# UNIVERSITÀ DEGLI STUDI DI NAPOLI

# **"FEDERICO II"**

# Scuola di Dottorato in Medicina Molecolare

# Dottorato di Ricerca in Patologia e Fisiopatologia Molecolare XX ciclo



# "Alterazione dell'espressione genica determinata dalla trisomia 21 in cuori di feti umani con la sindrome di Down"

Coordinatore: Prof. Vittorio Enrico Avvedimento Candidato: Dott. Fabbrini Floriana

Anno

2007

# **INDICE**

## INTRODUZIONE

1.1 Sindrome di Down (DS)				
1.1.1	La Down Syndrome Critical Region (DSCR)	pag. 4		
1.2 Model	pag. 5			
1.3 Cardio	pag. 9			
1.3.1	Sviluppo embrionale del cuore	pag. 11		
1.3.2	Morfogenesi dei setti	pag. 12		
1.3.3	Morfologia del cuore	pag. 14		
1.3.4	Geni implicati nella morfogenesi del cuore	pag, 15		
1.3.5	MicroRNA e cuore	pag. 19		
1.4 Espres	pag. 21			
1.4.1	Tecniche di valutazione dell'espressione genica:	pag. 22		
	microarray e qRT-PCR			
SCOPO I	pag. 31			
MATERIALI E METODI				
RISULTA	pag. 43			
DISCUSS	pag. 58			
CONCLUSIONI				
BIBLIOGRAFIA				

## **INTRODUZIONE**

#### 1.1 Sindrome di Down.

La Sindrome di Down (DS) è la più frequente aneuploidia autosomica compatibile con la vita; si verifica per una completa o parziale trisomia del cromosoma 21 (Hsa21) con una frequenza alla nascita di circa 1:700.

Fu individuata per la prima volta nel 1866 da un medico inglese, John Longdong Down, che descrisse questa patologia usando il termine "mongolismo" per la forma degli occhi degli individui affetti simile a quella delle popolazioni orientali (Down et al,1866). La causa di questa patologia fu scoperta però molti anni più tardi, nel 1959, per merito di Lejeune che identificò, nelle cellule dei pazienti affetti, la presenza di un cromosoma 21 sovrannumerario (Lejeune et al, 1959).

La DS è caratterizzata da un complesso fenotipo, variabile in termini di penetranza ed espressività, che include più di 80 caratteristiche cliniche. Nel soggetto trisomico sono segni costanti il ritardo mentale, di grado variabile, e l'ipotonia; inoltre i soggetti affetti presentano un ritardo nella crescita, un parziale deficit immunologico, soprattutto a carico del sistema timo-dipendente, e un rischio maggiore di ammalarsi di leucemia. Circa la metà dei soggetti con DS è affetto da cardiopatia congenita.

Vi sono, all'interno del gruppo Down, profonde differenze individuali nelle caratteristiche psichiche, biologiche e fisiche. Si pensa, infatti, che sia l'intero patrimonio genetico a determinare le diverse manifestazioni individuali. Si ritiene, inoltre, che l'espressione dei geni in triplice copia condizioni il funzionamento di moltissimi altri geni, attivandoli o inibendoli in maniera variabile in base alla costituzione genetica individuale. Sarebbe questa la base teorica per la non prevedibilità a priori degli effetti dell'extracromosoma 21, sia per quanto riguarda l'aspetto fenotipico in senso lato, sia per quanto riguarda il manifestarsi delle varie situazioni patologiche (Reeves et al, 2001).

#### 1.1.1 Down Syndrome Critical Region.

Sebbene la DS sia causata dalla presenza di un cromosoma 21 sovrannumerario, si ritiene che non sia la triplice copia del cromosoma 21 a determinare le complesse alterazioni fenotipiche, bensì un numero ristretto di geni localizzati sul cromosoma 21 in corrispondenza della banda 21q22 (Korenberg et al, 1990). I prodotti di tali geni risulterebbero sensibili al cosiddetto "effetto dosaggio genico", ovvero alla presenza di una copia extra del cromosoma 21 (Hattori et al, 2000).

Dallo studio di pazienti con parziale trisomia 21, è stato possibile definire, da un punto di vista molecolare, la regione minima responsabile del fenotipo Down (Korenberg et al, 1990). Tale regione, denominata *Down Syndrome Critical Region* (DSCR), si estende per circa 5.4 Mb e comprende parte della banda 21q22.2 e la banda 21q22.3. I geni contenuti in essa sarebbero causa dei dismorfismi faciali, della bassa statura, dell'ipotonia muscolare e del ritardo mentale (**fig.1**). Tuttavia i geni del 21 che mappano al di fuori di tale regione non devono essere ignorati poichè alcuni di essi contribuiscono ad alcune manifestazioni cliniche della DS.

Attraverso l'approccio della regione critica la Korenberg nel 2001 ha individuato all'interno della banda 21q22 una regione "critica" ristretta per lo sviluppo della cardiopatia, la DS-CHD (*Down Syndrome Congenital Heart Disease*).



Fig. 1 A: La DSCR (Down Syndrome Critical Region) comprende parte della banda 21q22.2 e la banda 21q22.3; B: DS-CHD, regione critica per lo sviluppo della cardiopatia. La DS-CHD è solo una parte della DSCR.

Tale regione, demarcata da D21S3 e PFKL è solo una frazione della regione definita critica per la DS; essa abbraccia parte della banda 21q22.3 (**fig.1**) e, tra i geni in essa localizzati, DSCAM (Down Sindrome cell adhesion molecule) è stato proposto come uno dei geni candidati per lo sviluppo della cardiopatia (Barlow et al, 2001). L'ipotesi DSCR è stata testata nei topi (Olson et al. 2004, 2006) e si è visto che la trisomia della DSCR da sola è necessaria ma non sufficiente per i fenotipi cerebrali nei topi trisomici. Questi risultati suggeriscono che le origini dei fenotipi trisomici sono più complicate di quanto si sosteneva in passato e che probabilmente coinvolgono interazioni tra più geni (Olson et al. 2006). Si è pensato, infatti, che le complesse alterazioni fenotipiche della DS potessero dipendere da un'interazione tra geni Hsa21 e geni presenti altrove nel genoma (Reeves et al, 2001) e che la perdita dell'equilibrio genetico nei processi chiave che regolano lo sviluppo potrebbe aumentare la suscettibilità a disordini genetici e ambientali (Shapiro et al. 1997).

#### 1.2 Modelli murini per lo studio della DS

I modelli animali di trisomia 21 hanno dato un contributo importante nella comprensione dell'effetto dell'espressione dei geni trisomici sullo sviluppo e sulla funzione dei geni.

Il cromosoma 21 presenta un'elevata omologia di sequenza con il cromosoma 16 murino; ciò ha reso possibile la creazione di topi transgenici trisomici per il cromosoma 16 per lo studio della DS. La regione del cromosoma 21 sintenica con il cromosoma 16 murino comprende circa 30 Mb e si estende dalla regione quasi centromerica del cromosoma 21 fino al gene MX1, il cui locus genico è omologo alla regione telomerica del cromosoma 16 murino (Gardiner, 2000). La regione di sintenia tra il cromosoma umano e quello di topo include la DSCR.

Altre regioni di sintenia sono state riscontrate tra il cromosoma 21 umano ed i cromosomi 10 e 17 murini. Nella regione telomerica del cromosoma 21 vi è un segmento di circa 2 Mb che è omologo ad un segmento interno del cromosoma 10 murino, mentre un altro segmento di circa 2 Mb del cromosoma 21 risulta omologo alla regione quasi centromerica del cromosoma 17 murino (Gardiner et al, 2000) (**fig. 2**).



**Fig. 2** A sinistra sono illustrate le regioni di sintenia tra il cromosoma 21 umano e i cromosomi 16, 17 e 10 murini, a destra le regioni di sintenia trisomiche nei due modelli murini, Ts65Dn e Ts1Cje

Il primo modello murino della DS, il topo Ts16, trisomico per l'intero cromosoma 16 contenente 154 ortologhi umani (**fig. 3**), presenta molte caratteristiche del fenotipo Down, inclusi i difetti cardiaci che sono presenti nel 100% dei topi transgenici. Poiché nell'uomo la percentuale di cardiopatia negli individui Down è all'incirca il 50%, si ritiene che i difetti cardiaci nel topo Ts16 possono essere causati da geni al di fuori della DSCR regione omologa al cromosoma 21 umano. Lo studio di tale modello è stato però fortemente limitato dal fatto che i topi generalmente non sopravvivono, muoiono in utero in tarda età gestazionale. Inoltre, la correlazione genotipo-fenotipo risulta ambigua per la presenza in triplice copia, sul cromosoma 16 murino, di ortologhi dei cromosomi 3, 16 e 22 umani (Moore, 2006).



Fig. 3 Modelli murini della DS

I successivi modelli murini, Ts65Dn e Ts1Cje, sono topi con parziale trisomia 16, in cui cioè il segmento cromosomico murino in eccesso e sintenico con il cromosoma 21 umano è solo una porzione della regione del cromosoma 16 murino comprendente la DSCR.

Nei topi Ts65Dn il segmento cromosomico presente in triplice copia e sintenico con il cromosoma 21 umano si estende per circa 15 Mb ed è delimitato dai geni APP (21q21.2) e MX1 (21q22.3), mentre nei topi Ts1Cje si estende per circa 10 Mb ed è delimitato dai geni SOD (21q22.1), non attivo funzionalmente, e MX1 (Sago et al,1998). Questi topi transgenici manifestano tra loro diverse caratteristiche fenotipiche equiparabili solo ad alcune delle alterazioni fenotipiche della DS. Recentemente è stato riscontrato che il topo Ts65Dn, seppure in piccola percentuale (8%), sviluppa la cardiopatia. Si ipotizza anche che siano i difetti cardiaci, in particolare difetti di settazione, ad incrementare il tasso di mortalità post-natale nei topi Ts65Dn (Moore, 2006).

L'utilizzo del modello Ts65Dn per lo studio della DS è stato però fortemente limitato dal fatto che la regione cromosomica presente in triplice copia contiene alcuni geni che sono ortologhi al cromosoma 17 umano. Ciò suggerisce che vi possono essere anche altre regioni responsabili delle caratteristiche fenotipiche della DS (fig. 3).

I topi Ts1Rhr trisomici per i geni del cromosoma 16 ortologhi a quelli contenuti nella DSCR non presentano le anomalie craniofaciali tipiche dei topi Ts65Dn (Olson, 2004). Lo studio di tale modello murino porta a concludere che l'overespressione di geni critici non è sufficiente a determinare il complesso fenotipo Down. Ulteriori studi hanno dimostrato che la trisomia della DSCR da sola è necessaria ma non sufficiente per i fenotipi cerebrali nei topi trisomici Ts1Rhr (Olson et al, 2007) (**fig. 3**). Questi risultati suggeriscono che la causa del fenotipo Down è più complessa di quanto si sosteneva in passato e che probabilmente sono coinvolte interazioni tra più geni (Olson et al, 2007). E' stato suggerito, infatti, che le complesse alterazioni fenotipiche della DS dipendano da un' interazione tra geni Hsa21 overespressi e geni regolati durante lo sviluppo presenti altrove sul genoma (Reeves et al, 2001) e che la perdita dell'equilibrio genetico nei processi chiave che regolano lo sviluppo potrebbe aumentare la suscettibilità a disordini genetici e ambientali ed influenzare il grado di penetranza ed espressività del fenotipo Down.

I topi "transcromosomici" Tc1, portatori di un cromosoma 21 umano (**fig. 3**), presentano anomalie craniofaciali ed alterazioni cardiache tipiche della DS. I difetti morfogenetici includono il difetto del setto ventricolare ed il difetto del setto atrioventricolare, quest'ultimo dovuto all'incompleta fusione dei cuscinetti endocardici (O'Doherty et al, 2005).

Nel più recente modello murino, Dp(16)1Yu/+, la regione cromosomica in eccesso e sintenica con il cromosoma 21 si estende dalla banda 21q11 alla banda 21q22.3 ed i geni che la delimitano sono D9300038D03Rik, ortologo del gene LIPI, e Zfp295, ortologo del gene ZNF295 (**fig. 4**). Essa comprende, quindi, parte della DS-CHD. Questi topi transgenici sviluppano cardiopatia congenita in una percentuale paragonabile a quella riscontrata nell'uomo. Le malformazioni cardiache riscontrate in questi topi sono le stesse riscontrate negli individui Down (Li et al,2007): la tetralogia di Fallot, i difetti del setto interatriale ed interventricolare ed il canale atrioventricolare.



Fig. 4 Regione di sintenia trisomica nel modello murino Dp(16)1Yu/+

#### 1.3 Cardiopatie congenite e DS

Le cardiopatie congenite sono malattie che si riscontrano alla nascita con un'incidenza pari a 8/1000 nati vivi e rappresentano un problema importante per la salute nell'età infantile. Sebbene molte malformazioni cardiache siano state ben caratterizzate sia dal punto di vista anatomico che funzionale, i meccanismi patogenetici responsabili non sono stati ancora identificati.

I difetti congeniti del cuore (CHD –Congenital Heart Disease-) sono dovuti ad errori che si verificano durante lo sviluppo cardiaco nella vita embrionale e sono presenti in circa la metà dei neonati affetti dalla DS. Una così alta incidenza di CHD in questi pazienti suggerisce che i geni che mappano sul cromosoma 21 possano essere coinvolti nella morfogenesi del cuore e che la loro espressione anomala negli individui trisomici possa alterare il normale processo di sviluppo del cuore o direttamente o influenzando l'espressione di altri geni.

Le malformazioni cardiache presenti negli individui con DS sono soprattutto difetti di settazione: difetti del canale atrioventricolare (AVC), difetti del setto

ventricolare (VSD) o malformazioni più complesse, come la tetralogia di Fallot (**fig. 5**).





d

Fig. 5 a: Difetto del setto ventricolare; b: Difetto del setto striale; c: Difetto del setto atrioventricolare; d: Tetralogia di Fallot

С

Esistono diverse forme di AVC a seconda della complessità delle alterazioni anatomiche. Nella forma più semplice si ha un foro nella parte inferiore del setto interatriale che, associato ad una malformazione della valvola mitrale, costituisce la forma cosiddetta incompleta o parziale. Il malfunzionamento della valvola mitrale fa si che il sangue refluisca direttamente dal ventricolo sinistro nell'atrio destro, mentre, in condizioni normali il sangue dal ventricolo sinistro viene pompato nell'aorta dalla quale raggiunge tutto il corpo per mezzo di capillari.

Nella forma completa, oltre a queste anomalie, si ha un foro nella parte alta del setto interventricolare ed un'alterazione delle valvole atrioventricolari. Ne deriva una comunicazione tra i due atrii associata ad una comunicazione tra i due ventricoli ed alla presenza di una valvola atrioventricolare unica.

Il VSD è una cardiopatia congenita caratterizzata dalla presenza di una o più comunicazioni tra le cavità ventricolari attraverso il setto interventricolare. Il VSD può presentarsi come lesione isolata oppure fare parte di un'anomalia più complessa, come, per esempio, la Tetralogia di Fallot.

La Tetralogia di Fallot è una malformazione cardiaca complessa caratterizzata da quattro elementi anatomici fondamentali:

- La comunicazione fra i due ventricoli, le due parti pompanti del cuore (difetto interventricolare).
- L'origine biventricolare dell'aorta, che si trova a cavaliere fra i due ventricoli, sopra il difetto interventricolare.
- Una stenosi (restringimento) sottovalvolare e valvolare polmonare.
- Un'ipertrofia (cioè ingrossamento muscolare) del ventricolo destro.

#### 1.3.1 Sviluppo embrionale del cuore

Il sistema cardiocircolatorio è la prima unità funzionale dell'embrione, ed il cuore il primo organo funzionante. Il cuore origina dal mesoderma.

L'inizio dello sviluppo del cuore si manifesta al 18°-19° giorno nell'*area cardiogenica*; in questa zona le cellule cardiache progenitrici si aggregano a formare un paio di cordoncini allungati detti *cordoncini cardiogenici*. Questi cordoni poi si cavitano e formano due tubi endoteliali a parete sottile, detti *tubi endocardici* destro e sinistro, che gradualmente si avvicinano l'uno all'altro sino a fondersi a formare un unico *tubo endocardico*. Man mano che il singolo tubo cardiaco si allunga si originano una serie di dilatazioni e costrizioni che portano

alla formazione del *bulbo cardiaco* (BC), del *ventricolo primitivo* (V), del *tronco arterioso* (TA) e del *seno-atrio* (SA) o *atrio primitivo* (**fig. 6**). Dal momento che il bulbo ed il ventricolo si accrescono più rapidamente delle altre regioni, il tubo cardiaco si flette su se stesso, formando la *piega bulbo ventricolare ad U*; in seguito si forma un *cuore ad S* (**fig. 6c**).

Il bulbo cardiaco ed il ventricolo primitivo daranno origine rispettivamente al ventricolo destro e sinistro, il tronco arterioso darà vita all'aorta e all'arteria polmonare, ed il seno-atrio agli atrii destro e sinistro.



**Fig. 6** Sviluppo embriologico del cuore. A: stadio di tubo cardiaco primitivo. B, C, D: fasi successive dello sviluppo fino alla formazione di un cuore con 4 cavità (E).

#### 1.3.2 Morfogenesi dei setti

**Suddivisione del canale atrioventricolare**: Sulle pareti dorsale e ventrale del cuore, nella regione interna del canale atrioventricolare, si formano degli ispessimenti del tessuto subendocardico, detti rilievi o *cuscinetti endocardici*. Durante la quinta settimana di gestazione i cuscinetti endocardici crescono l'uno verso l'altro e si fondono, dividendo il canale atrioventricolare in due canali atrioventricolari, destro e sinistro (**fig. 7**).



Fig. 7 Morfogenesi dei canali atrioventricolare destro e sinistro.

**Suddivisione dell'atrio primitivo**: In corrispondenza dell'atrio primitivo si forma una sottile membrana, il *septum primum o setto primario*. Man mano che questo setto si accresce in direzione dei cuscinetti endocardici si forma un ampio foro, il *forame primum* che si restringe gradualmente perché il septum primum si accresce verso i cuscinetti endocardici. Prima che la chiusura del forame primum sia completa, il septum primum si perfora in più punti e queste lacune confluiscono a formare una nuova apertura, il *forame secundum*. Verso la fine della quinta settimana di gestazione si sviluppa un'altra membrana, il *septum secundum* che, accrescendosi verso i cuscinetti endocardici, ricopre gradualmente il forame secundum del septum primum. Il septum secundum forma una parete incompleta e lascia un'apertura ovale detta *forame ovale* (**fig. 8**). Prima della nascita il forame ovale permette alla maggior parte del sangue di passare dall'atrio destro nell'atrio sinistro escludendo il passaggio nei polmoni. Dopo la nascita, il forame ovale normalmente si richiude ed il setto interatriale diviene una parete completa.



Fig. 8 Morfogenesi dei setti atriale e ventricolare

**Suddivisione del ventricolo primitivo**: In prossimità del ventricolo primitivo si sviluppa una cresta o piega muscolare, il *setto interventricolare* che si accresce in lunghezza in prossimità dei cuscinetti endocardici (**fig. 8**). In seguito, si ha un'intensa crescita del tessuto del setto e si forma la *porzione muscolare del setto interventricolare*. Fra il margine libero del setto interventricolare ed i cuscinetti endocardici fusi vi è il forame interventricolare che permette la comunicazione tra ventricolo destro e sinistro fino al termine della 7° settimana. Alla chiusura di questo forame partecipa anche il setto che divide il bulbo ed il tronco arterioso in *arteria polmonare ed aorta*. L'estremità inferiore di questo setto discende fino a raggiungere il margine libero del setto interventricolare, la *parte membranosa del setto interventricolare*; quest'ultimo diviene una parete completa.

#### 1.3.3 Morfologia del cuore

Esternamente il cuore è costituito da 3 tonache: endocardio, miocardio ed epicardio. Internamente, invece, è suddiviso in una metà destra e sinistra; in ciascuna di esse si distingue una atrio ed un ventricolo comunicanti attraverso due ristretti orifizi provvisti di valvole atrioventricolari: la valvole tricuspide per la porzione destra del cuore e la valvola bicuspide o mitrale per la porzione sinistra (**fig. 9**).

Nella sezione venosa, quella di destra, la vena cava inferiore, la vena cava superiore ed ilo seno coronario immettono il sangue nell'atrio destro. Il sangue,

attraverso la valvola tricuspide, passa nel ventricolo destro che lo spinge nell'arteria polmonare, affinché sia condotto ai polmoni per ossigenarsi. La porzione sinistra del cuore può, a questo punto, raccogliere il sangue arterioso (ricco di  $O_2$ ) proveniente dai polmoni per distribuirlo ai tessuti. Nell'atrio sinistro si immettono singolarmente le 4 vene polmonari (2 per polmone) che portano al cuore il sangue ossigenato dai polmoni. Quest'ultimo, attraverso la valvola bicuspide, passa nel ventricolo sinistro che lo spinge nell'aorta dalla quale può raggiungere tutto il corpo.



Fig. 9 Morfologia del cuore

#### 1.3.4 Geni implicati nella morfogenesi del cuore

Lo sviluppo del cuore durante la vita embrionale è regolato da un complesso network di geni che codificano per fattori di trascrizione che sono conservati nel corso dell'evoluzione (Olson, 2006). Il *core* di fattori trascrizionali essenziale per lo sviluppo del cuore comprende i seguenti componenti: MEF2, NKx2.5, GATA, TBX e Hand2 (**fig. 10**). (Olson, 2006)



**Fig. 10** Networks di geni coinvolti nello sviluppo del cuore nei mammiferi. Segnali intracellulari attivano un set di geni codificanti per fattori di trascrizione responsabili della formazione delle diverse componenti del cuore. I prodotti di tali geni attivano a loro volta i geni del *core cardiac network*. Tali geni regolano l'attivazione di geni muscolo-specifici e di geni che inducono la crescita cellulare e la formazione delle regioni primarie e secondarie cardiache.

Nella regione primaria avranno origine il ventricolo sinistro e l'area atriale, mentre nella regione secondaria il ventricolo destro, la regione atriale e il tratto di efflusso.

Le componenti cardiache derivanti dalla regione primaria e secondaria sono mostrate in blu e rosa, rispettivamente. Gli atrii, che derivano da entrambe le regioni, sono mostrati in viola.

Il fattore di trascrizione MEF2 appartiene alla classe delle MADS-box protein, ossia a quella classe di proteine il cui dominio di legame al DNA è caratterizzato da una specifica sequenza di 56 amminoacidi. MEF2 è cruciale per il differenziamento in senso muscolare. Esso governa diversi processi cellulari a seconda degli stimoli e dei fattori proteici con cui interagisce. L'interazione di MyoD con i domini MADS di MEF2 guida il programma genetico del differenziamento muscolare scheletrico, mentre l'associazione della stessa regione di MEF2 con i fattori di trascrizione zinc-finger della famiglia GATA regola l'espressione di geni coinvolti nello sviluppo cardiaco. Nelle cellule muscolari

cardiache MEF2 coopera con gli altri fattori di trascrizione regolando l'espressione di geni che codificano per proteine contrattili (Olson, 2006).

Il fattore di trascrizione NKx2.5, appartenente alla classe dei geni homeobox, è espresso nelle cellule progenitrici cardiache. Mutazioni nel gene codificante per NKx2.5 causano difetti congeniti del cuore, in particolar modo difetti del setto interatriale e disfunzioni delle valvole atrioventricolari (Schiaffino et al, 1999).

Nei vertebrati solo tre dei sei geni che codificano per i fattori di trascrizione GATA (GATA4, GATA5, GATA6) sono espressi in cuore. Essi sono caratterizzati da due motivi zinc fingers in tandem separati da 29 amminoacidi. Questi motivi sono essenziali per il legame al DNA e per l'interazione con altri fattori regolatori (Ohara et al, 2005). Il motivo zinc finger C-terminale è quello principalmente coinvolto nelle interazioni proteina-proteina (Reamon-Buettner et al, 2007).

GATA4 gioca un ruolo chiave nello sviluppo embrionale del cuore; esso è indispensabile per la formazione del tubo cardiaco e per la morfogenesi ventrale dell'embrione (Molkentin et al, 1997). La regolazione dei geni cardiaci mediata da GATA4 è cooperativa, infatti esso interagisce con numerosi fattori di trascrizione tra cui NKx2.5, TBX5, MEF2 e GATA6.

Recenti studi dimostrano che TBX5 e TBX20 sono co-espressi sin dalle prime fasi dello sviluppo cardiaco nella sezione atriale, ma la loro espressione diverge nel corso della cardiogenesi (Plageman et al, 2004). La loro espressione è importante per la definizione dei ventricoli destro e sinistro.

L'espressione differenziale di TBX5 e TBX20 nel corso dello sviluppo cardiaco correla con ruoli differenti: TBX5 induce la formazione del ventricolo sinistro, mentre TBX20 la formazione del ventricolo destro e del canale atrioventricolare (Plageman et al, 2004). Inoltre, TBX20 regola anche la maturazione e proliferazione delle cellule costituenti i cuscinetti endocardici e l'espressione dei geni della matrice extracellulare (Shelton et al, 2006).

Sebbene questi due fattori abbiano un dominio di legame al DNA altamente conservato (dominio T-box) mostrano una diversa attività trascrizionale. TBX5 sembra attivare da solo e/o sinergicamente con NKx2.5 e GATA4 il gene cardiaco ANF, il cui promotore contiene la sequenza di legame a TBX5; TBX20, invece, sembra avere un effetto repressorio sull'attività trascrizionale di ANF (Plageman et al,2004).

Mutazioni nel gene che codifica per il fattore di trascrizione TBX5 sono causa della Sindrome di Holt-Oram, le cui alterazioni fisiologiche e fenotipiche includono difetti di settazione del cuore (difetti del setto atriale e ventricolare, e disfunzioni delle valvole atrioventricolari), conduzione cardiaca difettiva e difetti agli arti superiori (Tacheuchi et al,2003).

Di recente sono state individuate nel motivo T-box di TBX20 due mutazioni, una missenso ed una nonsenso, che sono causa dei difetti congeniti del cuore. La prima abbassa notevolmente l'attività trascrizionale di TBX20, la seconda, invece, porta alla formazione di una proteina tronca (Kirk et al,2007). Entrambe le mutazioni impediscono a TBX20 di interagire con i fattori trascrizionali che regolano l'espressione dei geni cardiaci provocando una serie di anomalie come difetti di settazione e formazione anomala delle valvole atrioventricolari.

I geni che codificano per i fattori di trascrizione Hand sono coinvolti nella morfogenesi dei ventricoli: Hand1 sembra essere coinvolto nella formazione del ventricolo sinistro, mentre Hand2 nella formazione del ventricolo destro. Evidenze sperimentali dimostrano che la delezione cotemporanea del gene che codifica per Hand2 e del gene che codifica per NKx2.5, che si è visto, a sua volta, regolare l'espressione genica di Hand1, porta alla mancata formazione della metà ventricolare del cuore (Olson 2006).

Un altro regolatore trascrizionale di geni cardiaci è SRF (*serum response factor*), anch'esso appartenente alla classe delle MADS-box protein. SRF è un attivatore trascrizionale di geni muscolo-specifici ed induce il differenziamento e la crescita dei cardiomiociti in associazione con altri fattori di trascrizione, come NKx2.5 e GATA4.

Recenti studi dimostrano che la maggior parte dei geni codificanti per le proteine contrattili nelle cellule muscolari cardiache sono sotto il controllo di SRF. Poiché SRF è un debole attivatore trascrizionale, si ritiene che l'attivazione dei geni responsabili del differenziamento in senso muscolare avvenga tramite il reclutamento di un coattivatore trascrizionale: la miocardina (Wang et al, 2001), la quale promuoverebbe l'attivazione dei geni muscolo-specifici attraverso un meccanismo dipendente dalla presenza dei siti target di SRF. E' il reclutamento di SRF in corrispondenza delle sequenze consenso che richiamerebbe l'azione della miocardina. Quest'ultima sembra indurre il differenziamento dei cardiomiociti in seguito alla sua interazione con il motivo MADS di SRF (Parlakian et al, 2004).

SRF sembra regolare anche l'espressione di microRNA. In alcuni promotori di geni codificanti per microRNA, infatti, sono stati individuati siti target di SRF. Evidenze sperimentali suggeriscono che la delezione e/o il knock-out genico di SRF comportano una serie di difetti cardiaci tra cui la dilatazione delle camere cardiache e il difetto del setto ventricolare (Miano et al, 2004; Parlakian et al, 2004).

#### 1.3.5 MicroRNA e cuore

Numerose evidenze sperimentali suggeriscono il ruolo dei microRNA (miRNA o miR) nella cardiogenesi.

**Biogenesi dei miR** (Cullen, 2004): I microRNA sono piccoli RNA non codificanti a singolo filamento lunghi circa 20-22 nucleotidi (nt). Essi regolano l'espressione genica determinando un'inibizione traduzionale e/o trascrizionale. I miRNA, infatti, possono indurre un blocco della traduzione e/o trascrizione appaiandosi al loro target in corrispondenza di sequenze bersaglio localizzate nella 3'UTR.

I miRNA sono trascritti dall'RNA pol II e sono il prodotto di geni endogeni. A seconda della loro localizzazione genomica si distinguono:

- 1. miR derivanti da esoni o introni di geni codificanti
- 2. miR derivanti da cluster policistronici



Il processamento dei microRNA ha inizio nel nucleo quando il primiRNA (trascritto primario) viene Drosha processato da in collaborazione con la proteina DGCR8. Drosha, una RNAsi III, lega il pri-miRNA in corrispondenza del doppio filamento e produce un taglio sfalsato lasciando al 3' del prodotto

che si forma due nt sporgenti. Il prodotto che ne deriva, il pre-miRNA (precursore

del miR), viene esportato nel citoplasma ad opera dell'esportina 5 e del cofattore RAN-GTP. Il pre-miRNA, lungo circa 60-70 nt, ha una particolare struttura *stem* and *loop* bersaglio dell'enzima Dicer che produce un miR a doppio filamento. In seguito solo un filamento verrà selezionato ed il miR maturo a singolo filamento verrà incorporato nel complesso del RISC. Il miR può a questo punto legare in maniera o perfettamente complementare il proprio target in corrispondenza della 3'UTR ed indurre il blocco della traduzione, oppure in maniera non perfettamente complementare ed indurre la degradazione del messaggero.

**miR e regolazione dell'espressione dei geni cardiaci**: E' recente la correlazione tra i fattori di trascrizione coinvolti nella morfogenesi del cuore ed i microRNA. L'espressione di Hand2, infatti, sembra essere inibita da miR-1 (Olson 2006), a sua volta regolato da SRF (Zhao et al,2005).

miR-1-1 e miR-1-2, appartenenti alla sottofamiglia miR-1, sono entrambi regolati da SRF durante lo sviluppo cardiaco; i promotori dei corrispettivi geni, infatti, contengono le sequenze target riconosciute da SRF (Zhao et al,2005).

Evidenze sperimentali dimostrano che il knock-out genico di miR-1-2 determina un incremento del livello proteico di Hand2 che è causa, a sua volta, di uno dei principali difetti cardiaci, il difetto del setto ventricolare (Zhao et al,2007). L'esatto dosaggio proteico di Hand2 è essenziale per la proliferazione dei cardiomiociti e per la morfogenesi del cuore secondo un meccanismo miR dipendente.

Il knock-out genico di miR-1-2 sembra essere causa anche dei difetti elettrofisiologici. La sua assenza, infatti, provoca un incremento sia del livello di mRNA che del livello proteico di Irx5. L'incremento del livello proteico di Irx5 corrisponde ad un decremento del livello di mRNA di Kcnd2, suo target diretto.

Questi risultati dimostrano il ruolo di miR-1-2 nella morfogenesi e nella conduzione elettrica del cuore (Zhao et al,2007) e postulano un probabile ruolo dei miR nella regolazione trascrizionale oltre che traduzionale.

In condizioni d'ipertrofia cardiaca miR-1 e miR-133, appartenenti allo stesso cluster bicistronico e responsabili della proliferazione e differenziamento del muscolo cardiaco e scheletrico, risultano downregolati (Carè et al, 2007). La downregolazione di miR-1 potrebbe essere dovuta da un meccanismo SRF-dipendente, essendo il miR target diretto del suddetto fattore di trascrizione (Sayed et al, 2007).

SRF è a sua volta regolato negativamente da miR-133; la downregolazione di miR-133 incrementa il livello proteico di SRF, regolatore positivo della proliferazione cellulare (van Rooij et al, 2006). L'eccessivo aumento nel numero e nelle dimensioni dei cardiomiociti induce l'apoptosi e quindi una errata morfologia del cuore.

#### 1.4 Espressione genica e DS

Il meccanismo attraverso il quale una copia extra del cromosoma 21 determina il fenotipo DS è sconosciuto e probabilmente complesso.

E' stato ipotizzato che tre copie del cromosoma 21 causano un aumento di circa il 50% dell'espressione di geni trisomici come effetto primario del dosaggio genico (de la Luna et al, 2006)). La dimostrazione di tale effetto primario sull'espressione dei geni del cromosoma 21 è stata resa possibile grazie all'utilizzo della tecnologia dei microarrays e altre tecniche ad alto rendimento, che hanno permesso di misurare i livelli di RNA in tessuti e cellule umani (Mao et al, 2003; Mao et al, 2005) e in tessuti provenienti da modelli murini per la DS (Amano et al, 2004; Dauphinot et al, 2005). Dall'applicazione di queste tecnologie è stato stabilito che la trisomia causa un aumento del 50% dei livelli di RNA dei geni Hsa21 (de la Luna et al, 2006); tutto ciò potrebbe, almeno in parte, determinare delle perturbazioni in pathways e processi cellulari in cui questi geni funzionano e intaccare lo sviluppo cardiaco, causando quindi CHD (Mao et al, 2005). E', quindi, probabile che l'upregolazione dei geni Hsa21 influenzi l'espressione di altri geni, e che siano questi ad intaccare lo sviluppo cardiaco. L'analisi dei pathways ci permette di stabilire delle correlazioni genotipo/fenotipo nella DS, dato il numero così alto di geni coinvolti, il numero di varianti di splicing alternative e il numero di pathways in cui questi geni funzionano (Gardiner et al. 2004).

In questo lavoro è stato determinato il profilo di trascrizione di 15.000 geni, tra cui anche i geni Hsa21, nel cuore di feti umani a 18-22 settimane di gestazione, con e senza trisomia di Hsa21. Lo scopo è stato quello di capire come l'upregolazione di geni che mappano sul cromosoma 21 potesse influenzare l'espressione di geni che mappano su altri cromosomi e che siano potenzialmente coinvolti nei CHD.

### 1.4.1 Tecniche di valutazione dell'espressione genica: microarray e Rt-qPCR

#### 🖊 La tecnologia dei microarray

I microarrays sono la più recente di una serie di tecniche che sfruttano le caratteristiche peculiari della doppia elica degli acidi nucleici, ovvero la natura complementare delle due catene e la specificità dell'accoppiamento delle basi. L'enorme vantaggio di questa tecnica è che consente di effettuare analisi di espressione genica su scala genomica, ovvero di esaminare contemporaneamente i trascritti di molte migliaia di geni.

Il concetto di base dei microarrays è semplice: miniaturizzando le tecniche d'ibridazione migliaia d'analisi genetiche possono essere eseguite simultaneamente con una drastica riduzione dei volumi dei reagenti da utilizzare e dei tempi d'analisi.

Sequenze specifiche di DNA, chiamate *probes*, oligonucleotidi o cDNA, vengono scelte in quanto riconoscono selettivamente geni o regioni di interesse a cui sono perfettamente complementari e sono legate su substrati inerti a formare un chip.

L'mRNA da analizzare (*target*) viene marcato con un fluorocromo ed ibridato al chip in condizioni ottimali per l'appaiamento specifico delle basi del probe e del target.

Quando l'ibridazione è stata completata e i lavaggi eseguiti, i composti marcati legati alla sonda molecolare possono essere opportunamente rilevati mediante una scansione dei segnali dei fluorocromi con laser specifici per la frequenza emessa. Un software d'analisi di immagini misura la fluorescenza rilevata dallo scanner calcolando l'intensità di segnale per ogni probe e la traduce in dati numerici grezzi di espressione. Questi dati vengono ulteriormente elaborati con softwares bioinformatici specifici in grado di valutare le differenze fra gruppi di campioni analizzati e di attribuirvi un significato biologico.

**Materiali e supporti.** I chips per microarrays si differenziano per supporto utilizzato e per caratteristiche del DNA depositato sul supporto. I supporti utilizzati possono essere membrane di nylon, vetrini portaoggetti, silicon chips. Il

materiale depositato può essere cDNA (intero o frammenti) o oligonucleotidi di varia lunghezza.

In particolare la ditta Affymetrix (Santa Clara, CA) utilizza un silicon chip su cui deposita gruppi di oligonucleotidi da 25 basi con la tecnica fotolitografica. Poiché 25 basi sono poche per valutare il livello di espressione di un gene, occorre che vi siano più oligonucleotidi per ogni gene o trascritto che deve essere riconosciuto. Infatti per ogni trascritto vengono depositati 11 oligonucleotidi "Perfect Match" (PM) e 11 "Mismatch" (MM), cioè oligonucleotidi che hanno la stessa sequenza del PM eccetto che per una base e che funzionano come controllo interno. L'insieme di questi 22 oligonucleotidi è definito *probe set*. Data la tecnica di deposizione delle sonde, la presenza di un controllo interno ed il procedimento standardizzato di ibridazione, i chip Affymetrix sono considerati tra quelli che forniscono risultati estremamente riproducibili.

**Spotting e fotolitografia.** Esistono di fatto due tecnologie per la deposizione di sonde su microarrays: la prima denominata a *spotting* e la seconda detta *in situ*.

Nella tecnologia dello *spotting*, le sonde da ancorare al supporto solido, normalmente un vetrino da microscopia, sono sintetizzate a parte e quindi depositate sul supporto. Selezionate le sequenze da studiare, il cDNA relativo viene prodotto mediante PCR ottenendo così sonde della dimensione da 600 a 2.400 bps. Più recentemente, le sonde che vengono depositate sono rappresentate da sequenze sintetiche di oligonucleotidi lunghe 50-70 paia di basi.

La deposizione viene effettuata da sistemi robotizzati che mediante l'utilizzo di pennini prelevano le sonde direttamente dalle piastre utilizzate per la PCR e le depositano sul vetrino formando piccoli spots di dimensioni e distanza stabilite. Durante la deposizione, il sistema di controllo del robot registra automaticamente tutte le informazioni necessarie alla caratterizzazione ed all'identificazione di ciascun punto della matrice (identità del cDNA, coordinate sul supporto, ecc.). Una volta sul supporto, il probe viene legato covalentemente ai gruppi amminici del supporto attraverso una reazione innescata dall'irraggiamento con luce ultravioletta, mentre la sonda in eccesso viene rimossa con semplici lavaggi dell'array. Infine il probe legato al supporto viene reso a catena singola attraverso una denaturazione termica o chimica.

L'altra tecnica utilizzata per la produzione di microarrays è quella detta *in situ* che, sviluppata da Affymetrix, è frutto dell'interazione di due tecnologie

23

particolari, la fotolitografia e la sintesi diretta in fase solida di oligonucleotidi. La sintesi delle sonde avviene direttamente sulla superficie del supporto solido costituito da un wafer di silicio che viene ricoperto di piccole sequenze di oligonucleotidi (oligo-starter). Questi oligo hanno la caratteristica di avere il gruppo reattivo protetto da gruppi fotosensibili e quindi, grazie ad una maschera fotolitografica, è possibile indirizzare la luce in specifiche posizioni dell'array e liberare i siti necessari per la sintesi della sequenza. Una volta deprotetti selettivamente i siti reattivi, è sufficiente incubare la superficie con desossiribonucleotidi protetti per allungare la catena in fase di sintesi. Ripetendo il ciclo di deprotezione grazie all'applicazione di maschere fotolitografiche diverse e di incubazione è quindi possibile aggiungere nucleotidi solo dove la sequenza lo richiede e sintetizzare tutte le sonde necessarie per l'analisi di un dato genoma.



Tecnica fotolitografica – Affymetrix

#### 🖊 Analisi bioinformatica dei dati dei microarrays

Per la *quantificazione* del dato di ibridazione è stato utilizzato il software MicroArraySuite 5.0 (MAS 5.0). L'analisi eseguita da questo software valuta due punti: l'effettiva presenza/assenza del segnale e il relativo livello di espressione. L'algoritmo di *detection* utilizza il confronto fra l'intensità delle coppie di oligo (PM-MM) per definire il segnale come presente, assente o marginale (*flag* di presenza). L'algoritmo di segnale, inoltre, in relazione all'intensità di fluorescenza, dà una stima del livello relativo di espressione di un trascritto.

Per *l'interpretazione* dei dati (*data mining*) è stato utilizzato il software Gene Spring (Silicon Genetics, Redwood City; CA). I dati grezzi di espressione dei 22.283 probe sets presenti sul chip sono stati analizzati per evidenziare differenze significative di espressione genica tra tessuto cardiaco trisomico e non trisomico. L'espressione differenziale tra gruppi di campioni è stata valutata calcolando il *Fold Change* (FC), che è la *misura della differenza di espressione di un trascritto tra due diverse condizioni sperimentali*. La significatività statistica delle differenze tra i gruppi è stata valutata mediante il test ANOVA con p<0.05.

Per effettuare confronti tra i diversi chips, i data set su ogni chip sono stati scalati (normalizzati) rispetto ad una fluorescenza totale di 100. Prima, però, di effettuare il confronto tra chips diversi i dati grezzi sono stati pre-filtrati allo scopo di eliminare dall'analisi geni definiti "inaffidabili" come livello di espressione mediante il Cross Gene Error model, un metodo basato sul concetto che la variabilità tra i replicati è simile per tutti i geni che presentano livelli simili di espressione. Fra i geni ritenuti affidabili sono stati considerati solo quelli risultati presenti, ossia espressi in almeno 2 dei 15 campioni analizzati. Inoltre, il pre-filtraggio dei dati grezzi ha consentito l'esclusione dall'analisi di tutti quei geni che, avendo un'intensità di fluorescenza inferiore a 10, potessero confondersi con il background.

Per l'*analisi funzionale* è stato utilizzato il software GOTM (Zhang et al, 2004) che è in grado di individuare, all'interno di una lista di geni differenzialmente espressi, geni espressi nel contesto di Gene Ontology (GO), considerando come set di geni tutte le categorie di GO per processi biologici, funzione molecolare e componenti cellulari.

Il Consortium Gene Ontology è un progetto nato nel 1998 che cerca di unificare le informazioni contenute nei database di tre organismi modello: *Flybase* (Drosophila), *Saccharomyces Genome Database* (SGD) e Mouse Genome Database (MGD).

Il sequenziamento completo di molti genomi ha evidenziato che un elevato numero, associati alle principali funzioni biologiche (replicazione, trascrizione, metabolismo etc.), sono condivisi da tutti gli eucarioti. Pertanto la conoscenza del ruolo biologico delle proteine condivise in un organismo potrebbe essere trasferita ad altri organismi.

L'analisi dei pathways è stata effettuata con il software Pathway Miner (Pandey et al, 2004) che cataloga i geni di una lista in base al loro ruolo in pathways

metabolici, cellulari e regolatori presi da tre diversi databases di pathways (GenMap, Encarta e KEGG).

In base al numero di geni della lista appartenenti ad uno stesso pathway, si scelgono i pathways principali e si creano poi delle connessioni tra geni che partecipano a molteplici pathways.

Il Gene Set Enrichement Analysis (GSEA), infine, è stato utilizzato per identificare sets di geni correlati con la condizione della DS. L'analisi operata da questo software raggruppa i geni per funzioni biologiche, posizione sui cromosomi e meccanismi di regolazione (Subramanian et al, 2005).

#### 🖊 Validazione dei dati dei microarrays

Nonostante i microarrays siano una delle tecniche più sofisticate e all'avanguardia per l'analisi dell'espressione genica, molte domande però restano sull' attendibilità, sull'analisi e sull'interpretazione dei dati dei microarray.

I due tipi di errore che si devono prevenire in un'analisi di microarray sono: errori di misurazione ed errori di campionamento.

Gli errori di misurazione, ossia di valutazione dell'espressione, possono dar vita ad un solo rischio più immediato, quello di ottenere dei falsi positivi. Per ovviare a questo inconveniente i ricercatori sono soliti validare i propri risultati utilizzando la Real-Time Quantitative PCR (qRT-PCR), una tecnica quantitativa molto sensibile che permette di avere risultati più accurati. Sostanzialmente questo tipo di validazione è eseguito più per zelo da parte dei ricercatori che per reale utilità, in quanto un errore di questo tipo in una tecnica sperimentale come il microarray si presenterebbe solo come sistematico e non random, quindi dovuto a variabili esterne più che a variazioni interne all'esperimento.

L'errore di campionamento invece si manifesta sottoforma di dati di espressione genica discordanti per uno stesso trascritto all'interno di uno stesso gruppo ed è eliminabile solo eseguendo un nuovo campionamento e ritestando la stessa ipotesi con la stessa tecnica.

Per la validazione dei risultati dei microarray di solito si presume che ci sia una buona correlazione tra i dati dei microarray e i livelli di mRNA nella cellula investigata al momento dell'analisi, anche se questo assunto è talvolta contestato. Ad oggi, la qRT-PCR è il metodo di validazione maggiormente utilizzato, probabilmente perché è rapida e sensibile; inoltre, è una procedura ad alto rendimento che richiede minori quantità di materiale, se confrontata con altre tecniche di analisi dell'mRNA (Chuaqui et al, 2002).

#### 🖊 RT-qPCR

La RT-PCR permette la rilevazione e misurazione del prodotto amplificato in "real time", quindi durante la reazione di amplificazione. La rilevazione dei prodotti PCR è possibile aggiungendo ai reagenti una molecola fluorescente che mostra l'aumento di quantità del DNA con un aumento proporzionale del segnale di fluorescenza mentre avviene l'amplificazione. La fluorescenza emessa durante la reazione di amplificazione viene captata da un sistema di rilevazione che la trasmette poi ad un computer.

Esistono diverse modalità di rilevazione del prodotto fluorescente. Esse prevedono l'utilizzo di molecole particolari:

- 1. Intercalanti fluorescenti, es. SYBR Green
- 2. Probe fluorescenti, Hydrolysis probes (sistema Taqman)

Gli intercalanti fluorescenti legano DNA double strand ed emettono fluorescenza. Essi sono sostanze aspecifiche nel senso che legano qualsiasi molecola di DNA a doppio filamento. Il più semplice degli intercalanti è il SYBR Green, il quale, colpito da raggi UV, emette luce a differente lunghezza d'onda nel visibile. Il vantaggio di tale composto sta proprio nella sua aspecificità che consente di rivelare una notevole varietà di molecole di DNA. Tuttavia durante la reazione di PCR possono originarsi prodotti aspecifici di amplificazione o artefatti vari (come dimeri di primers); il legame del SYBR Green a tali molecole con la successiva emissione di fluorescenza provoca una falsificazione della quantizzazione. Per ovviare a questo problema la maggior parte degli strumenti che fanno RT-PCR hanno anche la possibilità di effettuare delle curve di melting o di fusione dell'amplificato le quali forniscono un'indicazione della purezza del prodotto di reazione e rivelano la presenza di dimeri dei primers. In questo modo è possibile distinguere chiaramente il prodotto specifico dalla presenza di dimeri di primers in base al numero e all'altezza di picchi presenti.

Le sonde di idrolisi, invece, vengono utilizzate nel cosiddetto "saggio Taqman". La sonda ha al 5' un fluoroforo e al 3' un quencer. Durante la fase di annealing i primers e la sonda ibridano al template e la fluorescenza emessa dal fluoroforo viene catturata dal quencer. Durante la fase di polimerizzazione i primer vengono estesi e la sonda spiazzata ed idrolizzata dall'attività esonucleasica 5'-3' della polimerasi. L'idrolisi libera il fluoroforo la cui fluorescenza può a questo punto essere direttamente misurata. Le sonde Taqman hanno una temperatura di annealing intorno ai 70° C, ciò vuol dire che anche durante la reazione di polimerizzazione esse rimangono appaiate al target da amplificare.

In una tipica curva di amplificazione, in cui la fluorescenza emessa è in funzione al numero di cicli, si susseguono diverse fasi:

• fase esponenziale: l'amplificazione del prodotto PCR prosegue in maniera esponenziale

 fase di transizione: vi è un incremento del prodotto di PCR ma non in maniera esponenziale

 fase di plateau: pur continuando la reazione di PCR l'amplificazione del prodotto si arresta, raggiunge cioè una fase di saturazione.



Curva di amplificazione

In una reazione di qRT-PCR è bene tener presente due concetti di notevole importanza: la linea soglia ed il ciclo soglia.

La linea soglia definisce quel valore di fluorescenza al di sopra del quale le misurazioni della fluorescenza sono attendibili, in quanto si discostano dal rumore di fondo. Al di sotto di questa linea la misurazione della fluorescenza non è attendibile perché si confonde con il background.

Il ciclo soglia ( $C_T$ ), invece, rappresenta quel ciclo di reazione in cui la fluorescenza emessa dal prodotto supera la linea soglia. Attraverso il ciclo soglia è

possibile calcolare la quantità iniziale di templato presente nel campione, ossia il numero di copie iniziali di templato. Ad un  $C_T$  basso corrisponde una quantità iniziale di templato piuttosto alta, maggiore è il numero di copie iniziale di templato minore sarà il numero di cicli necessari a raggiungere quel valore di fluorescenza che va al di sopra della linea soglia allontanandosi dal background; ad un  $C_T$  alto corrisponde, invece, una quantità iniziale di templato piuttosto bassa. Una reazione di qRT-PCR che abbia buon esito richiede un'amplificazione del prodotto efficiente e specifica.

Per capire se la nostra qRT-PCR è ottimizzata, è necessario correre diluizioni seriali di un templato a concentrazione nota e utilizzare i risultati per generare delle Curve Standard. La curva standard si ottiene plottando il log della quantità di partenza del templato contro il valore di  $C_T$  ottenuto durante l'amplificazione di ciascuna diluizione.

L'equazione della retta di regressione lineare, insieme al coefficiente di correlazione di Pearson (r) o il coefficiente di determinazione ( $R^2$ ), possono essere usati per capire se l'analisi con qRT-PCR è ottimizzata.

I punti focali di una qRT- PCR ottimizzata sono:

- Curva Standard lineare (R<sup>2</sup>>0.980 o r>l-0.990l)
- Alta efficienza di amplificazione (90-150%)
- Pendenza della curva (- 3.32)
- Concordanza tra i replicati della reazione.

**Scelta dei primers**. Un'efficiente amplificazione del prodotto può essere influenzata sia dai primers scelti sia dalla regione target da amplificare (amplicone). E' per questo motivo che bisogna prestare molta attenzione quando si sceglie una sequenza target e si disegnano i primers.

Sono disponibili numerosi softwares gratuiti on line che permettono di disegnare i propri pimers nelle migliori condizioni possibili ed altri che permettono di verificare eventuali ripiegamenti che andrebbero ad inibire la reazione.

Per la scelta dell'amplicone, ci sono alcune linee guida da seguire:

 Per avere una buona efficienza di amplificazione il range di lunghezza dell'amplicone deve essere 75-200 bp. Un amplicone dovrebbe essere almeno di 75 bp per poterlo distinguere facilmente da possibili dimeri di primers; • Evitare strutture secondarie se possibile; utilizzare programmi come mfold per predire se un amplicone formerà strutture secondarie alla temperatura di annealing;

- Evitare templati con lunghe repeats (>4) di singole basi;
- Mantenere un contenuto di GC del 50-60%.

Per il disegno dei primers, le linee guida sono:

- Disegnare primers con un contenuto di GC del 50-60%;
- Mantenere la temperatura di melting (Tm) tra 50°C e 65°C;
- Evitare strutture secondarie;
- Evitare repeats di G o di C più lunghe di 3 basi;
- Posizionare le G e le C alla fine dei primers;
- Controllare la sequenza del forward e del reverse per accertarsi che non ci sia complementarietà al 3';
- Verificare la specificità dei primers usando strumenti come BLAST.

**Metodi di quantizzazione**. La quantificazione mediante qRT-PCR può essere assoluta o relativa. Con la quantificazione assoluta si può determinare esattamente la quantità iniziale di templato all'interno del campione test. Ciò richiede la costruzione di una curva standard in cui il valore di  $C_T$  è inversamente proporzionale al log del numero iniziale di copie dei campioni standard derivanti dalle diluizioni seriali di un dato campione di riferimento a concentrazione nota.

Il metodo della quantificazione relativa è più semplice poiché richiede la quantificazione di geni di controllo (calibratore) e geni housekeeping (reference) per normalizzare l'espressione del gene studiato. Infatti il  $C_T$  del bersaglio viene confrontato direttamente con il  $C_T$  del calibratore allo scopo di determinare cambiamenti di espressione del gene bersaglio.

Dopo che sono stati misurati i valori  $C_T$ , possono essere usati diversi metodi per determinare il livello di espressione del gene target nel campione test rispetto al calibratore: 1) il metodo Livak, anche conosciuto come metodo del 2<sup>- $\Delta\Delta$ Ct</sup>; 2) il metodo del  $\Delta C_T$  usando un gene reference e 3) il metodo PfaffI.

La quantizzazione relativa necessita di uno o più geni reference (come GAPDH, RPL13A...) che abbiano una costante espressione in tutti i campioni testati e la cui espressione non sia cambiata in seguito al trattamento effettuato sui campioni, soprattutto quando il materiale di partenza è limitato. Recentemente è stato

proposto che in molti casi, l'uso di più geni reference potrebbe essere necessario per una quantizzazione accurata (Vandesompele et al, 2002).

**RT-qPCR per la quantizzazione dei microRNA**. E' piuttosto recente la messa a punto di RT-qPCR per saggiare il livello di espressione di microRNA, essendo essi di piccole dimensioni.

Uno dei kit per la quantizzazione dei microRNA è il miScript System (Qiagen) che permette, in un singolo step, la trascrizione inversa di mRNA, microRNA ed altri piccoli RNA non coding.

## <u>SCOPO DEL LAVORO</u>

Circa la metà dei feti affetti dalla DS sviluppano la cardiopatia. Si ritiene che i geni che mappano sul cromosoma 21 umano possano essere coinvolti nella morfogenesi del cuore e che la loro espressione anomala negli individui trisomici possa alterare il normale processo di sviluppo del cuore. Il fenotipo Down costituisce quindi un buon modello per la comprensione della morfogenesi del cuore in termini di eventi molecolari e per la patogenesi delle cardiopatie.

Lo scopo del progetto è stato comprendere come l'upregolazione dei geni Hsa21 in cuori trisomici possa influenzare la morfogenesi del cuore in maniera diretta (effetto primario del dosaggio genico), o attraverso l'alterazione dell'espressione di geni che mappano su altri cromosomi ed in particolare di fattori di trascrizione coinvolti nella cardiogenesi (effetto secondario del dosaggio genico).

Al fine di individuare geni candidati per lo sviluppo della cardiopatia, abbiamo intrapreso lo studio del profilo di espressione genica nel cuore fetale umano di feti affetti da DS a 18-22 settimane di gestazione.

Abbiamo ipotizzato che in condizioni di trisomia 21, vi è una disregolazione dei geni coinvolti nella cardiogenesi nel tessuto cardiaco di feti con DS.

Per prima cosa è stato determinato il profilo di espressione genica nel cuore fetale umano e valutato l'overespressione dei geni Hsa21 nei cuori di feti affetti da DS rispetto ai cuori euploidi mediante microarray.

Successivamente abbiamo determinato i geni differenzialmente espressi tra Down/non Down e ho eseguito su di essi un'analisi funzionale allo scopo di catalogarli nel contesto di Gene Ontology (GO) e di comprendere eventuali correlazioni tra essi, analizzando i pathways in cui tali geni funzionano.

Infine per validare i dati ottenuti con i microarray, i livelli di espressione dei geni di maggiore interesse, compresi quelli coinvolti nello sviluppo del cuore, sono stati confermati attraverso la qRT-PCR.

Poiché i nostri risultati dimostrano una globale disregolazione dei geni Hsa21 e di geni localizzati su altri cromosomi nel cuore di feti in condizioni di trisomia 21, abbiamo avviato un ulteriore approccio alla comprensione dei meccanismi molecolari responsabili della disregolazione dell'espressione genica conseguente a trisomia. Abbiamo ipotizzato che geni di microRNA presenti sul cromosoma 21 possano essere upregolati nel cuore di feti trisomici e che questa upregolazione possa essere a sua volta la causa della disregolazione di geni, target dei microRNA e coinvolti nella cardiogenesi. La verifica di questa ipotesi, che è attualmente in corso, prevede la quantizzazione dell'espressione dei microRNA del cromosoma 21, la ricerca di potenziali target e la verifica degli stessi. Stiamo inoltre verificando l'ipotesi collaterale che microRNA, non del cromosoma 21, possano essere disregolati da geni che risultano upregolati nella sindrome di Down. Questi microRNA potrebbero a loro volta modificare l'espressione di geni convolti nella morfogenesi cardiaca.

# <u>MATERIALI E METODI</u>

## 🖊 Campioni

Sono stati analizzati 15 campioni di cuore fetale umano. Il tessuto cardiaco è stato prelevato durante l'analisi anatomapatologica di feti dopo aborto terapeutico e immediatamente conservato a -135° C per la successiva fase di estrazione dell'RNA.

I 15 cuori fetali raccolti sono stati distinti in due gruppi: il gruppo NH, costituito da 5 cuori euploidi, senza difetti cardiaci, e il gruppo DSH, costituito a sua volta da due categorie di cuori trisomici, 5 con la cardiopatia (CDH) e 5 senza cardiopatia (DH). I dati dei campioni analizzati sono indicati nella **tab 1**.

I campioni sono stati conservati nella *Banca di tessuti, cellule e DNA derivati da feti malformati* presso il servizio di Citogenetica e Diagnosi Prenatale. Nella banca informatica ad essa associata i campioni sono stati catalogati con un numero di codice e le informazioni riguardanti i familiari trattate in maniera riservata.

Il prelievo, la conservazione e l'utilizzazione di tessuto cardiaco fetale sono subordinati al consenso espresso dai genitori all'atto della richiesta dell'esame autoptico. La formula di consenso prevede l'informazione sulle modalità di conservazione e sugli scopi scientifici e diagnostici di utilizzazione dei tessuti fetali e la scelta da parte dei genitori di essere informati o meno sui risultati scientifici prodotti.

ID Banca	ID	Cariotipo	Settimana	Intervallo	Difetti
Telethon	Esperimento		di	post-	cardiaci
			gestazione	mortem	
				(ore)	
GRUPPO					
NH					
TB 26	H1	46,XY	20	3	NO
TB 21	H2	46,XY	21	3	NO
TB 30	H3	46,XY	20	1	NO
TB 32	H4	46,XX	21	1	NO
TB 23	H5	46,XX	21	4	NO
GRUPPO					
DSH					
TB 37	DH1	47,XY+21	21	2	NO
TB 43	DH3	47,XY+21	20	3	NO
TB 50	DH4	47,XY+21	20	3	NO
TB 47	DH5	47,XY+21	19	4	NO
TB 48	DH6	47,XY+21	21	3	NO
TB 9	CDH1	47,XY+21	22	2	SI
TB 33	CDH2	47,XY+21	20	4	SI
TB 55	CDH4	47,XX+21	21	2	SI
TB 22	CDH5	47,XX+21	18	3	SI
TB 57	CDH6	47,XX+21	20	3	SI

#### Tabella 1: Caratteristiche dei campioni analizzati nello studio

DSH= Campioni di cuori da feti con trisomia Hsa21

NH= Campioni di cuori da feti normali

Difetti cardiaci= TB9 Tetralogia di Fallot, TB33 VSD, TB55 AVSD, TB22 e TB57 Tetralogia di Fallot.

### 🔸 🛛 Sintesi e biotinilazione del cRNA

L'RNA totale da tessuto cardiaco è stato estratto con TRIzol (Gibco\BRL Life Technologies, Inc., Gaithersburg, MD) e usato per preparare il cRNA *target* biotinilato, secondo le raccomandazioni Affymetrix riportate nel sito:

http://www.affymetrix.com/support/technical/manual/expression manual.affx.

La purificazione del PolyA+mRNA dall'RNA totale è stata fatta usando l'Oligotex mRNA Kit (Qiagen).

Il primo filamento di cDNA è stato sintetizzato a partire da 1 µg di RNA con una RNA polimerasi T7 legata ad un primer oligo-dT; la sintesi del secondo filamento, invece, è stata condotta in vitro con UTP e CTP biotinilati utilizzando l' Enzo BioArray High Yeld RNA Transcript Labelling Kit (Affymetrix).

Il cDNA è stato poi convertito in cRNA che, a sua volta, è stato marcato con biotina mediante l'uso dell'Enzo Kit.

In seguito, mezza aliquota di reazione è stata purificata attraverso le colonnine RNeasy spin columns (Qiagen), l'altra metà è stata conservata a -80°C. Il cRNA purificato è stato quantizzato allo spettrofotometro e digerito mediante una soluzione di frammentazione allo scopo di ottenere *targets* di 35-200 bps.

#### 📥 Ibridazione

I cRNA marcati sono stati ibridati sul chip Affymetrix HG-U133A (Human Genome U133 Array), che contiene 22.283 probe sets corrispondenti a circa 15.000 geni umani ben caratterizzati.

Dopo ibridazione overnight, gli arrays sono stati sottoposti a lavaggio e colorazione con streptavidina-phycoeritrina prima di essere scannerizzati con il GeneChip Laser Confocal Fluorescent Scanner (Hawlett Packard).

#### Disegno dell'amplicone e dei primers per la RT-PCR

Una volta scelto il gene target da amplificare, è stata presa la sequenza del trascritto con tutte le informazioni sulla dimensione e la disposizione di esoni e introni dal sito di Ensembl (<u>http://www.ensembl.org/Homo\_sapiens/index.html</u>). E' stata scelta una regione a cavallo di due esoni, preferibilmente separati da un introne più grande, di circa 2 Kb, per evitare di amplificare eventuale DNA gnomico. Tale regione rappresenta l'amplicone, cioè la regione del gene da amplificare.

La sequenza scelta è stata inserita nel programma per il disegno dei primers Primer3 (<u>http://frodo.wi.mit.edu/cgi-bin/primer3/primer3 www slow.cgi</u>) impostando i seguenti parametri:

- product size: min 80; opt 130; max 150
- primer size: min 18; opt 20; max 25
- primer T<sub>m</sub>: min 59,5; opt 60; max 60,5
- primer GC%: min 20; max 80
- max 3' self complementary: 3
- max self complementary: 4

Tutti gli altri parametri sono stati lasciati di default.

Una volta cliccato su Pick Primers il programma ha fornito la sequenza dei primers forward e reverse e la regione dell'amplicone alla quale si legano per complementarietà.

Il passo successivo è consistito nell' effettaure un BLAST (http://www.ncbi.nlm.nih.gov/genome/seq/BlastGen/BlastGen.cgi?taxid=9606) per verificare che i primers scelti non si trovino sullo stesso esone o che riconoscano più geni appartenenti alla stessa famiglia. Se così fosse è necessario scegliere un'altra regione del gene da amplificare.

Un altro passaggio fondamentale è stato quello di verificare che il primer non dimerizzi con se stesso o con l'altro primer, altrimenti la reazione sarebbe inibita. Sul sito dell'Oligo Analyzer (<u>http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/</u>) è stato possibile controllare se i primers formano hairpin, self-dimeri o etero-dimeri. A tale proposito è stata valutata la  $T_m$  delle strutture eventuali che i primers possono formare, considerando che la temperatura alla quale avviene l'amplificazione è  $60^{\circ}$ C.

Un altro parametro che è stato valutato è la variazione di energia libera delle strutture, cioè il  $\Delta G$  deve essere compreso tra 0 e -6, valori che indicano che la reazione di formazione delle strutture è sfavorevole, e il numero di basi appaiate all'interno della sequenza non deve essere superiore a 2-3.

E' stato poi verificato che l'amplicone non formi hairpin che possano inibire la reazione di elongazione e questo è stato fatto mediante il software mFold (http://www.bioinfo.rpi.edu/applications/mfold/dna/form1.cgi). La sequenza dell' amplicone è stata inserita nel formato della pagina iniziale e sono stati impostati i seguenti parametri: temperatura di folding 60°C e concentrazioni ioniche (espresse in mM) di [Na<sup>+</sup>]=50 e [Mg<sup>++</sup>]=3. Il programma ci ha dato come risultato le immagini delle possibili strutture secondarie che l'amplicone può formare, ma anche in questo caso sono stati valutati la valutando la T<sub>m</sub> e il  $\Delta$ G. Sono stati scelti ampliconi con una T<sub>m</sub> di 60° C e con valori di  $\Delta$ G compresi tra 0 e -6.

I primer disegnati sono stati forniti dall'MWG Biotech, Ebersberg, Germany.
### 🖊 🛛 Diluizione dei primers

I primers liofilizzati sono stati sottoposti ad una rapida centrifugazione, dopodichè nei tubini è stata aggiunta la quantità di  $H_2O$  indicata sul protocollo allegato per una concentrazione finale di 100 pmol/µl. A questo punto è stata eseguita una diluizione 1:10 per ottenere i primers ad una concentrazione di 10 pmol/µl.

### 👃 🛛 Sintesi di cDNA

Per la reazione di sintesi del cDNA è stato utilizzato il kit Biorad "iScript<sup>TM</sup> cDNA Synthesis Kit". La reazione è stata svolta in un volume di 20  $\mu$ l per 1  $\mu$ g di RNA.

Componenti della reazione	Volume per reazione
5x iScript Reaction Mix	4 µl
iScript Reverse Transcriptase	1 µl
Nucleare-free water	xμl
RNA template (1µg)	xμl
Volume totale	20 µl

La reazione è stata condotta utilizzando l'apparecchiatura

con il seguente protocollo di cicli:

- 5' a 25°C
- 30' a 42°C
- 5' a 85°C

### 4

### 🖊 Curve standard e RT-qPCR

Prima di procedere all'indagine con la RT-PCR, è stato necessario testare l'efficienza dei primers. A tale scopo sono state costruite curve standard a partire da diluizioni seriali di cDNA di cuori fetali selezionati, che esprimono i geni target ad un livello appropriato.

I reagenti della reazione sono:

- iQ Supermix SYBR Green 2X (www.bio-rad.com);
- Primer forward e reverse;
- Campione di cDNA templato retrotrascritto a partire da RNA totale umano.

Lo stock di partenza di cDNA di 40  $\mu$ l è stato diluito con diluizioni seriali 1:2 per ottenere altri 4 Standard, ciascuno di 40  $\mu$ l.

Il volume unitario dei reagenti della miscela è:

- iQ Supermix Sybr 2X  $12,5 \mu l$
- Primer For 2,5  $\mu$ M 1  $\mu$ l
- Primer Rev 2,5  $\mu$ M 1  $\mu$ l
- H<sub>2</sub>O 8,5 μl

per un volume totale della miscela di 23  $\mu$ l alla quale sono stati aggiunti 2  $\mu$ l di campione.

La reazione di amplificazione per ciascuno standard è stata effettuata in duplicato, dispensando 25  $\mu$ l della mix in ciascun pozzetto della piastra da 96 wells.

Lo stesso protocollo di reazione è stato seguito per la RT-PCR in triplice copia per ciascun gene, scegliendo, per le prime reazioni di amplificazione, più geni housekeeping come GAPDH, RPL13A,  $\beta$ -actina, TBP. In seguito la  $\beta$ -actina è stata scartata perché si esprime in maniera variabile in cuore, pertanto gli esperimenti sono stati condotti mantenendo come reference solo GAPDH e RPL13A.

Le reazioni di amplificazione sono state effettuate in un volume finale di 25  $\mu$ l, con lo strumento iCycler Bio-Rad, selezionando il profilo termico PPAgamma che è il seguente:

Cycle	Repeats	Step	Time	Set-Point
1	1	1	3:00	95°
2	40	1	00:15	95°
		2	0:30	60°
3	1	1	1:00	55°
4	80	1	0:10	55°

Melt curve +0,5

### 븆 Analisi dell'espressione genica relativa

La quantizzazione relativa necessita di uno o più geni reference (come GAPDH, RPL13A...) che abbiano una costante espressione in tutti i campioni testati e la cui espressione non sia cambiata in seguito al trattamento effettuato sui campioni, soprattutto quando il materiale di partenza è limitato. Recentemente è stato proposto che in molti casi, l'uso di più geni reference potrebbe essere necessario per una quantizzazione accurata (Vandesompele et al, 2002).

Quando si confrontano più campioni nella quantizzazione relativa, uno dei campioni è generalmente scelto come "calibratore" e l'espressione del gene target in tutti gli altri campioni è espressa come un aumento o una diminuzione rispetto a quella che esso ha nel calibratore. Nel nostro caso, i campioni euploidi sono stati scelti come calibratori.

Per determinare l'espressione relativa di un gene target nel campione test (nel nostro caso i trisomici) e nel calibratore usando un gene reference come normalizzatore, è necessario determinare i valori dei  $C_T$ :

- i C<sub>T</sub> del target nel test e nel calibratore;
- i C<sub>T</sub> del reference nel test e nel calibratore.

Dopo che sono stati misurati i valori  $C_T$ , possono essere usati diversi metodi per determinare il livello di espressione del gene target nel campione test rispetto al calibratore: 1) il metodo Livak, anche conosciuto come metodo del 2<sup>- $\Delta\Delta$ Ct</sup>; 2) il metodo del  $\Delta$ C<sub>T</sub> usando un gene reference e 3) il metodo PfaffI.

Noi abbiamo scelto il metodo Livak. Questo metodo assume che sia il gene target che il reference siano amplificati con un'efficienza quasi del 100% e con una differenza del 5% tra l'uno e l'altro.

Una volta stabilita l'efficienza di amplificazione, è stata determinata la differenza relativa nel livello di espressione del gene target con i seguenti steps:

1) Normalizzare il  $C_T$  del gene target con quello del reference (ref), sia per il test che per il calibratore:

 $\Delta C_{T(test)} = C_{T(target, test)} - C_{T(ref, test)}$ 

 $\Delta C_{T(calibratore)} = C_{T(target, calibratore)} - C_{T(ref, calibratore)}$ 

- 2) Normalizzare il  $\Delta C_T$  del campione test con il  $\Delta C_T$  del calibratore:  $\Delta \Delta C_T = \Delta C_{T(test)} - \Delta C_{T(calibratore)}$
- 3) Infine calcolare il rapporto di espressione:

### $2^{-\Delta\Delta Ct}$ = Rapporto di espressione normalizzato

Il risultato ottenuto è l'aumento o la diminuzione del livello di espressione del gene target nel campione test rispetto al calibratore ed è normalizzato all'espressione del gene reference.

Normalizzando l'espressione del gene target a quella del reference, si compensa qualsiasi differenza nella quantità del tessuto campione di partenza.

### **w** miScript System: RT-qPCR per la quantizzazione dei microRNA

miScript System (QIAGEN) è un sistema a tre componenti. Esso comprende:

- miScript Reverse Transcription Kit
- miScript Primer Assay
- miScript SYBR Green PCR Kit

Il miScript Reverse Transcription Kit comprende la miScript Reverse Transcriptase Mix and miScript RT Buffer. La miScript Reverse Transcriptase Mix contiene una poly(A) polimerasi ed una trascrittasi inversa. Il miScript RT Buffer, invece, contiene Mg<sup>2+</sup>, dNTPs, oligo-dT primers e random primers. Il prodotto di amplificazione viene poi rilevato con il SYBR Green (miScript SYBR Green PCR Kit).

La reazione di sintesi del cDNA è stata svolta in un volume di 20  $\mu$ l per 1  $\mu$ g di RNA.

Reagenti	Volume della	Concentrazion
	reazione	e finale
miScript RT Buffer 5x	4 µl	1x
miScript Reverse Transcriptase Mix	1 µl	
RNase-free water	Variabile	
Template RNA	Variabile	
Volume totale	20 µl	

I campioni sono stati incubati in un termociclizzatore e sottoposti ai seguenti trattamenti:

- 60 minuti a 37° C
- 5 minuti a 95° C per l'inattivazione della miScript Reverse Transcriptase Mix

I cDNA sintetizzati sono stati dapprima diluiti 1:4 e poi utilizzati come stampo per la RT-qPCR.

La reazione di RT-qPCR è stata eseguita in un volume di 10 µl

Reagenti	Volume della reazione (384-well)	Concentrazione finale
2x QuantiTect SYBR	5 µl	1x
Green PCR Master Mix		
10x miScript Universal	1 µl	1x
Primer		
10x miScript Primer Assay	1 µl	1x
RNase-free water	1 µl	
Template cDNA (25 ng)	2 µl	≤ 100 ng per reazione
Volume totale	10 µl	

e con il seguente protocollo di cicli:

Step	Time	Temperatura	Commenti addizionali
Step iniziale di PCR	15 min	95°C	Attivazione dell'HotStarTaq DNA polimerasi
Denaturazione	15 sec	94°C	
Annealing	30 sec	55°C	
Polimerizzazione	30 sec	70°C	Misurazione della fluorescenza
N° di cicli	35-40 cicli		Il numero di cicli dipende dalla quantità di cDNA template e dall'abbondanza del target.

Il livello di espressione dei miR è stato valutato con il metodo Livak, già descritto precedentemente.

## <u>RISULTATI</u>

Allo scopo di determinare il profilo di espressione genica nel tessuto cardiaco fetale umano e di individuare geni differenzialmente espressi tra Down/non Down è stato utilizzato il chip Affymetrix HG-U133A (Human Genome U133 Array).

Attraverso la valutazione del numero delle "presence calls" (flag di presenza) eseguita dal software MicroArraySuite (MAS 5.0) sono stati individuati 7.500 geni espressi nel cuore fetale umano a 18-22 settimane di gestazione.

Il chip Affymetrix HG-U133A include 262 probe sets che corrispondono a 168 geni noti che mappano sul cromosoma 21. Gli 81 probe sets, corrispondenti a 63 geni individuali, sono presenti in tutti i campioni analizzati. Il software MAS 5.0 sottostima il flag di presenza, infatti se si considerano i geni presenti in almeno 2/3 dei campioni, si può ritenere che 117 probe sets, rappresentativi di 87 geni Hsa21, sono espressi nel cuore umano a 18-22 settimane di gestazione.

Questi geni sono globalmente upregolati nei campioni trisomici, infatti, il rapporto tra la somma delle medie dei dati di espressione genica nei campioni con DS (gruppo DSH) e la media dei valori di espressione genica nei campioni euploidi (gruppo NH) è  $1,48 \pm 0,35$  per i geni Hsa21 e circa 1 per tutti gli altri cromosomi (**fig. 11**). Questa differenza è altamente significativa (p<0.0001; test ANOVA con Bonferroni post-hoc).



**Fig. 11** Rapporto DSH/NH dell'espressione genica calcolato per ciascun cromosoma e per tutti i cromosomi.

Tuttavia, il livello di variazione di espressione dei geni Hsa21 non è molto alto, con un FC che è compreso tra 1.2 a 3. Alcuni geni Hsa21 sono normoregolati nel confronto Down vs non Down.

Per identificare i geni differenzialmente espressi, sono stati confrontati i livelli di espressione genica nei feti DSH vs NH.

Un totale di 473 geni risultano significativamente disregolati: 278 geni risultano downregolati, mentre 195 upregolati nel gruppo DSH vs il gruppo NH (**fig. 12**). Tra i geni upregolati 32 sono localizzati sull' Hsa21, e 441 geni disregolati si trovano su cromosomi differenti.



Fig.12: *Volcano plot* dei geni differenzialmente espressi tra campioni trisomici e controlli.

Il log2 del FC tra campioni trisomici e controlli è rappresentato sull'asse delle x mentre il log negativo del p-value dal t-test è rappresentato sull'asse delle y. I geni upregolati nei trisomici si trovano a destra del valore 0 dell'asse orizzontale, i geni downregolati a sinistra. I puntini rossi indicano i 473 geni significativamente up- o downregolati nei trisomici rispetto ai controlli (p<0.05). I puntini gialli indicano geni senza variazione significativa.

### > Analisi funzionale dei geni differenzialmente espressi

Poiché i geni differenzialmente espressi tra Down/non Down e potenzialmente interessanti per lo sviluppo della cardiopatia sono un numero troppo elevato per poterli analizzare singolarmente, è stato necessario utilizzare metodi di analisi che raggruppino i geni per categorie funzionali.

La lista dei geni differenzialmente espressi tra DSH e NH, con un FC>1.2 e p<0.05, è stata sottoposta ad analisi funzionale sfruttando i softwares GOTM e Pathway Miner.

In base alle categorie funzionali di Gene Ontology, la lista dei geni differenzialmente espressi è stata confrontata con il set completo di geni presenti sul chip Affymetrix HG-U133A, scelti come lista di riferimento allo scopo di identificare categorie più rappresentate nella lista dei geni differenzialmente espressi che nel set di geni di riferimento.

Sono state così identificate 44 categorie arricchite per i geni downregolati (**tab. 2**) e 19 categorie arricchite per i geni upregolati (**tab. 3**).

GO Category	Observed	Expected	Ratio	p-value
Biological Process				
gener. of precursor metabolites and				
energy	30	9.78	3.07	3.31E-08
oxidative phosphorylation	9	1.28	7.03	4.43E-06
ATP synth. Coupled electron transp.	6	0.49	12.24	6.55E-06
coenzyme metabolism	12	2.65	4.53	1.26E-05
mitochondrial electron transport	5	0.41	12.2	3.84E-05
acetyl-CoA metabolism	5	0.56	8.93	0.000193
electron transport	15	5.51	2.72	0.000408
main pathways of carbohydrate				
metabolism	8	1.92	4.17	0.00065
glucose catabolism	6	1.15	5.22	0.000966
monosaccharide catabolism	6	1.32	4.55	0.001998
hexose catabolism	6	1.32	4.55	0.001998
energy derivation by oxid. of organic				
comp.	9	2.84	3.17	0.00214
alcohol catabolism	6	1.35	4.44	0.00217
Glycolysis	5	0.94	5.32	0.002329
porphyrin biosynthesis	3	0.28	10.71	0.002341
porphyrin metabolism	3	0.34	8.82	0.004372
group transfer coenzyme metabolism	5	1.11	4.5	0.004859
cellular carbohydrate catabolism	6	1.73	3.47	0.00756
carbohydrate catabolism	6	1.73	3.47	0.00756
glucose metabolism	6	1.77	3.39	0.008491
Molecular Function				
electron carrier activity	9	1.14	7.89	1.7E-06
NADH dehydrogenase activity	7	0.69	10.14	4.11E-06
hydrogen ion transporter activity	12	2.39	5.02	4.48E-06
oxidoreductase activity acting on				
NADH or NADPH	8	1.02	7.84	6.63E-06
sodium ion transporter activity	7	0.75	9.33	7.61E-06
primary active transporter activity	13	3	4.33	8.97E-06
monovalent inorganic cation transporter				
activity	12	2.58	4.65	9.83 <u>E-0</u> 6
metal ion transporter activity	9	1.56	5.77	2.34E-05

23	10.05	2.29	0.000172
12			
14	4.37	2.75	0.001446
16	6.91	2.32	0.001523
4	0.52	7.69	0.001633
4	0.52	7.69	0.001633
8	2.62	3.05	0.004727
48	10.2	4.71	3E-20
10	2.45	4.08	0.000163
8	1.61	4.97	0.000194
9	2.07	4.35	0.000217
13	4.21	3.09	0.000296
13	4.25	3.06	0.000326
8	1.8	4.44	0.000417
19	8.37	2.27	0.000692
19	8.37	2.27	0.000692
6	1.82	3.3	0.009655
	12 16 4 4 8 	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$

**Tab. 2**: Categorie di GO arricchite per i geni downregolati, ordinate per p-value. Sono riportati i rapporti tra i geni osservati e quelli attesi per le categorie di Gene Ontology più rappresentate nella lista di geni disregolati.

GO categories	Observed	Expected	Ratio	p-value	
Biological process					
phosphate transport	8	0.98	8.16	5.27E-06	
cell adhesion	22	8.15	2.7	1.81E-05	
anion transport	9	2.15	4.19	0.000297	
inorganic anion transport	8	1.77	4.52	0.000378	
axonogenesis	5	0.66	7.58	0.000471	
neuron morphogen. during different.	5	0.74	6.76	0.000807	
neurite morphogenesis	5	0.74	6.76	0.000807	
neuron development	5	0.98	5.1	0.002889	
cell development	6	1.45	4.14	0.003197	
neuron differentiation	5	1.14	4.39	0.005565	
neurogenesis	5	1.19	4.2	0.006627	
Molecular Function					
extracellular matrix struct. const.	10	1.25	8	4.11E-07	
integrin binding	4	0.43	9.3	0.000833	
copper ion binding	5	0.82	6.1	0.00128	
Cellular Component					
extracellular matrix	22	22 3.55		5E-12	
collagen	8	0.5	16	2.08E-08	
extracellular region	35	14.17	2.47	2.83E-07	
fibrillar collagen	4	0.16	25	1.24E-05	
actin cytoskeleton	8	2.78	2.88	0.006602	

**Tab. 3:** Categorie di GO arricchite per i geni upregolati, ordinate per p-value. Sono riportati i rapporti tra i geni osservati e quelli attesi per le categorie di Gene Ontology più rappresentate nella lista di geni disregolati.

Le categorie delle componenti cellulari di GO risultate più arricchite sono quelle legate ai mitocondri, per i geni downregolati, e quelle della matrice extracellulare (ECM), per i geni upregolati. Almeno 65 geni codificanti per enzimi mitocondriali risultano downregolati nel tessuto cardiaco DS, e 40 geni codificanti per proteine ECM sono upregolati nei campioni DSH.

L'analisi dei pathways con il software Pathway Miner ha rivelato una globale downregolazione dei geni mitocondriali, in particolare di quelli che codificano per gli enzimi della fosforilazione ossidativa, ed una iperespressione di geni che codificano per molecole della matrice extracellulare per la maggior parte appartenenti al pathway delle adesioni focali (**fig. 13**).

Per un'ulteriore conferma dei dati ottenuti da GOTM e Pathway Miner, è stato utilizzato il software Gene Set Enrichment Analysis (GSEA). Dall'analisi con GSEA, effettuata allo scopo di identificare sets di geni che correlano con la condizione DS, sono stati ottenuti 5 sets di geni biologicamente informativi: tra questi ci sono due sets di geni la cui upregolazione è altamente correlata con DS (ECM e Adesione Cellulare) e tre sets di geni la cui downregolazione è altamente correlata con DS (Mitocondrio, Catena di Trasporto degli Elettroni e OXPHOS)(**fig. 14**). Quindi i risultati ottenuti con GSEA sono in accordo con l'analisi funzionale eseguita con GOTM e Pathway Miner.



### Fig. 13: Geni e pathway genici affetti dalla trisomia Hsa21.

L'analisi dei pathways è stata effettuata con Pathway Miner sui 473 geni disregolati nei campioni trisomici. Il pathway della Fosforilazione ossidativa, globalmente downregolato, è rappresentato dal Cluster 1 in verde; il pathway "focal adhesion" ed ECM, globalmente upregolato, è rappresentato dal Cluster 2 in rosso; il Cluster 3 indica un network di geni del pathway "Cell Adhesion", soprattutto upregolato nei cuori trisomici.

### **OXIDATIVE PHOSPHORYLATION**





ELECTRON TRANSPORT CHAIN



EXTRACELLULAR MATRIX



#### **CELL ADHESION**



a: Geni downregolati nei campioni DSH vs NH

GENE SET NAME	SIZ	ES	p-value
	Ε		
MITOCHONDRIA	62	-0,3865	< 0.001
ELECTRON TRANSPORT CHAIN	29	-0,5448	< 0.001
OXIDATIVE	15	-0,5340	< 0.001
PHOSPHORYLATION			

b: Geni upregolati nei campioni DSH vs NH

GENE SET NAME	SIZE	ES	p-value
EXTRACELLULAR	53	0.5660	< 0.001
MATRIX			
CELL ADHESION	15	0.5872	< 0.001

### Validazione dei geni Hsa21 e dei geni differenzialmente espressi

Allo scopo di confermare i risultati ottenuti con i microarray, 14 geni Hsa21, scelti tra quelli upregolati, quelli normoregolati e quelli che possono essere maggiormente implicati nello sviluppo della cardiopatia, sono stati validati con la RT-PCR. Sono stati valicati, inoltre, 9 geni OXPHOS e 7 geni ECM.

La validazione con la RT-PCR è stata estesa ad una serie di geni candidati per la cardiopatia: SRF, HAND2, TBX20 e NFATc, la cui aploinsufficienza può causare difetti cardiaci e disfunzioni mitocondriali (Bushdid et al, 2003).

SRF ed HAND2 sono risultati downregolati nei campioni trisomici, cardiopatici e non, rispetto ai controlli normali (**tab. 4**).

Poiché TBX20 non è presente sul chip HG-U133A la sua espressione è stata quantizzata solo attraverso RT-PCR. I nostri risultati dimostrano che TBX20 è downregolato in tutti i cuori fetali DS (**tab. 4**).

Functional Category					p-value		r	Primers for RT-PCR
	Probe ID	Gene Name	GenBank	Micro-array Fold change		RT-PCR		
				Fold change		1410		LEFT-
Oxydative Phospho-	201226_at	NDUFB8	NM_005004	0.65	0.0160	0.76	0.78	GCCAAGAAGTATAATATGCGTGTG <b>RIGHT-</b> GTCAGGGAGCTTCGGGTAG
rylation	201740 at	NDUFS3	NM 004551	0.76	0.0294	0.60	0.88	LEFT-GATTATGGCTTCGAGGGACA RIGHT-ACCCGCTTCACTTCATCATC
								LEFT-CCGAGCAGTCCTCTCAGC
	201903_at	UQCRC1	NM_003365	0.77	0.0142	0.89	0.84	RIGHT- TGTTCCCTTGAAAGCCAGAT
								LEFT-GAATGGGCACAGCAGTTTG RIGHT-
	201966_at	NDUFS2	NM_004550	0.67	0.0010	0.58	0.89	GGCCCAAAGTTCAGGGTAAT
	203606_at	NDUFS6	NM_004553	0.75	0.0464	0.63	0.84	RIGHT-CACGGGCTGCTCTGCTAT
								LEFT-AACTGGGCCTTGACAAAGTG RIGHT
	203663_s_ at	COX5A	NM 004255	0.79	0.0400	1.04	0 79	GGTAACTGTTCACACTCAAGTAGC
	203858_s_	CONSIL	NR 604202	0.75	0.0400	0.66	0.75	LEFT-CTTTTGACTGGCCCTGTTTC
	at	COX10	NM_001303	0.77	0.0101	0.66	0.92	RIGHT-ACCAGCGGTCTGTTCTTTGT LEFT-GTCATGCCGGGGATAGTG
								RIGHT- TTAAGCACAGCAGAACTAATTTTC
	218160_at	NDUFA8	NM_014222	0.73	0.0445	0.71	0.94	A
	218190_s_ at	UCRC	NM_013387	0.76	0.0017	0.76	0.89	LEFT- GACGCTATCTACGACCACATCA RIGHT-GGTCCTTCTGGCCTGGAT
Extra- cellular	201069_at	MMP2	NM_004530	1.59	0.0069	1.89	0.93	LEFT- TCCACCACCTACAACTTTGAGA RIGHT-AACTTGCAGGGCTGTCCTT
Matrix	202310_s_	COLIAI	K01228	1.55	0.0333	5.86	0.90	LEFT-TGTTCAGCTTTGTGGACCTC RIGHT- CTGTACGCAGGTGATTGGTG
	at	COLIM	<b>R</b> 01220	1.55	0.0555	5.80	0.90	LEFT-CACATGCCGTGACTTGAGAC
	202403_s_ at	COL1A2	NM_000089	1.60	0.0192	3.59	0.77	RIGHT- TAGCATCCATAGTGCATCCTTG
	202994 s							LEFT-GCCACAAGTGCGAGAACAC
	at	FBLN1	Z95331	1.72	0.0271	2.38	0.89	TAGACGTTGGCACACTCCTG
	209081_s_ at	COL18A1	AF018081	1.57	0.0079	1.91	0.95	LEFT-GIGGCCCTCTACGIGGACT RIGHT-TCTGAGTCATCGCCTTCCTC
	213428_s_ at	COL6A1	AA292373	1.62	0.0493	2.20	0.78	LEFT-AGGAGACCCTGGTGAAGCTG RIGHT-AGGTCCTGGGGGCTCCTCT
	200156							LEFT-GACCTGGTCGCTGAGAAGTT
	209136_s_ at	COL6A2	AY029208	2.39	0.0002	2.18	0.79	GCCTTGTGGAAGTTCTGCTC
Other genes	205548 s							LEFT-GAGGCAGTTGAGAGGTTTGC RIGHT-
Other genes	at	BTG3	NM006806	1.68	0.0134	1.62	0.83	GAGTGAGCTCCTTTGGCAAG
	35776_at	ITSN1	AF064243	1.72	0.0003	1.57	0.88	<b>RIGHT-</b> GATCATGCTTCGCTCTTTCC
	205593_s_ at	PDE9A	NM 002606	2.05	0.0067	2.05	0.89	LEFT-CAGAACGCACTCCGTACAAA RIGHT-TGGGCTCTACCTGTCCACTT
	211065_x_	DEVI	PC006422	1 4 2	0.0000	1.07	0.01	LEFT-GGTGGACCTGGAGAGCTG
	al	rfkl	DC000422	1.42	0.0088	1.9/	0.81	LEFT-GCATCATCAATTTCGAGCAG
	200642 at	SOD1	NM 000454	1.16	>0.05	0.65	0.83	RIGHT- CAGCCTCTGTATTATCTCCAA
					> 0.05			LEFT-AAAGACTCGGCGTGTTGTC
	203405_at	DSCR2	NM_003720	1.15	>0.05	0.85	0.78	GAATTGCTGGGATTTTCCAT

1.1								
								LEFT-CATACAGAGCCACGTCATCC
	202671_s_		224 000 001		>0.05	1.00	0.00	RIGHT-
	at	PDXK	NM_003681	1.12		1.00	0.98	GCATAGCCIGIGIGGGITIGA
	202325_s_		NR 001605	1.00	0.05	1.17	0.70	LEFT-IGTTIGGCTTCIGTCICACC
	at	ATP5J	NM_001685	1.00	>0.05	1.17	0.78	RIGHT-GGCTGACCGAATGACAGAG
					0.05			LEFT-CICAGCIICGIGIIIIGGAIG
	202740 -+	WDD	NIM 004607	1 17	>0.05	1.05	0.95	KIGHI-
	202749_at	WKB	NM_004627	1.17		1.25	0.85	
					0.0001			LEFT-
					0.0001			
	200022							
	209033_8_ at	DVRK1A	D86550	1 53		1 71	0.84	GC
	at	DIKKIA	D00550	1.55		1.71	0.04	I FFT
					>0.05			
	208370 .				20.05			RIGHT.
	at	DSCR1	NM 004414	1.25		4.35	0.83	GGGAGGGGAGATCAGAAACT
								LEFT-CTTTGCAATGGCAAGAGGA
	210555 s				>0.05			RIGHT-
	at	NFATC3	U85430	0.45		0.54	0.97	GATGAGGCACAGGCAAAGAT
								LEFT-
					0.019			GAGTTCACATCCCAGAGTCCA
								RIGHT-
	217526_at	NFATC2	AI478300	0.77		0.67	0.84	GAGCACTCGATGGGGTTAGA
								LEFT-AAAACAGCCCCTATCGAAGC
								<b>RIGHT-</b> CATTTCTGTGCCTTCATTGC
		SRF				0.71	0.85	
								LEFT-ACATCGCCTACCTCATGGAC
								RIGHT-TCTTGTCGTTGCTGCTCACT
		HAND2				0.65	0.75	
								I FFT-CCTGGAGACCAAGGACCTTT
								RIGHT-CACCCCCGAAAAGGACAC
	Assente	TBX20		Assente	Assente	0.87	Assente	

**Tab. 4:** Rapporti della media di espressione (campioni DS/controlli) calcolati dai dati dei microarray e dell'RT-PCR e coefficiente di correlazione di Pearson (r) tra i microarray e l'RT-PCR. I geni Hsa21 sono indicati in grassetto.

Infine, l'ultimo pathway analizzato, importante regolatore della cardiogenesi, è stato quello del signaling Calcineurina/NFAT. La scelta di questo pathway è scaturita dal fatto che i fattori di trascrizione NFATc inducono la trascrizione di geni cuore-specifici e sono regolati negativamente da DSCR1 e DYRK1A. I geni DSCR1 e DYRK1A, entrambi localizzati sull'Hsa21, all'interno della regione critica per la DS, cooperano per prevenire la traslocazione nucleare dei fattori NFATc, riducendo in questo modo la trascrizione di geni target critici per la morfogenesi del cuore (Arron et al, 2006).

E' stata quindi esaminata l'espressione di DYRK1A, DSCR1 e dei geni del pathway della Calcineurina (NFATc1-4) nei campioni DSH ed NH.

DYRK1A è upregolato ed NFATc downregolato nei campioni trisomici con p<0.05 (tab. 4); anche NFATc3 è downregolato (p<0.06). L'espressione di

DSCR1 nei cuori fetali varia molto tra i campioni individuali; in generale non è significativamente disregolato, ma il suo pattern di espressione è inversamente correlato con quello di NFATc3, indipendentemente dalla trisomia o dalla cardiopatia (**fig. 15**).

L'analisi, mediante RT-PCR, di tutti i geni dei pathway, ha evidenziato sempre con una correlazione tra le due tecniche superiore all'80%. (**tab. 4**)



# Fig. 15: Livello di espressione dei geni DSCR1 ed NFATC3 nei campioni individuali.

I valori di espressione dei microarray di DSCR1 (linea rossa) e di NFATC3 (linea verde) in ogni singolo campione mostrano una correlazione inversa (r=-0.66) tra i due geni.

# Geni differenzialmente espressi tra feti trisomici con e senza cardiopatia

L'analisi di espressione genica differenziale di feti DS con (CDH) e senza (DH) cardiopatia ha rivelato una forte omogeneità tra i 2 gruppi. Solo 42 geni risultano differenzialmente espressi, 19 upregolati e 23 downregolati nel gruppo CDH, con un FC compreso tra 1.2 e 2 e p<0.5 (**fig. 16**). Nessuno di questi geni mappa su HSA21. La stessa analisi funzionale usata per confrontare campioni trisomici e controlli non ha prodotto risultati significativi quando sono stati considerati i geni differenzialmente espressi tra CDH e DH. Allo stesso modo non sono state riscontrate differenze nell'espressione dei geni mitocondriali e della ECM tra feti DS con e senza cardiopatia. Non c'è una sostanziale differenza nell'espressione di SRF; HAND2 sembra essere downregolato 1.3 volte nei CDH, ma comunque non in maniera statisticamente significativa.



# Fig. 16: Volcano plots ottenuti confrontando feti DS con e senza difetti cardiaci.

42 dei 15.000 geni analizzati sono differenzialmente espressi nei 5 campioni DS con anomalie cardiache (CDH) rispetto 5 campioni DS senza anomalie cardiache (DH) (16 a). Nel confronto tra CDH vs DH solo 3 geni dei 600 circa codificanti per proteine mitocondriali (16 b) sono downregolati e 1 dei 700 circa codificanti per proteine della ECM (16 c) è upregolato. I puntini rossi indicano i geni differenzialmente espressi con un FC>1.2 e p<0.05; i puntini gialli indicano geni senza variazione significativa.

# Ricerca dei miR del cromosoma 21 ed identificazione dei corrispettivi targets

Attraverso il miRBase::Sequence (http://microrna.sanger.ac.uk/sequences/search.shtml) sono stati identificati 5 miR del cromosoma 21. Essi sono: miR-99a, miR-125b, miR-155, let-7c e miR-802. I nostri studi dimostrano che let-7c, un miR espresso in cuore (Thum et al, 2007) è upregolato nel tessuto cardiaco di feti Down. Qui di seguito sono mostrati i risultati dell'RT-PCR di let-7c (**tab. 5 e fig. 17**).

Name	let-7c	<b>U6</b>	DCt	DDCt	2 <sup>^</sup> -DDCt
NH	23,85	28,97	-5,12	0	1
DH	15,27±1	22,00	-6,73	-1,61	3,07±0.6
CDH	15,82±1.9	23,24	-7,42	-2,30	4,94±1.3

**Tab. 5** Media dei livelli di espressione di let-7c nei gruppi NH, DH e CDH. Questi dati dimostrano che let-7c è upregolato di circa 3 volte nei campioni Down rispetto ai non Down e di circa 5 volte nei campioni Down con cardiopatia.



Fig. 17: Plot dei livelli di espressione di let-7c in cuori di feti Down con e senza cardiopatia.

I valori sono riportati in scala logaritmica.

In attesa della conferma che anche gli altri miR del 21 siano overespressi in cuori di feti trisomici per effetto primario del dosaggio genico e postulando un eventuale meccanismo di regolazione trascrizionale da parte dei miR, per ciascuno dei miR del cromosoma 21 è stato eseguito un laborioso approccio bioinformatico allo scopo di individuare possibili targets.

Per l'identificazione dei targets il Sanger Database utilizza 3 softwares: Miranda, Target Scan e Pictar. Il principio su cui si basano questi softwares è pressocchè simile ma con piccole differenze (Sethupathy et al, 2006).

Miranda e Target Scan riconoscono l'importanza della regione *seed* al 5' del miR. Si ritiene, infatti, che uno dei criteri di validazione di targets di microRNA sia il perfetto appaiamento dei primi 8 nucleotidi al 5' del miR (regione seed) con il sito target nella 3' UTR dell'mRNA (Zhao et al, 2005).

Miranda non richiede una perfetta complementarietà con il target in corrispondenza di tale regione, contrariamente a Target Scan. Pictar, invece, dà la possibilità di individuare targets comuni a più miR e, per quei targets che non

hanno una complementarietà perfetta in corrispondenza della regione seed del miR, stabilisce una soglia energetica ( $\Delta G$ ) per la quale è più possibile che siano effettivamente targets.

Per ogni miR sono state generate 3 diverse liste di geni ognuna corrispondente ad un software differente. Le liste dei geni sono state poi inserite nel sito Affymetrix per assegnare a ciascun gene il corrispettivo probe set. Le liste dei probe set sono state poi sovrapposte (*merge*) e ordinate in ordine crescente. Infine sono stati selezionati solamente i probe sets e quindi i geni target presenti in tutte e tre liste. La lista finale di probe sets di ciascun miR è stata inserita nel software Gene Spring per valutare l'espressione differenziale di questi geni fra Down/non Down. Sono stati isolati solo quei geni che risultano downregolati nei campioni trisomici.

Poiché i nostri dati dimostrano che let-7c è overespresso in cuori di feti trisomici, è stata stilata, attraverso il software Gene Spring, una lista di geni target downregolati espressi in cuore (**tab. 6**). La downregolazione di questi geni potrebbe essere in qualche modo correlata a difetti cardiaci.

	Pres.				Locatio
Probe set	Call	Gene name	GenBank	Description	n
			NM_00631		
204564_at	15	PCGF3	5	ring finger protein 3	4p16.3
				creatine kinase,	
			NM_00182	mitochondrial 2	
205295_at	15	CKMT2	5	(sarcomeric)	5q13.3
				glutamic-oxaloacetic	
				transaminase 1, soluble	
				(aspartate aminotransferase	10q24.1
208813_at	15	GOT1	BC000498	1)	-q25.1
221636_s_a		DKFZp586		hypothetical protein	
t	15	G2122	AL136931	FLJ20605	1q42.11
		PIASY;			
		Piasg;			
		FLJ12419;		protein inhibitor of	
		MGC3529		activated STAT protein	
212881_at	15	6	AK022481	PIASy	19p13.3
204185_x_a			NM_00503	peptidylprolyl isomerase D	
t	15	PPID	8	(cyclophilin D)	4q31.3

**Tab. 6** Targets di let-7c espressi in cuore e downregolati nei campioni trisomici. Qui di seguito è mostrata l'espressione, attraverso microarray, di tali geni in cuore fetale umano nei diversi campioni.



**Fig. 18** Livelli di espressione dei targets di let-7c indicati in tabella nei campioni Down con cardiopatia (CDH), nei campioni Down (DH) e nei campioni euploidi (NH)

E' probabile, però, che geni target di let-7c espressi in cuore non siano downregolati, in quanto il miR potrebbe agire sul meccanismo della traduzione influenzando il livello proteico. Per avere un quadro completo sarà necessario, quindi, saggiare i livelli proteici dei targets dei miR di interesse.

Sono in corso alcuni esperimenti non ancora conclusi per valutare esattamente quali altri miR del cromosoma 21 sono espressi in cuore, e se sono effettivamente upregolati.

Si sta valutando inoltre il livello d'espressione dei miR coinvolti nella cardiogenesi come miR-1 e miR-133.

## <u>DISCUSSIONE</u>

**Effetto primario del dosaggio genico: overespressione dei geni Hsa21.** Attraverso l'utilizzo della tecnologia dei microarrays è stato possibile dimostrare che nel tessuto cardiaco fetale di soggetti con trisomia del cromosoma 21 vi è una upregolazione di circa 1.5 volte dei geni Hsa21. Questa upregolazione è indipendente dalla presenza di malformazioni cardiache.

Questo risultato è un'apparente conferma della predizione che la presenza di una terza copia del cromosoma 21 determini un aumento del 50% nell'espressione dei geni trisomici. Questo rappresenterebbe un semplice modello di regolazione genica, in cui il livello di trascritto è proporzionale al numero di copie del gene.

L'aumento globale di 1.5 volte, tuttavia, non risulta semplicemente da una upregolazione di ciascun singolo gene. Infatti ci sono geni che sono upregolati di più di 2 volte e geni che non sono affatto upregolati. Questi risultati suggeriscono effetti compensatori o un'eterogeneità nei meccanismi regolatori dei geni Hsa21. Il livello di espressione genica di singoli Hsa21 nel cuore di soggetti con DS è di centrale importanza per capire come la trisomia Hsa21 causi CHD.

Vi potrebbe essere l'upregolazione di un singolo gene o la combinazione di geni disregolati alla base della cardiopatia.

# Effetti secondari del dosaggio genico: disregolazione di geni al di fuori del cromosoma 21.

Dall'analisi funzionale è risultato che 441 geni che non mappano sull'Hsa21 sono disregolati nei campioni trisomici. Questo dato supporta l'ipotesi secondo cui i geni al di fuori dell'Hsa21 possano avere un ruolo nel determinare le caratteristiche della DS (Reeves et al, 2001). La disregolazione di geni al di fuori del cromosoma 21, quale effetto secondario alla alterazione nei livelli di espressione dei geni presenti in triplice copia, non è sorprendente ed è giustificata dall'interazione tra prodotti genici del cromosoma 21 e geni nel resto del genoma. L'analisi funzionale dei geni differenzialmente espressi tra individui trisomici e controlli ha mostrato una globale downregolazione dei geni codificanti per proteine mitocondriali, specialmente enzimi coinvolti nel pathway della fosforilazione ossidativa. I geni di tutti i 5 complessi della fosforilazione editori di supporte dei geni proteine ed

attività enzimatiche possano essere ridotte, e che la funzione mitocondriale potrebbe essere di conseguenza danneggiata.

Precedenti lavori hanno dimostrato una riduzione dei livelli proteici dei complessi I, III e V nel cervello e nel cervelletto di individui con DS (Kim et al, 2000; Kim et al, 2001) e una riduzione dell'attività enzimatica mitocondriale nelle piastrine dei pazienti con DS (Prince et al, 1994). Recentemente è stato riportato che il cervello del topo Ts1Cje, un modello murino DS, ha un ridotto potenziale di membrana mitocondriale e una ridotta produzione di ATP (Shukkur et al, 2006). Questi risultati sono indicativi di una disfunzione mitocondriale molto estesa nella DS; l'indagine, però, non era mai stata estesa a livello di espressione genica.

I nostri dati suggeriscono che la disfunzione mitocondriale si potrebbe verificare anche nei cuori DS e che la ridotta espressione dei geni mitocondriali potrebbe essere la base molecolare di questa disfunzione nel cuore così come in altri tessuti DS. Fra i geni che contribuiscono alla funzione mitocondriale, l'apoptosis inducing factor (AIF), è una proteina che contribuisce al processo apoptotico mitocondriale. Recentemente è stato dimostrato che l'assenza di AIF compromette la fosforilazione ossidativa anche nei cardiomiociti, riducendo l'espressione delle subunità dei complessi I e IV. L'inattivazione specifica di AIF nel cuore causa una pronunciata frammentazione miofibrillare e una disorganizzazione determinando cardiomiopatia AIF mitocondriale. una dilatativa. è significativamente downregolato nei cuori di feti con DS.

L'alterazione della funzionalità mitocondriale non appare un effetto primario del dosaggio genico perchè i geni per la funzione mitocondriale, che mappano sull'Hsa21, non sono affatto espressi oppure non sono espressi in maniera differenziale nei cuori di feti con DS rispetto ai controlli.

L'analisi funzionale dei 473 geni differenzialmente espressi ha dimostrato anche una globale upregolazione dei geni che codificano per proteine della ECM. Questo gruppo include geni che mappano sull'Hsa21, come ADAMTS1, ADAMTS5, APP, JAM2, COL6A1, COL6A2 e COL18A1, upregolati nei campioni trisomici, e geni che non mappano sull'Hsa21, come fibronectina, fibulina, collagene di tipo I, tipo III, tipo V, tipo XV, metalloproteasi (MMPs). L'overespressione di questo gruppo di geni verosimilmente influenza le proprietà dell'adesione cellulare, determinando un aumento dell'adesività. In accordo con questa ipotesi, è la dimostrazione che le strutture derivanti da cellule dei cuscinetti endocardici di feti con trisomia Hsa21 sono più aderenti *in vitro* rispetto a quelle dei controlli (Wright et al, 1984).

E' stato proposto un modello stocastico per i difetti del setto nella DS, secondo il quale la maggiore adesività determina deficienze nello sviluppo del canale atrioventricolare, associato ad una variabilità clinica tra gli individui (Kurnit et al, 1985).

Molti dei geni della ECM, upregolati nei campioni trisomici, in particolare il collagene di tipo VI e le MMPs, sono stati analizzati in relazione al loro potenziale ruolo nella cardiopatia nella DS.

Il collagene di tipo VI è espresso nei cuscinetti endocardici durante la formazione del setto, persiste durante la differenziazione delle valvole cardiache ed è implicato nella differenziazione dei cuscinetti endocardici. La sua espressione pronunciata nella trisomia Hsa21 potrebbe portare ad AVSD (Gittenberger et al, 2003).

Le MMPs sono coinvolte nell'organogenesi cardiaca mediante la regolazione della proliferazione cellulare, la transizione epitelio-mesenchima, il rimodellamento e la degradazione della ECM. MMP2, in particolare, coordina gli stadi precoci della cardiogenesi, ovvero la formazione del tubo cardiaco, e la degradazione selettiva della ECM (Cai et al, 2000).

# Come può l'upregolazione dei geni Hsa21 influenzare l'espressione di altri geni?

Un problema chiave della DS è capire come l'upregolazione dei geni Hsa21 possa disregolare geni localizzati su cromosomi differenti. Recentemente è stata suggerita una relazione tra due geni Hsa21, *DYRK1A* e *DSCR1*, con la famiglia dei geni *NFATc*.

I fattori di trascrizione NFATc svolgono un ruolo chiave nella morfogenesi delle valvole cardiache e nella settazione del cuore (de la Pompa et al, 1998). Essi sono espressi in modo transiente nei cuscinetti endocardici e la loro attivazione (defosforilazione) è mediata dalla fosfatasi calcineurina, a sua volta regolata dal Ca<sup>+2</sup>. L'attivazione dei fattori trascrizionali NFATc ad opera della calcineurina promuove la loro traslocazione nucleare e l'induzione della trascrizione di geni cuore-specifici (Fuentes et al, 2000).

Evidenze sperimentali dimostrano che una overespressione, anche modesta, di DYRK1A determina una riduzione dei livelli e dell'attività delle proteine NFATc e potrebbe indurre difetti cardiaci e vascolari (Arron et al, 2006). Anche DSCR1, che codifica per una proteina regolatoria espressa nel tessuto cardiaco, esercita un effetto inibitorio sul pathway di signaling calcineurina/NFAT (Fuentes et al, 2000), anche se diverse isoforme della proteina possono avere effetti opposti (Qin et al, 2006).

La nostra analisi, sia mediante microarray che RT-PCR, ha mostrato che DYRK1A è upregolato e che NFATc2 ed NFATc3 sono downregolati nei campioni trisomici Hsa21. DSCR1 mostra livelli di espressione molto variabili tra campioni anche appartenenti allo stesso gruppo. Si è visto però che, nei singoli campioni, un aumento dell'espressione di DSCR1 nei cuori correla con una diminuzione di espressione di NFATc3 (figura 5).

Recenti dati mostrano che l'attività enzimatica del complesso II e IV della catena respiratoria e l'attività mitocondriale sono ridotte nei cardiomiociti Nfatc3-/-;Nfatc4-/-, suggerendo che il pathway calcineurina/NFAT possa influenzare l'attività mitocondriale durante lo sviluppo cardiaco (Bushidid et al, 2003).

Anche l'upregolazione del gene del collagene di tipo VI può essere indicativa di una relazione tra i geni Hsa21 e geni di cromosomi differenti. Il collagene di tipo VI, insieme con altri collageni, è un attivatore del "discoidin domain receptor tyrosine kinases" (DDRs) (Vogel et al, 1997). I DDRs attivati inducono l'espressione delle metalloproteasi come MMP1 e MMP2 (Vogel et al, 1997) e di altre proteine della ECM.

I nostri dati rivelano un'upregolazione di DDR1 e DDR2 nei campioni trisomici.

### Geni candidati per lo sviluppo della cardiopatia.

E' noto che SRF, HAND2 e TBX20 sono essenziali per lo sviluppo del cuore e che mutazioni nei geni che codificano per tali fattori sono la causa dei principali difetti cardiaci, quindi, è stato saggiato il livello di espressione dei suddetti fattori nei nostri campioni. I nostri risultati dimostrano che SRF e TBX20 sono downregolati nei campioni trisomici.

Se alla downregolazione di SRF corrispondesse un basso livello proteico potremmo postulare un meccanismo patogenetico alla base della cardiopatia. Infatti, in condizioni normali, SRF regola miR-1, un miR cuore-specifico, che inibisce traduzionalmente HAND2 (Olson, 2006). Quindi, se SRF è downregolato e alla sua downregolazione corrisponde un basso livello proteico, miR-1 sarebbe poco espresso e di conseguenza HAND2 non inibito traduzionalmente. Ciò comporterebbe un'eccessiva sintesi della proteina HAND2, il cui non corretto dosaggio proteico è causa di difetti cardiaci che interessano soprattutto la sezione ventricolare del cuore (Zhao et al, 2007).

L' interesse verso il fattore trascrizionale TBX20 nasce, invece, non solo dal fatto che è uno dei componenti del "core cardiac network" e che mutazioni del gene per esso codificante sono causa dei principali difetti di settazione (Kirk et al, 2007), ma anche dal suo legame con i geni della matrice extracellulare. Infatti, durante il rimodellamento delle valvole cardiache, TBX20 regola l'espressione delle MMPs che, a loro volta, inducono la degradazione e il rimodellamento delle componenti della matrice extracellulare nel cuore (Shelton et al, 2006).

La nostra ipotesi è che la downregolazione di TBX20 possa provocare gli stessi effetti dell' aploinsufficienza: difetti di settazione ed alterazione dell'espressione dei geni della matrice extracellulare con conseguente formazione anomala delle valvole atrioventricolari.

#### miR Hsa21 e regolazione genica.

Dopo aver dimostrato che la triplice copia del cromosoma 21 disregola l'espressione dei geni Hsa21 e di quelli localizzati altrove nel genoma, la nostra attenzione si è focalizzata sui microRNA sulla base del ruolo che essi svolgono nella regolazione dell'espressione genica.

I nostri dati dimostrano che let-7c è upregolato nei cuori Down verso i non Down di circa 3 volte, mentre risulta upregolato di circa 5 volte nei campioni Down con cardiopatia rispetto ai controlli. La differenza di espressione di let-7c tra Down e non Down è altamente significativa, mentre quella tra Down con e senza cardiopatia non è altamente significativa essendo il livello di espressione più variabile nel gruppo dei campioni Down con cardiopatia (CDH); è necessario, pertanto, aumentare il numero di campioni analizzati data la forte variabilità interdividuale.

Poichè let-7c è upregolato nei cuori di feti Down, sono stati presi in esame i geni target espressi in cuore downregolati poiché, la loro downregolazione potrebbe essere causata da let-7c e in qualche modo correlata a difetti cardiaci.

Se alla downregolazione corrispondesse anche un basso livello proteico si potrebbe provare a definire un meccanismo patogenetico alla base della cardiopatia.

# **CONCLUSIONI**

Questo studio ha dimostrato che nel cuore di feti con trisomia Hsa21 molti geni del cromosoma 21 sono espressi in cuore e globalmente upregolati, inoltre molti geni localizzati su cromosomi differenti sono disregolati.

L'analisi funzionale ha mostrato che la disregolazione coinvolge due importanti pathways: il pathway mitocondriale della Fosforilazione Ossidativa ed il pathway "focal adhesion", comprendente le proteine della ECM. In senso più generale, gli esperimenti di microarray e di RT-PCR hanno evidenziato una downregolazione dei geni codificanti per gli enzimi mitocondriali ed un'upregolazione dei geni codificanti per le proteine della ECM in tutti i cuori trisomici, indipendentemente dai CHD. E' possibile ipotizzare che ciò potrebbe costituire una base per i difetti cardiaci.

Alcuni geni del core cardiac network come SRF e TBX20 sono disregolati. Tale disregolazione può, in maniera diretta o indiretta, attraverso l'influenza di altri fattori di trascrizione coinvolti nello sviluppo del cuore, essere la causa dei principali difetti cardiaci congeniti.

Il fatto che non siano state riscontrate differenze significative tra i due gruppi di feti DS, con e senza cardiopatia, potrebbe essere dovuto all'eterogeneità dei difetti cardiaci nei nostri campioni CDH, al numero dei campioni e dei geni analizzati o all'influenza di fattori esterni o genetici che interferiscono con il fenotipo.

Per questo motivo l'analisi dovrà essere estesa ad un numero maggiore di campioni (10-15 per gruppo) separando i campioni cardiopatici in base allo specifico difetto cardiaco (Tetralogia di Fallot, AVC o VSD) e aumentando il numero di geni valutati.

Fattori come le differenze nel background genetico, diversi aplotipi Hsa21, fattori stocastici e/o ambientali, l'interazione di fattori di trascrizione cardiaci e l'azione di microRNA, potrebbero giocare un ruolo critico nel determinare il quadro patogenetico.

Poiché è forte la correlazione esistente tra fattori di trascrizione coinvolti nella morfogenesi del cuore e miR, i nostri studi futuri sono proiettati nel valutare quali altri miR del 21 sono espressi in cuore e se sono tutti effettivamente upregolati in condizioni di trisomia 21.

Inoltre, poiché la nostra analisi si è basata principalmente sulla valutazione dell'espressione genica, ci siamo proposti di estendere l'indagine sui targets dei miR anche a livello proteico. Infine, allo scopo di comprendere le basi molecolari della cardiopatia pensiamo di estendere la nostra analisi ai miR cuore-specifici e ai corrispettivi targets saggiando sia l'espressione genica sia i livelli proteici.

# <u>BIBLIOGRAFIA</u>

\* Amano K., Sago H., Uchikawa C., Suzuki T., Kotliarova SE, Nukina N., Epstein CJ, and Yamakawa K. Dosage-dependent over-expression of genes in the trisomic region of Ts1Cje mouse model for Down sindrome. Human Molecular Genetics, 2004, (13) Vol.13.

\* Arron JR, Winslow MM, Polleri A., Chang CP, Wu H., Gao X., Neilson JR, Chen L., Heit JJ, Kim SK, Yamasaki N., Miyakawa T., Francke U., Graef I., and Crabtree GR. NFAT dysregulation by increased dosage of DSCR1 and DYRK1A on chromosome 21. Nature 2006. Vol. 441.

\* Barlow GM, Chen XN, Shi ZY, Lyons GE, Kurnit DM, Celle L, Spinner NB, Zackai E, Pettenati MJ, Van Riper AJ, Vekemans MJ, Mjaatvedt CH, Korenberg JR. Down syndrome congenital heart disease: a narrowed region and a candidate gene. Genet Med 2001, 3(2):91-101.

"\* Bushdid PB, Osinska H., Waclaw RR, Molkentin JD and Yutzey KE. NFATc3 and NFATc4 are required for cardiac development and mitochondrial function. Circulation Research 2003; 92; 1305-1313."

\* Cai DH, Vollberg TM, Hahn-Dantona E., Quigley JP, and Brauer PR. MMP-2 expression durino early avian cardiac and neural crest morphogenesis. The anatomical record 259: 168-179, 2000.

\* Carè A., Catalucci D., felicetti F., Bonci D., Addario A., Gallo P., Bang ML, Segnalini P., Gu Y., Dalton ND, Elia L., Latronico MVG, Hoydal M., Autore C., Russo MA, Dorn II GW, Ellingsen O., Ruiz-Lozano P., Peterson KL, Croce CM, Peschle C., and Condorelli G. Micro-RNA 133 controls cardiac hypertrophy. Nature Medicine, 2007. Vol 13 (5).

\* Chuaqui RF, Bonner RF, Best CJM, Gillespie JW, Flaig MJ, Hewitt SM, Phillips JL, Krizman DB, Tangrea MA, Ahram M., Linehan WM, Knezevic V and Emmert-Buck MR. Post-analysis follow-up and validation of microarray experiments. Nature Genetics, 2002, Vol.32.

\* Conti A., Fabbrini F., D'Agostino P., Negri R., Greco D., Genesio R., D'Armiento M., Olla C., Paladini D., Zannini M., and Nitsch L. Altered expression of mitochondrial and extracellular matrix genes in the heart of human fetuses with chromosome 21 trisomy. BMC Genomics, 2007 (8) 1:268.

\* Cullen BR. Transcription and processing of human microRNA precursors. Molecular Cell, 2004. Vol.16 861-865.

\* Dauphinot L., Lyle R., Rivals I., Dang MT, Moldrich RX, Golfier G., Ettwiller L., Toyama K., Rossier J., Personnaz L., Antonarakis SE, Epstein CJ, Sinet P-M and Potier M-C. The cerebellar transcriptome during postnatal development of the Ts1Cje mouse, a segmental trisomy model for Down syndrome. Human Molecular Genetics, 2005, (3) Vol.14. \* de la Luna S. and Estivill X. Cooperation to amplify gene-dosage-

\* De la Pompa JL, Timmerman LA, Takimoto H., Yoshida H., Elia AJ, Camper E., Potter J., Wakeman A., Marengere L., Langille BL, Crabtree GR, and Mak TW. Role of the NF-ATc transcription factor in

morphogenesis of cardiac valves and septum. Nature, 1998. Vol 392.

imbalance effects. Trends in Molecular Medicine 2006. Vol.30.

"\* Down JL Observation on an ethnic classification of idiots. London Hospital Reports 3:259-262; 1866 "

\* Fuentes JJ, Genescà L., Kingsbury TJ, Cunningham KW, Pèrez-Riba M., Estivill X., and de la Luna S. DSCR1, overexpressed in Down syndrome, is an inhibitor of calcineurin-mediated signalling pathways. Human Molecular Genetics, 2000. Vol.9 (11) 1681-1690.

\* Gardiner K. and Davisson M. The sequence of human chromosome 21 and implications for research into Down syndrome. Genome Biology 2000, I (2):review0002.1-0002.9.

\* Gardiner K. Gene-dosage effects in Down Syndrome and trisomic mouse models. Genome Biology 2004, 5:244.

\* Gittenberger-de groot AC, Bartram U., Oosthoek PW, Bartelings MM, hogers B., Poelmann RE, Jongewaard IN, and Klewer SE. Collagen type VI expression during cardiac development and in human fetuses with trisomy 21. The anatomical record 2003, 275A:1109-1116.

"\* Kim Sh, Fountoulakis M., Dierssen M., and Lubec G. Decresead protein levels of complex I 30-kDa subunit in fetal Down syndrome brains. J Neural Transm Suppl 2001;(61):109-16."

"\* Kim SH, Vlkolinsky R., Cairns N., and Lubec G. Decreased levels of complex III core protein 1 and complex V beta chain in brains from patients with Alzheimer's disease and Down syndrome. Cell Mol Life Sci 2000; 57(12):1810-6."

"\* Kirk EP, Sunde M., Costa MW, Rankin SA, Wolstein O., Castro ML, Butler TL, Hyun C., Guo G., Otway R., Mackay JP, Waddel LB, Cole AD, Hayward C., Keogh A., Macdonald P., Griffiths L., Fatkin D., Sholler GF, Zorn AM, Feneley MP, Winlaw DS, and Harvey RP. Mutations in cardiac T-box factor gene Tbx20 are associated with diverse cardiac pathologies, including defects of septation and valvulogenesis and cardiomyopathy. Am.J.Hum.Genet. 2007; 81:280-291."

\* Kola I. and Hertzog P. Animal models in the study of the biological function of genes on human chromosome 21 and their role in the pathophysiology of Down syndrome. Human Molecular Genetics, 1997, Vol.6, 1713-1727.

\* Korenberg JR, Chen XN, Schipper R, Sun Z, Gonsky R, Gerwehr S, Carpenter N, Daumer C, Dignan P, Disteche C, Graham JM, Hugdins L., McGillivray B., Miyazaki K., Ogasawara N., Park JP, Pagon R., Pueschel S., Sack G., Say B., Schuffenhauer S., Soukup S., and Yamanaka T. Down syndrome phenotypes: the consequences of chromosomal imbalance. Proc Natl Acad Sci U S A 1994, 91(11):4997-5001.

\* Korenberg JR, Kawashima H., Pulst SM, Ikeuchi T., Ogasawara N., Yamamoto K., Schonberg SA, West R., Allen L., Magenis E., Ikawa K., Taniguchi N. and Epstein CJ. Molecular definition of a region of chromosome 21 that causes features of the Down Syndrome Phenotype. Am.J.Human Genet 1990, 47:236-246.

"\* Kurnit DM, Aldridge JF, Matsuoka R. and Matthysse S. Increased adhesiveness of trisomy 21 cells and atrioventricular canal malformations in Down syndrome: a stochastic model. Am.J.Med.Genet. 1985 Feb;20(2):385-99."

"\* Lejeune J., Gautier M. and Turpin R. Etudes des chromosomes somatique de neuf enfants mongoliens. Comptes Rendus Academie des Sciences Paris 248: 1721-1722; 1959 \* Li Z, Yu T, Morishima M, Pao A., LaDuca J., Conroy J., Nowak N., Matsui S-I, Shiraishi I. and Yu YE. Duplication of the entire 22.8 Mb human chromosome 21 syntenic region on mouse chromosome 16 causes cardiovascular and gastrointestinal abnormalities. Hum. Mol. Genet. 2007, 16(11): 1359-1366

\* Liu G., Loraine AE, Shigeta R., Cline M., Cheng J, VaimeeKam V., Sun S., Kulp D., and Siani-Rose MA. NetAffx: Affymetrix probesets and annotations. Nucleic Acids Research, 2003, (1) Vol.31.

\* M. Hattori, A. Fujiyama, T.D. Taylor, H. Watanabe, T. Yada, H.S. Park, A. Toyoda, K. Ishii, Y. Totoki, Choi D-K, Soeda E., Ohki M., Yaspo ML et al. The DNA sequence of human chromosome 21. Nature, 2000.

\* Mao R., Wang X., Spitznagel Jr EL, Frelin LP, Ting JC, Ding H., Kim J-W, Ruczinski I., Downey TJ and Pevsner J. Primary and secondary transcriptional effects in the developing human Down syndrome brain and heart. Genome Biology, 2005, 6:R107.

\* Mao R., Zielke CL, Zielke HR, and Pevsner J. Global up-regulation of chromosome 21 gene expression in the developing Down syndrome brain. Genomics 81 (2003) 457-467.

\* Miano JM, Ramanan N., Georger MA, de Mesy Bentley KL, Emerson RL, Balza RO, Xiao Q., Weiler H., Ginty DD, and Misra RP. Restricted inactivation of serum response factor to the cardiovascular system. PNAS 2004, vol. 101, 17132-17137.

\* Molkentin JD, Lin Q., Duncan SA, and Olson EN. Requirement of the transcription factor GATA4 for hert tube formation and ventral morphogenesis. Genes & Development 1997, 11:1061-1072.

\* Moore CS. Postnatal lethality and cardiac anomalies in Ts65Dn Down syndrome mouse model. Mammalian Genome 2006, (17) 1005-1012.

\* O'Doherty A, Ruf S., Mulligan C., Hildreth V., Errington ML, Cooke S., Sesay A., Modino S., Vanes L., Hernandez D., Linehan JM, Sharpe PT, Brandner S., Bliss TVP, Henderson DJ, Nizetic D., Tybulewicz VL, and Fisher EM. An aneuploid mouse strain carrying human chromosome 21 with Down Syndrome phenotypes. Science 2005, 309: 2033-2037 \* Ohara Y., Atarashi T., Ishibashi T., Ohashi-Kobayashi A., and Maeda M. GATA4 gene organization and analysis of its promoter. Biol. Pharm. Bull. 2006, 29(3) 410-419.

\* Olson EN. Gene regulatory networks in the evolution and development of the heart. Science 2006,Vol 313.

\* Olson LE, Richtsmeier JT, Leszl J, and Reeves RH. A chromosome 21 critical region does not cause specific Down syndrome phenotypes. Science 2004, 306(5696):687-690.

"\* Olson LE, Roper RJ, Sengstaken CL, Peterson EA, Aquino V, Galdzicki Z, Siarey R, Pletnikov M, Moran TH, and Reeves RH. Trisomy for the Down syndrome ""critical region"" is necessary but not sufficient for brain phenotypes of trisomic mice. Hum Mol Genet 2007, 16(7):774-82."

\* Pandey R., Guru RK, and Mount DW. Pathway Miner: extracting gene association networks from molecular pathways for predicting the biological significance of gene expression microarray data. Bioinformatics 2004 (13) Vol.20.

\* Parlakian A., Tuil D., Hamard G., Tavernier G., Hentzen D., Concordet JP, Paulin D., Li Z. and Daegelen D. Targeted inactivation of serum response factor in the developing heart results in myocardial defects and embryonic lethality. Molecular and Cellular Biology, 2004. Vol.24 (12) 5281-5289.

\* Plageman TF, and Yutzey KE. Differential expression and function of Tbx5 and Tbx20 in cardiac development. The journal of biological chemistry, Vol.279 (18)19026-19034, 2004.

"\* Prince J., Jia S., Bave U., Annerèn G., and Oreland L. Mitochondrial enzyme deficiencies in Down's syndrome. J Neural Transm Suppl 1994;8(3):171-81."

\* Reamon-Buettner SM, Cho S-H and Borlak J. Mutations in the 3'untranslated region of GATA4 as molecular hotspots for congenital heart disease (CHD). BMC Medical Genetics 2007, 8:38.

\* Reeves RH., Baxter LL, Richtsmeier JT. Too much of good thing: mechanism of gene action in Down Syndrome. Trends Genet. 2001, 17(2): 83-88.

\* Sago H., Carlson EJ, Smith DJ, Kilbridge J., Rubin EM, Mobley WC, Epstein CJ and Huang T-T. Ts1Cje, a partial trisomy 16 mouse model for Down syndrome, exhibits learning and behavioral abnormalities. Proc Natl Acad Sci USA1998, 95(5) 6256-6261.

"\* Sayed D., Hong C., Chen IY, Lypowy J., and Abdellatif M. MicroRNAs play an essential role in the development of cardiac hypertrophy. Circ. Res. 2007; 100; 416-424."

"\* Schiaffino S., Dallapiccola B. and Di Lisi R. Molecular genetics of congenital heart disease. A problem of faulty septation. Circ Res.1999;84:247-249."

\* Sethupathy P., Megraw M., and Hatzigeorgiou AG. A guide through present computational approaches for the identification of mammalian microRNA targets. Nature Methods 2006 (11) Vol.3

\* Shelton EL, and Yutzey KE. Tbx20 regulation of endocardial cushion cell proliferation and extracellular matrix gene expression. Developmental Biology, 2007, 302: 376-388.

\* Shukkur EA, shimohata A., Akagi T., Yu W., Yamaguchi M., Murayama M., Chui D., Takeuchi T., Amano K., Subramhanya KH, Hashikawa T., Sago H., Epstein CJ, Takashima A. and Yamakawa K. Mitochondrial dysfunction and tau hyper phosphorylation in Ts1Cje, a mouse model for Down syndrome. Human Molecular Genetics, 2006 (18) Vol.15.

\* Stennard FA, Costa MW, Elliot DA, Rankin S., Haast SJP, Lai D., McDonald LPA, Niederreither K., Dolle P., Bruneau BG, Zorn AM, and Harvey RP. Cardiac T-box factor Tbx20 directly interacts with Nkx2.5, GATA4, and GATA5 in regulation of gene expression in the developing heart. Developmental Biology 262 (2003), 206-224.

\* Subramanian A., Tamayo P., Mootha VK, Mukherjee S., Ebert BL, Gillette MA, Paulovich A., Pomeroy SL, Golub TR, Lander ES, and Mesirov JP. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. PNAS, 2005 (43) Vol. 102

\* Takeuchi JK, Ohgi M., Koshiba-Takeuchi K., Shiratori H., Sakaki I., Ogura K., Saijoh Y., and Ogura T. Tbx5 specifies the left/right ventricles and ventricular septum position during cardiogenesis. Development 2003,130, 5953-5964.

"\* Thum T., Galuppo P., Wolf C., Fiedler J., Kneitz S., van Laake LW, Doevendans PA, Mummery CL, Borlak J., Haverich A., Gross C., Engelhardt S., Ertl G., and Bauersachs J. MicroRNAs in the human heart. A clue to fetal gene reprogramming in heart failure. Circulation 2007;116:258-267"

\* Van Rooij E., Sutherland LB, Liu N., Williams AH, McAnally J., Gerard RD, Richardson JA, and Olson EN. A signature pattern of stressresponsive microRNAs that can evoke cardiac hypertrophy and heart failure. PNAS, 2006, vol. 183, 18255-18260.

\* Vandesompele J., De Preter K., Pattyn F., Poppe B., van Roy N., De Paepe A., and Speleman F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biology 2002, 3(7).

\* Vogel W., Gish GD, Alves F., and Pawson T. The discoidin domain receptor tyrosine kinases are activated by collagen. Molecular Cell, 1997. Vol.1 13-23.

\* Wang DZ, Chang PS, Wang Z., Sutherland L., Richardson JA, Small E., Krieg PA, and Olson EN. Activation of cardiac gene expression by miocardin, a transcriptional cofactor for serum response factor. Cell, 2001. Vol.105 851-862.

"\* Webb S., Brown NA and Anderson RH. Formation of the atrioventricular septal structures in the normal mouse. Circ Res. 1998;82:645-656."

\* Wright TC, Orkin RW, Destrempes M. and Kurnit DM. Increased adhesiveness of Down syndrome fetal fibroblasts in vitro. Proc. Natl. Acad. Sci. USA, 1984 Vol.81 2426-2430.

\* Zhang B., Schmoyer D., Kirov S., and Snoddy J. GOTree Machine (GOTM): a web-based platform for interpreting sets of interesting genes using Gene Ontology hierarchies. BMC Bioinformatics 2004, 5:16.

\* Zhao Y., ransom JF, Li A., Vedantham V., von Drehle M., Muth AN, Tsuchihashi T., McManus MT, Schwartz RJ, and Srivastava D. Dysregulation of cardiogenesis, cardiac conduction, and cell cycle in mice lacking miRNA-1-2. Cell, 2007.

\* Zhao Y., Samal E., and Srivastava D. Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. Nature, 2005. Vol.436.