di questi polipeptidi varia con conseguente produzione di catene emoglobiniche aventi proprietà specifiche per ogni epoca della vita. Nel periodo embrionale nel sacco vitellino è prodotta una Hb denominata Gower I che è composta da due catene globiniche  $\zeta$  e da due catene  $\epsilon$  ( $\zeta_2\epsilon_2$ ) [1]. Un primo processo di transizione globinica si verifica nel periodo tra la 5<sup>a</sup> e la 10<sup>a</sup> settimana, quando l'ematopoiesi si sposta dal sacco vitellino al fegato fetale, in questo periodo le catene  $\zeta$  vengono sostituite da catene  $\alpha$  e le catene  $\epsilon$  da catene  $\gamma$  ( $\alpha_2\gamma_2$  o HbF). Il periodo fetale è caratterizzato dall'HbF ( $\alpha_2^G\gamma_2$  e  $\alpha_2^A\gamma_2$ ) che costituisce il 90% dell'emoglobina in questo stadio, la cui produzione continua anche dopo la nascita andando a costituire il 5% dell'emoglobina totale per i primi mesi di vita dell'individuo.

Nel periodo perinatale ha inizio il secondo processo di switching emoglobinico, che viene completato nel corso del 2° anno di vita, durante il quale la produzione di catene  $\beta$ , e in minor misura di catene  $\delta$ , si sostituisce alle catene  $\gamma$  ( $\alpha_2\beta_2$  o HbA e  $\alpha_2\delta_2$  o HbA2). La sintesi di globine  $\gamma$  diminuisce gradualmente dopo la nascita fino a essere quasi completamente soppiantata dalla produzione di globine  $\beta$  alla fine del primo anno di età.

In un individuo adulto la forma prevalente di emoglobina è la HbA ( $\alpha_2\beta_2$ ), mentre la forma detta HbA2, costituita da due catene  $\alpha$  e da due catene  $\delta$ , rappresenta il 2-3% dell'Hb totale contro una frazione variabile, ma spesso inferiore all'1%, rappresentata dall'HbF (emoglobina fetale). I geni globinici  $\epsilon$   $\gamma$   $\beta$ , disposti dal 5' al 3' nell'ordine in cui sono attivati durante l'ontogenesi, sono localizzati nel cromosoma 11 (chr 11p15.5), i geni  $\alpha$  sono localizzati nel cromosoma 16 (chr 16p13.3). L'attività dei geni globinici è controllata da sequenze localizzate a monte, denominate LCR nel locus  $\beta$  e MRE nel locus  $\alpha$  rispettivamente (Fig. 1). Le catene  $G\gamma$  e  $A\gamma$ , prodotte da due geni distinti localizzati sul cromosoma 11 nel cluster  $\beta$ -globinico, sono strutturalmente omogenee tra loro, con l'unica diversità in posizione 136, dove può essere presente glicina (catene  $G\gamma$ ) o alanina (catene  $A\gamma$ ).



**Fig. 1.** Schema dell'organizzazione dei loci  $\alpha$  e  $\beta$ -globinici ed espressione temporale dei vari tipi di emoglobina (Wilber A. et al. Blood 2011) [2].

L'HbF è distribuita eterogeneamente nei globuli rossi in quanto la sua sintesi è limitata ad un numero ristretto di cellule, chiamate F-cellule, che rappresentano approssimativamente il 3-7% degli eritrociti e contengono circa il 20-25% di HbF [3].

Da un punto di vista funzionale, le emazie con elevato contenuto di HbF presentano un'affinità per l'ossigeno più elevata rispetto a quelle che hanno esclusivamente HbA a causa della scarsa interazione dell'HbF con il 2,3-difosfoglicerato (DPG): in gravidanza, questa caratteristica facilita infatti il trasporto dell'ossigeno dagli eritrociti materni a quelli fetali [4].

Livelli elevati di HbF nella vita adulta possono essere dovuti ad effetti primari o secondari di patologie acquisite o ereditarie. In presenza di disordini acquisiti, il rapporto  $\gamma_G : \gamma_A$  è sempre di "tipo adulto" (2:3) e l'incremento di HbF è indirettamente determinato dalla iperstimolazione del midollo osseo. Al contrario, valori elevati di HbF con un rapporto di "tipo fetale" (3:1) oppure con una forte prevalenza del tetramero  $\alpha_2\gamma_2$  (60-95%), indicano la presenza di determinanti genetici quali delezioni, polimorfismi e mutazioni puntiformi a carico dei geni  $\gamma$  e/o di altre porzioni del cluster  $\beta$ -globinico coinvolte nella regolazione dell'espressione dei geni fetali.

*Disordini ereditari*

Sindromi talassemiche

- $\beta$ -talassemia omo- od eterozigote
- $\delta\beta$ -talassemia omo- od eterozigote
- persistenza ereditaria di emoglobina fetale (HPFH) omo- ed eterozigote

Altre emoglobinopatie

- anemia falciforme
- HbC, HbE, Hb Lepore, alcune emoglobine instabili (effetto variabile)
- alcune varianti emoglobiniche (Hb Rainier, Hb Bethesda) alcali-resistenti<sup>a</sup>
- varianti emoglobiniche con pl uguale a quello dell'HbF<sup>b</sup>
- varianti emoglobiniche con tempo di ritenzione simile a quello della HbF<sup>b</sup>

Sferocitosi ereditaria

*Disordini acquisiti*

Malattie del sangue (non neoplastiche)

- anemia aplastica
- anemia normoblastica (refrattaria e non)
- anemia perniziosa
- anemia sideroblastica
- emoglobinuria parossistica notturna
- eritroblastopenia pura

Neoplasie

- epatocarcinoma
- eritroleucemia
- leucemia acuta
- leucemia mieloide cronica giovanile
- neoplasie con metastasi midollari

Condizioni associate a specifici trattamenti terapeutici

- chemioterapia per leucemie
- terapia con idrossiurea, 5-aza-2'-deossicitidina, butirrati, eritropoietina

Altre condizioni

- gammopatia monoclonale benigna
- gravidanza
- insufficienza renale cronica
- ipertiroidismo
- recupero dopo trapianto di midollo osseo o ipoplasia midollare
- trisomia 13 (trisomia D)

**Tab. 1.** Principali situazioni che possono determinare un aumento delle concentrazioni ematiche di HbF (Paleari R biochimica clinica 2009) [5].

Nella maggior parte delle condizioni acquisite (gravidanza, leucemie, trapianto di midollo, imponenti perdite ematiche, alcune malattie del sangue non neoplastiche), l'aumentata produzione di HbF sembra essere un effetto secondario in genere dovuto alla accelerazione del normale processo eritropoietico (eritropoiesi da stress) [6]. Nelle condizioni ereditarie l'aumento di HbF può essere il diretto risultato di alterazioni genetiche (Hereditary Persistence of Fetal Hemoglobin o HPFH) o la conseguenza di stati di stress eritropoietico o emolitico quali si verificano nelle sindromi  $\beta$ -talassemiche e nell'anemia falciforme.

Le HPFH e le  $\delta\beta$ -talassemie sono un gruppo eterogeneo di disordini caratterizzati da aumentati livelli di HbF nella vita adulta. La distinzione tra i due gruppi viene effettuata sulla base del quadro clinico ed ematologico. Le HPFH sono condizioni solitamente benigne, caratterizzate da singole sostituzioni nucleotidiche nel promoter del gene  $\gamma$ -globinico o da delezioni nel cluster  $\beta$ -globinico che rimuovono i geni  $\delta$  e  $\beta$ -globinico [7]. In generale, i soggetti eterozigoti HPFH, hanno livelli di HbF compresi tra il 15 e il 25% con indici eritrocitari normali: le HPFH da delezione (dHPFH) hanno livelli di HbF, nella maggior parte dei casi, molto più alti rispetto alle HPFH da non delezione (ndHPFH).

Gli eterozigoti per  $\delta\beta$ -talassemia hanno invece livelli di HbF più bassi (5-15%), indici globulari microcitici e ipocromici ed un lieve sbilanciamento della sintesi globinica. Allo stato omozigote i pazienti affetti da  $\delta\beta$ -talassemia e dHPFH presentano solo HbF poiché sia il gene  $\delta$  che quello  $\beta$  risultano deleti in tutto o in parte.

Da un punto di vista molecolare, nelle dHPFH la delezione coinvolge l'estremità 3' del cluster beta e può estendersi fino a rimuovere i geni  $\delta$  e  $\beta$ , nelle  $\delta\beta$ -talassemie la delezione può rimuovere, oltre ai geni  $\delta$  e  $\beta$ , anche il gene  $A\gamma$ . Nelle ndHPFH, le mutazioni puntiformi sono localizzate nel promotore, a monte del Cap, intorno alle posizioni -114, -175 e -200. Alterazioni in queste posizioni modificano elementi regolatori che agiscono in *cis*, con conseguente alterazione del legame di vari fattori trascrizionali eritroide-specifici o ubiquitari che agiscono in *trans* nei promotori dei geni gamma globinici [8].

La ndHPFH  $G\gamma$  -158 C→T (o XmnI- $G\gamma$ ; rs7482144), data la sua frequenza ( $f = 0,32 - 0,35$ ) in tutti i gruppi etnici, è considerata un polimorfismo [9].

Diversi studi hanno confermato l'associazione tra l'allele "T" (XmnI) e livelli elevati di HbF, così come l'attenuazione del fenotipo negli individui con anemia falciforme e  $\beta$ -talassemia in diverse popolazioni; ciò giustifica, in parte, l'esistenza di fenotipi clinici differenti associati alla stessa mutazione  $\beta$ -talassemica (la sua frequenza risulta molto bassa nei pazienti con talassemia major rispetto a quelli con talassemia intermedia) [10].

Contrariamente alle altre mutazioni ndHPFH, l'XmnI- $G\gamma$  non contiene sequenze di legame riconoscibili dai fattori di trascrizione noti. Inoltre, ha effetti minimi nei soggetti normali e negli eterozigoti  $\beta$ -tal, mentre può essere responsabile di un significativo incremento di HbF (>60%  $G\gamma$ ) nelle condizioni di stress emolitico cronico tipiche degli omozigoti per  $\beta$ -talassemia [11]. La presenza dell'allele "T" non è, però sempre indicativa della presenza di un fenotipo HPFH, così come livelli elevati di HbF associati a aplotipi beta non sempre presentano questo allele [12,13].

L'HbF è un tratto quantitativo caratterizzato da notevole eterogeneità genetica di cui XmnI è stato il primo determinante identificato; l'elevata produzione di HbF, può dunque essere spiegata dalla presenza anche di altri fattori genetici.

Studi di associazione genomica (GWAS) e studi familiari hanno portato all'identificazione di altre regioni al di fuori del cluster  $\beta$ -globinico, quali i loci 2q16, 6q23, 8q e Xp22.2, coinvolte nella regolazione dei livelli di HbF (Fig. 2). Circa il 45% delle variazioni nei livelli di HbF è associato alla presenza di tre principali QTL: il polimorfismo XmnI nel cromosoma 11 (11p15),

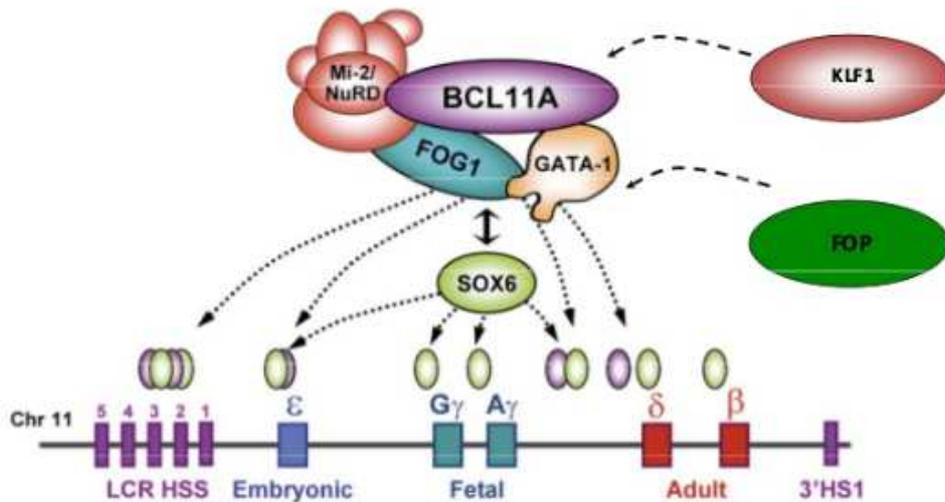
il locus HMIP nel cromosoma 6 e i polimorfismi (SNP) del gene BCL11A nel cromosoma 2 (2q16) [9,14].



**Fig. 2.** Loci identificati in studi di GWAS associati a variazioni di emoglobina fetale (Cao A. et al. Pediatric reports 2011) [15].

Il QTL che ha evidenziato la correlazione più forte con l'espressione di HbF, è rappresentato da varianti localizzate in una regione di circa 14 kb nel secondo introne del gene BCL11A [16-18].

Studi recenti hanno dimostrato, in precursori eritroidi umani e in topi knockout, che l'oncogene BCL11A è un regolatore (negativo) dell'espressione delle catene  $\gamma$  [19]. BCL11A (B cell lymphoma-leukemia 11A), denominato anche EV19 (Ecotropic Viral Integration site 19) è un fattore di trascrizione tipo zinc-finger necessario per la linfopoiesi B e T (Fig. 3) [20].



**Fig. 3.** Modello d'interazione del BCL11A per il silenziamento dei geni  $\gamma$ -globinici (Jian Xu, et al. Genes Dev. 2010) [21].

BCL11A si lega al sito HS3-LCR e alla regione intergenica  $A\gamma$ - $\delta$  [19,22] la quale si sovrappone ad una delezione umana precedente associata ad alti livelli di HbF [23]. Il BCL11A mostra una elevata espressione in precursori eritroidi umani, nei quali il suo silenziamento determina una marcata induzione di HbF. Analisi di proteomica hanno evidenziato che BCL11A interagisce fisicamente con diversi fattori tra i quali NURD (Nucleosome Remodelling Deacetylation complex) contenente due deacetilasi degli istoni (HDAC1 e 2), un componente della matrice nucleare denominata Matrix 3, i fattori di trascrizione eritroidi GATA1 e FOG1 e la proteina SOX3 associata alla cromatina [19,22,24]. Diversi studi sottolineano il coinvolgimento di BCL11A nel silenziamento del gene  $\gamma$ -globinico e il suo conseguente ruolo attivo nel processo di transizione HbF $\rightarrow$ HbA [23]. Studi di associazione genomica (GWAS) condotti da Uda M. e coll., su 4000 individui sardi, hanno portato all'identificazione del polimorfismo rs11886868, nel secondo introne del gene BCL11A, fortemente associato ( $p < 6.7 \times 10^{-35}$ ) a elevati livelli di HbF [17]. Analoghi risultati sono stati osservati anche in altre etnie con livelli elevati di F cellule [16], in individui con  $\beta$ -talassemia e con anemia falciforme [25,18,26]. La genotipizzazione del polimorfismo rs11886868 del gene BCL11A nella popolazione sarda oltre ad evidenziare una piú alta frequenza dell'allele C nei soggetti con HPFH ( $0.8\% < \text{HbF} < 5\%$ ) rispetto ai soggetti normali, ha permesso di indicare la variante C dell'rs11886868 come responsabile dell'attenuazione del fenotipo talassemico nella talassemia intermedia (non trasfusione-dipendente) rispetto a quelli affetti da talassemia major (trasfusione-dipendente) [17,27]. Lettre G. e coll., hanno riportato risultati analoghi in pazienti afro-americani e brasiliani con anemia falciforme, in cui si è osservata una forte associazione tra variazioni

dell'emoglobina fetale e la presenza del QTL BCL11A [26]. Questi studi suggeriscono che il prodotto del gene BCL11A, un fattore di trascrizione zinc-finger, è un mediatore chiave nel processo di silenziamento dell'espressione dei geni  $\gamma$ -globinici e nello switching verso il gene  $\beta$ -globinico [19]. Bauer DE. e coll., nel tentativo di capire come varianti comuni del BCL11A influenzano i livelli di HbF e la gravità del fenotipo  $\beta$ -talassemico, confrontano, in 1178 individui con anemia falciforme, 38 polimorfismi del BCL11A, associati a variazioni di HbF, a tre siti ipersensitivi alla DNase I [28]. La variante con maggiore associazione, rs1427407, distrugge il motivo E-box/GATA, un sito di riconoscimento di fattori di trascrizione altamente legato al complesso GATA1/TAL1 nelle cellule eritroidi. Questa variante è associata con una modesta riduzione dell'espressione del BCL11A nei precursori eritroidi umani [29]. Studi sulle cellule K562 hanno mostrato il legame di BCL11A al promotore del gene  $\gamma$ -globinico, sito da cui eserciterebbe l'attività di repressore sul gene adiacente [30]. Xu e altri descrivono come il silenziamento dei geni  $\gamma$ -globinici sia il risultato di un'interazione tra il fattore di trascrizione BCL11A e il cluster  $\beta$ -globinico, in particolare formando un *loop* della cromatina, in associazione al fattore di trascrizione SOX6, sulla regione promotoriale del gene  $\gamma$ -globinico. Il fattore di trascrizione BCL11A potrebbe essere, quindi, un regolatore dello sviluppo e del controllo del silenziamento dei geni  $\gamma$ -globinici [21].

Il locus HMIP, nella regione intergenica HBS1L-MYB, rappresenta un altro QTL in grado di influenzare i livelli di HbF. Il gene HBS1L è coinvolto nella regolazione di diversi processi cellulari, ed è espresso nelle cellule ematopoietiche, mentre il gene MYB è coinvolto nell'oncogenesi e gioca un ruolo chiave nei processi eritropoietici [3,31]. Craig JE e coll., ipotizza la presenza di un gene localizzato nella regione 6q, codificante un fattore coinvolto nel pathway di maturazione eritroide, implicato nella produzione di F cellule e nell'aumento di livelli di HbF [32]. L'associazione del locus 6q23 sull'aumentata produzione di HbF viene, successivamente, confermata da Garner in pazienti Indiani con  $\beta$ -talassemia, anche se non è chiaro se l'effetto sia diretto, agendo come fattore che interagisce con i geni  $\gamma$ -globinici, o indiretto [33,32]. Nella popolazione Caucasica questa regione è costituita da tre blocchi in LD (Linkage Disequilibrium), e il secondo blocco è quello più fortemente associato a variazioni di HbF. La variabilità del secondo blocco ha un effetto pleiotropico su diversi parametri ematologici in relazione al numero di eritrociti, suggerendo la presenza di sequenze regolatorie distali nella regione intergenica HBS1L-MYB, che agisce sui precursori della linea eritroide [32].

Tale associazione è stata dimostrata anche in soggetti sani del nord europa, in cui la presenza del QTL 6q23 contribuisce alla variazione dei livelli di emoglobina fetale per circa il 19% [3,31]. Questo contributo è risultato inferiore in altre popolazioni quali soggetti sani afro-americi (3%) e in pazienti brasiliani e britannici (7%) con anemia falciforme [26].

Thein SL e coll., in uno studio di GWAS condotto su gemelli nord europei, identifica nel polimorfismo rs9399137 ( $p=10^{-75}$ ), all'interno del secondo blocco della regione intergenica HBS1L-MYB, la maggiore associazione ai livelli di F cellule [34].

Si è visto che, in cellule eritroleucemiche umane, la sovraespressione di MYB inibisce l'espressione dei geni  $\gamma$ -globinici, mentre la sovraespressione di HBS1L non ha alcun effetto. Gli studiosi ipotizzano che livelli bassi di espressione di MYB determinino una riduzione del numero di cicli cellulari nelle prime fasi dell'eritropoiesi e una maturazione precoce degli eritroblasti che produce globuli rossi con livelli elevati di HbF [35]. MYB ha un ruolo importante nell'eritropoiesi e recenti studi di risequenziamento e di genotipizzazione supportano il coinvolgimento di MYB sulla modulazione dei livelli di HbF [25]. Farrell JJ et coll., hanno riportato una delezione di 3 nucleotidi nel locus HMIP in completo linkage disequilibrium con l'rs9399137, il polimorfismo più significativo associato ad HbF tra le popolazioni cinese, europea ed africana. Studi di immunoprecipitazione della cromatina dimostrano che questa delezione è localizzata in una regione di legame importante per fattori di trascrizione implicati nell'eritropoiesi [36].

Il KLF1 (Krupper-likefactor 1) è un fattore di trascrizione eritroide-specifico, così denominato per l'omologia di sequenze col fattore "Krupper" prodotto da un gene omeotico della *Drosophila melanogaster* [37]. Il KLF1 appartiene ad una famiglia estesa di fattori di trascrizione, caratterizzati dalla presenza di domini del tipo zinc-finger (dita di zinco: una sequenza aminoacidica contenente 4 aminoacidi conservati tra cui cisteina e/o istidina che lega un atomo di zinco), il KLF1 agisce modificando la cromatina ed attivando la trascrizione del gene  $\beta$ -globinico [37].

KLF1 si lega ad una sequenza specifica (CACCC) del promoter del gene  $\beta$ -globinico. Knockout nel topo del gene codificante, l'omologo murino *klf1*, produce un fenotipo letale risultante essenzialmente da una forma grave di  $\beta$ -talassemia [38,39].

Nell'uomo mutazioni della sequenza CACCC, che impediscono il suo legame con KLF1 determinano il fenotipo di una lieve  $\beta$ -talassemia [40,41].

Ricerche successive hanno dimostrato in realtà che il KLF1 ha un ruolo più esteso nell'eritropoiesi, controllando oltre il gene  $\beta$ -globinico, anche l'espressione di numerosi altri geni [42]. Nell'uomo il fenotipo "In(Lu)" nel quale l'espressione del gruppo sanguigno

Lutheran nel globulo rosso è marcatamente ridotta è risultato essere legato a mutazioni del gene KLF1 [43].

Una grave anemia nel topo denominata NAN è risultata essere dovuta a mutazioni localizzate nel secondo zing-finger di KLF1 che danno luogo alla produzione di una proteina con modificazioni della sua specificità. In particolare va sottolineato che nel topo NAN l'espressione della globina embrionaria  $\beta 1$  persiste nell'adulto indicando un'influenza di KLF1 nel processo di transizione globinica [44].

Infine studi recenti nel topo knockout per *klf1* [45] ed in famiglie con persistenza ereditaria di Hb fetale (HPFH) hanno dimostrato che KLF1 gioca anche un ruolo nella regolazione dell'espressione dell'HbF.

Studi sulla regolazione del gene codificante KLF1 in precursori eritroidi umani e murini ottenuti con l'uso di shRNA (short interfering RNA, molecola di RNA a doppia elica contenenti 21-22 nucleotidi che bloccano la traduzione del RNA messaggero bersaglio), hanno dimostrato accanto ad un effetto sul gene  $\beta$ -globinico, un aumento dell'espressione delle catene  $\gamma$ -globiniche associate ad una marcata riduzione dell'RNA messaggero e della proteina BCL11A che, come abbiamo detto sembra essere un regolatore negativo dell'espressione delle catene  $\gamma$ . Inoltre con l'uso della tecnica ChIP quantitativa (metodica secondo la quale un anticorpo specifico per un fattore di trascrizione consente di isolare accanto al fattore specifico anche il DNA ad esso legato le cui sequenze possono essere analizzate), è stato osservato che il KLF1 lega il CACCC del promoter del gene BCL11A, indicando che il KLF1 regola direttamente l'espressione di BCL11A [45]. Studi di una estesa famiglia di origine Maltese con HPFH, tramite analisi di associazione e linkage, hanno evidenziato che il tratto HPFH mappa nella regione cromosomica 19p13.12-13 ove si trova il locus codificante il KLF1 [46]. L'analisi delle sequenze di KLF1 ha mostrato una mutazione nonsense (K288X) in tutti i membri della famiglia con HPFH. Studi di espressione in precursori eritroidi dei soggetti con HPFH hanno dimostrato una ridotta espressione dei geni regolati da KLF1 incluso BCL11A e un'aumentata espressione dei geni  $\gamma$ -globinici. Inoltre è stato confermato che il KLF1 si lega al promotore di BCL11A. Lo studio di questa famiglia indica che il fenotipo HPFH dipenda da uno stato di aploinsufficienza del KLF1 [46].

Studi in un'altra famiglia di origine sarda, hanno tuttavia osservato che per avere il fenotipo HPFH occorre che entrambi gli alleli KLF1 siano mutati. In questa famiglia il fenotipo HPFH era associato a uno stato di composto genetico per una mutazione missenso (K332Q) e una mutazione nonsense (S270\*) [47].

Le famiglie Indiane studiate da Craig e coll., sono state nuovamente prese in esame per un altro studio di associazione da Garner. In questo GWAS basato su un modello a due locus genetici uno dei quali conteneva la regione codificante l'XmnI, è stato identificato un altro QTL sul cromosoma 8 (Chr 8q), legato ad alti livelli di HbF ed F cellule. Questo studio suggerisce che il QTL sul Chr 8q potrebbe contenere la regione codificante un fattore regolatore che agirebbe come attivatore o repressore trascrizionale legandosi direttamente al sito dell'XmnI o ad una regione vicino ad esso, regolando in tal modo le variazioni di HbF [48].

Dover e coll., analizzando soggetti sani e pazienti affetti da anemia falciforme di origine africana, hanno visto che la percentuale di F cellule presenti nelle donne era significativamente più elevata rispetto agli uomini, suggerendo il possibile coinvolgimento del cromosoma X. Questi autori hanno trovato l'associazione tra il locus Xp22.2 e la produzione di F cellule e hanno suggerito che circa il 70% della variazione nei livelli di HbF in pazienti affetti da anemia falciforme è associata a variazione nella percentuale di F cellule [49].

Recentemente Danjou F. e coll., [50], hanno studiato l'associazione tra le tre forme di emoglobina HbA1, HbA2 e HbF e SNPs (circa 10,9 milioni) imputati o genotipizzati su 6000 volontari sardi (studi di GWAS integrati con sequenziamenti ad alta definizione e Whole Genome Sequencing).

Da questo studio sono emerse 23 varianti significative su 10 loci, alcune già note in letteratura ed altre completamente nuove. Considerando l'emoglobina fetale, identificano un nuovo polimorfismo l'rs183437571 ( $p=1.61 \times 10^{-8}$ ) localizzato nel primo introne del gene NFIX (Chr19), che si trova raramente nella popolazione sarda (MAF=0.01); il polimorfismo rs67385638 ( $p=1.09 \times 10^{-25}$ ) situato nel secondo introne del gene HBE1 (Chr 11), che più fortemente associato rispetto all'XmnI, alla produzione di HbF; il polimorfismo rs2855122 ( $p=2.57 \times 10^{-11}$ ) a monte del gene HBG2 (Chr 11), coinvolto nello switching dell'emoglobina. Inoltre, quest'ultimo polimorfismo determina, oltre ad una riduzione dell'HbF, un aumento dell'HbA1 e dell'HbA2. In letteratura il polimorfismo rs2855122 con l'rs2855121 e l'rs10128653 è usato per costruire un aplotipo definito "pre Gamma" [12].

Nello stesso lavoro, Danjou F. e coll., valutano l'aplotipo pre G $\gamma$  "migliore" in grado di spiegare l'associazione con le variazioni di HbF in 1164 individui sardi non imparentati tra loro. Un aplotipo è stato costruito secondo il modello descritto da Pissard e coll.,( costituito dagli rs2855122, rs2855121, rs10128653) ad esclusione dell'rs10128653 che in Sardegna è monomorfo; mentre l'altro modello comprende il polimorfismo rs2855122 e il nuovo polimorfismo rs67385638.

Dal confronto dei due modelli risulta che l'aplotipo rs2855122 e rs67385638 è più fortemente associato a variazioni di HbF ( $1.1 \times 10^{-4}$  versus  $7.1 \times 10^{-4}$ ) rispetto all'aplotipo descritto da Pissard e coll., e modificato da Danjou e coll.

## Scopo della tesi

Lo studio dei determinanti genetici responsabili dell'espressione di HbF è molto utile nello studio e nella diagnosi di alcune importanti disordini a carico dei geni globinici dove le concentrazioni di HbF possono aumentare considerevolmente, in particolare  $\beta$ - e  $\delta\beta$ -talassemia, elevata persistenza di HbF (HPFH) ed in alcune patologie acquisite, che sono associate ad un modesto incremento di HbF.

Recenti ricerche di associazione genomica hanno definito alcuni loci cromosomici responsabili della regolazione della produzione di HbF. Uno di questi loci mappa nella regione regolatrice LCR-HS3 del cluster beta globinico. Un secondo locus sul cromosoma 6, nella regione intergenica tra i geni HBS1L e il gene MYB, ha funzione regolatrice nell'espressione del gene MYB i cui livelli sono ridotti in soggetti con HPFH. Il terzo e più importante locus riguarda il gene BCL11A, un regolatore negativo della produzione di HbF. L'allele minore a questo locus è associato ad HPFH e più frequentemente nei soggetti con forme attenuate di beta talassemia, in cui l'indipendenza dalle trasfusioni di sangue dipende da un aumento della produzione di HbF.

Studi in famiglie con HPFH causate da mutazioni nel gene KLF1 e analisi nei topi con inattivazione del gene *klf1* hanno dimostrato che KLF1 oltre ad attivare direttamente la produzione di  $\beta$ - e  $\gamma$ -globina, attiva anche il BCL11A inibendo indirettamente la produzione di  $\gamma$ -catene. KLF1 e BCL11A sono pertanto i fattori principali che regolano lo switching emoglobinico.

La recente individuazione di nuove varianti localizzate nel cluster beta globinico, quali quelle descritte da Danjou et al., [50] coinvolte nella regolazione dei livelli delle diverse emoglobine (HbA, HbA2, HbF), ha posto le basi per lo scopo di questa tesi che si propone di studiare in particolare i polimorfismi rs67385638 e rs2855122 in soggetti con HPFH, per spiegare i livelli aumentati di emoglobina fetale nella vita adulta, e in pazienti con beta talassemia major e beta talassemia intermedia con l'intento di spiegare il diverso fenotipo (trasfusione dipendente e non trasfusione dipendente).

## **Materiali e Metodi**

### **Casistica**

Sono stati presi in considerazione 87 soggetti sardi adulti di età compresa tra 10 e 73 anni ( $32 \pm 13$ ), con valori di HbF  $\geq 1.5$  % (campione HPFH); 58 soggetti sono di sesso femminile e 29 maschile e 41 soggetti sani adulti di controllo, 37 femmine e 4 maschi con valori di HbF  $< 1$ % (campione Normali); 454 pazienti con beta talassemia: 395 Major (campione Talassemie Major) e 59 Intermedie (campione Talassemie Intermedie). Sui soggetti in studio è stata prelevata un'aliquota di circa 5 ml di sangue periferico in provette con EDTA per l'esame emocromocitometrico, l'analisi del pattern emoglobinico e l'estrazione del DNA per la genotipizzazione dei polimorfismi rs67385638 localizzato nel gene HBE1, rs2855122 e rs7482144 localizzati nel gene HBG2, rs1427407 nel gene BCL11A, rs9399137 nel gene HBS1L-MYB e la ricerca delle mutazioni del gene KLF1 più frequenti (o già identificate) nella popolazione sarda (p.S270\*, p.R319Efs\*34, p.K332Q, p.T327S).

I campioni di sangue sono stati prelevati presso l'ambulatorio di ematologia pediatrica e presso il servizio di screening per le talassemie, dell'Ospedale Pediatrico Microcitemico "A. Cao" di Cagliari. Tutti i soggetti hanno firmato un modulo di consenso informato.

### **Analisi ematologica**

La misurazione degli indici eritrocitari MCV (fl), MCH (pg), Hb (g/dl) è stata effettuata con il contaglobuli elettronico (LH750 Hematology Analyzer, Beckman Coulter Inc. Miami, FL). La determinazione quantitativa dei livelli di emoglobina e l'analisi del pattern emoglobinico (HbF% e HbA<sub>2</sub>%) è stata eseguita mediante cromatografia liquida ad alta risoluzione (HPLC: High Performance Liquid Chromatography – Variant II, Dual kit analyzer, Bio-Rad, Segrate, Mi, Italia).

### **Analisi molecolare**

**Estrazione DNA.** Il DNA genomico è stato estratto sia con estrazione salina (salting-out) sia mediante il protocollo del kit di estrazione di acidi nucleici (Blood DNA 500, DiaSorin). La

qualità e la concentrazione del DNA sono state valutate mediante lettura spettrofotometrica (Nanodrop).

**Genotipizzazione XmnI.** L'identificazione del polimorfismo rs7482144 (XmnI) nel gene HBG2, è stata fatta tramite una reazione di PCR classica, con i primer EA59-TAAAAAAATACTTTTGTGA, R10-GGACTAGGAGCTTATTGAT, seguita da digestione (37°C O/N) con l'enzima di restrizione PDMI (XmnI) della Thermo Fisher Scientific.

**Genotipizzazione polimorfismi rs67385638, rs2855122, rs1427407 e rs9399137.** I polimorfismi rs67385638 (HBE1), rs2855122 (HBG2), rs1427407 (BCL11A), rs9399137 (HBS1L-MYB) sono stati genotipizzati utilizzando i saggi custom per discriminazione allelica (TaqMan 7500 Real Time PCR System, Applied Biosystems, Warrington, UK). I dati sono stati analizzati con il software Sequence Detection Software (SDS v.1.2.3) Applied Biosystems.

**Studio mutazioni gene KLF1.** L'analisi delle mutazioni del gene KLF1 più frequenti in Sardegna e localizzate nella porzione terminale del secondo esone e in quella iniziale del terzo esone [Chr19: da nt 12884837 a nt 12885641 (GR Ch38 p.2)] è stata effettuata mediante amplificazione (PCR) del DNA genomico e successivo sequenziamento automatico sulla piattaforma ABIPRISM 3130XL Genetic Analyzer, Applied Biosystems, Warrington, UK ABI PRISM, utilizzando i seguenti primers:

KLF1-Ex2-CF (TACGGGCTACTGTCCGGGTA)

KLF1-Ex2-BR (GCCCTCTGCAACCCTTCTTC)

KLF1-Ex3-F (TGCGGCAAGAGCTACACCA)

KLF1-Ex3-R (CTTGTCCCATCCCCAGTCACT)

### **Analisi Statistica**

I test di associazione degli SNPs tra casi e controlli sono stati elaborati con il software PLINK (v1.7) [51]. Lo studio della dipendenza tra i valori di HbF e HbA2 con ogni SNPs nella popolazione HPFH è stato effettuato con il software SPSS (IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp.). Per ogni SNPs è stata eseguita una analisi di regressione lineare considerando HbF e HbA2 come variabile dipendente e ogni SNPs come predittore.

## Risultati

### Polimorfismo rs67385638 HBE1

#### Distribuzione e correlazione dei polimorfismi con le variazioni dell'HbF in gruppi di individui con diverse condizioni cliniche

Il polimorfismo rs67385638 è stato genotipizzato in 22 soggetti sardi con valori di HbF nella norma (HbF <1%). Tredici individui sono risultati omozigoti per l'allele wild type (C), 6 eterozigoti e 3 omozigoti per l'allele minore (G). La MAF (Minor Allele Frequency) dell'rs67385638 è quindi 0,27 cioè simile a quella attesa dall'analisi 1000 Genomes che è pari a 0.30. La MAF dello stesso polimorfismo ottenuta genotipizzando circa 10,9 milioni di SNPs in 6602 volontari sani della popolazione sarda (Studio Progenia), e usata da noi come riferimento, è risultata solo leggermente più bassa (MAF=0.24) rispetto ai dati ottenuti da noi in laboratorio (Tab.2).

Soggetti sani di controllo			
Genotipo rs67385638	CC	CG	GG
%	59	27	14
n	13	6	3

**Tab.2** Genotipi rs67385638 dei soggetti sani di controllo (Normali)

In tabella 3 sono mostrati i risultati della genotipizzazione dell'rs67385638 di 395 Talassemie Major (TD) e 59 Talassemie Intermedie (NTD).

Genotipo rs67385638	Talassemie Major			Talassemie Intermedie		
	CC	CG	GG	CC	CG	GG
%	94	6	0	74	24	2
n	373	22	0	44	14	1

**Tab.3** Genotipi delle Talassemie Major e delle Talassemia Intermedie

La frequenza dei genotipi CC, CG e GG nelle Talassemie Major è risultata rispettivamente 94%, 6% e 0%, mentre nei pazienti con Talassemia Intermedia le frequenze sono pari a 74%(CC), 24%(CG) e 2%(GG). La frequenza dell'allele minore nelle Talassemie Major (0.03) è quindi marcatamente inferiore rispetto ai soggetti sani di controllo con un p di  $3.54 \times 10^{-8}$ . La frequenza allelica dell'allele G nelle Talassemie Intermedie (0.14) è diminuita rispetto ai soggetti sani di controllo (0,27), con un p di 0.08. Il confronto delle frequenze tra i due gruppi di talassemie, NTD versus TD, è risultato altamente significativo con un p di  $2.86 \times 10^{-11}$ .

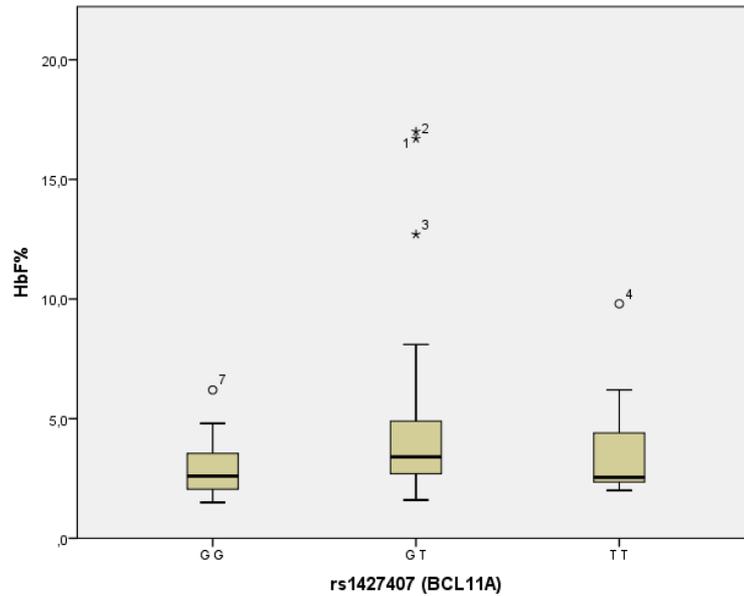
Lo studio è proseguito con la genotipizzazione del polimorfismo rs67385638 in una popolazione di 87 soggetti adulti con HPFH. In questo gruppo la frequenza allelica dell'allele minore G, rispetto alla popolazione di controllo, è aumentata (0.43), con la seguente distribuzione dei diversi genotipi: 32% CC; 50% CG e 18% GG (p=0.05).

Tali individui sono stati inoltre genotipizzati per gli altri determinanti genetici noti per essere capaci d'influenzare i livelli di HbF, vale a dire gli rs7482144 (XmnI), rs1427407 e rs9399137 rispettivamente nei loci HBG2, BCL11A e HBS1L-MYB (Tab.4).

	rs67385638 HBE1			rs7482144 HBG2			rs1427407 BCL11A			rs9399137 HBS1L-MYB		
	C/C	C/G	G/G	C/C	C/T	T/T	G/G	G/T	T/T	T/T	T/C	C/C
%	33	49	18	36	49	15	42	49	9	42	37	21
n	28	43	16	31	43	13	36	42	8	36	32	17

**Tab.4** Frequenze genotipiche dei determinati genetici di HbF in soggetti con HPFH

La distribuzione dei diversi genotipi in rapporto alle variazioni dell'HbF rivela un'associazione statisticamente significativa solo col polimorfismo rs1427407 (BCL11A) con un p=0.03. (Fig.4)



**Fig. 4** Distribuzione HbF nei diversi genotipi dell'rs1427407 del BCL11A

In sostanza, l'analisi della distribuzione dei genotipi rs67385638 nelle popolazioni major e intermedie mostrava una differenza statisticamente significativa che faceva presumere un effetto modificatore positivo del polimorfismo sulla severità clinica della talassemia, mentre la distribuzione dei genotipi al variare dell'HbF negli HPFH, in cui non si osservava nessuna associazione statisticamente significativa, sembrava sconfiggere l'ipotesi che l'effetto favorevole fosse ottenuto tramite l'incremento dell'HbF.

Abbiamo pensato che questa discordanza dipendesse dal fatto che la popolazione HPFH fosse troppo eterogenea, giacché conteneva, sia individui con che individui senza mutazioni del gene KLF1, e che un'eventuale effetto d'associazione potesse essere coperto dall'effetto dell'associata mutazione e svelato dall'analisi separata dei 2 gruppi di soggetti senza e con mutazione KLF1.

			rs67385638 HBE1		
			CC	CG	GG
HPFH	<i>KLF1 neg</i> (n°=32)	n°	9	17	6
		m ± ds	4.6 ± 4.8	2.9 ± 0.9	2.4 ± 0.7
	<i>KLF1 mut</i> (n°=55)	n°	19	26	10
		m ± ds	3.3 ± 1.8	4.2 ± 3.1	4.6 ± 3.1

**Tab.5** Distribuzione dei genotipi dell'rs67385638 e valori medi di HbF (%) nei gruppi KLF1 mutati e non mutati

L'analisi genetica separata indicata in tabella 4 mostra che 55 degli 87 soggetti HPFH hanno evidenziato mutazioni nel gene KLF1 (46 la S270\*, 7 la R319Efs\*34, 1 la K332Q e 1 la T327S) trovando quindi una plausibile giustificazione per l'incremento dell'HbF, un noto effetto dell'eterozigosi KLF1.

La distribuzione dei genotipi del polimorfismo rs67385638 e dei valori medi di HbF nei gruppi KLF1 mutati vs KLF1 non mutati, è descritta nella stessa Tab.5, dove si osserva un trend in progressiva riduzione dei genotipi permissivi per l'HbF nei KLF1 negativi e uno speculare aumento degli stessi genotipi nei soggetti KLF1 mutati.

Le variazioni dei genotipi in rapporto alla HbF nei non mutati rivela un trend del polimorfismo rs673856638 verso la quasi significatività, passando dal valore di  $p=0,7$  nella popolazione globale al valore di  $p=0,08$  e di  $p=0.1$  rispettivamente nei soggetti privi di e con mutazione KLF1. Il polimorfismo rs7482144 (XmnI) non risulta statisticamente significativo sia negli HPFH in toto ( $p=0.6$ ) che nei KLF1 mutati e non mutati, dove in entrambi i gruppi la significatività è  $p=0.2$ . Interessante osservare che anche la significatività dell'associazione col polimorfismo BCL11A subisce dei cambiamenti evidenti passando da un  $p=0.03$  nella popolazione HPFH globale al valore di  $p=0.8$  non significativo nei non mutati e a un  $p=1.8E-04$  molto più significativo nei mutati. Il polimorfismo del gene MYB a sua volta passa da un  $p=0.6$  negli HPFH globali a un  $p=0.3$  e  $p=0.01$  rispettivamente nei non mutati e mutati per KLF1. In tutti i casi si osserva o un evidente miglioramento della significatività o un trend verso la quasi significatività nelle popolazioni separate (Tab.6).



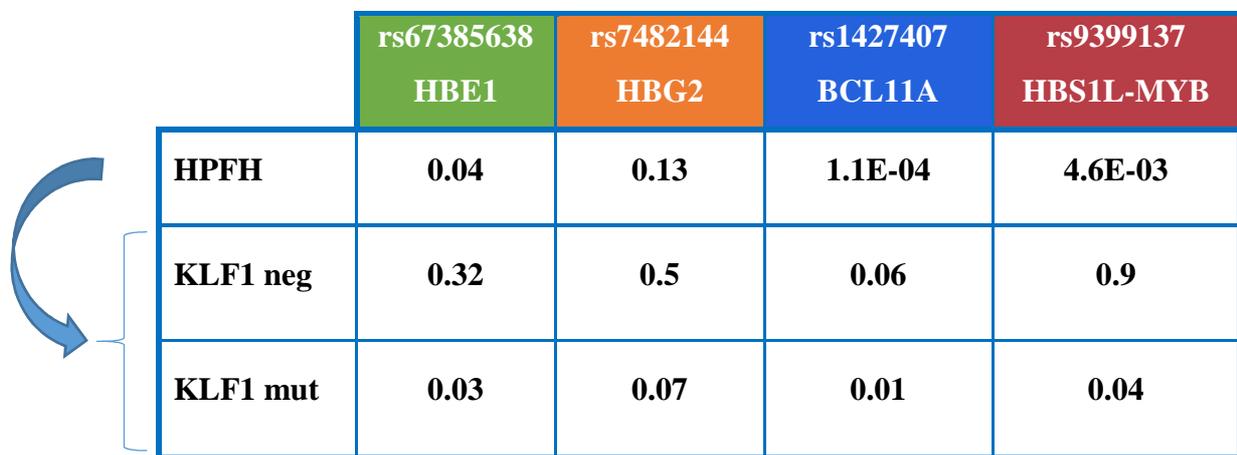
	rs67385638 HBE1	rs7482144 HBG2	rs1427407 BCL11A	rs9399137 HBS1L-MYB
<b>HPFH</b>	<b>0.7</b>	<b>0.6</b>	<b>0.03</b>	<b>0.6</b>
<b>KLF1 neg</b>	<b>0.08</b>	<b>0.2</b>	<b>0.8</b>	<b>0.3</b>
<b>KLF1 mut</b>	<b>0.12</b>	<b>0.2</b>	<b>1.84E-04</b>	<b>0.01</b>

**Tab 6.** Valori di significatività (p) dei diversi determinanti genetici in relazione all'HbF e alla presenza di mutazioni nel gene KLF1

## Correlazioni dei polimorfismi con le variazioni dell'HbA2 in soggetti HPFH

Oltre all'analisi dell'associazione dei genotipi con le variazioni dell'HbF, abbiamo anche condotto un'analisi di associazione tra gli stessi determinanti genetici e le variazioni di distribuzione dell'HbA2 trovando variazioni statisticamente significative per il polimorfismo rs67385638 che passa da  $p=0.04$  nel campione HPFH indistinto al  $p=0.32$  nei non mutati e al  $p=0.03$  nei mutati. Quindi a differenza delle variazioni dell'HbF, le variazioni dell'HbA2 in rapporto ai diversi genotipi rs67385638 sono più significative nei soggetti mutati. Il polimorfismo rs7482144 (XmnI) non ha mostrato significatività in nessuno dei campioni HPFH, KLF1 non mutati e KLF1 mutati con  $p=0.13$ ,  $p=0.5$  e  $p=0.07$  rispettivamente.

Il polimorfismo rs1427407 del gene BCL11A ha andamento opposto rispetto al polimorfismo rs67385638, in quanto passa da un'associazione chiaramente significativa nella popolazione HPFH indistinta con  $p=1.1E-04$  a perdita di significatività nei non mutati ( $p=0.06$ ) e riduzione di significatività nei KLF1 mutati ( $p=0.01$ ). Il polimorfismo rs9399137 dell'HBS1L-MYB si comporta in maniera analoga passando da una significatività iniziale di  $p=4.6E-03$  a  $p=0.9$  nei non mutati e  $p=0.04$  nei KLF1 mutati (Tab.7).



	rs67385638 HBE1	rs7482144 HBG2	rs1427407 BCL11A	rs9399137 HBS1L-MYB
<b>HPFH</b>	<b>0.04</b>	<b>0.13</b>	<b>1.1E-04</b>	<b>4.6E-03</b>
<b>KLF1 neg</b>	<b>0.32</b>	<b>0.5</b>	<b>0.06</b>	<b>0.9</b>
<b>KLF1 mut</b>	<b>0.03</b>	<b>0.07</b>	<b>0.01</b>	<b>0.04</b>

**Tab.7.** Valori di significatività (p) dei diversi determinanti genetici in relazione all'HbA2 e alla presenza di mutazioni nel gene KLF1

## Polimorfismo rs2855122 HBG2

Il polimorfismo rs2855122 è stato genotipizzato in 37 soggetti sardi normali con valori di HbF < 1%. Dodici individui sono risultati omozigoti per l'allele wild type (T), 13 eterozigoti e 12 omozigoti per l'allele minore (C) la cui MAF è perciò pari a 0.5. Le MAF dell'rs2855122, tratte da 1000 Genomes e dallo Studio Progenia, sono invece inferiori essendo in entrambi gli studi pari a 0.39 (Tab.8).

Soggetti sani di controllo			
Genotipo rs2855122	TT	TC	CC
%	33	35	33
n	12	13	12

**Tab.8** Genotipi rs2855122 dei soggetti sani di controllo (Normali)

Come per il polimorfismo rs67385638, abbiamo analizzato una possibile differenza di distribuzione del polimorfismo rs2855122 anche nei soggetti con talassemia. In tabella 9 sono mostrati i risultati della genotipizzazione in 394 pazienti con Talassemia Major e in 58 pazienti con Talassemia Intermedia.

Genotipo rs2855122	Talassemie Major			Talassemie Intermedie		
	TT	TC	CC	TT	TC	CC
%	8	34	58	9	29	62
n	30	133	231	5	17	36

**Tab.9** Genotipi delle Talassemie Major e delle Talassemie Intermedie

La frequenza genotipica è risultata sovrapponibile nelle categorie di Talassemia Major e intermedia con valori TT 8%, TC 34% e CC 58% nei soggetti con talassemia Major e di TT 9%, TC 29%, CC 62% in quelli con Talassemia Intermedia. La frequenza dell'allele C è sovrapponibile nelle due categorie studiate TD (0.76), NTD (0.77).

Le frequenze appaiono invece significative nel confronto tra soggetti talassemici e individui normali. In dettaglio, il confronto NTD-normali mostra una differenza con una significatività  $p=2.5 \times 10^{-4}$ , mentre il confronto TD -normali è ancora più significativo con  $p=8.2 \times 10^{-7}$ .

Lo studio è proseguito con la genotipizzazione del polimorfismo rs2855122 in una popolazione di 87 soggetti adulti con HPFH. In questo gruppo la frequenza allelica dell'allele minore C, rispetto alla popolazione di controllo non è risultata significativa con un valore di 0.52 e  $p=0.8$ . Nello stesso gruppo HPFH è stata valutata l'associazione del polimorfismo rs2855122 con le variazioni di HbF o con le variazioni di HbA2 ( $p=0.18$  e  $p=0.07$  rispettivamente). Il gruppo HPFH è stato poi suddiviso in KLF1 mutati e KLF1 non mutati; nei KLF1 mutati è risultata esserci un'associazione sia tra l'rs2855122 e le variazioni dei livelli di HbF ( $p=0.03$ ), che con le variazioni dei valori di HbA2 ( $p=0.03$ ). Nei KLF1 non mutati le stesse associazioni non sono risultate significative (Tab.10).

HPFH	KLF1	HbF	HbA2
	neg (n=32)	0.9	0.7
	mut (n=55)	<b>0.03*</b>	<b>0.03*</b>

**Tab.10** Valori di significatività (p) nel gruppo di soggetti con HPFH, suddiviso per KLF1 non mutati e mutati (HPFH) in relazione alle variazioni di HbF e HbA2

## Discussione e Conclusioni

Diverse distribuzioni di frequenza dei genotipi dei determinanti genetici dell'HbF, nei soggetti normali, TD, NTD e HPFH, fanno presupporre un effetto clinico favorevole verosimilmente attribuibile a variazioni dell'HbF. Nell'impossibilità di analizzare le variazioni di HbF nei talassemici per la presenza del sangue trasfuso e negli intermedi per la non confrontabilità dei livelli d'emoglobina nelle diverse epoche in cui sono stati diagnosticati, l'analisi della correlazione dei polimorfismi (rs67385638 sul gene HBE1, rs7482144 e rs2855122 del gene HBG2, rs1427407 sul gene BCL11A e l'r9399137 sull'HBS1L-MYB) con le variazioni dell'HbF, è stata eseguita solo nella popolazione di soggetti con HPFH.

Tuttavia, contrariamente alle attese, in questi soggetti non abbiamo trovato ad una prima analisi, alcuna associazione significativa. Abbiamo perciò provveduto a determinare in questa popolazione la presenza di eventuali mutazioni KLF1 che potessero di per se stesse spiegare il fenotipo HPFH. Dopo selezione dei soggetti con mutazioni del gene KLF1 abbiamo proceduto ad un'analisi distinta di associazione dei genotipi con le variazioni di HbF nei soggetti HPFH mutati (1/3) e non mutati (2/3).

Come evidenziato nella tabella 11 riassuntiva delle variazioni di significatività, l'analisi nelle popolazioni separate dei polimorfismi ha permesso di svelare per alcuni associazioni significative prima nascoste, mentre per altri ha attenuato o eliminato le precedenti associazioni.

	rs67385638 HBE1		rs7482144 HBG2		rs1427407 BCL11A		rs9399137 HBS1L-MYB		rs2855122 HBG2	
	HbF	HbA2	HbF	HbA2	HbF	HbA2	HbF	HbA2	HbF	HbA2
<b>HPFH</b>	0.7	<b>0.04*</b>	0.7	0.1	<b>0.03*</b>	<b>1.1E-04*</b>	0.6	<b>4.6E-03*</b>	0.18	0.06
<b>KLF1 neg</b>	0.08	0.32	0.2	0.5	0.8	0.06	0.3	0.9	0.96	0.7
<b>KLF1 mut</b>	0.12	<b>0.03*</b>	0.2	0.07	<b>1.84E-04*</b>	<b>0.01*</b>	<b>0.01*</b>	<b>0.04*</b>	<b>0.03*</b>	<b>0.02*</b>

**Tab.11** Valori di significatività (p) dei diversi determinanti genetici in relazione a HbF, HbA2 e presenza di mutazioni nel gene KLF1

A seguito dell'analisi separata, il polimorfismo rs67385638 ha evidenziato solo un trend verso la significatività di associazione con l'HbF passando da p=0.7 a p=0.08 nei non mutati e p=0.12 nei KLF1 mutati, mentre ha confermato la sua associazione significativa con variazioni

dell'HbA2 passando da  $p=0.04$  a  $p=0.03$  nei mutati, ma con una contemporanea perdita di significatività a  $p=0.32$  nei non mutati.

Il polimorfismo rs2855122 diventa significativamente associato solo nei soggetti con KLF1 mutato sia per l'HbF ( $p=0.03$ ) che per l'HbA2 ( $p=0.02$ ).

Il polimorfismo BCL11A che era già significativamente associato alla HbF nella popolazione in toto ( $p=0.03$ ) perde la significatività nei KLF1 negativi e acquista notevole significatività nei KLF1 mutati ( $p=1.8E-04$ ). Contemporaneamente si assiste a una perdita e attenuazione di significatività per l'HbA2 che passa da un  $p=1.1E-04$  a  $p=0.06$  nei KLF1 negativi e  $p=0.01$  nei KLF1 positivi.

Il polimorfismo MYB non significativamente associato negli HPFH in toto ( $p=0.6$ ), rimane non significativo nei KLF1 non mutati ( $p=0.28$ ) e diventa significativo nei KLF1 mutati ( $p=0.01$ ). Riguardo all'associazione MYB-HbA2 si passa da una forte associazione iniziale con  $p=4.6E-03$  nelle HPFH in toto, a una perdita totale di significatività nei KLF1 non mutati con  $p=0.96$  e un'attenuazione della significatività nei KLF1 mutati ( $p=0.04$ ).

Il polimorfismo rs7482144 (XmnI) nelle nostre analisi di associazione in relazione all'HbF e all'HbA2 non ha dato valori statisticamente significativi. Nel dettaglio, per l'HbF i valori di significatività negli HPFH in toto ( $p=0.6$ ) si avvicinano a valori leggermente più significativi ( $p=0.2$ ) nelle due categorie KLF1 mutati e KLF1 non mutati. In relazione all'HbA2 i valori di associazione si avvicinano alla significatività passando da  $p=0.13$  (HPFH tutti) a  $p=0.5$  (KLF1 non mutati) a  $p=0.7$  (KLF1 Mutati).

I dati suggeriscono che i vari determinanti correlino in modo diverso con le variazioni emoglobiniche nei soggetti positivi rispetto ai soggetti negativi per mutazioni del gene KLF1. In particolare, l'analisi della significatività dell'associazione mostra un evidente aumento della significatività per quasi tutti i determinanti nella popolazione KLF1 mutata. I polimorfismi rs67385638 e XmnI non raggiungono la significatività nelle popolazioni HPFH distinte, tuttavia l'rs67385638 mostra un trend verso la significatività che fa presupporre che in futuro aumentando il campione esaminato si evidenzino valori significativi. Appare importante rilevare che il polimorfismo rs67385638 è l'unico a mostrare un trend verso la quasi significatività nella popolazione di soggetti HPFH non mutati. Questo risultato, se confermato in una casistica più ampia, sarebbe a favore di un ruolo causale diverso di questo polimorfismo. E' possibile che i polimorfismi rs2855122 (HBG2), rs1427407 (BCL11A), rs9399137 (MYB) abbiano necessità di una preattivazione del gene gamma globinico per esplicare il loro effetto stimolante sulla sua espressione che richiederebbe l'intervento regolatore dei fattori di trascrizione BCL11A, KLF1 e forse MYB. Questa preattivazione potrebbe essere innescata da

un'interazione preferenziale dei geni gamma con la LCR, evento dimostrato con lo studio di espressione dell'mRNA nei modelli murini transgenici, per il cluster globinico umano e knockout per il gene KLF1 [52]. Secondo la nostra ipotesi, il polimorfismo rs67385638, l'unico a correlare con variazioni HbF negli HPFH non mutati, potrebbe agire direttamente o indirettamente in un pathway d'attivazione diverso che non richieda la preattivazione dei geni gamma e che perciò potrebbe agire anche su geni gamma che siano completamente silenziati. Poiché l'analisi in silico predice nella regione del polimorfismo il legame differenziale dei fattori di trascrizione C/EBP $\beta$  e CP2, entrambi fattori implicati nei meccanismi di switching e di regolazione dell'emoglobina, è possibile che siano proprio questi fattori i mediatori della stimolazione dell'espressione dei geni fetali ascrivibile al polimorfismo rs67385638 [53]. Il polimorfismo rs2855122, che era anch'esso oggetto del presente studio, ha invece un comportamento in linea con i polimorfismi di BCL11A e MYB e quindi potrebbe essere coinvolto nello stesso pathway di attivazione dell'HbF mediata dai geni KLF1>BCL11A e/o MYB.

## Bibliografia

- 1) Weatherall DJ. and Clegg JB. *The thalassemia syndromes*. Blackwell Scientific Publication 2001.
- 2) Wilber A. et al., *Transcriptional regulation of fetal to adult hemoglobin switching: new therapeutic opportunities*. Blood.2011;117(15):3945-53.
- 3) Thein SL., *Control of fetal hemoglobin: new insights emerging from genomics and clinical implications*. Human Molecular Genetics 2009 ;18(R2):R216-23.
- 4) Schechter AN. *Hemoglobin research and origins of molecular medicine*. Blood.2008;112:3927-38.
- 5) Paleari R. et al., *Utilità clinica della misura dell'emoglobina fetale*. Biochimica Clinica 2009; vol.33 n°6.
- 6) Wood WG. "*Increased HbF in adult life*". Baillière's Clinical Haematology. 1993;vol.6, Issue 1:177–213.
- 7) Olave IA., et al., "*Purification and identification of proteins that bind to the hereditary persistence of fetal hemoglobin -198 mutation in the gamma-globin gene promoter*". J Biol Chem. 2007;282(2):853-62.
- 8) Ottolenghi S. et al., "*DNA sequences regulating human globin gene transcription in non deletional hereditary persistence of fetal hemoglobin*". Hemoglobin. 1989;13(6):523-41.
- 9) Gilman JG., et al., "*DNA sequence variation associated with elevated fetal G gamma globin production*". Blood. 1985;66(4):783-7.
- 10) Labie D., et al., "*The -158 site 5' to the G gamma gene and G gamma expression*" Blood. 1985;66(6):1463-5.
- 11) Thein SL. "*Genetic modifiers of beta-thalassemia*". Haematologica. 2005;90(5):649-60.
- 12) Pissard S. et al., "*A potential regulatory region for the expression of fetal hemoglobin in sickle cell disease*". Blood.1994;84:331-338.
- 13) Kulozik AE., et al., "*Fetal hemoglobin levels and  $\beta$  (s) globin haplotypes in an Indian population with sickle cell disease*". Blood.1987;69:1742-1746.
- 14) Thein SL., *Association of thalassaemia intermedia with a beta-globin gene haplotype*. Br J Haematol 1987;65:367-373.
- 15) Cao A. et al., *Recent advances in  $\beta$ -thalassemias*. Pediatric Reports 2011; v3:e17.
- 16) Menzel S., et al. *A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15*. Nat Genet. 2007;39:1197-1199.
- 17) Uda M., et al. *Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassemia*. Proc Natl Acad Sci. USA 2008;105:1620-1625.
- 18) Sedgewick AE. et al., *BCL11A is a major HbF quantitative trait locus in three different populations with beta-hemoglobinopathies*. Blood Cells Mol Dis. 2008;41:255-258.
- 19) Sankaran VG., et al. *Human fetal hemoglobin expression is regulated by the development stage-specific repressor BCL11A*. Science. 2008; 332:1839-42.

- 20) Liu P., et al. *Bcl11a is essential for normal lymphoid development*. Nat. Immunol. 2003; 4:525-32.
- 21) Xu J. et al., *Transcriptional silencing of  $\gamma$ -globin by BCL11A involves long-range interactions and cooperation with SOX6*. Genes Dev. 2010; 24(8): 783-98.
- 22) Sankaran VG. et al. *Advances in the understanding of haemoglobin switching*. Br J Haematol. 2010; 149: 181-94.
- 23) Bieker JJ. et al., *Putting a finger on the switch*. Nature Genet. 2010; 42:733-46.
- 24) Sankaran VG., et al., *Developmental and species-divergent globin switching are driven by BCL11A*. Nature. 2009; 460:1093-7.
- 25) Galarneau G., et al., *Fine-mapping at three loci known to affect fetal hemoglobin levels explains additional genetic variation*. Nat Genetics.2010; 42(12):1049-51.
- 26) Lettre G., et al. *DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease*. Proc Natl Acad Sci U S A. 2008; 105(33):11869-74.
- 27) Galanello R. et al., *Amelioration of Sardinian beta0 thalassemia by genetic modifiers*. Blood.2009; 114(18):3935-7.
- 28) Bauer DE. et al., *An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level*. Science.2013; 342(6155):253-7.
- 29) Bauer DE. et al., *Hemoglobin switching's surprise: the versatile transcription factor BCL11A is a master repressor of fetal hemoglobin*. Current Opinion in Genetics & Development.2015; 33:62-70.
- 30) Chen Z .et al., *BCL11A represses HBG transcription in K562 cells*. Blood Cell Mol Dis. 2009; 42(2): 144-9.
- 31) Wahlberg K. et al., *The HBS1L-MYB intergenic interval associated with elevated HbF levels shows characteristics of a distal regulatory region in erythroid cells*. Blood.2009; 114(6):1254-62.
- 32) Craig JE. et al., *Dissecting the loci controlling fetal haemoglobin production on chromosomes 11p and 6q by the regressive approach*. Nature Genetics.1996;12(1):58-64.
- 33) Garner C. et al., *Haplotype mapping of a major quantitative-trait locus for fetal hemoglobin production, on chromosome 6q23*. American Journal of Human Genetics. 1998; 62(6):1468-74.
- 34) Thein SL.et al., *Intergenic variants of HBS1L-MYB are responsible for a major quantitative trait locus on chromosome 6q23 influencing fetal hemoglobin levels in adults*. Proc Natl Acad Sci USA. 2007. 3;104(27).
- 35) Jiang J, et al., *cMYB is involved in the regulation of fetal hemoglobin production in adults*. Blood 2006;108(3):1077-1083.
- 36) Farrell JJ. et al., *A 3-bp deletion in the HBS1L-MYB intergenic region on chromosome 6q23 is associated with HbF expression*. Blood. 2011.5;117(18):4935-45.
- 37) Armstrong JA. et al., *A SWI/SNF-related chromatin remodeling complex E-RC1, is required for issue-specific transcriptional regulation by EKLF in vitro*. Cell. 1998;95:93-104.

- 38) Feng WC. et al., *Analyses of beta-thalassemia mutant DNA interactions with erythroid Kruppel-like factor (EKLF), an erythroid cell-specific transcription factor.* J Biol Chem. 1994;269 2: 1493–500.
- 39) Perkins AC. et al., *Lethal  $\beta$ -thalassemia in mice lacking the erythroid CACCC-transcription factor EKLF.* Nature 1995;375, 318-22.
- 40) Nuez B. et al., *Defective haematopoiesis in fetal liver resulting from inactivation the EKLF gene.* Nature.1995;375:316-18.
- 41) Moi P. al., *A novel silent  $\beta$ -thalassemia mutation in the distal CACCC box affects the binding and responsiveness to EKLF.* Br J Haematol.2004;126:881-4.
- 42) Hodge D, et al., *A global role for EKLF in definitive and primitive erythropoiesis.* Blood. 2006;107 8: 3359–70.
- 43) Singleton BK. et al., *Mutations in EKLF/KLF1 form the molecular basis of the rare blood group In(Lu) phenotype.* Blood.2008;112:2081-8.
- 44) Siatecka M. al., *Severe anemia in the Nan mutant mouse caused by sequence-selective disruption of erythroid Krüppel-like factor.* Proc Natl Acad Sci USA. 2010;107(34):15151-6.
- 45) Zhou D. et al., *KLF1 regulates BCL11A expression and gamma to beta-globin gene switching.* Nat Genet. 2010; 42: 742-4.
- 46) Borg J. et al., *Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin.* Nat Genet. 2010; 42: 801-5.
- 47) Satta S. et al., *Compound heterozygosity for KLF1 mutations associated with remarkable increase of fetal hemoglobin and red cell protoporphyrin.* Hematologica. 2011; 96:2-5.
- 48) Garner CP. et al., *Evidence of genetic interaction between the beta-globin complex and chromosome 8q in the expression of fetal hemoglobin.* Am J Hum Genet. 2002;70(3):793-9.
- 49) Dover GJ,et al., *Fetal hemoglobin levels in sickle cell disease and normal individuals are partially controlled by an X-linked gene located at Xp22.2.* Blood. 1992 1;80(3):816-24.
- 50) Danjou F. et al., *Genome-wide association analyses based on whole-genome sequencing in Sardinia provide insights into regulation of hemoglobin levels.* Nat Genet. 2015;47(11):1264-71.
- 51) Purcell S et al., PC (2007) *PLINK: a toolset for whole-genome association and population-based linkage analysis.* American Journal of Human Genetics, 81.
- 52) Porcu S. et al., *Beta-Minor globin gene expression is preferentially reduced in EKLF Knock-Out mice.* Gene. 2005; 351:11-7.
- 53) Stephen M Jane et al., *Methylation-Enhanced Binding of Spl to the Stage Selector Element of the Human  $\gamma$ -Globin Gene Promoter May Regulate Developmental Specificity of Expression stage selector element.* Molecular and cellular biology. 1993, p. 3272-3281.

La presente tesi è stata prodotta durante la frequenza del corso di dottorato in Terapia pediatrica e farmacologia dello sviluppo, dell'Università degli Studi di Cagliari, a.a. 2014/2015 - XXVIII ciclo, con il supporto di una borsa di studio finanziata con le risorse del P.O.R. SARDEGNA F.S.E. 2007-2013 - Obiettivo competitività regionale e occupazione, Asse IV Capitale umano, Linea di Attività 1.3.1 "Finanziamento di corsi di dottorato finalizzati alla formazione di capitale umano altamente specializzato, in particolare per i settori dell'ICT, delle nanotecnologie e delle biotecnologie, dell'energia e dello sviluppo sostenibile, dell'agroalimentare e dei materiali tradizionali".

"Laura Manunza gratefully acknowledges Sardinia Regional Government for the financial support of her PhD scholarship (P.O.R. Sardegna F.S.E. Operational Programme of the Autonomous Region of Sardinia, European Social Fund 2007-2013 - Axis IV Human Resources, Objective 1.3, Line of Activity 1.3.1.)".