



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova  
Dipartimento di Pediatria

SCUOLA DI DOTTORATO DI RICERCA IN: MEDICINA DELLO SVILUPPO E SCIENZE  
DELLA PROGRAMMAZIONE  
INDIRIZZO: GENETICA BIOCHIMICA E MOLECOLARE  
CICLO 22°

***FUNCTIONAL SPLICING ASSAY* MEDIANTE L'UTILIZZO DI  
MINIGENI PLASMIDICI NEL GENE *NF1***

**Direttore della Scuola:** Ch.mo Prof. Giuseppe Basso

**Coordinatore d'indirizzo:** Dott. Maurizio Scarpa

**Supervisore:** Ch.mo Prof. Maurizio Clementi

**Dottorando:** Monica Forzan



# Indice

<b>Abstract</b>	I
<b>1. Introduzione</b>	1
<b>1.1 Generalità</b>	3
<b>1.2 Neurofibromatosi di tipo 1 (NF1)</b>	4
1.2.1 Manifestazioni cliniche	5
1.2.2 Criteri diagnostici	7
<b>1.3 Il gene <i>NF1</i></b>	8
<b>1.4 Il prodotto genico: la neurofibromina</b>	11
1.4.1 <i>Splicing</i> alternativo	13
1.4.2 Distribuzione tissutale e localizzazione subcellulare	15
<b>1.5 Funzioni della neurofibromina</b>	15
1.5.1 Attività GAP	15
1.5.2 Neurofibromina e tubulina	17
1.5.3 Neurofibromina e cAMP	18
1.5.4 Neurofibromina e chinasi	19
<b>1.6 Il gene <i>NF1</i> è un oncosoppressore</b>	19
<b>1.7 Le mutazioni nel gene <i>NF1</i></b>	20
<b>1.8 Espressione fenotipica</b>	25
<b>1.9 Analisi di mutazione nel gene <i>NF1</i></b>	25
<b>1.10 Elementi di regolazione dello <i>splicing</i></b>	27
<b>2. Scopo della tesi</b>	33
<b>3. Materiali e metodi</b>	37
<b>3.1. Campione in esame</b>	39
<b>3.2. Estrazione DNA genomico</b>	39
<b>3.3 Analisi di mutazione del gene <i>NF1</i></b>	39

3.3.1 <i>Denaturing High Performance Liquid Chromatography</i> (DHPLC)	39
3.3.1.1 Amplificazione e analisi di mutazione mediante DHPLC	42
3.3.2 <i>High-Resolution Melting Analysis</i> (HRMA)	49
3.3.2.1 Amplificazione e analisi di mutazione mediante HRMA	53
3.3.3 Sequenziamento	59
3.3.4 Risultati analisi mutazione	61
3.3.4.1 Mutazioni di <i>splicing</i>	63
<b>3.4 Costruzione del <math>\beta</math>-globin minigene</b>	66
3.4.1 Amplificazione del gene della $\beta$ -globina	66
3.4.2 Clonaggio del gene della $\beta$ -globina in un vettore plasmidico	69
3.4.2.1 Digestione enzimatica	69
3.4.2.2 Ligazione	70
3.4.3 Trasformazione dei batteri	70
3.4.4 Trasfezione transiente di cellule eucariotiche	72
3.4.5 Estrazione dell'RNA totale e analisi del trascritto	73
3.4.5.1 Estrazione dell'RNA totale	73
3.4.5.2 Sintesi del cDNA mediante retrotrascrizione	73
3.4.5.3 Amplificazione del cDNA e sequenziamento	74
3.4.6 Sintesi della regione “adattatore” e inserimento nel vettore	75
3.4.7 Osservazione dei trascritti	79
3.4.7.1 Mutazione IVS 18+1 G>T	80

<b>4. Analisi e risultati</b>	85
<b>4.1 Mutazione <i>NFI</i> IVS 5-5 T&gt;G</b>	87
<b>4.2 Un caso particolare <i>CFTR</i> IVS 4+3 A&gt;G</b>	90
4.2.1 Dati di popolazione	91
4.2.2 Clonaggio della regione genomica che comprende l'esone 4 di <i>CFTR</i>	93
4.2.3 Espressione del minigene e analisi dei trascritti	95
4.2.4 Risultati di epidemiologia	96
4.2.5 Risultati dell'analisi molecolare	97
4.2.6 Conclusioni	99
<b>5. Conclusioni</b>	103
<b>6. Bibliografia</b>	109
<b>Appendice</b>	121
<b>Pubblicazioni</b>	



# **Abstract**





Le recenti conoscenze sulla complessità del processo di *splicing*, hanno rivelato l'esistenza di importanti elementi di regolazione dello stesso. Variazioni di questi nuovi elementi, che si trovano sia nelle regioni codificanti che non codificanti dei geni, possono manifestarsi con effetti deleteri sullo *splicing* del pre-mRNA. Forse una delle ipotesi più accreditate di questi anni è che anche variazioni silenti del modulo di lettura o variazioni delle sequenze introniche siano da considerare responsabili di patologie. L'attivazione di pseudo-esoni o di esoni criptici in seguito alla creazione di nuovi siti di *splicing* o all'alterazione di siti di *splicing* costitutivi sono eventi già osservati come causa di patologie da difetti dello *splicing*. Tuttavia, le informazioni attualmente disponibili sulle patologie genetiche rivelano che le conoscenze su molti dei meccanismi fondamentali che regolano lo *splicing* del pre-mRNA sono ancora molto scarse.

I difetti di *splicing* costituiscono il 30-50% del totale di mutazioni del gene *NF1*. Circa il 30% di queste mappa lontano dalle giunzioni esone-introne dove possono causare la creazione di nuovi siti di *splicing*, attivare siti criptici, o portare alla perdita o formazione di sequenze segnale esoniche e/o introniche, importanti per il riconoscimento del messaggero da parte dei fattori di *splicing*.

Sulla base di queste premesse l'obiettivo di questo lavoro è stato di ricercare mutazioni potenzialmente alteranti il meccanismo di *splicing* e di sviluppare un minigene plasmidico per caratterizzare *in vitro* l'effetto di queste variazioni genomiche sul meccanismo e sulle sue alterazioni.

Inizialmente è stata effettuata un'analisi di *screening* di mutazioni nel gene *NF1* su 310 pazienti affetti da Neurofibromatosi di tipo 1 afferiti presso il servizio di Genetica Clinica ed Epidemiologia del Dipartimento di Pediatria dell'Università degli Studi di Padova. La ricerca di mutazione è stata eseguita mediante DHPLC e HRMA. Ne sono emerse 196 mutazioni a carico del suddetto gene di cui 41 che alterano, o potrebbero alterare, il meccanismo di *splicing*.

Nel caso delle mutazioni osservate *de novo* sono stati eseguiti dei test *in silico* mediante l'utilizzo di *software* disponibili in rete come: *BDGP: Splice Site Prediction by Neural Network* ([http://www.fruitfly.org/seq\\_tools/splice.html](http://www.fruitfly.org/seq_tools/splice.html)) e *NetGene 2 Server* (<http://www.cbs.dtu.dk/services/NetGene2>) che analizzano variazioni a carico delle sequenze *consensus* GT-AG dello *splicing*, *Human*

*Splicing Finder Version 2.4* (<http://www.umd.be/HSF/>) e *ESEfinder Release 3.0* (<http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home>) che, basandosi su modelli di calcolo predittivi, prendono in considerazione anche le alterazioni delle sequenze ESE (*Exonic Splicing Enhancer*), ESS (*Exonic Splicing Silencer*), ISE (*Intronic Splicing Enhancer*) e ISS (*Intronic Splicing Silencer*).

Per la valutazione in vitro è stato costruito un *β-globin minigene*. Clonando all'interno di questo vettore la regione genomica interessata dalla mutazione, questo diventa un veicolo per trasfettare in modo transiente cellule HeLa in modo da poter estrarre i trascritti e valutare l'alterazione dello *splicing*.

L'analisi dei prodotti di trascrizione ha confermato la funzionalità del sistema. Nel caso della variazione *NF1* IVS 5-5 T>G si è potuto osservare che tale mutazione provoca lo *skipping* dell'esone 6 del gene *NF1*.

In un caso particolare è stata utilizzata la metodica per valutare gli effetti della variazione: *CFTR* IVS 4+3 A>G. L'analisi ha evidenziato che dal costrutto mutato si ottengono tre prodotti di trascrizione: un trascritto wild type, un trascritto che presenta la delezione di 93 nucleotidi all'estremità 3' dell'esone 4 di *CFTR* e un trascritto che presenta il completo *skipping* dell'esone 4 di *CFTR*. In aggiunta sono stati effettuati anche degli studi epidemiologici valutando mediante HRMA la presenza di tale variante in 250 individui sani di controllo. Il calcolo delle frequenze alleliche e l'analisi dei trascritti ha portato ad affermare che la variazione 621+3 A>G non sia da considerarsi la causa di forme gravi di Fibrosi Cistica, andando a smentire quanto precedentemente riportato in letteratura. Il lavoro è stato di recente pubblicato sulla rivista internazionale *Journal of Human Genetics*.

---

Recent knowledge of the splicing process complexity have unfolded the existence of important regulating elements of the process itself. Variations of these new elements, located both on coding and non-coding gene regions, can exert detrimental effects on the pre-mRNA splicing process. One of the most credited hypotheses in last years is that even silent variations of the reading frame or

variations of intronic sequences are to be considered potentially involved in the development of pathologies. The activation of pseudo-exons or of cryptic exons following the creation of new splice sites or the alteration of constitutive splice sites is already known to be causing pathologies related to a defective splicing process. Nevertheless, the informations on genetic pathologies available to date reveal that the knowledge on most of the fundamental mechanisms regulating the pre-mRNA splicing is still very poor.

Splicing defects amount to the 30-50% of *NFI* total mutations. About the 30% of these mutations map far away from exon-intron junctions where they can produce new splice sites, activate cryptic sites or lead to the loss or formation of intronic and/or exonic signal sequences important for the messenger recognition by the splicing factors.

Based on these premises, the goal of this study has been to search for mutations potentially responsible of altering the splicing mechanism and to develop a plasmid minigene in order to perform an in vitro characterization of the effect of these genomic variations on the mechanism and its alterations.

Initially, a screening analysis of the *NFI* gene mutations has been performed on 310 patients with Neurofibromatosis type 1 received at the University of Padova, Department of Pediatrics – Clinical Genetics and Epidemiology Service. The mutation search has been performed by means of DHPLC and HRMA techniques. The analysis revealed a total of 196 mutations of the *NFI* gene, of which 41 were implied, or potentially implied, in the alteration of the splicing mechanism.

In the case of de novo mutations, a series of in silico test has been performed, by means of softwares available on the net, such as: BDGP: Splice Site Prediction by Neural Network ([http://www.fruitfly.org/seq\\_tools/splice.html](http://www.fruitfly.org/seq_tools/splice.html)) and NetGene 2 Server (<http://www.cbs.dtu.dk/services/NetGene2>), both analysing variations in the splicing consensus sequences GT-AG; Human Splicing Finder Version 2.4 (<http://www.umd.be/HSF/>) e *ESEfinder Release 3.0* (<http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home>), that based on predictive models, take into account also the alterations of ESE sequences (Exonic Splicing Enhancer), of ESS sequences (Exonic Splicing Silencer), of ISE

sequences (Intronic Splicing Enhancer) and of ISS sequences (Intronic Splicing Silencer).

For the in vitro evaluation a  $\beta$ -globin minigene has been built. By cloning the genomic region including the mutation inside this vector, the vector itself become a vehicle to temporarily transfect HeLa cells, in order to extract the transcripts and evaluate the splicing alteration. The analysis of transcription products confirmed the functionality of the system.

In the case of the *NFI* IVS 5-5 T>G variation it has been observed that the mutation is responsible of causing the *NFI* exon 6 skipping.

In a particular case the method has been used to valuate the effects of the variation *CFTR* IVS 4+3 A>G. The analysis highlighted that from the mutated construct three transcription products are obtained: a wild type transcript, a transcript bearing the deletion of 93 nucleotides at the 3' end of the *CFTR* exon 4, and a transcript showing the complete skipping of the *CFTR* exon 4. In addition, an epidemiological study has been performed, evaluating the occurrence of the aforementioned variation in 250 healthy control individuals by means of the HRMA technique. Frequency computation and transcripts analysis allow to affirm that the variation 621+3 A>G is not to be considered the cause of severe forms of Cystic Fibrosis, differently from what previously reported in the literature. The study has been recently published on the international magazine Journal of Human Genetics.

# **Introduzione**



## 1.1 Generalità

Fino agli anni '70, i termini “neurofibromatosi”, “malattia di Von Recklinghausen” o “neurofibromatosi multipla” venivano utilizzati indistintamente per descrivere pazienti che presentassero alcuni segni clinici, come macchie caffè-latte e tumori del sistema nervoso centrale e/o periferico, combinati tra loro in diversa maniera.

Attualmente il termine “neurofibromatosi” si riferisce ad un gruppo di sindromi, aventi in comune delle caratteristiche ma rappresentanti, dal punto di vista clinico, delle entità distinte.

Sulla base dei criteri stabiliti dal “*National Institute of Health*” si definiscono due forme principali di neurofibromatosi: la **Neurofibromatosi tipo 1** (NF1) (MIM#162200) e la **Neurofibromatosi tipo 2** (NF2) (MIM#101000), che presentano una sovrapposizione fenotipica e genotipica minima.

Le due malattie sono causate da mutazioni a carico di due geni diversi che sono localizzati su cromosomi diversi.

**NF1:** il gene è localizzato sul cromosoma 17; è caratterizzata dalla presenza di numerose macchie caffè-latte (>6), lentiginosi in aree ben determinate del corpo, noduli iridei di Lisch e neurofibromi (tumori dei nervi periferici) che a seconda della sede si distinguono in cutanei e sottocutanei o nodulari.

**NF2:** il gene è localizzato sul cromosoma 22; è una forma piuttosto rara (frequenza di 1:33000 nati); è caratterizzata principalmente dalla presenza di tumori bilaterali del nervo acustico, detti schwannomi vestibolari.

Esistono altre forme considerate varianti alleliche sia della NF1 che della NF2:

**Sindrome di Watson:** macchie caffè-latte, bassa statura, *deficit* cognitivo e stenosi dell'arteria polmonare (MIM#193520).

**Neurofibromatosi spinale familiare:** neurofibromi paravertebrali e sottocutanei multipli (MIM#162210).

**Neurofibromatosi intestinale familiare:** neurofibromi gastrointestinali multipli (MIM#162220).

**Neurofibromatosi segmentale:** macchie caffè-latte e neurofibromi solo in alcune aree del corpo.

**Forma con macchie caffè-latte a trasmissione autosomica dominante:** presenza di sole macchie caffè-latte a trasmissione familiare (MIM#114030).

**Forma con neurofibromi a trasmissione autosomica dominante:** presenza di soli neurofibromi a trasmissione familiare.

**Sindrome Neurofibromatosi/Noonan:** manifestazioni sovrapposte di sindrome di Noonan e NF1 (MIM#601321).

**Schwannomatosi:** presenza di schwannomi multipli in assenza di altre manifestazioni cliniche della NF2; può essere sporadica ma anche familiare (MIM#162091).

**Schwannomatosi segmentale:** schwannomi multipli localizzati in aree corporee circoscritte in assenza di altre manifestazioni cliniche di NF2.

## 1.2 Neurofibromatosi tipo 1 (NF1)

La NF1 (MIM#162200, <http://www.ncbi.nlm.nih.gov/Omim/>), chiamata anche Malattia di Von Recklinghausen o Neurofibromatosi periferica o multipla, è la più comune tra le forme di neurofibromatosi.

Si tratta di una sindrome autosomica dominante con frequenza di 1 su 3000-3500 nati vivi in tutte le popolazioni.

Presenta penetranza completa età dipendente: infatti la maggior parte dei segni clinici sono presenti nel 90% dei pazienti entro la pubertà (Mattocks *et al.*, 2004). L'espressione fenotipica è estremamente variabile anche all'interno dello stesso nucleo familiare.

Le cause di questa variabilità si possono attribuire a diversi fattori:

- *background* genetico (polimorfismi che modulano l'entità delle manifestazioni);
- tipo di mutazione-malattia;
- fattori ambientali;
- presenza di mosaicismi.

Il tasso di mutazione spontanea è uno dei più elevati per quanto concerne i disordini genetici umani, ed è stimato tra  $1 \times 10^{-4}$  e  $3,1 \times 10^{-5}$  gameti per generazione



(Clementi *et al.*, 1990); da 10 a 100 volte maggiore del tasso abituale per un singolo locus.

Inoltre in circa il 50% degli affetti da NF1 non si tratta di una forma familiare della malattia: la frequenza di mutazioni *de novo* è infatti una delle più elevate tra le patologie genetiche.

### 1.2.1 Manifestazioni cliniche

- *Macchie caffè-latte (Cafè au lait spot, CLS)*: macule asintomatiche, piatte, omogeneamente pigmentate derivanti da un difetto di pigmentazione dei melanociti della pelle. Interessano circa il 95% degli affetti e sono presenti alla nascita o entro il 1°-2° anno di vita. Aumentano in numero e dimensioni fino ai 5-6 anni di età, poi tendono a depigmentarsi e possono quindi essere difficilmente riscontrabili negli individui adulti (Ruggieri e Tenconi, 2001). La bassa incidenza di melanomi tra gli individui affetti da NF1 suggerisce che le CLS non vadano incontro a cambiamenti neoplastici maligni.
- *Lentiginosi*: lesioni simili alle CLS ma di dimensioni più piccole (pochi mm). Compaiono verso i 6-7 anni di età; sono disposte “a spruzzo” alle ascelle, all’inguine, alla base del collo o anche in altre sedi. Possono aumentare di numero fino alla pubertà ma non di dimensione (Ruggieri e Tenconi, 2001).  
*Noduli di Lisch*: amartomi, noduli asintomatici, rotondeggianti visibili con lampada a fessura. Compaiono poco prima della pubertà nell’iride. Sono presenti nell’85-90% degli adulti affetti.
- *Neurofibromi*: tumori benigni del sistema nervoso; aumentano di dimensioni e numero con l’età o con variazioni ormonali come la pubertà o la gravidanza. Si distinguono in:
  - cutanei: piccoli noduli sollevati, di consistenza molle con una pigmentazione diversa da quella circostante. Istologicamente sono tumori benigni delle estremità terminali dei piccoli nervi cutanei composti da cellule di Schwann, fibroblasti, cellule perineurali, mastociti e cellule endoteliali. Generalmente compaiono intorno alla pubertà e sono presenti in un’alta percentuale di individui adulti affetti;

- nodulari (o sottocutanei): appaiono come noduli di consistenza duro-elastica palpabili nel tessuto sottocutaneo. Interessano i tronchi nervosi periferici principali o le radici dei nervi. Causano spesso deficit sensitivi o motori nel tratto interessato dal nervo colpito o dolore alla palpazione nella regione coinvolta.

E' inoltre possibile osservare un insieme di “segni clinici minori” presenti in una discreta percentuale di individui affetti ma non abbastanza da poterli considerare segni tipici della malattia.

Possono comunque essere un valido aiuto nella diagnosi soprattutto in età pediatrica quando i segni tipici non sempre sono già comparsi.

Sono considerati tali:

- *macrocefalia,*
- *bassa statura,*
- *malformazioni toraciche,*
- *ipertelorismo.*

La gravità della patologia spesso è determinata dal manifestarsi di “complicanze” rare ma invalidanti:

- *disturbi cognitivi e dell'apprendimento:* nel 40-60% degli affetti si riscontrano disturbi dell'apprendimento con deficit visuo-spaziali, dell'attenzione e della memoria a breve termine. In circa il 40% dei casi sono associati a deficit della coordinazione motoria fine e grossolana. Ritardo mentale vero e proprio (QI<70) si riscontra solo nel 3-5% dei casi e spesso è associato a particolari dismorfismi;
- *neurofibroma plessiforme:* si manifesta nel 25-30% degli affetti da Neurofibromatosi tipo 1. Appaiono come masse cutanee di consistenza molle con dimensioni che possono essere di pochi cm fino ad un'intera area del corpo. Le lesioni più gravi ed altamente sfiguranti compaiono nelle regioni del viso e del collo entro i due anni di età, mentre in altre sedi possono comparire verso i 20-25 anni causando danni estetici e funzionali minori. Le lesioni comparse in età infantile possono estendersi verso l'interno coinvolgendo grossi tronchi nervosi o importanti vasi venosi,

arteriosi o linfatici, oppure dislocando organi interni incontrati lungo il tragitto. Circa il 10-15% dei casi di neurofibroma plessiforme possono evolvere in tumore maligno della guaina dei nervi periferici (MPNST, *malignant peripheral nerve sheath tumor*) (Ferner and Gutmann, 2002);

- *neurofibroma diffuso*: si presenta come un placca di aspetto molliccio, tipicamente presente nella regione del cuoio capelluto o degli arti. Si distingue dagli altri neurofibromi per l'aspetto diffuso e per la sua elevata matrice di collagene fibrillare non sedimentata che non coinvolge un'area di un tronco nervoso. Si sviluppano diffondendosi tra i setti di tessuto connettivo e le cellule adipose. E' tipico dell'età infantile o dell'adolescenza;
- *complicanze ortopediche*: scoliosi, cifoscoliosi, lordoscoliosi, displasia congenita della tibia e pseudoartrosi, fibromi non ossificanti, spondilolistesi, ginocchio valgo o varo, difetti della corticale ossea, ipertrofia ossea focale, displasia dello sfenoide;
- *complicanze neurologiche*: lesioni iperintense nelle immagini pesate alla risonanza magnetica cerebrale (chiamate UBO *Unidentified bright objects*), tumori cerebrali (i più frequenti sono gliomi delle vie ottiche, pontini e cerebellari), tumori del midollo spinale e paravertebrali, convulsioni, cefalea, stenosi dell'acquedotto di Silvio, malattia cerebrovascolare, meningocele spinale, neuropatia da NF1;
- *altre complicanze*: ipertensione arteriosa, feocromocitoma, malformazioni cardiache congenite, xantogranulomi (Ruggieri e Tenconi, 2001).

### 1.2.2 Criteri diagnostici

L'espressione clinica della Neurofibromatosi tipo 1 è composta da uno spettro molto vario di manifestazioni che possono differire notevolmente da un paziente all'altro, tra famiglie diverse e anche tra membri di una stessa famiglia portatori della stessa mutazione.

Per permettere una diagnosi accurata, nel 1988, il NIH (*National Institute of Health*) definì i criteri diagnostici standard pubblicando un "*Consensus*

*Statement*".

La diagnosi di NF1 viene confermata se nel paziente si riscontrano due o più dei seguenti criteri:

- un numero di CLS non inferiore a 6;
- almeno due neurofibromi cutanei o un neurofibroma plessiforme;
- lentiginosi ascellare e/o inguinale;
- 2 o più noduli di Lisch;
- una lesione ossea tipica;
- un parente di primo grado con NF1 diagnosticata secondo i criteri sopra citati.

A causa di questa estrema variabilità dell'espressione clinica non è possibile prevedere la gravità della malattia sulla base della familiarità o delle analisi molecolari; non si è al momento a conoscenza di una relazione tra gravità e età di insorgenza, stato del parente affetto, ordine di nascita, età dei genitori o fattori ambientali.

La maggior parte degli affetti presenta sintomi lievi che permettono di condurre una vita relativamente normale; tuttavia l'aspettativa di vita di questi è ridotta in seguito all'aumentato rischio di sviluppare tumori maligni.

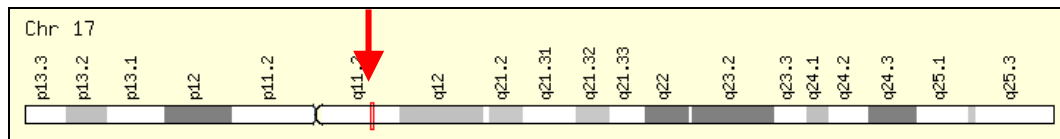
La Neurofibromatosi tipo 1 è infatti una malattia che ha un'alta predisposizione a tumore.

### **1.3 Il gene *NF1***

Mediante clonaggio posizionale il gene *NF1* fu mappato sul braccio lungo del cromosoma 17. I primi risultati furono ottenuti da studi di *linkage* del gene *NF1* con il gene *NGFR* (*nerve growth factor receptor*) localizzato in posizione 17q22 (Seizinger *et al.*, 1987). L'analisi con marcatori, tra cui pHHH202, ha permesso di localizzare il gene nella posizione 17q11.2 (White *et al.*, 1987).

La conferma dell'esatta localizzazione fu ottenuta dallo studio di anomalie cromosomiche riscontrate in due pazienti affetti da NF1. Si osservarono due traslocazioni indipendenti: t(1;17) (Schmidt *et al.*, 1987) e t(17;22) (Ledbetter *et al.*, 1989). I cromosomi traslocati presentavano i punti di rottura del cromosoma

17, nella regione in cui era stato mappato il gene *NFI*; si ritenne quindi che le traslocazioni fossero la causa della malattia perché interrompevano il gene *NFI* e causavano perdita della funzione genica (Cawthon *et al.*, 1990).



**Figura 1.** Localizzazione del gene *NFI* sul cromosoma 17 (17q11.2).

Il gene *NFI* ricopre una regione di 350 kb, è costituito da 60 esoni che codificano per un trascritto di 11-13 kb. Quest'ultimo comprende un' *Open Reading Frame* (ORF) di 8457 bp tradotta in una proteina, la neurofibromina, costituita da 2818 aminoacidi e avente una massa molecolare di 327 kDa.

All'interno dell'esone 1 del gene, oltre alla regione codificante, è contenuta anche la regione non tradotta in 5' (5' UTR); analogamente la regione non tradotta in 3' (3' UTR) si presenta in contiguità con il codone di stop della proteina nell'esone 49 e si estende per circa 3,5 kb.

L'esame dei siti di *splicing* degli esoni rispecchia una generale aderenza alla regola GT-AG delle sequenze *consensus* di *splicing* (Li *et al.*, 1995).

Le dimensioni dell'introne 1 sono stimate attorno a 140 kb, si presume che viste le notevoli dimensioni all'interno di questo siano presenti elementi regolatori che in qualche modo possano influenzare l'espressione del gene.

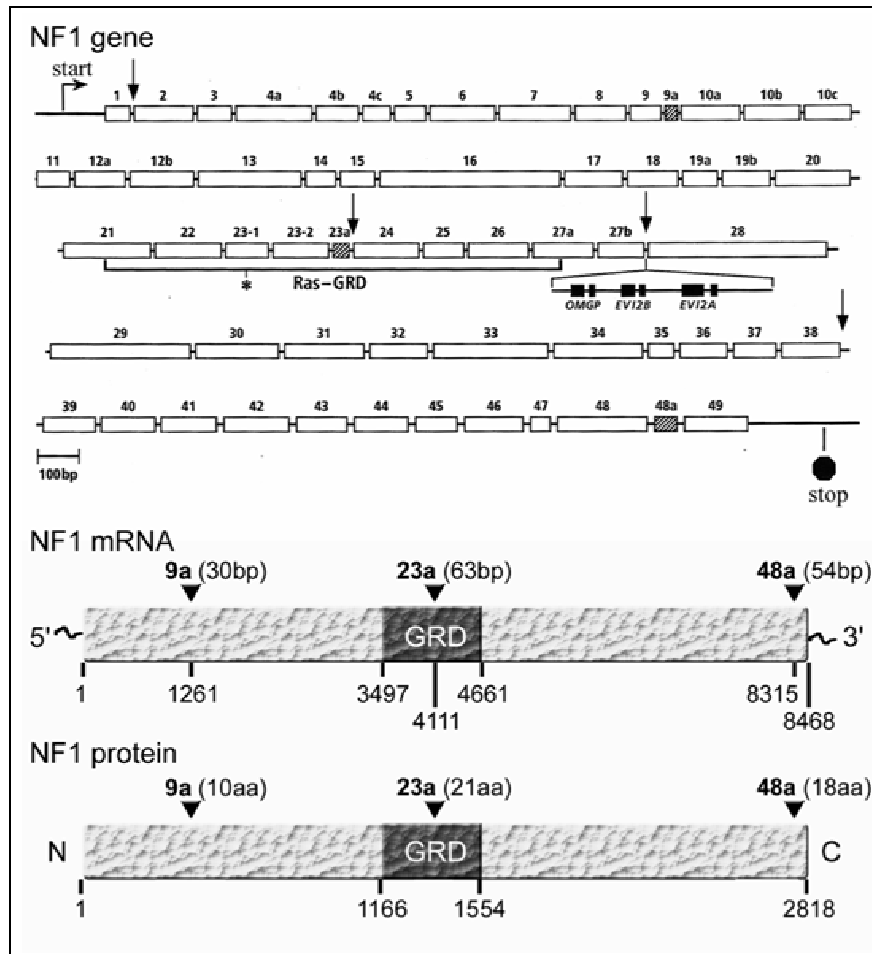
Nel genoma umano sono presenti diversi loci che mostrano più del 90% di identità con le sequenze del gene *NFI*; queste sequenze si trovano vicino alle regioni centromeriche di sette diversi cromosomi: 2, 12, 14, 15, 20, 21 e 22. Questi pseudogeni di *NFI* probabilmente si sono diffusi per riarrangiamento pericentromerico di parziali copie del gene *NFI* (Cummings *et al.*, 1996).

All'interno di un altro introne, il 27b, anch'esso relativamente esteso (45-50 kb), sono inseriti tre piccoli geni che codificano sul filamento complementare in direzione opposta a *NFI* (Upadhyaya *et al.*, 1994). *EVI2A* e *EVI2B* sono due geni che rappresentano gli omologhi umani di geni murrini attivati da inserzione retrovirale e sono coinvolti nella genesi di leucemie murrine. Nell'uomo *EVI2A* è espresso nel cervello, nel midollo osseo e nel sangue periferico, mentre *EVI2B* nel

midollo osseo e nel sangue periferico. Il terzo gene *OMGP* (*oligodendrocyte myelin glycoprotein*), codifica per una glicoproteina mielinica oligodendrocitica presente nelle cellule di Schwann e negli oligodendrociti. Tale proteina ha effetti oncosoppressori e può regolare negativamente le vie di segnale mitogeniche in modo correlato a quello della neurofibromina. *OMGP* è stato mappato entro 4 kb dal punto della traslocazione t(1;17), mentre *EVI2A* e *EVI2B* si trovano rispettivamente circa 20 e 5 kb da *OMGP* (Shen *et al.*, 1996).

E' stato ipotizzato un ruolo per questi geni nella patogenesi della Neurofibromatosi di tipo 1; si presume infatti che elevati livelli nucleari dei loro trascritti possano permettere l'ibridazione alla corrispondente sequenza nel trascritto complementare, non processato, del gene *NF1* e agiscano quindi come antisense naturale interferendo quindi con la processazione dell'RNA e con la sua traduzione.

Vi è inoltre, nell'introne 37, uno pseudogene della adenilato chinasi, *AK3*, che codifica nella stessa direzione di *NF1* (Li *et al.*, 1995; Shen *et al.*, 1996).



**Figura 2.** Organizzazione del gene *NF1* in esoni. 9a, 23a e 48a sono gli esoni aggiuntivi osservati nelle diverse isoforme del trascritto, dovute a *splicing* alternativo nel gene.

## 1.4 Il prodotto genico: la neurofibromina

Il prodotto della traduzione del trascritto principale è una proteina idrofilica: la neurofibromina, che presenta una sequenza primaria prevista di 2818 amminoacidi ed una massa molecolare stimata in 327 kDa. La massa molecolare osservata di questa proteina è di circa 220 kDa, una dimensione inferiore a quella prevista sulla base dell'*open reading frame* del gene. Questa discrepanza sembra essere dovuta a fenomeni di *protein folding* durante l'elettroforesi, visto che non si è a conoscenza di eventi di glicosilazione post-traduzionale o processazione della proteina stessa (Shen *et al.*, 1996).

Analisi della sequenza amminoacidica della neurofibromina hanno indicato la presenza di un dominio centrale, di circa 360 residui amminoacidici, che

evidenzia una forte omologia di sequenza al dominio catalitico delle proteine appartenenti alla superfamiglia delle GAP (*GTPase activating protein*), compresa la p120-GAP dei mammiferi, le ira1, ira2 e sar1 di lievito e la Gap-1 di *Drosophila*. Questo dominio isolato, così come la struttura completa della proteina, dimostra una attività *ras*-GAP *in vitro* e *in vivo*; questa regione è stata denominata *GAP-related domain* (GRD). Il dominio GRD è localizzato nella porzione di gene compresa tra gli esoni 21 e 27a e rappresenta circa il 10% dell'intera molecola (Hattori *et al.*, 1992)



**Figura 3.** Struttura secondaria tridimensionale del dominio GRD della neurofibromina.

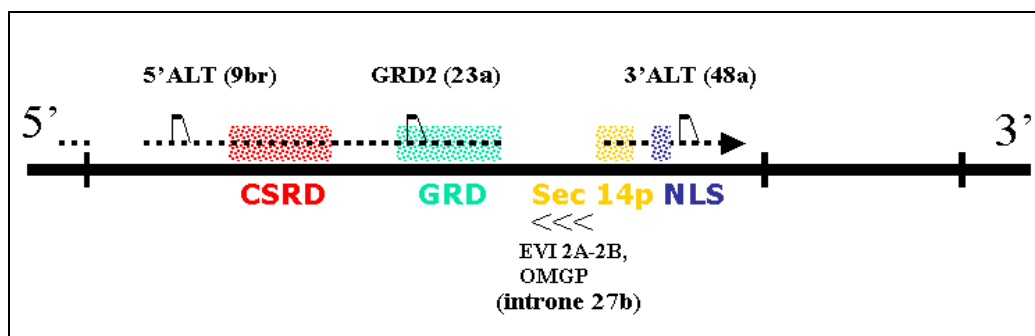
Esiste un segmento nel gene *NF1*, a monte di GRD, tra gli esoni 11 e 17, che è sede di numerose mutazioni. Se il raggruppamento di mutazioni nel dominio GRD è coerente con la sua rilevanza funzionale, si può assumere che questa regione a monte, definisca un secondo dominio funzionale. Infatti, l'analisi della distribuzione delle mutazioni note, pesata per ogni esone, evidenzia che anche tra gli esoni 11 e 17 si concentrano mutazioni missenso e delezioni *in frame* (Mattocks *et al.*, 2004).

Questa regione coincide con un dominio ricco di cisteina e serina (CSRD, *cysteine-serine rich domain*) già descritto all'interno della neurofibromina (Izawa *et al.*, 1996). Il dominio CSRD si estende dal residuo amminoacidico 543 al



residuo 909; tra questi vi sono tre coppie di cisteina che potrebbero essere compatibili con altrettanti siti di legame per ATP. Inoltre, questa sequenza ospita tre potenziali siti di riconoscimento per la proteina chinasi A (PKA) cAMP-dipendente. Essi sono soggetti a fosforilazione da parte di PKA e uno dei tre è posizionato in una breve regione con sequenza simile a quella di proteine che si associano ai microtubuli. Questa osservazione porta a formulare l'ipotesi che la fosforilazione in questa zona possa regolare l'interazione tra neurofibromina e microtubuli (Fashold *et al.*, 2000).

Vi sono inoltre due regioni considerate possibili domini funzionali nella neurofibromina: una regione compresa tra gli esoni 27b e 28, definita dominio *Sec14p homology* (Sec14p), che presenta omologia con i domini di legame di lipidi come fosfatidil-colina e fosfatidil-inositolo (Aravind *et al.*, 1999); una regione definita *nuclear localization signal* (NLS) nell'esone 43 (Vandenbroucke *et al.*, 2004).



**Figura 4.** Schema della sequenza del gene *NF1*; 5' ALT, GRD II e 3' ALT sono isoforme del trascritto; NLS, segnale di localizzazione nucleare; Sec 14p, dominio *Sec14p homology*; *EVI2A-2B* e *OMGP*, geni codificati sul filamento complementare all'interno dell'introne 27b.

### 1.4.1 *Splicing* alternativo

Il gene *NF1* viene trascritto in più isoforme, in seguito a *splicing* alternativo, che sono espresse secondo un *pattern* tessuto e stadio di sviluppo specifico. In aggiunta all'isoforma di 13 kb originariamente riportata, chiamata tipo I o GRD I, è stato identificato un trascritto alternativo, presente indifferentemente in individui affetti e non affetti, denominato tipo II o GRD II.

Esso contiene un'inserzione di 63 bp tra gli esoni 23 e 24 (esone 23a), all'interno del dominio GRD. La composizione amminoacidica dell'inserito di GRD II è conservata attraverso le specie; condivide caratteristiche con i segnali di localizzazione nucleare e codifica una sequenza di amminoacidi basici, che prevede la presenza di 6 lisine su 21 residui. Un alto contenuto di lisine potrebbe alterare significativamente la conformazione del peptide, in una regione presumibilmente critica per l'interazione con la proteina p21<sup>ras</sup> (Viskochil *et al.*, 1993). Questo spiega come tale forma alternativa di GRD dimostri una drastica riduzione nella sua attività catalitica di GAP (se comparata con GRD I) e una inferiore capacità di legame con la tubulina. Le cellule cerebrali normali differenziate esprimono in modo predominante l'isoforma di tipo I, mentre cellule cerebrali tumorali esprimono prevalentemente il tipo II, suggerendo che alterazioni del rapporto tra l'espressione di tipo I e tipo II, a favore di quest'ultima, sembrano essere associate alla genesi di tumori del sistema nervoso centrale.

Un'altra isoforma, con un'inserzione di 54 bp (esone 48a) vicino all'estremità 3' del gene, chiamata 3' ALT o tipo III, ha un alto livello di espressione nel muscolo cardiaco, scheletrico e liscio.

Un trascritto di circa 2,9 kb, 5' ALT-1, possiede una ORF che codifica per una proteina di 551 amminoacidi, che condivide i primi 547 residui aminoterminali con il prodotto originale di tipo I, ed è priva del dominio GRD (Shen *et al.*, 1996).

Un'ulteriore isoforma, 5' ALT-2 contiene un esone extra di 30 bp (esone 9br o esone 9a) tra gli esoni 9 e 10, che aggiunge 10 residui tra gli amminoacidi 420 e 421. E' stata identificata nell'uomo in elevate quantità nel sistema nervoso periferico e specialmente nei neuroni localizzati nella parte anteriore del cervello, in un *pattern* correlato alla maturazione post-natale dei neuroni, implicando così questa isoforma nel differenziamento neuronale (Gutmann *et al.*, 1995).

## **1.4.2 Distribuzione tissutale e localizzazione subcellulare**

Il gene *NF1* è espresso ubiquitariamente, nei mammiferi, nelle primissime fasi dello sviluppo. La neurofibromina è presente nella maggior parte dei tessuti; tuttavia, negli stadi più avanzati dello sviluppo e negli adulti, la sua espressione diventa rilevabile per lo più in neuroni, cellule di Schwann e oligodendrociti (Daston *et al.*, 1992).

Le isoforme del trascritto di *NF1*, ottenute da *splicing* alternativo, sono state investigate per determinare il loro *pattern* di espressione in tessuti differenti. I trascritti di tipo I e II sono risultati rappresentati in tutti i tessuti esaminati, ma con una variazione nel rapporto tra le due isoforme: è stato osservato che la loro espressione varia in relazione allo stato di differenziazione di un particolare tessuto.

La disponibilità di anticorpi contro porzioni della proteina ha reso possibile la localizzazione subcellulare della neurofibromina. La colorazione immunostochimica ha rilevato la presenza citoplasmatica della proteina in vari tessuti. Ulteriori studi hanno dimostrato la co-localizzazione della neurofibromina con i microtubuli citoplasmatici (Shen *et al.*, 1996).

## **1.5 Funzioni della neurofibromina**

### **1.5.1 Attività GAP**

La prima GAP di mammifero ad essere caratterizzata è stata una proteina citosolica con una massa molecolare di 120 kDa, p120-GAP, che è stata identificata per la sua proprietà di stimolare la conversione della forma attiva delle proteine *ras*, legata a GTP, alla forma inattiva, legata a GDP. Tutte le proteine simili alle GAP (dunque, in via teorica, anche il dominio GRD della neurofibromina) possono regolare negativamente l'attività delle proteine *ras* normali (McCormick, 1989).

Tra più di 30 geni *ras*, tre, strettamente correlati, sono conosciuti come *H-ras*, *K-ras* e *N-ras* e codificano una proteina di 21 kDa,  $p21^{ras}$ , la quale lega i nucleotidi guanidinici con alta affinità e possiede una intrinseca attività GTPasica: è attiva nello stato legato a GTP, spenta nello stato legato a GDP. Questa sorta di interruttore molecolare gioca un ruolo critico nel controllo della crescita e della differenziazione cellulare; mutazioni attivanti in *ras* portano a un segnale aberrante di proliferazione cellulare e sono coinvolte nella genesi di molti tumori maligni umani.

Analisi biochimiche hanno dimostrato che il dominio GRD della neurofibromina (NF1-GRD) ha la capacità di accelerare la conversione della  $p21^{ras}$  attiva alla forma inattiva, stimolando un aumento nel tasso di idrolisi del GTP, e in questo modo ne riduce l'attività. Tuttavia, sebbene sia le GAP umane sia NF1-GRD possono legare i mutanti oncogenici di  $p21^{ras}$ , nessuna è in grado di incentivarne l'attività GTPasica.

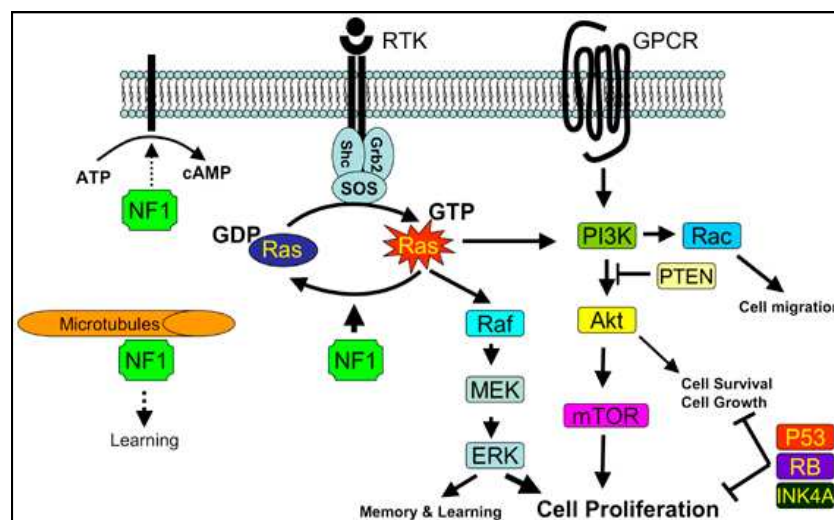
In aggiunta, o in alternativa, la neurofibromina può funzionare da effettore a valle di  $p21^{ras}$ , ovvero come effettore allosterico positivo che, complessato alla proteina inattiva, trasmette nel citoplasma il segnale di inibizione della crescita cellulare. In realtà, fino ad ora nessuna prova diretta ha stabilito questo ruolo per la neurofibromina, ma, per l'analogia con p120-GAP, è probabile che lo ricopra. Sia p120-GAP sia NF1-GRD sono in grado di legarsi con efficienza alle  $p21^{ras}$  oncogeniche costitutivamente segnalanti, ma interagiscono poco con i componenti successivi nella via di trasmissione del segnale, che viene così alterata.

NF1-GRD e p120-GAP sono diverse per l'affinità e per il legame a  $p21^{ras}$  e per l'attività specifica. Come regolatori negativi, sia la neurofibromina sia p120-GAP disattivano la proteina bersaglio; come effettori della propagazione del segnale a valle, invece, è probabile che queste due proteine siano abbastanza differenti. Ad ogni modo, a basse concentrazioni del prodotto dei geni *ras*, le attività di NF1-GRD e di p120-GAP sono comparabili. La neurofibromina, quindi, può essere un regolatore significativo dell'azione di  $p21^{ras}$ , in particolare a basse concentrazioni di quest'ultima.

NF1-GRD e le GAP umane mostrano inoltre differenze nella sensibilità

all'inibizione da lipidi: il dominio GRD è inibito da alcuni lipidi, la p120-GAP non lo è. Questa inibizione selettiva può essere sfruttata per distinguere i due tipi di meccanismi attivanti la GTPasi nei tessuti umani (Shen *et al.*, 1996).

Un'ulteriore funzione della neurofibromina è quella di regolare la crescita cellulare attraverso un meccanismo indipendente dalla sua funzione GAP. E' stato osservato che, nei neuroblastomi, nelle cellule di melanoma e nella leucemia mieloide, le mutazioni di *NF1* associate ad una espressione difettiva o ridotta della neurofibromina causano oncogenesi senza alterare i livelli delle proteine *ras* attive, suggerendo così che la funzione della neurofibromina in questi casi sia indipendente dalla regolazione dell'attività di *ras* (Lakkis *et al.*, 2000). Va ricordato, tuttavia, che la risposta ad aumentati livelli di *ras* cambia a seconda del tipo cellulare, portando alla crescita oppure alla differenziazione. Poiché il significato di un segnale *ras* potenziato è variabile, i meccanismi interpretativi del sistema di trasduzione devono trovarsi a valle del componente *ras* (Viskochil *et al.*, 1993).



**Figura 5.** Schema grafico dell'attività GAP della neurofibromina e delle diverse vie cellulari correlate (www.nature.com).

### 1.5.2 Neurofibromina e tubulina

E' stato dimostrato che la tubulina è un altro inibitore dell'attività GAP, sia

del peptide NF1-GRD sia della neurofibromina completa (Bollag *et al.*, 1993), mentre non influenza l'azione della p120-GAP. Alcuni siti di legame alla tubulina sono stati localizzati in un segmento di 80 residui in posizione N-terminale al GRD; è stato inoltre osservato che la tubulina e i fosfolipidi sembrano interagire con porzioni distinte della neurofibromina. La relazione osservata tra la neurofibromina e la tubulina è coerente con la co-localizzazione neurofibromina-microtubuli. La tubulina, infatti, è il componente primario dei microtubuli e, poiché questi rivestono un ruolo importante in molti aspetti dell'organizzazione cellulare, l'interazione tra neurofibromina e tubulina può essere implicata quasi in qualsiasi funzione cellulare. È stato ipotizzato che la neurofibromina sia un collegamento tra la tubulina e il regolatore della crescita *ras* (Bollag *et al.*, 1993).

### **1.5.3 Neurofibromina e cAMP**

Studi in *Drosophila* indicano che la neurofibromina potrebbe svolgere anche funzioni indipendenti dalla regolazione di *ras*, che coinvolgerebbero il controllo del cAMP intracellulare (Guo *et al.*, 1997). Negli astrociti, l'aumento della concentrazione interna di cAMP determina un forte segnale di inibizione della crescita poiché porta all'attivazione di Rap1, una GTPasi antagonista anti-mitogena della via di segnale di *ras* (Dugan *et al.*, 1999). La perdita o il deficit della neurofibromina negli astrociti risulta in un'attenuazione della produzione di cAMP (Dasgupta *et al.*, 2003); la neurofibromina, infatti, è richiesta per la generazione di cAMP indotta da PACAP (l'adenilato ciclasi pituitaria) e agisce a livello dell'attivazione della adenilato ciclasi. Inoltre, le vie di trasmissione del segnale a valle, innescate da cAMP negli astrociti, sono parimenti alterate dalla mancanza della neurofibromina, che ha come conseguenza, quindi, anche una diminuzione dell'influsso nelle cellule di calcio e dell'attivazione di Rap1.

Ma se negli astrociti cAMP agisce da segnale anti-mitogeno, nelle cellule di Schwann esso funge da attivatore della divisione cellulare e la neurofibromina piuttosto che stimolare la produzione di cAMP, sembra inibirla (Kim *et al.*, 2001). Considerando il fatto che i livelli di cAMP regolano la crescita e il

differenziamento delle cellule cerebrali, c'è da aspettarsi che la neurofibromina possa essere coinvolta in altri importanti meccanismi di trasmissione dei segnali intracellulari.

#### **1.5.4 Neurofibromina e chinasi**

La neurofibromina può subire la regolazione da parte di proteine chinasi. Analisi biochimiche, infatti, hanno dimostrato la presenza di almeno sei potenziali siti di riconoscimento per una serina-treonina chinasi cAMP-dipendente (ruolo per cui la PKA è il candidato più accreditato) e uno per tirosina chinasi. Tutti questi siti risultano pesantemente fosforilati sia costitutivamente sia in risposta a fattori di crescita epidermici e derivati dalle piastrine (EGF e PDGF) (Izawa *et al.*, 1996). I siti di fosforilazione per PKA situati all'interno del dominio CSRD suggeriscono che questa regione potrebbe essere coinvolta nella trasmissione del segnale mediata da cAMP (Fashold *et al.*, 2000). Il significato funzionale della fosforilazione della neurofibromina da parte della PKA non è stato ancora del tutto chiarito, ma l'evidenza del coinvolgimento della neurofibromina nel *signaling* mediato dal secondo messaggero cAMP potrebbe spiegare il ruolo della neurofibromina nel collegamento tra la via di segnale di PKA e quella di *ras*.

### **1.6 Il gene *NF1* è un oncosoppressore**

L'identificazione di oncogeni e oncosoppressori ha confermato il coinvolgimento di componenti genetici nell'insorgenza dei tumori. L'ipotesi più accreditata risulta essere quella del *two hits* di Knudson per geni oncosoppressori (Knudson, 1971), che implica l'inattivazione biallelica del gene in due eventi separati.

Il gene *NF1* è considerato un oncosoppressore e la teoria più valida per spiegare l'insorgenza di tumori nei pazienti NF1, si basa sul fatto che la

neurofibromina è un regolatore negativo di *ras*. Infatti, l'attività GAP ridotta, dovuta a proteina difettosa o assente, porta a livelli aumentati di p21<sup>ras</sup> nel suo stato attivo, che a sua volta genera *signaling* anormale e oncogenesi. Mutazioni che inattivano *ras*, o la rendono resistente alle GAP, sono infatti implicate in più del 30% dei tumori umani (Viskochil *et al.*, 1993). Molte mutazioni di *NF1* nella regione GRD, che inattivano o modificano l'attività GAP della neurofibromina, sono state riportate in diversi cancro correlati a NF1, fornendo una prova diretta che il danno all'attività GAP della proteina è critico nella patogenesi della Neurofibromatosi di tipo 1 (Upadhyaya *et al.*, 1997). Inoltre, sono stati riscontrati elevati livelli di *ras*-GTP attiva in tumori dei nervi periferici in pazienti NF1, in associazione alla perdita di espressione della neurofibromina.

## 1.7 Le mutazioni nel gene *NF1*

Le analisi del gene *NF1* per la ricerca di mutazioni sono ostacolate da molti fattori (Lakkis *et al.*, 2000):

- a) le grandi dimensioni del gene rendono difficile la scansione dell'intera sequenza;
- b) la presenza dell'allele normale e l'esistenza nel genoma di molti pseudogeni *NF1* interferiscono nelle analisi con PCR e *Southern blot*, a causa della co-amplificazione delle sequenze omologhe dei *loci* pseudo-*NF1*;
- c) la mancanza di *hot spots* mutazionali ben definiti, dato che le mutazioni insorgono abbastanza casualmente lungo tutto il gene *NF1*;
- d) l'ampia variazione nel tipo e nelle dimensioni delle mutazioni richiede l'utilizzo di molti diversi e complementari metodi di indagine;
- e) nessuna apparente correlazione genotipo-fenotipo è stata stabilita con certezza, che possa aiutare a prevedere il tipo o il sito della mutazione, con l'eccezione degli individui con microdelezioni, che mostrano un fenotipo



più grave (Kayes *et al.*, 1994).

Le mutazioni a carico del gene *NF1* risultano essere molteplici e di diversa tipologia, tra le quali estese delezioni genomiche che rimuovono l'intero gene e le sue regioni fiancheggianti (più di 1,5 Mb, responsabili di meno del 5% dei casi di NF1) oppure sostituzioni di singole basi (circa il 67% delle variazioni riscontrate) (Mattocks *et al.*, 2004).

E' possibile escludere che la causa principale dell'alto tasso di mutazioni misurato sia la presenza dei numerosi pseudogeni; la responsabilità maggiore si può imputare alle grandi dimensioni del gene: il tasso di mutazioni, infatti, è paragonabile, anche se più elevato, a quello di altri geni di dimensioni simili (Fashold *et al.*, 2000).

In tabella 1 sono riportate le mutazioni attualmente note (dato aggiornato a Gennaio 2010) secondo l'*Human Genome Mutation Database* (HGMD, [www.hgmd.org](http://www.hgmd.org)).

<b>Tipo di mutazione</b>	<b>n° di mutazioni</b>
Sostituzioni nucleotidiche (missenso / non senso)	220
Sostituzioni nucleotidiche in siti di <i>splicing</i>	160
Sostituzioni nucleotidiche in siti regolatori	0
Piccole delezioni	238
Piccole inserzioni	113
Piccole indels	14
Grosse delezioni	77
Grosse inserzioni	8
Riarrangiamenti complessi	8
Variazione dei repeat	0
<b>Totale</b>	<b>838</b>

**Tabella 1.** Mutazioni note del gene *NF1* riportate in HGMD.

In generale, lo spettro di mutazioni del gene *NF1* sembra molto simile a

quello di altri oncosoppressori, caratterizzati da una proporzione rilevante di mutazioni troncanti.

Le mutazioni missenso sembrano raggrupparsi in due regioni: nel dominio GRD e in quello CSRD. Una percentuale rilevante (circa il 39%) (Fashold *et al.*, 2000) è costituita da sostituzioni con prolina, un amminoacido conosciuto per la sua proprietà di destabilizzare le strutture secondarie, come  $\alpha$ -eliche e foglietti  $\beta$ .

Buona parte delle mutazioni riscontrate sono puntiformi, oppure altri tipi di piccole lesioni accomunati dalla caratteristica di introdurre prematuri codoni di terminazione (PTCs) della sintesi proteica (Fashold *et al.*, 2000).

La quota di mutazioni *frameshift* e non senso che ci si aspetta interrompano il codice di lettura, si aggira attorno al 49% delle mutazioni totali e sembrano essere omogeneamente distribuite lungo tutta la sequenza di *NF1*, fatta eccezione per i tre esoni originati dallo *splicing* alternativo, 9a, 23a, e 48a, finora apparentemente privi di mutazioni (Fashold *et al.*, 2000; Messiaen *et al.*, 2000; Han *et al.*, 2001; Ars *et al.*, 2003).

Rilevante è l'effetto che le mutazioni comportano a livello dell'mRNA: circa la metà delle mutazioni causano alterazioni nello *splicing*; i difetti di processamento del messaggero si propongono come il difetto molecolare più comune nella *NF1* (Mattocks *et al.*, 2004). Le alterazioni possono interessare sia i canonici siti donatori (GT) e accettori (AG), sia posizioni meno conservate, oppure possono creare nuovi siti di *splicing* (ss). Mutazioni al 5'ss e al 3'ss non sembrano avere la stessa frequenza: il 65% delle mutazioni delle sequenze *consensus* dei siti di *splicing* coinvolge il sito donatore, in particolare in posizione +1, +2 e +5 (Pros *et al.*, 2008). Inoltre, circa la metà delle mutazioni a carico dello *splicing* dell'mRNA indurrebbero la formazione di trascritti *in frame* che risulterebbero in una proteina presunta leggermente più corta o più lunga dell'autentica neurofibromina, ma non troncata (Ars *et al.*, 2003).

Appare probabile che tutte le regioni del gene *NF1* siano soggette a mutazioni allo stesso modo, ma, se i dati relativi alla presenza di alterazioni in un determinato esone vengono pesati rispetto alla percentuale di sequenza codificante che occupa, alcune regioni risulteranno siti a notevole densità di mutazione, ad

esempio gli esoni 4b e 37 (Fashold *et al.*, 2000) e 10a, 10b e 10c (Messiaen *et al.*, 2000). Il test  $\chi^2$ , applicato per comparare la frequenza osservata di mutazioni rispetto a quella attesa per ogni esone, ha dimostrato che, effettivamente, ci sono aree del gene *NFI* che hanno una maggiore tendenza ad accumulare alterazioni (Ars *et al.*, 2003). Una spiegazione potrebbe trovarsi nella presenza in tali esoni di alcuni elementi strutturali inclini alla mutazione, come brevi sequenze ripetute in tandem (esone 4b) o elementi quasi simmetrici (esone 37). Pertanto, data la ricorrenza di mutazioni missenso e delezioni *in frame*, le regioni GRD e CSRD rappresenterebbero degli *hot spot* mutazionali, coerentemente con il loro importante ruolo funzionale (Mattocks *et al.*, 2004).

Se è vero che la maggioranza delle mutazioni sono uniche, vale a dire sono presenti in un solo paziente (Fashold *et al.*, 2000), è anche vero che le mutazioni ricorrenti sono più comuni (40% circa delle mutazioni identificate) di quanto precedentemente descritto (Ars *et al.*, 2003). Le ragioni della ricorrenza potrebbero essere le seguenti (Messiaen *et al.*, 2000):

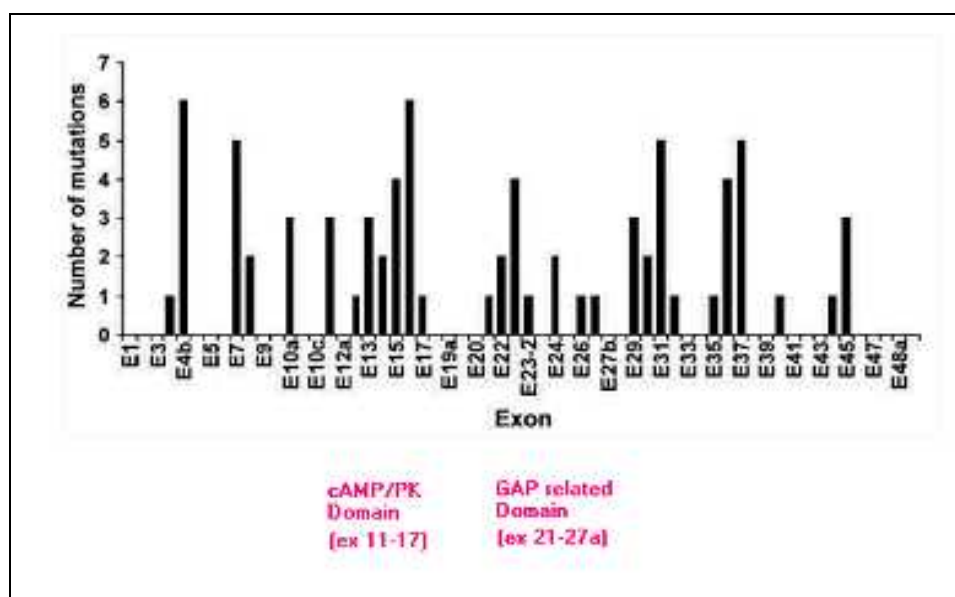
- a) transizioni (C>T e G>A) ai dinucleotidi CpG;
- b) slittamento delle polimerasi in una serie di basi uguali o in zone contenenti sequenze ripetute o palindrome.

Nei geni umani, le mutazioni non avvengono casualmente rispetto alle sequenze di DNA circostanti e questa regola è valida anche per le mutazioni osservate in *NFI*, dove la composizione della sequenza influenza la formazione di alterazioni. Ad esempio, sono stati individuati almeno 118 dinucleotidi CpG (Fashold *et al.*, 2000); questi siti mostrano un alto tasso di mutazione, dovuto alla deaminazione spontanea della 5-metilcitosina: la metilazione, dunque, è un prerequisito indispensabile per questo tipo di mutazioni. Circa il 41% (Messiaen *et al.*, 2000) delle sostituzioni di una singola base sono dovute a transizioni C>T o G>A in dinucleotidi CpG. Inoltre, particolari *pattern* di sequenza, come ripetizioni dirette, sequenze palindrome e quasi palindrome, elementi simmetrici e serie di basi identiche, sono associati alla formazione preferenziale di inserzioni e/o delezioni.

Uno studio eseguito da De Luca *et al.* (2004) ha previsto l'analisi

molecolare di 110 pazienti non correlati. Sono state evidenziate mutazioni nel 68% dei pazienti analizzati, di cui 22 de novo e 23 ricorrenti (a conferma dell'ipotesi che la presenza di mutazioni ricorrenti non è poi così inusuale in affetti da NF1). Il 53% di queste mutazioni ricorrenti è dovuto a transizioni C>T o G>A in dinucleotidi CpG. (De Luca *et al.*, 2004). Dai risultati ottenuti si ricava che, nonostante le mutazioni siano distribuite lungo tutta la sequenza genica, si possono comunque evidenziare dei *cluster* di mutazioni: il 68% delle mutazioni osservate infatti sono risultate localizzate in 12 esoni e/o introni fiancheggianti: 4b, 7, 10a, 11, 15, 16, 23-1, 29, 31, 36, 37, e 45.

La maggior parte delle mutazioni missenso sono risultate a carico dei due domini più importanti della neurofibromina: sugli esoni 11-17 (dominio CSRD) e sugli esoni 21-27a (dominio GRD)



**Figura 6.** Distribuzione delle mutazioni lungo la sequenza del gene *NF1* (De Luca *et al.*, 2004).

Un' ulteriore studio di analisi molecolare sul gene *NF1* in 374 pazienti riporta l'identificazione di 282 differenti mutazioni distribuite uniformemente lungo tutto il gene, con l'eccezione di pochi esoni: 9a, 23a, 47, 48a e 49 (Pros *et al.*, 2008).

## 1.8 Espressione fenotipica

Non è tutt'ora possibile stabilire una correlazione genotipo-fenotipo sulla base del tipo di mutazione osservata o sulla sede della mutazione. Infatti mutazioni identiche possono non causare lo stesso fenotipo in pazienti diversi sia che siano correlati o meno.

L'unica correlazione sembra essere quella tra la presenza di microdelezioni che coinvolgono l'intero gene e le sue sequenze fiancheggianti (hanno una frequenza del 5% e sono evidenziabili tramite ibridazione *in situ* fluorescente) ed un fenotipo particolarmente grave che colpisce tali pazienti.

La variabilità del fenotipo non può essere spiegata solo con l'aploinsufficienza del gene *NF1*, si pensa vi sia un'associazione anche con la perdita di funzione dei geni contigui, nel caso delle microdelezioni, e/o dei tre geni *EVI2A*, *EVI2B* e *OMPG* anche se non è ancora noto se questi abbiano un ruolo nel determinare l'espressione fenotipica di *NF1* (Shen *et al.*, 1996). La spiegazione più accreditata per spiegare questa variabilità implica il possibile coinvolgimento di geni modificatori, presenti nello stesso locus o in loci vicini ad *NF1* (Wiest *et al.*, 2003).

## 1.9 Analisi di mutazione nel gene *NF1*

Le mutazioni a carico del gene *NF1* sono molto varie per tipologia e dimensioni: variano da grandi delezioni (>15 Mb) a sostituzioni di una singola base e non sembrano esserci veri e propri *hot spot* mutazionali lungo la sequenza genica. Questo rende difficile la ricerca di mutazioni nel gene e non ci si può aspettare che una singola tecnica possa evidenziarle tutte.

Per l'analisi del gene *NF1* sono state utilizzate numerose metodiche, tra le quali: *Single Strand Conformation Polymorphism* (SSCP), *Heteroduplex Analysis* (HA), *Denaturing Gradient Gel Electrophoresis* (DGGE), *Protein Truncation Test*

(PTT) e *Direct Genome Sequencing* (DGS), ma indipendentemente dalla metodica applicata, l'efficienza nell'evidenziare le mutazioni non ha mai superato il 50% dei pazienti analizzati.

Alcuni lavori hanno portato all'identificazione di circa il 95% delle mutazioni, nel campione in analisi, eseguendo uno studio che prevedeva l'uso combinato di diverse tecniche: PTT, HA, FISH (*fluorescent in situ hybridization*), *Southern blotting* e analisi citogenetiche; le variazioni di sequenza sono state studiate, oltre che a livello genomico, anche a livello dell'RNA (Messiaen *et al.*, 2000).

Un'altro studio di mutazione sul gene *NF1* effettuato su un notevole numero di casi (521 pazienti non correlati), ha riportato che circa l'80% delle mutazioni osservate causavano la formazione di un codone di stop prematuro con la conseguente sintesi di una proteina inefficiente (Fahsold *et al.*, 2000).

Ars *et al.* (2003) hanno eseguito uno studio su 474 pazienti affetti da NF1; su 189 mutazioni osservate circa il 45% era costituito da mutazioni ricorrenti.

Pros *et al.* (2008) riportano invece che di 282 mutazioni identificate su 374 pazienti non correlati, solo 37 si sono rivelate ricorrenti (il 13%).

Questi risultati possono essere importanti per lo sviluppo di procedure di diagnostica atte a riscontrare questo tipo di mutazioni.

Dal punto di vista diagnostico fino ad ora le metodiche più accreditate per uno studio del gene *NF1* sembrano essere:

- l'analisi FISH per lo studio di microdelezione, soprattutto in quei soggetti che presentano un particolare fenotipo associabile a tale mutazione (Kayes *et al.*, 1994);
- l'utilizzo della metodica MLPA (*Multiplex Ligation-dependent Probe Amplification*) per l'esecuzione di un'analisi di tipo quantitativo su tutta la lunghezza del gene per evidenziare inserzioni e/o delezioni che coinvolgono porzioni di gene o singoli esoni;
- l'analisi DHPLC (*Denaturing High Performance Liquid Chromatography*)

per lo studio molecolare del gene;

- infine, una metodica sviluppata recentemente: l'HRMA (*High-Resolution Melting Analysis*) che può essere utilizzata per individuare, quantificare e genotipizzare campioni con una *detection rate* delle mutazioni comparabile o superiore alle tecniche disponibili attualmente per lo *screening* (White *et al.*, 2006).

## **1.10 Elementi di regolazione dello *splicing***

Lo *splicing* del pre-mRNA è un evento critico nell'espressione dei geni eucariotici: è un processo complesso che consente la giunzione delle parti codificanti le proteine (esoni) che sono separate da sequenze non codificanti (introni).

E' considerato uno dei principali meccanismi di generazione della flessibilità e diversità genomica che porta alla formazione di proteine diverse a partire da uno stesso gene. Le recenti conoscenze sulla complessità del processo di *splicing* (Matlin *et al.*, 2005) hanno rivelato l'esistenza di importanti elementi di regolazione dello stesso. Variazioni di questi elementi, che si trovano sia nelle regioni codificanti che non codificanti dei geni, possono manifestarsi con effetti deleteri sullo *splicing* del pre-mRNA

Un tipico gene di mammifero è composto da numerosi esoni relativamente piccoli separati da sequenze introniche molto più lunghe. La corretta maturazione dell'mRNA, cioè il corretto riconoscimento delle regioni esoniche e la loro unione, è un processo che richiede l'assemblamento di un complesso proteico (lo spliceosoma) derivante dall'azione coordinata di 5 piccoli snRNPs (*small nuclear ribonucleoproteins*: U1,U2,U4,U5 e U6) e di più di 60 polipeptidi.

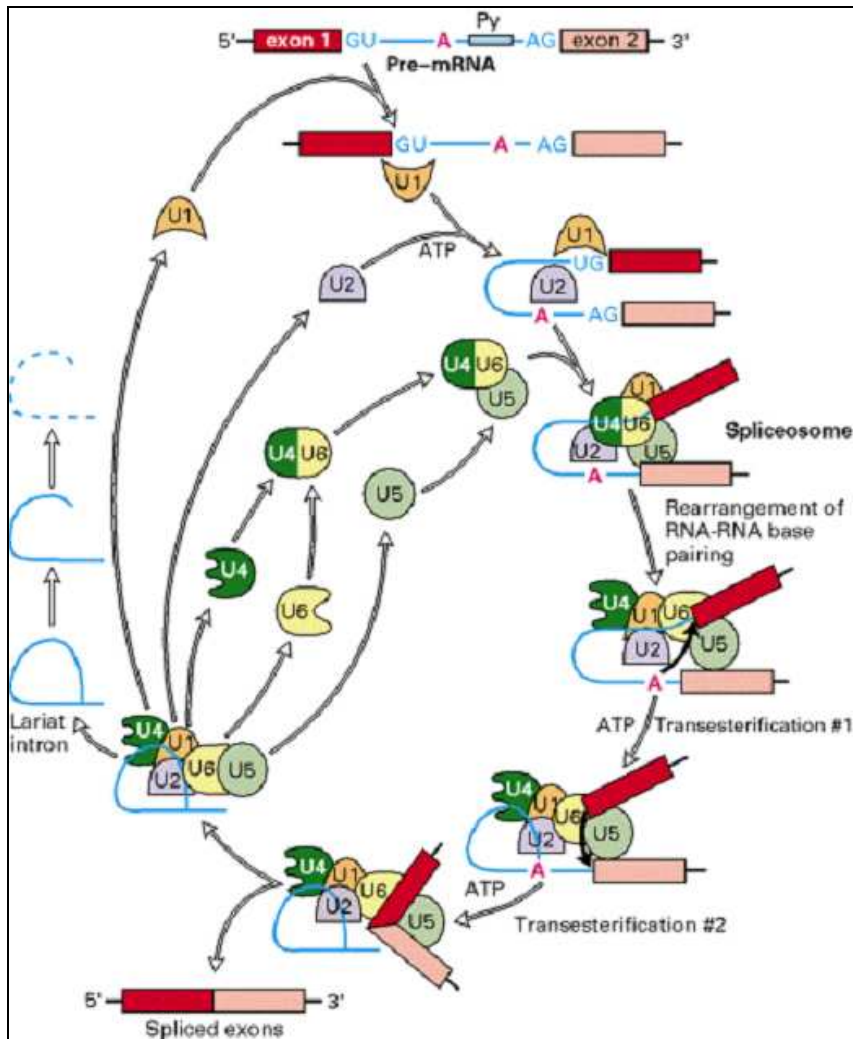


Figura 7. Schematizzazione del meccanismo di *splicing*.

Un errato riconoscimento dei confini tra esone e introne, o il fallimento della rimozione di uno o più esoni, porta alla formazione di mRNA aberranti che possono essere instabili e quindi venir degradati oppure può portare alla formazione di isoforme proteiche difettose o addirittura dannose.

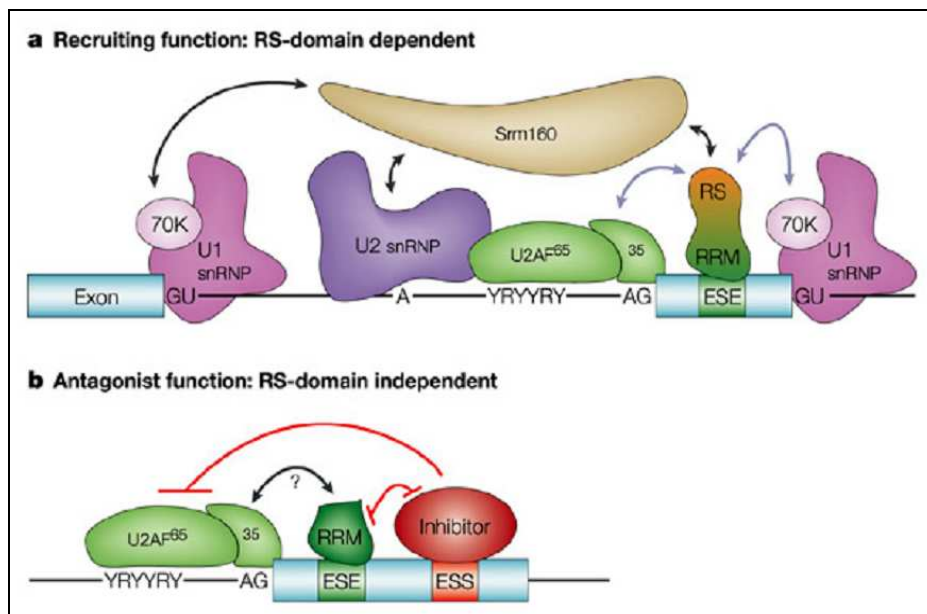
Gli elementi di regolazione dello *splicing* sono:

- Elementi in *cis*: sito di *splicing* al 5' (sito donatore)  
 sito di *splicing* al 3' (sito accettore)  
*branch site*  
*enhancers* (ESEs)





In particolare, gli *Exonic Splicing Enhancer* (ESE) sono corte sequenze di RNA (10-15 basi) che sembrano essere presenti nella maggior parte, se non in tutti, gli esoni. Risultano essere siti di legame per specifiche proteine ricche in serina-arginina (*serine/arginine-rich: SRprotein*); queste proteine sono dei fattori di *splicing* appartenenti ad una famiglia di proteine correlate strutturalmente e altamente conservate. Sono caratterizzate da due motivi: uno che riconosce l'RNA (RRM: *RNA-recognition motifs*) e l'altro C-terminale ricco di dipeptidi serina-arginina (*RS domain*). I domini RRM si legano a specifiche regioni dell'RNA (sequenze ESE), mentre il dominio RS è coinvolto nel legame con altre proteine e agisce da tramite tra la sequenza nucleotidica dell'acido ribonucleico e le proteine coinvolte nell'apparato di *splicing*. In particolare queste proteine RS si legano alle sequenze ESE con le loro subunità RRM agendo da segnale per il riconoscimento dell'esone e attraverso la subunità RS inviano l'*input* per l'aggregazione dello spliceosoma. Inoltre agiscono anche da antagonisti per possibili agenti inibitori del meccanismo come le sequenze ESS (*Exonic Splicing Silencer*) che sono siti legame per le proteine HnRNP che sono considerate dei regolatori negativi dello *splicing*. Sono meno conosciuti degli *enhancer*.



**Figura 9.** Modello d'azione delle proteine SR (Nature Reviews Genetics).

Riassumendo, la corretta divisione degli esoni dagli introni in un determinato mRNA dipende da due fattori:

- la forza intrinseca dei siti di *splicing* determinata dalle sequenze fiancheggianti GU e AG;
- la combinazione di effetti positivi e negativi dovuta alla presenza di *enhancers* e *silencers*.

E' stato ormai dimostrato che mutazioni puntiformi possono causare alterazioni dello *splicing*. La maggior parte di queste sono variazioni che provocano lo *skipping* dell'esone in prossimità della mutazione. In altri casi la mutazione può generare o attivare un "sito criptico" di *splicing* provocando profonde modificazioni del trascritto.

Le mutazioni responsabili possono essere sia esoniche (missenso, nonsenso e silenti) che introniche.

Queste variazioni possono influenzare il sottile equilibrio tra le isoforme di una proteina prodotte dallo *splicing* alternativo dello stesso gene e di conseguenza essere la causa di malattia (Caceres and Kornblihtt, 2002; Pagani and Baralle, 2004). Forse una delle ipotesi più accreditate di questi anni è che variazioni silenti del modulo di lettura o variazioni delle sequenze introniche sono seriamente considerate responsabili di patologie (Sun and Chasin, 2000; Cartegni *et al.*, 2002; Pagani *et al.*, 2003; Baralle and Baralle, 2005).

L'attivazione di pseudo-esoni o di esoni criptici in seguito alla creazione di nuovi siti di *splicing* o all'alterazione di siti costitutivi, sono eventi già osservati come causa di patologie da difetti dello *splicing* (Treisman *et al.*, 1983; Metherell *et al.*, 2001; Beroud *et al.*, 2004). Di recente nuovi tipi di mutazioni sono stati descritti come causa di malattie da difetti dello *splicing* (Pagani *et al.*, 2002; Lin *et al.*, 2005).

Tuttavia, le informazioni attualmente disponibili su tali patologie rivelano che le nostre conoscenze su molti dei meccanismi che regolano lo *splicing* sono ancora molto scarse.

I difetti di *splicing* costituiscono il 30-50% del totale di mutazioni del gene

*NFI*. Circa il 30% di queste mappa lontano dalle giunzioni esone-introne dove possono causare la creazione di nuovi siti di splicing (5'ss e 3'ss), attivare siti criptici, o portare alla perdita o formazione di sequenze segnale esoniche e/o introniche, importanti per il riconoscimento del messaggero da parte dei fattori di *splicing* (ESE e ISE) (Bottillo *et al.*, 2005).

Risultano quindi fondamentali studi approfonditi sul meccanismo di *splicing* del pre-mRNA e soprattutto sugli effetti che possono generare su questo le variazioni di sequenza osservate a livello genomico.

Un valido approccio sperimentale per eseguire test diagnostici sul trascritto, in quelle situazioni in cui non è possibile estrarre l'RNA dalle cellule del paziente, sembra essere la trasfezione transiente di cellule eucariote mediante minigeni plasmidici: si tratta di costrutti determinati da DNA ricombinante progettato artificialmente, nel quale la sequenza codificante di un gene d'interesse è associata con elementi di regolazione presenti nel vettore, che consentono l'espressione dell'inserito e la produzione del polipeptide desiderato.

## **Scopo della tesi**



Visto l'alto numero di mutazioni a carico del gene *NFI* che generano anomalie nel meccanismo di *splicing* (30-50%) (Messiaen *et al.*, 2000) e che è generalmente difficile poterne stabilire gli effetti solo con analisi a livello genomico, ci si è posti l'obiettivo di definire un protocollo di ricerca che prevede:

3. selezione di pazienti affetti da Neurofibromatosi di tipo 1 nei quali, mediante analisi sul genoma (*screening* del gene con DHPLC, HRMA e sequenziamento diretto degli esoni), siano state osservate variazioni che potenzialmente possono alterare il meccanismo di *splicing*;
4. analisi *in silico* delle sequenze mediante opportuni *software* (*BDGP Splice Site Prediction*, *NetGene 2*, *ESEfinder*, *Human Splicing Finder*) per prevedere gli effetti della modifica, formazione o scomparsa di sequenze regolatrici dello *splicing* in seguito ad alterazioni nucleotidiche della sequenza genomica;
- valutazione funzionale delle mutazioni che potenzialmente possono alterare il meccanismo di *splicing*, mediante trasfezioni transienti di cellule HeLa con un minigene plasmidico e saggio *in vitro* per comparare lo *splicing* in sequenze introniche e/o esoniche mutate rispetto a controlli *wild type*.





## **Materiali e metodi**



### **3.1 Campione in esame**

Il campione in oggetto è costituito da 310 individui inseriti nell'analisi di ricerca di mutazione nel gene *NF1*, in seguito a diagnosi clinica, o sospetta diagnosi, di Neurofibromatosi tipo 1.

L'anamnesi familiare e la successiva diagnosi sono state eseguite presso il Servizio di Genetica Clinica ed Epidemiologica del Dipartimento di Pediatria dell'Università degli Studi di Padova, secondo i protocolli diagnostici in uso presso il servizio.

Ad ogni paziente sono stati prelevati campioni di sangue periferico venoso conservato in EDTA. Ai pazienti, così come ai controlli sani coinvolti nella ricerca, è stato chiesto di fornire il consenso informato per l'analisi genetica e l'uso dei risultati per eventuali ricerche epigenetiche.

### **3.2 Estrazione del DNA genomico**

Il DNA genomico è stato estratto a partire da cellule nucleate di sangue intero periferico in EDTA conservato a  $-20^{\circ}\text{C}$ , secondo il protocollo del kit commerciale *High Pure PCR Template Preparation Kit* (Roche).

Per valutare la qualità e la quantità del DNA estratto è stata effettuata una corsa elettroforetica su gel d'agarosio allo 0,8% colorato con etidio bromuro, un intercalante della doppia elica che rivela la presenza di acido nucleico quando esposto alla luce ultravioletta.

Il DNA viene conservato in freezer a  $-20^{\circ}\text{C}$ .

### **3.3 Analisi di mutazione del gene NF1**

#### **3.3.1 *Denaturing High Performance Liquid Chromatography* (DHPLC)**

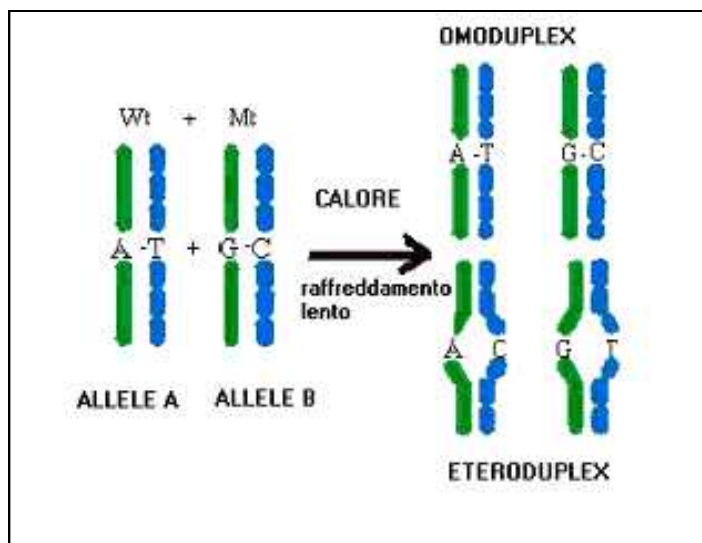
La DHPLC è stata introdotta fin dagli anni '90, per lo studio di mutazioni

puntiformi. E' una tecnica rapida e ad alta sensibilità per analisi di mutazioni di modeste dimensioni e attualmente sembra essere uno degli strumenti più efficaci per l'analisi di *screening* del gene *NF1* in pazienti affetti, o sospetti tali, da Neurofibromatosi tipo 1 (O'Donovan *et al.*, 1998).

Si basa sul principio della cromatografia liquida a scambio ionico in fase inversa. La fase stazionaria della struttura cromatografia è costituita da una matrice non porosa di polimeri C<sub>18</sub> alchilati di polistirene-divinilbenzene, idrofobica ed elettricamente neutra. I gruppi fosfato acidi delle molecole di DNA non sono in grado, quindi, di legarsi direttamente alla colonna. L'interazione con la matrice è resa possibile dal trietilammonio acetato (TEAA), un reagente carico positivamente. Un gradiente lineare crescente di acetonitrile permette l'eluizione dei frammenti di DNA, riducendo progressivamente il legame idrofobico tra le catene alchiliche del TEAA e la matrice di polistirene. La rivelazione dei frammenti eluiti avviene infine tramite un rivelatore a raggi UV che effettua lettura a 260 nm.

La ricerca di mutazione viene eseguita ad una temperatura critica, sequenza dipendente, che favorisce la parziale denaturazione dei filamenti in molecole che contengono *mismatch* tra le basi non complementari (*eteroduplex*).

Gli *eteroduplex* sono frammenti di DNA a doppio filamento formati per l'appaiamento tra un filamento *wild type* ed uno recante la mutazione. Gli *omoduplex* sono il risultato dell'unione tra due filamenti perfettamente complementari. Queste due diverse molecole di DNA si ottengono per denaturazione termica e seguente rinaturazione graduale di una miscela di DNA *wild type* e DNA mutato.



**Figura 10.** Molecole *omoduplex* e molecole *eteroduplex*.

La risoluzione tra il picco di eluizione degli *eteroduplex* e quello degli *omoduplex* è resa possibile dalle condizioni parzialmente denaturanti a cui avviene la cromatografia. Alla temperatura di risoluzione, gli *eteroduplex* sono leggermente più destabilizzati rispetto agli *omoduplex*, a causa del *mismatch* tra i filamenti in corrispondenza della mutazione; vengono pertanto eluiti più velocemente. I picchi nel cromatogramma corrispondenti all'eluizione degli *eteroduplex* sono, quindi, anticipati, o visibilmente modificati.

La DHPLC è una tecnica veloce per il rilevamento di piccole mutazioni in amplificati di dimensioni comprese tra le 100 e le 700 paia di basi; l'analisi di ogni campione, infatti, richiede circa 8 minuti, incluse la rigenerazione e l'equilibratura della colonna. L'analisi basata sull'identificazione degli *eteroduplex* è un rapido mezzo di *screening* quando si devono valutare grandi numeri di pazienti; si presta, dunque, ad un utilizzo sistematico in diagnostica.

La sensibilità della tecnica è stata valutata con uno studio retrospettivo (Han *et al.*, 2001) di una coorte di 111 pazienti NF1, non correlati tra loro, con una mutazione germinale nota. Gli esoni contenenti le specifiche mutazioni sono stati analizzati mediante DHPLC e 108 lesioni sono state correttamente confermate, dimostrando che la metodica ha una sensibilità del 97%. Tuttavia, nell'analisi prospettica la DHPLC riconosce fino al 72% delle mutazioni presenti

(De Luca *et al.*, 2003). Questa discrepanza nel tasso di identificazione delle lesioni si spiega con la considerazione che alcuni tipi di alterazioni sfuggono all'analisi con DHPLC, tra cui ampie delezioni, che rimuovono l'intero gene, delezioni multiesoniche (15% degli affetti) (De Luca *et al.*, 2003), estese duplicazioni o inversioni ed alcuni errori nei siti di *splicing*, perché troppo interni negli introni o troppo vicini alle estremità dell'amplificato.

### 3.3.1.1 Amplificazione e analisi di mutazione mediante DHPLC

Gli esoni del gene *NFI* e le rispettive regioni fiancheggianti sono state amplificate utilizzando coppie di *primer* specifiche per ogni amplificato. I *primer* sono stati reperiti per lo più dalla letteratura (De Luca *et al.*, 2003; Han *et al.*, 2001), oppure sono stati disegnati utilizzando il *software*, disponibile *online*, *Primer 3* ([http://frodo.wi.mit.edu/cgi-bin/primer3/primer3\\_www.cgi](http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi)). In tabella 2 sono riportate le sequenze delle coppie di *primer* (*forward* e *reverse*), con le relative temperature di *annealing* e dimensioni degli amplificati.

Esone	<i>Primer Forward (5'-3')</i>	<i>Primer Reverse (5'-3')</i>	bp	T <sub>a</sub> (°C)
1	cag acc ctc tcc ttg cct ctt	gga tgg agg gtc gga ggc tg	439	62
2	aag ctg tta acg tgt ttt ttt ttt c	aag aaa aga aag caa att ccc c	228	54
3	ttt cac ttt tca gat gtg tgt tg	tgg tcc aca tct gta ctt tg	245	55
4a	tta aat cta ggt ggt gtg t	aaa ctc att tct ctg gag	517	57
4b	tca agt ggt cct cct gcc tt	cat tca tga tac tag ttt ttg ac	283	56
4c	ttt cct agc aga caa cta tcg a	agg atg cta aca aca gca aat	306	58
5	gaa gga agt tag aag ttt gtg aca	caa tgg tat cct tac cag cca t	308	57
6	cat gtt tat ctt tta aaa atg ttg cc	ata atg gaa ata att ttg ccc tcc	301	56
7	gct aca tct gga ata gaa gaa act tca	cca ttt agg ctg atg aac aca	399	58
8	gga ttt tac tgc cat ttg tgt g	taa cag cat cag taa ata tag tta gat a	275	60
9	tagagattaataggttcacttta	tccctcggtaagacttaa	460	58
9a	ctg tgg ctc aga aca cta aaa	cac atg cag tgc tca tta ga	308	54
10a	acg taa ttt tgt act ttt tet tcc	caa tag aaa gga ggt gag att c	222	55
10b	gcaaaaacgatttcattg	taaaccaattaactagtgtgat	393	58
10c	att gaa gtt tcc ttt ttt tcc ttg c	catggaattcatttcccctta	164	56
11	cca aaa atg ttt gag tga gtc t	acc ata aaa cct ttg gaa gtg	256	54

<b>Esone</b>	<b>Primer Forward (5'-3')</b>	<b>Primer Reverse (5'-3')</b>	<b>bp</b>	<b>T<sub>a</sub> (°C)</b>
<b>12a</b>	aaaccttacaagaaaaactaagct	attaccattccaatattcttcca	303	53
<b>12b</b>	ctc ttg gtt gtc agt gct tc	cag aaa aca aac aga gca cat	260	58
<b>13</b>	cac agt tta ttg cat tgt tag	gcc atg tgc ttt gag gca g	375	55
<b>14</b>	gct ctt cct act cct ttt gg	ttt ctg ttg cta agg gca ta	191	58
<b>15</b>	act tgg ctg tag ctg att ga	act tta ctg agc gac tct tga a	247	57
<b>16</b>	tgg ata aag cat aat ttg tca agt	tag aga aag gtg aaa aat aag ag	549	53
<b>17</b>	tct cta ggg ggt ctg tct tc	cac cct agt ttg tgt gca gt	326	60
<b>18</b>	aga agt tgt gta cgt tct ttt ct	ctc ctt tct acc aat aac cgc	367	58
<b>19a</b>	tcatgtcacttaggtatctgg	taaaaccactaataactgaagg	242	55
<b>19bext</b>	tagaactcactgatacagagggc	tatcctagtctctgcatgggtatt	509	58
<b>19bint</b>	agagatttgaccaggcaag	tatcctagtctctgcatgggtatt	218	58
<b>20</b>	cca ccc tgg ctg att atc g	taa ttt ttg ctt ctc tta cat gc	402	52
<b>21</b>	tgg cat gta aga gaa gca aaa a	tct tcc ccg ctt act cta atc a	312	56
<b>22</b>	tgc tac tct tta gct tcc tac	cct taa aag aag aca atc agc c	331	58
<b>23-1</b>	ttt gta tca ttc att ttg tgt gta	aaa aac agc ggt tct atg tga aaa g	282	55
<b>23-2</b>	ctt aat gtc tgt ata aga gtc tc	act tta gat taa taa tgg taa tct c	268	54
<b>23a</b>	agc cag aaa tag tag aca tga ttg ggt	cta ttt tgt gcc aga att agt aga	446	58
<b>24</b>	tgc aaa gtt tga cct ttg aac t	aaa gca cat aac tga aaa cca tag g	369	61
<b>25</b>	cct gtt tta ttg tgt aga tac ttc a	taa gtg gca aga aaa tta cct	134	55
<b>26</b>	aat tct aat gac ttt gca ttt ttg	atc taa att taa acg gag agt g	226	56
<b>27a</b>	gtt aca agt taa aga aat gtg tag	cta aca agt ggc ctg gtg gca aac	298	59
<b>27b</b>	ttt att gtt tat cca att ata gac tt	tcc tgt taa gtc aac tgg gaa aaa c	296	55
<b>28</b>	aaa ata aaa ttg att agt ggc atc tg	caa aac aca gac tgg aac tta ctt t	536	56
<b>29</b>	tct gga gcc ttt tag aat ttt atg	gca aca aac ccc aaa tca aac tga	460	57
<b>30</b>	caa ctt cat ttg tgt ttt ctc cta g	ctt tga att ctc tta gaa taa ttg tta	282	56
<b>31</b>	tttctcattgaccatcacatgc	ccaatgtggcaccagataaa	304	56
<b>32</b>	atc tag tat ttt tga ggc ctc ag	gggactcaaagtttagcacia	391	60
<b>33</b>	tctgcttctttacaggttattg	gca act tgg tgt tag agc ac	403	59
<b>34</b>	ttt tct gtc ttt act tgt tcc ttt a	cag tcc atg caa gtg ttt tt	384	58
<b>35</b>	gca tgg act gtg tta ttg gta	tgc aat taa aag atc cac aga	319	53
<b>36</b>	gtt ctg tgg atc ttt taa ttg c	cat tga cct caa att taa acg	238	53
<b>37</b>	cat tcc gag att cag ttt agg ag	aag taa cat tca aca ctg ata ccc	236	58
<b>38</b>	cta tgt cat gat tca tct tac tag c	cta aat ttg agt aat cta gga acc tc	233	60
<b>39</b>	cta ctg tgt gaa cct cat caa cc	gta aga cat aag ggc taa ctt act tc	284	62
<b>40</b>	tca ggg aag aag acc tca gca gat gc	tga act ttc tgc tct gcc acg caa cc	328	59
<b>41</b>	gtg cac att taa cag gta cta t	ctt cct agg cca tct cta gat	373	65
<b>42</b>	ctt gga agg agc aaa cga tgg ttg	caa aaa ctt tgc tac act gac atg g	356	60
<b>43</b>	ttt tct ttt tag tgt att ccc att	gat tct aag aaa tgg ctg gaa	287	53

<b>Esone</b>	<b>Primer Forward (5'-3')</b>	<b>Primer Reverse (5'-3')</b>	<b>bp</b>	<b>T<sub>a</sub> (°C)</b>
<b>44</b>	cac gtt aat tcc cta tct tgc	tga gaa gta gaa gac tgt atc c	268	56
<b>45</b>	cat gaa tag gat aca gtc ttc tac	cac att act ggg taa gca ttt aac	269	61
<b>46</b>	aaa tgt tcc tct gtt gac tt	cat caa cca tcc ttc tcc aga	211	53
<b>47</b>	ctg tta caa tta aaa gat acc ttg c	tgt gtg ttc tta aag cag gca tac	185	62
<b>48</b>	ttt tgg ctt cag atg ggg att tac	aag gga att cct aat gtt ggt gtc	351	66
<b>48a</b>	att cct tct gaa aac caa gg	aag gca gac tga gct tac ag	280	57
<b>49</b>	aga atg tgt ccc cgt tgt taa	caa acc gga tgg gtt cat ta	369	55

**Tabella 2.** Sequenze dei primers utilizzati per la PCR, lunghezza dell'amplificato e temperatura di annealing.

L'amplificazione mediante PCR è stata eseguita, in un volume di 50 µl, secondo il protocollo riportato in tabella 3, utilizzando l'enzima *Taq Gold Polimerase* (Amplitaq Gold, Applied Biosystem) e sottoponendo la soluzione di reazione al ciclo di temperature descritto in tabella 4.

<b>T (°C)</b>	<b>Tempo</b>	<b>N° cicli</b>
95	10'	1
95	10"	40
T <sub>a</sub>	40"	
72	1'	
72	30'	1
4	10'	1

**Tabella 3.** Profilo termico della reazione.

<b>Mix di reazione</b>	<b>Volume per una reazione (µl)</b>
H <sub>2</sub> O sterile	31,7
Buffer 10X	5
MgCl <sub>2</sub>	3
dNTPs (10mM)	1
pForward (10 µM)	2,5
pReverse (10 µM)	2,5
Taq Gold	0,3
Volume finale	46 µl mix+ 4 µl DNA

**Tabella 4.** Mix della reazione di amplificazione.



La resa della reazione, sia qualitativa che quantitativa, è stata valutata sottoponendo 5 µl di prodotto di amplificazione a corsa elettroforetica su gel d'agarosio all'1,5% colorato con etidio bromuro, dopo aver sottoposto i campioni a denaturazione e rinaturazione termica. L'intensità luminosa della banda su gel è stata utilizzata come parametro per valutare la quantità di amplificato da sottoporre ad analisi DHPLC (da 8 a 15 µl a seconda dell'intensità della banda).

Per ogni esone è stato amplificato anche il DNA di soggetti campione, presi casualmente nella popolazione e negativi ad una diagnosi clinica di Neurofibromatosi tipo 1, per poter confrontare le sequenze dei soggetti in analisi con quelle di controlli sani.

I campioni ottenuti dall'amplificazione dei singoli esoni del gene *NF1* da DNA genomico dei pazienti in analisi, sono stati sottoposti a denaturazione e rinaturazione termica per consentire la formazione di eventuali molecole *eteroduplex*. In tabella 5 è descritto il profilo termico per la denaturazione e la successiva graduale rinaturazione (uguale per tutti gli esoni).

<b>T (°C)</b>	<b>Tempo</b>	<b>N° cicli</b>
95	10'	1
T - 1 <sup>(*)</sup>		75
20	10'	1

**Tabella 5.** Profilo termico per il ciclo di denaturazione / rinaturazione.

<sup>(\*)</sup> La temperatura viene abbassata di 1°C ad ogni ciclo, da 95°C fino a 20°C.

L'analisi dei prodotti di PCR è stata eseguita con l'ausilio dell'apparecchiatura per DHPLC: *2100 WAVE DNA fragment analysis system* (Transgenomic).

Per il calcolo teorico della temperatura di parziale denaturazione specifica per ogni amplificato sono stati utilizzati i seguenti metodi:

1. con l'ausilio del *software* disponibile *online*: <http://insertion.stanford.edu/melt.html>;
2. utilizzando il *software* NavigatorTM (Transgenomic) specifico dello strumento;
3. mediante informazioni riportate in letteratura.

Sperimentalmente, la messa a punto delle condizioni ottimali è stata eseguita valutando i cromatogrammi ottenuti eseguendo l'analisi su controlli sani.

Per ogni esone sono state utilizzate due temperature di analisi in modo che tutta la lunghezza della sequenza in questione, potesse trovarsi in condizioni di parziale denaturazione: domini diversi, all'interno della stessa sequenza, possono comportarsi in modo differente alla stessa temperatura di parziale denaturazione. Un parametro testato solo sperimentalmente è stato il valore da assegnare al *time shift* (tabella 6), valore numerico che induce una variazione nella quantità di acetonitrile utilizzata per l'eluizione e di conseguenza una variazione temporale nel rilevamento dei picchi cromatografici.

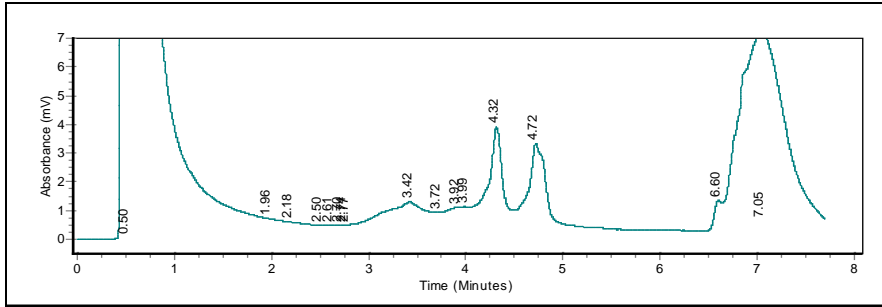
<b>Esone</b>	<b>T° dhplc (°C)</b>	<b>Time shift</b>	<b>T° dhplc (°C)</b>	<b>Time shift</b>
<b>1</b>	62,3	-0,5	63,3	0
<b>2</b>	55	-0,5	56	0
<b>3</b>	54,9	-1	55,7	-0,5
<b>4a</b>	53	-1	58	0,5
<b>4b</b>	51	-1	54,6	-1,5
<b>4c</b>	54	-1,5	55	-1,5
<b>5</b>	53,6	-1,5	54,6	-1
<b>6</b>	53	-0,5	56	0,5
<b>7</b>	54,2	-0,5	56,2	0,5
<b>8</b>	55,5	0	57,5	-0,5
<b>9a</b>	53	-1	54,1	-0,5
<b>10a</b>	53	-1	58	0
<b>10b</b>	53,6	0	54,3	-1
<b>10c</b>	57,3	0	59	1
<b>11</b>	52	-1	54	-0,5
<b>12a</b>	52	-2	56,2	-0,5
<b>12b</b>	57,5	0	59	0,5
<b>13</b>	58	-1,5	60	-1
<b>14</b>	60,3	-1	61,7	-0,5
<b>15</b>	56,6	-1,5	57,5	-1
<b>16</b>	55,5	0	58,2	1
<b>17</b>	54,6	-1	55,6	-1,5
<b>18</b>	55,5	1	56,5	1,5
<b>19b int</b>	54,8	-0,5	56,8	0
<b>19b ext</b>	51,5	-1	/	/

<b>Esone</b>	<b>T° dhplc (°C)</b>	<b>Time shift</b>	<b>T° dhplc (°C)</b>	<b>Time shift</b>
<b>20</b>	55	-1	57	0
<b>21</b>	55,8	-1,5	56,8	-1
<b>22</b>	56,5	0	58	0,5
<b>23-1</b>	55,6	-0,5	56,6	-0,5
<b>23-2</b>	53	-0,5	58	0,5
<b>23a</b>	53	-1	54,5	-0,5
<b>24</b>	57	0	58	1
<b>25</b>	56	-1	57	-0,5
<b>26</b>	54,8	-0,5	56,8	0
<b>27a</b>	55	0	57	0,5
<b>27b</b>	53,6	-1,5	54,6	-1,5
<b>28</b>	54,3	-1,5	55,3	-1,5
<b>29</b>	57,5	0	60	1
<b>30</b>	54,4	0	56,4	0
<b>31</b>	56,7	-1	57,7	-0,5
<b>32</b>	56	0,5	57	1
<b>33</b>	55,5	0	59	1,5
<b>34</b>	57,5	0	59,5	1
<b>35</b>	50,2	-1,5	51,2	-1,5
<b>36</b>	55	-0,5	59	0,5
<b>37</b>	54,1	-1	56	-0,5
<b>38</b>	57	0	59,5	1
<b>39</b>	55,5	0	56,5	0,5
<b>40</b>	53,9	-2,5	55,9	-2
<b>41</b>	54,7	-2	55,7	-1
<b>42</b>	52	-0,5	55	0,5
<b>43</b>	54,8	0	55,8	0
<b>44</b>	56,1	-1,5	58,1	-0,5
<b>45</b>	56	-0,5	57	-0,5
<b>46</b>	55	-1	56	-1,5
<b>47</b>	55,4	0	56,4	0
<b>48</b>	55,1	-0,5	58,1	0
<b>48a</b>	54,2	-1	58	-0,5
<b>49</b>	58	-1	60	0

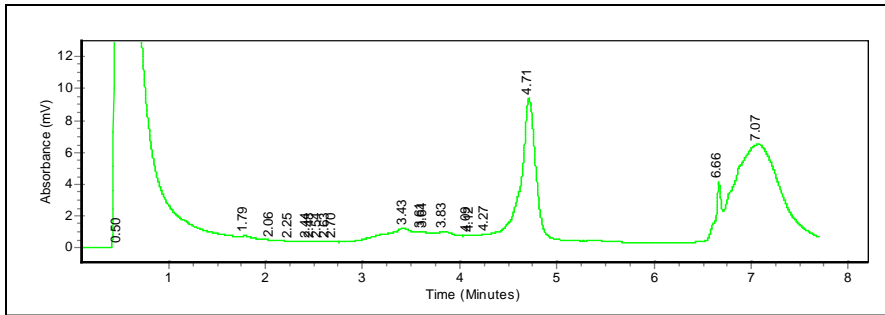
**Tabella 6.** Temperature di parziale denaturazione e rispettivo *time shift* per l'analisi DHPLC di ogni esone del gene *NFI*.

Il confronto dei cromatogrammi ottenuti dall'analisi degli amplificati dei pazienti, con quelli dei controlli sani ha messo in evidenza la presenza di anomalie

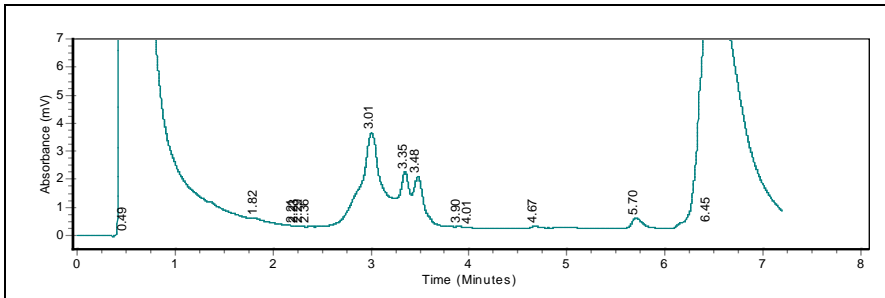
grafiche (numero e/o forma dei picchi) nelle sequenze di alcuni esoni in alcuni pazienti. Le figure 11-14 mostrano alcuni esempi di cromatogrammi anomali rispetto al corrispondente cromatogramma del controllo sano:



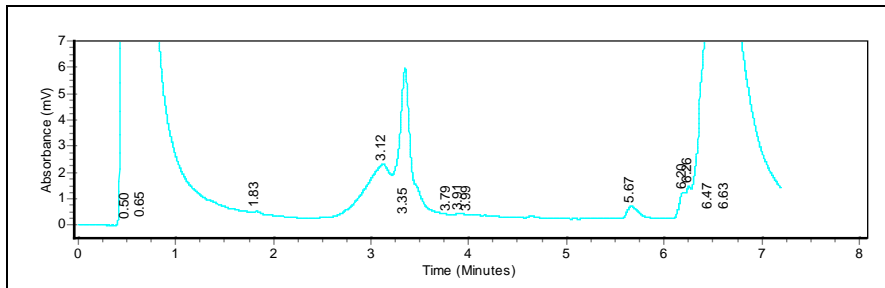
**Figura 11.** cromatogramma dell'esone 14 nel campione 2882.



**Figura 12.** cromatogramma dell'esone 14 nel controllo sano.



**Figura 13.** cromatogramma dell'esone 31 nel campione 2891.



**Figura 14.** cromatogramma dell'esone 31 nel controllo sano.

I campioni in cui è stata valutata un'anomalia del cromatogramma sono stati sottoposti a sequenziamento diretto della regione genica in questione. Il cromatogramma, infatti non fornisce alcuna informazione sul tipo di alterazione, ne indica solo la presenza; di conseguenza è necessario sequenziare il prodotto di PCR.

### **3.3.2 High-Resolution Melting Analysis (HRMA)**

L'analisi di *melt* ad alta risoluzione (*High-Resolution Melting Analysis*) è una tecnica, introdotta recentemente tra le tecniche di biologia molecolare, che presenta numerose applicazioni come:

- ricerca di mutazioni;
- genotipizzazione degli SNPs (*Single Nucleotide Polymorphism*);
- ricerca della perdita di eterozigotità;
- analisi di metilazione del DNA;
- caratterizzazione degli aplotipi.

Questa metodica permette di distinguere gli ampliconi sfruttando la diversa temperatura di *melting* delle sequenze; questa temperatura è legata alla lunghezza dell'amplicato (numero di basi), alla composizione in basi e al contenuto di GC della sequenza stessa. Basandosi proprio su questo aspetto la tecnica permette di rilevare molteplici tipi di variazioni di sequenza del DNA come: cambiamenti di singola base, inserzioni, delezioni e duplicazioni.

Per eseguire l'*High Resolution Melting Analysis* è stato utilizzato lo strumento Rotor-Gene 6000 (Corbett).

La caratteristica principale di questo apparecchio è quella di essere dotato di un sistema rotante dove vengono posizionati i campioni; il dispositivo è alloggiato all'interno di una camera dove attraverso un sofisticato sistema di riscaldamento – raffreddamento è possibile sviluppare accurati profili termici per l'amplificazione e il rilevamento degli ampliconi. La peculiarità del sistema rotante, a differenza delle normali piattaforme Peltier, è quella di assicurare l'uniformità della temperatura di lavoro in tutti i campioni contemporaneamente; i campioni ruotano continuamente durante un esperimento alla velocità di 400 rpm

(giri per minuto).

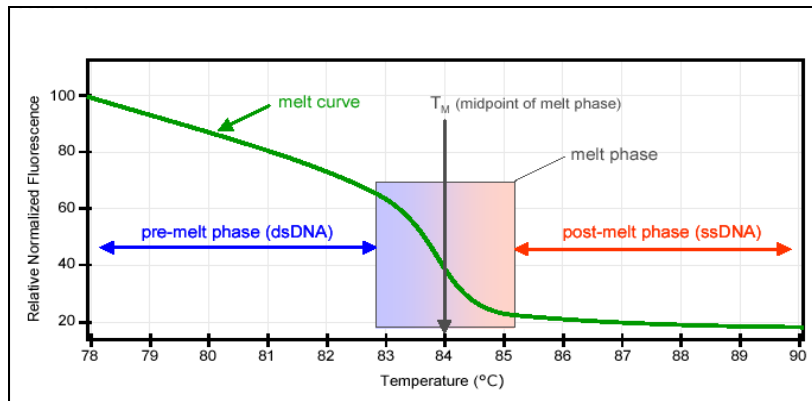
Lo strumento possiede un sistema di rilevazione in fluorescenza per l'analisi e la quantificazione dei prodotti di amplificazione. I campioni vengono eccitati dal fondo della camera attraverso un LED (*Light Emitting Diode* o diodo ad emissione luminosa) e l'energia viene trasmessa attraverso la base delle provette. La fluorescenza emessa passa attraverso i filtri di emissione ai lati della camera e viene catturata da un fotomoltiplicatore. Il percorso ottico fisso assicura l'eccitazione uniforme di ogni singolo campione durante la rotazione nella camera. L'apparecchio è collegato ad un computer che, tramite un apposito *software*, consente di seguire l'andamento della reazione di amplificazione in tempo reale e successivamente di analizzare la curva di *melting*.

L'amplificazione deve avvenire in presenza di un fluoroforo che si lega al DNA a doppia elica ed emette fluorescenza se esposto a fonte luminosa di opportuna lunghezza d'onda.

Perciò durante la reazione di amplificazione la concentrazione di DNA e quindi la fluorescenza aumentano in maniera esponenziale per poi arrivare rapidamente a *plateau* intorno al 35°- 40° ciclo.

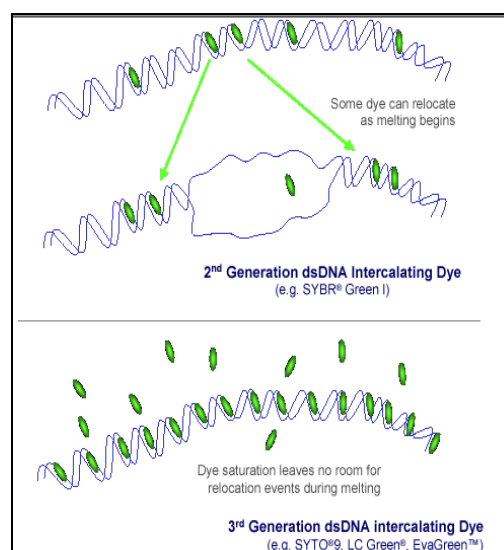
Dopo l'amplificazione, i campioni subiscono una denaturazione progressiva da 70° a 90°C e la variazione dell'emissione di fluorescenza viene monitorata per generare una curva di *melt* (figura 15). Questa diminuirà con l'aumentare della temperatura perché il fluoroforo viene rilasciato dal dsDNA (*double strand DNA*) che si dissocia in ssDNA (*single strand DNA*).

La temperatura in corrispondenza della quale si ha un repentino decremento della fluorescenza costituisce la temperatura di *melting* alla quale il 50% del DNA si trova in forma denaturata. Ad ogni prodotto di PCR corrisponde una temperatura di *melting* caratteristica, perciò tramite questo tipo di analisi è possibile distinguere i diversi genotipi.



**Figura 15.** Curve di *melt* dopo analisi HRM. La curva mostra una diminuzione della fluorescenza con l'aumentare della temperatura mano a mano che il fluoroforo è rilasciato dal dsDNA che si dissocia in ssDNA. Si possono distinguere una fase di *pre-melt* (in cui la maggior parte del DNA è ancora a doppio filamento) e una fase di *post-melt* (in cui il DNA è quasi tutto a singolo filamento).  $T_M$  indica la temperatura di *melting* alla quale il 50% del DNA si trova in forma denaturata (HRM Assay Design and Analysis CorProtocol™ 6000-1-July06).

Ciò che rende questa analisi di *melting* ad alta risoluzione è l'utilizzo di fluorofori intercalanti di terza generazione, introdotti sul mercato nel 2003. Questi possono essere usati ad alte concentrazioni in modo da saturare il doppio filamento di DNA, senza per questo andare a modificare la temperatura di *melting*. Una più alta concentrazione di intercalante consente di avere una più alta fedeltà del segnale, perché non c'è una redistribuzione del fluoroforo nelle regioni non ancora denaturate durante la *melting*, come avviene con i fluorofori di seconda generazione (figura 16).

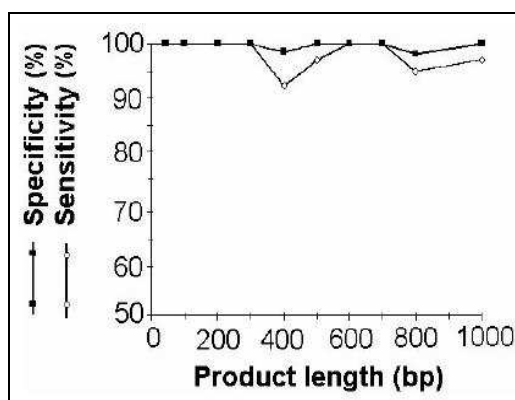


**Figura 16.** Ridistribuzione del fluoroforo durante la dissociazione del DNA (White *et al.*, 2006).

Il tempo richiesto, per campione, per l'HRMA con Rotor-Gene 6000 è variabile poiché dipende dal rotore utilizzato e dal numero di campioni caricati; l'intera corsa richiede circa due ore, tra amplificazione e dissociazione. Se si utilizza il rotore da 36 posizioni si può stimare un tempo di analisi per campione di circa 3 minuti, mentre utilizzando il rotore da 72 posizioni questo si riduce fino a circa 1,5 minuti.

Reed G. e Wittwer C. hanno pubblicato nel 2004 un lavoro di valutazione della sensibilità e specificità dell'*High Resolution Melting Analysis*. Lo studio è stato condotto valutando la capacità della metodica di rilevare variazioni di sequenza in funzione della lunghezza del prodotto di PCR (da 50 a 1000 bp), della posizione dello SNP all'interno dell'amplificato e della percentuale di GC contenuta (40%, 50% e 60%).

Dall'analisi di 1632 casi viene riportato che il tasso di errore nel rilevare la variazione dipende principalmente dalla lunghezza del prodotto di PCR (figura 17) e dal tipo di cambiamento molecolare: è stato osservato che la maggior parte degli SNPs non rilevati, sono localizzati in sequenze ricche in AT e coinvolgono un cambiamento T>A o A>T. Al contrario la posizione della variazione non produce alcun effetto sull'accuratezza del metodo.



**Figura 17.** Effetto della lunghezza del prodotto di PCR sulla sensibilità e la specificità della tecnica di identificare SNPs (Reed *et al.*, 2004).

L'analisi di variazione di singolo nucleotide può essere limitata anche dalla sensibilità e dall'efficienza dello strumento nel controllo della temperatura e nell'acquisizione dei dati. Dallo studio di White *et al.* del 2006, in cui sono state testate la sensibilità e la specificità di tre macchine per HRMA (*Rotor-Gene 6000*,



Corbett Life Science; *HR-1*, Idaho Technology; *LightScanner 384 well*, Idaho Technology) emerge infatti una sensibile differenza tra gli strumenti (tabella 7) (White *et al.*, 2006).

Strumento	Sensibilità	Specificità
Rotor-Gene 6000	100%	95%
HR-1	98,4%	95%
LighScanner 384 well	99%	88%

**Tabella 7.** Confronto della percentuale di sensibilità [veri positivi / (veri positivi + falsi negativi)] e di specificità [veri negativi / (veri negativi + falsi positivi)] tra tre macchine specifiche per HRMA (White *et al.*, 2006).

Questo dato avvale la scelta di utilizzare il Rotor-Gene 6000 come strumento per effettuare l'analisi di mutazione nel gene *NFI*.

### 3.3.2.1 Amplificazione e analisi di mutazione mediante HRMA

La reazione è stata condotta utilizzando l'enzima *Takara Ex Taq Polimerase* (Takara) secondo il protocollo riportato in tabella 8 e con il profilo termico riportato in tabella 9, in un volume totale di 25  $\mu$ l.

Mix di reazione	Volume per 1 reazione ( $\mu$ l)
H <sub>2</sub> O sterile	12.95
Buffer 5X	5
MgCl <sub>2</sub> (50 $\mu$ M)	1.5
dNTPs 10mM	0.5
p <i>Forward</i> (10 $\mu$ M)	0.75
p <i>Reverse</i> (10 $\mu$ M)	0.75
EvaGreen 20X	1.25
Taq Takara	0.3
Volume finale	23 $\mu$ l mix + 2 $\mu$ l DNA

**Tabella 8.** Mix della reazione di amplificazione.

T (°C)	Tempo	N° cicli
95	3'	1
95	30"	40
Ta	30"	
72	30"	
72	5'	1

**Tabella 9.** Profilo termico della reazione.

Il protocollo è stato applicato per tutti gli amplificati ad eccezione di quelli per gli esoni 5, 10a, 10c, 14, 25 e 26, i quali sono stati eseguiti con un profilo di temperature *shuttle*, riportato in tabella 10.

T (°C)	Tempo	N° cicli
95	3'	1
95	10"	40
Ta	20"	
72	5'	1

**Tabella 10.** Profilo termico della reazione *shuttle*.

In contemporanea all'analisi dei campioni è stato amplificato in triplice copia anche il DNA di soggetti sani, presi casualmente nella popolazione e negativi ad una diagnosi clinica di Neurofibromatosi di Tipo 1, per poter confrontare i risultati ottenuti dai soggetti in analisi con quelli dei controlli.

I campioni ottenuti dall'amplificazione dei singoli esoni del gene *NFI* da DNA genomico dei pazienti in analisi, sono stati di seguito sottoposti a denaturazione termica progressiva, con un *degrees* di 0,1°C. Per ogni esone il *range* di temperatura applicato per l'analisi della curva di dissociazione è stato ottimizzato attorno alla temperatura di *melt* caratteristica di ogni singolo amplificato, come riportato in tabella 11.

Sperimentalmente, la messa a punto delle condizioni ottimali è stata eseguita valutando i grafici relativi alla curva di *melt* ottenuti eseguendo l'analisi su controlli sani.

<b>Esone</b>	<b>Tm 1 (°C)</b>	<b>Tm 2(°C)</b>	<b>Range HRM (°C)</b>
<b>2</b>	78,2	/	75-83
<b>3</b>	77	/	73-82
<b>4b</b>	77,5	/	73-82
<b>4c</b>	77,6	/	74-82
<b>5</b>	76,8	/	72-81
<b>6</b>	79,3	/	73-83
<b>7</b>	77,8	81,1	74-86
<b>9</b>	76,8	/	72-83
<b>9a</b>	76,8	83,2	73-88
<b>10a</b>	81,2	/	77-86
<b>10b</b>	76,4	77,5	72-82
<b>10c</b>	82,3	/	78-87
<b>11</b>	75,2	77,8	70-83
<b>12b</b>	80,1	/	75-85
<b>13</b>	85	/	80-89
<b>14</b>	82,8	/	79-87
<b>15</b>	79,4	/	76-84
<b>17</b>	78,8	/	75-84
<b>18</b>	75,5	78,7	71-84
<b>19a</b>	79,1	/	74-84
<b>21</b>	82,3	/	78-86
<b>22</b>	81	/	76-86
<b>23-1</b>	79,5	/	75-84
<b>23-2</b>	75,6	81,4	71-87
<b>23a</b>	78	80	74-85
<b>24</b>	77,7	81,7	72-86

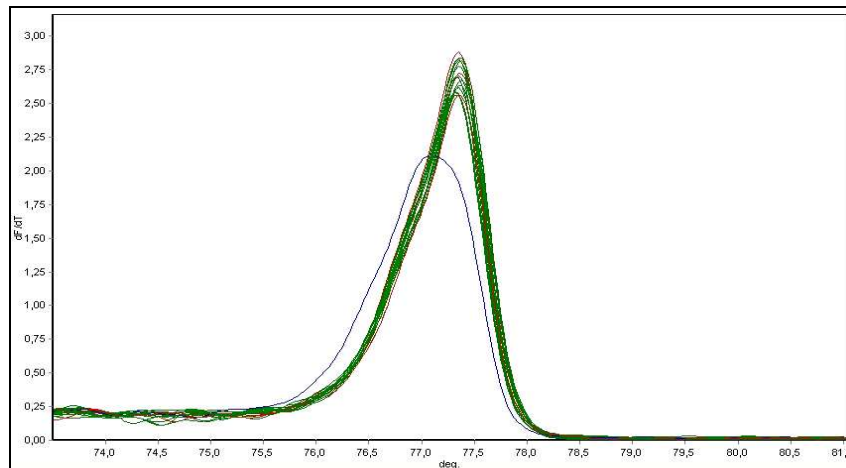
<b>Esone</b>	<b>Tm 1 (°C)</b>	<b>Tm 2(°C)</b>	<b>Range HRM (°C)</b>
<b>25</b>	78,5	/	74-84
<b>26</b>	80,4	/	76-85
<b>27a</b>	81,5	/	77-86
<b>27b</b>	78,2	/	73-83
<b>29</b>	82,6	84,7	79-89
<b>30</b>	80,4	/	76-85
<b>31</b>	79,8	/	75-84
<b>32</b>	76,4	80,5	72-85
<b>33</b>	79,6	81,9	76-87
<b>34</b>	74,6	81,6	71-87
<b>36</b>	80,4	/	77-85
<b>37</b>	77,8	/	73-83
<b>38</b>	80,2	82,8	78-88
<b>39</b>	75,7	79,4	72-85
<b>40</b>	79,7	/	76-85
<b>41</b>	79,1	/	75-84
<b>42</b>	74,8	84,7	72-87
<b>43</b>	79,1	/	73-84
<b>44</b>	80,3	/	77-84
<b>45</b>	79,4	/	75-84
<b>46</b>	79,6	/	76-84
<b>47</b>	78,2	/	74-83
<b>48</b>	82,4	/	78-87
<b>48a</b>	81,4	/	77-84
<b>49</b>	82,3	83,6	79-90

**Tabella 11.** Temperatura di *melt* caratteristica di ogni esone del gene *NF1* e *range* di temperatura applicato per l'analisi della curva di dissociazione.

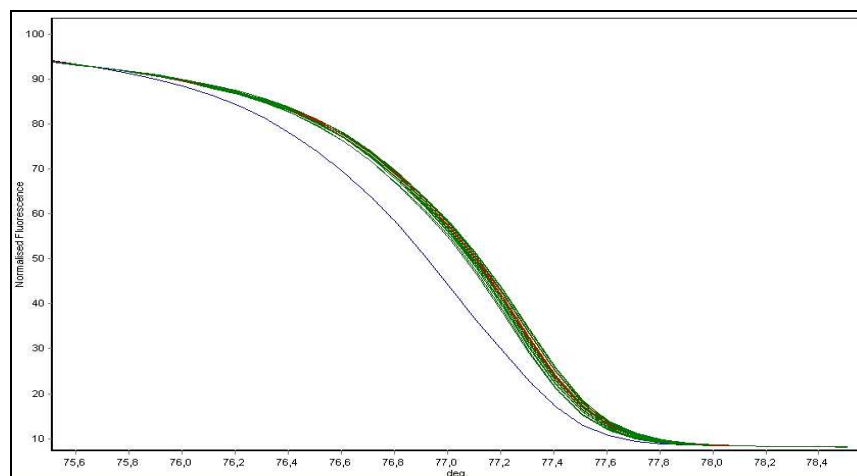
Il *software* di analisi delle curve di dissociazione estrapola tre grafici per l'analisi dei risultati: un grafico delle temperature di *melt* e due grafici specifici dell'HRM (*Normalized Graph* e *Difference Graph*).

Il confronto dell'andamento di tutte e tre le curve dei campioni rispetto a quello

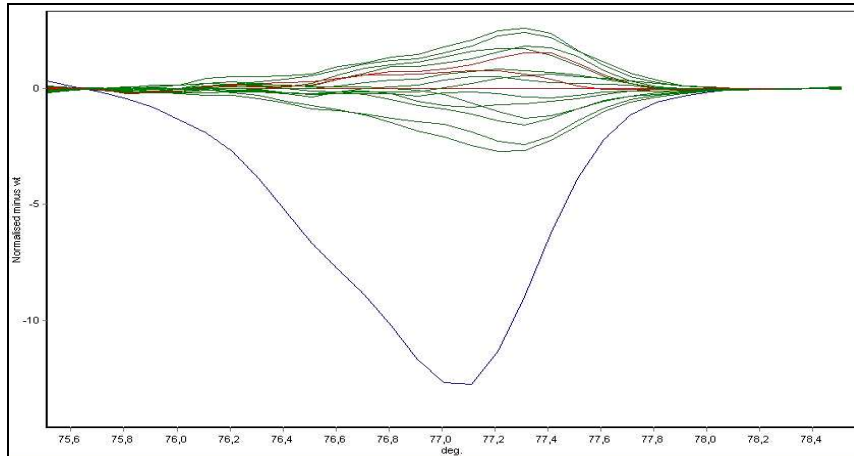
dei controlli sani di riferimento ha evidenziato delle anomalie nella curva di dissociazione di alcuni pazienti. Nelle figure 18-23 sono riportati alcuni esempi:



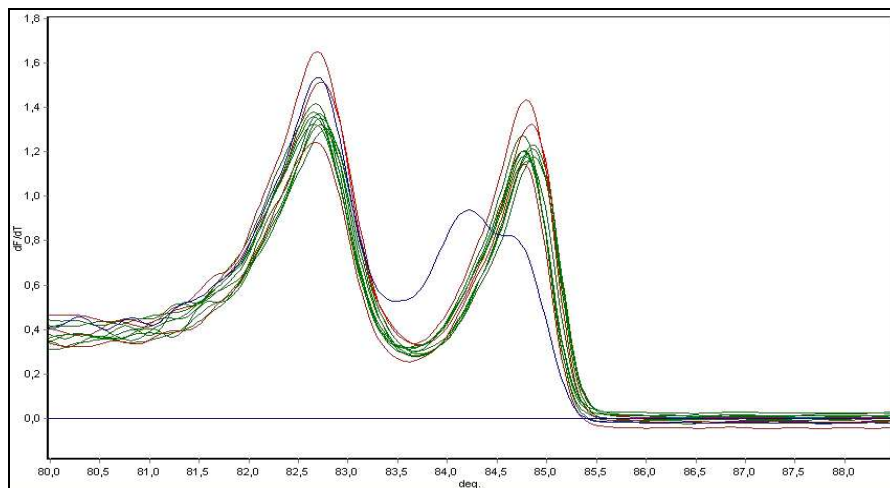
**Figura 18.** Grafico della Tm dell'esone 3 di *NF1*. In rosso è evidenziato il controllo sano in triplice copia; in verde campioni risultati *wild type* e in blu il campione 3569.



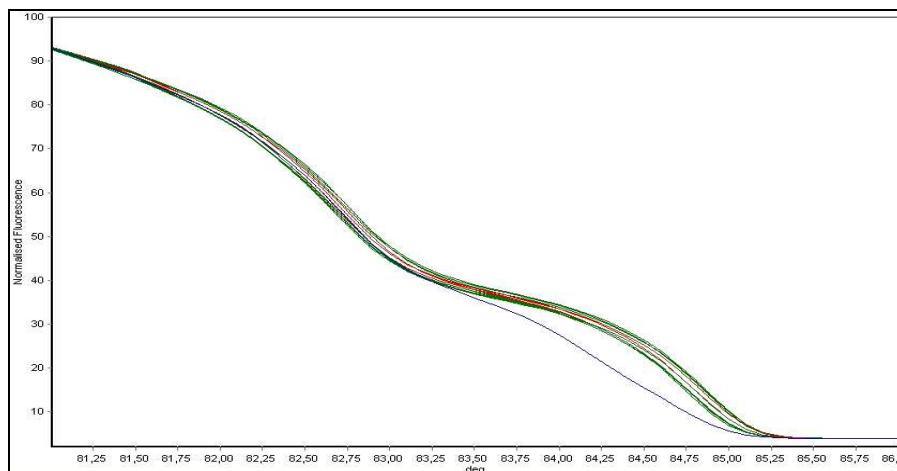
**Figura 19.** *Normalized graph* dell'esone 3 di *NF1*. In rosso è evidenziato il controllo *sano* in triplice copia; in verde campioni risultati *wild type* e in blu il campione 3569.



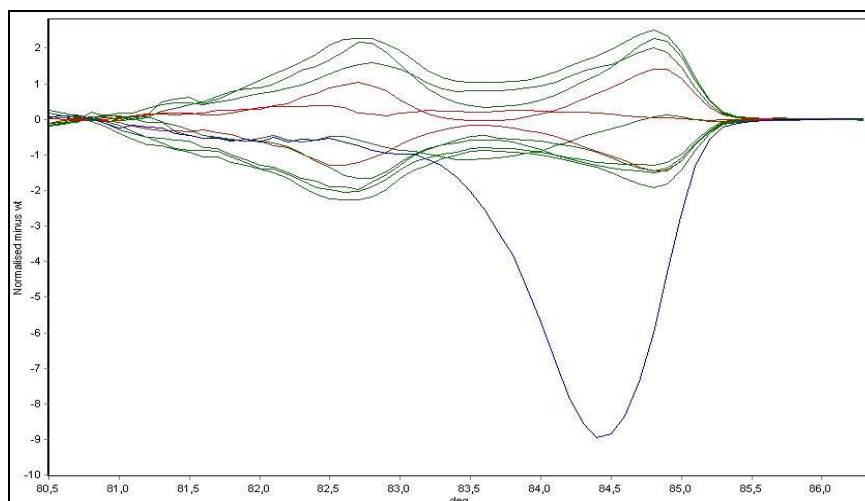
**Figura 20.** *Difference graph* dell'esone 3 di *NF1*. In rosso è evidenziato il controllo *sano* in triplice copia; in verde campioni risultati *wild type* e in blu il campione 3569.



**Figura 21.** Grafico della due  $T_m$  dell'esone 29 di *NF1*. In rosso è evidenziato il controllo *sano* in triplice copia; in verde campioni risultati *wild type* e in blu il campione 3024.



**Figura 22.** *Normalized graph* dell'esone 29 di *NF1*. In rosso è evidenziato il controllo *sano* in triplice copia; in verde campioni risultati *wild type* e in blu il campione 3024.



**Figura 23.** *Difference graph* dell'esone 29 di *NF1*. In rosso è evidenziato il controllo *sano* in triplice copia; in verde campioni risultati *wild type* e in blu il campione 3024.

I campioni in cui è stata osservata una variazione nell'andamento delle curve sono stati sottoposti a sequenziamento diretto della regione genica in questione: i grafici rivelano la presenza di un'alterazione nella sequenza dell'amplificato, ma non forniscono alcuna informazione sulla natura di questa; si rende quindi necessario sequenziare il prodotto di amplificazione.

### 3.3.3 Sequenziamento

L'analisi del campione è stata eseguita utilizzando un sequenziatore automatico *ABI-PRISM 3100 Genetic Analyzer* (Applied Biosystem).

I prodotti di PCR sono stati sottoposti a purificazione mediante l'uso di *Microcon Centrifugal Filter Devices* (Millipore). La concentrazione di acido nucleico nel prodotto di purificazione è stata valutata mediante lettura spettrofotometrica a 260 nm.

La reazione di sequenza è stata eseguita a partire da 80 ng di acido nucleico secondo il protocollo riportato nel kit *Big Dye 3.1 Cycle Sequencing Kit* (Applied Biosystem);

Mix di reazione	Volume per 1 reazione (µl)
PCR purificata (10 ng/µl)	8
H <sub>2</sub> O sterile	5,4
buffer 5X	3
primer <i>Forward/Reverse</i> (10 µM)	0,6
BigDye 3.1	3
Volume finale	20

**Tabella 12.** Mix di reazione di amplificazione per il sequenziamento.

T (°C)	Tempo	N° cicli
95	10"	25
50	5"	
60	2'	

**Tabella 13.** Ciclo della reazione di sequenza.

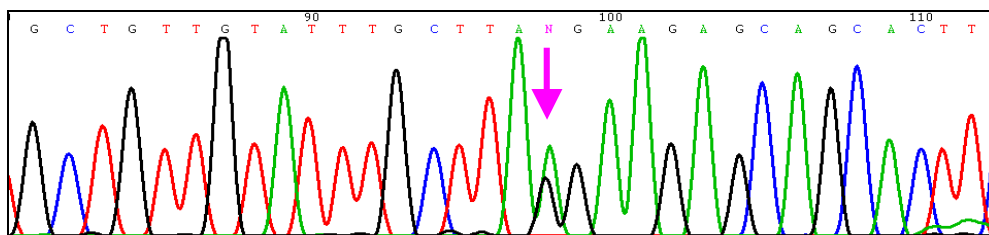
Gli amplificati ottenuti sono stati sottoposti ad ulteriore purificazione con colonnine *Centrisept Spin Columns* (Applied Biosystem); 10 µl di purificato sono stati mescolati a 10 µl di formammide e, dopo rapida denaturazione (10 minuti a 95° C), caricati nel sequenziatore automatico.

Gli elettroferogrammi ottenuti sono stati analizzati con il *software Chromas* (WinImp self-Extractor Copyright 2000 Technelysium Pty Ltd) e le sequenze dei campioni allineate alla sequenza *wild type* del gene utilizzando il *software BLAST* (<http://www.ncbi.nih.gov/BLAST>).

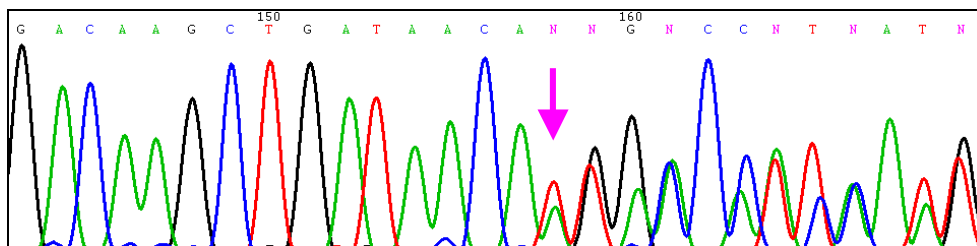
Gli amplificati degli esoni che presentavano anomalie sono stati sequenziati in entrambi i sensi di lettura, con i *primer forward e reverse*.

Dall'analisi delle sequenze e il loro allineamento in *BLAST* si sono rilevate 196 alterazioni di sequenza (patogenetiche o sospette tali) e diversi polimorfismi. Nelle figure 24-27 sono riportati alcuni esempi di elettroferogrammi in cui sono evidenziate le alterazioni di sequenza.

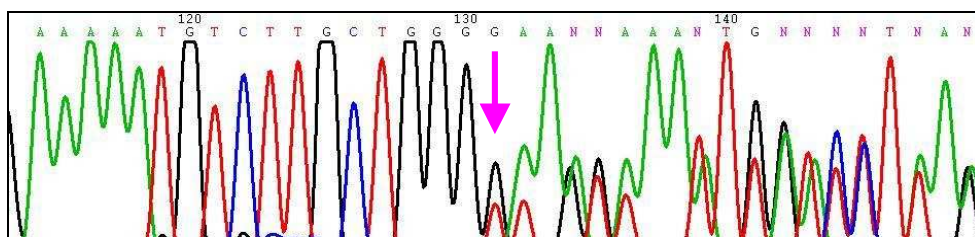




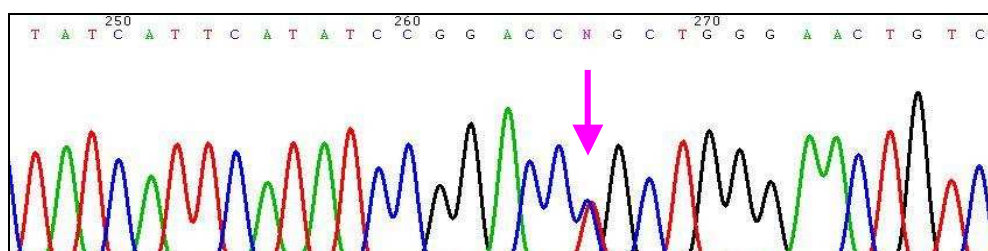
**Figura 24.** Parte dell'elettroferogramma della sequenza *forward* dell'introne 13/esone 14 di *NF1* nel campione 2882. La mutazione osservata è la **transizione G>A**.



**Figura 25.** Parte dell'elettroferogramma della sequenza *forward* dell'esone 31 di *NF1* nel campione 2891. La mutazione osservata è la **delezione di una A**.



**Figura 26.** Parte dell'elettroferogramma della sequenza *forward* dell'esone 3/introne 3 di *NF1* nel campione 3569. La mutazione osservata è la **delezione di una G**.

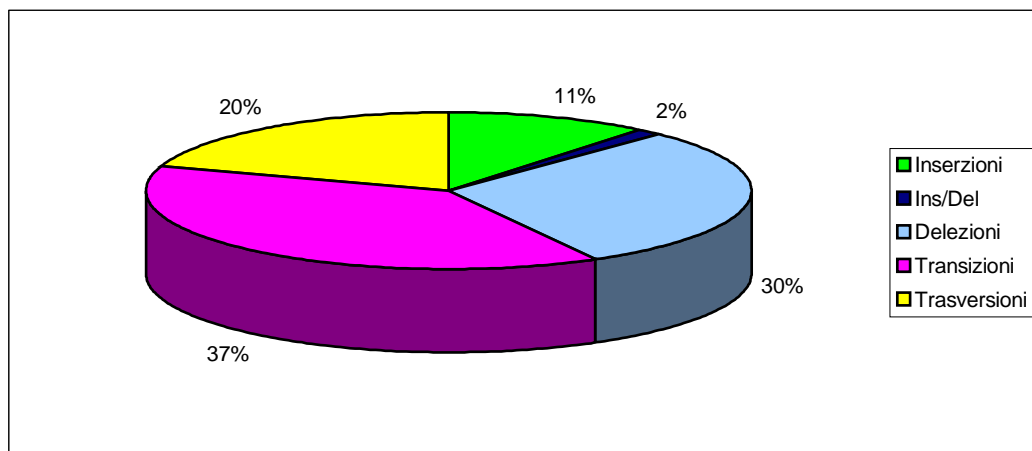


**Figura 27** Parte dell'elettroferogramma della sequenza *forward* dell'esone 29 di *NF1* nel campione 3024. La mutazione osservata è la **transizione C>T**.

### 3.3.4 Risultati dello *screening* di mutazione

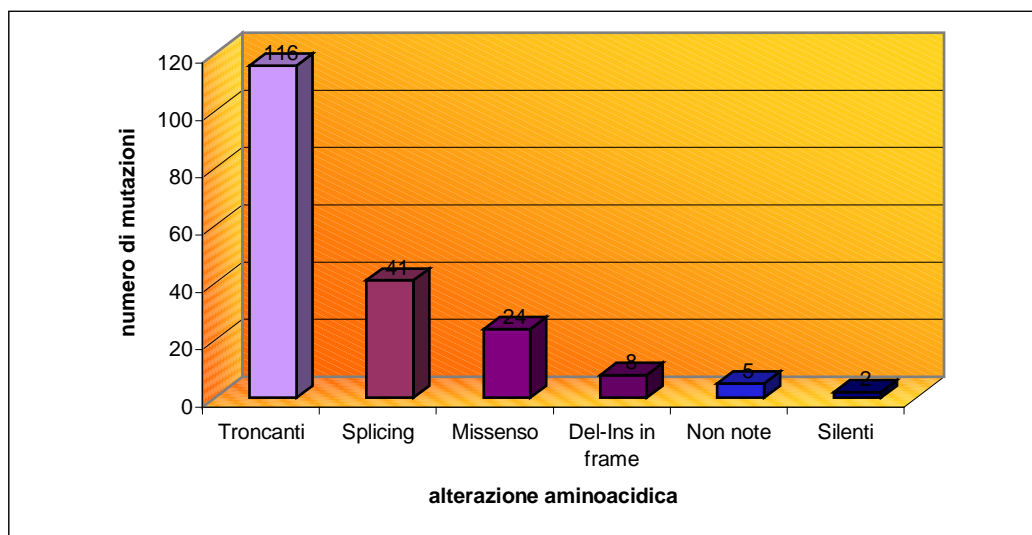
Dalle analisi delle sequenze sono emerse 196 mutazioni su 310 pazienti analizzati; l'elenco delle mutazioni osservate è riportato in tabella 14 (in appendice). A livello nucleotidico le mutazioni osservate sono state: 21 inserzioni, 58 delezioni, 75 transizioni, 39 trasversioni e 3 mutazioni che derivano da

un'inserzione combinata con una delezione (ins/del).



**Figura 28.** Percentuale delle diverse alterazioni nucleotidiche sul totale delle mutazioni osservate.

Considerando le alterazioni a livello proteico, le mutazioni osservate si traducono in: 116 mutazioni troncanti, 24 mutazioni missenso, 2 variazioni silenti, 7 delezioni *in frame*, 1 inserzione *in frame*, 41 mutazioni che coinvolgono i siti di *splicing* e 5 di cui non è noto l'effetto a livello aminoacidico.



**Figura 29.** Mutazioni classificate sulla base dell'alterazione aminoacidica che ne deriva.

A conferma dell'alto tasso mutazionale del gene, la percentuale di mutazioni *de novo* osservate nella casistica esaminata è stata del 48%.

In accordo con i dati riportati in letteratura (Fashold *et al.*, 2000), la porzione maggiore di mutazioni (59,18%) sono alterazioni che provocano la formazione di una proteina tronca: sia in seguito a cambio aminoacidico (mutazioni nonsense), sia per inserzione e/o delezione di una o più basi che, generando una variazione nel modulo di lettura, causano la formazione di un codone di stop prematuro.

### 3.3.4.1 Mutazioni di *splicing*

Le mutazioni osservate che inducono un'alterazione del meccanismo di *splicing* dell'RNA costituiscono il 20,92% delle mutazioni osservate e risultano essere le più frequenti dopo quelle troncanti. In tabella 15 sono riportate le variazioni che alterano, o potrebbero alterare, il meccanismo di *splicing* dell'mRNA.

Esone Introne	Numero Identificativo	Alterazione nucleotidica	Alterazione aminoacidica	Previsione prodotto genico	Tipo di mutazione	Ricorrenza di mutazione	F/S	Referenze
IVS2	3112	204+1 G>T	/	splice site skip exon 2	trasversione	1	S	Fashold <i>et al.</i> , 2000
IVS 2	3446	205-23 G>A	/	splice site	transizione	1	U	this report
IVS 3	3366	288+1 G>A	/	splice site	transizione	1	U	this report
IVS 4b	3718	587-134 G>A (in omozigosi)	/	splice site	transizione	1	U	this report
IVS 4c	3715	655-8 delATTTAT	/	splice site	delezione	1	U	Serra <i>et al.</i> , 2001
IVS 5	3236	730-5 T>G	/	splice site	trasversione	1	S	this report
IVS 6	3608	888+2 T>A	/	splice site	trasversione	1	U	this report
IVS 7	3165	1062+3 A>G	/	splice site	transizione	1	U	this report
IVS 8	2620	1185+1 G>A	/	splice site skip exon 8 (IF-41aa)	transizione	1	F	Hoffmeyer <i>et al.</i> , 1995
IVS 8	3548	1185+2 T>G	/	splice site	trasversione	1	U	this report
IVS 9	3418	1261-13 A>G	/	splice site	transizione	1	S	this report
IVS 9a	3158	1261-19 G>A	/	splice site	transizione	1	U	this report
IVS 10a	4821	1392+1 G>A	/	splice site	transizione	1	S	De Luca <i>et al.</i> , 2004
IVS 10b	3458	1527+1 delGTAA	/	splice site	delezione	1	F	this report
11	3195	1721 G>A	S547N	skip of exon11(?)out of frame;PTC 560(missense) splice site	transizione	1	U	Fashold <i>et al.</i> , 2000

Esoni Introne	Numero Identificativo	Alterazione nucleotidica	Alterazione aminoacidica	Previsione prodotto genico	Tipo di mutazione	Ricorrenza di mutazione	F/S	Referenze
IVS 11	2991	1722-3 C>T	/	splice site	transizione	1	U	this report
IVS 12b	3335	2002-14 C>T	/	splice site	transizione	1	F	this report
IVS 13	2882	2252-1 G>A	/	splice site	transizione	1	S	Ruggieri, <i>in press</i>
IVS 13	3296	2252-2 A>G	/	splice site	transizione	1	U	this report
IVS 14	3506	2325+2insT	/	splice site	inserzione	1	U	this report
16	2235	2764 G>A	G922S	Cryptic 5' ss/2761del90 missense/splice 2788/(IF -30aa)	trasversione	1	F	Ars <i>et al.</i> , 2003
IVS 17	3028	2991-1 G>A	/	splice site skip exon18 (IF-41aa)	transizione	2	S	Fashold <i>et al.</i> , 2000
IVS 17	4040	2991-1 G>A	/	splice site skip exon18 (IF-41aa)	transizione	2	S	Fashold <i>et al.</i> , 2000
IVS 18	2844	3113+1 G>T	/	splice site skip exon18 (IF-41aa)	trasversione	2	S	Ars <i>et al.</i> , 2003
IVS 18	3204	3113+1 G>T	/	splice site skip exon18 (IF-41aa)	trasversione	2	U	Ars <i>et al.</i> , 2003
IVS 18	3260	3114-27 G>C	/	splice site	trasversione	1	U	this report
IVS 20	3089	3496+1 G>A	/	splice site	transizione	1	S	this report
IVS 20	3580	3496+3 G>T	/	splice site	trasversione	1	U	this report
IVS 21	2572	3709-1 G>A	/	splice site	transizione	1	S	this report
IVS 24	3192	4269+1 G>A	/	splice site	transizione	1	U	this report
IVS 27b	2939	4772+2 T>G	/	splice site	trasversione	1	U	this report
IVS 31	3049	5943+1 G>C	/	splice site	trasversione	1	F	this report
IVS 31	3055	5944-5 A>G	/	splice site skip exon 32 (IF-47aa)	transizione	2	S	Fashold <i>et al.</i> , 2000
IVS 31	2900	5944-5 A>G	/	splice site skip exon 32 (IF-47aa)	transizione	2	S	Fashold <i>et al.</i> , 2000
IVS 34	2893	6579+1 G>C	/	splice site	trasversione	1	S	this report
IVS 35	3437	6641+1 G>A	/	splice site	transizione	1	U	De Luca <i>et al.</i> , 2004
IVS 35	3625	6641+1 G>C	/	splice site	trasversione	1	F	this report
IVS 36	2616	6756+1 G>A	/	Skip of exon 36; out of frame; PTC 2220	transizione	1	U	Fashold <i>et al.</i> , 2000
IVS 41	3719	7394+1 G>C	/	splice site	trasversione	1	U	this report
IVS 43	3637	7676-2 A>T	/	splice site	trasversione	1	U	Osborn <i>et al.</i> , 1999
IVS 47	3234	8097+55 T>C	/	splice site	transizione	1	U	this report

**Tabella 15.** Mutazioni a carico dei siti regolatori dello *splicing*.

Nel caso delle mutazioni osservate *de novo*, per verificare che siano effettivamente variazioni che inducono modificazioni al meccanismo di *splicing*, sono stati eseguiti dei test *in silico* mediante l'utilizzo di *software* disponibili in rete come: *BDGP: Splice Site Prediction by Neural Network* ([http://www.fruitfly.org/seq\\_tools/splice.html](http://www.fruitfly.org/seq_tools/splice.html)) e *NetGene 2 Server* (<http://www.cbs.dtu.dk/services/NetGene2>) che analizzano variazioni a carico delle sequenze consensus GT-AG dello *splicing*, *Human Splicing Finder Version 2.4* (<http://www.umd.be/HSF/>) e *ESEfinder Release 3.0* (<http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home>) che, basandosi su modelli di calcolo predittivi, prendono in considerazione anche le alterazioni delle sequenze ESE (*Exonic Splicing Enhancer*), ESS (*Exonic Splicing Silencer*), ISE (*Intronic Splicing Enhancer*) e ISS (*Intronic Splicing Silencer*).

In particolare quest'ultimo *software*, sulla base di algoritmi matematici, esegue un'analisi in grado di riconoscere le sequenze *enhancer* (elementi in **cis** presenti nel genoma) che sono riconosciute in modo specifico da proteine coinvolte nella costruzione del complesso proteico dello *splicing*. Nello specifico del *software* vengono valutate sequenze ESE per le proteine; SF2/ASF, SC35, SRp40 e SRp55.

In figura 30 sono riportati come esempio i risultati dell'analisi con *ESEfinder* ottenuti confrontando la sequenza *wild type* con quella contenente la mutazione 1261-19 G>A nell'introne 9a.

SC35 Thr=2.383			SRp40 Thr=2.67			SC35 Thr=2.383			SRp40 Thr=2.67		
Position	Motif	Score	Position	Motif	Score	Position	Motif	Score	Position	Motif	Score
4624	GTTTTATA	2.518274	4741	TTAAAAC	3.032857	4612	GCTCACTA	2.924335	4726	ACAAAAG	3.137931
4646	AGTTAGTG	2.512089	4747	CTACAGT	3.273917	4624	GTTTTATA	2.518274	4741	TTAAAAC	3.032857
4718	AGTTCCCG	3.255373	4759	ACAGAGC	2.707824	4646	AGTTAGTG	2.512089	4747	CTACAGT	3.273917
4719	GTTCCCGA	2.465149	4766	ATACAAC	3.552263	4718	AGTTCCCG	3.255373	4759	ACAGAGC	2.707824
4745	AACTACAG	2.612367	4771	ACTCACG	3.921889	4719	GTTCCCGA	2.465149	4766	ATACAAC	3.552263
4803	GGTCTTIG	3.394470	4798	CTATTGG	3.025295	4745	AACTACAG	2.612367	4771	ACTCACG	3.921889
4831	GATTGGTG	3.478045	4815	TCTCTAG	4.463574	4831	GATTGGTG	3.478045	4815	TCTCTAG	4.463574
4838	GGCCTAAG	3.392510	4863	TCACCTCG	5.773057	4838	GGCCTAAG	3.392510	4863	TCACCTCG	5.773057
4845	GATTGATG	3.116558	4927	GCACACC	2.671808	4845	GATTGATG	3.116558	4927	GCACACC	2.671808
4887	GTTTGGTG	3.191219	4942	CGAATGG	3.331411	4887	GTTTGGTG	3.191219	4942	CGAATGG	3.331411

A)

B)

**Figura 30.** A) Sequenza *wild type*: sono evidenziate in rosa le potenziali sequenze ESE in riferimento al punto di mutazione. B) Sequenza mutata: le sequenze ESE presenti sulla sequenza *wild type* sono scomparse in seguito alla variazione 1261-19 G>A

Le mutazioni *de novo* riportate in tabella 15 hanno evidenziato tutte, un'alterazione delle regioni coinvolte nel meccanismo di *splicing* quando sono state sottoposte ad analisi *in silico* con i *software* citati, ma è fondamentale ricordare che questi risultati sono frutto di analisi predittive che si basano totalmente su calcoli statistici.

L'unico modo possibile per poter valutare l'effetto reale di queste variazioni nucleotidiche, a livello del DNA genomico, è quello di eseguire un'analisi di espressione dei trascritti.

### 3.4 Costruzione del *beta-globin minigene*

#### 3.4.1 Amplificazione del gene della $\beta$ -globina

L'amplificato è stato ottenuto a partire dal DNA genomico di un soggetto sano. I *primer* utilizzati sono stati disegnati con il *software Primer 3* (<http://frodo.wi.mit.edu/primer3>).

**primer Forward:** cttctGCTAGCagtcagggcagagccatcta

**primer Reverse:** cttctGGGCCctttgcagcctcacctctttt

L'amplificato che ne deriva è di 1782 bp

I *primer* sono stati disegnati a monte e a valle del gene della  $\beta$ -globina, in modo da amplificare tutta la regione codificante e i relativi introni (figura 31).

### Homo sapiens $\beta$ -globin Genomic Location: 11p15.5

```
ATATATCTTAGAGGGAGGGCTGAGGGTTTGAAGTCCAACCTCCTAAGCCAGTGCCAGAAGAGCCAAG
GACAGGTACGGCTGTTCATCACTTAGACCTCACCCCTGTGGAGCCACACCCTAGGGTTGGCCAATCTA
CTCCCAGGAGCAGGGAGGGCAGGAGCCAGGGCTGGGCATAAAAAGTCAGGGCAGAGCCATCTATTGC
TTACATTTGCTTCTGACACAACCTGTGTTCACTAGCAACCTCAAACAGACACCATGGTGCATCTGAC
TCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGA
GGCCCTGGGCAGGTTGGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAACTGGGCATG
TGGAGACAGAGAAGACTCTTGGGTTTCTGATAGGCACTGACTCTCTCTGCCTATTGGTCTATTTTC
CCACCCTTAGGCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTCCCTTTGGGGATCTGT
CCACTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGGTGCCCT
TTAGTGATGGCCTGGCTCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCAC
GTGACAAGCTGCACGTGGATCCTGAGAACTTCAGGGTGAGTCTATGGGACGCTTGATGTTTTCTTT
CCCCTTCTTTTCTATGGTTAAGTTCATGTTCATAGGAAGGGGATAAGTAACAGGGTACAGTTTAGAA
TGGGAAACAGACGAATGATTGCATCAGTGTGGAAAGTCTCAGGATCGTTTTAGTTTCTTTTATTGGC
TGTTTATAACAATTGTTTTCTTTTGTTTAATTCTTGCTTTCTTTTTTTTTCTTCTCCGCAATTTTT
ACTATTATACTTAATGCCTTAACATTGTGTATAACAAAAGGAAATATCTCTGAGATACATTAAGTA
ACTTAAAAAAAACCTTTACACAGTCTGCCTAGTACATTACTATTTGGAATATATGTGTGCTTATTT
GCATATTCATAATCTCCCTACTTTATTTTTCTTTTATTTTTAATTGATACATAATCATTATACATAT
TTATGGGTTAAAGTGTAATGTTTTAATATGTGTACATATTGACCAAATCAGGGTAATTTTGCAT
TTGTAATTTTAAAAAATGCTTTCTTTTAAATATACTTTTTTGTATTATCTTATTTCTAATACTTT
CCCTAATCTCTTTCTTTTCAAGGCAATAATGATACAATGTATCATGCCTCTTTGCACCATTCTAAAG
AATAACAGTGATAATTTCTGGGTTAAGGCAATAGCAATATCTCTGCATATAAAATATTTCTGCATAT
AAATTGTAAGTGTGTAAGAGGTTTCATATTGCTAATAGCAGCTACAATCCAGCTACCATTCTGCT
TTTATTTTATGGTTGGGATAAGGCTGGATTATTCTGAGTCCAAGCTAGGCCCTTTTGCTAATCATG
TTCATACCTCTTATCTTCTCCACAGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCATCAC
TTTGGCAAAGAATTCACCCACCAGTGCAGGCTGCCTATCAGAAAAGTGGTGGCTGGTGTGGCTAAT
GCCCTGGCCACAAGTATCACTAAGCTCGCTTTCTTGCTGTCCAATTTCTATTAAGGTTCCCTTTG
TTCCCTAAGTCCAACCTACTAACTGGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTA
ATAAAAAACATTTATTTTTCATTGCAATGATGTATTTAAATTTATTTCTGAATATTTTACTAAAAAGG
GAATGTGGGAGGTCAGTGCATTTAAACATAAAGAAATGAAGAGCTAGTTCAAACCTTGGGAAAAT
ACACTATATCTTAAACTCCATGAAAAGAAGGTGAGGCTGCAAAACAGCTAATGCACATTGGCAACAGC
CCCTGATGCATATGCCTTATTCATCCC
```

**Figura 31.** Sequenza dell'amplificato: in **azzurro** sono evidenziati gli esoni, in **nero** le regioni non codificanti, in **verde** i *primer* utilizzati per amplificare il gene della  $\beta$ -globina, in **rosa** la sequenza del sito di taglio dell'enzima di restrizione BsrGI.

I *primer* presentano alle loro estremità 5' la regione di taglio per due specifici enzimi di restrizione (*primer forward*: NheI, *primer reverse*: ApaI) necessarie per clonare il gene della  $\beta$ -globina nel vettore plasmidico.

La scelta del gene della  $\beta$ -globina è stata motivata dal fatto che essendo un gene relativamente piccolo, con pochi esoni (3 esoni codificanti), è stato possibile amplificare, in un'unica reazione, tutto il gene.

L'amplificazione mediante PCR (*Polymerase Chain Reaction*) è stata eseguita, in un volume di 50  $\mu$ l, secondo il protocollo riportato in tabella 16, utilizzando *Expand Long Template PCR System* (Roche) e sottoponendo la soluzione di reazione al ciclo di temperature descritto in tabella 17.

Mix di reazione	Volume per 1 reazione ( $\mu$ l)
H <sub>2</sub> O sterile MilliQ	31,25
Buffer 1 10X	5
dNTP (10 mM)	1,75
pForward (10 $\mu$ M)	2,5
pReverse (10 $\mu$ M)	2,5
BSA 10X	2
Taq Polimerase	1
Volume finale	46 $\mu$ l mix + 4 $\mu$ l DNA

**Tabella 16.** Mix della reazione di amplificazione.

T (°C)	Tempo	N° cicli
94	2'	1
94	10"	15
65	30"	
68	3'	
94	15"	20
65	30"	
68	3' (+ 20"/ciclo)	
68	7'	1
4	10'	1

**Tabella 17.** Profilo termico della reazione.



La resa della reazione è stata valutata sottoponendo 5 µl di prodotto di amplificazione a corsa elettroforetica su gel d'agarosio all'1% colorato con etidio bromuro (figura 32).



**Figura 32.** Amplificazione del gene della  $\beta$ -globina.

Per escludere la presenza di eventuali mutazioni, indotte dalla reazione di PCR, l'amplificato è stato sequenziato (vedi paragrafo 3.3.3).

### 3.4.2 Clonaggio del gene della $\beta$ -globina in un vettore plasmidico

**3.4.3.1 Digestione enzimatica:** l'amplificato del gene della  $\beta$ -globina purificato e quantificato e il vettore plasmidico commerciale *pcDNA3.1/Hygro(+)* (Invitrogen) sono stati sottoposti a digestione enzimatica con gli enzimi NheI e ApaI (New England BioLabs) secondo protocollo:

Mix di reazione	Volume per singola reazione
DNA $\beta$ -globina/pcDNA3.1 Hygro(+)	4 µg
Nebuffer 4 10X	1X
BSA 100X	1X
NheI (10U/µl)	10U x 1 µg
Incubare a 37°C per 4 ore	
ApaI (50U/µl)	50U x 1 µg
Incubare 25°C over night	
Dephosphorylation buffer 10x (Roche)	1X
Alkaline Phosphatase, shrimp 1U/µl (Roche)	3 µl
Incubare a 37°C per 1 ora e 65°C per 15 minuti	

**Tabella 18.** Protocollo digestione e defosforilazione.

Per verificare l'esito della digestione 5 µl di prodotto di digestione sono stati sottoposti ad elettroforesi orizzontale su gel d'agarosio all'1% colorato con etidio bromuro.

Il vettore *pcDNA3.1/Hygro(+)* e l'amplificato del gene della β-globina digeriti sono stati purificati con *Microcon Centrifugal Filter Devices* (Millipore) e quantificati mediante lettura spettrofotometrica a 260 nm.

**3.4.2.2 Ligazione:** la reazione di ligazione è stata eseguita considerando un rapporto molare 3:1 tra inserto e vettore (per un totale di 300-400 ng), applicando la formula:

$$(\text{ng di vettore} \times \text{k basi di inserto} / \text{k basi di vettore}) \times (3/1) = \text{ng di inserto}$$

Per la reazione è stato usato l'enzima *T4 DNA Ligase* (New England BioLabs) seguendo il protocollo di reazione fornito dal produttore nel quale è indicato di utilizzare 1 µl di enzima in 20 µl di reazione. La reazione di ligazione è stata eseguita mescolando il vettore e l'inserto digeriti, all'enzima e al suo *buffer* (1X del volume finale di reazione) e sono stati lasciati incubare a 16°C *over night*.

### 3.4.3 Trasformazione dei batteri

Con 2 µl del prodotto di ligazione sono stati trasformati 50 µl di cellule batteriche *E. Coli DH5α T1R* (Invitrogen) secondo il protocollo che prevede:  
30 minuti in ghiaccio,  
30 secondi *heat shock* a 42° C,  
2 minuti in ghiaccio,  
aggiunta di 250 µl di *SOC medium* a temperatura ambiente,  
incubazione a 37° C per 1 ora a 300-400 rpm di oscillazione.

Al termine dell'incubazione la coltura è stata suddivisa in due aliquote corrispondenti al 10% e al 90% del volume finale di coltura, distribuita su due piastre di LB agar contenenti 50-100 µg/ml di ampicillina e incubata a 37°C *over night*.

Avvenuta la crescita delle colonie, queste sono state isolate e lasciate in coltura per una notte in 5 ml di LB liquido + 5 µl di Ampicillina 1000X a 300-400 rpm.

I plasmidi sono stati estratti dalle colture liquide di cellule batteriche seguendo il protocollo fornito con *Illustra plasmidPrep Mini Spin Kit* (GE Healthcare) e sono stati risospesi in 100 µl di H<sub>2</sub>O MilliQ.

Per verificare la presenza dell'insero nei plasmidi estratti da diverse colonie, è stata eseguita una PCR di controllo utilizzando, per l'amplificazione, dei *primer* complementari a dei siti specifici del plasmide commerciale: T7 *promoter/primer binding site* e BGH *reverse priming site*:

T7 *promoter primer*: TAATACGACTCACTATAGGG

BGH *reverse primer*: CCTCGACTGTGCCTTCTA

Il protocollo di amplificazione è stato:

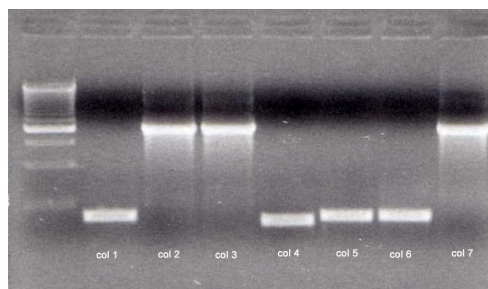
Mix di reazione	Volume per 1 reazione (µl)
Acqua MilliQ sterile	21,5
Buffer 10X	5
dNTP (10mM)	1
pF T7 (1µl)	10
pR BGH (1µl)	10
Taq polimerase (Roche)	0,5
Volume finale	48 µl di mix + 2 µl di miniprep

**Tabella 19.** Mix di amplificazione.

T. (°C)	Tempo	N° di cicli
94	12 min	1
94	1 min	35
55	1min	
72	2 min	
72	10 min	1
4	10 min	1

**Tabella 20.** Profilo termico della reazione.

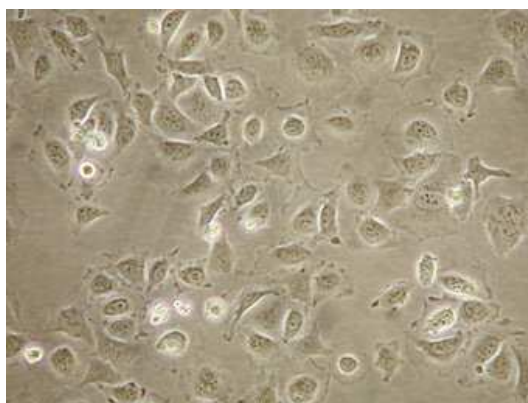
Le dimensioni dell'amplificato, se presente nel plasmide il gene della  $\beta$ -globina, corrispondono a 1863 bp ed è stato verificato su gel d'agarosio all'1% mediante corsa elettroforetica (figura 33); definitiva conferma sulla qualità del vettore plasmidico creato si è ottenuta mediante sequenziamento diretto dell'inserito a partire dagli amplificati del plasmide.



**Figura 33.** L'amplicone derivato dai plasmidi con inserito il gene della  $\beta$ -globina hanno una dimensione di 1863 bp.

### 3.4.4 Trasfezione transiente di cellule eucariotiche

Selezionato il vettore con inserito il gene della  $\beta$ -globina, 400 ng di plasmide purificato, sono stati utilizzati per trasfettare colture di  $2 \times 10^5$  cellule eucariotiche (cellule HeLa) in adesione secondo quanto indicato dal protocollo *Effectene Transfection Reagent* (Qiagen).



**Figura 34.** Cellule HeLa in adesione (ingrandimento 20X).

Dopo aver costituito il complesso DNA-*Effectene* e averlo distribuito sulle cellule HeLa le piastre sono state incubate a 37°C per 48 ore. In coltura sono state

messe anche cellule non trasfettate con il minigene per verificare che sia effettivamente avvenuta la formazione del trascritto del gene della  $\beta$ -globina.

### **3.4.5 Estrazione dell'RNA totale e analisi del trascritto**

**3.4.5.1 Estrazione dell'RNA totale:** dalle colture cellulari trasfettate è stato estratto l'RNA totale con *TRIzol Reagent* (Invitrogen).

Il *TRIzol* è una soluzione contenente fenolo in grado di fissare le cellule a temperatura ambiente e che permette l'estrazione degli acidi nucleici. Usato in eccesso il *TRIzol* minimizza il problema della contaminazione proteica molto frequente nella procedura di estrazione degli acidi nucleici. Il protocollo consiglia l'uso di 1,5 ml di *TRIzol* ogni 10 milioni di cellule.

Uno dei passaggi chiave della metodica di estrazione di RNA con *TRIzol*-cloroformio è al momento dell'aggiunta del cloroformio in cui, in seguito a centrifuga fredda (4°C) per 15 minuti a 12000 rpm, si ha la formazione di un surnatante limpido in cui è disciolto l'RNA, una fase sottostante opaca in cui è disciolto il DNA e alla base della provetta i residui di estrazione.

Isolato, l'RNA è stato precipitato a freddo con isopropanolo e purificato con etanolo 70% e 100%, infine viene risospeso in acqua sterile trattata con DEPC (dietilpirocarbonato) per inibire le RNasi.

**3.4.5.2 Sintesi del cDNA mediante retrotrascrizione:** 1  $\mu$ g di RNA totale è stato retrotrascritto usando l'enzima *SuperScript II Reverse Transcriptase* (Invitrogen) secondo il protocollo fornito con l'enzima che prevede in un volume finale di 20  $\mu$ l:

Mix di reazione	Volume per 1 reazione
RNA totale	1 µg
Oligo (dT) <sub>12-18</sub> (500 µg/ml) o 50-250 ng random primer	1 µl
dNTP (10 mM)	1 µl
H <sub>2</sub> O-DEPC sterile	a 12 µl
incubare a 65°C per 5 minuti	
5X First-Strand Buffer	4 µl
0,1 M DTT	2 µl
RnaseOUT™ (40 units/µl)	1
mescolare e incubare a 42°C per 2 minuti	
SuperScript™ II RT	1 µl (200 unità)
incubare a 25°C per 10 min. (solo per random primer) - 42°C per 50 min. - 70°C per 15 min.	

**Tabella 21.** Protocollo di retrotrascrizione.

**3.4.5.3 Amplificazione del cDNA e sequenziamento:** il cDNA, ottenuto per retrotrascrizione dall'RNA estratto dalle cellule trasfettate e dalle cellule controllo non trasfettate, è stato amplificato con primer specifici per regioni codificanti nel gene della  $\beta$ -globina, in particolare il *primer forward* all'interno dell'esone 2 e il *primer reverse* nell'esone 3 del gene:

*Primer ex2  $\beta$ -glob Forward:* TCTGTCCACTCCTGATGCTG

*Primer ex3  $\beta$ -glob Reverse:* CACTGGTGGGGTGAATTCTT

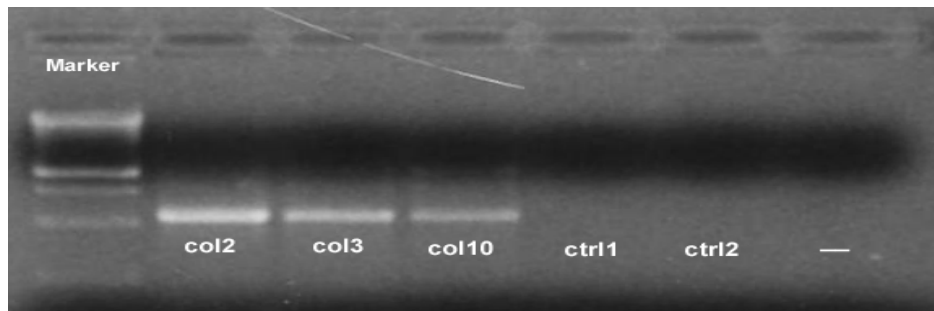
Mix di reazione	Volume per 1 reazione (µl)
H <sub>2</sub> O sterile	39,5
buffer 10X	5
primer ex2 Forward (10 µM)	1
primer ex 3 Reverse (10 µM)	1
dNTP (10 mM)	1
Taq Polimerase (Roche)	0,5
Volume finale	48 µl di mix + 2 µl di RT-PCR

**Tabella 22.** Mix di amplificazione.

T. (°C)	Tempo	N° di cicli
94	3'	1
94	1'	35
55	1'	
72	45'	
72	7'	1
4	10'	1

**Tabella 23.** Profilo termico della reazione.

L'esito della retrotrascrizione è stato controllato su gel d'agarosio all'1% mediante corsa elettroforetica.



**Figura 35.** cDNA ottenuto dalla retrotrascrizione di RNA prodotto da cellule HeLa trasfettate con il  $\beta$ -globin minigene estratto da tre differenti colonie batteriche (col 2, col 3, col 10); ctrl 1 e ctrl 2 sono i controlli per confermare che le cellule non trasformate non sono in grado di produrre il trascritto per la  $\beta$ -globina.

La presenza del trascritto, solo nelle cellule trasfettate, ha quindi confermato la funzionalità del minigene plasmidico per lo studio dei meccanismi coinvolti nei processi di espressione genica.

### 3.4.6 Sintesi della regione “adattatore” e inserimento nel vettore

Il passo successivo è stato sintetizzare e inserire nel plasmide una regione "adattatore", generata dall'*annealing* di due oligonucleotidi di 33 basi nella cui sequenza sono stati inseriti i siti di taglio per quattro enzimi di restrizione (BsrGI, XhoI, NotI, HindIII):

**SITI DI RESTRIZIONE NELLA SEQUENZA ADATTATORE:**

**BsrGI**      **XhoI**                      **NotI**                                      **HindIII**      **BsrGI**  
**TgtacaCTCGAGttatGCGGCCGCtactAAGCTTGTACA**

Si è creato in questo modo un sito specifico per l'inserimento degli amplificati da testare.

I due oligonucleotidi:

*BsrGI Adaptor Forward*      [Phos]GTACACTCGAGTTATGCGGCCGCTACTAAGCTT

*BsrGI Adaptor Reverse*      [Phos]GTACAAGCTTAGTAGCGGCCGCATAACTCGAGT

sono stati mescolati tra loro (una concentrazione di 1  $\mu$ M per entrambi) in una soluzione tampone (*10X Annealing buffer* Promega) e sottoposti ad un protocollo di denaturazione - rinaturazione che prevede un riscaldamento iniziale di 5 minuti a 70°C e poi un graduale raffreddamento fino a 37°C , abbassando la temperatura di 1,5°C al minuto.

**ADATTATORE:**

GTACACTCGAGttatGCGGCCGCtactAAGCTT  
tgagctcaatacgcgcgatgattcgaacatg

Alle estremità della regione adattatore, così sintetizzata, sono stati generati i siti di taglio dell'enzima BsrGI (5'...t▼gtaca...3') in quanto tale enzima, presenta un unico sito di restrizione all'interno del  *$\beta$ -globin minigene* (figura 31) in particolare, nel mezzo dell'introne 2 del gene della  $\beta$ -globina.

In pratica: 5  $\mu$ g del plasmide  $\beta$ -globina-pcDNA 3.1(+) sono stati digeriti con l'enzima BsrGI (New England BioLabs) e defosforilati secondo il protocollo:



Mix di reazione	Volume per singola reazione
plasmide $\beta$ -globina-pcDNA3.1 Hygro(+)	4 $\mu$ g
Nebuffer 2 10X	1X
BSA 100X	1X
BsrGI 10U/ $\mu$ l	10U x 1 $\mu$ g
Incubare a 37°C over night	
Dephosphorylation buffer 10x (Roche)	1X
Alkaline Phosphatase, shrimp 1U/ $\mu$ l (Roche)	3 $\mu$ l
Incubare a 37°C per 1 ora e 65°C per 15 minuti	

**Tabella 24.** Protocollo digestione e defosforilazione.

In un volume finale di 20  $\mu$ l, per azione della *T4 DNA Ligase* (New England BioLabs) sono stati ligati tra loro 450 ng di prodotto di digestione con diverse quantità di “adattatore” sintetizzato (mix 1=10 ng; mix 2= 20 ng; mix 3= 40 ng) lasciando avvenire la reazione per 3 ore a temperatura ambiente (20-25°C).

Le ligazioni sono poi servite a trasformare 3 aliquote di cellule batteriche *E. Coli DH5 $\alpha$  T1R* (Invitrogen) con la metodica descritta precedentemente (paragrafo 3.4.3).

Isolati i plasmidi, derivanti da numerose colonie formatesi sulle piastre di coltura e purificati con *Illustra plasmidPrep Mini Spin Kit* (GE Healthcare), sono stati sottoposti a reazione di PCR con *primer* specifici per le regioni fiancheggianti l'adattatore inserito all'interno dell'introne 2 del gene della  $\beta$ -globina:

*primer*  $\beta$ glob INTR *Forward*: CAGTCTGCCTAGTACATTACTATTTGG

*primer*  $\beta$ glob INTR *Reverse*: TTGCCCTGAAAGAAAGAGATT

Le condizioni di lavoro applicate sono state:

Mix di reazione	Volume per 1 reazione (µl)
H <sub>2</sub> O sterile	39,5
buffer 10X	5
primer ex2 <i>Forward</i> (10 µM)	1
primer ex 3 <i>Reverse</i> (10 µM)	1
dNTP (10 mM)	1
Taq Polimerase (Roche)	0,5
Volume finale	48 µl di mix + 2 µl di plasmide

**Tabella 25.** Mix di amplificazione.

T. (°C)	Tempo	N° di cicli
94	3'	1
94	1'	35
55	1'	
72	45'	
72	7'	1
4	10'	1

**Tabella 26.** Profilo termico della reazione.

Oltre a verificarne la presenza all'interno del plasmide è stato necessario valutare che “l'adattatore” fosse all'interno della sequenza in singola copia e che non vi fossero errori nella sequenza all'interno del minigene.

Per verificare la presenza in singola copia della sequenza adattatore è stata eseguita una digestione enzimatica, a 37°C *over night*, con 10Ux1µg dell'enzima di restrizione HindIII (New England Biolabs) che presenta un unico sito di taglio all'interno del minigene proprio all'interno della sequenza dell'adattatore.

Infine, per valutare la corretta sequenza all'interno dei plasmidi selezionati, gli amplificati sono stati sottoposti a sequenziamento diretto.

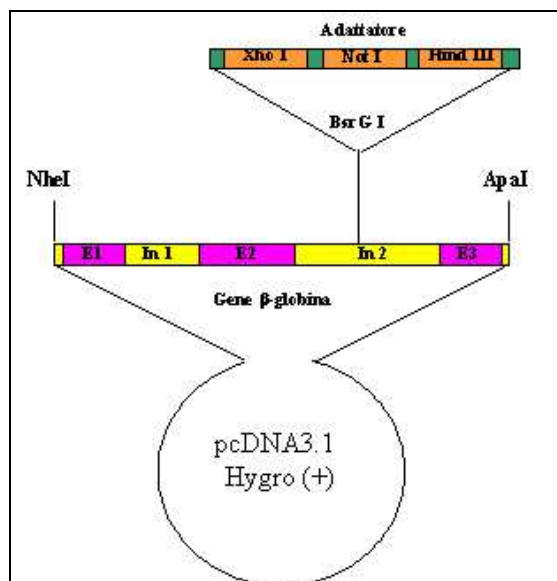


Figura 36. Beta-globin minigene.

### 3.4.7 Osservazione dei trascritti

Terminata la prima fase del progetto in cui si è provveduto alla costruzione del minigene plasmidico è stato necessario valutare la funzionalità della metodica.

Per analizzare *in vitro* le mutazioni genomiche osservate nei pazienti, per ogni singolo caso, è stato necessario disegnare dei *primer* di amplificazione complementari alle regioni, esoniche e introniche, fiancheggianti la mutazione nel gene interessato.

Fondamentale è che questi *primer* presentino alla loro estremità 5' la sequenza dei siti di taglio per due degli enzimi di restrizione che digeriscono nella regione “adattatore”. Nella scelta dei due enzimi è importante considerare che non vi siano ulteriori siti di taglio, per quest'ultimi, all'interno della regione da amplificare. La presenza di queste regioni, alle estremità dell'amplificato, permette di clonare la sequenza da studiare all'interno del minigene in modo sito specifico e direzionalmente specifico.

Digerendo, sia il minigene plasmidico che l'amplicone del gene da studiare, con i due enzimi selezionati: si asporta la sequenza “adattatore” e si linearizza il vettore; mentre nei prodotti di PCR, ottenuti da DNA genomico del paziente in analisi, si creano le estremità complementari per quelle del vettore linearizzato.



In questo caso particolare al *primer Reverse* non è stata aggiunta artificialmente la sequenza per il sito di taglio dell'enzima, perchè tali nucleotidi sono già presenti nel tratto di introne 18 del gene *NF1* e quindi è stato sfruttato il naturale sito di taglio per l'enzima (figura .37).

L'amplificazione è stata eseguita con l'enzima *AmpliTaq Gold Polimerase* (Applied Biosystem) applicando le condizioni descritte nelle tabelle 26 e 27:

Mix di reazione	Volume per 1 reazione (µl)
H <sub>2</sub> O MilliQ sterile	31,7
buffer 10X	5
MgCl <sub>2</sub>	3
dNTP (10 mM)	1
pFOR (10 µM)	2,5
pREV (10 µM)	2,5
Taq Gold	0,3
Volume finale	46 µl di mix + 4 µl di DNA

**Tabella 26.** Mix di amplificazione.

T. (°C)	Tempo	N° di cicli
95	10 min	1
95	30 sec	35
60	40 sec	
72	1 min	
72	10 min	1
4	10 min	1

**Tabella 27.** Profilo termico della reazione.

L'esito dell'amplificazione è stato verificato su gel di poliacrilammide al 10% e colorato con etidio bromuro (figura 38).



**Figura 38.** Amplificazione del DNA genomico del paziente.

L' amplificato è stato purificato con *Microcon Centrifugal Filter Devices* (Millipore) e quantificato con lettura spettrofotometrica a 260 nm.

Sia il prodotto di amplificazione purificato che il  $\beta$ -globin minigene sono stati digeriti prima con l'enzima HindIII e successivamente con l'enzima XhoI, secondo il protocollo:

Mix di reazione	Volume per singola reazione
amplificato ex18 <i>NF1</i> / $\beta$ -globin minigene	4 $\mu$ g
Nebuffer 2 10X	1X
HindIII 20U/ $\mu$ l	20U x 1 $\mu$ g
Incubare a 37°C per 4 ore	
XhoI 20U/ $\mu$ l	20U x 1 $\mu$ g
BSA 100X	1X
Incubare 37°C per 4 ore	

**Tabella 28.** Protocollo digestione.

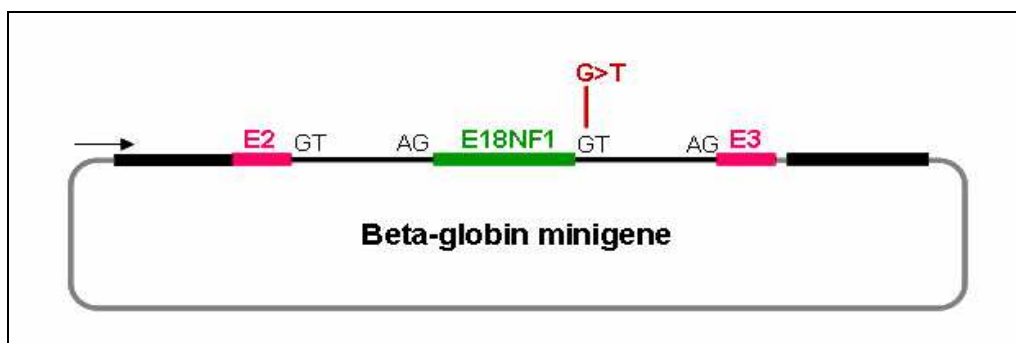
Le digestioni sono state purificate con *Microcon Centrifugal Filter Devices* (Millipore) e sottoposti a corsa elettroforetica su gel di acrilammide al 10 % per verificare l'avvenuta digestione.

Linearizzato il plasmide e costituite le estremità complementari nell'amplicone dell'esone 18 del gene di *NF1*, è stato possibile unirli insieme mediante reazione di ligazione per azione della *T4 DNA Ligase* (New England

BioLabs), come da protocollo (paragrafo 3.4.2.2).

Il vettore ottenuto è quindi un potenziale minigene che esprime un trascritto costituito dai 3 esoni della  $\beta$ -globina insieme all'esone 18 di *NF1*.

Come nella fase di preparazione del vettore (paragrafo 3.4.3) si sono trasformate le cellule batteriche *E. Coli DH5 $\alpha$  T1R* (Invitrogen) per poter isolare diverse colonie, estrarne il plasmide mediante *Illustra plasmidPrep Mini Spin Kit* (GE Healthcare) e verificarne la presenza dell'inserito di *NF1* mediante reazione di PCR con specifici *primer* complementari a regioni nell'introne 2 di  $\beta$ -globina, a monte e a valle dell'inserito.



**Figura 39.** Schema d'inserimento dell'esone 18 del gene *NF1* all'interno della regione adattatore del minigene plasmidico.

Stabilito quali fossero i plasmidi con inserito l'esone 18 di *NF1*, gli amplificati sono stati sottoposti a sequenziamento diretto per assicurarsi che non vi fossero errori nella sequenza indotti da eventi di ricombinazione con il genoma batterico e per verificare la presenza della mutazione da testare.

Questo è stato necessario perchè il DNA dei pazienti affetti da Neurofibromatosi di Tipo 1 presentano in eterozigosi la mutazione e di conseguenza danno origine a popolazioni di colonie con la sequenza *wild type* e colonie con la sequenza mutata.

Una volta selezionati i plasmidi, come descritto in precedenza, sia il minigene con la sequenza dell'esone 18 di *NF1* mutata, che quello *wild type*, sono stati usati per trasfettare cellule HeLa da cui estrarre l'RNA ed osservare il trascritto mediante retrotrascrizione e amplificazione del cDNA con *primer* complementari a regioni dell'esone 2 e 3 dell'RNA della  $\beta$ -globina (paragrafo 3.4.4 e 3.3.5).

E' possibile notare nel cDNA ottenuto dalla trasfezione effettuata con il minigene *wild type*, la presenza di un trascritto le cui dimensioni corrispondono a quelle del prodotto genico della  $\beta$ -globina sommato dell'esone 18 del gene *NFI* (tot 361 bp); mentre il trascritto ottenuto dal minigene mutato presenta un cDNA corrispondente alla lunghezza del prodotto genico della sola  $\beta$ -globina (238 bp).



**Figura 40.** Amplificati ottenuti da cDNA ottenuto per retrotrascrizione del prodotto di espressione del  $\beta$ -globin minigene contenente la sequenza dell'esone 18, e relative sequenze introniche, *wild type* e mutata.

**Amplificato ottenuto dal trascritto *wild type* (361 bp):**

TCTGTCCACTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGG  
 TGCCTTTAGTGATGGCCTGGCTCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCT  
 GCACTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGGGTATGTTTCGTGTGCTTGGGAATATGG  
 TCCATGCAATTCAAATAAAAAACGAAACTGTGTCAATTAGTTGAAGTAATGATGGCAAGGAGAGATG  
 ACCTCTCATTTTGCCAAGAGATGAAATTTAGCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCA  
 TCACTTTGGCAAGAATTCACCCCACCAGTG

**Amplificato derivante dal trascritto mutato (238 bp):**

TCTGTCCACTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGG  
 TGCCTTTAGTGATGGCCTGGCTCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCT  
 GCACTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGGCTCCTGGGCAACGTGCTGGTCTGTGT  
 GCTGGCCCATCACTTTGGCAAGAATTCACCCCACCAGTG

**Figura 41.** Sequenza in basi ottenuta mediante sequenziamento degli amplificati ottenuti dal cDNA ricavato dalle cellule trasfettate con il minigene; in **azzurro** gli esoni del gene della  $\beta$ -globina, in **nero** la sequenza dell'esone 18 di *NFI*, in **rosa** le sequenze dei *primer* usati per l'amplificazione.



# **Analisi e Risultati**



L'obiettivo preposto è stato quello di utilizzare il *β-globin minigene* per testare mutazioni, non riportate in letteratura, a carico del gene *NF1* osservate in pazienti affetti da Neurofibromatosi di Tipo 1.

#### 4.1 Mutazione *NF1* IVS5-5 T>G

Il DNA genomico del campione è stato amplificato con la seguente coppia di *primer*:

*primer Forward-XhoI*: cttctCTCGAGcacaactgcaaggcagagaa

*primer Reverse NotI*: cttctGCGGCCGCatgcgagagccttaccttc

che introducono rispettivamente un sito di taglio per l'enzima XhoI e l'enzima NotI.

L'amplificazione è stata eseguita, in un volume di 50 µl, secondo il protocollo riportato in tabella 29, utilizzando *Expand Long Template PCR System* (Roche) e sottoponendo la soluzione di reazione al ciclo di temperature descritto in tabella 30:

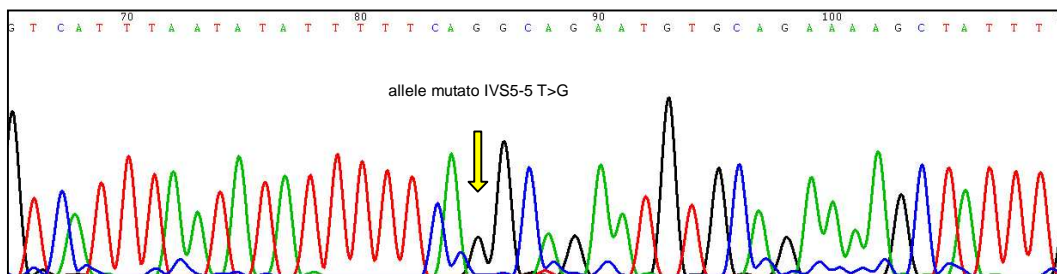
Mix di reazione	Volume per 1 reazione (µl)
H <sub>2</sub> O MilliQ sterile	27,25
Buffer 1 10X	5
dNTP (10 mM)	1,75
p <i>Forward</i> (10 µM)	2,5
p <i>Reverse</i> (10 µM)	2,5
BSA 10X	2
Taq Polimerase	1
Volume finale	42 µl mix + 8 µl DNA

**Tabella 29.** Mix della reazione di amplificazione.

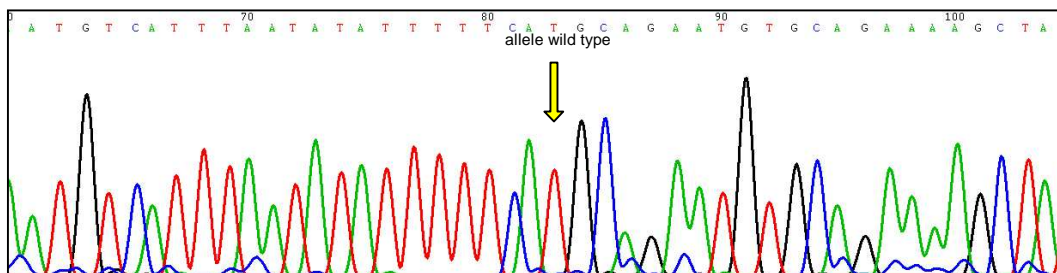
T (°C)	Tempo	N° cicli
94	2'	1
94	10"	15
67	30"	
68	3'	
94	15"	20
67	30"	
68	3' (+ 20"/ciclo)	
68	7'	1
4	10'	1

**Tabella 30.** Profilo termico della reazione.

Come da protocollo, mediante digestione e ligazione enzimatiche, l'amplificato con la mutazione in eterozigosi è stato inserito nel  $\beta$ -globin *minigene* e con il costrutto si sono trasformate cellule batteriche (*E. Coli* DH5 $\alpha$  T1R, Invitrogen) per isolare i cloni mutati e i cloni *wild type*.



**Figura 42.** Sequenza dell'amplicone ottenuto da un plasmide mutato.

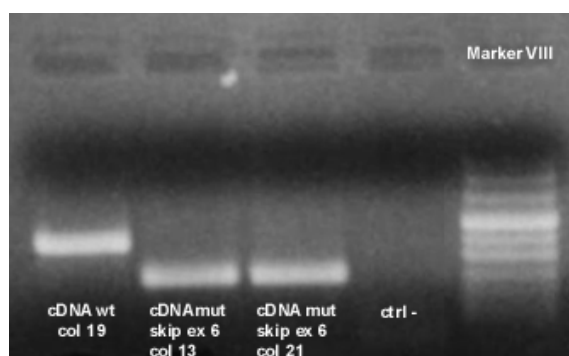


**Figura 43.** Sequenza dell'amplicone ottenuto da un plasmide wild type.

In seguito a sequenziamento diretto degli amplificati dei plasmidi estratti dalle colonie (figure 42 e 43), sono stati selezionati tre costrutti (uno con l'allele *wild type*, due con l'allele mutato) e utilizzati per trasformare cellule HeLa secondo il protocollo *Effectene Transfection Reagent* (Quiagen).

L'RNA totale estratto dalle cellule con *TRIzol* è stato retrotrascritto con *SuperScript II Reverse Transcriptase* (Invitrogen) e amplificato utilizzando un *primer forward* complementare all'esone 2 del gene della  $\beta$ -globina ed un *primer reverse* complementare all'esone 3 della  $\beta$ -globina. I prodotti di trascrizione sono poi stati visualizzati su gel di agarosio al 1,5%.

I risultati ottenuti dalla retrotrascrizione e sequenziamento dell'RNA estratto hanno evidenziato che il costrutto mutato presenta la delezione dell'esone 6 del gene *NFI*.



**Figura 44.** Amplificati derivati dal cDNA ottenuto per retrotrascrizione del prodotto di espressione del  *$\beta$ -globin minigene* contenente la sequenza dell'esone 6 di *NFI*.

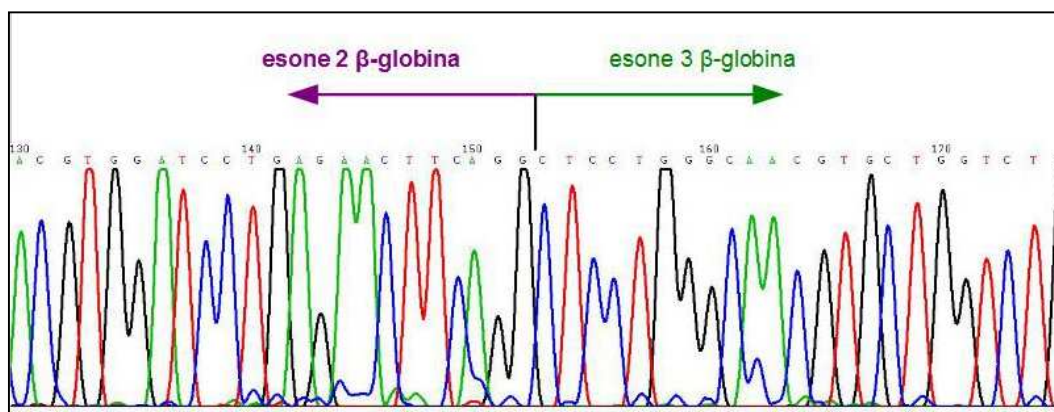
**Amplificato derivante dal trascritto wild type (396 bp):**

```
TCTGTCCACTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGG
TGCCTTTAGTGATGGCCTGGCTCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCT
GCACTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGGAATGTGCAGAAAAGCTATTTGACTTG
GTGGATGGTTTTGTGAAAGCACCAAACGTAAAGCAGCAGTTTGGCCACTACAAATCATTCCTCCTT
ATCTTGTGTCCAGAAATAATCCAGGATATATCCAAAGACGTGGTTGATGAAAACAACATGAATAAG
CTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCATCACTTTGGCAAAGAATTCACCCACCAGTG
```

**Amplificato derivante dal trascritto mutato (239 bp):**

```
TCTGTCCACTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGG
TGCCTTTAGTGATGGCCTGGCTCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCT
GCACTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGGCTCCTGGGCAACGTGCTGGTCTGTG
TGCTGGCCATCACTTTGGCAAAGAATTCACCCACCAGTG
```

**Figura 45.** Sequenza in basi ottenuta mediante sequenziamento degli amplificati ottenuti dal cDNA ricavato dalle cellule trasfettate con il minigene; in **azzurro** gli esoni del gene della  $\beta$ -globina, in **nero** la sequenza dell'esone 6 di *NFI*, in **rosa** le sequenze dei *primer* usati per l'amplificazione.



**Figura 46.** Parte dell'elettroferogramma della sequenza *forward* del trascritto derivante dal plasmide in cui era presente la mutazione IVS5-5 T>G a monte dell'esone 6 del gene *NFI*. E' possibile vedere il punto di unione dei due esoni del gene della  $\beta$ -globina in cui è avvenuto lo *skipping* dell'esone 6.

## 4.2 Un caso particolare: *CFTR* IVS 4+3 A>G

Durante un'indagine di *screening*, in un collaboratore del Servizio di Genetica Clinica ed Epidemiologica del Dipartimento di Pediatria, Università degli Studi di Padova, è stata osservata la variazione 621+3 A>G del gene *CFTR* in eterozigosi; è nata quindi la curiosità di valutare mediante l'uso del  $\beta$ -globin

*minigene*, oggetto di questo lavoro di dottorato, l'azione funzionale di questa variazione a livello del trascritto.

Tale variante è riportata come polimorfismo in *Cystic Fibrosis Mutation Database* (<http://www.genet.sickkids.on.ca/cftr/app>). Successivamente in letteratura è stato riportato il caso di quattro pazienti greci, gravemente affetti da Fibrosi Cistica, in cui 621+3 A>G induce un'alterazione nel meccanismo di *splicing* dell'esone 4 del gene *CFTR* con conseguente formazione di una proteina non funzionante (Tzetis *et al.*, 2001).

Ci si è posti quindi l'obiettivo di valutare, sia dal punto di vista epidemiologico che funzionale, se 621+3 A>G sia in realtà una variante che non si può considerare mutazione patogenetica.

#### 4.2.1 Dati di popolazione

La frequenza della variante 621+3 A>G è stata stimata sulla base dei dati ottenuti prendendo in considerazione un campione di 250 individui sani di controllo ai quali è stato chiesto di fornire il consenso informato per l'analisi genetica e l'uso dei risultati per eventuali ricerche epigenetiche.

Lo *screening* dei controlli è stato eseguito mediante *High Resolution Melting Analysis* (HRMA) effettuata su Rotorgene 6000 (Corbett Life Sciences). Le condizioni ottimali per l'amplificazione e l'analisi HRMA del frammento sono state messe a punto sperimentalmente.

I *primer* sono stati disegnati con il *software Primer 3*:

*primer CFTRex4rotor Forward*: CACATTGGAATGCAGATGAGA

*primer CFTRex4rotor Reverse*: TCCCTTACTTGTACCAGCTCACT

le condizioni per l'HRMA sono state:

Hold: 95° 5min; Cycling: 95° 10sec, 60° 20 sec per 50 cicli

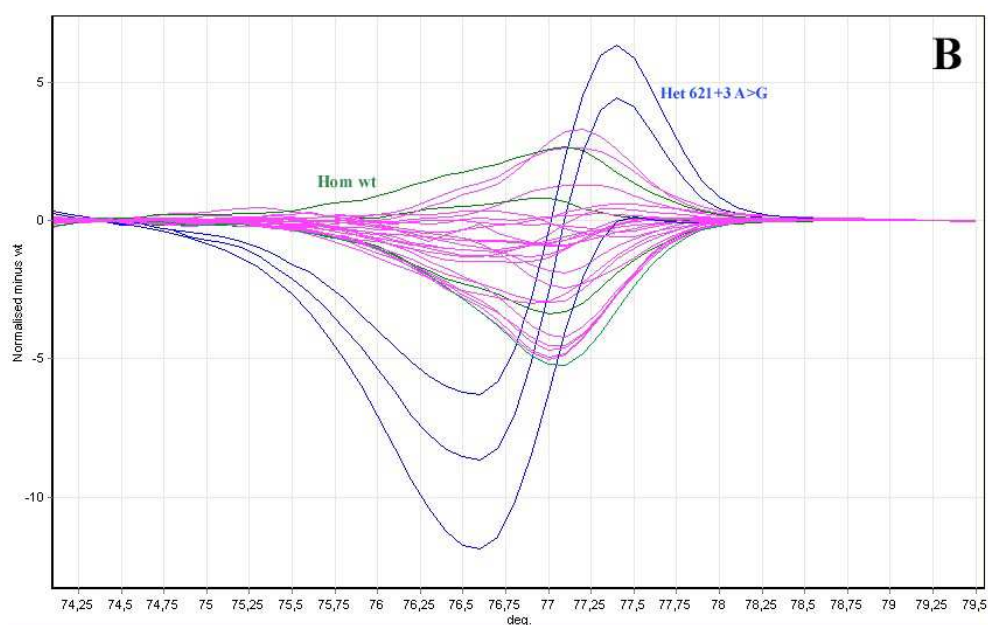
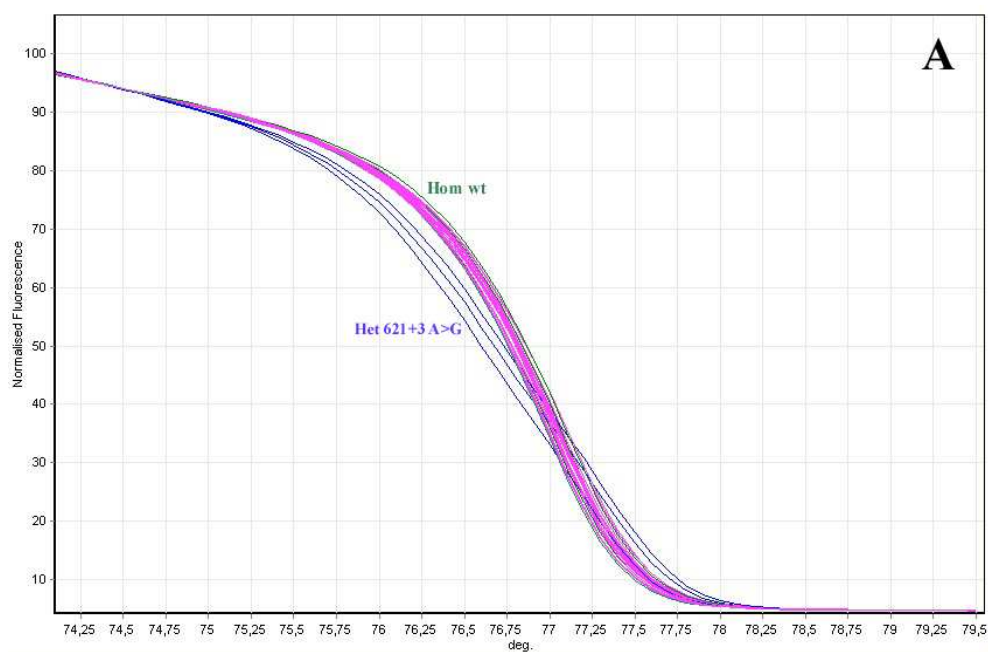
Hold 2: 72° 5 min

HRM: 71°-83°

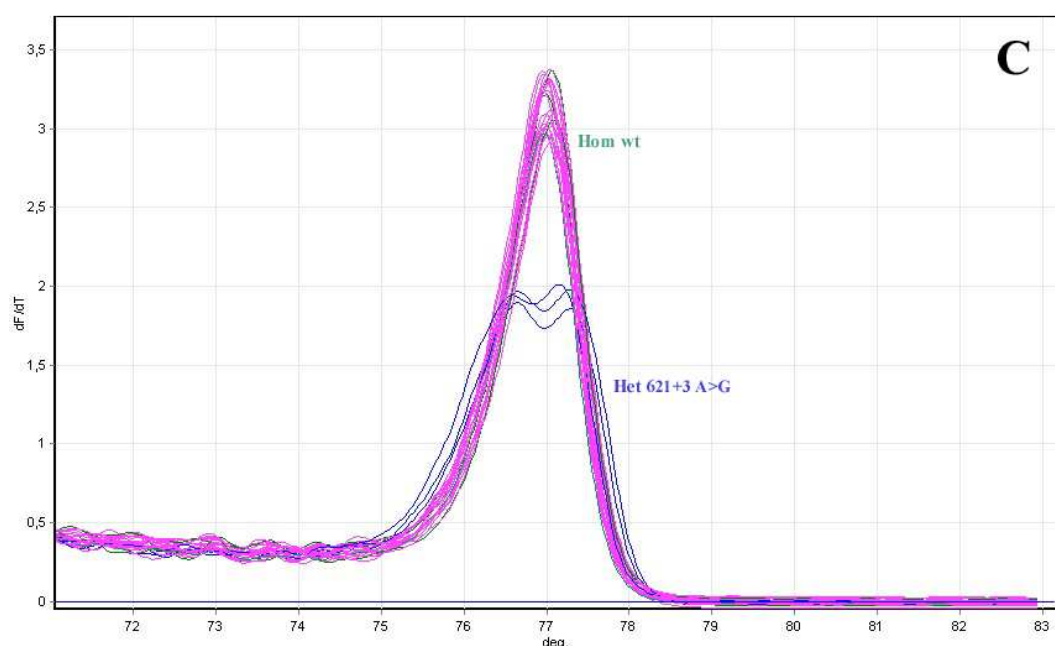
In contemporanea ai campioni, sono stati analizzati anche un controllo eterozigote per 621+3 A>G e un controllo *wild type* omozigote per l'allele A, che

erano stati precedentemente sottoposti a sequenziamento diretto dell'amplicone della regione coinvolta.

Dall'osservazione dei grafici ottenuti nell'analisi HRMA, è stato possibile isolare i soggetti che presentavano un profilo di reazione simile all'eterozigote da quelli che seguivano l'andamento grafico dell'omozigote *wild type*.







**Figura 47.** Normalized graph (A); Difference graph (B); Melt graph (C).

I campioni che hanno evidenziato un profilo simile all'eterozigote di controllo sono stati sequenziati utilizzando gli stessi *primer* usati per l'analisi HRMA.

Le frequenze attese per il genotipo CF nella popolazione sono state calcolate secondo l'equazione di Hardy-Weinberg (Hardy G.H., 1908), l'intervallo di confidenza (CI) al 95% è stato calcolato sulla base di quanto è riportato in letteratura (Newcombe R.G., 1998).

#### **4.2.2 Clonaggio della regione genomica che comprende l'esone 4 di *CFTR***

Il DNA genomico di un soggetto eterozigote per 621+3 A>G è stato amplificato con *primer* che introducono nel prodotto di amplificazione un sito di restrizione per gli enzimi XhoI e HindIII:

*primer* Mgivs4CFTRXhoI Forward: ctctCTCGAGaaaagggaaatgcttagaaactg

*primer* MGivs4CFTRHindIIIReverse: ctctAAGCTTtgagtcatcttaacaggaaca

secondo le condizioni:

Mix di reazione	Volume per 1 reazione (µl)
H <sub>2</sub> O MilliQ sterile	31,7
buffer 10X	5
MgCl <sub>2</sub>	3
dNTP (10 mM)	1
pFOR (10 µM)	2,5
pREV (10 µM)	2,5
Taq Gold	0,3
Volume finale	46 µl di mix + 4 µl di DNA

**Tabella 31.** Mix di amplificazione.

T. (°C)	Tempo	N° di cicli
95	10 min	1
95	30 sec	35
60	40 sec	
72	1 min	
72	10 min	1
4	10 min	1

**Tabella 32.** Profilo termico della reazione.

L'amplificato che ne deriva è di 853 bp:

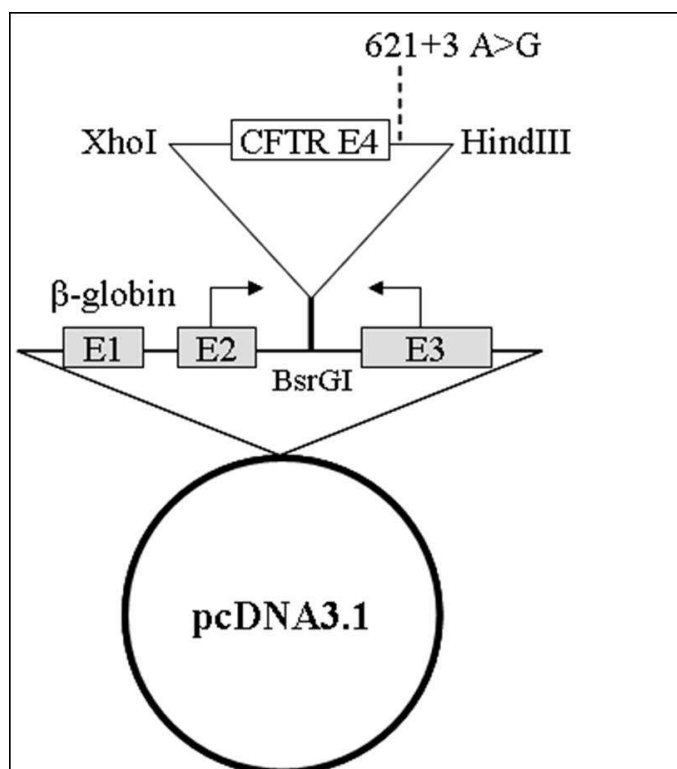
```

aaaagggaatgcttagaaactgttaagtgctatgtagatgttactaattaacaaccatttcagaaactatactttta
ttttatggccactattcactgttaactaaatacctcatatgtaaactgtctcccactgttgctataacaaatcccgaagct
tattcaaagtaccaagatattgaaaatagtgtaagagtttcacatggtagaccctctatataaactcatttaagctc
cctctaaagatgaaaagcttgttgaaattctcagggtattttatgagaaataaatgaaatttaatttctctgttttccct
ttttagGAAGTCACCAAAGCAGTACAGCCTCTCTTACTGGGAAGAATC
ATAGCTTCCTATGACCCGGATAACAAGGAGGAACGCTCTATCGCG
ATTTATCTAGGCATAGGCTTATGCCTTCTCTTTATTGTGAGGACAC
TGCTCCTACACCAGCCATTTTGGCCTTCATCACATTGGAATGCA
GATGAGAATAGCTATGTTTAGTTTGATTTATAAGAAGgtaataactccttgca
caggccccatggcacatatattctgtatcgtacatgttttaatgcatataaattaggtagtgagctggtagaaggaaggaat
aatgctgaaattaatgctcctattaataaatggcaggaataaataatgctccttaattaccttgataatgaaatgac
ttaaactgataattattgagatcttctgtaaactgcctctgttgtagtttttttctcctaatcatgttatcattttttggaatc
catggtttctgtaagatgactca

```

**Figura 48.** Sequenza dell'amplificato del gene *CFTR*; in **minuscolo** le regioni introniche, in **maiuscolo** l'esone 4 di *CFTR*. Evidenziato in **verde** il sito di variazione A>G.

Dopo controllo su gel di acrilamide all'8%, purificazione con *Microcon Centrifugal Filter Device* (Millipore) e quantificazione con spettrofotometro a 260 nm: l'amplificato, che comprende una porzione dell'introne 3, l'intero esone 4 e una porzione dell'introne 4 del gene *CFTR*, è stato digerito e clonato nel *beta-globin minigene*.



**Figura 49.** Costrutto *beta-globin minigene* + esone 4 di *CFTR*.

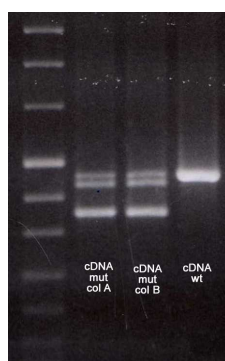
L'analisi dei costrutti al sequenziatore ha permesso di selezionare un clone con l'allele *wild type* e due con l'allele mutato.

### 4.2.3 Espressione del minigene e analisi dei trascritti

I due cloni selezionati sono stati usati per trasfettare  $2 \times 10^5$  cellula HeLa con il protocollo *Effectene Transfection Reagent* (Qiagen); dopo 48 ore è stato estratto l'RNA con *TRIzol*; retrotrascritto con *SuperScript II Reverse Transcriptase* (Invitrogen) e sequenziato.

Il cDNA è stato amplificato con *primer* localizzati sull'esone 2 e 3 del gene della  $\beta$ -globina (paragrafo 3.4.4 e 3.4.5).

I prodotti di trascrizione sono poi stati visualizzati su gel di agarosio al 3% o acrilamide al 12% (figura 50).



**Figura 50.** Amplificati ottenuti dal cDNA ricavato dalle cellule trasfettate con il minigene; si vede molto bene come dai costrutti mutati si siano ottenuti tre tipi di trascritti di lunghezza differente.

Per poter sequenziare i tre trascritti ottenuti con il costrutto mutato del minigene è stato eseguito con un clonaggio come descritto nel protocollo del *TOPO TA Cloning* (Invitrogen). Questo ha permesso di isolare i tre trascritti e di procedere al sequenziamento diretto delle tre bande, visualizzate su gel, con i *primer* complementari all'esone 2 e 3 del gene della  $\beta$ -globina.

Gli amplificati del cDNA sono stati inoltre marcati con  $\alpha^{32}\text{P}$ -dATP usando il protocollo “*last cycle hot*” (Sacconi *et al.*, 2008), separati su gel, visualizzati e quantificati usando *Storm Phosphorimager* (Molecular Dynamics) (figura 53-54).

#### 4.2.4 Risultati di epidemiologia

Analizzando 250 soggetti sani (500 alleli) nella popolazione, solo 2 sono risultati eterozigoti per 621+3 A>G con una frequenza allelica pari a 0,4% [95% Intervallo di Confidenza (CI) 0,11-1,45] per l'allele G.

La Fibrosi Cistica (CF; MIM no. 219700) è una delle più frequenti malattie autosomiche recessive presenti nella popolazione caucasica con una frequenza di 1:2500 individui (Davies J. C. *et al.*, 2007). Considerando che la

frequenza degli eterozigoti nella popolazione europea è di 1:25 individui, sulla base dei risultati ottenuti ci si aspetta che la frequenza degli eterozigoti composti per 621+3 A>G e un'altra mutazione in *CFTR* sia di 1:7812 (30% dei pazienti CF) mentre quella degli omozigoti sia di 1:62500 (4% dei pazienti CF); anche se si considera il più basso valore di 95%CI le frequenze che si possono osservare sono di 1:24010 per l'eterozigote (10% dei pazienti CF) e di 1:800000 per l'omozigote (0,3% dei pazienti CF).

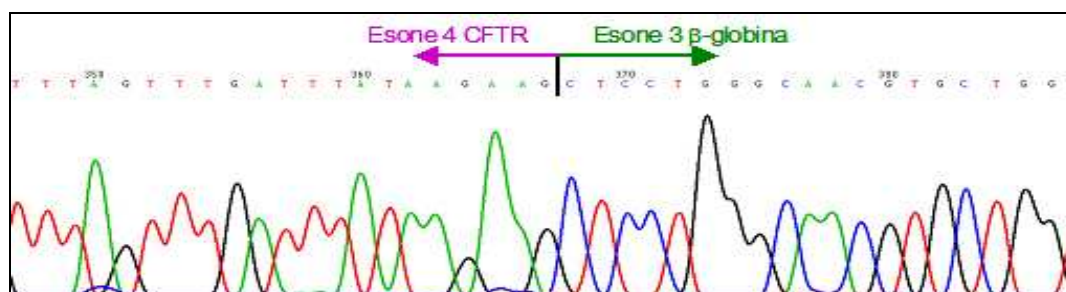
Il database italiano delle mutazioni di *CFTR* (<http://spazioinwind.libero.it/laboratoriCF/frequenze.htm>) riporta un singolo paziente eterozigote composto per 621+3 A>G, su un totale di 2253 individui affetti (4506 alleli).

#### 4.2.5 Risultati dell'analisi molecolare

L'analisi dei trascritti espressi dal minigene in cellule HeLa, ha evidenziato che il costrutto *wild type* subisce uno *splicing* corretto producendo un singolo trascritto contenente l'intero esone 4 (in totale 454 bp):

**Amplificato derivante dal trascritto wild type (454 bp):**

TCTGTCCACTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGG  
 TGCCTTTAGTGATGGCCTGGCTCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCT  
 GCCTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGGGAAGTCACCAAAGCAGTACAGCCTCT  
 CTTACTGGGAAGAATCATAGCTTCTATGACCCGGATAACAAGGAGGAACGCTCTATCGCGATTTA  
 TCTAGGCATAGGCTTATGCCTTCTCTTTATTGTGAGGACACTGCTCCTACACCAGCCATTTTTGG  
 CCTTCATCACATTGGAATGCAGATGAGAATAGCTATGTTTAGTTTGATTTATAAGAAGCTCCTGGG  
 CAACGTGCTGGTCTGTGTGCTGGCCCATCACTTTGGCAAGAATTCACCCACCAGTG

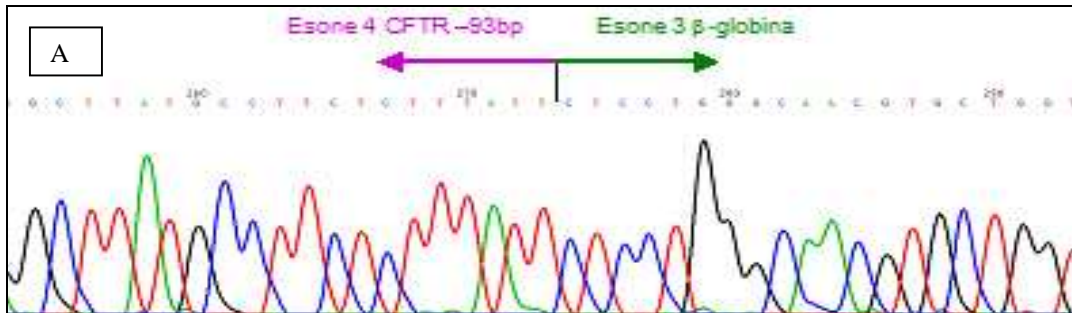


**Figura 51.** Sequenza in basi e parte del grafico ottenuti mediante sequenziamento dell'amplificato ottenuto dal cDNA ricavato dalle cellule trasfettate con il minigene *wild type*; in **azzurro** gli esoni del gene della  $\beta$ -globina, in **nero** la sequenza dell'esone 4 di *CFTR*, in **rosa** le sequenze dei *primer* usati per l'amplificazione.

Il costrutto mutato presenta tre differenti prodotti: uno corrispondente al prodotto normale, uno risultante dall'attivazione di un sito criptico di *splicing* che porta alla perdita di 93 bp dell'esone 4 di *CFTR* e uno che presenta il completo *skipping* dell'esone 4:

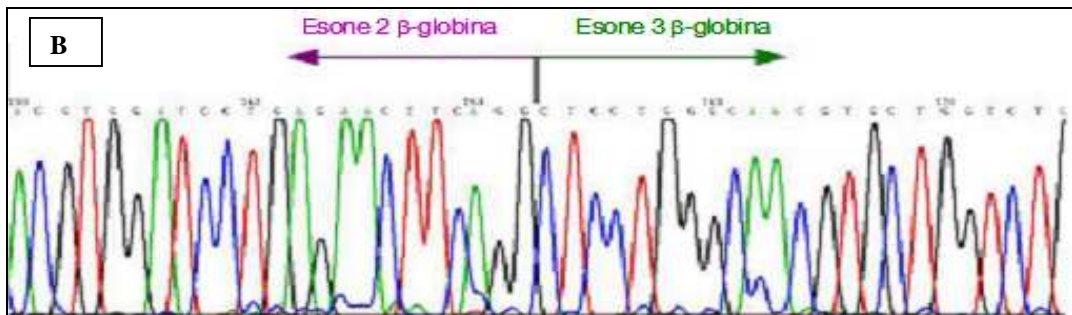
**Amplificato derivante dal trascritto mutato [del 93 bp alla fine dell'esone 4 di *CFTR*] (361 bp):**

**TCTGTCCA**CTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGG  
 TGCCTTTAGTGATGGCCTGGCTCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCT  
 GCACTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGGGAAGTCACCAAAGCAGTACAGCCTCT  
 CTTACTGGGAAGAATCATAGCTTCCTATGACCCGGATAACAAGGAGGAACGCTCTATCGCGATTTA  
 TCTAGGCATAGGCTTATGCCTTCTCTTTATTCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCA  
 TCACTTTGGCA**AAGAATTCACCCACCAGTG**

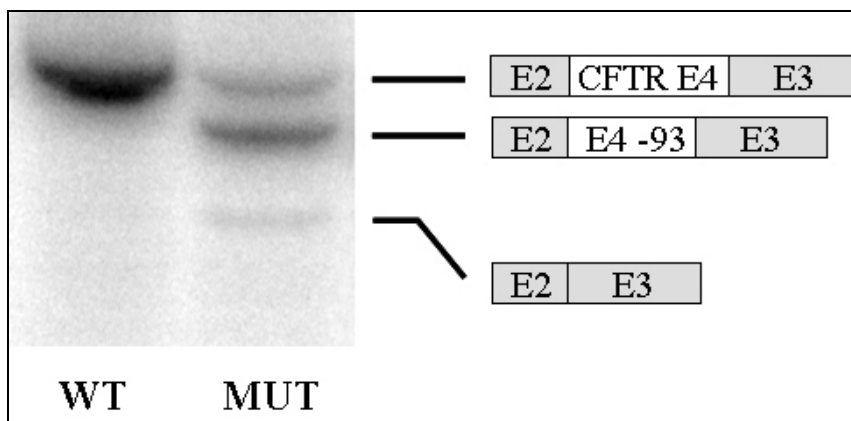


**Amplificato derivante dal trascritto mutato [skipping dell'esone 4 di *CFTR*] (238 bp):**

**TCTGTCCA**CTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGG  
 TGCCTTTAGTGATGGCCTGGCTCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCT  
 GCACTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGG•CTCCTGGGCAACGTGCTGGTCTGTG  
 TGCTGGCCCATCACTTTGGCA**AAGAATTCACCCACCAGTG**

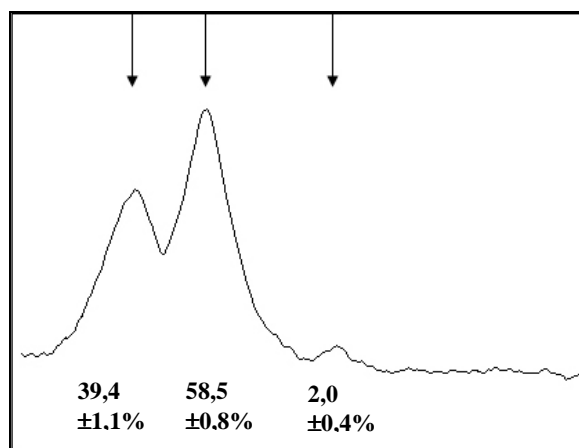


**Figura 52.** Parte degli elettroferogrammi e sequenze in basi ottenuti dal sequenziamento del trascritto derivante dal plasmide in cui era presente la variazione 612+3 A>G nel gene *CFTR*. **A:** trascritto che presenta la delezione di 93 bp alla fine dell'esone 4 di *CFTR*; **B:** punto di unione dei due esoni del gene della  $\beta$ -globina in cui è avvenuto lo *skipping* dell'esone 4.



**Figura 53.** Prodotti di espressione del  $\beta$ globin-minigene + esone 4 CFTR.

L'analisi densitometrica ottenuta sull'immagine acquisita da gel, eseguita con il software *Image J*, ha evidenziato un  $39,4 \pm 1,1\%$  di trascritto *wild type*, un  $58,5 \pm 0,8\%$  di trascritto delecto di 93 bp e un  $2,0 \pm 0,4\%$  di trascritto completamente privo dell'esone 4 di *CFTR*.



**Figura 54.** Analisi densitometrica con *Image J*.

#### 4.2.6 Conclusioni

Poter stabilire l'esatta patogenicit  di varianti come 621+3 A>G   essenziale per poter eseguire una corretta consulenza genetica ai soggetti identificati mediante *screening* nella popolazione.

Tale variante è stata riportata come causa di gravi forme di Fibrosi Cistica in alcuni pazienti di nazionalità greca (Tzetis *et al.*, 2001). Tuttavia dai risultati ottenuti in questo studio la sua patogenicità sembra ancora discutibile.

I dati epidemiologici hanno dimostrato che la frequenza di questa variante è più elevata nella popolazione di controllo che nei pazienti. La frequenza allelica osservata nel campione di controllo preso in considerazione in questo lavoro è risultata dello 0.4% (CI 95% 0.11-1.45). Ci si aspetterebbe quindi di trovare 225 pazienti con questa mutazione all'interno del Database Nazionale Italiano (<http://spazioinwind.libero.it/laboratoriCF/frequenze.htm>) mentre è riportato solo un caso di 621+3 A>G su 2253 pazienti.

Inoltre, nonostante vi sia un'evidenza di alterazione del meccanismo di *splicing* nel costrutto mutato, una porzione considerevole (39,4±1,1%) di trascritto *wild type* viene comunque prodotta. Studi effettuati su un'altra variazione a carico dello stesso sito di *splicing*, 621+1 G>T, hanno dimostrato che questa variante attiva lo stesso sito criptico che porta alla delezione di 93 bp, ma non porta alla formazione di nessuna percentuale di trascritto *wild type* (Zielenski *et al.*, 1993).

Anche se i risultati ottenuti sull'alterazione del meccanismo di *splicing* sono in accordo con quanto riportato in letteratura (Tzetis *et al.*, 2001), si ritiene che lo studio funzionale mediante minigene sia l'approccio più corretto per valutare l'esatta conseguenza della variante (Zuccato *et al.*, 2004; Pagani *et al.*, 2003; Hefferon *et al.*, 2004). Infatti la tradizionale RT-PCR sull'mRNA del paziente non permette di distinguere i trascritti derivanti dai singoli alleli da quelli derivanti dall'eterozigote nel suo insieme, a meno che non venga associata ad altre analisi come ad esempio la digestione selettiva di uno o due alleli con specifici enzimi di restrizione (Martella *et al.*, 2006).

Sulla base dei dati raccolti ne risulta quindi che la mutazione 621+3 A>G non è da considerarsi la causa di forme gravi di Fibrosi Cistica perché è frequente nella popolazione e perché vi è comunque la sintesi di una quantità significativa di proteina CFTR normale.

La presenza di questa variante in soggetti con fenotipo grave (4 casi riportati in pazienti greci e un singolo caso in Italia) probabilmente è da attribuire



alla presenza di una seconda mutazione in *cis* con 621+3 A>G non riconosciuta con le metodiche utilizzate per l'analisi dei suddetti casi.



# **Conclusioni**



Lo sviluppo di sistemi per caratterizzare lo *splicing* e le sue alterazioni, quando l'RNA non è facilmente disponibile, rappresenta un obiettivo importante per la genomica funzionale.

La possibilità di lavorare con RNA estratto dalle cellule di individui malati richiede metodiche e attrezzature che spesso non sono alla portata di tutte le strutture diagnostiche, come ad esempio:

- la possibilità di trasferire il materiale biologico del paziente dal punto di prelievo al laboratorio d'analisi nell'immediato e con le dovute condizioni (esempio: trasporto di tessuto in azoto liquido);
- la possibilità di usufruire di congelatori a  $-80^{\circ}\text{C}$  per stoccare l'RNA una volta estratto;
- la disponibilità di attrezzature di laboratorio (come cappe aspiranti, micropipette, materiale plastico *Rnase free*, ecc.) adibite ad uso esclusivo per analisi di RNA in modo da ridurre al minimo la contaminazione da Rnase;
- la presenza dei trascritti di molti geni solo in determinati tessuti non accessibili facilmente (esempio: il prodotto di espressione di *CFTR*).

Per questi motivi l'utilizzo di sistemi di minigene per caratterizzare *in vitro* lo *splicing* sembra essere un modello efficiente per sopperire a una parte di queste difficoltà.

I risultati ottenuti in questo lavoro confermano che il  *$\beta$ -globin minigene* sintetizzato è uno strumento valido per l'osservazione degli effetti sullo *splicing* di mutazioni a livello genomico.

Tuttavia la metodica ha presentato dei limiti non indifferenti per le applicazioni diagnostiche; l'aspetto che incide maggiormente è sicuramente la tempistica necessaria per poter eseguire l'analisi di una singola mutazione. Questa infatti risulta essere molto lunga, anche senza intoppi nell'esecuzione, e spesso non compatibile con i tempi richiesti per la refertazione dei risultati. Vi sono, poi, eventi non prevedibili che complicano la metodica, tra cui:

- la formazione di variazioni nucleotidiche all'interno della sequenza introdotta nel  *$\beta$ -globin minigene* in seguito ad errori di amplificazione

delle polimerasi. Generalmente questo fattore è trascurabile nelle normali reazioni di amplificazione in quanto i pochi ampliconi che presentano errori nella sequenza sono trascurabili sul totale della resa di reazione, ma nel caso del clonaggio i singoli ampliconi vengono inseriti nei diversi costrutti e quindi risultano selezionabili, rendendo quindi necessario sequenziare ogni singolo costrutto isolato.

- spesso si sono verificati, durante la fase di trasformazione delle cellule batteriche, fenomeni di ricombinazione tra il genoma batterico e il DNA del minigene inserito. Questo, ovviamente, ha impedito di poter utilizzare poi quei vettori plasmidici come mezzo per trasfettare le cellule eucariotiche, rendendo necessario ripetere più volte la trasformazione delle cellule batteriche, l'isolamento dei plasmidi e il loro sequenziamento per verificare l'esatta sequenza.

Per ovviare a quest'ultimo problema è stato preso in considerazione un diverso sistema di clonaggio basato sulla *Gateway® Technology* secondo il protocollo *Gateway® Vector Conversion System with One Shot® ccdB Survival™ 2 T1® Component Cells* (Invitrogen). Tale metodica prevede l'inserimento di una cassetta di ricombinazione all'interno del vettore plasmidico in modo da trasformarlo in un *Gateway® destination vector*.

Nel caso del *β-globin minigene* non è stato possibile, nonostante diversi tentativi, sostituire la regione "adattatore" con la cassetta di ricombinazione. Probabilmente, non essendo la regione un semplice sito di clonaggio tipico di un vettore plasmidico commerciale e trattandosi di una regione inserita all'interno dell'introne 2 della *β-globina*, intervengono fattori di disturbo che non hanno permesso la corretta ligazione della cassetta.

E' stato inoltre presa in considerazione l'opportunità di creare un minigene di dimensioni inferiori trasferendo il gene della *β-globina* con la regione adattatore al suo interno, dal vettore *pcDNA 3.1/Hygro (+)* al vettore *pEGFP-Δ* (privato della sequenza del gene EGFP). Si presume che l'uso di un plasmide più piccolo possa ridurre i rischi di ricombinazione durante la fase di trasformazione delle cellule batteriche.

Senza dubbio la metodica dell'analisi dei trascritti d'espressione mediante minigene plasmidico è valida e i risultati della trasfezione sono attendibili ma bisogna considerarne i limiti ai fini dell'utilizzo a cui è destinata.

L'uso di questa procedura per una standardizzazione del metodo a scopo diagnostico non è la scelta consigliabile in quanto i tempi per arrivare al risultato non sono prevedibili perché soggetti a troppe variabili come ad esempio il comportamento della cellula batterica fine a se stessa.

L'analisi di espressione mediante minigene plasmidico resta comunque un approccio valido in protocolli di ricerca ed è in grado di fornire importanti informazioni sul meccanismo di *splicing* dell'RNA; la caratterizzazione *in vitro* dello *splicing* di geni coinvolti in fenomeni patologici, può evidenziare l'esistenza di meccanismi di alterazione di questo diversi da quelli finora descritti.

Tali studi risulterebbero fondamentali, non solo per la definizione della patogenesi delle malattie associate, ma anche per fornire delle indicazioni utili per possibili approcci terapeutici alternativi.





# **Bibliografia**



- Aravind L, Neuwald AF, Ponting CP. *Sec14p-like domains in NF1 and Dbl-like proteins indicate lipid regulation of Ras and Rho signaling*. *Curr. Biol* (1999), 9 (6): R 195-7.
- Ars E, Kruyer H, Morell M, Pros E, Serra E, Ravella A, Estivill X, Lázaro C. *Recurrent mutations in the NF1 gene are common among neurofibromatosis type 1 patients*. *J Med Genet* 2003; 40: e82.
- Ars E, Serra E, Garcia J, Kruyer H, Gaona A, Lazaro C, Estivill X. *Mutations affecting mRNA splicing are the most common molecular defects in patients with neurofibromatosis type 1*. *Human Molecular Genetics*, (2000) 9(2): 237-247.
- Baralle D, Baralle M. *Splicing in action-assessing disease causing sequence variations*. *JMed Genet* (2005) 42:737-748.
- Beroud C, Carrie A, Beldjorf C, Deburgrave N, Llense S, Carelle N, Peccate C, Cuisset JM, Pandit F, Carre-Pigeon F, Mayer M, Bellance R, Recan D, Chelly J, Kaplan JC, Leturcq F. *Dystrophinopathy caused by mid-intronic substitutions activating cryptic exons in the DMD gene*. *Neuromuscul Disord* (2004) 14(1):10-8.
- Bollag G, McCormick F, Clark R. *Characterization of full-length neurofibromin: tubulin inhibits Ras GAP activity*. *EMBO J* 1993; 12: 1923-1927.
- Bottillo I, De Luca A, Colapietro P, Schirinzi A, Guida V, Pizzetti A, Larizza L, Dalla Piccola B. *Mutazioni di splicing nell'esone 7 del gene NF1 associate all'eliminazione o all'introduzione di siti ESE: evidenze per la presenza di un'isoforma alternativa NF1-ΔE7 della neurofibromina*. (2005) Poster, SIGU 2005.

- Caceres JF and Kornblihtt AR. *Alternative splicing: multiple control mechanisms and involvement in human disease*. Trends Genet. (2002) 18, 186-193.
- Caputi M, Casari G, Guenzi S, Tagliabue R, Sidoli A, Melo CA, Baralle FE. *A novel bipartite splicing enhancer modulates the differential processing of human fibronectin EDA exon*. Nucleic Acids Research (1994) 22:1018-1022.
- Cartegni L, Chew SL, and Krainer AR. *Listening to silence and understanding nonsense: exonic mutations that affect splicing*. (2002) Nat Rev Genet 3, 285-298.
- Cawthon RM, O'Connell P, Buchberg AM, et al. *Identification and characterization of transcripts from the neurofibromatosis 1 region: the sequence and genomic structure of EVI2 and mapping of other transcripts*. Genomics (1990); 7: 555-65.
- Cawthon RM, Weiss R, Xu G, Viskochil D, Culver M, Stevens J, Robertson M, Dunn D, Gesteland R, O'Connell P, White R. *A major segment of the neurofibromatosis type 1 gene: cDNA sequence, genomic structure, and point mutations*. Cell 1990; 62(1): 193-201.
- Clementi M, Barbujani G, Turolla L, Tenconi R. *Neurofibromatosis-1: a maximum likelihood estimation of mutation rate*. Hum. Genet. 1990; 84: 116-118.
- Cummings LM., Trent JM, Marchuck DA. *Identification and mapping of type 1 neurofibromatosis (NF1) homologous loci*. Cytogenet Cell Genet 1996; 73: 334-340.
- Dasgupta B, Dugan LL, Gutmann DH. *The neurofibromatosis 1 gene product neurofibromin regulates pituitary adenylate cyclase-activating polypeptide-mediated signaling in astrocytes*. The J of Neuroscience (2003); 23: 8949-8954.

- Daston MM, Scrabble H, Nordlund M, Sturbaum AK, Nissen LM, Ratner N. *The protein product of neurofibromatosis type 1 gene is expressed at the highest abundance in neurons, Schwann cells and oligodendrocytes.* Neuron (1992) 8: 415-428.
- Davies JC, Alton EW and Bush A. *Cystic fibrosis.* BMJ (2003) 335, 1255-1259.
- De Luca A, Buccino A, Gianni D, Mangino M, Giustizi S, Richetta A, Divona L, Calvieri S, Mingarelli R, Dalla Piccola B. *NF1 gene analysis based on DHPLC.* Hum Mutat. (2003) 21(2): 171-2.
- Dugan LL, Kim JS, Zhang Y, Bart RD, Sun Y, Holtzman DM, Gutmann DH. *Differential effects of cAMP in neurons and astrocytes. Role of B-raf.* J Biol Chem 1999; 274: 25842-25848.
- Fashold R, Hoffmeyer S, Mischung C, Gille C, Ehlers C, Küçükceylan N, Abdel-Nour M, Gewies A, Peters H, Kauffman D, Buske A, Tinschert S, Nürnberg P. *Minor lesion mutational spectrum of the entire NF1 gene does not explain its high mutability but points to a functional domain upstream of the GAP-related domain.* Am J Hum Genet. (2000); 66: 790-818.
- Ferner RE, Gutmann DH. *International consensus statement on malignant peripheral nerve sheath tumors in neurofibromatosis 1.* Cancer Res (2002) 62:1573-1577.
- Guo HF, The I, Hannan F, Bernardis A, Zhong Y. *Requirement of Drosophila NF1 for activation of adenylyl cyclase by PACAP-like neuropeptides.* Science 1997; 276: 795-798.
- Gutmann DH, Geist RT, Douglas E, Wright E, Snider WD. *Expression of the neurofibromatosis type 1 (NF1) isoforms in developing and adult rat tissues.* Cell growth and Differentiation (1995); 6:315-323.

- Han S, Cooper D, Upadhyaya M. *Evaluation of denaturing high performance liquid chromatography (DHPLC) for the mutational analysis of the neurofibromatosis type 1 (NF1) gene.* Hum. Genet. (2001) 109:487-497.
- Hardy GH. *Mendelian Proportions in a Mixed Population.* Science (1908) 28, 49–50.
- Hattori S, Maekawa M, Nakamura S. *Identification of neurofibromatosis type 1 gene product as an insoluble GTP-ase-activating protein toward ras p21.* Oncogene (1992); 7:481-5.
- Hefferon TW, Groman JD, Yurk CE and Cutting GR. *A variable dinucleotide repeat in the CFTR gene contributes to phenotype diversity by forming RNA secondary structures that alter splicing.* Proc Natl Acad Sci USA (2004) 101, 3504–3509.
- Huber CG, Oefner PJ, Bonn GK. *High resolution liquid chromatography of oligonucleotides on nonporous alkylated styrene-divinylbenzene copolymers.* Anal. Biochem. (1993); 212: 351-358.
- Izawa I, Tamaki N, Saya H. *Phosphorylation of neurofibromatosis type 1 gene product (neurofibromin) by cAMP-dependent protein kinase.* FEBS Lett (1996); 382: 53-59.
- Kayes LM, Burke W, Riccardi VM, Bennett R, Ehrlich P, Rubenstein A, Stephens K. *Deletions spanning the neurofibromatosis 1 gene: identification and phenotype of five patients.* Am J Hum Genet. 1994; 54: 424-436.
- Kim HA, Ratner N, Roberts TM, Stiles CD. *Schwann cell proliferative responses to cAMP and NF1 are mediated by Cyclin D1.* J Neurosci (2001); 21: 1110-16.

- Knudson AG Jr. *Mutation and cancer: statistical study of retinoblastoma*. Proc Natl Acad Sci USA (1971); 68: 820-823.
- Lakkis MM, Tennekoon GI. *Neurofibromatosis type 1. I. General overview*. Journal of Neuroscience Research 2000; 62: 755-763.
- Ledbetter DH, Rich DC, O'Connell P, Leppert M, Carey JC. *Precise localization of NF1 to 17q11.2 by balanced translocation*. Am J Hum Genet. 1989 Jan; 44(1): 20-24.
- Li Y, O'Connell P, Breidenbach HH, Cawthon R, Stevens J, Xu G, Neil S, Robertson M, White R, Viskochil D. *Genomic organization of the neurofibromatosis 1 gene (NF1)*. Genomics 1995; 25: 9-18.
- Lin Z, Thomas NJ, Wang Y, Guo X, Seifart C, Shakoob H, Floros J. *Deletions within a CA-Repeat-rich region of intron 4 of the human SP-B gene affect mRNA splicing*. (2005) Biochem J. 389:403-12.
- Martella M, Salviati L, Casarin A, Trevisson E, Opocher G, Polli R et al. *Molecular analysis of two uncharacterized sequence variants of the VHL gene*. J Hum Genet (2006) 51, 964-968.
- Matlin AJ, Clark F, Smith CW. *Understanding alternative splicing: towards a cellular code*. (2005) Nat Rev Mol Cell Biol 6:386-398.
- Mattocks C, Baralle D, Tarpey P, French-Constant C, Bobrow M, Whittaker J. *Automated comparative sequence analysis identifies mutations in 89% of NF1 patients and confirms a mutation cluster in exons 11-17 distinct from the GAP related domain*. J Med Genet (2004); 41: e48.
- McCormick F. *Ras GTPase activating protein: signal transmitter and signal terminator*. Cell (1989); 56: 5-8.

- Messiaen LM, Callens T, Mortier G, Beysen D, Vandenbroucke I, Van Roy N, Speleman F, De Paepe A. *Exhaustive mutation analysis of the NF1 gene allows identification of 95% of mutations and reveals a high frequency of unusual splicing defects.* Human Mutation. (2000); 15: 541-555.
- Metherell LA, Akker SA, Munroe PB, Rose SJ, Caulfield M, Savage MO, Chew SL, Clark AJ. *Pseudoexon activation as a novel mechanism for disease resulting in atypical growth-hormone insensitivity.* (2001) Am. J. Hum. Genet. 69, 641-646.
- Newcombe RG. *Two-sided confidence intervals for the single proportion: comparison of seven methods.* Stat Med (1998) 17, 857–872.
- O'Donovan MC, Oefner PJ, Roberts SC, Austin J, Hoogendoorn B, Guy C, Speight G, Upadhyaya M, Sommer SS, McGuffin P. *Blind analysis of denaturing high-performance liquid chromatography as a tool for mutation detection.* Genomics 1998; 52: 44-49.
- Pagani F, Baralle FE. *Genomic variants in exons and introns: identifying the splicing spoilers.* Nat Rev Genet. (2004) May;5(5):389-96. Review.
- Pagani F, Buratti E, Stuani C, and Baralle FE. *Missense, nonsense, and neutral mutations define juxtaposed regulatory elements of splicing in cystic fibrosis transmembrane regulator exon 9.* (2003) J Biol Chem 278: 26580-26588.
- Pagani F, Buratti E, Stuani C, Bendix R, Dork T, and Baralle FE. *A new type of mutation causes a splicing defect in ATM.* Nat Genet (2002) 30: 426-429.
- Pagani F, Buratti E, Stuani C, Romano M, Zuccato E, Niksic M, Giglio L, Faraguna D, Baralle FE. *Splicing factors induce cystic fibrosis transmembrane regulator exon 9 skipping through a nonevolutionary conserved intronic element.* J Biol Chem. (2000) 275:21041-7.



- Pagani F, Stuani C, Zuccato E, Kornblihtt AR and Baralle FE. *Promoter architecture modulates CFTR exon 9 skipping*. J Biol Chem (2003) 278, 1511–1517.
- Pros E, Gòmez C, Martìn T, Fàbregas P, Serra E, Làzaro C. *Nature and mRNA effect of 282 different NF1 point mutation: focus on splicing alteration*. Human Mutation (2008); 26: E173-E193.
- Raponi M, Upadhyaya M and Baralle D. *Functional splicing assay show a pathogenetic intronic mutation in Nuerofibromatosis type 1 (NF1) due to intronic sequence exonization*. Hum. Mut. (2006) 27:294-295.
- Reed GH, Wittwer CT. *Sensitivity and Specificity of Single-Nucleotide Polymorphism Scanning by High-Resolution Melting Analysis*. Clinical Chemistry 50 (2004):10; 1748–1754.
- Ruggieri M e Tenconi R. *Le Neurofibromatosi*. Associazione Linfa, Lottiano Insieme per la Neurofibromatosi- ONLUS- (2001).
- Sacconi S, Salviati L, Nishigaki Y, Walker WF, Hernandez-Rosa E, Trevisson E et al. *A functionally dominant mitochondrial DNA mutation*. Hum Mol Genet (2008) 17, 1814–1820.
- Schmidt MA, Michels VV, Dewald GW, Opitz JM, Reynolds JF. *Cases of neurofibromatosis with rearrangements of chromosome 17 involving band 17q11.2*. Am J Med Genet 1987; 28: 771-777.
- Seizinger BR, Rouleau GA, Ozelius LJ, Lane AH, Faryniarz AG, Chao MV, Huson S, Korf BR, Parry DM et al. *Genetic linkage of von Recklinghausen neurofibromatosis to the nerve growth factor receptor gene*. Cell 1987; 49(5): 589-594.

- Shen MH, Harper PS, Upadhyaya M. *Molecular genetics of neurofibromatosis type 1 (NF1)*. J Med Genet (1996); 33: 2-17.
- Sun H and Chasin LA. *Multiple splicing defect in an intronic false exon*. Mol. Cell. Biol. (2000) 20: 6414–6425.
- Treisman R, Orkin SH, Maniatis T. *Specific transcription and RNA splicing defects in five cloned beta-thalassemia genes*. Nature (1983) 302, 591-596.
- Tzetis M, Efthymiadou A, Doudounakis S and Kanavakis E. *Qualitative and quantitative analysis of mRNA associated with four putative splicing mutations (621+3A–4G, 2751+2T–4A, 296+1G–4C, 1717-9T–4C-D565G) and one nonsense mutation (E822X) in the CFTR gene*. Hum Genet (2001) 109, 592–601.
- Upadhyaya M, Osborn MJ, Maynard J, Kim MR, Tamanoi F, Cooper DN. *Mutational and functional analysis of the neurofibromatosis type 1 (NF1) gene*. Hum Genet (1997); 99: 88-92.
- Upadhyaya M, Shaw DJ and Harper PS. *Molecular basis of neurofibromatosis type 1 (NF1): mutation analysis and polymorphisms in the NF1 gene*. Human Mutation (1994) 4:83-101.
- Vandenbroucke I, Van Oostveldt P, Coene E, De Paepe A, Messiaen L. *Neurofibromin is actively transported to the nucleus*. FEBS Lett. (2004); 560 (1-3): 98-102.
- Viskochil D, White R, Cawthon R. *The Neurofibromatosis type 1 gene*. Annu. Rev. Neurosci. (1993) 16:183-205.

- White H, Potts G. *Mutation scanning by high resolution melt analysis. Evaluation of RotorGene™ 6000 (Corbett Life Science), HRI™ and 384 well LightScanner™ (Idaho Technology). NGRLW\_HRM\_1.0 (2006).*
- White R, Nakamura Y, O'Connell P, Leppert M, Lalouet JM, Barker D, Golgar D, Skolnick M, Carey J, Wallis CE. *Tightly linked markers for the neurofibromatosis type 1 gene. Genomics (1987) 1:364-7.*
- Wiest V, Eisenbarth I, Schmegner C, Krone W, Assum G. *Somatic NF1 mutation spectra in a family with Neurofibromatosis type 1: toward a theory of genetic modifiers. Hum Mutat. (2003); 22: 423-427.*
- Zielenski J, Bozon D, Markiewicz D, Aubin G, Simard F, Rommens JM et al. *Analysis of CFTR transcripts in nasal epithelial cells and lymphoblasts of a cystic fibrosis patient with 621+1G-4T and 711+1G-4T mutations. Hum Mol Genet (1993) 2, 683-687.*
- Zuccato E, Buratti E, Stuani C, Baralle FE and Pagani F. *An intronic polypyrimidine-rich element downstream of the donor site modulates cystic fibrosis transmembrane conductance regulator exon 9 alternative splicing. J Biol Chem (2004) 279, 16980-16988.*



## Appendice

Esone Introne	Numero identificativo	Alterazione nucleotidica	Alterazione aminoacidica	Previsione prodotto genico	Tipo di mutazione	Ricorrenza di mutazione	F/S	Referenze
1	2484	insT 1-2 delCATGG 1-1to4	/	/	ins/del	1	F	this report
1	2847	51 C>G 54insT	F17L aa18 PTC 37	truncation	inserzione	1	F	this report
IVS 1	3629	61-12 delT	/	/	delezione	1	F	this report
2	3085	185delT	aa 62 PTC 62	truncation	delezione	1	F	this report
2	2901	82C>T	Q28X	truncation	transizione	1	S	this report
2	2987	79C>T	Q27X	truncation	transizione	1	U	this report
IVS 2	3112	204+1 G>T	/	splice site skip exon2	trasversione	1	S	Fahsold et al, 2000
IVS 2	3446	205-23 G>A	/	splice site	transizione	1	U	this report
3	2954	207delAA	aa 69-70 PTC 75	truncation	delezione	1	U	this report
3	3604	240 T>G	Y80X	truncation	trasversione	1	U	this report
3	3084	248 A>C	Q83P	missense	trasversione	1	U	this report
3	3800	278 G>A	C 93 Y	missense	transizione	1	S	Messiaen et al, 2000
IVS 3	3569	288+1delG	aa 96 PTC 102	truncation	delezione	1	S	this report
IVS 3	3366	288+1 G>A	/	splice site	transizione	1	U	this report

Esone Introne	Numero identificativo	Alterazione nucleotidica	Alterazione aminoacidica	Previsione prodotto genico	Tipo di mutazione	Ricorrenza di mutazione	F/S	Referenze
4a	3292	313delGATGAAACGATGCTGG	aa 105-110 PTC 164	truncation	delezione	1	U	this report
4a	3419	332insCA	aa111 PTC165	truncation	inserzione	1	S	this report
4b	3237	496delGT	aa 166 PTC 172	truncation	delezione	1	S	Toliat et al,1999
4b	2551	499-502 del TGTT	aa 167-168 PTC 176	truncation	delezione	4	F	Ars et al, 2003
4b	3052	499-502 del TGTT	aa 167-168 PTC 176	truncation	delezione	4	U	Ars et al, 2003
4b	3435	499-502 del TGTT	aa 167-168 PTC 176	truncation	delezione	4	S	Ars et al, 2003
4b	3466	499-502 del TGTT	aa 167-168 PTC 176	truncation	delezione	4	U	Ars et al, 2003
4b	3443	574C>T	R192X	truncation	transizione	1	U	Toliat et al,1999
IVS 4b	3718	587-134 G>A (omozigosi)	/	splice site	transizione	1	U	this report
4c	3051	607del GCCCTAA	aa 203-205 PTC 210	truncation	delezione	1	U	this report
4c	3054	620 ins G	aa 207 PTC 215	truncation	inserzione	1	S	this report
4c	3183	625C>T (*)	Q209X	truncation	transizione	1	S	this report
4c	3808	654insAAAA	aa 218 PTC 231	truncation	inserzione	1	U	this report
IVS 4c	3715	655-8 delATTTAT	/	splice site	delezione	1	U	Serra et al, 2001
5	3844	662 G>A	W221X	truncation	transizione	1	S	this report

Esone Introne	Numero identificativo	Alterazione nucleotidica	Alterazione aminoacidica	Previsione prodotto genico	Tipo di mutazione	Ricorrenza di mutazione	F/S	Referenze
5	3056	663 G>A	W221X	truncation	transizione	1	U	this report
IVS 5	3236	730-5 T>G	/	splice site	trasversione	1	S	this report
6	3795	750 del T	aa 250 PTC 280	truncation	delezione	1	U	this report
6	3475	801 del G	aa 267 PTC 280	truncation	delezione	1	S	this report
6	2965	878delA	aa 293 PTC 294	truncation	delezione	1	F	Ars et al, 2003
IVS 6	3608	888+2 T>A	/	splice site	trasversione	1	U	this report
7	3205	1019delCT	aa 340 PTC 351	truncation	delezione	1	U	Fahsold et al, 2000
7	3391	983delGT	aa 328 PTC 328	truncation	delezione	1	U	Abernathy et al, 1997
IVS 7	3165	1062+3 A>G	/	splice site	transizione	1	U	this report
IVS 7	3368	1062+113 A>G	/	/	transizione	1	U	this report
8	3568	1094 C>G	S365X	truncation	trasversione	1	U	this report
8	3504	1183 ins T	aa 395 PTC 395	truncation	inserzione	1	S	this report
IVS 8	2620	1185+1 G>A	/	splice site skip exon 8 (IF -41 aa)	transizione	1	F	Hoffmeyer et al, 1995
IVS 8	3548	1185+2 T>G	/	splice site	trasversione	1	U	this report
IVS 9	3418	1261-13 A>G	/	splice site	transizione	1	S	this report

Esone Introne	Numero identificativo	Alterazione nucleotidica	Alterazione aminoacidica	Previsione prodotto genico	Tipo di mutazione	Ricorrenza di mutazione	F/S	Referenze
<b>IVS 9a</b>	3158	1261-19 G>A	/	splice site	transizione	1	U	this report
<b>10a</b>	2988	1278 G>A	W426X	truncation	transizione	2	S	this report
<b>10a</b>	3261	1278 G>A	W426X	truncation	transizione	2	S	this report
<b>10a</b>	3187	1307delG	aa 436 PTC 472	truncation	delezione	1	U	this report
<b>10a</b>	3206	1318 C>T	R440X	truncation	transizione	2	S	Heim et al., 1995
<b>10a</b>	4324	1318 C>T	R440X	truncation	transizione	2	U	Heim et al., 1995
<b>10b</b>	3216	1466 A>G	Y489C	missense	transizione	1	U	Han et al, 2001
<b>IVS 10a</b>	4821	1392+1 G>A	/	splice site	transizione	1	S	De Luca et al, 2004
<b>IVS 10a</b>	3003	1392+5 G>T 1392+6 A>T	/	/	trasversione	1	U	this report
<b>IVS10b</b>	3458	1527+1delGTAA	/	splice site	delezione	1	F	this report
<b>10c</b>	3306	1557 delA	aa 519 PTC 525	truncation	delezione	1	S	this report
<b>10c</b>	3415	1541-1542delAG	aa 514 PTC 556	truncation	delezione	3	U	Robinson et al, 1996
<b>10c</b>	3159	1541-1542delAG	aa 514 PTC 556	truncation	delezione	3	U	Robinson et al, 1996
<b>10c</b>	3460	1541-1542delAG	aa 514 PTC 556	truncation	delezione	3	U	Robinson et al, 1996
<b>11</b>	3195	1721G>A	S574N	skip of exon11 out of frame; PTC 560 (missense) splice site	transizione	1	U	Fahsold et al, 2000



Esone Introne	Numero identificativo	Alterazione nucleotidica	Alterazione aminoacidica	Previsione prodotto genico	Tipo di mutazione	Ricorrenza di mutazione	F/S	Referenze
<b>IVS 11</b>	2991	1722-3 C>T	/	splice site	transizione	1	U	this report
<b>12a</b>	3717	1724 C>G	S575X	truncation	trasversione	1	U	this report
<b>12a</b>	3039	1725-1733delACAAATGCT	SQMLF575-579S	delezione in frame di 4 aa	delezione	1	S	this report
<b>12b</b>	2890	1919 ins A	aa 640 PTC 648	truncation	inserzione	1	F	this report
<b>12b</b>	3648	1986delA	aa 662 PTC 687	truncation	delezione	1	U	this report
<b>12b</b>	3260	1945 G>T	E649X	truncation	trasversione	1	U	this report
<b>12b</b>	3487	1996delT	aa 666 PTC 687	truncation	delezione	1	U	this report
<b>IVS 12b</b>	3335	2002-14 C>T	/	splice site	transizione	1	F	this report
<b>13</b>	2506	2177delT	aa726 PTC 747	truncation	delezione	1	F	this report
<b>13</b>	2745	2041 C>T	R681X	truncation	transizione	1	S	Ars et al, 2003
<b>13</b>	3041	2034 insC	aa678 PTC699	truncation	inserzione	2	S	Heim et al., 1995
<b>13</b>	3417	2034 insC	aa678 PTC699	truncation	inserzione	2	U	Heim et al., 1995
<b>IVS 13</b>	2882	2252-1 G>A	/	splice site	transizione	1	S	Ruggieri
<b>IVS13</b>	3296	2252-2 A>G	/	splice site	transizione	1	U	this report
<b>14</b>	2502	2273 ins T	aa 758 PTC 767	truncation	inserzione	1	F	this report

Esone Introne	Numero identificativo	Alterazione nucleotidica	Alterazione aminoacidica	Previsione prodotto genico	Tipo di mutazione	Ricorrenza di mutazione	F/S	Referenze
IVS14	3506	2325+2 ins T	/	splice site	inserzione	1	U	this report
15	2947	2330 G>C	W777S	Missense,regular size	trasversione	1	S	Fahsold et al, 2000
15	3502	2335delGATACA	aa 779-780 del D-T	deletion 2 aa (779-780) in frame	delezione	1	U	this report
15	3038	2388 del A	aa 796 PTC 820	truncation	delezione	1	S	this report
15	3134	2331 G>A	W777X	truncation	transizione	1	S	this report
15	3490	2342 A>C	H781P	missense	trasversione	1	U	Fahsold et al, 2000
15	3390	2353delG	aa 785 PTC 790	truncation	delezione	1	U	this report
16	3794	2446 C>T	R816X	truncation	transizione	1	U	Fahsold et al, 2000
16	3329	2488insT	aa 830 PTC 830	truncation	inserzione	1	U	this report
16	2450	2543-2557del GGGGAGTGTGCCTCC	GGVCLQ848-853E	delezione in frame di 5 aa (missense)	delezione	1	S	this report
16	2907	2543insG	aa 848 PTC 864	truncation	inserzione	1	U	this report
16	3032	2674delA	aa 892 PTC 901	truncation	delezione	1	U	Fahsold et al, 2000
16	2443	2685insTAT	M 895 I+M	inserzione in frame	inserzione	1	U	this report
16	3144	2689insGGAT	aa 897 PTC 906	truncation	inserzione	1	S	this report

Esone Introne	Numero identificativo	Alterazione nucleotidica	Alterazione aminoacidica	Previsione prodotto genico	Tipo di mutazione	Ricorrenza di mutazione	F/S	Referenze
16	2235	2764 G>A	G922S	Cryptic 5¢ ss/2761del90 Missense/splice 2788 / IF (-30)	trasversione	1	F	Ars et al, 2003
16	3527	2810 T>A	L937X	truncation	trasversione	1	U	this report
IVS 17	3028	2991-1 G>A	/	splice site skip exon 18 (IF -41 aa)	transizione	2	S	Fahsold et al, 2000
IVS 17	4040	2991-1 G>A	/	splice site skip exon 18 (IF -41 aa)	transizione	2	S	Fahsold et al, 2000
18	3010	3062 ins G	aa 1021 PTC 1027	truncation	inserzione	1	U	this report
18	2875	3112 A>G	R1038G	missense	transizione	1	S	this report
IVS 18	2844	3113+1 G>T	/	splice site skip exon 18 (IF -41 aa)	trasversione	2	S	Ars et al, 2003
IVS 18	3204	3113+1 G>T	/	splice site skip exon 18 (IF -41 aa)	trasversione	2	U	Ars et al, 2003
IVS 18	3260	3114-27 G>C	/	splice site	trasversione	1	U	this report
19a	3133	3165 ins T del AG	aa 1055-1056 Q 1055 H PTC 1061	truncation	ins/del	1	U	this report
20	3095	3456delACTC	aa 1152-1153 PTC 1156	truncation	delezione	1	S	Upadhyaya et al,1997
IVS 20	3089	3496+1 G>A	/	splice site	transizione	1	S	this report
IVS 20	3580	3496+3 G>T	/	splice site	trasversione	1	U	this report
21	2999	3525-3526 del AA	aa 1175-1176 PTC 1193	truncation	delezione	2	U	Fahsold et al, 2000
21	3090	3525-3526 del AA	aa 1175-1176 PTC 1193	truncation	delezione	2	U	Fahsold et al, 2000

<b>Esone Introne</b>	<b>Numero identificativo</b>	<b>Alterazione nucleotidica</b>	<b>Alterazione aminoacidica</b>	<b>Previsione prodotto genico</b>	<b>Tipo di mutazione</b>	<b>Ricorrenza di mutazione</b>	<b>F/S</b>	<b>Referenze</b>
21	3463	3528 del A	aa 1176 PTC 1183	truncation	delezione	1	U	Mattocks et al, 2004
21	3375	3600 del AT	aa 1200-1201 PTC 1202	truncation	delezione	1	U	this report
21	3374	3586 C>T	L1196F	missense	transizione	1	U	this report
21	3467	3649 G>C	D1217H	missense	trasversione	1	S	this report
<b>IVS 21</b>	2572	3709-1 G>A	/	splice site	transizione	1	S	this report
22	3715	3715dup3715-3787	aa 1239 PTC 1272	truncation	inserzione (duplicazione)	1	F	this report
22	2462	3721 C>T (*)	R1241X	truncation	transizione	1	S	Fahsold et al, 2000
22	2845	3728 del T	aa 1243 PTC 1265	truncation	delezione	1	S	Upadhyaya et al,1997
22	3464	3826 C>T	R1276X	truncation	transizione	2	U	Klose et al,1999
22	3465	3826 C>T	R1276X	truncation	transizione	2	U	Klose et al,1999
22	3654	3827 G>A	R1276Q	missense	transizione	2	U	Fahsold et al, 2000
22	3412	3827 G>A	R1276Q	missense	transizione	2	U	Fahsold et al, 2000
22	2279	3864T>A	C1288X	truncation	trasversione	1	F	this report
22	3238	3805-3809delGACTC	aa 1269-1270 PTC 1281	truncation	delezione	1	U	this report
<b>23-1</b>	3374	3883 A>G	T1295A	missense	transizione	1	U	this report

Esone Introne	Numero identificativo	Alterazione nucleotidica	Alterazione aminoacidica	Previsione prodotto genico	Tipo di mutazione	Ricorrenza di mutazione	F/S	Referenze
23-1	3048	3911T>A	L1304X	truncation	trasversione	1	U	this report
23-2	3445	3989 del AGAG	aa 1330-1331 PTC 1341	truncation	delezione	1	U	this report
23-2	3803	4043 A>C	H1348P	missense	trasversione	1	F	this report
23-2	3227	4084 C>T	R1362X	truncation	transizione	2	U	Fahsold et al, 2000
23-2	3693	4084 C>T	R1362X	truncation	transizione	2	S	Fahsold et al, 2000
IVS 23a	3258	4111-8 delGTT	/	/	delezione	1	F	Mattocks et al, 2004
24	3344	4265 C>G	S1422X	truncation	trasversione	1	U	this report
IVS 24	3192	4269+1 G>A	/	splice site	transizione	1	U	report ?
25	1745	4312-4314 del GAA	1438 del E	aa deletion, IF (-1)	delezione	1	F	Ars et al, 2003
25	3394	4309G>A	E1437K	missense	transizione	2	F	this report
25	4012	4309G>A	E1437K	missense	transizione	2	S	this report
26	3792	4461 T>G	H1487Q	missense	trasversione	1	F	this report
27a	2876	4537C>T	R1513X	truncation	transizione	5	S	Ars et al, 2003
27a	3040	4537C>T	R1513X	truncation	transizione	5	S	Ars et al, 2003
27a	3082	4537C>T	R1513X	truncation	transizione	5	U	Ars et al, 2003

Esone Introne	Numero identificativo	Alterazione nucleotidica	Alterazione aminoacidica	Previsione prodotto genico	Tipo di mutazione	Ricorrenza di mutazione	F/S	Referenze
27a	3632	4537C>T	R1513X	truncation	transizione	5	U	Ars et al, 2003
27a	3733	4537C>T	R1513X	truncation	transizione	5	U	Ars et al, 2003
27a	3235	4538-4540 del GAC	R-P1513-1514P	aa deletion, IF (-1)	delezione	1	F	this report
27a	3442	4614 del G	aa 1538 PTC 1552	truncation	delezione	1	U	Lee et al, 2006
27b	3251	4680 del A	aa 1560 PTC 1566	truncation	delezione	1	U	this report
27b	3459	4760 ins T	aa 1587 PTC 1600	truncation	inserzione	1	U	this report
IVS 27b	2939	4772+2 T>G	/	splice site	trasversione	1	U	this report
28	2964	4794 del T	aa 1598 PTC 1602	truncation	delezione	1	U	this report
28	3503	4984 ins T	aa 1662 PTC 1662	truncation	delezione	1	S	this report
29	3638	5207 T>A	V1736D	missense	trasversione	1	U	this report
29	3053	5220 delTGTinsA	aa 1740-1741 PTC 1759	truncation	ins/del	1	U	this report
29	3016	5242 C>T	R1748X	truncation	transizione	1	F	Ars et al, 2003
29	3663	5248 A>G	K1750E	missense	transizione	1	U	this report
29	3063	5425 C>T	R1809C	Missense	transizione	2	F	this report
29	3024	5425 C>T	R1809C	Missense	transizione	2	U	this report

Esone Introne	Numero identificativo	Alterazione nucleotidica	Alterazione aminoacidica	Previsione prodotto genico	Tipo di mutazione	Ricorrenza di mutazione	F/S	Referenze
29	3718	5482 G>T	D1829Y	Missense	trasversione	1	U	this report
30	3582	5553delTGCCTATAATCT	AAYNLL1851-1855A	deletion 5 aa (1852-1855) in frame	delezione	1	U	this report
30	2927	5624C>G	S1875X	truncation	trasversione	1	S	this report
30	3569	5694 G>A	E1898E	silente	transizione	1	U	this report
31	2879	5839 C>T (*)	R1947X	truncation	transizione	4	S	Klose et al,1999
31	2899	5839 C>T (*)	R1947X	truncation	transizione	4	U	Klose et al,1999
31	3622	5839 C>T (*)	R1947X	truncation	transizione	4	F	Klose et al,1999
31	3633	5839 C>T (*)	R1947X	truncation	transizione	4	U	Klose et al,1999
31	2891	5878 del A	aa 1960 PTC 1960	truncation	delezione	1	F	this report
IVS 31	3049	IVS 31+1 G>C	/	splice site	trasversione	1	F	this report
IVS 31	3055	5944-5 A>G	/	splice site skip ex 32 (IF -47 aa)	transizione	2	S	Fahsold et al, 2000
IVS 31	2900	5944-5 A>G	/	splice site skip ex 32 (IF -47 aa)	transizione	2	S	Fahsold et al, 2000
33	3317	6312 del G	aa 2104 PTC 2128	truncation	delezione	1	S	this report
34	2967	6410 ins T	aa 2137 PTC 2147	truncation	inserzione	1	U	this report
34	2361	6492 delG	aa 2164 PTC 2178	truncation	delezione	1	F	this report

Esone Introne	Numero identificativo	Alterazione nucleotidica	Alterazione aminoacidica	Previsione prodotto genico	Tipo di mutazione	Ricorrenza di mutazione	F/S	Referenze
34	2639	6513T>G	Y2171X	truncation	trasversione	1	F	this report
34	3177	6524insGT	aa 2175 PTC 2179	truncation	inserzione	1	F	this report
IVS 34	2893	6579+1 G>C	/	splice site	trasversione	1	S	this report
35	3002	6637 C>T	Q2213X	truncation	transizione	1	U	this report
IVS 35	3437	6641+1 G>A	/	splice site	transizione	1	U	De Luca et al, 2004
IVS 35	3625	6641+1 G>C	/	splice site	trasversione	1	F	this report
36	3525	6733 C>T	Q2245X	truncation	transizione	1	F	this report
IVS 36	2616	6756+1 G>A	Skip of ex 36(?); out of frame; PTC2220	splice site	transizione	1	U	Fahsold et al, 2000
37	3147	6789 del TTAC	aa 2263-2264 PTC 2268	truncation	delezione	2	S	Robinson et al,1995
37	3659	6789 del TTAC	aa 2263-2264 PTC 2268	truncation	delezione	2	U	Robinson et al,1995
37	3101	6792 C>A	Y2264X	truncation	trasversione	3	U	Fahsold et al, 2000
37	3226	6792 C>A	Y2264X	truncation	trasversione	3	U	Fahsold et al, 2000
37	3566	6792 C>A	Y2264X	truncation	trasversione	3	U	Fahsold et al, 2000
37	3194	6792insA	aa 2264 PTC 2264	truncation	inserzione	1	U	De Luca et al, 2004
38	3099	6882 C>G	L2294L	silente	trasversione	1	U	this report



Esone Introne	Numero identificativo	Alterazione nucleotidica	Alterazione aminoacidica	Previsione prodotto genico	Tipo di mutazione	Ricorrenza di mutazione	F/S	Referenze
38	3677	6866 del C	aa 2289 PTC 2297	truncation	delezione	1	U	this report
38	3734	6933-6934delAG	aa 2311-2312 PTC 2317	truncation	delezione	1	U	this report
38	3724	6947 T>C	L2316P	missense	transizione	1	S	Wu et al, 1999
39	3184	7090 ins T	aa 2364 PTC 2364	truncation	inserzione	1	U	Report (fogli Clem.)
39	3924	7096-7101 delAACTTT	del aa 2364-2365 N-F	delezione di 2 aa in frame	delezione	1	F	Ars, et al 2003
40	3579	7199 A>G	N2400S	missense	transizione	1	U	this report
40	3163	7218-7231 delCAAATTTGAAGTGA	aa 2406-2411 PTC 2421	truncation	delezione	1	U	this report
41	3247	7285 C>T	R2429X	truncation	transizione	2	S	Fashold et al, 2000
41	3268	7285 C>T	R2429X	truncation	transizione	2	U	Fashold et al, 2000
IVS 41	3719	7394+1 G>C	/	splice site	trasversione	1	U	this report
IVS 43	3637	7676-2 A>T	/	splice site	trasversione	1	U	Osborn et al, 1999
44	2686	7711 C>G	H2571D	missense	trasversione	1	F	this report
44	3368	7682-7683delAG	aa 2561 PTC2573	truncation	delezione	1	U	De Luca et al, 2004
45	3303	7846 C>T	R2616X	truncation	transizione	3	F	Fashold et al, 2000
45	3030	7846 C>T	R2616X	truncation	transizione	3	U	Fashold et al, 2000

Esone Introne	Numero identificativo	Alterazione nucleotidica	Alterazione aminoacidica	Previsione prodotto genico	Tipo di mutazione	Ricorrenza di mutazione	F/S	Referenze
45	3550	7846 C>T	R2616X	truncation	transizione	3	F	Fashold et al, 2000
IVS 47	3234	8097+55 T>C	/	splice site	transizione	1	U	this report
49	3024	8332 G>A	V2778I	missense	transizione	1	U	this report

**Tabella 14.** Elenco delle mutazioni riscontrate in 196 pazienti su 310 analizzati. Oltre a riportare l'alterazione nucleotidica e aminoacidica della sequenza (la numerazione è stata formulata sulla base del cDNA di riferimento: NM\_000267.1), in tabella è evidenziato il grado di familiarità della mutazione (F: familiare; S: sporadica; U: non noto) ed eventuali riferimenti in letteratura.

## ORIGINAL ARTICLE

# Is CFTR 621+3 A>G a cystic fibrosis causing mutation?

Monica Forzan<sup>1,3</sup>, Leonardo Salviati<sup>1,2,3</sup>, Vanessa Pertegato<sup>1,2</sup>, Alberto Casarin<sup>1,2</sup>, Alice Bruson<sup>1</sup>, Eva Trevisson<sup>1,2</sup>, Elena Di Gianantonio<sup>1</sup> and Maurizio Clementi<sup>1</sup>

The 621+3 A>G variant of the *CFTR* gene was initially detected in four Greek patients with a severe form of cystic fibrosis, and it is reported to impair *CFTR* mRNA splicing. We present three lines of evidence that argue against the pathogenicity of this variant. First, its allelic frequency in the Italian population was 0.4%. Even considering the lowest value in the confidence interval we would expect 10% of Italian CF patients to be heterozygotes for this variant, whereas it has been reported only in one patient (0.04% of Italian CF patients). Second, expression of the 621+3 A>G variant in HeLa cells using a hybrid minigene showed that  $39.5 \pm 1.1\%$  of transcripts were correctly spliced, indicating that its effects on mRNA splicing are similar to those of the *CFTR* intron 8 5T variant, associated with congenital bilateral absence of vas deferens (CBAVD), but not with CF. Third, we have identified an asymptomatic individual who harbored the 621+3 A>G variant in *trans* with the Q552X mutation. Because 621+3 A>G is often included in population-screening programs, this information is critical to provide adequate counseling to patients. Further work should be aimed at investigating whether this variant may have a role in CBAVD or atypical CF.

Journal of Human Genetics advance online publication, 6 November 2009; doi:10.1038/jhgc.2009.115

**Keywords:** cystic fibrosis; hybrid minigene; population screening; splicing

## INTRODUCTION

Cystic fibrosis (CF; MIM no. 219700) is one of the most frequent autosomal recessive disorder in the Caucasian population, affecting 1:2500 individuals and it is caused by mutations of the *CFTR* gene.<sup>1</sup> The frequency of heterozygous carriers in the European population is about 1 in 25 individuals; therefore several countries have instituted specific population-screening programs to identify carriers. Because more than 1600 different *CFTR* mutations have been reported in literature, these programs usually provide testing for a panel which includes the 20–50 most common *CFTR* mutations in that particular population.<sup>2</sup> However, the pathogenicity of some mutations included in these panels has been disputed. An example is the I148T variant initially reported as a CF mutation, which is now considered a non-pathogenic nucleotide change.<sup>3</sup>

We have chosen to study the 621+3 A>G nucleotide change because one of us had tested positive for this variant during a CF-screening program. The 621+3 A>G variant was first reported to the Cystic Fibrosis Mutation Database (<http://www.genet.sickkids.on.ca/cftr/app>) as a neutral polymorphism. Subsequently, it was described as a severe disease-causing mutation in four CF patients. It was reported to affect the splicing of exon 4, leading to the formation of a non-functional CFTR protein.<sup>4</sup> We now present epidemiological,

functional and clinical evidence suggesting that the 621+3 A>G variant should not be considered a CF disease-causing mutation.

## MATERIALS AND METHODS

### Analysis of 621+3 A>G variant in the general population

We studied a cohort of 250 healthy individuals of northern Italian origin, aged 20–45 years. The screening of the 621+3 A>G variant has been performed by High Resolution Melting (HRM) analysis using the Rotorgene 6000 (Corbett Life Sciences, Concorde, NSW, Australia). PCR primers were: FOR 5'-CAC ATTGGAATGCAGATGAGA-3' and REV 5'-TCCCTTACTTGTACCAGCTC ACT-3'; amplification conditions were 95 °C 5 min, 50 cycles of 95 °C 10 s, 60 °C 20 s and a final step of 72 °C 5 min. HRM analysis was performed between 71 and 83 °C. One 621 +3 A>G heterozygote and one wild-type homozygote were used as controls. All samples showing the HRM profile of the heterozygote have been sequenced using the amplification primers.

Expected frequencies of CF genotypes in the population were calculated using the Hardy–Weinberg equation.<sup>5</sup> 95% confidence interval (CI) of the proportion was calculated as described.<sup>6</sup>

### Generation of the minigene construct

To obtain the backbone of the artificial minigene, a 1.8 kb fragment encompassing the entire human beta-globin gene was amplified from genomic DNA of a healthy volunteer using primers FOR 5'-CTTCTGCTAGCAGTCAGGG

<sup>1</sup>Clinical Genetics Unit, Department of Pediatrics, University of Padova, Padova, Italy and <sup>2</sup>Hematology-Oncology Laboratory, Department of Pediatrics, University of Padova, Padova, Italy

Correspondence: Professor M Clementi, Department of Pediatrics, Clinical Genetics Unit, University of Padova, via Giustiniani 3, Padova, Italy.  
E-mail: maurizio.clementi@unipd.it

<sup>3</sup>These authors contributed equally to this work.

Received 20 July 2009; revised 13 October 2009; accepted 14 October 2009

CAGAGCCATCTA-3' REV 5'-CTTCTGGGCCCTTTGCAGCCTCACCTTCTT T-3' (which introduce *NheI* and *ApaI* sites in the product) and the Expand Long Template PCR System (Roche Diagnostics GmbH, Mannheim, Germany). PCR conditions were 94 °C 2 min 15 cycles of 94 °C 10 s, 65 °C 30 s, 68 °C 3 min, followed by 20 cycles of 94 °C 15 s, 65 °C 30 s, 68 °C 3 min, with a 15 s increase every cycle, and a final extension step of 68 °C 7 min.

After enzymatic digestion it was cloned into the corresponding sites of pCDNA3.1 (Invitrogen, Carlsbad, CA, USA).

An artificial multiple cloning site containing the *XhoI*, *NotI* and *HindIII* sites (in capital letters in the following sequence) tgtacaCTCGAGttatGCGGCCGC tactAAGCTTGtACA was then cloned into the *BsrGI* site within intron 2 of the  $\beta$ -globin gene.

A 853 bp fragment encompassing 341 bp of intron 3, exon 4 and 296 bp of intron 4 of the CFTR gene was amplified from genomic DNA of a healthy individual identified as carrier of the 621+3 A>G through population screening, using primers: FOR 5'-CTTCTCTCGAGAAAAGGGAAATGCTTTAGAA ACTG-3' and REV 5'-CTTCTAAGCTTTGAGTCATCTTAACAGGAAACCA-3' (which introduce *XhoI* and *HindIII* restriction sites in the PCR product) and Applied Biosystems AmpliTaq Gold DNA Polymerase. PCR conditions were: 94 °C 10 min, 30 cycles of 95 °C 30 s, 60 °C 40 s, 72 °C 1 min and a final extension step 72 °C 10 min. After digestion the fragment was cloned into the minigene vector, and sequenced. One clone containing the wild-type allele, and one with the mutation were retained for expression experiments.

### Minigene expression and analysis of transcripts

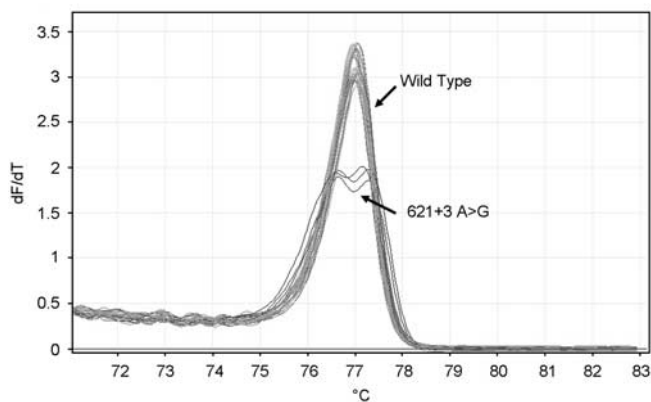
HeLa cells ( $2 \times 10^5$ ) were transfected with 400 ng of the wild type or the mutated minigene using Effectene reagent (Qiagen GmbH, Hilden, Germany) according to the manufacturer's instruction. After Total RNA was extracted and retro-transcribed as reported<sup>7</sup> using primers FOR 5'-TCTGTCCACTCCTGATGCTG-3' and REV 5'-CACTGGTGGGGTGAATCTT-3' located on  $\beta$ -globin exon 2 and 3 and Roche Taq DNA polymerase. PCR conditions were: 94 °C 3 min, 30 cycles of 94 °C 1 min, 55 °C 1 min, 72 °C 45 s and a final extension step 72 °C 7 min. This set of primers does not amplify eventual ectopically expressed CFTR transcripts.

PCR products were separated either on a 3% agarose gel or on a 12% acrilamide gel. Individual bands were excised and sequenced using amplification primers as described.<sup>8</sup> PCR products were also labeled with  $\alpha^{32}$ P-dATP using the "last cycle hot" procedure<sup>9</sup> and after gel separation were visualized and quantitated using a Storm Phosphorimager (Molecular Dynamics, Sunnyvale, CA, USA). The quantitation was performed on three independent experiments.

## RESULTS

### Epidemiology

Two out of 250 normal individuals (500 alleles) were heterozygous for the 621+3 A>G mutation giving an allele frequency of 0.4%, 95% CI (0.11–1.45). Figure 1 shows the abnormal HRM profile in these individuals.



**Figure 1** HRM profiles of the wild-type individuals, and of heterozygous carriers of the 621+3 A>G variant (see Materials and methods section for details).

Based on these data the expected frequency of individuals harboring the 621+3 A>G allele in compound heterozygosity with another CF allele would be 1:7812 (30% of CF patients), whereas the frequency of homozygotes would be 1:62500 (4% of CF patients). Even considering the lowest value in the 95% CI the expected frequencies would be 1:24010 (10% of CF patients) for heterozygotes, and 1:800000 (0.3% of CF patients) for homozygotes.

The Italian CFTR mutation database reports only one patient who was compound heterozygous for the 621+3A>G allele out of a total of 2253 affected individuals (4506 alleles) (<http://spazioinwind.libero.it/laboratoriCF/frequenze.htm>).

### Molecular studies

Because epidemiological data were arguing against the pathogenicity of the 621+3 A>G, we investigated CFTR mRNA splicing using a minigene-based approach.

Constructs containing wild-type and mutant CFTR exon 4 and the intron–exon boundaries were cloned into the pCDNA3.1- $\beta$ -globin construct and expressed in HeLa cells. mRNA was analyzed by RT-PCR using  $\beta$ -globin-specific primers (Figure 2a), that do not amplify eventual ectopically expressed CFTR transcripts. We also did not find evidence of expression of the endogenous  $\beta$ -globin gene (not shown).

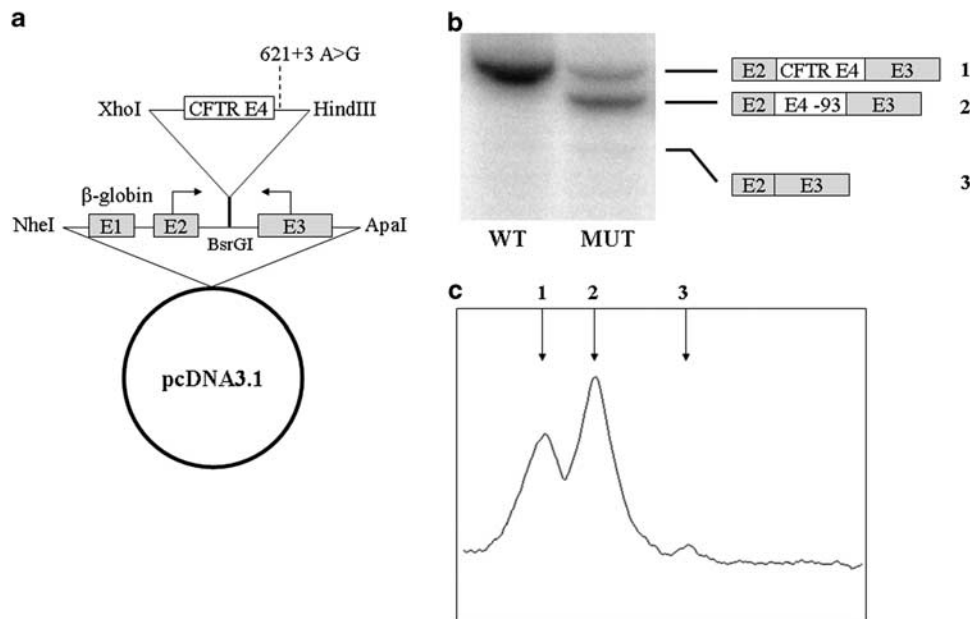
As seen in Figure 2b the transcript derived from the wild-type construct is correctly spliced, producing a single transcript containing the entire CFTR exon 4. Instead, the mutated construct yielded three different bands, one corresponding to the correctly spliced transcript, one resulting from the activation of a cryptic splice site within exon 4, 93 bp upstream of the exon 4/intron 4 boundary, and a faint third band corresponding to a transcript lacking the entire CFTR exon 4. Although these transcripts maintain the reading frame of the mRNA, the resulting proteins are likely to be non-functional.<sup>10</sup> Relative quantification of radiolabelled RT-PC products (Figure 2c) was performed on three independent experiments and revealed that the aberrantly spliced forms account for  $58.5 \pm 0.8\%$  (transcript 2 in Figure 1b) and  $2.0 \pm 0.4\%$  (transcript 3 in Figure 2b), whereas the correctly spliced product accounts for  $39.4 \pm 1.1\%$  of total transcripts (transcript 1 in Figure 2b).

### Patient data

In the course of the study we identified by chance an individual who harbored the 621 +3A>G variant in *trans* with the pathogenic Q552X-truncating mutation. She is an 8-year-old girl, first child of healthy unrelated parents. Immunoreactive trypsinogen at birth was normal. She had an unremarkable medical history in particular; she never experienced severe respiratory infections, episodes of wheezing, sinusitis or recurrent diarrhea. Stature and weight were both above the 75th centile. She was identified to be the carrier of both alterations because her parents underwent CF screening during a second pregnancy: the mother was found to harbor 621+3 A>G, whereas the father had tested positive for Q552X. After the molecular diagnosis she underwent a thorough medical evaluation. Chest X-Ray was normal, serum amylase, lipase and fecal chymotrypsin levels were all within normal ranges. Chloride sweat test was repeated three times. In one occasion chloride sweat test was pathologic (75 mEq/l), and in two occasions it yielded borderline results (49 mEq/l).

## DISCUSSION

In Southern European populations the proportion of non F508del alleles in CF carriers is high,<sup>11</sup> but the pathogenicity of some of these variants is disputed. Defining the exact role of these nucleotide



**Figure 2** (a) Structure of the hybrid minigene (see Materials and methods section for details). The arrows indicate the primers used for RT-PCR analysis. Boxed in gray,  $\beta$ -globin exons 1–3, in white *CFTR* exon 4. Expression is driven by the CMV promoter of pcDNA3.1. (b) RT-PCR analysis of mRNA of cell expressing the wild-type (WT) minigene, or the construct harboring the 621+3 A>G variant (MUT). Fragments were labeled with  $^{32}\text{P}$  using the last cycle hot procedure (see Materials and methods section for details) and visualized with a phosphor-imager. On the right the composition of each transcript. The experiment was performed in triplicate. (c) Densitometric profile above the gel. Arrows indicate the peaks corresponding to each fragment. Quantitation of three independent experiments revealed that fragment 1 represents  $39.4 \pm 1.1\%$  of total transcripts, fragment 2  $58.5 \pm 0.8\%$  and fragment 3  $2.0 \pm 0.4\%$ .

substitutions is essential to provide genetic counseling to subjects identified through population screening.

The 621+3 A>G variant has been reported in Greek patients with severe form of CF. We now present three lines of evidence that argue against the pathogenicity of the 621+3 A>G mutation. The first set of data comes from epidemiological analysis. We observed that this variant is more frequent in the general population than in patients. Its allelic frequency in our sample of healthy Italian controls was 0.4% (CI 95% 0.11–1.45), making it one of the most frequent CF alleles. In other words, even considering the lowest value in the 95% CI, we would have expected to find at least 225 patients with this mutation in the Italian CF registry. In contrast, it was reported only in 1/2253 CF patients. It should be noted that in most Italian centers mutation screening in CF patients is performed by denaturing gradient gel electrophoresis or DHPLC, therefore this mutation should have been identified if present.

The second line of evidence derives from our minigene experiments. Our data clearly show that although this nucleotide change does affect mRNA processing, a considerable portion of transcripts ( $39.4 \pm 1.1\%$ ) are correctly spliced, thus giving rise to functional CFTR proteins. Another mutation targets the same splice site in position +1 and does not allow formation of any residual correctly spliced transcripts.<sup>10</sup> Interestingly, the 621+1 G>T mutation causes the activation of the same cryptic splice within exon 4. It should be noted that position +3 in the 5' splice site is less critical than positions +1 or +2, with both A and G nucleotides being almost equally represented in this position,<sup>12</sup> therefore it is not surprising that 621+3 A>G has a less detrimental effect on mRNA splicing than 621+1 G>T. Although mRNA splicing efficiency can vary between different tissues,<sup>13</sup> these data suggest that the effect of 621+3 A>G on mRNA splicing is similar to that of the *CFTR* intron 8 5 T variant,<sup>14</sup> involved in congenital bilateral absence of vas deferens, but not in CF.

Our data are not in contrast with the results obtained by Tzetzis *et al.*<sup>4</sup> In fact, we identified the same splicing alteration qualitatively. The discrepancy in the conclusions lies in the fact that methodology employed by Tzetzis *et al.*<sup>4</sup> could not discriminate the relative contribution of each individual allele to the total pool of correctly spliced transcripts, and therefore they could not rule out the presence of correctly spliced transcripts deriving from the 621+3 A>G allele. In fact, traditional RT-PCR analysis of patient mRNA cannot distinguish transcripts deriving from individual alleles of compound heterozygotes (as were the four reported 621+3 A>G patients), unless the analysis is coupled to either selective digestion of one of the two alleles by a specific restriction endonuclease,<sup>15</sup> or if the amplification protocol is allele-specific.<sup>14</sup> This type of analysis is not always possible because of the absence of informative polymorphic sites in the transcript. Hybrid minigenes overcome these limitations, and in fact they have been widely used to study *CFTR*-splicing alterations.<sup>16–18</sup>

Finally, we have identified an asymptomatic 8-year-old girl harboring 621+3 A>G allele in *trans* with a truncating *CFTR* mutation. Although she had a borderline sweat test she did not present with any symptom of CF. It will be important to monitor the patient to determine if in the future she will develop an atypical form of CF, because the quantity of correctly spliced transcripts may differ among various organs of the same patient, contributing to differential organ disease manifestations.

A likely explanation of the severe phenotype of the four reported patients (and of the single Italian patient with harboring 621+3 A>G) is that another undetected mutation was in *cis* with 621+3 A>G. A complex allele comprising 621+3 A>G and a 4332delTG has been identified in a patient of north-African origin.<sup>19</sup> It is likely that the mutation analysis protocol, carried out by denaturing gradient gel electrophoresis, would have identified this mutation (if present) in the four Greek patients with 621+3 A>G (in any case we ruled out the presence of this variant in all individuals carrying the 621+3 A>G

allele included in our study). However, this technique cannot detect heterozygous deletions of individual exons, that may be present in up to 60% of classic CF patients in whom a single-point mutation is detected.<sup>20</sup> This case is similar to that of the I148T variant, which by itself is not a CF disease-causing mutation, but in some patients it is found in *cis* with the 3199del6 mutation.<sup>21</sup> Extensive analysis of *CFTR*, including gene rearrangements, should be performed in all patients with typical CF symptoms harboring the 621+3 A>G allele, to identify a possible other mutation in *cis* with this variant, which could account for the severe phenotype.

In conclusion, our data suggest that 621+3 A>G should not be considered a severe CF mutation because of its frequency in the population, because it still allows the synthesis of significant amounts of functional CFTR protein, and the absence of clinical disease in a compound (621+3 A>G/ Q552X) heterozygote individual. Although a large study performed in the Italian population did not detect the presence of this mutation in infertile men,<sup>22</sup> further epidemiological studies should be aimed at determining whether 621+3 A>G in isolation may have a role in congenital bilateral absence of vas deferens or in atypical forms of CF.

#### ACKNOWLEDGEMENTS

Dr Salviati is supported by a grant from AFM and by Telethon Italia Grant GGP06256.

- Davies, J. C., Alton, E. W. & Bush, A. Cystic fibrosis. *BMJ* **335**, 1255–1259 (2007).
- Tomaiuolo, R., Spina, M. & Castaldo, G. Molecular diagnosis of cystic fibrosis: comparison of four analytical procedures. *Clin Chem Lab Med* **41**, 26–32 (2003).
- Claustres, M., Altieri, J. P., Guittard, C., Templin, C., Chevalier-Porst, F. & Des Georges, M. Are p.I148T, p.R74W and p.D1270N cystic fibrosis causing mutations? *BMC Med Genet* **5**, 19 (2004).
- Tzetis, M., Efthymiadou, A., Doudounakis, S. & Kanavakis, E. Qualitative and quantitative analysis of mRNA associated with four putative splicing mutations (621+3A>G, 2751+2T->A, 296+1G->C, 1717-9T->C-D565G) and one nonsense mutation (E822X) in the *CFTR* gene. *Hum Genet* **109**, 592–601 (2001).
- Hardy, G. H. Mendelian Proportions in a Mixed Population. *Science* **28**, 49–50 (1908).
- Newcombe, R. G. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med* **17**, 857–872 (1998).
- Trevisson, E., Salviati, L., Baldoiu, M. C., Toldo, I., Casarin, A., Sacconi, S. *et al.* Argininosuccinate lyase deficiency: mutational spectrum in Italian patients and identification of a novel ASL pseudogene. *Hum Mutat* **28**, 694–702 (2007).
- Salviati, L., Trevisson, E., Baldoiu, M. C., Toldo, I., Sartori, S., Calderone, M. *et al.* A novel deletion in the *GJA12* gene causes Pelizaeus-Merzbacher-like disease. *Neurogenetics* **8**, 57–60 (2007).
- Sacconi, S., Salviati, L., Nishigaki, Y., Walker, W. F., Hernandez-Rosa, E., Trevisson, E. *et al.* A functionally dominant mitochondrial DNA mutation. *Hum Mol Genet* **17**, 1814–1820 (2008).
- Zielenski, J., Bozon, D., Markiewicz, D., Aubin, G., Simard, F., Rommens, J. M. *et al.* Analysis of *CFTR* transcripts in nasal epithelial cells and lymphoblasts of a cystic fibrosis patient with 621+1G->T and 711+1G->T mutations. *Hum Mol Genet* **2**, 683–687 (1993).
- Chillon, M., Casals, T., Gimenez, J., Ramos, M. D., Palacio, A., Morral, N. *et al.* Analysis of the *CFTR* gene confirms the high genetic heterogeneity of the Spanish population: 43 mutations account for only 78% of CF chromosomes. *Hum Genet* **93**, 447–451 (1994).
- Cartegni, L., Chew, S. L. & Krainer, A. R. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* **3**, 285–298 (2002).
- Mak, V., Jarvi, K. A., Zielenski, J., Durie, P. & Tsui, L. C. Higher proportion of intact exon 9 *CFTR* mRNA in nasal epithelium compared with vas deferens. *Hum Mol Genet* **6**, 2099–2107 (1997).
- Rave-Harel, N., Kerem, E., Nissim-Rafinia, M., Madjar, I., Goshen, R., Augarten, A. *et al.* The molecular basis of partial penetrance of splicing mutations in cystic fibrosis. *Am J Hum Genet* **60**, 87–94 (1997).
- Martella, M., Salviati, L., Casarin, A., Trevisson, E., Opocher, G., Polli, R. *et al.* Molecular analysis of two uncharacterized sequence variants of the *VHL* gene. *J Hum Genet* **51**, 964–968 (2006).
- Zuccato, E., Buratti, E., Stuani, C., Baralle, F. E. & Pagani, F. An intronic polypyrimidine-rich element downstream of the donor site modulates cystic fibrosis transmembrane conductance regulator exon 9 alternative splicing. *J Biol Chem* **279**, 16980–16988 (2004).
- Pagani, F., Stuani, C., Zuccato, E., Kornblihtt, A. R. & Baralle, F. E. Promoter architecture modulates *CFTR* exon 9 skipping. *J Biol Chem* **278**, 15111–15117 (2003).
- Hefferon, T. W., Groman, J. D., Yurk, C. E. & Cutting, G. R. A variable dinucleotide repeat in the *CFTR* gene contributes to phenotype diversity by forming RNA secondary structures that alter splicing. *Proc Natl Acad Sci USA* **101**, 3504–3509 (2004).
- Loumi, O., Ferec, C., Mercier, B., Creff, J., Fercot, B., Denine, R. *et al.* *CFTR* mutations in the Algerian population. *J Cyst Fibros* **7**, 54–59 (2008).
- Hantash, F. M., Redman, J. B., Starn, K., Anderson, B., Buller, A. & McGinniss, M. J. *et al.* Novel and recurrent rearrangements in the *CFTR* gene: clinical and laboratory implications for cystic fibrosis screening. *Hum Genet* **119**, 126–136 (2006).
- Monaghan, K. G., Highsmith, W. E., Amos, J., Pratt, V. M., Roa, B., Friez, M. *et al.* Genotype-phenotype correlation and frequency of the 3199del6 cystic fibrosis mutation among I148T carriers: results from a collaborative study. *Genet Med* **6**, 421–425 (2004).
- Morea, A., Cameran, M., Rebuffi, A. G., Marzenta, D., Marangon, O., Picci, L. *et al.* Gender-sensitive association of *CFTR* gene mutations and 5T allele emerging from a large survey on infertility. *Mol Hum Reprod* **11**, 607–614 (2005).