



Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN BIOCHIMICA

Ciclo XXII

Settore scientifico disciplinare: MED/03

LO STUDIO DI FAMIGLIE CON PATOLOGIA
GENETICA PER L'IDENTIFICAZIONE DEL GENE
MALATTIA:
SINDROME DI MUNGAN E TROMBOCITOPENIA
AUTOSOMICA DOMINANTE (THC2)

Coordinatore:
Prof. Giorgio Lenaz

Relatore:
Prof. Giovanni Romeo

Candidato:
Giovanni Castegnaro

Esame finale anno 2010

1. INTRODUZIONE

1.1 L'ANALISI DI LINKAGE

I più semplici caratteri genetici sono quelli la cui presenza o assenza dipende dal genotipo in un singolo locus; ciò non implica per forza che il carattere sia controllato da un singolo gene, può essere anche controllato da due o più geni, i quali vengono ereditati assieme; questi caratteri sono definiti mendeliani.

I caratteri che invece sono controllati da più geni indipendenti tra loro e la cui espressione può essere influenzata dall'ambiente sono definiti caratteri multifattoriali.

I caratteri mendeliani sono chiamati così in onore di Gregor Mendel, il quale, verso la metà del 1800, dedusse le leggi che governano la trasmissione da una generazione all'altra dei suddetti caratteri studiando la trasmissione di fenotipi quali il colore del fiore, del seme, la sua rugosità, eccetera, nelle piante di pisello; inoltre egli fu il primo ad intuire l'esistenza degli alleli come determinanti genotipici di un fenotipo; le leggi da egli ricavate possono essere così riassunte:

- **Prima Legge Di Mendel**, o della segregazione degli alleli: i due alleli di uno stesso gene portati da un determinato individuo non possono mai essere trasmessi insieme alla progenie dell'individuo; i figli riceveranno o l'uno o l'altro allele (gli alleli cioè segregano alla meiosi nei gameti dell'individuo e quindi nella generazione successiva).
- **Seconda Legge Di Mendel**, o dell'assortimento indipendente: la trasmissione simultanea di due o più coppie di alleli appartenenti a geni indipendenti avviene in modo indipendente, ossia altro non è che il risultato di due o più segregazioni indipendenti che avvengono simultaneamente (come se ciascuna coppia di alleli fosse situata su una coppia diversa di cromosomi).

Ancora oggi le fondamenta teoriche della mappatura dei geni per i caratteri mendeliani sono interamente contenute nell'enunciazione delle due leggi di Mendel.

La prima legge consente di definire le diverse modalità di trasmissione dei caratteri: un carattere è definito dominante quando è fenotipicamente visibile nell'eterozigote; è invece definito recessivo quel carattere che si manifesta solamente nell'omozigote. Di conseguenza le modalità di trasmissione dei caratteri saranno definite Dominante o Recessiva.

La seconda legge di Mendel consente di formulare i concetti di ricombinazione e concatenazione (o linkage) tra loci.

Da un punto di vista statistico, perché la seconda legge venga rispettata, gli alleli di geni diversi dovrebbero segregare assieme il 50% delle volte, e ciò è vero se si considerano loci posizionati su cromosomi diversi.

Nel caso invece di due loci "sintetici" (che si trovano cioè sullo stesso cromosoma) si potrebbe ipotizzare che essi segreghino sempre insieme. Tuttavia, questa ipotesi non tiene conto del fenomeno della ricombinazione. Durante la ricombinazione infatti, due loci sintetici possono essere separati qualora un crossing over si verifichi tra di essi (figura 1.1), e questo fenomeno sarà tanto più frequente quanto più i due loci saranno lontani lungo il cromosoma. Nell'eventualità che la distanza sia considerevole, si origineranno anche in questo caso mediamente un 50% di cromatidi ricombinanti e un 50% di cromatidi non ricombinanti, e la seconda legge di Mendel è nuovamente rispettata.

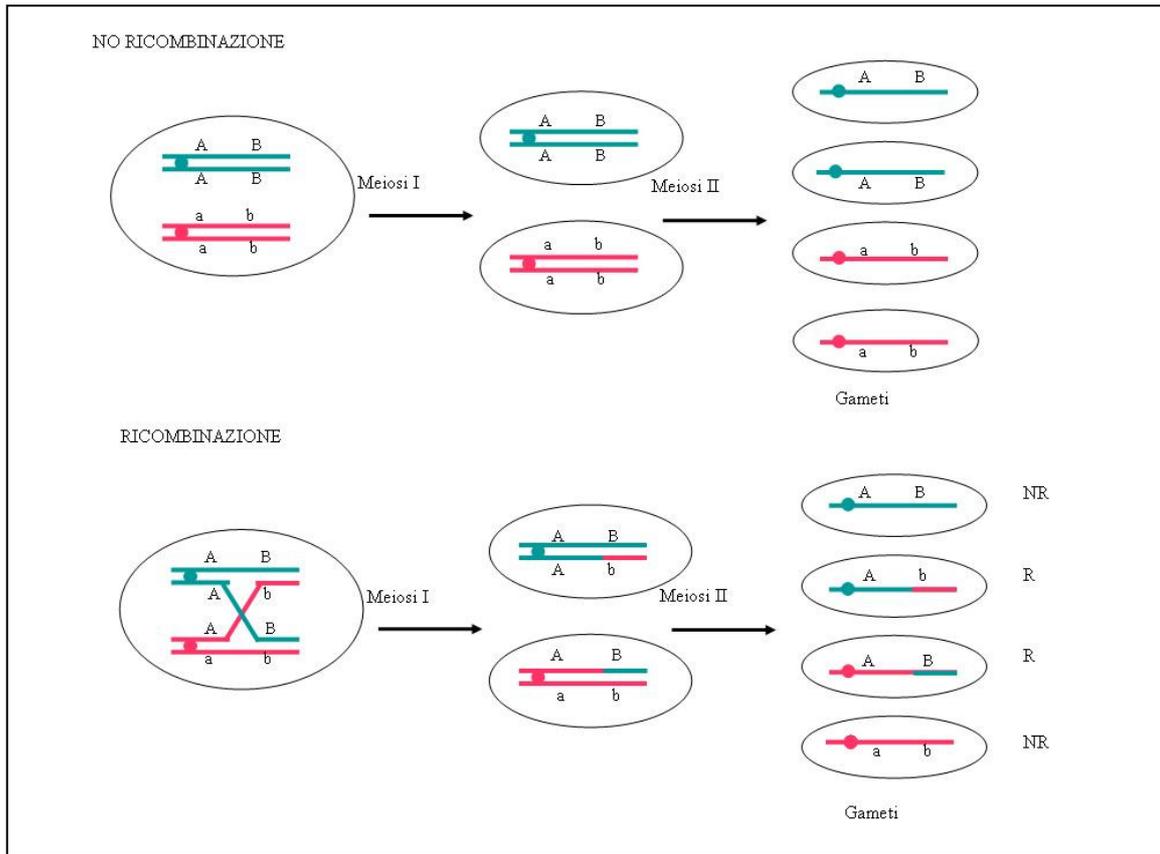


Figura 1.1. Il fenomeno della ricombinazione. Nel caso in cui durante la meiosi si abbia un evento di ricombinazione, i gameti che si originano potranno essere ricombinanti (R) o non ricombinanti (NR).

In termini di frazione di ricombinazione θ (definita come il rapporto tra il numero di ricombinanti e il numero di ricombinanti sommato al numero di non ricombinanti) quindi, il valore che essa assume sempre tra due loci su cromosomi diversi o lontani sul medesimo cromosoma è dunque di 0,5, che è il valore massimo che la frazione di ricombinazione può assumere. Infatti, per quanti eventi di crossing over possano verificarsi tra due loci pur molto distanti sul cromosoma, l'effetto complessivo risulterà sempre della stessa proporzione di 50 a 50.

Il fenomeno del linkage (o concatenazione) è un'eccezione alla seconda legge di Mendel, e riguarda il comportamento di loci posizionati vicini sullo stesso cromosoma. Tra due loci in linkage, θ assume valori superiori o uguali a 0 e inferiori a 0,5. Questa

deviazione di θ dal valore 0,5 è direttamente proporzionale alla distanza tra i loci sintenici: la frazione di ricombinazione sarà tanto più bassa quanto più sono vicini i loci, e viceversa. Questo fenomeno consente di utilizzare la frazione di ricombinazione come indice della distanza genetica tra due loci, infatti, la misura universalmente riconosciuta della distanza genetica è il centimorgan (cM) o unità di mappa, che corrisponde all'1% di ricombinazione, cioè a $\theta = 0,01$.

Si può quindi intuire la differenza tra mappa fisica (che mostra l'ordine dei loci sul cromosoma e la loro distanza in termini di coppie di basi) e mappa genetica: quest'ultima dovrebbe mostrare lo stesso ordine dei loci lungo il cromosoma, ma le distanze tra i loci sono espresse in funzione della frequenza di ricombinazione. Poiché la frazione di ricombinazione non è una misura additiva, mentre l'unità di mappa lo è, si rende necessaria una funzione matematica che le metta in relazione, che è la cosiddetta funzione di mappa. Vi sono varie funzioni di mappa, distinguibili a seconda che siano in grado o meno di prendere in considerazione il fenomeno dell'interferenza, secondo cui la formazione di un chiasma influenza la formazione di un altro, e ciò ha ripercussioni sulla distribuzione dei crossing over lungo i cromosomi.

L'intento dell'analisi di linkage è quello di stimare la frequenza di ricombinazione θ e di provare che l'eventuale sua deviazione dal valore di 0,5 sia statisticamente significativa, vale a dire sia frutto di un'effettiva concatenazione e non sia invece casuale.

Il metodo più utilizzato per determinare questa significatività statistica è il calcolo dei *lod scores* o del rapporto di massima verosimiglianza, che sfrutta la regola secondo cui in un reincrocio in cui i rapporti di indipendenza o concatenazione fra due geni (o due marcatori) sono ignoti, la probabilità che uno o più figli mostrino un fenotipo parentale sarà sempre più alta in presenza di concatenazione piuttosto che in assenza di essa. Per convenzione, si considera significativo un risultato quando alla stima di θ corrisponde un valore di lod-score superiore a 3,

che indica che l'ipotesi di concatenazione in corrispondenza di quel determinato valore di θ è più di 1000 (10^3) volte maggiormente probabile di quella di indipendenza. Per escludere la concatenazione e invece sostenere l'indipendenza dei loci, invece, il valore di lod-score deve essere inferiore o uguale a -2.

Per poter effettuare un'analisi di linkage però è preliminarmente necessaria la disponibilità di due risorse fondamentali: un marcatore genetico e una o più famiglie in cui ricorre la malattia mendeliana che si intende mappare.

Per marcatore genetico si intende un qualsiasi carattere mendeliano polimorfico attraverso cui si può tracciare il passaggio dei cromosomi lungo un albero genealogico. Il requisito più importante che un marcatore genetico deve possedere al fine dell'analisi di linkage è l'informatività, che è espressa dall'eterozigosità, che indica la probabilità che un individuo a caso della popolazione sia eterozigote per il marcatore. Una meiosi è definita informativa se è possibile ricostruire con esattezza la segregazione degli alleli di un marcatore dai genitori ai figli, e ciò è sempre possibile se i genitori sono doppi eterozigoti (non uguali) per il marcatore; d'altra parte nei casi in cui uno od entrambi i genitori sono omozigoti per il marcatore e nella metà dei casi in cui entrambi i genitori sono eterozigoti uguali, non si ha la possibilità di ricostruire con esattezza la segregazione degli alleli, di conseguenza le meiosi sono dette non informative.

I polimorfismi del DNA che sono stati maggiormente utilizzati nell'analisi di linkage sono i microsatelliti, ripetizioni di sequenze di-, tri- o tetranucleotidiche altamente polimorfiche e ampiamente diffuse in tutto il genoma. Sono queste caratteristiche ad averli resi marcatori d'elezione rispetto agli RFLP (Restriction Fragment Length Polymorphisms) e ai minisatelliti. Infatti da una parte gli RFLP, pur essendo ben distribuiti sul genoma, forniscono una bassa informatività, visto che formando un sistema biallelico il valore più alto per l'eterozigosità è di 0,5; dall'altra i minisatelliti, pur con la loro elevata

eterozigotità, sono mal distribuiti sul genoma e inoltre sono tipizzabili soltanto attraverso Southern Blot e non attraverso PCR, il che rende il metodo estremamente più laborioso. Di recente, l'introduzione di nuove piattaforme tecnologiche ha consentito di passare agli SNP (Single Nucleotide Polymorphisms) come marcatori di elezione per l'analisi di linkage. Benché questi siano molto meno informativi dei microsatelliti, sono decisamente più diffusi nel genoma, e le nuove tecnologie consentono di effettuare analisi di centinaia di migliaia o milioni di SNP per singolo individuo, con tempi e prezzi inferiori rispetto all'analisi dei microsatelliti.

L'analisi di caratteri mendeliani nell'uomo si basa sulla raccolta dei dati provenienti dall'osservazione della distribuzione dei caratteri ereditari in alberi genealogici più o meno estesi; Lo studio inizia generalmente con l'osservazione di una persona affetta (il probando o l'index-case/il caso-indice), a partire dal quale viene ricostruito l'albero genealogico della famiglia su quante più generazioni possibile, cercando di trarre tutte le informazioni fenotipiche utili sui componenti della famiglia, Nella figura 1.2 sono spiegati i simboli utilizzati nella rappresentazione degli alberi genealogici.

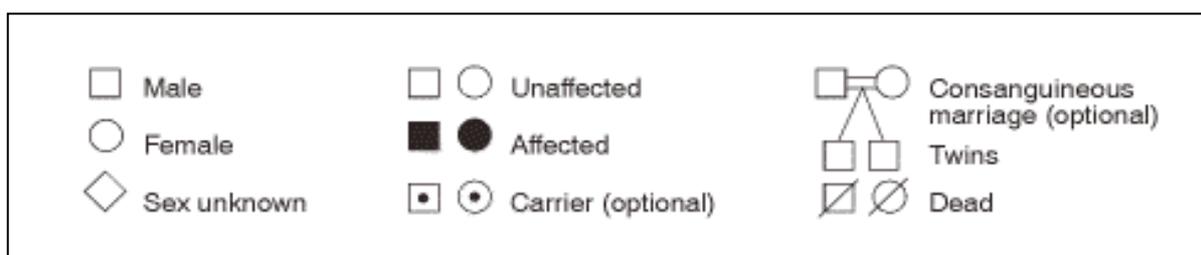


Figura 1.2: Simboli utilizzati nella rappresentazione dei pedigree

Il passaggio successivo è quello di cercare di stabilire come viene ereditato il carattere nel pedigree (albero genealogico) in esame.

I caratteri mendeliani possono essere determinati da *loci* su cromosomi autosomici o su cromosomi sessuali. Perciò si distingue tra eredità autosomica (quando il locus è su un cromosoma autosomico) ed ereditarietà legata ai cromosomi sessuali.

La trasmissione dei caratteri mendeliani è classificata in base al cromosoma su cui si trova il locus ed in base alla dominanza o recessività del carattere; vi sono quattro tipi principali di trasmissione per i caratteri mendeliani, conoscerli e conoscere le loro caratteristiche è fondamentale per definire il tipo di trasmissione di una malattia osservando il pedigree di una famiglia con membri affetti.

- Malattie Autosomiche Dominanti

La trasmissione di un carattere autosomico dominante non prevede salti generazionali. Questo comporta che un individuo sano non possa generare prole affetta e che il genotipo del portatore coincida con il genotipo dell'individuo affetto. Un genitore affetto trasmette la malattia alla prole col 50% di probabilità. Ci sono alcuni fenomeni che complicano notevolmente l'interpretazione di un albero genealogico in cui ricorre una malattia autosomica dominante. In caso di *penetranza incompleta*, si possono osservare individui sani che sono in realtà portatori del gene malattia, e che possono quindi generare prole affetta. Talvolta la penetranza varia sensibilmente tra i componenti di una famiglia. Esempio classico è il caso dell'*esordio tardivo*, in cui la manifestazione clinica insorge in età avanzata. In caso di *espressività variabile*, il fenotipo può comparire in varie gradazioni tra gli affetti, distinguendosi di caso in caso anche per la differenza delle caratteristiche cliniche. Un fenomeno tipico di espressività variabile è l'*anticipazione*, che è la tendenza di alcune malattie a divenire più gravi col procedere delle generazioni. È una caratteristica comune alle malattie da mutazioni dinamiche, come quelle generate dall'instabilità del DNA.

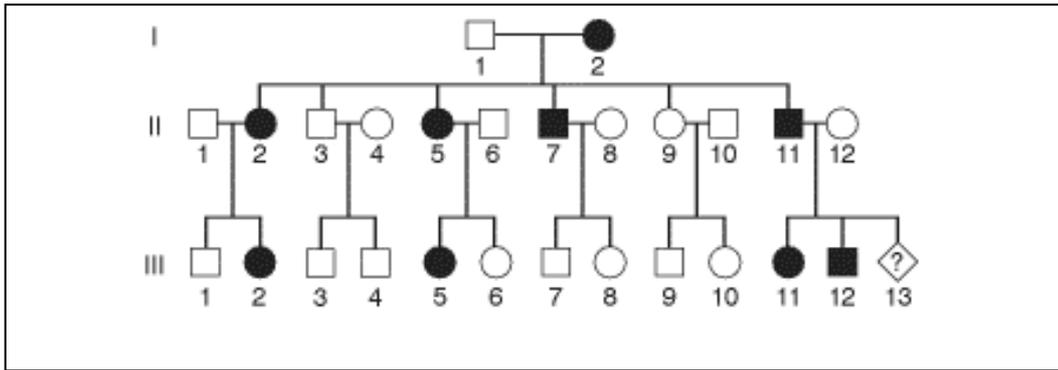


Figura 1.3 : esempio di un pedigree che presenta un carattere con trasmissione autosomica dominante

- Malattie autosomiche recessive

La trasmissione di un carattere autosomico recessivo avviene attraverso l'incrocio di due portatori sani, eterozigoti per il gene-malattia, che hanno probabilità del 25% di avere prole sana e non portatrice, il 50% di avere prole portatrice sana, il 25% di avere prole affetta. In generale, la frequenza estremamente rara del gene-malattia nella popolazione preserva da incroci tra individui affetti e tra un individuo affetto e un individuo portatore sano. Se queste eventualità si dovessero verificare comunque, nel primo caso si produrrebbe inevitabilmente progenie affetta, nel secondo si produrrebbe invece progenie affetta e portatrice sana con una proporzione di 50:50.

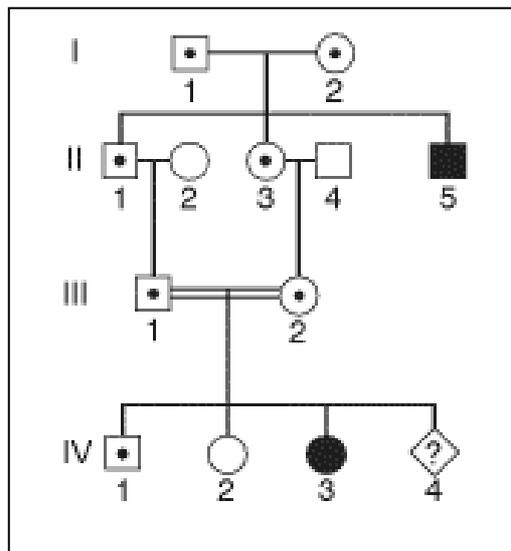


Figura 1.4: esempio di un pedigree che presenta un carattere a trasmissione autosomica recessiva

- Malattie X-linked

La manifestazione dei caratteri X-linked è influenzata dal fatto che negli individui di sesso maschile essi sono sempre in una situazione di emizigotità. La trasmissione di una malattia X-linked può seguire esclusivamente la via dalla madre alla prole di sesso maschile; il gene-malattia invece può propagarsi dal padre o dalla madre attraverso la prole di sesso femminile. Se l'ereditarietà X-linked è di tipo dominante, il fenotipo si manifesterà sia nei maschi che nelle femmine portatrici, anche se in misura più lieve. Questo comportamento è una conseguenza del fenomeno della *lyonizzazione*, che prevede l'inattivazione di uno dei due cromosomi X nelle femmine. Se invece l'ereditarietà è di tipo recessivo, si manifesterà soltanto nei maschi e obbligatoriamente salterà la generazione nel sesso femminile.

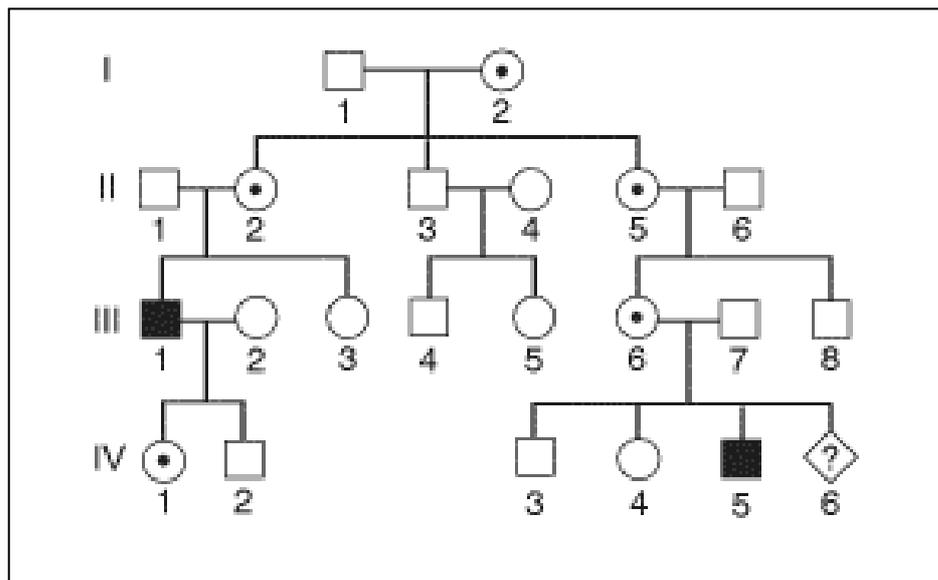


Figura 1.5 : esempio di un pedigree che presenta un carattere a trasmissione X-linked recessiva

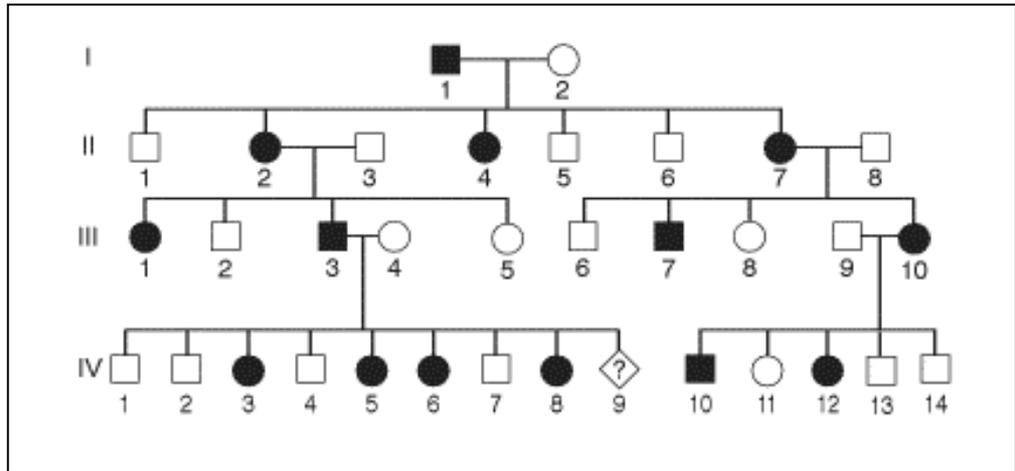


Figura 1.6 : esempio di un pedigree che presenta un carattere a trasmissione X-linked dominante

1.2 L'IDENTIFICAZIONE DI GENI MALATTIA

Durante gli ultimi anni, la ricerca in Genetica Medica si è indirizzata verso l'identificazione e la caratterizzazione di geni coinvolti in malattie genetiche e/o della componente genetica nelle malattie multifattoriali. Questo approccio è il più diretto per comprendere le basi molecolari coinvolte nella patogenesi di malattie ereditarie e presenta un riscontro anche a livello clinico. L'identificazione di un gene malattia, infatti, permette di poter svolgere una diagnosi molecolare diretta con una ricaduta immediata sia per quanto riguarda gli aspetti di prevenzione della malattia che per un'accurata Consulenza Genetica e un corretto calcolo del rischio di ricorrenza. L'analisi molecolare di un gene può consentire di determinare il fenotipo atteso sulla base della correlazione genotipo/fenotipo e/o la prognosi della malattia. La comprensione dei meccanismi patogenetici responsabili della patologia, inoltre, può talvolta indirizzare verso nuovi approcci terapeutici per la cura di difetti biochimici alla base delle malattie genetiche.

Questo aspetto della ricerca ha subito negli ultimi anni un enorme sviluppo grazie al completamento del progetto genoma umano, all'allestimento di banche dati pubbliche e, di recente, all'introduzione delle nuove tecnologie di sequenziamento.

Gli approcci sperimentali per giungere all'identificazione di geni malattia sono numerosi, ma con il progredire delle tecnologie e l'abbattimento dei costi, l'approccio d'elezione del futuro prossimo sarà il sequenziamento genomico.

Il clonaggio posizionale è una strategia sviluppata durante gli anni '80 che permette di determinare il meccanismo patogenetico di molte malattie ereditarie per cui è precluso l'approccio più tradizionale dell'analisi biochimica diretta del tessuto danneggiato. Di solito l'analisi biochimica del tessuto alterato è possibile quando il difetto genetico altera una proteina di cui si conosce il ruolo metabolico nel tessuto normale. La grande maggioranza dei geni malattia sono stati

clonati partendo dalla conoscenza della proteina codificata e del difetto biochimico patogenetico. In parecchi casi (ed un esempio è dato dall'emofilia A) ciò ha consentito la purificazione del prodotto proteico e la determinazione anche parziale della sequenza aminoacidica, da cui mediante il codice genetico sono disegnati oligonucleotidi, o meglio, per la degenerazione del codice, miscele di oligonucleotidi. Queste possono essere utilizzate come sonda per pescare il gene in librerie di cDNA o come primer per amplificare i cDNA mediante PCR.

In altri casi (quali ad esempio la fenilchetonuria, la sindrome di Lesch-Nyhan, etc.) si sono utilizzati anticorpi per la proteina che sono serviti come sonda per vagliare sistematicamente una libreria di cDNA in vettori di espressione. In una minoranza di casi (quali ad esempio gli oncogeni o i geni implicati nella riparazione del DNA) la “funzione” del gene (la capacità trasformante o la complementazione funzionale del difetto di cellule di pazienti dopo fusione in vitro) è stata utilizzata direttamente per clonarlo.

Benché tali strategie di clonaggio “funzionale” siano state utilizzate con successo, esse non sono applicabili alla maggioranza delle malattie mendeliane, elencate nel catalogo di McKusick (OMIM www.ncbi.nih.gov), per le quali non c'è alcuna informazione sul difetto genetico di base. Ne sono esempi malattie ereditarie che causano ritardo mentale, psicosi, malformazioni congenite, tumori maligni e altri fenotipi complessi. Per malattie di questo tipo è applicabile la strategia di clonaggio “posizionale”, che presuppone come unica informazione la localizzazione cromosomica del gene che si vuole clonare.

A tutt'oggi più di un centinaio di geni che, quando mutati sono alla base di patologie nell'uomo, sono stati identificati mediante clonaggio posizionale. Di solito l'iniziale mappaggio del gene malattia è effettuato mediante studi di linkage in famiglie con parecchi membri affetti. Per malattie autosomiche dominanti è sufficiente un unico pedigree (come esemplificato bene dal grande pedigree venezuelano con più di 500 membri, utilizzato per il mappaggio della corea di

Huntington, Gusella J.F. et al., 1984), mentre per malattie autosomiche recessive sono raccolti parecchi piccoli pedigree (ben 50 sono stati utilizzati per il mappaggio della fibrosi cistica). Questa seconda situazione è meno favorevole perché l'eterogeneità genetica, ovvero la presenza di più loci genici implicati nello stesso fenotipo clinico, presente in molte patologie genetiche, potrebbe rendere difficile l'interpretazione dei dati. La ricerca di linkage viene effettuata tipizzando i membri delle famiglie in esame per molteplici marcatori polimorfici di DNA fino a quando non si identifichi un marcatore che mostra un lod score maggiore di 3. Lo studio di altri polimorfismi nella regione di interesse può servire a confermare o rigettare il linkage. Il passaggio successivo è quello di restringere la regione candidata ad un intervallo genetico che, idealmente, sia al massimo di 2 cM, corrispondente ad un intervallo fisico di circa 2 Megabasi. Ciò è possibile là dove sono disponibili parecchi pedigree (malattie genetiche abbastanza comuni) o là dove il mappaggio fine è accelerato dalla disponibilità di rari pazienti con riarrangiamenti citogeneticamente visibili, che interrompono (traslocazioni) o eliminano (delezioni) il locus responsabile. Quest'ultima evenienza è relativamente comune per malattie X-linked. Per malattie autosomiche recessive rare è possibile un mappaggio preciso utilizzando famiglie in cui sono avvenuti matrimoni tra consanguinei (*homozygosity mapping*). Questo approccio consiste nella ricerca, mediante marcatori opportunamente spazati, di regioni ridotte ad omozigosi per discendenza nella progenie affetta di genitori imparentati: tali regioni indicano il mappaggio del gene malattia. Esempi rappresentativi di geni malattia mappati con questo metodo sono quelli alla base della sindrome di Bloom (MIM 210900, Lander, E. S., Botstein D., 1987), la sindrome di Werner (MIM 277700, Schellenberg, et al., 1992) e l'ataxia di Friedreich (MIM 229300, Shaw J. Et al., 1990).

Nei casi in cui la regione in cui cercare il gene sia superiore a 2 Mb l'effettiva identificazione del gene malattia si profila come un arduo obiettivo tenendo presente che al clonaggio della regione candidata deve seguire l'esame di tutte le sequenze codificanti contenute nella

regione. E' a questo punto che una mappa integrata genetica e fisica della regione critica è fondamentale. Identificare il gene responsabile di una malattia genetica senza una mappa è come cercare un ago in un pagliaio. E naturale pertanto che gli enormi progressi nel mappaggio genetico e fisico, e soprattutto i risultati del progetto genoma umano (http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml) abbiano semplificato notevolmente nel tempo la procedura di clonaggio posizionale.

Gli approcci del gene candidato non richiedono l'isolamento di nuovi geni, ma si basano sull'informazione disponibile su geni già isolati. La conferma o l'esclusione di un gene "candidato" si basa sull'identificazione positiva o negativa di mutazioni nei pazienti. Naturalmente geni candidati possono essere suggeriti per alcune malattie senza conoscere la localizzazione cromosomica del gene malattia e/o del gene candidato (**approccio del gene candidato indipendente dalla posizione**). Un esempio è rappresentato dalla sindrome di predisposizione al cancro di Li Fraumeni (MIM151623), dovuta a mutazioni germinali del gene TP53, uno dei soppressori più frequentemente coinvolto nei tumori. Tuttavia la forza di questo approccio è notevolmente aumentata quando si sappia che il gene candidato mappa nella stessa regione cromosomica del gene malattia (**approccio del gene candidato per posizione**). Sono in numero crescente i geni malattia identificati con tale approccio combinato. Esempi rappresentativi sono la sindrome di Waardenburg e il deficit di glicerolchinasasi. Per quest'ultima malattia il gene è stato isolato indipendentemente per clonaggio posizionale e per mappaggio di una sequenza espressa isolata casualmente (EST), omologa alla glicerolchinasasi batterica, alla banda p21 del cromosoma X umano, dove mappa la malattia: l'informazione combinata sulla funzione e la posizione ha portato all'identificazione del gene. Ma l'esempio più emblematico di approccio combinato è rappresentato dai geni implicati nella predisposizione al cancro coloretale senza poliposi (HNPCC o Hereditary Non Polyposis Colon Cancer, MIM 120435). Studi di linkage in famiglie affette avevano localizzato uno dei geni coinvolti

nella predisposizione al carcinoma colorettalesu 2p e nella stessa regione si osservavano gli effetti dell'instabilità di sequenze di DNA nei tumori di pazienti affetti, convergenza questa che additava candidati plausibili da testare nella stessa regione.

Benché l'arricchimento e saturazione della mappa dei trascritti potrebbe rendere sempre più facile nel futuro più prossimo l'identificazione di geni malattia con l'approccio posizionale del gene candidato, l'introduzione delle tecnologie di *next generation sequencing* segnerà sicuramente un cambio di direzione nella ricerca di geni malattia. Le nuove piattaforme tecnologiche consentiranno infatti, nei prossimi anni, di sequenziare interi genomi ad un prezzo ed in tempi decisamente inferiori a quelli delle odierne tecnologie, che consentono di sequenziare piccole porzioni di genoma in molto tempo ed a caro prezzo. E' proprio di quest'anno (2010) il primo lavoro nel quale viene descritta la prima applicazione di successo dell'*exome sequencing* (Buckingham et al. 2010); gli autori descrivono come, senza conoscere dati sulla localizzazione del gene malattia e senza avere idea di geni candidati per funzione, attraverso il sequenziamento di tutte le regioni codificanti del genoma di quattro pazienti affetti dalla Sindrome di Miller in tre pedigree separati, abbiano identificato mutazioni nel gene *DHODH*, confermando poi la presenza di mutazioni anche in altri casi sporadici.

1.3 LE MALATTIE NELLE FAMIGLIE ANALIZZATE

1.3.1 LA SINDROME DI MUNGAN

La Motilità Intestinale

Le funzioni digestive sono regolate da un complesso network neuronale, noto come Sistema Nervoso Enterico (ENS), “immerso” nelle pareti del tratto intestinale e con un’estensione che va dall’esofago fino allo sfintere anale interno. L’ENS comprende una miriade di neuroni (almeno quanti quelli contenuti nel midollo spinale, circa 10^{10} negli umani) organizzati in due maggiori strutture ganglioniche, il plesso mienterico (o di Auerbach) ed il plesso submucosale (o di Meissner). Ciascun ganglio sottoclassi di neuroni funzionalmente distinte, che includono neuroni primari afferenti intrinseci, interneuroni ascendenti e discendenti, motoneuroni eccitatori ed inibitori, vasomotoneuroni e secretomotoneuroni, organizzati in circuiti al fine di coordinare le funzioni motorie, la più nota delle quali è la peristalsi.

Ci sono evidenze che l’ENS sia coinvolto oltre che nella secrezione endocrina ed esocrina, la microcircolazione e la motilità, anche nel controllo dei processi immunitari ed infiammatori intestinali (Collins et al. 1996).

Ciò rende plausibile il fatto che ogni danno ai meccanismi del sistema nervoso enterico risulti in un ampio spettro di disordini intestinali, tra i quali ci sono le disfunzioni a carico della motilità, che sono caratterizzate da alta morbilità, con una forte riduzione della qualità di vita dei pazienti, con occasionali complicazioni che posso portare al decesso.

A parte poche eccezioni, i meccanismi attraverso i quali disfunzioni a carico dell’ENS causano patologie gastrointestinali, rimangono ad oggi poco chiari. Alcuni dati indicano che alterazioni nel sistema nervoso enterico, tra le quali la riduzione del numero di neuroni, la

loro degenerazione morfologica e/o funzionale, sono correlate con perdita di coordinazione motoria intestinale e, di conseguenza, con alterato transito di contenuto intestinale (Goyal et al. 1996, De Giorgio et al. 2004) (Tabella 1.1).

Disease	Neuropathological features	Proposed aetiological factors
Achalasia	Defective inhibitory innervation with or without inflammatory neuropathy (myenteric ganglionitis)	Immune-mediated (including antineuronal antibodies) Neurotropic viruses (e.g. <i>Varicella zoster</i> ; <i>Herpes simplex</i> type 1)
	Complete loss of myenteric ganglia	Genetic factors
Gastroparesis	Degenerative neuropathy Inflammatory neuropathy (myenteric ganglionitis)	Immune-mediated (including antineuronal antibodies) Neurotropic viruses (undetermined) Genetic factors
Congenital hypertrophic pyloric stenosis	Defective inhibitory innervation	Selective alteration of nitrergic neurons
Chronic idiopathic intestinal pseudo-obstruction	Degenerative neuropathy Inflammatory neuropathy (myenteric ganglionitis)	Immune-mediated (including antineuronal antibodies) Neurotropic viruses (undetermined) Genetic factors
Slow transit constipation	Impairment of neurotransmission without evidence of neuronal damage Degenerative neuropathy Inflammatory neuropathy (myenteric ganglionitis)	Immune-mediated Neurotropic viruses (undetermined) Genetic factors
Hirschsprung's disease	Congenital defect in enteric neuron migration and maturation (aganglionosis)	Genetic factors

Tabella 1.1: Caratteristiche neuropatologiche e ipotetici fattori eziologici delle neuropatie enteriche idiopatiche (da Di Nardo et al, 2008)

Un ampio spettro di patologie sono note per essere in grado di colpire l'integrità morfo-funzionale del sistema nervoso enterico, causando disordini digestivi neuropatici. In particolare, alcune condizioni patologiche come il diabete o l'amiloidosi, possono essere responsabili

di alterata motilità intestinale. D'altro canto, le manifestazioni idiopatiche rappresentano ancora una sfida per la clinica, poiché vi è ancora una forte carenza di informazioni sulla loro eziopatogenesi.

Un sempre maggior numero di studi, accumulati negli ultimi anni, indicano che alterazioni a carico della muscolatura liscia intestinale, delle cellule di Cajal (ICC) e delle cellule gliali dell'ENS, possono avere un ruolo importante nella patogenesi delle dismotilità intestinali (Sanders et al. 2006, Sarna et al. 2007, Bassotti et al. 2007, Savidge et al. 2007).

La pseudo-ostruzione intestinale cronica

Il termine pseudo-ostruzione intestinale (CIP) è usato per indicare una sindrome caratterizzata da disfunzione della propulsione intestinale, con un quadro clinico simile a quello di una ostruzione meccanica, in assenza di una reale lesione occludente il lume intestinale.

Le sindromi pseudo-ostruttive possono presentarsi in forma acuta o in forma cronica.

La pseudo-ostruzione acuta, o ileo paralitico, si manifesta nella maggior parte dei pazienti che si sottopone a chirurgia addominale e si risolve spontaneamente nel giro di pochi giorni. Essa può essere causata anche da peritonite, da gravi infezioni (specialmente se indotte da batteri Gram-negativi), carenza di potassio ematico (e gravi carenze di elettroliti in generale), emorragie retroperitoneali, fratture spinali o pelviche ed anche da infarto del miocardio. Il meccanismo attraverso il quale queste differenti condizioni patologiche possono temporaneamente sopprimere la motilità intestinale, non sono ancora del tutto chiariti. Una delle ipotesi è che sia implicata una prolungata attività dei neuroni intrinseci inibitori dovuta all'azione di mediatori neuromorali (Stanghellini et al., 1988).

Il termine pseudo-ostruzione intestinale cronica è stato per la prima volta coniato da Maldonado e colleghi nel 1970, benché fossero già

stati descritti casi di pazienti con ostruzione intestinale in assenza di reali lesioni occludenti il lume fin dalla fine del XIX secolo.

Il trattamento chirurgico, per la CIP, è spesso inefficace, in quanto i sintomi ricompaiono anche in seguito a resezione intestinale.

La CIP può derivare da anomalie sia neurogeniche che miogeniche del controllo della motilità intestinale; nel primo caso la mancata propulsione del contenuto intestinale è dovuta alla perdita della coordinazione delle contrazioni della muscolatura intestinale, nel secondo caso è dovuta ad una riduzione della forza delle contrazioni muscolari (Stanghellini et al., 1988).

Sia la forma miogenica che neurogenica di CIP possono presentarsi come secondarie ad una conosciuta condizione patologica, o come idiopatiche (pseudo-ostruzione intestinale cronica idiopatica, CIIP).

La pseudo-ostruzione intestinale, dal punto di vista clinico, è caratterizzata da ricorrente dolore addominale, distensione e/o gonfiore addominali, nausea e/o vomito, costipazione, e si può associare a perdita di peso e sensazione di sazietà precoce durante i pasti. I sintomi variano, sia nella loro presentazione che nella loro gravità, tra soggetto e soggetto.

Nel lasso di tempo compreso tra due episodi occlusivi i pazienti possono non essere soggetti a sintomi, ma più frequentemente presentano sintomi che possono variare per intensità.

Quando la malattia colpisce principalmente la porzione prossimale dell'intestino, i sintomi predominanti sono nausea, vomito e perdita di peso, mentre quando ad essere colpita è la porzione distale, i sintomi sono dolore addominale diffuso, grave distensione (quanto quella di una gravidanza di 5-9 mesi) e costipazione.

Diarrea e steatorrea possono presentarsi in seguito a colonizzazione batterica del piccolo intestino, la disfagia è presente in una minoranza di casi dei pazienti affetti da CIP (Stanghellini et al., 1988).

La comparsa della CIP è generalmente insidiosa; il primo episodio pseudo-ostruttivo può essere preceduto da un quadro clinico di dispepsia e costipazione durato diversi anni con gravità crescente dei

sintomi. Più raramente le manifestazioni della patologia possono cominciare in seguito ad un intervento chirurgico.

Benché la maggior parte dei pazienti affetti da CIIP sono casi sporadici, non mostrano cioè ricorrenza familiare, esistono forme sindromiche note di CIIP ad ereditarietà autosomica dominante, autosomica recessiva e X-linked.

Alcuni geni e loci sono stati associati a forme sindromiche di pseudo-obstruzione; tra questi vi è il fattore di trascrizione *SOX10*, sul cromosoma 22p12, il gene della DNA polimerasi gamma (*POLG*) sul cromosoma 21q17 ed un locus sul cromosoma 8q23-24.1 (De Giorgio et al. 2004, De Giorgio et al. 2001, Deglincerti et al. 2007). Per quanto riguarda l'ereditarietà X-linked, è stata identificata di recente una delezione di due paia di basi nell'esone 2 del gene *FLNA*, codificante una proteina citoscheletrica, in un soggetto affetto di un pedigree nel quale segrega una forma sindromica di CIIP (Gargiulo et al. 2007).

I casi familiari di CIIP possono essere associati a MNGIE (encefalomiopatia neurogastrointestinale mitocondriale); mutazioni nel gene timidina fosforilasi (*TP* o endothelial cell growth factor-1, *ECGF1*) possono causare MNGIE con CIIP (Hirano et al. 1994, Nishino et al. 2000).

Esofago di Barrett

L'esofago di Barrett (da Norman Barrett, studioso che per primo lo descrisse nel 1950) è una forma di metaplasia in cui il normale epitelio esofageo viene sostituito da epitelio di tipo intestinale. Vi sono significative evidenze del fatto che la formazione dell'esofago di Barrett sia un processo che prevede due passaggi. Il processo metaplastico nella giunzione gastroesofagea, inizia con la conversione della mucosa squamosa esofagea distale in epitelio di tipo cardiace; questo epitelio subirebbe poi una "intestinalizzazione" trasformandosi in epitelio colonnare di tipo gastrico, colonico o simile a quello presente nel piccolo intestino (Peters et al., 2004).

L'eziologia dell'esofago di Barrett non è ancora stata chiarita, ma vi sono varie ipotesi sulla sua origine; oggi è generalmente accettata la teoria che l'esofago di Barrett origini da cellule multipotenti dell'esofago, che possono essere ritrovate nello strato basale dell'epitelio squamoso (Mueller et al., 2004).

Il motivo per il quale si instauri l'esofago di Barrett non è ancora chiarito; nonostante vi siano ipotesi secondo le quali la metaplasia sia una condizione congenita dovuta ad una incompleta epitelizzazione embrionale esofagea, oggi si ritiene più probabile che l'esofago di Barrett sia dovuto ad una condizione di reflusso gastroesofageo cronico (von Rahden et al., 2004); secondo questa teoria l'epitelio intestinale sarebbe più resistente ai danni dovuti al reflusso gastroesofageo, che danneggia ed infiamma il normale epitelio esofageo; la correlazione tra malattia da reflusso ed esofago di Barrett, rimane ancora non completamente chiara, infatti più dell'80% dei pazienti con malattia da reflusso gastroesofageo non sviluppano l'esofago di Barrett (Mueller et al., 2004).

La classificazione dell'esofago di Barrett è effettuata in base all'estensione della zona di esofago interessata da metaplasia, si parla di esofago di Barrett a segmento lungo (sopra i 3-5 cm), a segmento breve (2-3 cm) mentre sotto i 2 cm si parla di segmento ultra-corto (in alcuni casi è anche considerato una variante normale) (Mueller et al., 2004).

La complicazione più importante dell'esofago di Barrett consiste nello sviluppo di adenocarcinoma esofageo; essere portatori di esofago di Barrett corrisponde ad un aumento del rischio di sviluppare adenocarcinoma del 30-40% , per i pazienti con esofago di Barrett a segmento lungo (Ramel, 2003).

La patologia si evolve da metaplasia a displasia a basso grado e poi ad alto grado, per poi arrivare allo sviluppo del carcinoma; questo processo è caratterizzato da un aumento del tasso proliferativo cellulare, perdita del differenziamento, espressione di fattori di crescita e molecole di adesione, alterazione dell'apoptosi; ciò è favorito

dall'accumulo di mutazioni nelle cellule della lesione (Peters et al., 2004).

Sono stati descritti anche casi di esofago di Barrett congenito, o dovuto a forme congenite familiari di malattia da reflusso gastroesofageo; ciò supporta la teoria che l'esofago di Barrett sia conseguenza di reflusso gastroesofageo, ma rimane una componente congenita non ancora chiarita.

La Sindrome di Mungan

Descritta per la prima volta nel 2003 da Mungan e colleghi, la sindrome di Mungan (MGS, MIM:611376) è caratterizzata da pseudo-obstruzione intestinale cronica idiopatica, esofago di Barrett ed anomalie cardiache, più una serie di segni e sintomi di minore rilevanza (tabella 1.2).

La patologia, descritta in un'unica famiglia di origine turca, ha ereditarietà di tipo autosomico recessivo. In particolare, la famiglia è caratterizzata da elevata consanguineità, con i membri affetti che sono figli di cugini di primo grado (si veda materiali e metodi per una descrizione più dettagliata della famiglia).

<i>Family member</i>	<i>Gastrointestinal features</i>	<i>Cardiac features</i>	<i>Other relevant clinical features</i>	<i>Outcome</i>
IV-1, ^a F	Diarrhoea, recurrent abd. distension/ pain	Unknown	Unknown	Died at 14 years
IV-2, ^a F	Diarrhoea, recurrent abd. distension/pain	Unknown	Unknown	Died at 15 years
IV-9, M				
(proband)	CIIP, megaduodenum, long Barrett	Multiple valve/septal defects	Growth retardation	Alive at 28 years
IV-10, M	CIIP, megaduodenum, long Barrett	Multiple valve/septal defects	Growth retardation	Died at 29 years
IV-11, F	CIIP, megaduodenum, long Barrett	Multiple valve/septal defects	Growth retardation	Alive at 30 years
IV-14 ^a , M	CIIP	Unknown	Unknown	Died at 19 years
IV-15 ^a , F	CIIP	Unknown	Unknown	Died at 15 years
V-4, F	One acute abd. pain attack	None	Growth retardation	Alive at 4 years
V-5, M	No obvious digestive symptoms	None	Hyperkinetic behaviour	Alive at 18 years

Note: abd., abdominal; F, female; M, male; CIIP, chronic idiopathic intestinal pseudo-obstruction; yrs., years.

^aClinical details in these members were obtained indirectly from clinical records.

Tabella 1.2: da Deglincerti et al. 2007; principali caratteristiche cliniche dei soggetti analizzati della famiglia turca.

E' stato eseguito un esame istopatologico delle biopsie intestinali a pieno spessore dei soggetti affetti (Deglincerti et al, 2007) che ha mostrato anomalie rispetto ai controlli (tessuto intestinale ottenuto dai margini di resezione intestinale da soggetti affetti da cancro del colon retto non ostruttivo e privo di complicazioni). Tali anomalie riguardavano la innervazione intrinseca, che mostrava riduzione neuronale insieme a severa riduzione neurochimica di NOS (nitric oxide synthase), VIP (vasoactive intestinal pelyptide), sostanza P/tachichinina, neuropeptide Y, *calcitonin gene-related peptide* e galanina, e risultavano quindi suggestive di disorganizzazione dei neuroni sia eccitatori che inibitori. L'espressione di Bcl-2 (proteina coinvolta nei meccanismi regolatori della sopravvivenza cellulare) era notevolmente ridotta nei neuroni enterici degli individui affetti rispetto ai controlli. L' esame della KIT immunoreactivity, marker per le cellule interstiziali di Cajal (ICC), ha rivelato un ridotto ICC network attorno al plesso mienterico e all'interno dello strato di muscolatura liscia; infine, l'immunoreattività dell' α -actina, comparata a quella dei controlli, era diminuita sia nello strato circolare che in quello longitudinale, suggerendo un coinvolgimento anche della muscolatura liscia (figura 1.7).

In conclusione le indagini di immunoistochimica condotte sulle biopsie a tutto spessore della parete intestinale indicano la presenza di una neuromiopia enterica diffusa.

La presenza di esofago di Barrett lungo è stata confermata dall'evidenza di goblet cells con metaplasia intestinale in biopsie ottenute fino a 18 cm dalla giunzione esofago-gastrica.

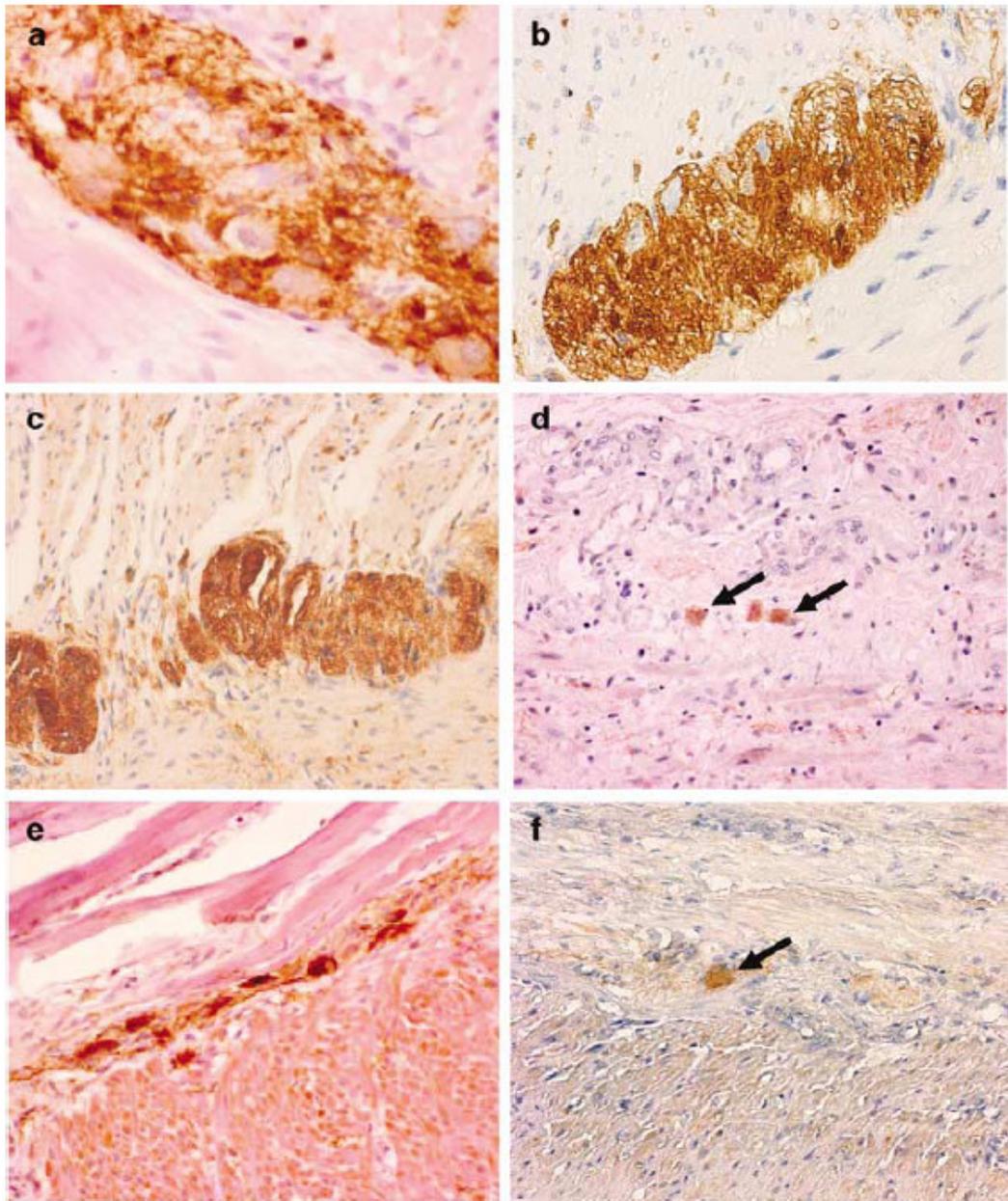


Figura 1.7: Esempi rappresentativi delle analisi immunoistologiche. A destra (b, d, f) il paziente, a sinistra (a, c, e) il controllo. **a e b:** proteina marker della glia enterica S-100; **c e d:** il marker neuronale NSE; **e e f:** immunoreattività di nNOS.

Il nostro gruppo di ricerca ha avuto modo di mappare il gene malattia responsabile di questa sindrome. Tramite homozygosity mapping è stato infatti possibile stabilire una chiara concatenazione tra la patologia ed un locus sul cromosoma 8q23-24.1 (Deglincerti et al, 2007; vedere risultati).

L'obiettivo principale del progetto era quello di giungere all'identificazione del gene responsabile della patologia, analizzando i geni presenti nella regione. Infatti, l'identificazione del gene responsabile aprirebbe interessanti e promettenti possibilità di ricerca. Tramite studi di tipo funzionale, infatti, sarebbe possibile acquisire nuove fondamentali conoscenze che consentano di spiegare i meccanismi genetici e molecolari alla base della pseudo-ostruzione intestinale cronica idiopatica. Lo studio del gene responsabile, inoltre, sarebbe di grande interesse per quanto concerne l'esofago di Barrett, infatti, i meccanismi che stanno alla base di questa metaplasia sono ancora da chiarire, e nuove conoscenze in questo ambito potrebbero rappresentare la chiave per comprendere la patogenesi di questa diffusa condizione patologica, nota essere l'unico precursore conosciuto dell'adenocarcinoma esofageo (Peters et al. 2004).

1.3.2 TROMBOCITOPENIA AUTOSOMICA DOMINANTE THC2

Le Trombocitopenie Ereditarie

Le trombocitopenie (o piastrinopenie) ereditarie sono un eterogeneo gruppo di patologie caratterizzato da ridotto numero di piastrine ed una tendenza al sanguinamento da molto moderata fino a molto grave (Dowton et al. 1985). Sono stati proposti differenti criteri al fine di classificare le diverse forme, come il volume delle piastrine, il numero e la morfologia dei megacariociti o il tipo di ereditarietà. Sono state descritte forme a trasmissione autosomica dominante, autosomica recessiva ed X-linked.

Oltre alle sindromi cliniche nelle quali la piastrinopenia si presenta in associazione ad altri segni e sintomi, come l'anemia di Fanconi, la sindrome di Wiskott-Aldrich e la sindrome TAR (Derry et al. 1994; Villa et al. 1995; Savoia et al. 1996; Strippoli et al. 1998), in molte famiglie la trombocitopenia è l'unico sintomo. In questi casi, alcune trombocitopenie sono caratterizzate da un ridotto numero di megacariociti nel midollo osseo con coinvolgimento di un unico o di più lineage, mentre altre hanno un numero di megacariociti nella norma. Sulla base del volume medio delle piastrine, le forme non sindromiche di piastrinopenia, possono essere ulteriormente classificate in macro, micro o normo-trombocitopenia (Najean and Lecompte 1995; Bellucci 1997).

Benché ritenute rare, il numero totale di pazienti potrebbe essere considerevole, poiché più di 1000 famiglie con diagnosi specifica sono state descritte in letteratura.

L'introduzione dei contatori automatizzati di cellule ha reso la conta delle piastrine un test routinario, rivelando il fatto che la trombocitopenia è una condizione molto più comune di quel che si pensasse, benché non sia mai stato pubblicato un dato epidemiologico.

La diagnosi di piastrinopenia rimane tutt'oggi problematica; sono infatti molteplici le cause di piastrinopenia, sia ereditaria che acquisita (malattie autoimmuni, aumentato consumo di piastrine, splenomegalia, ipo/aplasia midollare...) e benché il quadro clinico del paziente possa semplificare la diagnosi, spesso questa rimane molto difficoltosa, se non impossibile.

Uno dei problemi più frequenti è quello della diagnosi differenziale tra Porpora Trombocitopenica Idiopatica (ITP) e le trombocitopenie ereditarie, poiché la prima è diagnosticata solo in seguito all'esclusione delle forme ereditarie (Geddis and Balduini, 2007).

Se si considera il fatto che i geni responsabili per le forme ereditarie (codificanti per glicoproteine di membrana, componenti del citoscheletro ed elementi delle vie di segnalazione intracellulare) identificati coprono circa il 60% dei pazienti con piastrinopenia sindromica e non (Nurden and Nurden, 2006), il 40 % dei pazienti rischia di incorrere in errori di diagnosi. Di conseguenza a diversi pazienti con trombocitopenia ereditaria viene diagnosticata una ITP, e ricevono terapie non solo inutili, ma anche dannose. Alcuni studi hanno mostrato come la rivalutazione di due serie di casi di pazienti con diagnosi di ITP, ha messo in evidenza che il 7% di questi era in realtà affetto da trombocitopenia ereditaria (Gohda et al, 2007; Bader–Meunier et al, 2003), e che il 15% di una terza serie di casi di pazienti con piastrinopenie ereditarie aveva subito splenectomia a causa dell'errata diagnosi di ITP (Noris et al, 2004).

Il fatto che il 40% dei pazienti non rientra nei criteri diagnostici di nessuna malattia conosciuta, indica che un numero relativamente alto di nuove forme deve ancora essere descritto.

Recentemente sono stati descritti nuovi geni responsabili di trombocitopenia: *CYCS* (Morison et al, 2008) per la trombocitopenia 4 (THC4, OMIM 612004), *ITGB3* (Kunishima et al, 2009) per la macrotrombocitopenia con GPIIb-IIIa gain of function, *TUBB1* (Ghevaert et al, 2008) per una macrotrombocitopenia autosomica dominante (OMIM, 613112). Tuttavia, analisi mutazionali in ampie coorti di pazienti italiani, ancora senza diagnosi, hanno individuato la

presenza di mutazioni in questi geni in solo due famiglie (Gresele et al, 2009; Savoia et al, 2009), indicando che probabilmente mutazioni in questi geni sono molto rare, almeno in Italia.

Trombocitopenia THC2

Nel 1999 Savoia e colleghi descrissero un' ampia famiglia italiana nelle quale segregava trombocitopenia con ereditarietà autosomica dominante, caratterizzata da moderata piastrinopenia con sintomi minimi, normale morfologia delle piastrine e megacariociti midollari nella norma (figura 1.8). In seguito ad analisi di linkage estesa a tutto il genoma, il gene fu mappato sul cromosoma 10p11.2-12 (Savoia et al, 1999).

La regione identificata misurava circa 6 cM ed era compresa tra i marcatori D10S586 e il D10S1639. Il gruppo ottenne un LOD score massimo di 8,12 a $\theta=0$ per il marcatore D10S588 (figura 1.9).

Nel 2000 un altro gruppo, in maniera indipendente, mappò il gene responsabile di piastrinopenia autosomica dominante nella stessa regione sul cromosoma 10, in seguito all'analisi di una famiglia americana (Figura 1.10; Drachman et al, 2000). Lo stesso gruppo nel 2003 identificò una mutazione missenso nel gene *MASTL*, (microtubule associated serine/threonine kinase-like) posizionato all'interno del locus THC2 (Gandhi et al, 2003). Il potenziale ruolo di questo gene nella patologia è supportato da un recente lavoro, che mette in evidenza la correlazione tra knockdown transiente del gene e riduzione dei trombociti circolanti in zebrafish (Johnson et al, 2009).

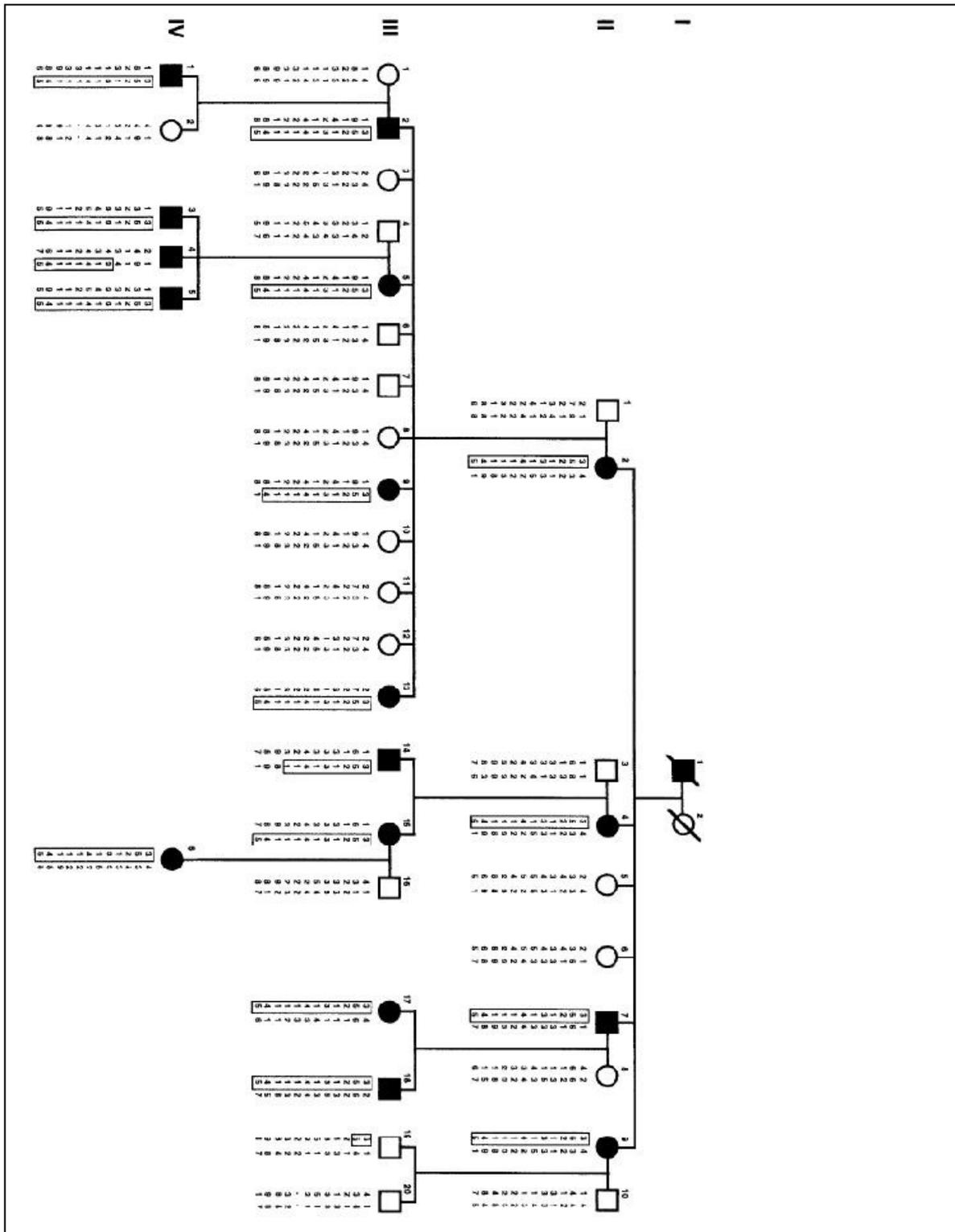


Figura 1.8: La famiglia studiata da Savoia et al, 1999

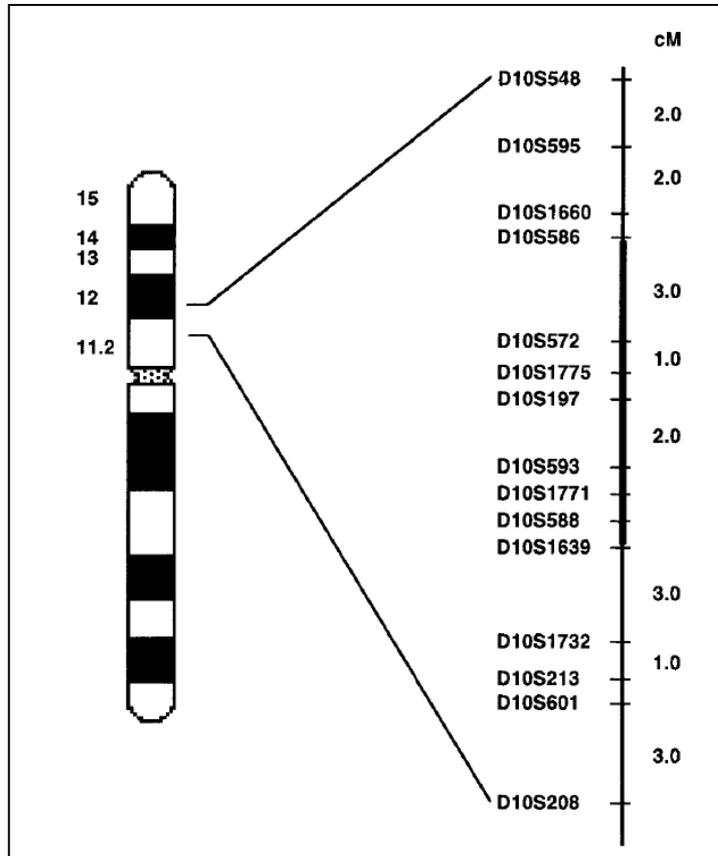


Figura 1.9: Il locus THC2 (da Savoia et al, 1999)

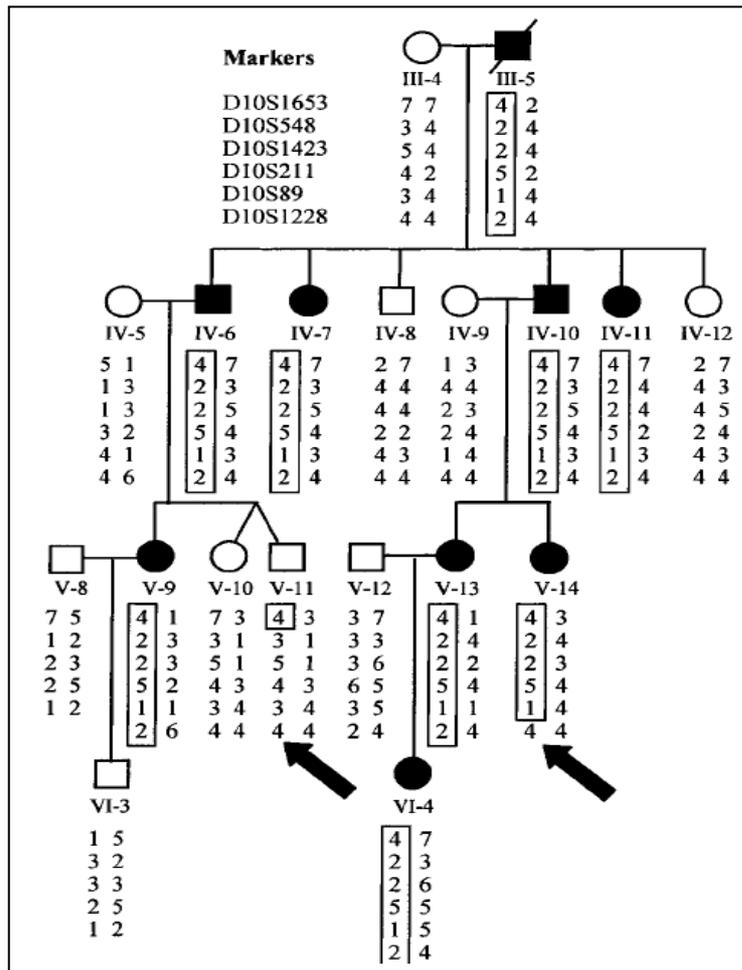


Figura 1.10: La famiglia studiata da Drachman e colleghi, 2000.

2. MATERIALI E METODI

2.1 LE FAMIGLIE ANALIZZATE

2.1.1 SINDROME DI MUNGAN: FAMIGLIA TKS

La famiglia è giunta all'osservazione clinica poiché vi erano tre fratelli affetti da pseudo-ostruzione. Questi soggetti risultavano essere figli di cugini di primo grado (figura 2.1) (Mungan et al. 2003; Deglincerti et al, 2007).

Il paziente 1 (IV:9 in figura 2.1), il caso indice, un maschio di 26 anni allora, è morto nell'estate del 2005 in seguito ad un intervento chirurgico per una recidiva sub-occlusiva. Egli lamentava dolore addominale ricorrente, distensione addominale e diarrea. È stato più volte ricoverato per ripetuti episodi di pseudo-ostruzione intestinale fin dall'età di 11 anni. Il paziente presentava inoltre ritardo di crescita (47 kg, 158 cm) e lentiggini al volto. Oltre all'interessamento gastrointestinale, il paziente mostrava esofago di Barrett a segmento lungo (18 cm), anomalie cardiache e altre condizioni patologiche in associazione.

Il paziente 2 (IV:10 in figura 2.1), un maschio di 28 anni fratello del paziente 1, ha lamentato fin dall'infanzia dolore addominale e diarrea. Oltre a un ritardo di crescita (51 kg, 155 cm), l'esame fisico ha evidenziato calvizie, ptosi bilaterale e lentiggini al volto. Come il fratello, presentava esofago di Barrett a segmento lungo (18 cm) e anomalie cardiache.

Il paziente 3 (IV:11 in figura 2.1), una femmina di 30 anni, sorella dei due pazienti precedentemente descritti, è risultata essere interessata da attacchi ricorrenti di pseudo-ostruzione intestinale. Anche lei presentava ritardo di crescita (40 kg, 145 cm), anomalie cardiache ed esofago di Barrett a segmento lungo (vedere la tabella 2.1 per maggiori dettagli).

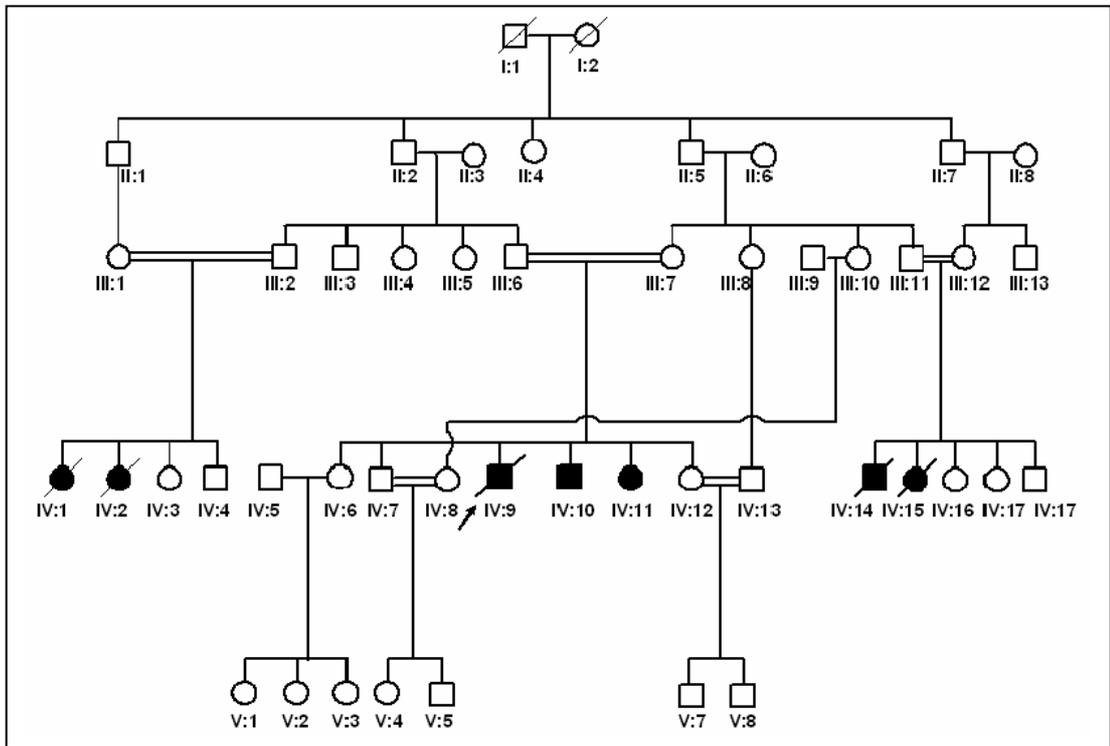


Figura 2.1: albero genealogico della famiglia turca con membri affetti da pseudo-ostruzione cronica intestinale; la freccia indica il probando

	Paziente 1	Paziente 2	Paziente 3
Endoscopia gastro-esofagea	Esofago di Barrett. Ipomotilità e svuotamento gastrico ritardato	Esofago di Barrett. Ipomotilità e svuotamento gastrico ritardato	Esofago di Barrett. Ipomotilità e svuotamento gastrico ritardato
Manometria esofagea	Aperistalsi. Pressione dello sfintere esofageo inferiore di 0 mm Hg	Aperistalsi. Pressione dello sfintere esofageo inferiore di 0 mm Hg	Aperistalsi e simultanee contrazioni. Pressione dello sfintere esofageo inferiore di 0 mm Hg
Piccolo intestino	Megaduodeno. Svuotamento ritardato	Megaduodeno. Svuotamento ritardato	Megaduodeno. Svuotamento ritardato
Anomalie cardiache	Stenosi prossimale dell'arteria polmonare. Rigurgito della valvola polmonare. Rigurgito della tricuspide.	Difetto del setto ventricolare	Stenosi della valvola polmonare
Condizioni addizionali	Otosclerosi. Glaucoma. Epilessia.	Ptosi.	-

Tabella 2.1: quadro clinico dei tre pazienti affetti della famiglia turca

L'analisi del pedigree rivela altri casi di matrimoni tra consanguinei legati a casi clinici di disordine digestivo.

Due cugini dell'index case (IV:14 e IV:15) sono morti rispettivamente a 19 e 15 anni, a causa di disordini digestivi compatibili con la diagnosi di pseudo-ostruzione; la cugina IV:15 aveva evidenze cliniche e radiologiche di CIIP, ritardo della crescita, malnutrizione, megaduodeno, dismotilità esofagea (aperistalsi e assenza di tono dello sfintere esofageo inferiore), ipoplasia renale, reflusso vescico-uretrale, asciti, ed epatite granulomatosa. E' morta al di fuori dell'ospedale per cause non note.

Il fratello IV:14 è morto di ciò che è stato descritto come “una malattia simile” (riferito alla sorella), ma non sono disponibili dati clinici; i fratelli e sorelle di questi pazienti sono tuttora in vita e godono di buona salute.

Altri due pazienti (IV:1 e IV:2), anch’essi figli di cugini di primo grado, sono morti rispettivamente all’età di 14 e 15 anni, prima di poter essere studiati dal punto di vista clinico. I parenti hanno riferito che essi erano affetti da distensione addominale, diarrea e ricorrente dolore addominale; i loro fratelli e sorelle sono vivi e godono di buona salute (Deglincerti et al, 2007).

L’albero genealogico della famiglia TKS (Figura 2.1) mostra chiaramente una trasmissione della malattia secondo uno schema di ereditarietà mendeliana autosomica recessiva. Si vede infatti come gli individui affetti appartengono ad entrambi i sessi, e sono figli di genitori non affetti (che saranno quindi portatori asintomatici).

2.1.2 PIASTRINOPENIA AUTOSOMICA DOMINANTE THC2

Sono giunti presso il nostro laboratorio i DNA dei componenti di diverse famiglie nelle quali segregava trombocitopenia autosomica dominante, per tutte queste famiglie era stato escluso il coinvolgimento di geni già noti, ed il nostro intento era quello di verificare l'eventuale segregazione del locus THC2.

Le famiglie sono tutte caratterizzate da moderata piastrinopenia con sintomi minimi, normale morfologia delle piastrine e megacariociti midollari nella norma. Nelle figure dalla 2.2 alla 2.6 sono rappresentati gli alberi genealogici.

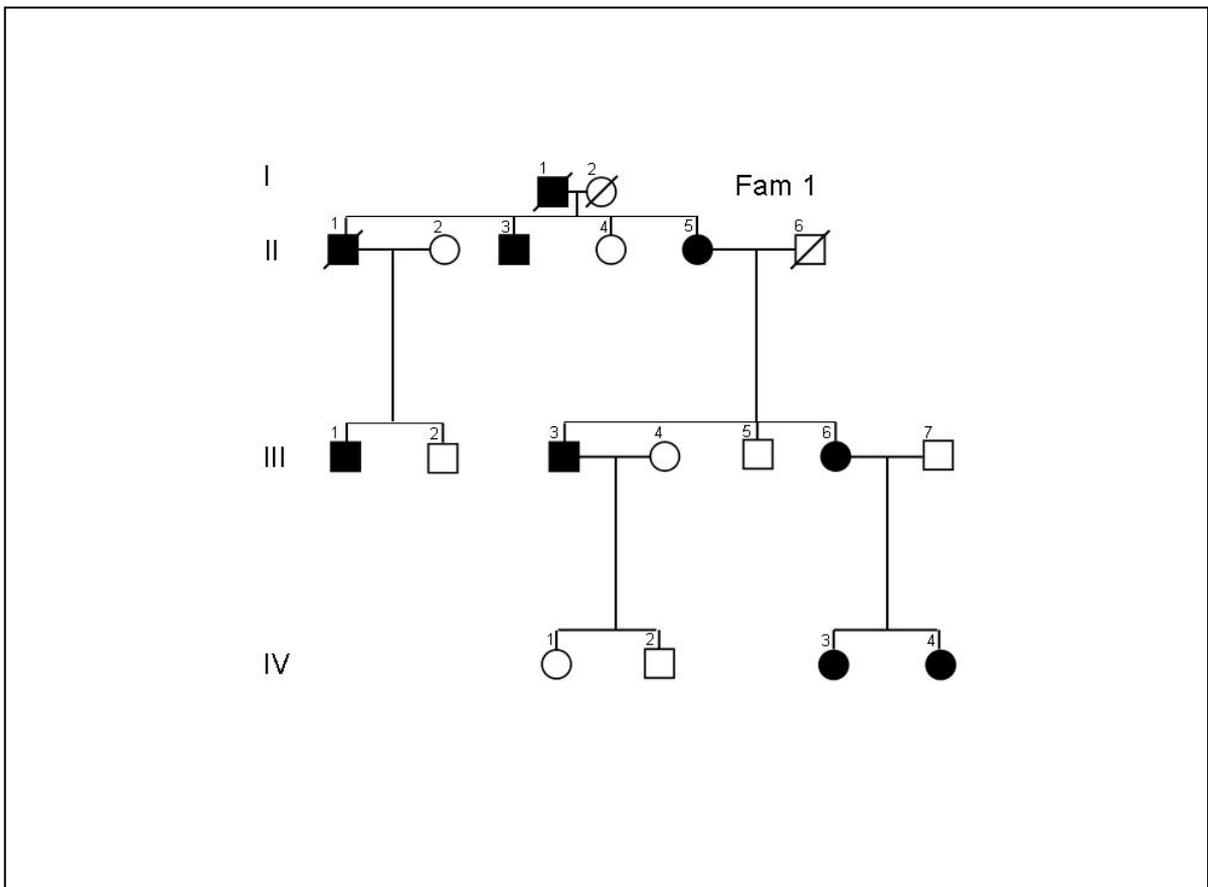


Figura 2.2: famiglia 1

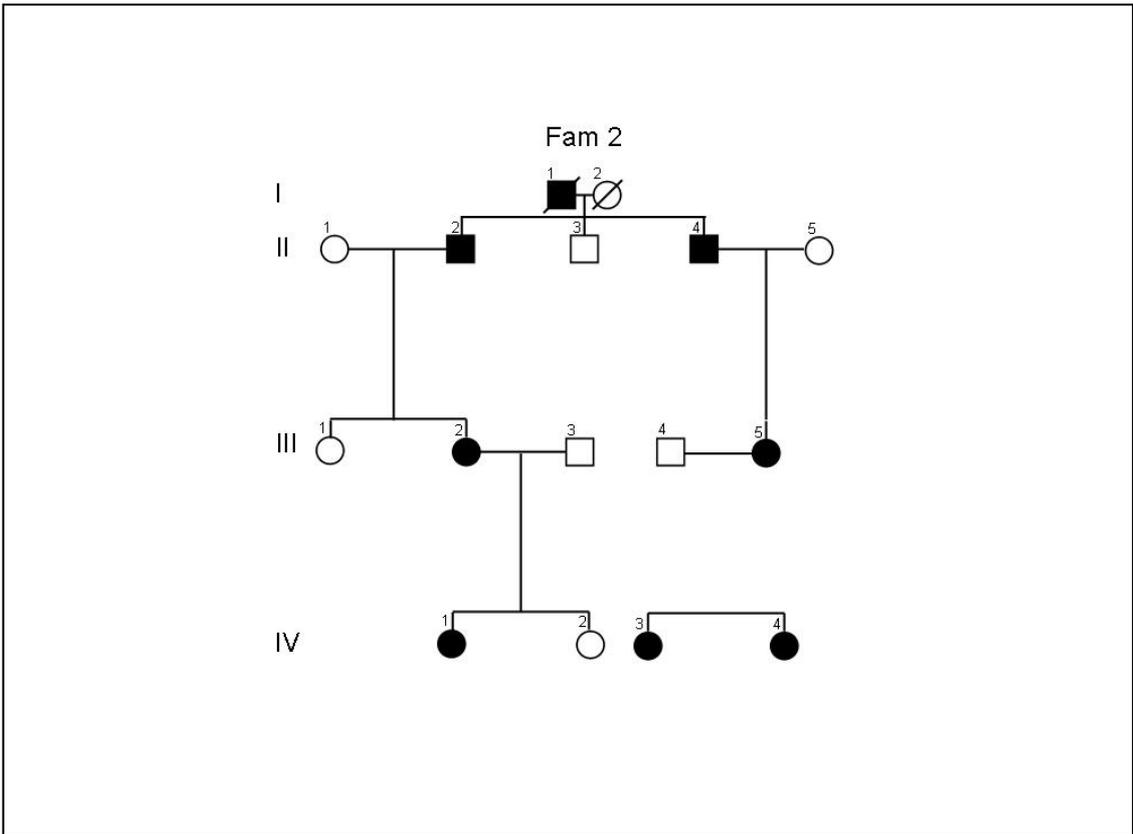


Figura 2.3: famiglia 2

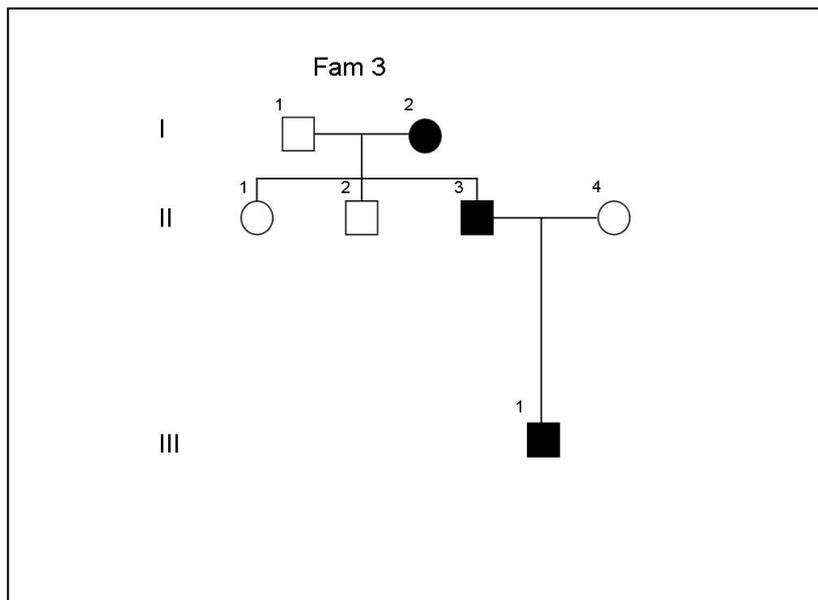


Figura 2.4: Famiglia 3

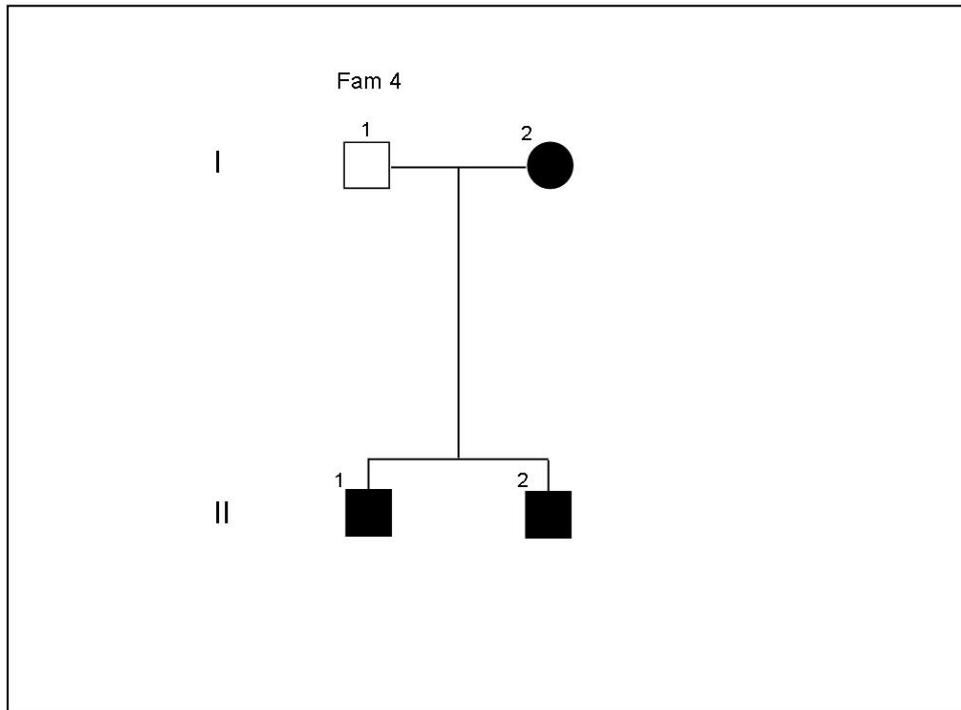


Figura 2.5: Famiglia 4

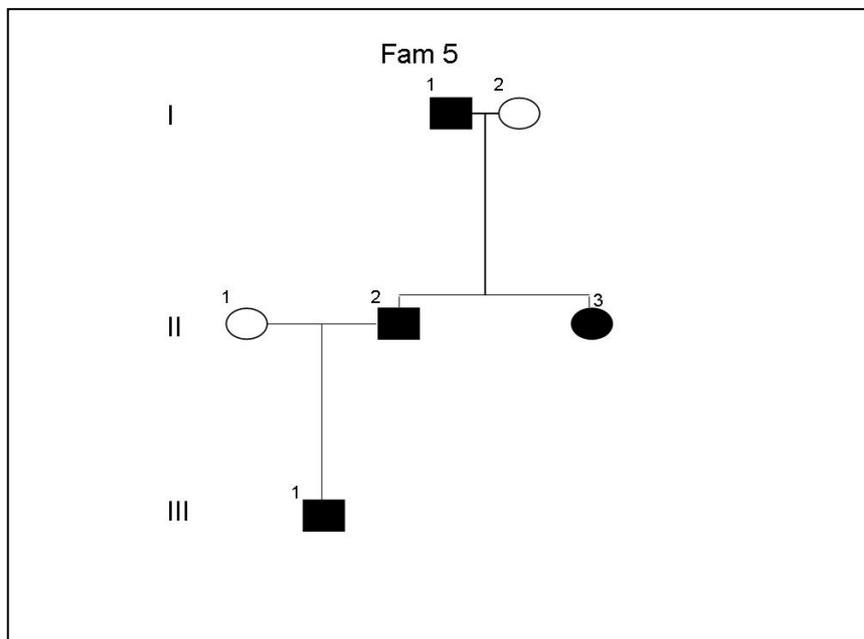


Figura 2.6: Famiglia 5

2.2 METODI DI LABORATORIO

Estrazione Del DNA

Il DNA genomico è stato estratto a partire da campioni di sangue congelato (una provetta di prelievo di sangue venoso contenente EDTA come anti-coagulante) tramite una lisi cellulare iniziale seguita da digestione delle proteine e dalla precipitazione del DNA nelle opportune condizioni di sali e alcool. Il metodo utilizzato per l'estrazione è stato il "Salting Out", e si compone di due giornate di lavoro.

Primo giorno

I campioni di sangue vengono scongelati a temperatura ambiente e si portano a volume (45 mL) con il LYSIS BUFFER (Tabella 3.1).

Si agitano delicatamente per due minuti e si pongono in ghiaccio per 15 minuti.

Si centrifugano a 3500 rpm a 4°C per 10 minuti.

Si elimina il surnatante, si porta a volume (45 mL) con FISIO BUFFER (Tabella 2.2) e si risospende energicamente.

Si centrifuga a 3000 rpm a 4°C per 10 minuti.

Le operazioni degli ultimi due punti (eliminare il surnatante, portare a volume con FISIO BUFFER, risospendere e centrifugare) vanno ripetute per altre due volte.

Si elimina il surnatante e si risospende il pellet in 3 mL di BUFFER A (Tabella 3.1). Si aggiungono 200 µL di SDS 10% e 50 µL di Proteinasi K e si lascia a incubare overnight a 37°C in bagno termostato.

Secondo giorno

Si aggiunge 1 mL di NaCl saturo, si inverte per 2 minuti e si centrifuga a 3000 rpm a 4°C per 10 minuti.

Si trasferisce il surnatante, prestando attenzione a non aspirare il fondo di SDS, in una Falcon da 15 mL.

Si aggiunge un ugual volume di cloroformio/alcool isoamilico (24/1), si inverte delicatamente per due minuti e si centrifuga a 3000 rpm per 10 minuti a temperatura ambiente.

Si preleva la fase acquosa e la si trasferisce in una Falcon da 15 mL. Si aggiunge il doppio del volume di EtOH 100% e si inverte delicatamente fino a rendere visibile la “medusa” di DNA.

Con un puntale o con l’uncino di una pasteur di vetro si raccoglie il DNA e lo si lava in una eppendorf contenente 1 mL di EtOH 70%. Si lascia per qualche minuto il puntale/la pasteur all’aria per farla asciugare. Si ripete il lavaggio in una eppendorf contenente 1 mL di EtOH 100%. Si lascia asciugare all’aria e si risospende il DNA in 500 µL di acqua sterile.

Il campione viene poi conservato a 4°C.

LYSIS BUFFER	1 Lt	100 mL
Saccarosio 0.32 M	109.5 gr	10.95 gr
Tris HCl pH 7.5 10 mM	10 mL (da 1M)	1 mL (da 1M)
MgCl 5mM	5 mL (da 1M)	0.5 mL (da 1M)
Triton X 100 1%	10 mL	1 mL
FISIO BUFFER	1 Lt	250 mL
NaCl 0.075 M	15 mL (da 5M)	3.75 mL (da 5M)
EDTA pH 8 0.025 M	50 mL (da 0.5M)	12.5 mL (da 0.5M)
BUFFER A	1 Lt	40 mL
Tris-HCl pH 8 10 mM	10 mL (da 1M)	400 µL (da 1M)
NaCl 400 mM	80 mL (da 5M)	3.2 mL (da 5M)
EDTA 2 mM	4 mL (da 0.5M)	160 µL (da 0.5M)

Tabella 2.2. Reagenti per l'estrazione del DNA con il metodo “Salting Out”

Valutazione della quantità di DNA estratto

La quantità di DNA estratto viene stimata tramite misurazione dell'assorbanza a 260 nm con uno spettrofotometro a luce UV. Si aggiunge una aliquota di 5 µL di DNA estratto a 995 µL di acqua e si registrano i valori di assorbanza. Generalmente la quantità di DNA genomico ottenuto da 1.5 mL di sangue intero è compresa tra 15 µg e 75 µg. Come conferma (o come alternativa) della lettura allo spettrofotometro, si valuta la quantità di DNA tramite corsa elettroforetica su gel di agarosio. Si allestisce un gel di agarosio alla concentrazione finale dell'1% in tampone TBE 1X (Tris 89mM, Acido Borico 89mM, EDTA 2mM) contenente 0.1 µg/mL di Bromuro di Etidio. Si caricano aliquote di 1-5 µL del DNA estratto previa miscela con 2 µL di gel-loading dye 6X (Blu di Bromofenolo 2.5%, Xylene Cianolo 2.5%, Glicerolo 25%) e acqua sterile per portare a volume finale di 10 µL. Come controlli vengono caricate aliquote di DNA a concentrazione nota. Si procede con la corsa elettroforetica a voltaggio costante (80-100 V).

Al termine della corsa il gel viene visualizzato su transilluminatore e fotografato tramite CCD camera, e si valuta l'intensità delle bande di DNA estratto rispetto a quella delle bande di DNA a concentrazione nota.

Genotipizzazione

La genotipizzazione è stata effettuata mediante una tecnica semi-automatizzata basata sulla fluorescenza, usando un sequenziatore modello 3730 (Applied Biosystems, Tokyo, Japan) corredato di specifici programmi computazionali (GeneMapper) per facilitare la ricostruzione degli aplotipi. Il mappaggio genetico su larga scala ha sfruttato i 28 pannelli di primers ABI PRISM Linkage Mapping Set Version 2 (Applied Biosystems), set che comprende circa 400

marcatori polimorfici che definiscono una mappa umana della risoluzione di 10 cM con una eterozigosità media di 0.81. Ogni pannello contiene dalle 10 alle 20 coppie di primer che generano prodotti di PCR che possono essere combinati insieme e rilevati su un singolo capillare. Ogni primer forward è marcato con un fluorocromo tra 6-FAM (blu), VIC (verde) e NED (giallo), così da poter distinguere alleli sovrapposti con la diversa marcatura tramite un metodo di rilevazione multicolore, mentre il reverse non è marcato. Per ogni coppia, i primer forward e reverse sono combinati nello stesso tubo, a una concentrazione finale di 5 μ M ciascuno (10 μ M totale) in buffer TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA). I loci sono stati scelti a partire dalle mappe di linkage del Genethon (Weissenbach, 1992; Gyapay, 1994). Per facilitare la genotipizzazione, la mix per la PCR è identica per tutti i marcatori, così come sono uguali le condizioni di amplificazione.

Le reazioni sono fatte in un volume finale di 7.5 μ L.

Per ogni campione da amplificare, si usano:

Buffer 10x 0.75 μ L

dNTP 10x 0.75 μ L

MgCl₂ 2mM 0.6 μ L

Taq 0.03 μ L

DNA 20 ng/ μ L 2.0 μ L

Primer 0.3 μ L

H₂O a volume di 7,5 μ L

I cicli di PCR sono i seguenti:

Numero ripetizioni	Condizioni del ciclo
1	Denaturazione iniziale a 95°C per 12'
	Denaturazione a 94°C per 25''
10	Ibridazione a 55°C per 20''
	Estensione a 72°C per 30''
	Denaturazione a 90°C per 25''
25	Ibridazione a 55°C per 20''
	Estensione a 72°C per 30''
1	Estensione finale a 72°C per 20'

I prodotti di PCR dello stesso pannello per un singolo campione vengono poi combinati insieme in quantità di 2 µL ogni prodotto marcato con 6-FAM o VIC e 4 µL ogni prodotto marcato con NED, eventualmente aggiungendo acqua per arrivare a un volume finale di 40 µL per rendere i prodotti diluiti rispettivamente di almeno 1:20, 1:20 e 1:10. 1 µL di questo pool viene risospeso in una mix di 9.7 µL di formamide e 0.3 µL di size-standard (LIZ, arancione) e viene caricato sul sequenziatore automatico a capillari previa denaturazione (2 minuti a 96°C).

L'analisi delle dimensioni dei frammenti di DNA viene fatta in prima battuta usando il software GeneMapper, per essere poi rivista manualmente dall'operatore. Il set finale di alleli individuati per ciascun marcatore viene quindi esportato per ulteriori analisi.

Ridefinizione di regioni critiche e Studio di segregazione del locus THC2

Quando lo studio lo studio tramite marcatori estesi su tutto il genoma ha individuato una regione candidata per le famiglie analizzate si può procedere cercando di definire le estremità dell'intervallo critico entro cui cercare il gene candidato. A tale scopo si aggiungono una serie di marcatori nella regione per mappare eventi di ricombinazione negli individui della famiglia, che definiscano il limite centromerico e telomerico della regione critica; l'utilizzo di marcatori *ad hoc* è utile anche nel caso in cui si voglia investigare la segregazione di un determinato locus cromosomico in una famiglia.

Per la scelta dei marcatori ci si può basare sulle mappe del Genethon oppure i primer vengono progettati a partire dalla sequenza genomica della regione critica fornita dalle banche dati pubbliche come l'NCBI, intorno a un sito che presenti ripetizioni bi- o trinucleotidiche; l'ausilio di software come Tandem Repeat Finder (<http://tandem.bu.edu/trf/trf.html>; G. Benson, 1999) può semplificare la ricerca dei repeats.

Per fare ciò, è buona norma tenere presenti alcuni punti; i primer (Innis e Gelfand, 1991):

- devono essere lunghi almeno 18-20 basi per assicurare la specificità di ibridazione, e non devono generalmente superare le 25-28 basi;
- devono contenere una quantità di C e G pari a circa la metà delle basi per avere una T_m (temperatura di melting) stabile;
- devono terminare all'estremità in 3' con una C o una G così da migliorare la stabilità di ibridazione e aumentare l'efficienza di priming;
- non devono essere complementari tra loro e non devono formare strutture secondarie quali forcine;
- non devono riconoscere sequenze ripetute sul genoma, per evitare amplificazioni aspecifiche;
- dovrebbero possibilmente avere una T_m di 55°C-60°C.

la T_m può essere calcolata in vari modi, si riportano di seguito i più usati:

$T_m = 69.1 + (0.41 * \%GC) - (650/n)$, dove n è la lunghezza del primer

$$T_m = 4^{\circ}\text{C} * (G+C) + 2^{\circ}\text{C} * (A+T).$$

Il primer forward sarà progettato con la stessa sequenza del filamento di DNA genomico, il reverse avrà la sequenza complementare al genomico. Con il programma BLAST è poi possibile verificare la specificità del primer ed eventuali amplificazioni aspecifiche che potrebbe innescare.

Ciascuna nuova coppia di primer dovrebbe poi venire marcata con un fluorocromo, in modo da poter rilevare il prodotto di PCR al sequenziatore.

È stata sviluppata una metodologia per contenere i costi dell'analisi; grazie a tale tecnica, chiamata UFP (che sta per Universal Fluorescence Primer), infatti, non è necessario marcare con un fluorocromo ogni singola coppia di primer. Ad uno dei due primer, il forward per semplicità, viene aggiunta all'estremo in 5' una coda universale di 12 basi che non riconosce sequenze sul genoma ma la cui sequenza è uguale alla porzione in 3' di un oligonucleotide di 18 basi marcato tramite fluorocromo (UFP-18). Si effettuano due reazioni di PCR (figura 2.7), la prima utilizzando una miscela di oligonucleotidi con primer forward modificato e il primer reverse a concentrazione finale equimolare di 0,5 mM.

Buffer 10x 0.75 μL

dNTP 10x 0.75 μL

MgCl₂ Variabile

Taq 0.05 μL

DNA 20 ng/ μL 2.0 μL

Primer 5 ng/ μL 0.75 μL

H₂O a volume finale di 7,5

Questo prodotto di PCR viene quindi purificato in piastre per la purificazione Millipore© da 96 pozzetti. Il prodotto di PCR purificato è utilizzato successivamente come stampo per una successiva PCR, in cui la miscela di oligonucleotidi è composta dal primer UFP fluorescente e dal primer reverse in concentrazione finale equimolare sempre di 0,5 mM.

Buffer 10x 0.75 µL

dNTP 10x 0.75 µL

MgCl₂ variabile

Taq 0.05 µL

Purificato a volume finale di 7,5

Primer 5 ng/ µL 0.75 µL

La concentrazione dei sali non è standardizzata e deve essere determinata tramite prove di PCR che sono controllate con corsa elettroforetica su gel di agarosio. Tali prove consistono in reazioni di PCR tra loro differenti per la concentrazione di MgCl₂ (con il magnesio concentrato 1.0 mM, 1.5 mM e 2.0 mM) ed eventualmente per l'aggiunta di particolari additivi.

Per esempio si può incorporare fino al 10% di dimetil-solfossido, dimetil-formamide, urea o formamide per diminuire la T_m; il dimetil-solfossido è spesso utile nell'amplificare sequenze lunghe (anche se può diminuire l'attività della Taq); la formamide sembra migliorare la specificità della PCR; il glicerolo aiuta l'amplificazione di sequenze con un alto contenuto di G e C; il polietilenglicole, infine, promuove l'associazione macromolecolare escludendo il solvente, e faciliterebbe la reazione di PCR con basse concentrazioni di DNA stampo.

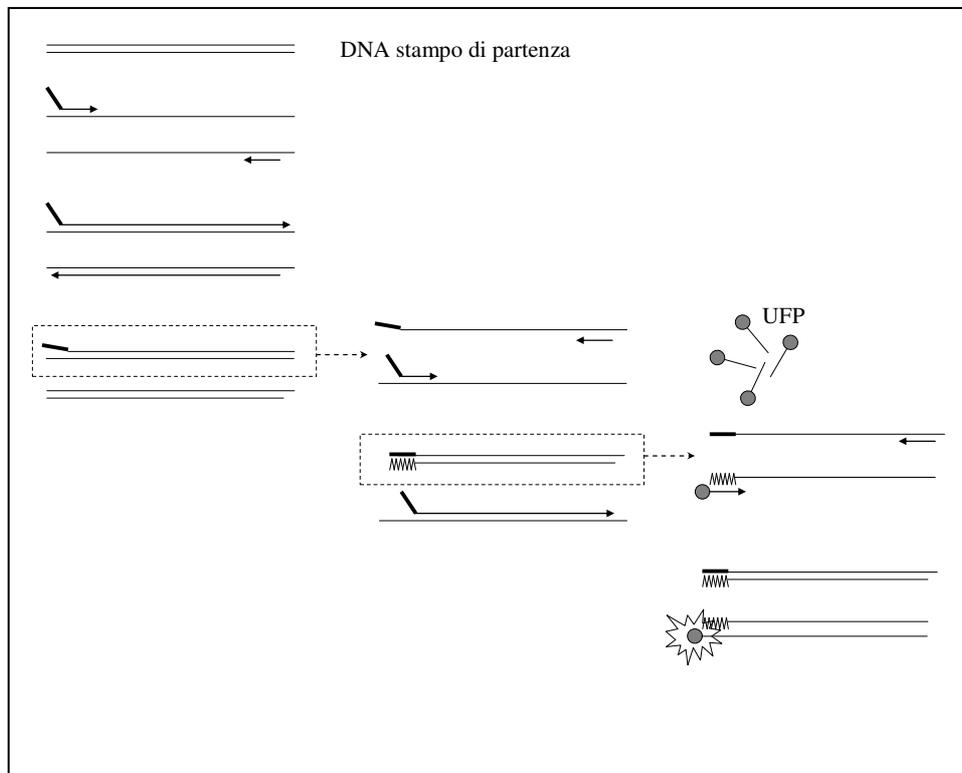


Figura2.7: Schematizzazione del metodo UFP-18. I primer forward e reverse si appaiono specificamente alle estremità della regione che interessa amplificare, e fanno da innesco per la polimerizzazione, come in qualunque reazione di PCR. Uno dei filamenti neosintetizzati, a questo punto, avrà una sequenza non complementare di 12 basi costituita dalla coda del primer forward (in figura, il tratto più marcato). Con un nuovo ciclo di PCR, i due filamenti saranno nuovamente amplificati. In particolare si verrà a creare un tratto di DNA che reca la sequenza complementare alla coda del primer (il tratto a zig-zag in figura), cui può legarsi il primer fluorescinato UFP-18 che, per competizione, diventa il nuovo primer forward perché è in eccesso. In questo modo viene incorporato un fluorocromo nel prodotto a doppio filamento di PCR.

Il controllo delle prove viene effettuato tramite corsa elettroforetica su gel di agarosio. Si allestisce un gel di agarosio alla concentrazione finale dell'1.2-1.5% in tampone TBE 1X (Tris 89mM, Acido Borico 89mM, EDTA 2mM) contenente 0.1 µg/mL di Bromuro di Etidio. Si carica la metà del prodotto di PCR di prova (circa 4 µL) previa miscela con 2 µL di gel-loading dye 6X (Blu di Bromofenolo 2.5%, Xylene Cianolo 2.5%, Glicerolo 25%) e acqua sterile per portare a volume finale di 10 µL. In una corsia si carica uno standard di peso molecolare (1 kb ladder). Si procede con la corsa elettroforetica a voltaggio costante (80-100 V). Al termine della corsa il gel viene visualizzato su transilluminatore fotografato tramite CCD camera, e si valuta la presenza delle bande di amplificato e anche approssimativamente la loro dimensione per conferma.

Il protocollo delle PCR è leggermente diverso rispetto a quanto detto per i marcatori dei pannelli industriali. Nella prima, la temperatura di ibridazione varia a seconda di come sono stati progettati gli oligonucleotidi forward e reverse; nella seconda invece è di 50°C, ottimale per l'appaiamento tra coda del primer forward e del primer UFP.

PCR I	
NUMERO DI CICLI	CONDIZIONI DEL CICLO
1	Denaturazione iniziale a 95°C per 12'
	Denaturazione a 95°C per 30''
35	Ibridazione a n°C per 20''
	Estensione a 72°C per 35''
1	Estensione finale a 72°C per 20'

PCR II	
NUMERO DI CICLI	CONDIZIONI DEL CICLO
1	Denaturazione iniziale a 95°C per 12'
	Denaturazione a 95°C per 30''
35	Ibridazione a 50°C per 20''
	Estensione a 72°C per 35''
1	Estensione finale a 72°C per 20'

L'altra differenza tra i due protocolli generali sta nella quantità di prodotto di PCR da usare per i pool, che sarà un poco più alta rispetto a prima: si utilizzano 4 µL di prodotto indifferentemente che sia marcato con 6-FAM, VIC o NED.

Sequenziamento

Per sequenziare una regione di DNA, il primo step consiste nella progettazione dei primer, che segue regole analoghe a quanto già detto in precedenza per la progettazione dei primer per il metodo UFP. Qualora si voglia sequenziare un esone, è meglio progettare i primer ad almeno 50/60 basi dagli estremi dell'esone in quanto le prime basi possono non venire lette in modo accurato.

Per aumentare il materiale genomico di partenza, si amplifica il DNA tramite PCR. A tal fine è però necessario, attraverso opportune prove, mettere a punto la reazione, sia per quanto riguarda la concentrazione di $MgCl_2$ che per la temperatura di ibridazione, eventualmente aggiungendo additivi quali dimetilsolfossido (5%) , formamide (0.3%) o glicerolo (5 - 10 %).

La reazione viene fatta in un volume finale di 25 μL .

Per ogni sequenza da amplificare, si usano:

Buffer 10x	2.5 μL
dNTPs 10x	2.5 μL
$MgCl_2$	Come da prove
Taq	0.1 μL
DNA 20 ng/ μL	2.5 μL
Primer	2.5 μL
H ₂ O	A volume

I cicli di PCR sono i seguenti:

Numero ripetizioni	Condizioni del ciclo
1	Denaturazione iniziale a 96°C per 10'
	Denaturazione a 96°C per 30''
30	Ibridazione a n°C per 30''
	Estensione a 72°C per 30''
1	Estensione finale a 72°C per 5'

Il prodotto di PCR così ottenuto viene quindi purificato per eliminare i primers e i sali che potrebbero disturbare nei passaggi successivi.

Purificazione del prodotto di PCR (con piastre Millipore)

In ogni pozzetto (dotato di filtro) della piastra, si aggiunge un volume di acqua sterile pari a 5 volte il volume del prodotto di PCR, poi si aggiunge in ciascun pozzetto il prodotto di PCR. Si chiude la piastra con il coperchio e la si mette ad asciugare sulla macchina a pressione Millipore, dopo aver fatto il vuoto (10 mmHg per circa 10 minuti).

Una volta asciugati, si trasferiscono nei pozzetti 50 µL di acqua, si lascia a temperatura ambiente per circa dieci minuti e si risospende con una pipetta la soluzione in ogni pozzetto facendo attenzione a non rompere il filtro sul fondo del pozzetto; si trasferisce il contenuto di ogni pozzetto in una eppendorf da 0.5 mL o in uno strip.

Prima di procedere con la reazione di sequenza vera e propria, generalmente si controlla su gel di agarosio all'1.2-1.5% la purificazione per vedere l'eventuale presenza di primer, per assicurarsi che durante la purificazione non sia stato perso l'amplificato e per stimare approssimativamente la quantità di prodotto da cui si parte.

Quindi si fa avvenire la reazione di sequenza, in un volume finale di 10 μL . A tal fine per ogni campione si prepara una reazione contenente:

Amplificato	5-6 μL ReadyReactionMix	1 μL
Buffer 5x	2.0 μL	
Primer 3.2 pmol/ μL	1.0 μL	
H ₂ O	A volume	

La ReadyReactionMix è una soluzione già pronta contenente l'enzima, i nucleotidi per la polimerizzazione e i dideossinucleotidi marcati tramite fluorescenza (BigDye V.1.1, Applied Biosystems).

Il primer che si aggiunge sarà il forward o il reverse, che funzionerà da innesco per la sintesi di sequenze forward o reverse; la sintesi di due sequenze dello stesso DNA di partenza è utile non solo come controllo, ma anche per poter sequenziare le prime 40-50 basi dopo il primer che di solito sono "perse" nella lettura.

Per far avvenire la reazione di sequenza si usano i termociclatori, impostando i cicli nel modo seguente:

Numero ripetizioni	Condizioni del ciclo
1	Denaturazione iniziale a 96°C per 10''
	Denaturazione a 96°C per 10''
30	Step facoltativo a 50°C per 5''
	Estensione a 60°C per 4'

Lo step intermedio a 50°C è facoltativo e serve per favorire l'appaiamento tra il primer e il DNA stampo.

Le sequenze vengono poi purificate prima di essere caricate sul sequenziatore automatico a capillari (modello 3730, Applied Biosystems, Tokyo, Japan).

Protocolli di purificazione di sequenza

Protocollo 1.

Ad ogni sequenza, si aggiunge 1/10 in volume di sodio acetato per salificare il DNA e 2.5 volumi di etanolo 100% così da far precipitare il DNA. Si lascia per 10 minuti con il tappo chiuso in ghiaccio (il freddo favorisce la precipitazione).

Si centrifuga 30 minuti a 3000 g e si elimina subito il surnatante.

Si centrifugano le provette capovolte per 15 secondi a 50 g per eliminare le gocce.

Si aggiungono 50 μ L di etanolo 70% per ogni campione, e si centrifuga nuovamente per 20 minuti a 3000 g.

Si elimina il surnatante e si centrifugano le provette capovolte per 15 secondi a 50 g (per eliminare le gocce); si lascia asciugare all'aria.

Si risospendono le sequenze in 10 μ L di formamide e poi vengono caricate sul sequenziatore previa denaturazione a 96°C per 2 minuti.

Protocollo 2.

Un altro protocollo utilizzato (di rado) per la purificazione delle sequenze è stato quello della Millipore (Sequencing Reaction Cleanup Kit). Questo protocollo, mantenendo una qualità dei purificati simile a quella del protocollo precedentemente descritto, consente di ottenere le sequenze purificate in circa 15 minuti di lavoro; essendo comunque più costoso è stato utilizzato solo quelle volte che era necessario per questioni di tempo.

In seguito alla reazione di sequenza, si aggiungono in ogni pozzetto 10 μ L di Injection Solution (soluzione fornita nel kit); dopo avere risospeso il contenuto del pozzetto, lo si trasferisce in una piastra Millipore Seq-96.

Si mette la piastra ad asciugare sulla macchina a pressione impostando il vuoto a 20 mmHg; si aspetta che i pozzetti si svuotino (circa 3 minuti). Dopo aver rimosso la piastra dalla macchina a vuoto, si asciuga il fondo della piastra con della carta assorbente e si aggiungono ad ogni pozzetto 20 μ L di Injection Solution.

Si mette la piastra sulla macchina a vuoto ad asciugare per circa 3-4 minuti (20 mmHg); una volta che i pozzetti si sono svuotati, si rimuove la piastra, si asciuga il fondo e si aggiungono ad ogni pozzetto 20 μ L di Injection Solution; dopo aver risospeso bene il contenuto di ogni pozzetto, 10 μ L di questo contenuto vengono caricati sulla piastra per il sequenziatore; si aggiungono ad ogni pozzetto 10 μ L di formamide e si carica la piastra sul sequenziatore dopo aver denaturato le sequenze per 2 minuti a 96°C.

Analisi delle sequenze

L'analisi delle sequenze ottenute può essere fatta tramite opportuni programmi di allineamento che confrontano tra loro due sequenze (BLAST); tuttavia il metodo più sicuro è la lettura a monitor dell'elettroferogramma (Figura 2.8).

Per visualizzare su computer i cromatogrammi forniti dal sequenziatore, abbiamo utilizzato il programma "Chromas Lite 2.0" prodotto dalla società Technelysium Pty Ltd., (freeware @ www.technelysium.com.au/chromas_lite.html), FinchTV (freeware @ www.fichtv.org) e Sequencher (Genes Code Corporation, DEMO version @ <http://www.genecodes.com/demos/>).

Utilizzando tali programmi abbiamo analizzato i cromatogrammi ed allineato le sequenze relative agli individui affetti e dei loro familiari ed abbiamo confrontato queste con le sequenze corrispondenti presenti nella banca dati *NCBI Entrez Nucleotide*.

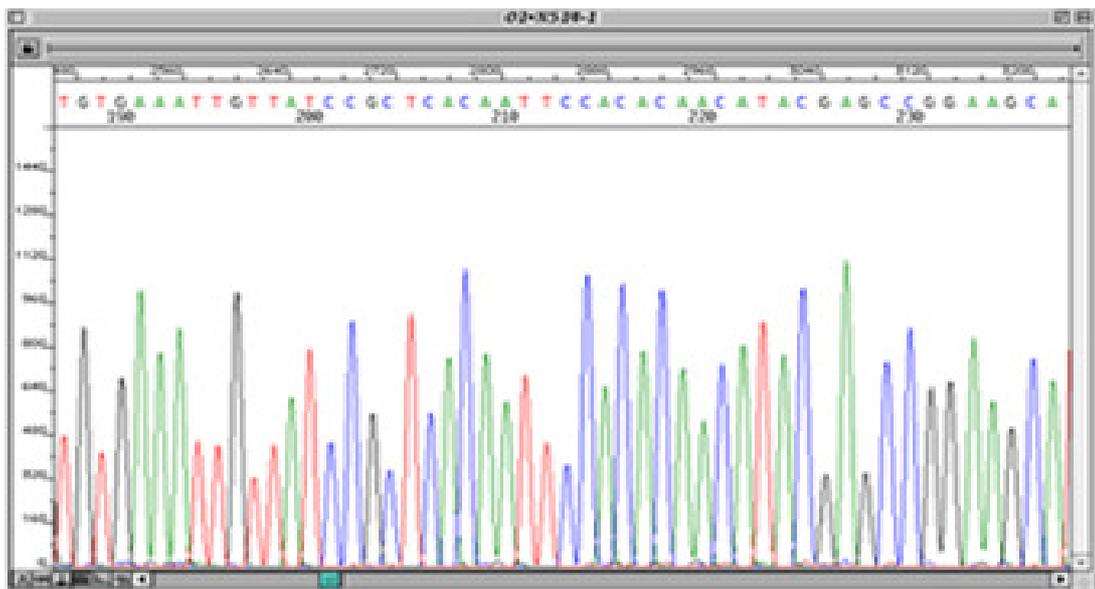


Figura 2.8. Esempio di diagramma che si ottiene come output di sequenziamento. La sequenza è rappresentata da un susseguirsi di picchi che rappresentano le basi nucleotidiche; i diversi colori sono attribuiti dal programma di analisi a seconda della base riconosciuta (rosso=T, verde=A, blu=C, nero=G) per facilitarne la lettura.

Calcolo Del LOD Score.

Sindrome di Mungan

Al fine di effettuare il calcolo dei LOD score, si è scelto per la malattia il modello di autosomica recessiva con penetranza completa, con frequenza della malattia dello 0,001.

Per l'analisi genome wide sul nucleo familiare del probando è stato utilizzato il software GeneHunter (Kruglyak et al, 1996). Gli eventuali errori di genotipizzazione sono stati rilevati con l'ausilio del programma Pedcheck (O'Connell et al, 1998).

Gli aplotipi sono stati ricostruiti con il programma GeneHunter e gli eventi di ricombinazione sono stati messi in evidenza nel pedigree della famiglia con il software HaploPainter (Thiele et al, 2005). Nell'estensione dell'analisi a tutta la famiglia il two point LOD score è stato generato dal programma MLINK, del LINKAGE package (Lathrop et al, 1984), ed il multipoint LOD score è stato calcolato utilizzando il programma SIMWALK2 (Sobel et al, 1996).

THC2

I LOD score a due punti e multipoint sono stati calcolati con il programma Merlin, Versione 1.1.2 (Abecasis et al., 2002), separatamente in ogni singola famiglia e nell'intero gruppo di famiglie. La proporzione stimata di famiglie in Linkage (α) ed il corrispondente HLOD (heterogeneity LOD scores) sono stati generati in maniera automatica dal software. L'analisi di linkage è stata effettuata impostando la trombocitopenia come un tratto a penetranza completa con una frequenza stimata della malattia dello 0,001. Tutti gli individui con piastrinopenia sono stati considerati affetti, poiché nessun dato di penetranza incompleta risulta né dalla letteratura né dai nostri campioni.

Gli aplotipi sono stati ricostruiti con l'ausilio delle funzioni del software Merlin.

Studi Funzionali

Clonaggio mediante TA Cloning

Il clonaggio è stato effettuato utilizzando il protocollo Topo TA Cloning (Invitrogen), che consente di inserire rapidamente un prodotto di PCR in un vettore plasmidico.

Il vettore plasmidico, fornito con il kit, risulta linearizzato e presenta una singola Timidina (T) protrudente al 3' al fine di consentire la ligazione specifica del prodotto di PCR con il vettore stesso.

Questo è reso possibile dal fatto che la Taq polimerasi, durante il processo di amplificazione, aggiunge al prodotto di PCR una singola deossiadenosina al 3'.

Inoltre il vettore è legato covalentemente alla Topoisomerasi I, che catalizza il processo di ligazione riducendo il tempo necessario a 5 minuti per ampliconi di dimensioni standard.

Per ogni trasformazione è necessaria un'aliquota di cellule chimicamente competenti (TOP10, DH5 α , TOP10F' o XLBlue).

Vengono quindi aliquotati:

- PCR 1-4 μ l
- Vettore 1 μ l
- Salt Solution 1 μ l
- Acqua a volume
- Volume totale 6 μ l

La reazione di legazione avviene a temperatura ambiente per 5-30 minuti. A questo punto la mix viene posta in ghiaccio e si procede al protocollo di trasformazione.

Si aggiungono 2 μ L della reazione di ligazione, ad un' aliquota di cellule competenti e si mescola delicatamente evitando di risospendere pipettando su e giu.

Le cellule vengono incubate in ghiaccio per 5-30 minuti.

A questo punto viene favorita la trasformazione con uno shock termico, ponendo le cellule in un bagnetto termico tarato a 42°C per 30 secondi.

Si trasferisce velocemente la provetta in ghiaccio e si aggiungono 250 μ L di SOC medium a temperatura ambiente.

A questo punto si pone la provetta in agitazione a 200rpm per 1 ora, a 37°C.

E' possibile, quindi piastrare 10-50 μ L della trasformazione, in piastre LB Agar, precedentemente preparate con un antibiotico e contenenti XGAL (40mg/ml) se viene scelta la selezione bianco-blu. Le piastre vengono incubate a 37°C per tutta la notte.

In genere è preferibile piastrare due volumi differenti in modo tale da essere sicuri che almeno una piastra presenti delle colonie ben spaziate per poter essere raccolte.

In genere l'efficienza del clonaggio risulta piuttosto elevata, si ottengono centinaia di colonie e, tra queste, è possibile selezionare e raccogliere solo le colonie bianche, relative alle cellule che contengono il vettore.

Clonaggio in vettore Reporter

I costrutti ottenuti mediante TA cloning sono stati trattati con l'enzima di restrizione prescelto in modo da separare l'inserto dal vettore. Il prodotto della digestione è stato purificato da gel (tramite colonnine QuiaQuick, Quiagen). Dopo aver digerito il vettore reporter con lo stesso/gli stessi enzimi di restrizione ed averlo defosforilato mediante incubazione con CIAP, vettore ed inserto vengono incubati con ligasi. In seguito si procede alla trasformazione come per TA cloning; l'unica differenza sta nella selezione delle colonie: in questo caso non si applicherà la selezione bianco/blu.

Estrazione DNA plasmidico

Per l'estrazione del DNA plasmidico è stata utilizzata una variante del protocollo Qiagen, protocollo di estrazione basato sulla Lisi Alcalina. Le colonie cresciute su piastra vengono prelevate singolarmente per essere poste in coltura liquida, in 1-5 mL di terreno, con l'aggiunta dell'antibiotico.

Vengono utilizzati tubi Falcon da 50 mL che vengono lasciati in agitazione, con tappo semiaperto, O.N. a 220-230 rpm/min a 37°C. A crescita ultimata le cellule vengono centrifugate a 3000 rpm per 10 minuti, il pellet viene risospeso in 300µL di soluzione P1 con aggiunta di RNase ad una concentrazione finale di 0,1 mg/µL e si trasferisce il tutto in un tubo eppendorf (1,5 mL).

A questo punto si aggiunge un uguale volume di soluzione P2, si mescola delicatamente per inversione e si lascia la mix a temperatura ambiente per 5 minuti esatti.

Si aggiungono, infine, 0,3 mL di soluzione P3 fredda, si mescola delicatamente e si lascia il flocculato ottenuto 10 minuti in ghiaccio. A questo punto si centrifuga a 13000 rpm per 20 minuti, si trasferisce il sopranatante in un nuovo tubo e, nell'eventualità che questo non sia limpido si centrifuga per altri 10 minuti.

Il DNA viene, quindi, precipitato aggiungendo 0,7 volumi di isopropanolo e centrifugando la mix a 13000 rpm per 30 minuti a 4°C. Si lava il pellet ottenuto con 200 µL-1 mL di etanolo 70% e lo si asciuga; infine il DNA viene risospeso in 20-30 µL di acqua.

Dual Luciferase Assay

Il vettore pGL3-basic (Promega) portante l'inserito da testare ed il vettore Renilla-TK (Promega), usato come controllo interno, sono stati cotrasfettati nelle cellule, utilizzando Lipofectamine 2000 (Invitrogen). L'attività delle luciferasi è stata misurata con il Dual Luciferase Assay kit (Promega) seguendo le istruzioni del costruttore.

Nel dual luciferase assay, le attività delle luciferasi firefly (*Photinus pyralis*) e renilla (*Renilla reniformis*) vengono misurate in maniera sequenziale nello stesso campione. Viene misurata per prima l'attività della firefly luciferasi aggiungendo il substrato specifico, in modo da generare un segnale stabile. In seguito alla quantificazione della luminescenza, la reazione viene stoppata e simultaneamente viene iniziata la reazione della renilla tramite l'aggiunta del substrato specifico.

Grazie all'utilizzo della renilla luciferasi come controllo interno è possibile ovviare a tutti quei problemi che generalmente si presentano negli esperimenti di trasfezione di geni reporter.

L'attività del reporter di controllo cotrasfettato garantisce un controllo interno al sistema che funge da "risposta" di base, normalizzando il dato, consentendo dunque di poter confrontare i dati di esperimenti diversi e di campioni diversi.

3. RISULTATI E DISCUSSIONE

3.1 RICERCA DEL GENE RESPONSABILE DELLA SINDROME DI MUNGAN

Nel 2007 il nostro gruppo di ricerca ha pubblicato un lavoro nel quale veniva descritto il locus cromosomico associato alla patologia (Deglincerti et al, 2007). Questo risultato è stato ottenuto tramite l'approccio della mappatura per omozigotità. Vista la natura autosomica recessiva della sindrome e l'alto grado di consanguineità presente nella famiglia, abbiamo cercato in un'analisi estesa a tutto il genoma le regioni di IBD (Identity By Descend), ipotizzando che il difetto genetico alla base della malattia si trovasse in una regione ereditata dai genitori degli affetti, eterozigoti per il tratto, da un progenitore comune, e quindi presente in omozigosi nei pazienti.

Per fare ciò è stata effettuata una prima analisi ristretta al nucleo familiare dell'index case (figura 3.1) utilizzando il kit ABI PRISM Linkage Mapping Set Version 2 (Applied Biosystems, vedere materiale e metodi). Questa analisi ha evidenziato una regione in 8q23–q24 tra i marcatori D8S285 e D8S272 (72 Mb, 68.9 cM) che dava un LOD score a due punti massimo di 1.96 ($\theta=0,0$) per il marcatore D8S1784 ed un multipoint LOD score massimo di 2.55 a 3.90cM dal marcatore D8S1784. Questi dati, benché non dessero una rilevanza statistica assoluta, potevano spingere ad approfondire l'analisi del locus individuato; la regione, infatti, era ancora molto estesa ed il numero di individui analizzato basso.

Per confermare il linkage al locus e possibilmente migliorare il dato statistico l'analisi è stata estesa a tutti i membri disponibili della famiglia, utilizzando 14 marcatori microsatellitari addizionali posizionati tra i marcatori D8S270 e D8S272. alcuni di questi marcatori non sono stati descritti in precedenza e sono stati progettati ad hoc per questo studio (tabella 3.1).

Grazie a questo approccio è stato possibile ottenere un LOD score a due punti massimo di 3.97 per il marcatore NT_008046_6 (tabella 3.2) ed un multipoint LOD score massimo di 5,01 (grafico 3.1).

La genotipizzazione di questi nuovi marcatori ha consentito inoltre di ridefinire la regione critica; il limite centromerico è stato posto tra il marcatore NT_008046_1 (ricombinante in III-1) ed il marcatore non ricombinante NT_008046_2. Il limite telomerico è stato invece posto tra il marcatore non ricombinante NT_008046_6 ed il NT_008046_7, ricombinante in V-4, come mostrato nella figura 3.2.

Come riportato in figura 3.3, la regione critica si estende per circa 13 Mb e contiene 8 marcatori (tutti in omozigosi), 23 geni noti, 16 geni ipotetici e 8 pseudogeni.

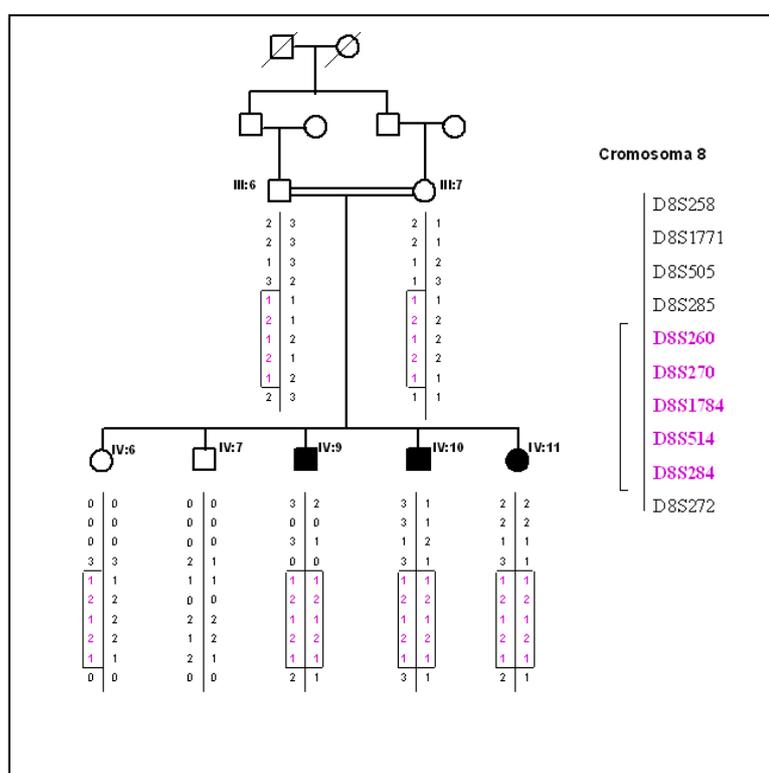


Figura 3.1: Il nucleo centrale della famigli utilizzato per l'analisi wide genome. I riquadri indicano la regione critica.

Polimorphism	Primers (Forward and Reverse)	Number of Alleles (*)	From Base... To Base... on chr 8 (according to the UCSC Database)
NT_008046_1	5'- GCTTGTATTTTGAGGTCCCC -3' 5'- AAGAGCCTTTCTGGAAGGTC -3'	5	109343276..109343522
NT_008046_2	5'- TGAATGCTTTGGATAGATAC -3' 5'- CTTGAGCTGATGATCATATAGGC -3'	6	112194539..112194744
NT_008046_3	5'- TGACATATGGCAGGATGCGG -3' 5'- CTCCTTTTATGTCCAAGCTGC -3'	5	121169470..121169740
NT_008046_4	5'- AACACATGAAAAAATGCATGCG -3' 5'- TAGATAAGCTTGCAGTGAAAG -3'	5	121415971..121416071
NT_008046_5	5'- TGGCCTTGACACAGTGCG -3' 5'- AGAACATAAGTGCTCTAGGGC -3'	3	121836515..121836706
NT_008046_6	5'- CATAGAAAAATGCATCAACAAGC -3' 5'- TTATCACTGCAAGCACCTCCAC -3'	4	122055257..122055376
NT_008046_7	5'- TATAGACGAATGGCTTGAGTCC -3' 5'- GTAGTAAGGAACCTATTGGC -3'	4	122431607..122431920

Tabella 3.1: I marcatori microsatellitari disegnati ad hoc per lo studio.

Markers	Recombination fraction							LOD _{max}	Physical position (kpb)
	0	0.01	0.05	0.1	0.2	0.3	θ		
D8S1830	-3.59	-0.77	0.33	0.62	0.63	0.41	0.15	0.67	107486
NT_008046_1	-∞	-3.97	-1.56	-0.65	-0.06	0.04	0.28	0.06	109343
NT_008046_2	2.4	2.36	2.18	1.95	1.46	0.95	0.00	2.4	112194
D8S1787	2.04	2	1.82	1.58	1.12	0.66	0.00	2.04	114520
D8S522	3.08	3.01	2.73	2.38	1.71	1.07	0.00	3.08	119629
D8S199	3.33	3.26	2.95	2.57	1.82	1.11	0.00	3.33	120349
NT_008046_3	2.62	2.58	2.39	2.13	1.56	0.98	0.00	2.62	121169
NT_008046_4	1.08	1.05	0.96	0.84	0.6	0.37	0.00	1.08	121416
NT_008046_5	0.31	0.3	0.25	0.2	0.11	0.05	0.00	0.31	121836
NT_008046_6	3.97	3.88	3.54	3.1	2.23	1.38	0.00	3.97	122055
NT_008046_7	3	2.94	2.71	2.4	1.75	1.08	0.00	3	122431
D8S514	0.97	0.95	0.87	0.76	0.55	0.34	0.00	0.97	123811
D8S1804	1.33	1.3	1.18	1.02	0.7	0.41	0.00	1.33	124934
D8S1799	-4.85	-1.83	-0.55	-0.09	0.18	0.17	0.21	0.24	125646

Tabella 3.2: LOD score a due punti per i marcatori della regione critica.

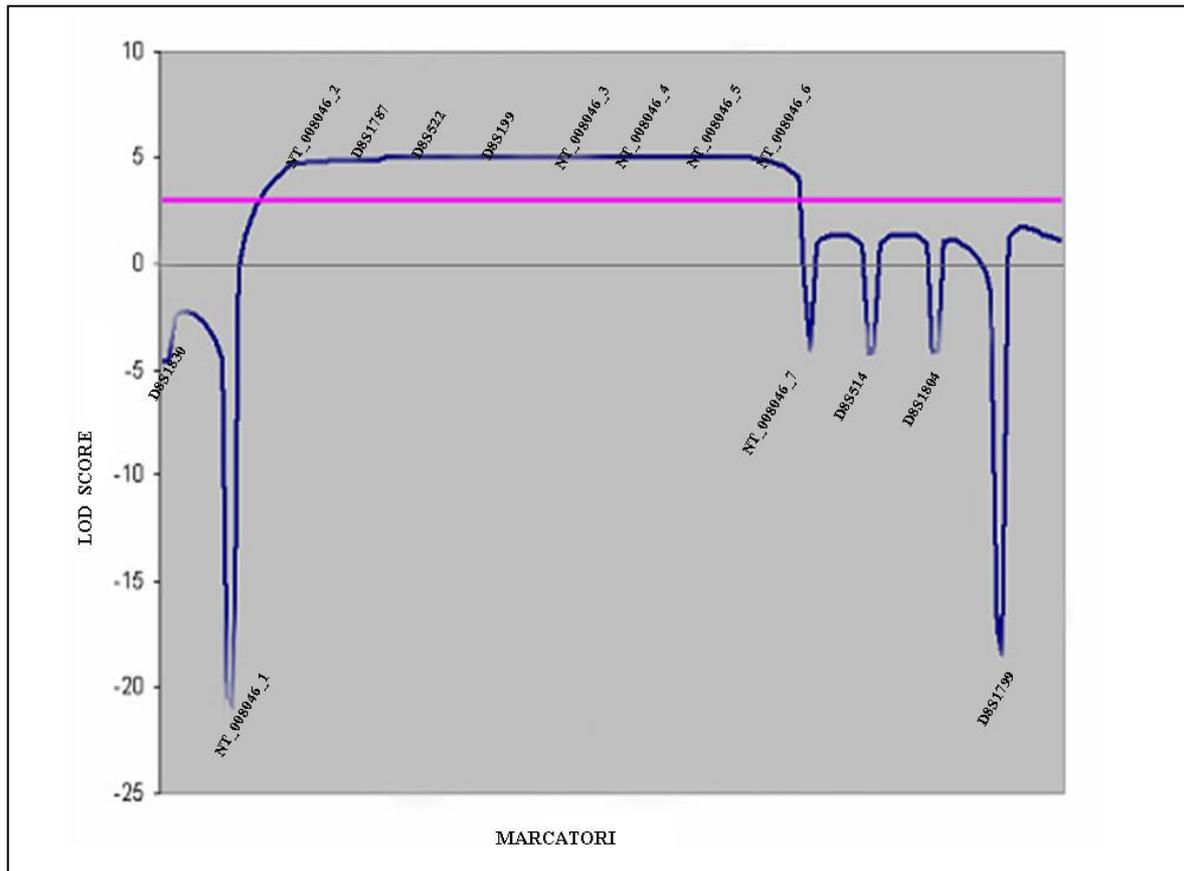


Grafico 3.1: Multipoint LOD score. Il valore rimane superiore a 5 per tutti i marcatori che definiscono la regione.

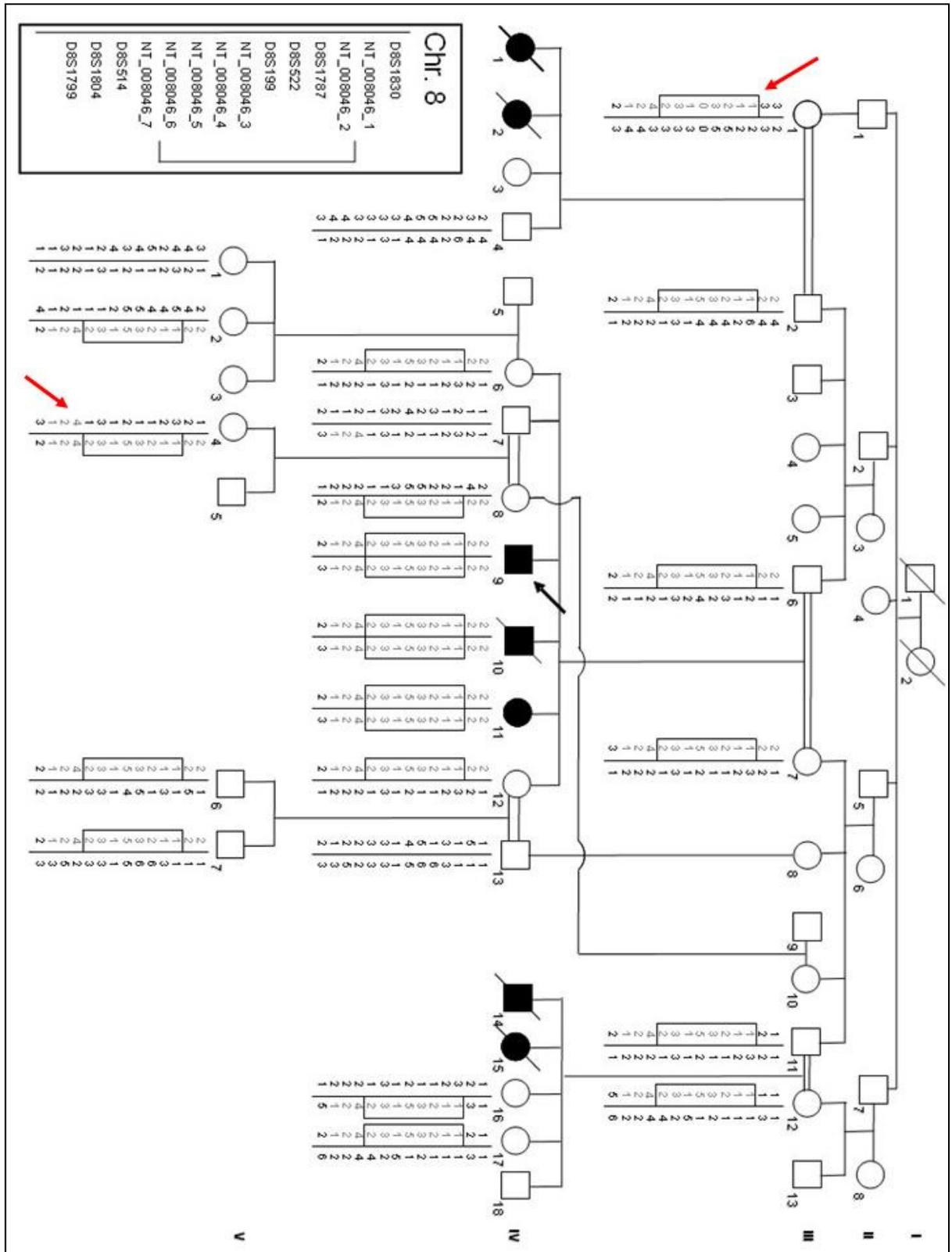


Figura 3.2: la freccia nera indica l'index case, le frecce rosse indicano gli eventi di ricombinazione che hanno consentito la ridefinizione della regione

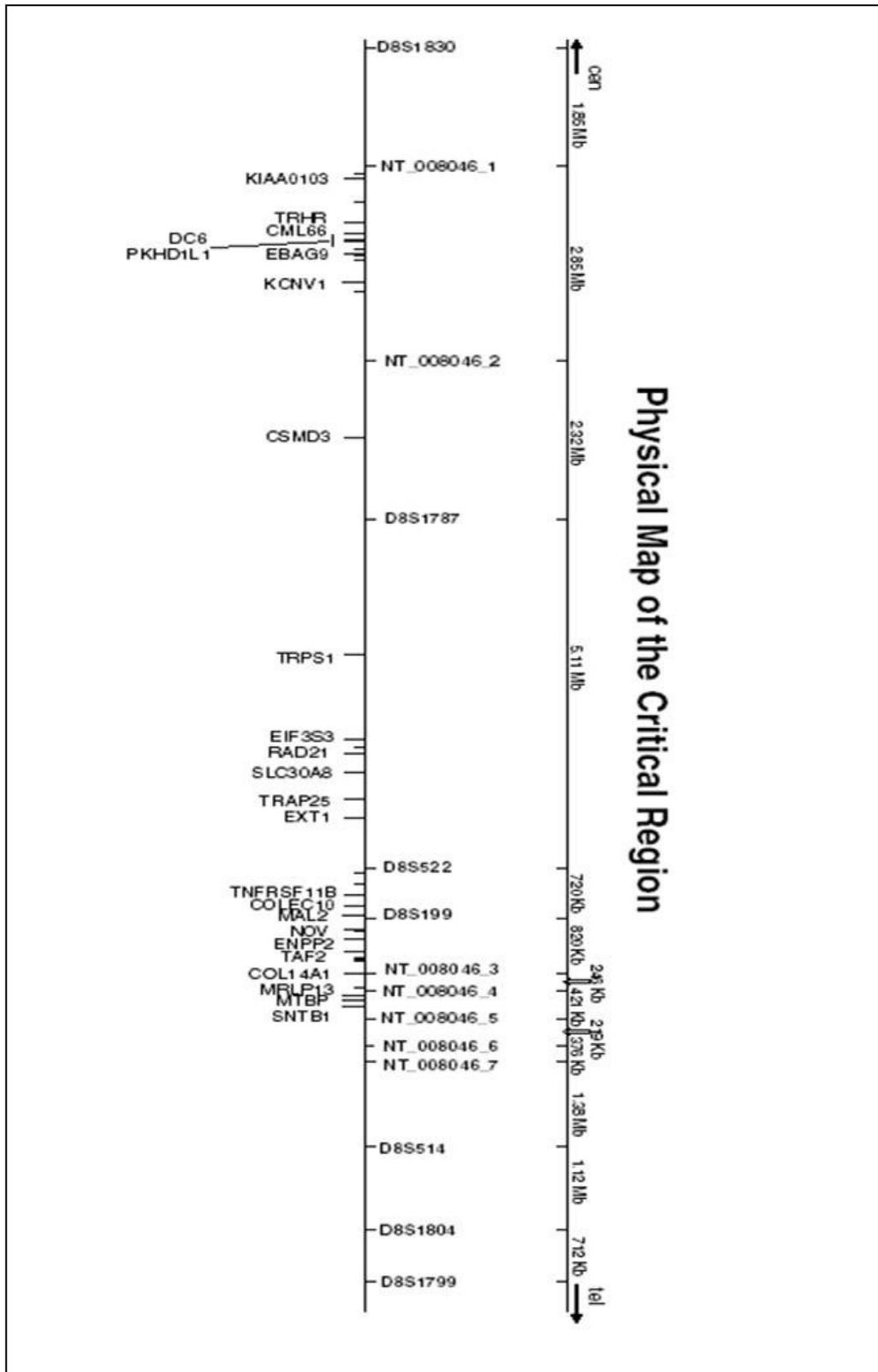


Figura 3.3: Mappa fisica della regione critica. Il limite centromerico è posto tra i marcatori NT_008046_1 e NT_008046_2; il limite telomerico tra i marcatori NT_008046_6 e NT_008046_7.

Con questo studio il nostro gruppo ha dimostrato l'esistenza di un gene recessivo predisponente alla sindrome di Mungan sul cromosoma 8q23-24.1.

Poiché i reperti dell'esame istopatologico delle biopsie intestinali dei pazienti suggerivano un coinvolgimento sia neurologico che muscolare, abbiamo cercato di individuare potenziali geni candidati, in base a questi dati, nella regione critica.

Il nostro interesse si è soffermato sui geni *SLC30A8* (un trasportatore dello zinco; Seve et al, 2004), *COLEC10* (un membro della famiglia delle C-lectine, che include proteine che posseggono sequenze collagen-like e domini di riconoscimento dei carboidrati; Ohtani et al, 1999), *SNTB1* (una proteina associata alla distrofina; Yoshizawa et al, 2003), *COL14A1* (proteina del Collagene di tipo XIV, undulina; Bauer et al, 1997), *KCNV1* (un canale neuronale del potassio; Hugnot et al, 1996), *MRPL13* (proteina ribosomiale mitocondriale; Kenmochi et al, 2001), *NOV* (proteina overespressa nel nefroblastoma e probabile fattore di crescita del tessuto connettivo; Kim et al, 1997).

Malgrado alcuni di questi geni avessero delle funzioni o pattern di espressione sovrapponibili con la sindrome, nessuna mutazione è stata identificata in seguito all'analisi di sequenza. Abbiamo così proseguito lo screening mutazionale di tutti i geni presenti nella regione critica, escludendo dall'analisi solo gli pseudogeni, ma le uniche varianti individuate si sono rivelate essere dei semplici polimorfismi presenti nelle banche dati pubbliche.

Tuttavia un risultato importante, benché inaspettato, è giunto dall'analisi dei geni *CSMD3* e *TRSP1*. L'analisi di sequenza ha messo in evidenza la presenza di SNP in eterozigosi negli affetti.

La tipizzazione di questi polimorfismi a livello del DNA dei pazienti e dei loro genitori ha consentito così di identificare un aplotipo diverso ai geni *CSMD3* e *TRPS1* tra quelli in comune tra i due genitori, verosimilmente generato da una doppia ricombinazione (figura 3.4).

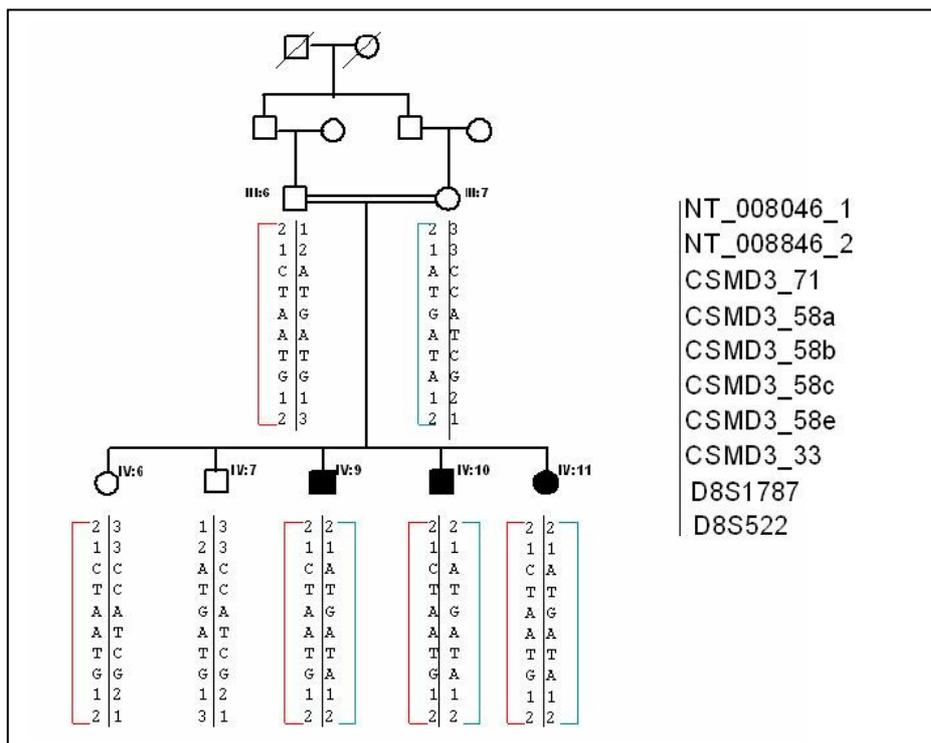


Figura 3.4: In questa figura sono riportati solamente gli SNP identificati nel gene CSMD3, a livello degli esoni 33, 58 e 71; il gene TRPS1 è posizionato tra il marcatore D8S1787 e D8S522.

Questo dato era in disaccordo con la tesi di partenza su cui si basa l'homozigosity mapping, cioè che la regione associata ad una patologia a trasmissione autosomica recessiva, oltretutto in una famiglia con elevato grado di consanguineità, deve essere in omozigosi negli affetti ed in eterozigosi nei genitori. Ciò ha reso necessaria una revisione della regione critica, che, in seguito a questi dati, si è trovata essere interrotta da molteplici eventi di ricombinazione.

In figura 3.5, viene mostrata la regione critica, mettendo in evidenza gli eventi di ricombinazione ancora da definire.

Infatti, come già detto, il primo evento di ricombinazione è tra i marcatori NT_008056_1 (ricombinante) e NT_008056_2 (non ricombinante). La distanza tra questi due marcatori è di circa 3 Mb, e comprende diversi geni; un'analisi più approfondita di questa regione sarebbe stata auspicabile, ma non è stato possibile individuare alcun

evento di ricombinazione, poiché i marcatori microsatellitari studiati non erano abbastanza informativi (dato non mostrato).

L'evento di ricombinazione successivo, non è più quello tra i marcatori NT_008056_6 e NT_008056_7, come detto in precedenza, ma bensì è localizzato tra il marcatore NT_008056_2 ed il gene *CSMD3*. Infatti la presenza di polimorfismi in eterozigosi negli affetti indica la presenza di ricombinazioni che hanno interrotto il locus malattia "ancestrale". Tra i due geni che presentavano le varianti in eterozigosi vi è un marcatore non ricombinante, il D8S1787, che segrega perfettamente ed in maniera informativa nella famiglia, e che quindi stabilisce nuovamente la presenza di altri due eventi di ricombinazione, tra *CSMD3* e D8S1787 il primo e tra D8S1787 e *TRPS1* il secondo. Inoltre è necessario che sia avvenuto un ulteriore evento di ricombinazione tra il gene *TRPS1* ed il marcatore D8S522, il quale segrega perfettamente (figura 3.2).

Osservando quindi la figura 3.5, si può notare dunque che l'unica regione, per così dire, certa (cioè che ha almeno due marcatori non ricombinanti consecutivi), è quella compresa tra *TRPS1* e NT_008056_7. Questa regione misura poco più di 5 Mb, meno della metà della regione originaria, e di queste più della metà sono comprese tra un marcatore non ricombinante ed uno ricombinante.

Se ancora non fosse stato eseguito il lavoro di sequenziamento, sarebbe necessario cercare di tipizzare maggiormente la regione, in modo da poterne comprendere a pieno l'estensione. Il fatto però che tutti i geni siano stati già analizzati semplifica di molto le cose e non rende necessario un lungo e dispendioso lavoro; sarebbe infatti necessario progettare molti marcatori e tipizzarli per tutti i componenti della famiglia.

La nostra scelta è stata invece quella di effettuare una nuova analisi estesa a tutto il genoma, per i motivi che andrò tra poco ad elencare.

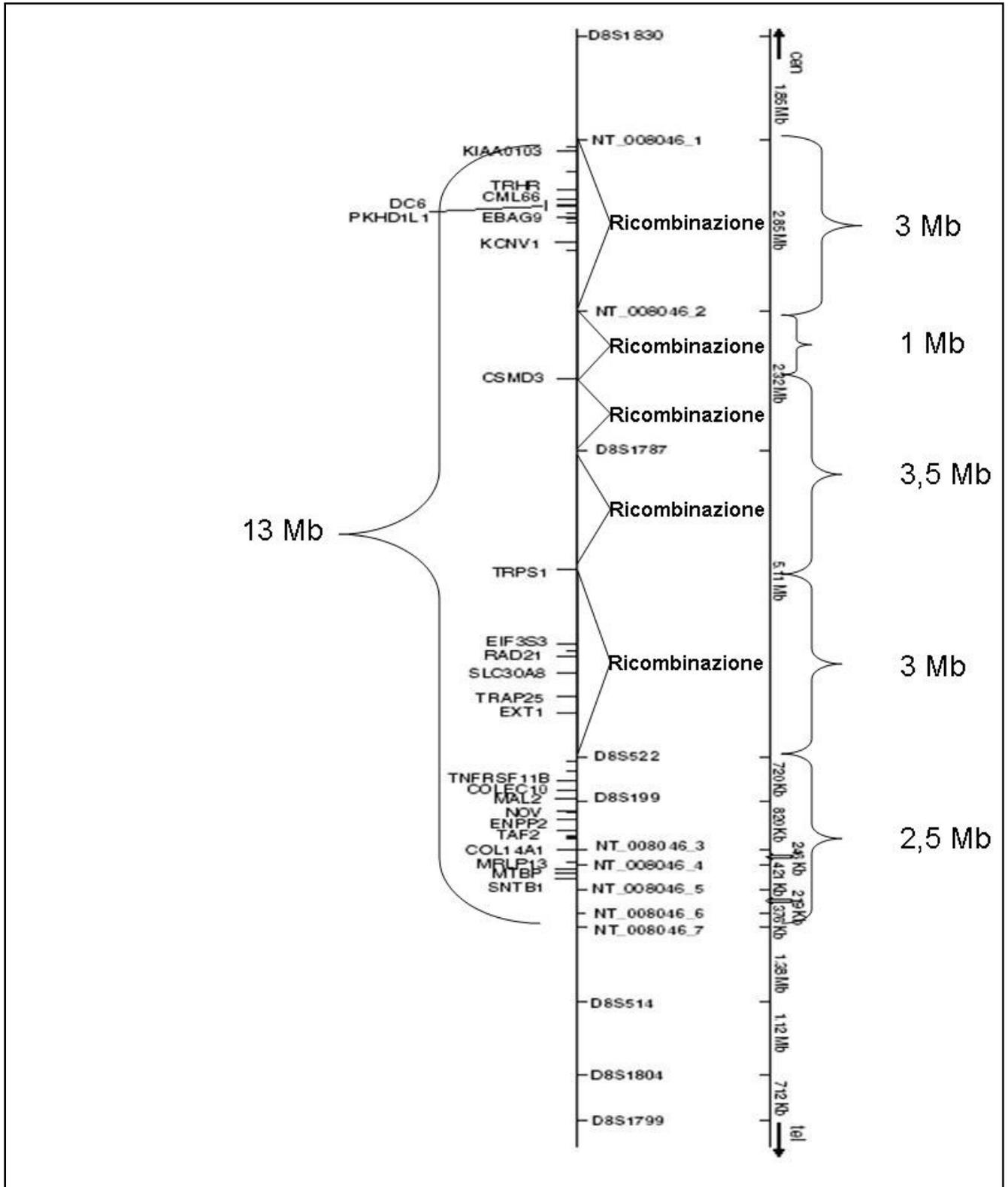


Figura 3.5: localizzazione degli eventi di ricombinazione e distanze relative tra i marcatori; per la descrizione vedere il testo.

Trovandosi di fronte una famiglia nella quale segrega una malattia rara recessiva e con alto grado di consanguineità l'approccio d'elezione, come già detto, è la mappatura per omozigosità, che in questo caso consiste nell'individuare le regioni di autozigosi; il termine autozigosi indica l'omozigosi per marcatori identici per discendenza (IBD), perché ereditati da un antenato comune recente.

È probabile che in famiglie consanguinee gli individui affetti da una malattia recessiva rara risultino autozigoti per i marcatori in linkage con la malattia. Se un membro affetto è omozigote per un allele marcatore ciò può essere conseguenza dell'autozigosi oppure dell'ingresso nell'albero genealogico di una copia dell'allele di diversa provenienza. D'altro canto, più l'allele è raro più è probabile che omozigosi ed autozigosi coincidano. Nel nostro caso l'allele è estremamente raro, ad oggi la famiglia TKS è l'unica nella quale è stata descritta la patologia, quindi l'ipotesi omozigosi=autozigosi è fortemente supportata.

E' tuttavia importante tenere conto anche del grado di parentela dei genitori degli affetti, infatti più la parentela è lontana, minore sarà la percentuale di patrimonio genetico in comune, e quindi la probabilità che le regioni di autozigosi coincidano con il locus malattia sarà maggiore. Viceversa una parentela stretta implicherà una alta percentuale di genoma identico, e quindi la possibilità che vi siano molte regioni IBD sarà alta.

Nel nostro caso i genitori degli affetti sono cugini di primo grado (quindi la parentela è stretta) e la percentuale di patrimonio genetico identico sarà di 1/16 (circa il 6%) e di conseguenza gli affetti saranno autozigoti in 1/32 (circa il 3%) di tutti i loro loci (il loro genoma).

E' possibile quindi ipotizzare che le regioni di autozigosi negli affetti della famiglia TKS siano molteplici.

La regione identificata da noi sul cromosoma 8, benché sia supportata dal dato statistico del LOD score, potrebbe non essere l'unica regione di autozigosi in linkage con la malattia. Questa ipotesi è fortemente supportata dal ritrovamento degli SNP in eterozigosi di cui si è ampiamente parlato.

Questo dato infatti mette in evidenza la necessità di una genotipizzazione ad alta densità di marcatori. Ho già accennato al fatto che i microsattelliti sono marcatori molto informativi, ma sono piuttosto rari.

L'utilizzo di marcatori come gli SNP avrebbe senz'altro messo fin da subito in evidenza la presenza di eterozigosi, ed avrebbe così consentito fin da subito di poter definire meglio la regione.

Un'altra considerazione importante riguarda la dimensione della regione. Se si prende in considerazione la regione critica così come è in figura 3.3, essa misura 13 Mb, ma in realtà la regione di autozigosi individuata dai marcatori microsattellitari era molto maggiore. Infatti se si osserva la figura 3.2 si noterà che i pazienti sono omozigoti anche per i marcatori esterni alla regione critica, ma esclusi per via di ricombinazioni in portatori obbligati (III:1) o soggetti sani (V:4).

La presenza di eterozigosi nelle sequenze dei geni *TRPS1* e *CSMD3*, ha dimostrato che la regione di omozigosi è in realtà più piccola (si veda la figura 3.5), e che il dato ottenuto dalla tipizzazione con i microsattelliti è falsato dalla relativamente bassa densità di marcatori.

Riteniamo sia ragguardevole la possibilità che vi siano altre regioni di autozigosi nel genoma degli affetti, e che possano essere sfuggite all'analisi iniziale; infatti, come già spiegato nella sezione dedicata in Materiali e Metodi, la densità di marcatori di questa metodica è di circa 1 marcatore ogni 10 cM, circa 10 Mb. E' quindi plausibile che diverse regioni inferiori alle 10 Mb possano semplicemente non essere state coperte dalla mappatura.

Per recuperare tali regioni abbiamo quindi deciso di effettuare una nuova genotipizzazione; abbiamo scelto come marcatori gli SNP ed abbiamo utilizzato la densità più elevata disponibile, circa 350.000 SNP distribuiti in tutto il genoma.

Per fare ciò ci siamo rivolti ad un laboratorio specializzato in questa metodica presso l'Università di Trieste il quale è dotato di una piattaforma per la genotipizzazione con tecnologia Illumina.

La genotipizzazione è già stata effettuata per i componenti del nucleo familiare del probando, ed attualmente sono in corso le analisi statistiche necessarie al fine di ottenere un dato comprensibile a partire dal dato grezzo.

Speriamo, con questa nuova genotipizzazione, di poter ottenere informazioni su tutte le regioni di possibile autozigosità nei pazienti. La nostra intenzione è quella di indagare eventuali geni candidati in queste regioni, dopo averne verificato la segregazione nella famiglia.

Se in seguito a questo approccio ancora non verrà identificato il gene responsabile di questa forma sindromica di CIIP, sarà necessario rivalutare l'approccio all'identificazione del gene.

Vi sono numerosi esempi di mutazioni di un singolo gene alla base di sindromi con CIIP. Bott e colleghi descrissero nel 2004 l'associazione di idrocefalo X-linked con stenosi dell'acquedotto del Silvio e CIIP in un bambino che presentava una mutazione nel gene *LICAM*, che colpiva un dominio importante della proteina (fibronectin type III). E' da notare che il gene *LICAM* controlla la differenziazione delle cellule interstiziali di Cajal (ICC), è quindi plausibile che questa mutazione sia responsabile del grave quadro di dismotilità intestinale riportato dagli autori. Una riduzione delle ICC è stata identificata anche nei pazienti della famiglia TKS, e quindi mutazioni in geni che controllano il differenziamento delle ICC potrebbero causare la patologia.

Vi sono altri esempi di mutazioni in singoli geni che possono causare quadri sindromici con CIIP, si pensi ai geni *TP*, *POLG*, *SOX10*, *FLNA1*, è quindi ipotizzabile che anche per la sindrome di Mungan una mutazione a carico di un singolo gene sia in grado di compromettere il differenziamento muscolare o neuronale intestinale

e/o l'integrità morfo-funzionale del comparto neuromuscolare enterico.

Inoltre gli esempi riportati in letteratura di mutazioni a singoli geni coinvolte in sindromi con CIIP, mettono in evidenza come le mutazioni stesse sono tutte collocate nelle regioni codificanti dei geni, quindi la nostra scelta di sequenziare le regioni codificanti dei geni candidati per posizione risulta motivata.

Tuttavia vanno tenuti in considerazione anche gli altri fattori che potrebbero determinare una patologia genetica. Infatti è possibile che mutazioni a carico di sequenze non codificanti alterino l'espressione del gene, aumentando o riducendo il prodotto genico, creando sbilanciamenti che possono risultare dannosi. D'altra parte è possibile che riarrangiamenti cromosomici alterino a loro volta l'espressione di un gene oppure causino la produzione di proteine difettose. Infatti traslocazioni, delezioni, duplicazioni e inversioni potrebbero andare a interrompere la sequenza genomica di un gene, senza alterarne le sequenze codificanti. Oppure nel caso di duplicazioni si potrebbe avere un aumento della dose genica, con conseguenti effetti nocivi.

Al fine di escludere il coinvolgimento di riarrangiamenti cromosomici, è stata effettuata un'analisi su un paziente della famiglia, tramite array-CGH (Comparative Genomic Hybridization), tecnica che consente di valutare le alterazioni comparando il DNA da analizzare con un DNA standard. Benché questa tecnica sia molto efficace nell'individuare delezioni e duplicazioni, non consente di individuare traslocazioni bilanciate e inversioni.

L'analisi sul soggetto affetto della famiglia turca è stata effettuata utilizzando una risoluzione di 100 Kb; dall'analisi non è risultata nessuna alterazione.

Benché siano stati effettuati notevoli sforzi per identificare la causa genetica alla base della sindrome, ancora manca un risultato positivo, che speriamo di ottenere dai risultati della genotipizzazione con SNP.

Sebbene la famiglia fosse relativamente grande e decisamente informativa, si è rivelata non essere un punto di partenza sufficiente all'individuazione del gene. Sicuramente sarebbe di grande aiuto poter analizzare un'altra famiglia con la stessa sindrome, ma ad oggi, l'unica famiglia descritta è quella descritta in questo studio.

Speriamo di poter trarre vantaggio dalle nuove tecnologie di *next generation sequencing*, che consentono di sequenziare ampie regioni genomiche, se non genomi interi, con costi e tempi relativamente contenuti. Sono sempre più diffusi i centri che dispongono di queste piattaforme tecnologiche, garantendo una facile accessibilità al servizio, e portando ad un graduale abbattimento dei costi.

Sarebbe auspicabile sequenziare tutti gli eventuali loci di autozigosità negli affetti della famiglia TKS ed in un genitore, al fine di valutare la presenza di eventuali mutazioni in regioni codificanti. In questa maniera inoltre sarebbe possibile indagare tutte le regioni non codificanti dei geni, come promotori, enhancers, eccetera, anche per i geni presenti nella regione in 8q.

Un ulteriore approccio possibile, basato sempre sulle nuove tecnologie di *deep sequencing*, sarebbe quello di analizzare tutte le regioni codificanti dell'intero genoma, il così detto *exoma*, oppure di sequenziare direttamente l'intero genoma dei pazienti.

L'identificazione del gene responsabile di una nuova sindrome con CIIP sarebbe di grande importanza per poter comprendere i meccanismi patologici alla base delle sindromi pseudo-ostruttive, ancora poco conosciute e comprese.

Inoltre, in questo studio, la pseudo-ostruzione intestinale viene per la prima volta associata ad esofago di Barrett. A differenza della CIIP, l'esofago di Barrett è una patologia diffusa, e rappresenta ad oggi l'unico precursore noto dell'adenocarcinoma esofageo. I meccanismi che portano alla formazione di questa metaplasia sono ancora poco conosciuti; e l'identificazione del gene della Sindrome di Mungan

potrebbe aiutare a comprenderli, e, possibilmente, aprire le porte a nuovi approcci terapeutici.

3.2 TROMBOCITOPENIA AUTOSOMICA DOMINANTE THC2: ESCLUSIONE DEL GENE MASTL IN 5 FAMIGLIE ED INDIVIDUAZIONE DI MUTAZIONI IN UN GENE POTENZIALMENTE COINVOLTO NELLA PATOGENESI DELLA SINDROME.

Il locus THC2, sul cromosoma 10p11.2-12, è stato associato a trombocitopenia autosomica dominante in due famiglie, una italiana ed una americana, da due gruppi in maniera indipendente (Savoia et al, 1999; Drachman et al, 2000). Nel 2003 il gruppo di Drachman ha individuato una mutazione nel gene *MASTL*, nel locus THC2, proponendola come mutazione causativa della patologia nella famiglia da loro studiata. Benchè recentemente lo stesso gruppo (Johnson et al, 2009) abbia prodotto prove a supporto del fatto che il gene possa essere coinvolto nella maturazione dei megacariociti e che, di conseguenza, possa giocare un ruolo nella patologia, ad oggi in nessuna altra famiglia con piastrinopenia sono state identificate mutazioni nel gene *MASTL*.

Il nostro gruppo di ricerca ha avuto modo di verificare la segregazione del locus THC2 in 4 famiglie italiane con trombocitopenia autosomica dominante, dimostrando, tramite sequenziamento, che in almeno cinque famiglie con piastrinopenia al locus THC2 (le quattro famiglie qui descritte e la famiglia descritta da Savoia e colleghi) il gene *MASTL* non è coinvolto nell'eziopatogenesi della malattia.

Ciò riduce le possibilità che il gene *MASTL* sia responsabile della patologia e apre le porte alla possibilità che il gene responsabile di piastrinopenia autosomica dominante THC2 sia ancora da identificare.

Al fine di verificare la segregazione del locus THC2 e di valutare il coinvolgimento del gene *MASTL* nella patogenesi di piastrinopenia autosomica dominante, abbiamo studiato cinque famiglie nelle quali ricorre tale patologia.

Tutti i soggetti disponibile delle famiglie sono stati genotipizzati per i seguenti CA repeats: D10S586, D10S572, D10S1775, D10S197, D10S111, D10S593, CArep01, CArep02, D10S174, che ricoprono il locus THC2 sul cromosoma 10p11-12.1.

Due di questi marcatori (D10S197 e D10S174) sono stati scelti dal kit ABI PRISM Linkage Mapping Set, versione 2.5 (Applied Biosystems). I restanti 7 sono stati genotipizzati progettando degli oligonucleotidi fiancheggianti CA repeats ed utilizzando il metodo UFP (Tabella3.3).

Di questi marcatori, 5 sono stati selezionati dalla mappa genetica Marshfield (www.bli.uzh.ch/BLI/Projects/genetics/maps/marsh.html), mentre i markers CArep01 e CArep02 sono stati individuati direttamente a partire dalla sequenza genomica, grazie all'ausilio del software Tandem Repeats Finder (<http://tandem.bu.edu/trf/trf.html>; Benson, 1999).

Marker	Forward Primer	Reverse Primer	Position on Chr10 according to the UCSC database
D10S586	5'-GAACAGCAGGATGTCTCG-3'	5'-TGAGACTTCATCACATGCGC-3'	24736107...24736260
D10S572	5'-TGGGTAGGAGGCCCTTAG-3'	5'-TTGTAGAATGCCAGGTAGGG-3'	25677181...25677532
D10S1775	5'-CCTAGTCCCAGAGTCTTG-3'	5'-CACAGGTATTGCTGAGTGAG-3'	26433775...26433947
D10S111	5'-AATGTGGTAGCCATCTC-3'	5'-ACATAGGTAGGGTCACACAG-3'	26810739...26810995
D10S593	5'-CAGGACCAACAACCAGAATA-3'	5'-CACATACCTGTTGCCATTC-3'	27163370...27163563
CArep01	5'-GTAAAGTACTACGCCAC-3'	5'-AATCATCTACTGTTGTAGCC-3'	27985712...27985881
CArep02	5'-CAGAGAGTCAACAAGTAG-3'	5'-GATGATTGTTTGTCTGCCA-3'	28391309...28391592

Tabella 3.3: I marcatori microsatellitari progettati per lo studio di segregazione del locus THC2 nelle famiglie analizzate.

I risultati della genotipizzazione sono riportati per ogni singola famiglia nelle figure dalla 3.6 alla 3.10, nella tabella 3.4 sono riportati i LOD scores.

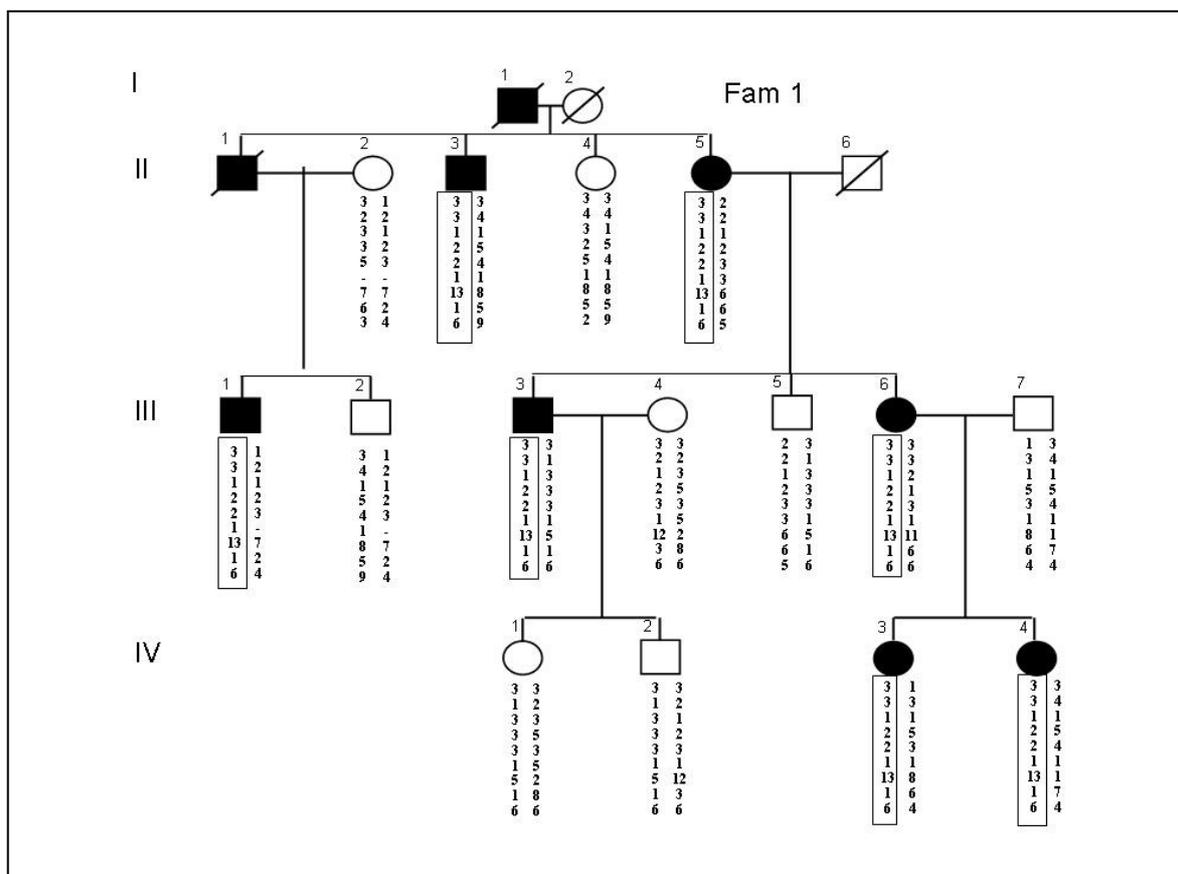


Immagine 3.6: Aplotipi dei componenti della famiglia 1 al locus THC2; si può notare come il locus segreghi perfettamente nella famiglia.

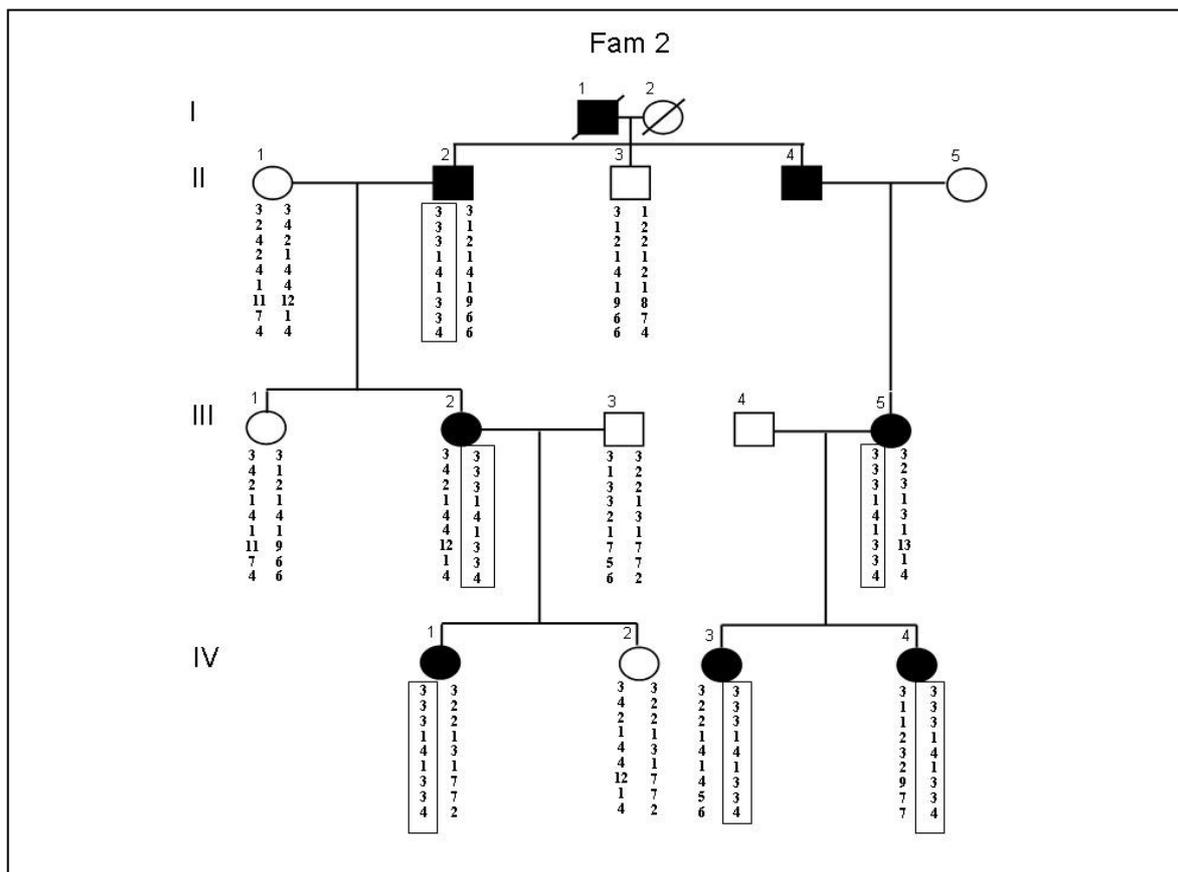


Figura 3.7: alleli della famiglia 2 al locus THC2; anche in questo caso la segregazione del locus è confermata,

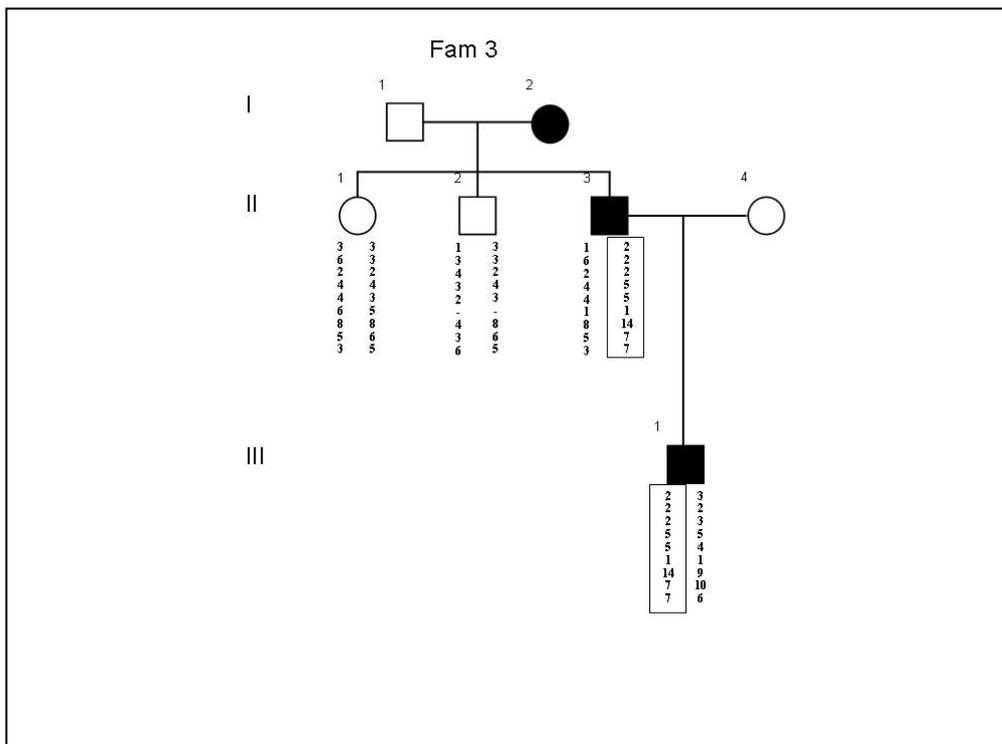


Figura 3.8: La segregazione del locus THC2 è confermata anche nella famiglia 3, in base alla ricostruzione degli aplotipi.

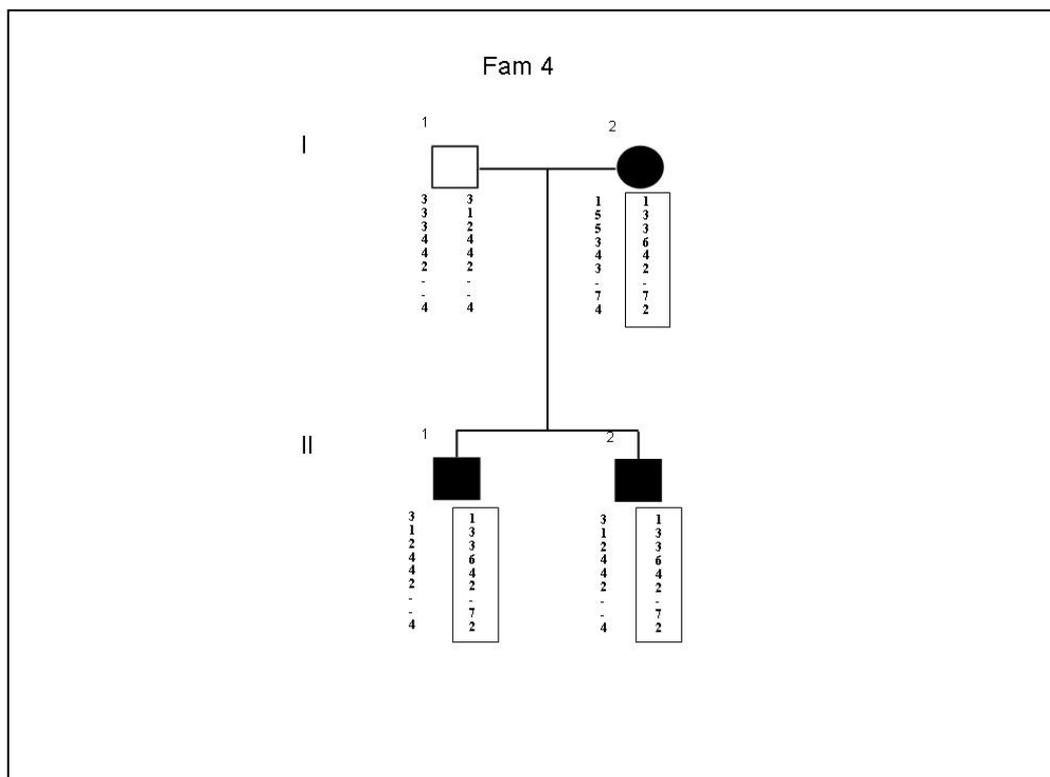


Figura 3.9: Alleli dei componenti della famiglia 4 al locus THC2; la segregazione è confermata.

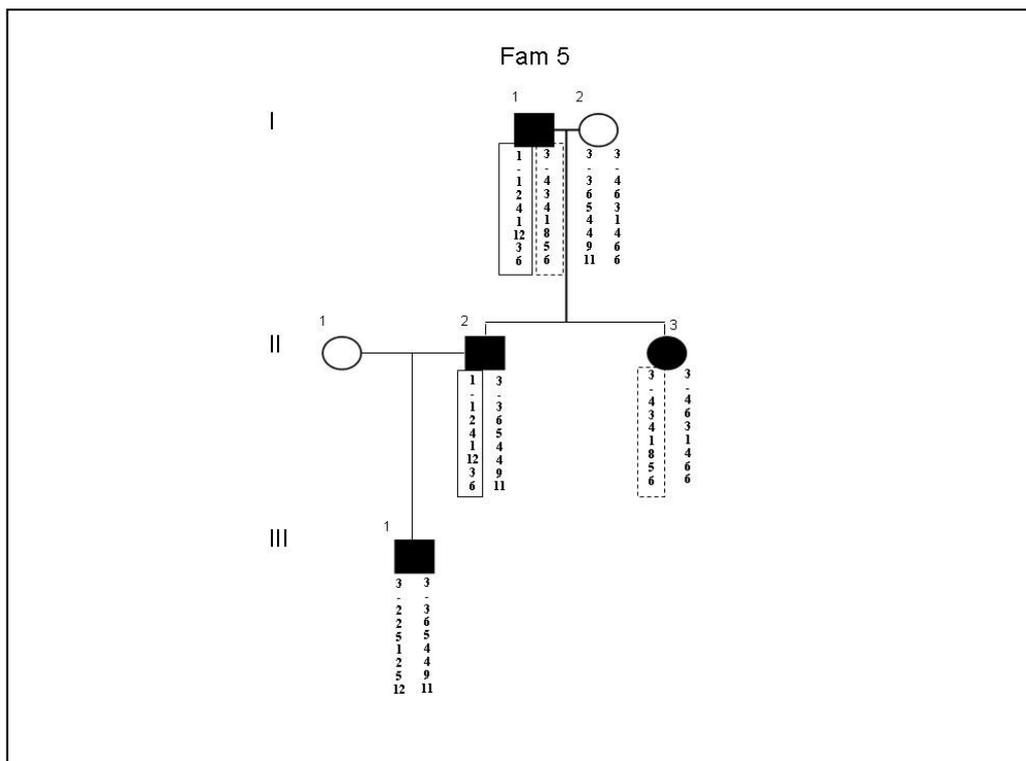


Figura 3.10: La ricostruzione degli aplotipi dei componenti della famiglia 5 per i marcatori del locus *THC2*. Come si può notare il padre affetto I:1 passa ai figli affetti II:2 e II:3 due cromosomi diversi, escludendo in questa famiglia il locus *THC2*.

Famiglia	1	2	3	4	5	1	2	3	4	5
Marker	Multi-point LOD scores					Single-point LOD scores				
D10S586	3.15	2.28	0.60	0.27	-2.64	0.47	0.30	0.60	0.00	-0.43
D10S572	3.31	2.39	0.60	0.30	-9.42	2.70	1.64	0.60	0.30	0.00
D10S1775	3.31	2.39	0.60	0.30	-∞	1.43	1.41	-0.48	0.30	-∞
D10S197	3.31	2.39	0.60	0.30	-10.50	1.50	-0.23	0.60	0.30	-3.95
D10S111	3.31	2.40	0.60	0.30	-∞	3.23	0.10	0.53	0.00	-∞
D10S593	3.31	2.40	0.60	0.30	-14.39	0.55	0.55	0.30	0.30	-0.74
CArep01	3.31	2.40	0.60	0.30	-∞	3.31	2.35	0.60	0.00	-∞
CArep02	3.31	2.40	0.60	0.30	-∞	1.59	2.30	0.60	0.00	-∞
D12S174	3.31	2.40	0.60	0.30	-∞	2.09	0.41	0.60	0.30	-∞

Tabella 3.4: Valori dei Single point LOD score e Multi point LOD score a $\theta=0$ per i marcatori al locus *THC2* nelle 5 famiglie analizzate.

Nelle famiglie dalla 1 alla 4 il locus sul cromosoma 10p11.2-12 co-segrega con la malattia nelle generazioni, mentre nella famiglia 5 la trasmissione degli aplotipi non supporta la colocalizzazione del locus THC2 con la piastrinopenia.

Per le famiglie 1 e 2 abbiamo ottenuto valori di LOD score che confermano il linkage al locus THC2: la soglia di significatività statistica è superata dal solo dato della famiglia 1, che raggiunge un LOD score di 3.3 per i marcatori D10S111 e CArep01; per la famiglia 2 il dato ottenuto suggerisce la presenza di linkage al locus, con un LOD score massimo di 2,4 per i marcatori CArep01 e CArep02 (tabella 3.4).

A causa delle loro piccole dimensioni, le famiglie 3 e 4 non riescono a contribuire in maniera significativa al LOD score, raggiungendo valori massimi di 0,6 e 0,3 rispettivamente.

L'esclusione della famiglia 5 può essere ricondotta all'eterogeneità genetica, e potrebbe quindi indirizzare questa famiglia al secondo locus nel quale è localizzato un gene responsabile di trombocitopenia autosomica dominante (THC4, *CYCS*), oppure ad altri loci ancora sconosciuti

Dall'analisi della segregazione degli aplotipi non è stato possibile individuare nessuna ricombinazione che consentisse di restringere l'intervallo cromosomico del locus THC2, i cui limiti rimangono gli stessi riportati in precedenza. La regione critica si estende per 5,84 cM dal marcatore D10S586 ed il D10S174 ed al suo interno vi sono 34 geni candidati per posizione.

Al fine di valutare il coinvolgimento del gene *MASTL*, abbiamo analizzato tramite sequenziamento diretto le parti esoniche codificanti e le sequenze introniche fiancheggianti. L'analisi è stata effettuata a partire dal DNA degli affetti delle famiglie 1-4 ed, inoltre, per un affetto della famiglia descritta da Savoia e colleghi nel 1999, che chiameremo famiglia 0 (zero).

Nessuna mutazione è stata identificata nel gene *MASTL* in nessuna delle 5 famiglie.

I dati da noi ottenuti supportano fortemente l'ipotesi che nella regione cromosomica 10p11.2-12 sia localizzato un gene coinvolto nella patogenesi di piastrinopenia autosomica dominante.

In questa regione è presente un gene *MASTL*, nel quale è stata identificata una sostituzione aminoacidica in una famiglia americana con un quadro patologico sovrapponibile a quello delle famiglie studiate da noi (Gandhi et al, 2003).

La mutazione identificata colpisce il residuo 167 della proteina, causando un cambio amminoacidico da aspartato a glutammato (p.E167D).

Questo gene codifica per una ipotetica kinasi (microtubule-associated serine/threonine kinase-like) con due domini chinasi, uno C-terminale ed uno N-terminale.

Recentemente, lo stesso gruppo che ha individuato la mutazione, ha pubblicato uno studio nel quale ha cercato una correlazione tra l'espressione di *MASTL* e la trombocitopenia (Johnson et al, 2009). Tramite KnockDown transiente del gene *MASTL*, utilizzando una sonda antisenso basata su morfolino, nel pesce Danio Rerio (Zebrafish), sono riusciti ad osservare una riduzione di trombociti circolanti.

Benchè dall'esperimento sia possibile evidenziare un meccanismo patogenetico dovuto a perdita di funzione, l'effetto della mutazione E167D non è stato chiarito. Poiché il residuo mutato si trova nel dominio chinasi N-terminale (che va dal residuo 35 al 835), gli autori ipotizzano che esso possa in qualche maniera essere cruciale per l'attività kinasica della proteina. Tuttavia non è stato effettuato alcun saggio in vitro che potesse dimostrare la perdita di funzione (per esempio un saggio di autofosforilazione).

E' da notare inoltre che la mutazione E167D è indicata come benigna dai software Polyphen (genetics.bwh.harvard.edu/pph/) e SIFT (blocks.fhrc.org/sift/SIFT.html), che rappresentano ottimi strumenti per la predizione dell'effetto funzionale di nuove mutazioni.

Inoltre il residuo 167 è l'unico, in una sequenza di 32 aminoacidi estremamente conservati in almeno 10 specie, per il quale una

variazione è tollerata (tabella 3.5); in particolare in questa posizione è presente una glutammina (Q) in Gallus Gallus ed una lisina (K) in Xenopus Laevis.

0	ref NP_116233.2	Homo Sapiens: <i>MASTL</i>	LHRHGIIHRDLKPDNMLISN	E	GHIKLTDFGLS
1	ref XP_001104386.1	Macaca mulatta: microtubule associated serine/thr...	LHRHGIIHRDLKPDNMLISN	E	GHIKLTDFGLS
2	ref XP_848866.1	Canis familiaris: similar to Microtubule associated se...	LHRHGIIHRDLKPDNMLISN	E	GHIKLTDFGLS
3	ref XP_592286.3	Bos Taurus: hypothetical protein isoform 1	LHRHGIIHRDLKPDNMLISN	E	GHIKLTDFGLS
4	ref XP_225629.4	Canis familiaris: similar to microtubule associated se...	LHRHGIIHRDLKPDNMLISN	E	GHIKLTDFGLS
5	ref NP_080255.2	Mus Musculus: Mastl.	LHRHGIIHRDLKPDNMLISN	E	GHIKLTDFGLS
6	ref XP_001376016.1	Monodelphis domestica: hypothetical protein ...	LHRHGIIHRDLKPDNMLISN	E	GHIKLTDFGLS
7	ref XP_418589.2	Gallus Gallus: hypothetical protein [Gallus gallus]	LHRHGIIHRDLKPDNMLISN	Q	GHIKLTDFGLS
8	gb AAH68968.1	Xenopus laevis MGC83260 protein >gi 49364656 gb...	LHRHGIIHRDLKPDNMLISN	K	GHIKLTDFGLS
9	ref NP_001003609.1	Danio Rerio (Zebrafish): <i>MASTL</i> .	LHRHSIIHRDLKPDNMLISN	E	GHIKLTDFGLS

Tabella 3.5: allineamento multiplo della regione intorno al residuo 167 della proteina MASTL: si può notare come l'unico residuo la cui variazione è tollerata è il 167 (freccia).

Gli autori commentano il fatto che non vi siano altre famiglie con piastrinopenia e mutazioni al gene *MASTL*, addebitando questa mancanza alla modestia dei sintomi della patologia, ed alla relativa difficoltà con cui i pazienti giungono all'attenzione clinica.

La nostra ipotesi è invece che la mutazione individuata da Gandhi e colleghi non sia altro che una rara variante privata della famiglia descritta da Drachman et al. nel 2000.

Il nostro gruppo ha dimostrato che in almeno cinque famiglie con trombocitopenia autosomica dominante, in cui la patologia co-segrega con il locus *THC2* sul cromosoma 10p11.2-12, il gene *MASTL* non è

coinvolto nella malattia, e che quindi vi deve essere un altro gene al locus *THC2* responsabile di piastrinopenia.

Al fine di identificare il gene responsabile della patologia nelle 5 famiglie con piastrinopenia *THC2* e per le quali era stato escluso il gene *MASTL*, abbiamo deciso di sequenziare i geni presenti nella regione critica, tramite sequenziamento diretto delle porzioni codificanti e delle regioni introniche fiancheggianti.

Benchè questo lavoro fosse già stato eseguito in parte da Savoia e colleghi nella famiglia descritta nel 1999, non era stata identificata alcuna mutazione nelle regioni codificanti dei geni presenti nel locus (Savoia, comunicazione orale). La nostra speranza era che, avendo a disposizione 5 famiglie distinte, la probabilità di individuare mutazioni diverse in uno stesso gene potesse aumentare. Infatti, se per lo screening mutazionale si ha a disposizione una sola famiglia, esiste la possibilità di non vedere la mutazione, e quindi di “perdere” il gene. Nel caso in cui la mutazione si trovi nel promotore di un gene, oppure vi sia una delezione o una amplificazione di sequenze introniche che altera lo splicing, l’espressione, eccetera, l’analisi di una sola famiglia non permetterebbe l’identificazione del gene, mentre la possibilità di analizzare più famiglie, aumenta la probabilità che almeno in una di esse la mutazione si trovi in una regione codificante, e quindi facilmente individuabile tramite sequenziamento.

L’analisi di sequenza di 34 geni per 5 famiglie, consiste in un lavoro molto lungo e dispendioso. Abbiamo così deciso di suddividere il lavoro, lavorando in collaborazione con il gruppo italiano di studio delle piastrine (in particolare i gruppi di Balduini C. e colleghi, Savoia A. e colleghi, Perrotta S. e colleghi).

Il lavoro è stato suddiviso in maniera da dare ad ogni laboratorio un uguale carico di lavoro.

Tra i geni analizzati dal nostro gruppo, vi è il gene *ANKRD26*.

In questo gene sono state individuate delle varianti non descritte in letteratura in 4 delle 5 famiglie analizzate.

Le varianti, tutte sostituzioni nucleotidiche in eterozigosi, non sono presenti nei database pubblici e non sono presenti in 200 cromosomi di controllo.

Le varianti sono tutte localizzate nella regione trascritta e non tradotta al 5' del gene (5'UTR), nelle posizioni -118C>T (Fam 3), -127A>T (Fam 2), -128G>A (Fam 1) e -134G>A (Fam 0) dall'inizio di traduzione.

Le sequenze codificanti del gene sono risultate essere wild-type, ed in nessuno degli altri geni del locus *THC2* sono state identificate mutazioni o varianti in nessuna delle famiglie.

Tutti i membri delle famiglie sono stati tipizzati per le mutazioni individuate, verificandone la segregazione nei singoli pedigree; tutti gli affetti presentano la mutazione in eterozigosi, mentre gli altri componenti sono wild-type.

Dall'analisi di sequenza di un affetto della famiglia 4 non è emersa alcuna mutazione nei geni analizzati.

Ciò potrebbe essere dovuto al fatto che, vista la ridotta dimensione della famiglia e l'eterogeneità genetica della patologia, il gene responsabile di trombocitopenia non sia localizzato nel locus *THC2*, benchè il locus segregasse nel pedigree.

D'altro canto è plausibile che anche la famiglia 4 abbia una mutazione nel gene *ANKRD26*, ma che questa sia localizzata in regioni non visibili nell'analisi di sequenza da noi effettuata (per esempio nel promotore). In futuro ci riproponiamo di vagliare queste ipotesi, verificando il coinvolgimento del gene anche in questa famiglia, oppure indirizzando la famiglia all'analisi di altri loci o geni causativi di trombocitopenia.

Il fatto che le mutazioni identificate nelle quattro famiglie, benché diverse tra loro, sono tutte collocate in una ristretta stringa di nucleotidi (figura 3.11), ci ha motivato fortemente a ritenere di avere individuato le mutazioni responsabili della patologia.

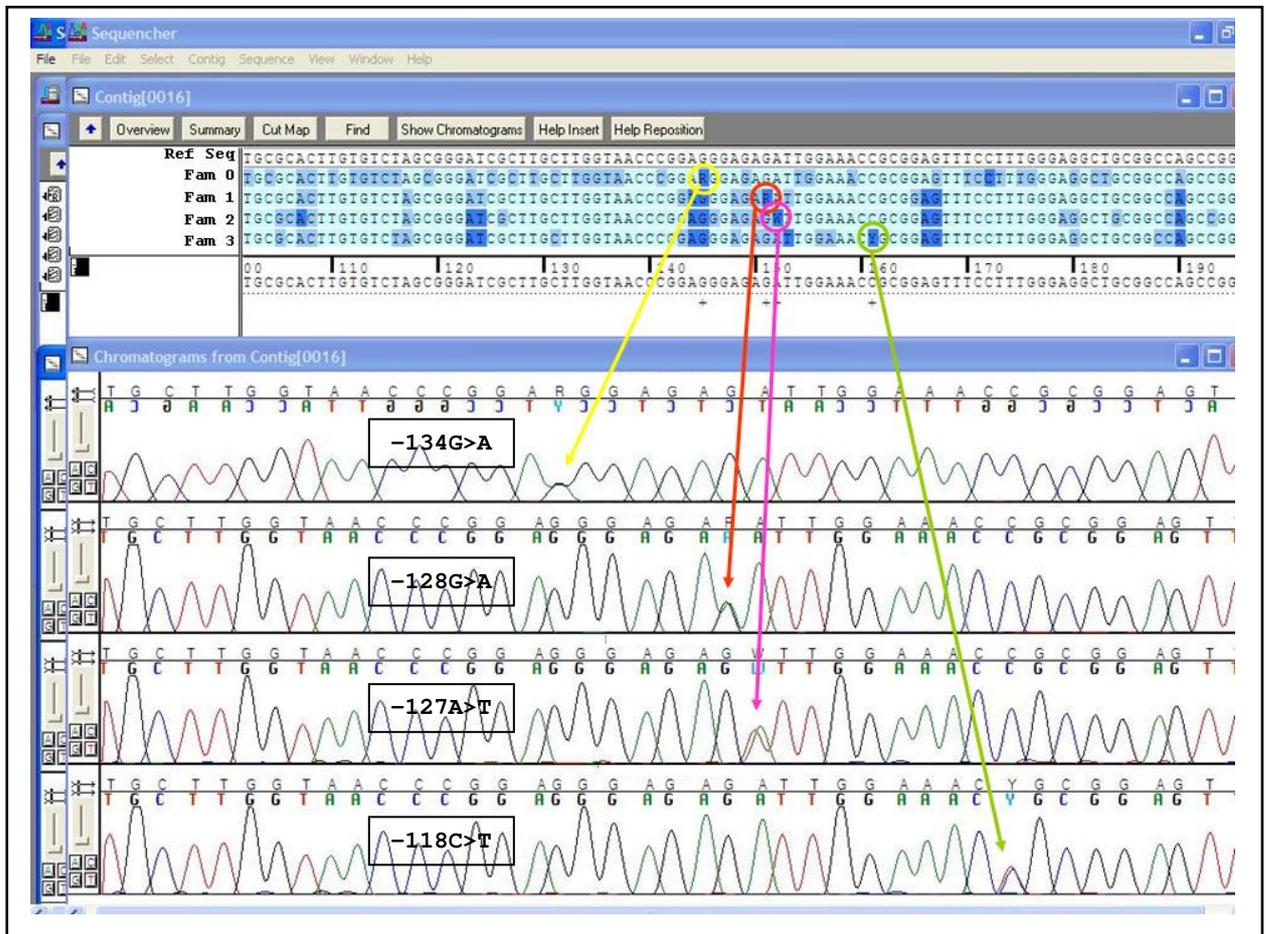


Figura 3.11: screenshot dal software Sequencer: nel riquadro in alto sono mostrati gli allineamenti delle sequenze dei componenti le famiglie con una sequenza di riferimento wild-type; nel riquadro in basso sono mostrati gli elettroferogrammi e sono indicate le mutazioni in eterozigosi.

Il 5'UTR di un gene rappresenta una regione trascritta e non tradotta a monte dell'ATG. Le funzioni di queste regioni sono prevalentemente di regolazione. Queste regioni possono giocare un ruolo di promotore della trascrizione, aumentando o diminuendo il reclutamento del macchinario proteico necessario alla trascrizione; possono soprattutto

giocare un ruolo a livello post-trascrizionale, aumentando il reclutamento di ribosomi per la traduzione, favorendo il legame di fattori promotori della traduzione o inibitori della stessa; possono inoltre agire nel controllo della produzione proteica formando strutture secondarie che aumentano o riducono la stabilità del messaggero, oppure mostrano o nascondono siti di legame per fattori di regolazione.

Le istruzioni per espletare queste funzioni sono “scritte” nella sequenza stessa del 5’UTR. Anche nella porzione trascritta e non tradotta a valle dello stop codon del messaggero sono presenti elementi che sono in grado di regolare l’espressione genica.

Il database UTRSITE (<http://www2.ba.itb.cnr.it/UTRSite/index.php>) raccoglie informazioni riguardo questi motivi regolatori e la loro funzione.

Alcuni esempi di questi elementi regolatori sono gli IRE (Iron Responsive Elements), particolari strutture a forcina (figura 3.12) localizzate nel 5’UTR o 3’UTR di diversi mRNA codificanti per proteine coinvolte nel metabolismo del ferro. Gli IRE sono riconosciuti da proteine note come Iron Regulatory Proteins (IRPs), che controllano il livello di traduzione dell’mRNA e la sua stabilità in funzione della quantità di ferro presente nella cellula.

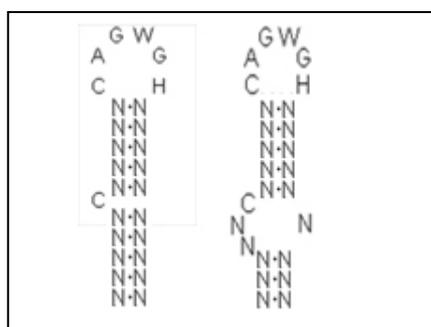


Figura 3.12: motivo tipico degli IRE

Un altro esempio è rappresentato dagli elementi di controllo del messaggero della Amiloid Precursor Protein (APP_SCE). Un’ aumentato livello di espressione di APP viene spesso ritrovato nei

pazienti affetti da morbo di Alzheimer. La stabilità del messaggero dell'APP è controllata da un elemento di 29 nucleotidi nel 3'UTR, che si è visto essere capace di interagire con diverse proteine citoplasmatiche. Se non vi sono proteine che interagiscono con questo elemento, esso agisce destabilizzando il messaggero ($t_{1/2}=4h$); Il legame di proteine all'elemento aumenta in maniera drammatica l'emivita del messaggero ($t_{1/2}=10h$; Zaidi e Malter, 1994). Questo elemento di 29 nucleotidi è altamente conservato nei mammiferi, sia per quanto riguarda la sequenza primaria (figura 3.13), sia per quanto riguarda la posizione (200 nucleotidi a valle dello stop codon).

Human (Y00264)	: UCUCUUUACAUUUUU-GGUCUCUAUACUACA
Rhesus (M58727)	:-.....C.....
Guinea Pig (X97631)	:-.....C.....
Mouse (X59379)	:C.-.....C.U....
Pig (NM_214372)	:CC.....C.....
Chicken (NM_204308)	:C-A....U..C.....
Consensus	: UCUCUUUACAUUUUU-GGUCUCUACACUACA

Figura 3.13: la sequenza primaria degli elementi APP_SCE è altamente conservata nei mammiferi.

Un ulteriore esempio dei meccanismi attraverso i quali gli elementi al 5'UTR possono interagire con l'espressione genica è rappresentato dagli elementi IRES (Internal Ribosome Entry Site). Questi elementi rappresentano un sistema alternativo al convenzionale ribosome scanning dipendente dal 5'-cap. Questi elementi sono caratterizzati da un motivo comune a "Y" che precede l'inizio di traduzione.

La nostra ipotesi è che le mutazioni trovate nei soggetti affetti delle famiglie 0, 1, 2 e 3 vadano a modificare un elemento regolatore dell'espressione genica, aumentando o diminuendo l'espressione del gene *ANKRD26*.

Un esempio nel quale un meccanismo di questo tipo è in grado di causare una patologia è rappresentato dalla Trombocitemia essenziale (MIM#187950). Kondo e colleghi descrissero una famiglia con trombocitemia, nella quale gli affetti presentavano livelli sierici di

TPO (trombopoietina) significativamente elevati, se comparati ai livelli dei sani. Essi analizzarono il gene *THPO*, codificante per la trombopoietina, ed identificarono una delezione di un singolo nucleotide al 5' del gene, 43 nucleotidi a monte dell' ATG. Il gruppo effettuò dei saggi funzionali di espressione e dimostrò che la mutazione causava un significativo aumento della produzione di proteina (Kondo et al, 1998).

In maniera indipendente, un altro gruppo, ha individuato un'altra mutazione nel 5'UTR del gene *THPO*, una trasversione da G a T in posizione 516 del messaggero; Essi dimostrarono che linee cellulari trasfettate con il cDNA di *THPO* mutato erano in grado di produrre una quantità di TPO 8 volte maggiore di linee cellulari trasfettate con cDNA wild-type (Ghilardi et al, 1999).

La regione al 5'UTR del gene *ANKRD26* nella quale abbiamo individuato le mutazioni è abbastanza conservata. E' infatti difficile avere un dato preciso sulla conservazione delle regioni trascritte e non tradotte, poiché possono cambiare a seconda del tessuto in cui il gene viene espresso, e per alcune specie non sono disponibili informazioni sufficienti. Utilizzando il tool UTRblast, disponibile nel database UTRSITE (<http://www.ba.itb.cnr.it/BIG/Blast/utrblast.cgi>), abbiamo "blastato" la porzione di 5'UTR da -113 a -139 (una regione di 27 nucleotidi che comprendono le mutazioni individuate, Figura 3.14) dall'ATG, utilizzando come database di riferimento UTRef, che raccoglie le sequenze di circa 1 milione di regioni 5' e 3' UTR.

Da questa analisi risultano 51 sequenze con elevata similitudine (tabella 3.6), 25 collocate in 5'UTRs (tra cui *ANKRD26* di *Homo Sapiens* e *Pan Troglodytes*, che hanno identità del 100%) e 26 in 3'UTRs.

La porzione più conservata è di 17 nt, e va dalla posizione -137 a -118, dove sono localizzate le mutazioni da noi individuate, presente in molte specie e geni: a parte gli eterologhi del gene *ANKRD26*, è presente in molte specie nel 3'UTR del Vascular Endothelial Growth Factor A (*VEGFA*). Nella figura 3.14 sono schematizzati i risultati.

5HSAR010257	5'UTR	in Homo sapiens ankyrin repeat domain 26 (ANKR...	54	2e-07
5PTRR021737	5'UTR	in PREDICTED: Pan troglodytes hypothetical LOC...	54	2e-07
3PTRR013277	3'UTR	in PREDICTED: Pan troglodytes phosphodiesteras...	42	0.001
5MMUR053086	5'UTR	in Mus musculus nuclear receptor subfamily 2, ...	34	0.22
5BTAR008823	5'UTR	in PREDICTED: Bos taurus ankyrin repeat domain...	34	0.22
5BTAR003698	5'UTR	in Bos taurus ankyrin repeat domain 26 (ANKRD2...	34	0.22
5HSAR018429	5'UTR	in Homo sapiens coiled-coil domain containing ...	32	0.89
5HSAR008350	5'UTR	in PREDICTED: Homo sapiens similar to Coiled-c...	32	0.89
5HSAR000698	5'UTR	in Homo sapiens coiled-coil domain containing ...	32	0.89
5HSAR000694	5'UTR	in Homo sapiens coiled-coil domain containing ...	32	0.89
5PTRR034392	5'UTR	in PREDICTED: Pan troglodytes hypothetical pro...	32	0.89
3RNOR020873	3'UTR	in Rattus norvegicus vascular endothelial grow...	30	3.5
3RNOR020872	3'UTR	in Rattus norvegicus vascular endothelial grow...	30	3.5
3RNOR020871	3'UTR	in Rattus norvegicus vascular endothelial grow...	30	3.5
3RNOR020867	3'UTR	in Rattus norvegicus vascular endothelial grow...	30	3.5
3RNOR020865	3'UTR	in Rattus norvegicus vascular endothelial grow...	30	3.5
3MMUR059224	3'UTR	in Mus musculus vascular endothelial growth fa...	30	3.5
3MMUR059222	3'UTR	in Mus musculus vascular endothelial growth fa...	30	3.5
3MMUR059215	3'UTR	in Mus musculus vascular endothelial growth fa...	30	3.5
3MMUR059207	3'UTR	in Mus musculus vascular endothelial growth fa...	30	3.5
3MMUR059205	3'UTR	in Mus musculus vascular endothelial growth fa...	30	3.5
3MMUR059202	3'UTR	in Mus musculus vascular endothelial growth fa...	30	3.5
3HSAR033407	3'UTR	in Homo sapiens vascular endothelial growth fa...	30	3.5
3HSAR033401	3'UTR	in Homo sapiens vascular endothelial growth fa...	30	3.5
3HSAR033397	3'UTR	in Homo sapiens vascular endothelial growth fa...	30	3.5
3HSAR033391	3'UTR	in Homo sapiens vascular endothelial growth fa...	30	3.5
3HSAR033388	3'UTR	in Homo sapiens vascular endothelial growth fa...	30	3.5
3HSAR033385	3'UTR	in Homo sapiens vascular endothelial growth fa...	30	3.5
3HSAR032760	3'UTR	in Homo sapiens vascular endothelial growth fa...	30	3.5
3MMUR014740	3'UTR	in PREDICTED: Macaca mulatta vascular endothel...	30	3.5
3CRER02454	3'UTR	in Chlamydomonas reinhardtii strain CC-503 cw9...	30	3.5
3PPAR006215	3'UTR	in Physcomitrella patens subsp. patens predict...	30	3.5
3ECAR002181	3'UTR	in PREDICTED: Equus caballus similar to myeloi...	30	3.5
3CLUR021945	3'UTR	in Canis lupus familiaris vascular endothelial...	30	3.5
3CLUR021944	3'UTR	in Canis lupus familiaris vascular endothelial...	30	3.5
3CLUR021943	3'UTR	in Canis lupus familiaris vascular endothelial...	30	3.5
5PTRR008411	5'UTR	in PREDICTED: Pan troglodytes poly(rC) binding...	30	3.5
5PTRR008410	5'UTR	in PREDICTED: Pan troglodytes poly(rC) binding...	30	3.5
5PTRR008397	5'UTR	in PREDICTED: Pan troglodytes poly(rC) binding...	30	3.5
5PTRR008395	5'UTR	in PREDICTED: Pan troglodytes poly(rC) binding...	30	3.5
5PTRR008394	5'UTR	in PREDICTED: Pan troglodytes poly(rC) binding...	30	3.5
5PTRR008392	5'UTR	in PREDICTED: Pan troglodytes poly(rC) binding...	30	3.5
5PTRR008388	5'UTR	in PREDICTED: Pan troglodytes poly(rC) binding...	30	3.5
5PTRR008387	5'UTR	in PREDICTED: Pan troglodytes poly(rC) binding...	30	3.5
5MMUR042437	5'UTR	in PREDICTED: Macaca mulatta similar to ankyri...	30	3.5
5MMUR042435	5'UTR	in PREDICTED: Macaca mulatta similar to ankyri...	30	3.5
5MMUR006703	5'UTR	in PREDICTED: Macaca mulatta similar to poly(r...	30	3.5
5MMUR006689	5'UTR	in PREDICTED: Macaca mulatta similar to poly(r...	30	3.5
5ATHR023512	5'UTR	in Arabidopsis thaliana RAP2.10 (related to AP...	30	3.5
5OSAR014188	5'UTR	in Oryza sativa (japonica cultivar-group) Os08...	30	3.5

Tabella 3.6: risultato dell'UTRblast di 27 nt nell'intorno delle mutazioni identificate nel 5'UTR del gene ANKRD26.

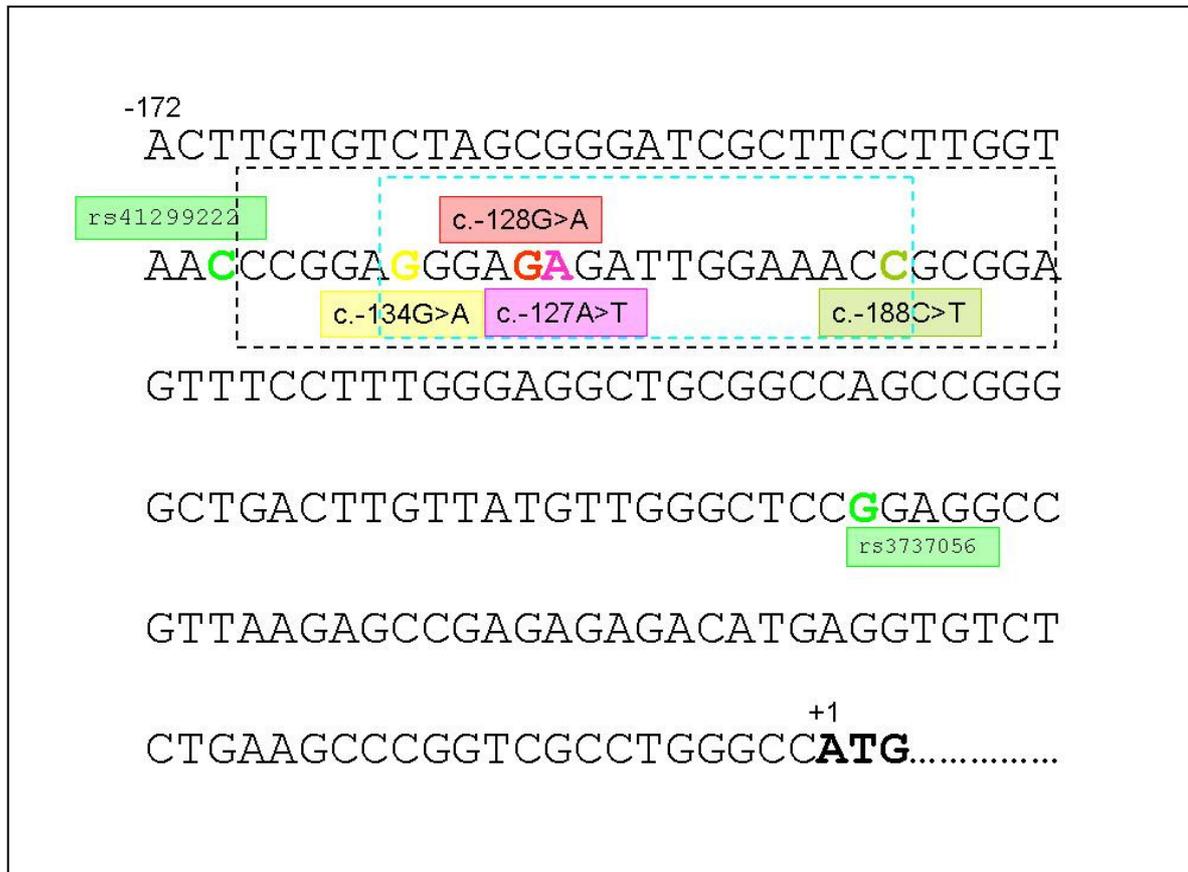


Figura 3.14: la regione 5'UTR del gene ANKRD26. Le mutazioni individuate sono indicate; sono indicati anche gli unici due SNP presenti nei database (rs41299222 e rs3737056). Il riquadro nero tratteggiato indica la regione utilizzata per effettuare l'UTRblast; il riquadro tratteggiato azzurro indica la regione maggiormente conservata.

Dal punto di vista del folding del mRNA, è possibile che sostituzioni nucleotidiche causino delle modificazioni nella struttura secondaria del messaggero, portando, eventualmente, alla formazione o distruzione di loop e forcine che possono essere funzionali dal punto di vista della produzione proteica.

Al fine di avere indicazioni sull'eventuale modificazione del folding del messaggero dovuta alle mutazioni identificate nei pazienti, abbiamo utilizzato software che sono in grado di predire la struttura secondaria del mRNA a partire dalla struttura primaria. Abbiamo utilizzato il programma mFOLD (<http://mfold.bioinfo.rpi.edu/>) per

analizzare i cambiamenti conformazionali a livello del solo 5'UTR, mentre il software VIENNA (Vienna RNA package v. 1.7, <http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>) è stato utile per ottenere informazioni sul folding di tutto il messaggero.

Benchè le informazioni ottenute siano da considerarsi puramente delle indicazioni, è interessante notare come nella situazione wild-type il 5'UTR è posizionato in una struttura complessa separata dal resto del messaggero e vicino al 3'UTR. Inoltre la stringa di basi nella quale sono collocate le mutazioni, forma una struttura a stem-loop particolarmente esposta (figure 3.15 e 3.16).

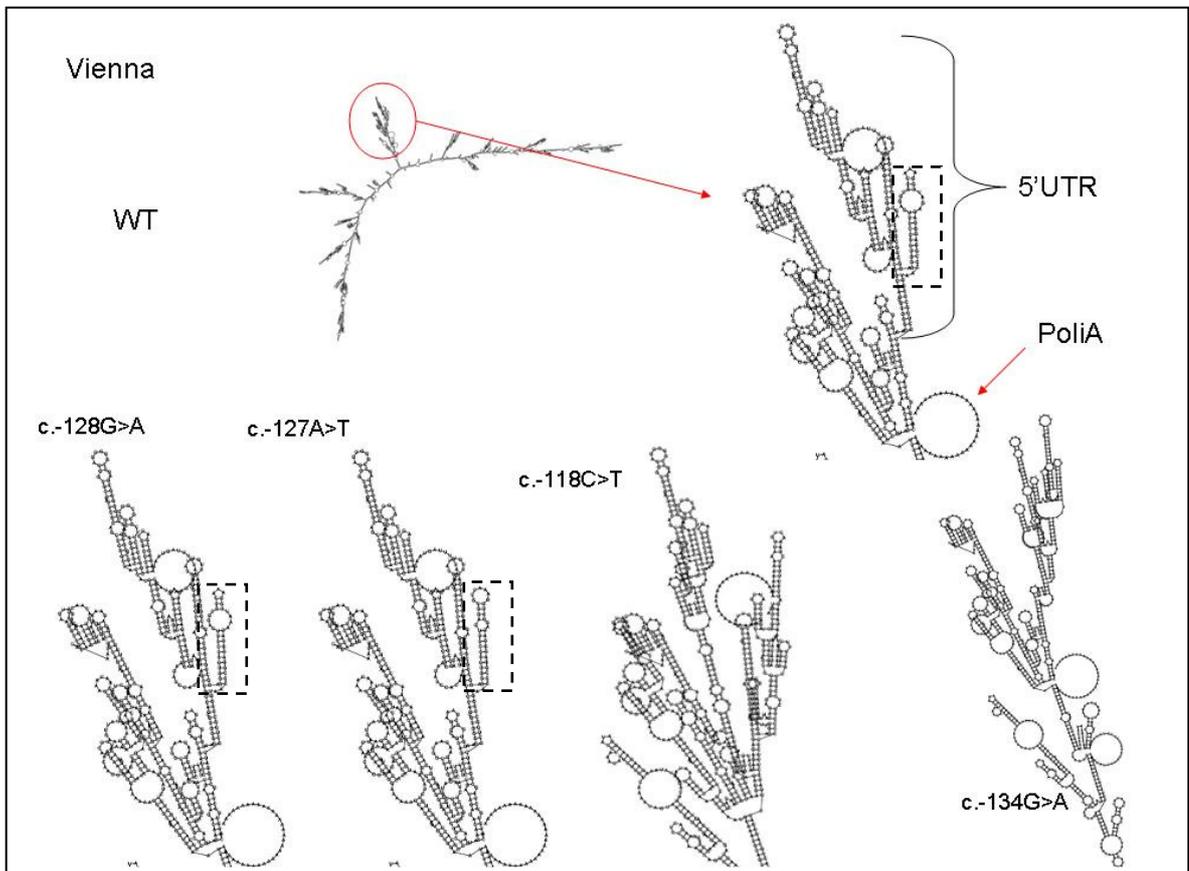


Figura 3.15: struttura secondaria del messaggero del gene ANKRD26 secondo la predizione del software Vienna; in alto è mostrato il wild-type, in basso le modificazioni causate dalle singole mutazioni. Il rettangolo tratteggiato mette in evidenza la struttura a stem-loop nella quale sono situate le mutazioni. Questa struttura è modificata ma mantenuta nelle mutazioni -128 e -127, mentre scompare completamente nelle mutazioni -118 e -134.

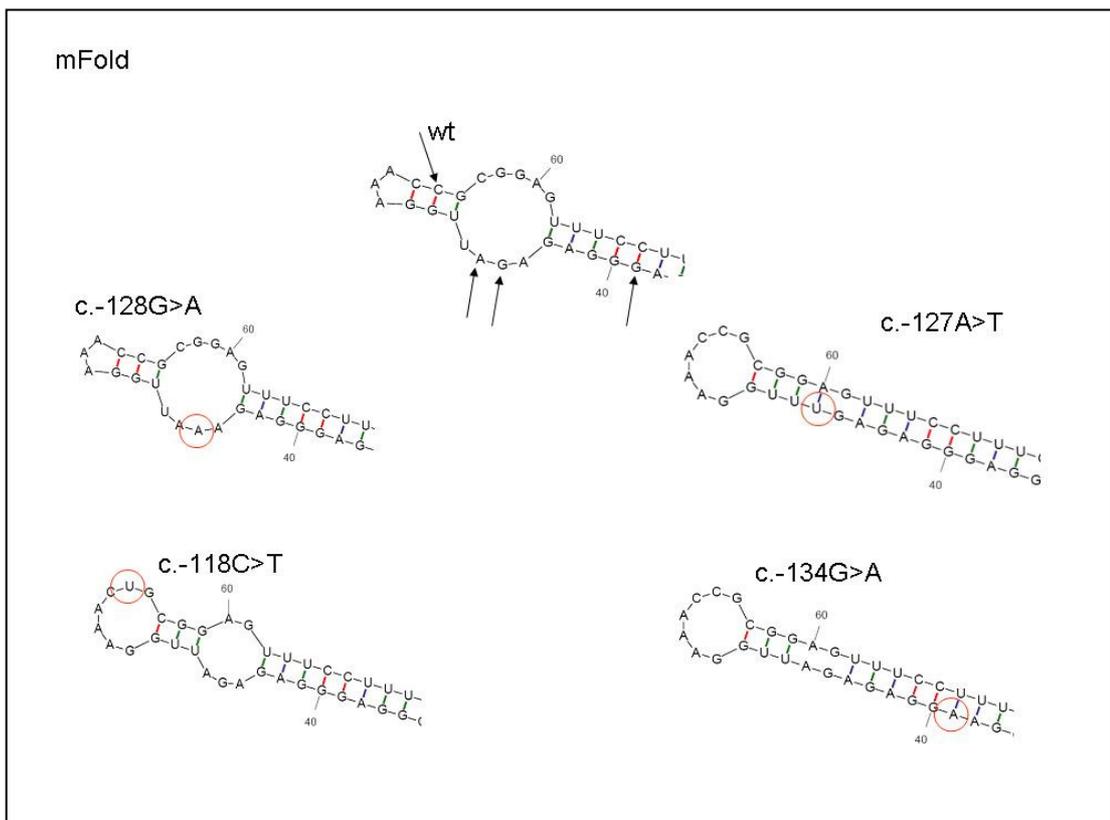


Figura 3.16: la conformazione della struttura a stem-loop nella quale sono localizzate le mutazioni, secondo la predizione del software mFold. Solo per la mutazione -128 la conformazione non viene modificata, cambia però un nucleotide esposto nel loop.

E' interessante notare come la sostituzione di un singolo nucleotide sia in grado di influire nella formazione della struttura secondaria; per quanto riguarda le mutazioni -127 e -128, l'effetto a livello conformazionale non sembra essere importante (figura 3.15), ma se si osserva la figura 3.16 si può notare che causano modificazioni nella sotto-struttura in cui sono collocate. Le mutazioni in posizione -118 e -134 sembrano a prima vista causare cambiamenti maggiori, causando un riarrangiamento importante a livello della struttura in cui è localizzato il 5'UTR del gene (figure 3.15 e 3.16).

L'insieme di questi dati ci ha spinto ad approfondire lo studio del gene *ANKRD26*.

Le informazioni a riguardo di questo gene non sono molte, benché si pensi sia coinvolto nei processi apoptotici e nell'obesità.

Questo gene è recentemente stato descritto come il capostipite di una famiglia di geni, POTE, implicati nel controllo dell'apoptosi (Xiu et al, 2009). Lo stesso gruppo aveva già descritto un modello murino di obesità e gigantismo generato in seguito a KnockDown del gene *Ankrd26* (Bera et al, 2008; figura 3.17).

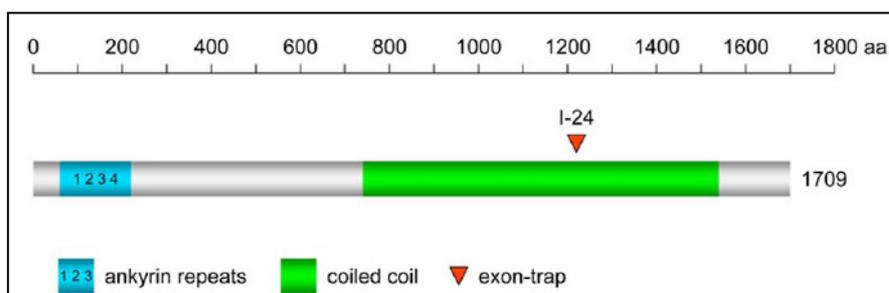


Figura 3.17: struttura del gene *Ankrd26* murino e schematica rappresentazione della sua inattivazione mediante la tecnica di exon-trapping.

La famiglia dei geni POTE composta da proteine che contengono ankirin repeats ed eliche simili a spectrina con domini N-terminali ricchi in cisteina. Il gene *ANKRD26*, benchè sia considerato il precursore di questa famiglia genica, manca del dominio ricco in cisteine (Bera et al, 2008).

Il gene è espresso nella maggior parte dei tessuti umani, compresi cervello, fegato, reni, cuore e tessuti embrionali, ma non è stata identificata la funzione del prodotto genico.

Al fine di comprendere la funzione del gene, il gruppo di Bera e colleghi ha generato un topo KO per il gene. Il topo così ingegnerizzato si è dimostrato essere un modello per obesità e gigantismo (Bera et al, 2008).

Il gene *ANKRD26* è composto da 34 esoni ed il suo messaggero misura 7,5 Kb. La proteina è costituita da 1709 aminoacidi e contiene 5 ANK repeats e cinque domini coiled coil; sono note tre isoforme della proteina.

Al fine di valutare l'effetto delle mutazioni da noi identificate nel 5'UTR del gene *ANKRD26*, abbiamo deciso di effettuare degli studi di espressione in un sistema con gene reporter.

La nostra intenzione era quella di testare la capacità del 5'UTR, mutato e wild-type, di modificare l'espressione genica. Abbiamo pensato di effettuare una prima serie di esperimenti nei quali saggiare la capacità in se del 5'UTR di influenzare l'espressione del reporter luciferasi. Se questi esperimenti non si fossero dimostrati efficaci, la nostra intenzione era quella di studiare, oltre al 5'UTR, anche il promotore endogeno del gene *ANKRD26*.

Abbiamo deciso di utilizzare il vettore pGL3-basic (Promega), che ha il gene della *firefly luciferase* ma non ha promotore che ne attivi la trascrizione (figura 3.18).

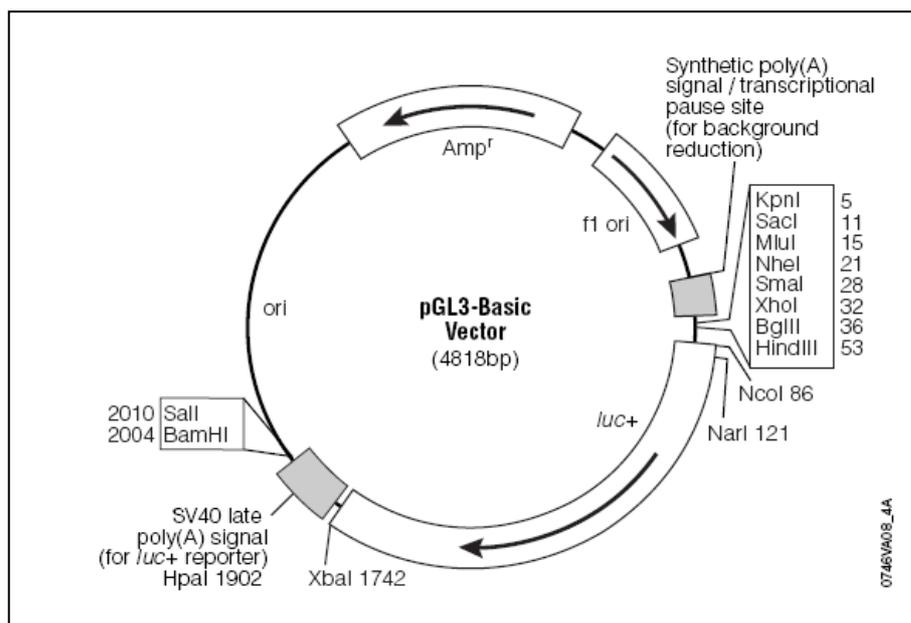


Figura 3.18: il vettore pGL3-basic (Promega).

Questo vettore è studiato apposta per il Dual-Luciferase Reporter Assay (Promega). L'assay prevede la cotrasfezione di due vettori, uno che porta la *firefly* luciferasi (*luc*) e la sequenza regolatoria da testare, ed uno che porta la *renilla* luciferasi (*ren*) posta sotto il controllo di un promotore costitutivo. Grazie a questa metodica è possibile

normalizzare i dati e, di conseguenza, comparare i risultati di esperimenti diversi.

Utilizzando il sito di restrizione NcoI del vettore, che è localizzato esattamente sull'ATG della luciferasi, abbiamo clonato il 5'UTR di *ANKRD26* esattamente in frame con l'ORF del gene reporter, ottenendo 5 cloni diversi, il wild-type e quattro vettori ognuno con una delle mutazioni identificate.

Ad oggi abbiamo effettuato solo pochi esperimenti nelle linee cellulari Hek293 e K562.

La scelta della linea cellulare Hek293 (cellule derivanti da un carcinoma renale) è dovuta prevalentemente alla estrema maneggevolezza di queste cellule; abbiamo quindi deciso di effettuare i primi esperimenti su questa linea cellulare per testare la funzionalità del nostro sistema.

Le cellule K562 sono una linea che deriva da una eritrioleucemia, sono cioè "bloccate" ad uno stato del lineage emopoietico multipotente, che è una via di mezzo tra il lineage eritroblastico e megacariocitoblastico. Il trattamento di queste cellule con esteri del forbolo (PMA) ne favorisce il differenziamento in senso megacariocitario.

Gli esperimenti hanno dato risultati incoraggianti, evidenziando un aumento di espressione genica nelle cellule trasfettate con i vettori con le mutazioni rispetto al vettore con la sequenza 5'UTR wild-type (grafici dal 3.2 al 3.6).

I valori di luminescenza della luciferasi sono stati normalizzati con la luminescenza della renilla divisi per il valore ottenuto dal wild-type.

Nella tabella 3.7 sono riportati i valori di luminescenza.

	Hek 293 12 h	Hek 293 24 h	K562 6h	K562 12h	K562 24
WT	1	1	1	1	1
-118	0,872438084	1,033605579	1,097946	1,333231	1,096109
-127	1,024605996	1,098542565	1,4926	1,49906	0,701139
-128	1,490406534	1,387193796	-	1,552093	0,894714
-134	2,849434391	2,949334445	1,425678	6,819132	6,290171

Tabella 3.7: Valori di luminescenza normalizzati e rapportati al wild-type. I valori di luminescenza ottenuti per la firefly luciferasi sono stati divisi per quelli ottenuti per la renilla luciferasi. E' stata calcolata una media dei valori e questa media è stata divisa per il valore medio del wild-type. Il wild-type di conseguenza è sempre uguale a 1, mentre i valori per le mutazioni variano a seconda che siano maggiori del wild-type (>1) o minori (<1). Un valore di 1,5 per esempio, indicherà che vi è il 50% in più di luminescenza rispetto al wild-type.

Le letture di luminescenza sono state effettuate a 12 e 24 ore, sia per le Hek293 che per le K562, per tutte le mutazioni. La lettura a 6 ore è stata rilevata solo per le mutazioni -118, -127 e -134; per la mutazione -128 non è stato possibile rilevare il dato di luminescenza, probabilmente a causa della scarsa espressione del reporter a 6 ore.

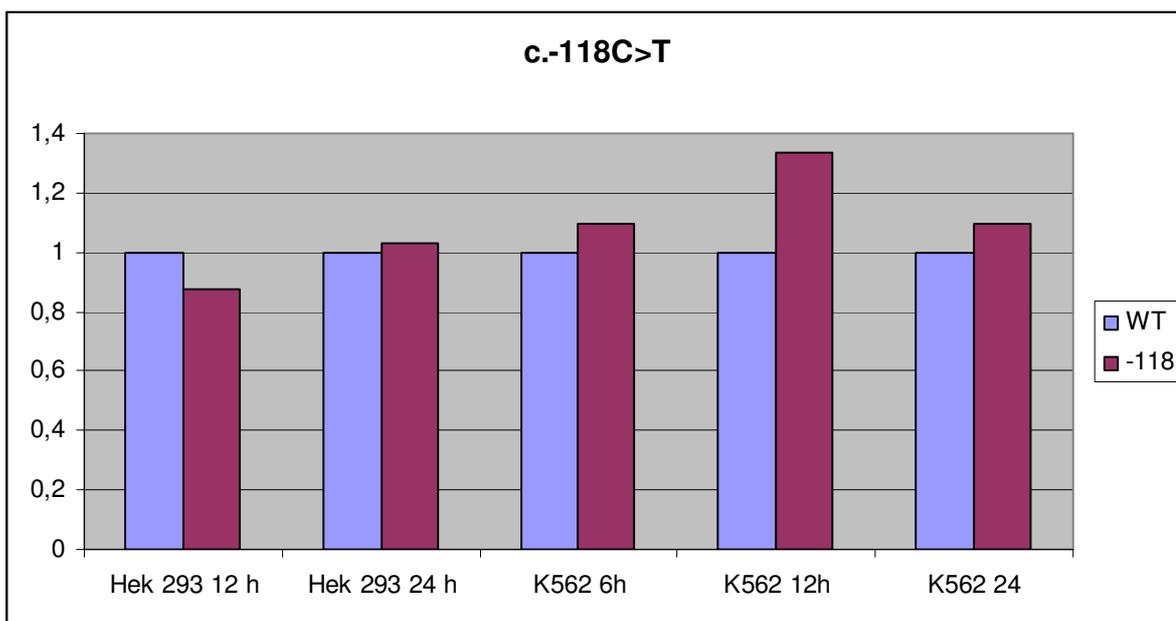


Grafico 3.2: valori di luminescenza per la mutazione c.-118C>T

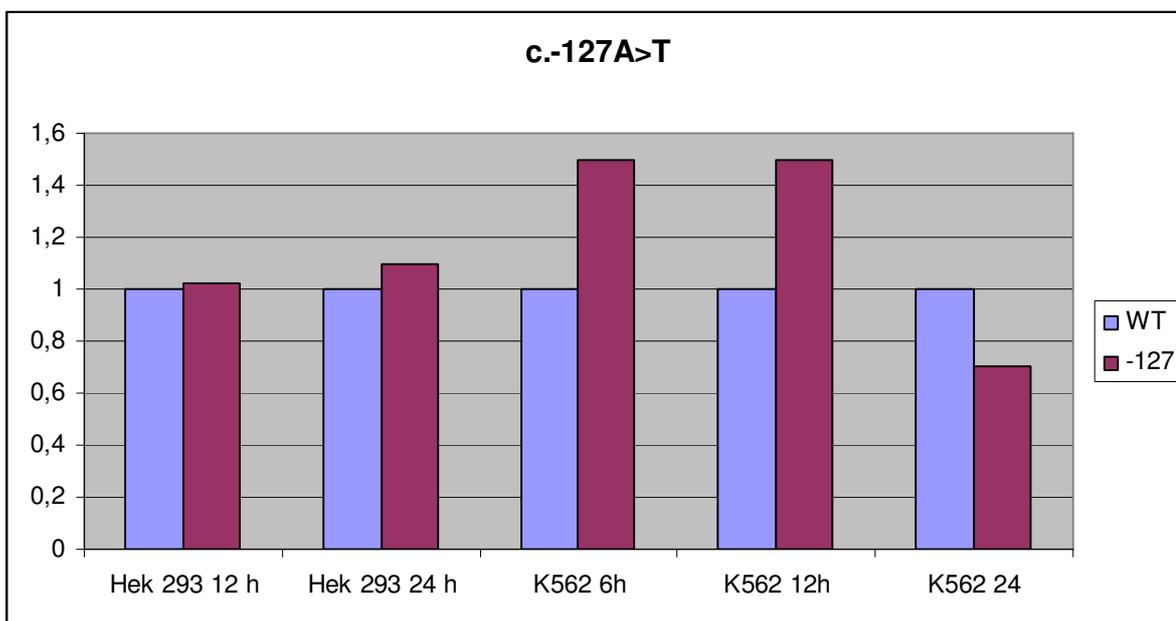


Grafico 3.3: valori di luminescenza per la mutazione c.-127A>T

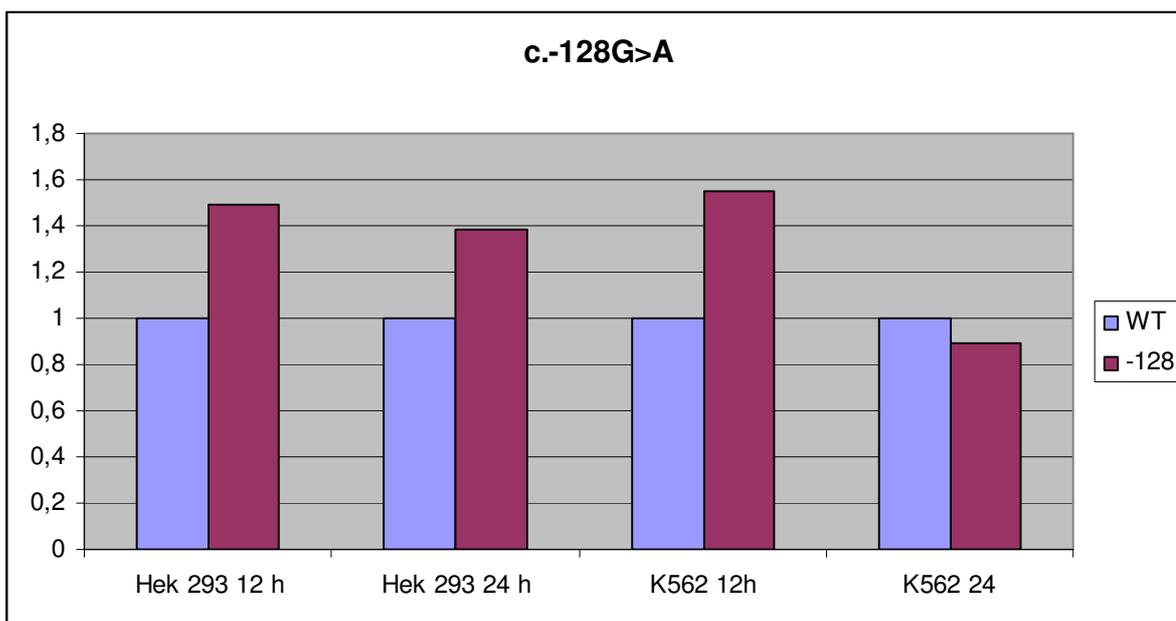


Grafico 3.4: valori di luminescenza per la mutazione c.-128G>A

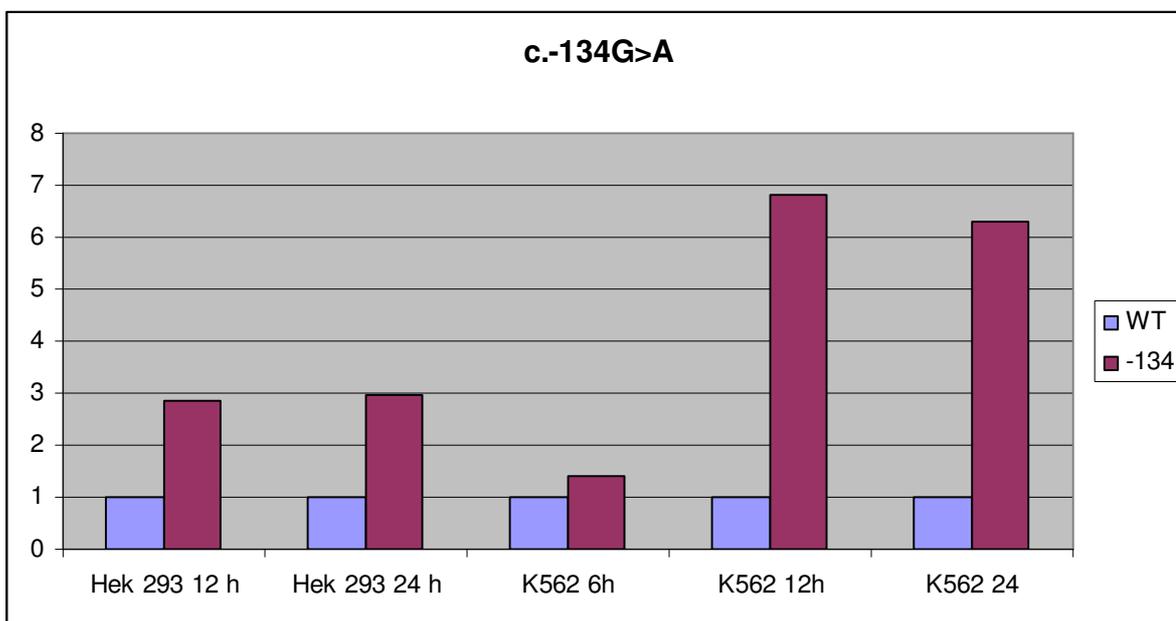


Grafico 3.5: valori di luminescenza per la mutazione c.-134G>A

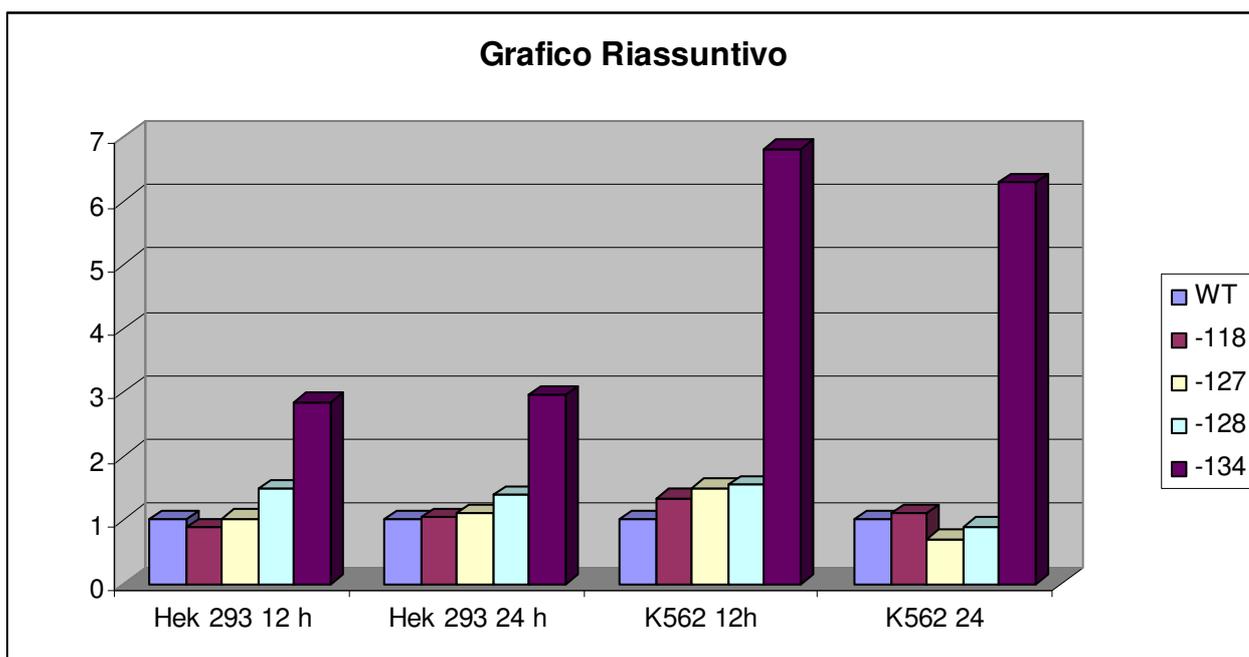


Grafico 3.6: *i valori di luminescenza delle diverse mutazioni rispetto al wild-type.*

Questi risultati, benchè parziali, indicano un' aumentata espressione del gene reporter nelle cellule trasfettate con i vettori con le mutazioni rispetto al wild-type. In particolare il dato di lettura della luminescenza a 12 ore nelle cellule K562 mostra un aumento significativo di luminescenza, e quindi di espressione per tutte le mutazioni, superando il dato del wild-type di più del 30%. E' particolarmente significativo il dato ottenuto dalla mutazione in posizione -134. Questa mutazione causa un aumento della luminescenza di quasi 3 volte nelle cellule Hek293, e di più di 6 volte nelle K562, rispetto al wild-type.

Il dato più significativo è sicuramente quello ottenuto dalla trasfezione delle cellule K562, poiché hanno caratteristiche simili ai megacariociti, cellule dalle quali si formano le piastrine, e che si pensa siano coinvolte nella patogenesi di trombocitopenia.

Sarà necessario effettuare ulteriori esperimenti al fine di confermare questo dato e dimostrare che le mutazioni identificate causano un

aumento di espressione del gene *ANKRD26*, portando alla manifestazione di piastrinopenia.

Per prima cosa, tramite esperimenti di rtPCR, valuteremo se il 5'UTR del gene *ANKRD26* viene trascritto con la luciferasi nel sistema cellulare K562. Questo dato è fondamentale al fine di comprendere il meccanismo attraverso il quale le mutazioni causano aumento di espressione. Se infatti il 5'UTR viene trascritto, è probabile che le mutazioni distruggano un sito di regolazione negativa della traduzione, oppure creino un sito di regolazione positiva. Viceversa, se il 5'UTR non è presente nel messaggero della luciferasi, il meccanismo d'azione più probabile è che il sito in cui sono localizzate le mutazioni agisca come promotore della trascrizione, con le mutazioni che causano un'alterazione in un sito regolatore, creando un sito attivatorio o distruggendo un sito inibitorio.

Al fine di confermare l'effetto delle mutazioni sull'espressione genica sarà necessario trasfettare assieme al 5'UTR mutato e wild-type, anche il promotore "fisiologico" del gene *ANKRD26*; poiché non vi sono in letteratura informazioni riguardo l'effettiva localizzazione e dimensione del promotore del gene, prevediamo di effettuare gli esperimenti clonando il 5'UTR del gene assieme a 1000 basi, circa, a monte di esso, al fine di avere la relativa certezza che vi sia compreso il promotore. La nostra ipotesi è che questo esperimento consenta di ottenere un dato ancora più significativo, che indichi un chiaro aumento di espressione genica.

Al fine di stabilire una netta correlazione tra le mutazioni e l'aumento di espressione, sarà necessario inoltre, clonare nei vettori la sola regione del 5'UTR nella quale sono localizzate le mutazioni. Se anche con questi costrutti sarà possibile vedere un aumento significativo dell'espressione genica, ciò significherà che le mutazioni colpiscono un sito fondamentale per il controllo d'espressione, e sarà possibile

effettuare dei saggi che consentano di individuare gli eventuali fattori che legano il sito mutato.

Per raggiungere questo obiettivo ci riproponiamo di effettuare un Band-Shift Assay. Questa metodica consiste nell'ibridare la regione di acido nucleico di interesse con un estratto proteico nucleare, e di far correre su gel di poliacrilamide il prodotto della reazione di ibridazione. In questa maniera sarà possibile identificare eventuali anomalie nella migrazione su gel, dovute al legame di proteine che riconoscono siti specifici nell'acido nucleico target. In seguito sarà possibile isolare la proteina legata alla sequenza target e sequenziarla, giungendo così all'identificazione di fattori coinvolti nella patogenesi della piastrinopenia, e, eventualmente, ottenendo informazioni fondamentali per futuri approcci terapeutici.

Riteniamo che i dati da noi ottenuti siano fortemente indicativi del coinvolgimento del gene *ANKRD26* nella patogenesi di trombocitemia autosomica dominante al locus *THC2*.

Benché sia stata identificata una mutazione nel gene *MASTL* nel medesimo locus, in una unica famiglia, riteniamo che il gene responsabile di piastrinopenia al locus sul cromosoma 10p11.2-12 sia *ANKRD26*. La nostra ipotesi è che, come già detto, la mutazione identificata da Drachman e colleghi, rappresenti una variante rara familiare, e non una mutazione causativa. Il nostro gruppo di ricerca, invece, ha individuato mutazioni nel gene *ANKRD26* in 4 famiglie distinte con trombocitopenia autosomica dominante al locus *THC2*, queste mutazioni sono localizzate in una ristretta stringa di nucleotidi nel 5'UTR del gene, ed i nostri dati indicano un coinvolgimento diretto delle mutazioni nell'aumento di espressione del gene in linee cellulari con caratteristiche megacariocitarie ed eritrocitarie.

I dati ottenuti dal nostro gruppo indicano un coinvolgimento delle mutazioni individuate nel gene *ANKRD26* nella patogenesi della

patologia. Altri dati egualmente importanti sono stati ottenuti dai gruppi di ricerca che collaborano con noi a questo progetto.

Infatti l'analisi di sequenza del 5'UTR del gene *ANKRD26* in 80 pazienti, sia sporadici che familiari, affetti da forme ereditarie di trombocitopenia, per i quali non era stato possibile effettuare una diagnosi definitiva, ha consentito di individuare mutazioni in 16 di essi. Sono state identificate 9 diverse mutazioni nel 5'UTR del gene (5 nuove mutazioni oltre alle 4 identificate in precedenza), tutte collocate nella regione da -113 a -134 nucleotidi a monte dell'ATG. Questo dato conferma fortemente l'ipotesi di coinvolgimento del gene *ANKRD26* nella patologia, infatti 16 pazienti su 80 corrispondono al 20%, collocando il gene tra i principali responsabili di trombocitopenia autosomica dominante.

Un dato rilevante è giunto dall'analisi del profilo di espressione del gene *ANKRD26*. L'analisi semi-quantitativa tramite rtPCR su RNA da diverse linee cellulari umane, tra le quali linee linfoblastiche, fibroblastiche ed emopoietiche, ha mostrato che il gene è ampiamente espresso, anche se a livelli basali, nella maggior parte dei tessuti analizzati. In particolare, l'analisi nei megacariociti ha mostrato che l'espressione del gene è molto debole, rilevabile solo con 35 cicli di amplificazione, suggerendo l'ipotesi che il gene possa essere down-regolato o silenziato nel processo di differenziamento megacariocitario.

Questo dato potrebbe spiegare il trend di luminescenza nel tempo da noi individuato. Infatti se si osservano i dati nei grafici, si nota che si ha un picco massimo di luminescenza a 12 ore nelle K562, che poi tende a scemare, come se si raggiungesse un plateau, o mancasse un qualche fattore necessario al mantenimento dell'espressione nel tempo. Questo dato potrebbe anche spiegare la differenza di attivazione genica riscontrata per le diverse mutazioni. E' plausibile infatti che in presenza del macchinario proteico necessario, l'effetto delle mutazioni sia simile o quantomeno paragonabile.

Al fine di verificare queste ipotesi abbiamo deciso di effettuare un dual-luciferase assay in cellule progenitrici megacariocitarie (K562 e Dami), trattandole con PMA (esteri del forbole, che sono in grado di indurre il differenziamento megacariocitario). Se effettivamente il gene *ANKRD26* viene down-regolato o silenziato nel corso del differenziamento megacariocitario, la nostra ipotesi è che, trattando le cellule trasfettate con PMA, si veda una diminuzione della produzione di luciferasi nel wild-type ed un aumento o una costante produzione nelle cellule trasfettate coi in vettori nei quali il gene luciferasi è posto sotto il controllo dei 5'UTR mutati.

Questi esperimenti verranno effettuati nei prossimi mesi; speriamo di ottenere risultati che ci permettano di affermare senza ombra di dubbio che le mutazioni individuate sono responsabili di piastrinopenia autosomica dominante.

Come già accennato, mutazioni al 5'UTR del gene *THPO* causano trombocitemia, dovuta ad un aumento di espressione del gene. Inoltre il topo KO per il gene *ANKRD26* è un modello di obesità e gigantismo, fenotipi completamente diversi dalla trombocitopenia. E' plausibile quindi che il gene abbia un ruolo nel controllo della crescita cellulare, e che una sua iper-espressione possa causare una riduzione della crescita o del differenziamento.

Riteniamo che le mutazioni identificate nella ristretta stringa di oligonucleotidi al 5'UTR del gene abbiano un ruolo nell'attivazione o inattivazione della differenziazione megacariocitaria, e quindi nella produzione di piastrine. La nostra ipotesi è che queste mutazioni agiscano su siti specifici riconosciuti da fattori specifici del lineage megacariocitario, causando piastrinopenia e non altri fenotipi a livello sistemico.

3.3 CONCLUSIONI

Il percorso che porta all'identificazione di un gene malattia è spesso lungo e dispendioso. Benché la recente introduzione delle nuove tecnologie di *Next Generation Sequencing*, stia rendendo il sequenziamento genomico il *gold standard* del futuro della ricerca di geni malattia, l'approccio del gene candidato per posizione si è dimostrato essere quello più "fruttifero" dell'era post-genomica. Questo approccio prevede la mappatura del gene in un determinato *locus* cromosomico, che deve essere il più ristretto possibile, al fine di consentire una rapida analisi dei geni. Tuttavia non è sempre possibile giungere all'identificazione di *loci* piccoli, e quindi facilmente analizzabili. Una risorsa fondamentale a tal fine è rappresentata dalle famiglie di cui si può disporre. Per malattie diffuse e con impatto relativamente moderato sulla salute degli affetti, spesso è possibile "collezionare" DNA di più membri affetti di molteplici famiglie; ciò agevola l'analisi di linkage, fornendo un alto numero di meiosi informative al fine del calcolo del LOD score ed aumentando la possibilità di individuare ricombinazioni utili al fine di ottenere una regione critica minima.

Riuscire a stabilire un linkage ad uno stesso locus in due o più famiglie nelle quali ricorre la stessa malattia, consente di affermare con ragionevole certezza che il gene responsabile della patologia si trova nel locus individuato.

Quando si ha a che fare con malattie estremamente rare, il cui effetto sulla salute dei pazienti è molto grave, se non mortale, può risultare complicato avere a disposizione più famiglie sulle quali effettuare l'analisi, e può quindi essere difficile, se non impossibile, riuscire a confermare il linkage ad uno stesso locus per due o più famiglie.

Inoltre la disponibilità di DNA da più famiglie aumenta la probabilità di individuare la mutazione causativa nel gene responsabile. La funzione di un gene può essere infatti alterata in diversi modi. Se la mutazione colpisce una tripletta codificante, causando un cambio amminoacidico o la creazione di uno stop codon, eccetera, allora con tutta probabilità sarà

relativamente semplice individuare la mutazione (anche in una sola famiglia) con l'analisi di sequenza. Se, invece, la mutazione colpisce regioni regolatorie del gene, come il promotore, gli enhancers, gli UTRs, segnali di splicing intronici, eccetera, potrebbe non essere sufficiente una sola famiglia per identificare la mutazione con l'analisi di sequenza.

Come la mia esperienza dimostra, le famiglie sono la risorsa più importante negli studi di linkage. Il fallimento nella ricerca del gene responsabile della Sindrome di Mungan, può essere ricondotto proprio al fatto che si aveva una sola famiglia a disposizione. D'altro canto, il successo nell'identificazione del gene della THC2 è dovuto proprio al numero di famiglie a disposizione: con molta probabilità non si sarebbe tenuto conto di una mutazione a livello del 5'UTR se si avesse avuto a disposizione una sola famiglia!!!

BIBLIOGRAFIA

- **Abecasis** GR, Cherny SS, Cookson WO and Cardon LR: Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* (2002) 30:97-101
- **Auricchio** A, Brancolini V, Casari G *et al*: The locus for a novel syndromic form of neuronal intestinal pseudoobstruction maps to Xq28. *Am J Hum Genet* 1996; **58**: 743-748.
- **Bader–Meunier** B, Proulle V, Trichet C, Debray D, Gabolde M, Yvart J, Dreyfus M. Misdiagnosis of chronic thrombocytopenia in childhood. *J Pediatr Hematol Oncol* 2003;25:548–52.
- **Balduini** A, d'Apolito M, D. Arcelli D, Conti V, Pecci A, Pietra D, Danova M, Benvenuto F, Perotti C, Zelante L, Volinia S, Balduini CL, Savoia A. Cord blood in vitro expanded CD41+ cells: identification of novel components of megakaryocytopoiesis. *J Thromb Haemost* 2006;4:848–60
- **Balduini** A, Malara A, Pecci A, Badalucco S, Bozzi V, Pallotta I, Noris P, Torti M, Balduini CL. Proplatelet formation in heterozygous Bernard–Soulier syndrome type Bolzano. *J Thromb Haemost* 2009;7:478–84.
- **Balduini** A, Pallotta I, Malara A, Lova P, Pecci A, Viarengo G, Balduini CL, Torti M. Adhesive receptors, extracellular proteins and myosin IIA orchestrate proplatelet formation by human Mks. *Thromb Haemost* 2008;6:1900–7.
- **Balduini** CL, Noris P, Belletti S, Spedini P, Gamba GI.. In vitro and in vivo effects of DDAVP on platelet function. *Haematologica* 1999;84:891–6.
- **Bassotti** G, Villanacci V, Fisogni S, et al. Enteric glial cells and their role in gastrointestinal motor abnormalities: introducing the neuro-gliopathies. *World J Gastroenterol* 2007; 13: 4035–41.
- **Bauer**, M.; Dieterich, W.; Ehnis, T.; Schuppan, D. : Complete primary structure of human collagen type XIV (undulin). *Biochim. Biophys. Acta* 1354: 183-188, 1997.

- **Bellucci S** (1997) Megakaryocytes and inherited thrombocytopenias. *Bailliere Clin Haematol* 10:149–162
- **Benson G**. "Tandem repeats finder: a program to analyze DNA sequences" *Nucleic Acids Research* (1999) Vol. 27, No. 2, pp. 573-580.).
- **Bera Tapan K., Xiu-Fen Liu, Masanori Yamada, Oksana Gavrilova, Eva Mezey, Lino Tessarollo, Miriam Anver, Yoonsoo Hahn, Byungkook Lee, and Ira Pastan**. A model for obesity and gigantism due to disruption of the *Ankrd26* gene. 270–275 _ *PNAS* _ January 8, 2008 _ vol. 105 _ no. 1
- **Bott L, Boute O, Mention K, Vinchon M, Boman F, Gottrand F**: Congenital idiopathic intestinal pseudo-obstruction and hydrocephalus with stenosis of the aqueduct of Sylvius. *Am J Med Genet* 2004; 130: 84–87.
- **Buckingham KJ, Lee C, Bigham AW, Tabor HK, Dent KM, Huff CD, Shannon PT, Jabs EW, Nickerson DA, Shendure J, Bamshad MJ**. Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet*. 2010 Jan;42(1):13-4.
- **Bussel JB, Provan D, Shamsi T, Cheng G, Psaila B, Kovaleva L, Salama A, Jenkins JM, Roychowdhury D, Mayer B, Stone N, Arning M**. Effect of eltrombopag on platelet counts and bleeding during treatment of chronic idiopathic thrombocytopenic purpura: a randomised, double-blind, placebo-controlled trial. *Lancet* 2009;373:641–8.
- **Camilleri M, Phillips SF**: Acute and chronic intestinal pseudo-obstruction. *Adv Intern Med* 1991; **36**: 287-306.
- **Collins SM**. The immunomodulation of enteric neuromuscular function: implications for motility and inflammatory disorders. *Gastroenterology* 1996; 111: 1683–9.
- **Coulie B, Camilleri M**: Intestinal pseudo-obstruction. *Ann Rev Med* 1999; **50**: 37-55.
- **De Giorgio R, Camilleri M**. Human enteric neuropathies: morphology and molecular pathology. *Neurogastroenterol Motil* 2004; 16: 515–31.
- **De Giorgio R, Guerrini S, Barbara G, Cremon C, Stanghellini V, Corinaldesi R**. New insights into human enteric neuropathies. *Neurogastroenterol Motil* 2004; 16(Suppl. 1): 143–7.
- **De Giorgio R, Sarnelli G, Corinaldesi R, Stanghellini V**: Advances in our

understanding of the pathology of chronic intestinal pseudo-obstruction. *Gut* 2004; **53**: 1549-1552.

- **De Giorgio R**, Seri M, Cogliandro R, et al. Analysis of candidate genes for intrinsic neuropathy in a family with chronic idiopathic intestinal pseudo-obstruction. *Clin Genet* 2001; 59: 131–3.
- **De Giorgio R**, Stanghellini V, Barbara G *et al*: Primary enteric neuropathies underlying gastrointestinal motor dysfunction. *Scand J Gastroenterol* 2000; **35**: 114-121.
- **Deglicerti A**, De Giorgio R, Cefle K, et al. A novel locus for syndromic chronic idiopathic intestinal pseudoobstruction maps to chromosome 8q23-q24. *Eur J Hum Genet* 2007; 15: 889–97.
- **Derry JMJ**, Ochs HD, Francke U (1994) Isolation of a novel gene mutated in Wiskott-Aldrich syndrome. *Cell* 78: 635–644
- **Deutsch V**, Bitan M, Friedmann Y, Eldor A, Vlodaysky I. Megakaryocyte maturation is associated with expression of the CXC chemokine connective tissue-activating peptide CTAP III. *Br J Haematol* 2000;111:1180–9.
- **Di Lorenzo C**: Pseudo-obstruction: Current approaches. *Gastroenterology* 1999; **116**: 980-987.
- **Di Nardo G**, Blandizzi C, Volta U, Colucci R, Stanghellini V, Barbara G, Del Tacca M, Tonini M, Corinaldesi R, De Giorgio R. Review article: molecular, pathological and therapeutic features of human enteric neuropathies. *Aliment Pharmacol Ther.* 2008 Jul;28(1):25-42. Epub 2008 Apr 11.
- **Dowton SB**, Beardsley D, Jamison D, Blattner S, Li FP (1985) Studies of a familial platelet disorder. *Blood* 65:557–563
- **Drachman JG**, Jarvik GP, Mehaffey MG. Autosomal dominant thrombocytopenia: incomplete Mk differentiation and linkage to human chromosome 10. *Blood* 2000;96:118–25. Project Description – Cited Literature 33
- **Filipovich AH**, Stone JV, Tomany SC, Ireland M, Kollman C, Pelz CJ, Casper JT, Cowan MJ, Edwards JR, Fasth A, Gale RP, Junker A, Kamani NR, Loechelt BJ, Pietryga DW, Ringdén O, Vowels M, Hegland J, Williams AV, Klein JP, Sobocinski KA, Rowlings PA, Horowitz MM.

Impact of donor type on outcome of bone marrow transplantation for Wiskott–Aldrich syndrome: collaborative study of the International Bone Marrow Transplant Registry and the National Marrow Donor Program. *Blood* 2001;97:1598–603.

- **FitzPatrick** DR, Strain L, Thomas AE *et al*: Neurogenic chronic idiopathic intestinal pseudo-obstruction, patent ductus arteriosus, and thrombocytopenia segregating as an X-linked recessive disorder. *J Med Genet* 1997; **34**: 666-669.
- **Flejou** JF: Barrett's oesophagus: from metaplasia to dysplasia and cancer. *Gut* 2005; 54 (Suppl 1): i6– i12.
- **Gandhi** MJ, Cummings CL, Drachman JG. FLJ14813 missense mutation: a candidate for autosomal dominant thrombocytopenia on human chromosome 10. *Hum Hered* 2003;55:66–70.
- **Gargiulo** A, Auricchio R, Barone MV, et al. Filamin A is mutated in X-linked chronic idiopathic intestinal pseudoobstruction with central nervous system involvement. *Am J Hum Genet* 2007; 80: 751–8.
- **Gautschi** OP, Frey SP, Zellweger R. Bone morphogenetic proteins in clinical applications. *ANZ J Surg* 2007;77:626–31.
- **Geddis** AE, **Balduini** CL. Diagnosis of immune thrombocytopenic purpura in children. *Curr Opin Hematol* 2007;14:520–5.
- **Ghevaert** C, Salsmann A, Watkins NA, Schaffner–Reckinger E, Rankin A, Garner SF, Stephens J, Smith GA, Debili N, Vainchenker W, de Groot PG, Huntington JA, Laffan M, Kieffer N, Ouwehand WH. A nonsynonymous SNP in the ITGB3 gene disrupts the conserved membrane–proximal cytoplasmic salt bridge in the alphaIIb beta3 integrin and cosegregates dominantly with abnormal proplatelet formation and macrothrombocytopenia. *Blood* 2008;111:3407–14.
- **Ghilardi**, N.; Wiestner, A.; Kikuchi, M.; Ohsaka, A.; Skoda, R. C. : Hereditary thrombocythaemia in a Japanese family is caused by a novel point mutation in the thrombopoietin gene. *Brit. J. Haemat.* 107: 310-316, 1999.
- **Gohda** F, Uchiumi H, Handa H, Matsushima T, Tsukamoto N, Morita K, Amagai H, Murakami M, Murakami H, Nojima Y, Karasawa M. Identification of inherited macrothrombocytopenias based on mean

platelet volume among patients diagnosed with idiopathic thrombocytopenia. *Thromb Res* 2007;119:741–6.

- **Goyal RK, Hirano I.** The enteric nervous system. *New Engl J Med* 1996; 25: 1106–15.
- **Gresele P, Falcinelli E, Giannini S, D'Adamo P, D'Eustacchio A, Corazzi T, Mezzasoma AM, Di Bari F, Guglielmini G, Cecchetti L, Noris P, Balduini CL, Savoia A.** Dominant inheritance of a novel integrin beta3 mutation associated with a hereditary macrothrombocytopenia and platelet dysfunction in two Italian families. *Haematologica* 2009;94:663–9.
- **Gusella, J. F.,** Gibbons, K.; Hobbs, W.; Heft, R.; Anderson, M.; Rashtchian, R.; Folstein, S.; Wallace, P.; Conneally, P. M.; Tanzi, R. : The G8 locus linked to Huntington's disease. (1984) *Am. J. Hum. Genet.* 36: 139S.
- **Hirano M, Silvestri G, Blake D, et al.** Mitochondrial neurogastrointestinal encephalomyopathy (MNGIE): clinical, biochemical, and genetic features of an autosomal recessive mitochondrial disorder. *Neurology* 1994; 44: 721–7.
- **Hu FZ, Preston RA, Post JC et al:** Mapping of a gene for severe pediatric gastroesophageal reflux to chromosome 13q14. *JAMA* 2000; 284: 325–334.
- **Hugnot, J.-P.; Salinas, M.; Lesage, F.; Guillemare, E.; de Weille, J.; Heurteaux, C.; Mattei, M.-G.; Lazdunski, M. :** Kv8.1, a new neuronal potassium channel subunit with specific inhibitory properties towards Shab and Shaw channels. *EMBO J.* 15: 3322-3331, 1996.
- **Innis MS, Myambo KB, Gelfand DH, Brown MA.** DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Biotechnology.* 1992;24:6-10
- **Iolascon A, Perrotta S, Amendola G, Altomare M, Bagnara GP, Del Vecchio ME, Savoia A.** Familial dominant thrombocytopenia: clinical, biologic, and molecular studies. *Pediatr Res* 1999;46:548–52.
- **Johnson HJ, Gandhi MJ, Shafizadeh E, Langer NB, Pierce EL, Paw BH, Gilligan DM, Drachman JG.** In vivo inactivation of MASTL kinase results in thrombocytopenia. *Exp Hematol* 2009;37:901–8.

- **Kenmochi**, N.; Suzuki, T.; Uechi, T.; Magoori, M.; Kuniba, M.; Higa, S.; Watanabe, K.; Tanaka, T. : The human mitochondrial ribosomal protein genes: mapping of 54 genes to the chromosomes and implications for human disorders. *Genomics* 77: 65-70, 2001.
- **Kim**, H.-S.; Nagalla, S. R.; Oh, Y.; Wilson, E.; Roberts, C. T., Jr.; Rosenfeld, R. G. : Identification of a family of low-affinity insulin-like growth factor binding proteins (IGFBPs): characterization of connective tissue growth factor as a member of the IGFBP superfamily. *Proc. Nat. Acad. Sci.* 94: 12981-12986, 1997.
- **King** S, Germeshausen M, Strauss G, Welte K, Ballmaier M. Congenital amegakaryocytic thrombocytopenia: a retrospective clinical analysis of 20 patients. *Br J Haematol* 2005;131:636–44.
- **Kondo**, T.; Okabe, M.; Sanada, M.; Kurosawa, M.; Suzuki, S.; Kobayashi, M.; Hosokawa, M.; Asaka, M. : Familial essential thrombocythemia associated with one-base deletion in the 5-prime-untranslated region of the thrombopoietin gene. *Blood* 92: 1091-1096, 1998.
- **Krishnamurthy** S, Schuffler MD: Pathology of neuromuscular disorders of the small intestine and colon. *Gastroenterology* 1987; **93**: 610-639.
- **Kruglyak** L, Daly MJ, Reeve-Daly MP, Lander ES: Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996; **58**:1347-1363.
- **Kunishima** S, Kobayashi R, Itoh TJ, Hamaguchi M, Saito H. Mutation of the beta1-tubulin gene associated with congenital macrothrombocytopenia affecting microtubule assembly. *Blood* 2009;113:458–61.
- **Kuter** DJ, Bussel JB, Lyons RM, Pullarkat V, Gernsheimer TB, Senecal FM, Aledort LM, George JN, Kessler CM, Sanz MA, Liebman HA, Slovick FT, de Wolf JT, Bourgeois E, Guthrie TH Jr, Newland A, Wasser JS, Hamburg SI, Grande C, Lefrère F, Lichtin AE, Tarantino MD, Terebelo HR, Viallard JF, Cuevas FJ, Go RS, Henry DH, Redner RL, Rice L, Schipperus MR, Guo DM, Nichol JL. Efficacy of romiplostim in patients with chronic immune thrombocytopenic purpura: a double-blind randomised controlled trial. *Lancet* 2008;371:395–403.

- **Lander, E. S.,** Botstein, D. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. *Science* 236: 1567-1570, 1987.
- **Lathrop GM,** Lalouel JM, Julier C, Ott J: Strategies for multilocus linkage analysis in humans. *Proc Natl Acad Sci USA* 1984; 81: 3443–3446.
- **Levay K,** Slepak VZ. Tescalcin is an essential factor in megakaryocytic differentiation associated with Ets family gene expression. *J Clin Invest* 2007; 117:2672–83.
- **Locatelli F,** Rossi G, Balduini C. Hematopoietic stem–cell transplantation for the Bernard–Soulier syndrome. *Ann Intern Med* 2003;138:79.
- **Long MW,** Heffner CH, Williams JL, Peters C, Prochownik EV. Regulation of megakaryocyte phenotype in human erythroleukemia cells. *J Clin Invest* 1990;85:1072–84.
- **Malagelada JR,** Camilleri M, Stanghellini V: Manometric diagnosis of gastrointestinal motility disorders. New York: Thieme, 1986.
- **Morison IM,** Cramer Bordé EM, Cheesman EJ, Cheong PL, Holyoake AJ, Fichelson S, Weeks RJ, Lo A, Davies SM, Wilbanks SM, Fagerlund RD, Ludgate MW, da Silva Tatley FM, Coker MS, Bockett NA, Hughes G, Pippig DA, Smith MP, Capron C, Ledgerwood EC. A mutation of human cytochrome c enhances the intrinsic apoptotic pathway but causes only thrombocytopenia. *Nat Genet* 2008;40:387–9.
- **Mueller J,** Werner M, Stolte M. Barrett’s esophagus: histopathologic definitions and diagnostic criteria. *World J Surg.* 2004 Feb;28(2):148-54.
- **Mungan, Z.;** Akyuz, F.; Bugra, Z.; Yonal, O.; Ozturk, S.; Acar, A.; Cevikbas, U. : Familial visceral myopathy with pseudo-obstruction, megaduodenum, Barrett's esophagus, and cardiac abnormalities. *Am. J. Gastroenterol.* 98: 2556-2560, 2003.
- **Najean Y,** **Lecompte T** (1990) Genetic thrombocytopenia with autosomal dominant transmission: a review of 54 cases. *Br J Haematol* 74:203- 208
- **Nishino I,** Spinazzola A, Hirano M: Thymidine phosphorylase gene mutations in MNGIE, a human mitochondrial disorder. *Science* 1999;

283: 689-692.

- **Nishino I**, Spinazzola A, Papadimitriou A, et al. Mitochondrial neurogastrointestinal encephalopathy: an autosomal recessive disorder due to thymidine phosphorylase mutations. *Ann Neurol* 2000; 47: 792–800.
- **Noris P**, Pecci A, Di Bari F, Di Stazio MT, Di Pumpo M, Ceresa IF, Arezzi N, Ambaglio C, Savoia A, Balduini CL. Application of a diagnostic algorithm for inherited thrombocytopenias to 46 consecutive patients. *Haematologica* 2004;89:1219–25.
- **Nurden AT**, Nurden P. Inherited disorders of platelets: an update. *Curr Opin Hematol* 2006;13:157–62.
- **O’Connell JR**, Weeks DE: PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 1998; 63: 259–266.
- **O’Brien JJ**, Spinelli SL, Tober J, Blumberg N, Francis CW, Taubman MB, Palis J, Seweryniak KE, Gertz JM, Phipps RP. 15–deoxy–delta12,14–PGJ2 enhances platelet production from Mks. *Blood* 2008;112:4051–60.
- **Ohtani K**, Suzuki Y, Eda S et al: Molecular cloning of a novel human collectin from liver (CL-L1). *J Biol Chem* 1999; 274: 13681 –13689.
- **Pallotta I**, Lovett M, Rice W, Kaplan DL, Balduini A. Bone marrow osteoblastic niche: a new model to study physiological regulation of megakaryopoiesis. *PLoS One*. 2009 Dec 21;4(12):e8359.
- **Pecci A**, Canobbio I, Balduini A, Stefanini L, Cisterna B, Marseglia C, Noris P, Savoia A, Balduini CL, Torti M. Pathogenetic mechanisms of hematological abnormalities of patients with MYH9 mutations. *Hum Mol Genet* 2005;14:3169–78.
- **Pecci A**, Malara A, Badalucco S, Bozzi V, Torti M, Balduini CL, Balduini A. Megakaryocytes of patients with MYH9–related thrombocytopenia present an altered proplatelet formation. *Thromb Haemost* 2009;102:90–6.
- **Peters JH**, Hagen JA, DeMeester SR: Barrett’s esophagus. *J Gastrointest Surg* 2004; 8: 1– 17.
- **Pingault V**, Girard M, Bondurand N *et al*: SOX10 mutations in chronic

intestinal pseudo-obstruction suggest a complex physiopathological mechanism. *Hum Genet* 2002; **111**: 198-206.

- **Ramel S.** Barrett's Esophagus: Model of neoplastic progression. *World J Surg.* 2003 Sep;27(9):1009-1013.
- **Rimoin DL** and Emery EH Principle and Practice of Medical Genetics IV Edition (Churchill Livingstone, 2002).
- **Romeo G e Devoto M** Dal DNA alle malattie ereditarie, II Edizione (Casa Editrice Ambrosiana, 1993).
- **Ronzoni L, Bonara P, Rusconi D, Frugoni C, Libani I, Cappellini MD.** Erythroid differentiation and maturation from peripheral CD34+ cells in liquid culture: Cellular and molecular characterization. *Blood Cells Mol Dis* 2008;40:148–55.
- **Roper EC, Gibson A, McAlindon ME et al:** Familial visceral neuropathy: a defined entity? *Am J Med Genet* 2005; **137**: 249-254.
- **Sabri S, Foudi A, Boukour S, Franc B, Charrier S, Jandrot–Perrus M, Farndale RW, Jalil A, Blundell MP, Cramer EM, Louache F, Debili N, Thrasher AJ, Vainchenker W.** Deficiency in the Wiskott–Aldrich protein induces premature proplatelet formation and platelet production in the bone marrow compartment. *Blood* 2006;108:134–40.
- **Sanders KM, Koh SD, Ward SM.** Interstitial cells of Cajal as pacemakers in the gastrointestinal tract. *Annu Rev Physiol* 2006; 68: 307–43.
- **Sarna SK.** Are interstitial cells of Cajal plurifunction cells in the gut? *Am J Physiol* 2007; 294: G372–90.
- **Savidge TC, Sofroniew MV, Neunlist M.** Starring roles for astroglia in barrier pathologies of gut and brain. *Lab Invest* 2007; 87: 731–6.
- **Savoia A, Del Vecchio M, Totaro A, Perrotta S, Amendola G, Moretti A, Zelante L, Iolascon A.** An autosomal dominant thrombocytopenia gene maps to chromosomal region 10p. *Am J Hum Genet* 1999;65:1401–5.
- **Savoia A, Noris P, Perrotta S, Punzo F, Rocco DD, Oostra BA, Balduini CL.** Absence of CYCS mutations in a large Italian cohort of patients with inherited thrombocytopenias of unknown origin. *Platelets* 2009;20:72–3.
- **Savoia A, Zatterale A, Del Principe D, JoenjeH(1996)** Fanconi anaemia in Italy: high prevalence of complementation group A in two geographic clusters. *Hum Genet* 97:599–603

- **Schellenberg EG**, Wasylenki D, Webster CD, Goering P. (1992); A review of arrests among psychiatric patients. *Int J Law Psychiatry*. 15(3):251-64.
- **Schuffler MD**, Pope CE: Studies of idiopathic intestinal pseudoobstruction. II. Hereditary hollow visceral myopathy: family studies. *Gastroenterology* 1977; **73**: 339-344.
- **Seve M**, Chimienti F, Devergnas S, Favier A: In silico identification and expression of SLC30 family genes: an expressed sequence tag data mining strategy for the characterization of zinc transporters' tissue expression. *BMC Genomics* 2004; 5: 32.
- **Shaw, J.**; Lichter, P.; Driesel, A. J.; Williamson, R.; Chamberlain, S. (1990) Regional localisation of the Friedreich ataxia locus to human chromosome 9q13-q21.1. *Cytogenet. Cell Genet.* 53: 221-224.
- **Sobel E**, Lange K: Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. *Am J Hum Genet* 1996; 58: 1323– 1337.
- **Stanghellini V**, Cogliandro R, De Giorgio R *et al*: Natural history of chronic idiopathic intestinal pseudo-obstruction in adults: a single center study. *Clin Gastroenterol Hepatol* 2005; **3**: 449-458.
- **Stanghellini V**, Corinaldesi R, Barbara L: Pseudo-obstruction syndromes. *Baillieres Clin Gastroenterol* 1988; **2**: 225-254.
- **Strachan T e Read AP** *Genetica Umana Molecolare* III Edizione (2004 Garland Publishing, 2006 UTET). Edizione italiana a cura di Gian Antonio Danieli.
- **Strippoli P**, Savoia A, Iolascon A, Tonelli R, Savino M, Giordano P, D'Avanzo M, et al (1998) Mutational screening of thrombopoietin receptor gene (c-mpl) in patients with congenital thrombocytopaenia and absent radii (TAR). *Br J Haematol* 103:311–314
- **Stroncek DF**, Rebutla P. Platelet transfusions. *Lancet* 2007;370:427–38.
- **Tanner MS**, Smith B, Lloyd JK: Functional intestinal obstruction due to deficiency of argyrophil neurones in the myenteric plexus. Familial syndrome presenting with short small bowel, malrotation, and pyloric hypertrophy. *Arch Dis Chil* 1976; **51**: 837-841.

- **Thiele H**, Nurnberg P: HaploPainter: a tool for drawing pedigrees with complex haplotypes. *Bioinformatics* 2005; 21: 1730– 1732.
- **Van Goethem G**, Schwartz M, Lofgren A, Dermaut B, Van Broeckhoven C, Vissing J: Novel POLG mutations in progressive external ophthalmoplegia mimicking mitochondrial neurogastrointestinal encephalomyopathy. *Eur J Hum Genet* 2003; **11**: 547-549.
- **Villa A**, Notarangelo L, Macchi P, Mantuano E, Cavagni G, Brugnani D, Strina D, et al (1995) X-linked thrombocytopenia and Wiskott-Aldrich syndrome are allelic diseases with mutations in the WASP gene. *Nat Genet* 9:414–417
- **Von Rahden BHA**, Stein HJ, Becker K, Liebermann-Meffert D, Siewert R. Heterotopic gastric mucosa of the esophagus: literature-review and proposal of a clinicopathologic classification. *Am J Gastroenterol.* 2004 Mar;99(3):543-51.
- **Xiu Fen Liu** & Tapan K. Bera & Lisa J. Liu & Ira Pastan: A primate-specific POTE-actin fusion protein plays a role in apoptosis. *Apoptosis* (2009) 14:1237–1244
- **Yoshizawa K**, Inaba K, Mannen H, Kikuchi T, Mizutani M, Tsuji S: Analyses of beta-1 syntrophin, syndecan 2 and gem GTPase as candidates for chicken muscular dystrophy. *Exp Anim* 2003; 52: 391– 396.
- **Zaidi Syed H.E.** and **Malter James S.**: Amyloid precursor protein mRNA stability is controlled by a 29-Base element in the 3'-Untranslated region. *Journal of Biological Chemistry* vol.269 n.39 pp.24007-24013 1994

INDICE

1. Introduzione	<i>pag</i>	1
1.1 L'analisi di Linkage	<i>pag</i>	2
1.2 L'identificazione di geni malattia	<i>pag</i>	12
1.3 Le malattie nelle famiglie analizzate	<i>pag</i>	17
1.3.1 La sindrome di Mungan	<i>pag</i>	17
1.3.2 Trombocitopenia Autosomica Dominante THC2	<i>pag</i>	27
2. Materiali e Metodi	<i>pag</i>	33
2.1 Le famiglie analizzate	<i>pag</i>	34
2.1.1 Sindrome di Mungan: Famiglia TKS	<i>pag</i>	34
2.1.2 Trombocitopenia Autosomica Dominante THC2	<i>pag</i>	38
2.2 Metodi di Laboratorio	<i>pag</i>	41
3. Risultati e Discussione	<i>pag</i>	63
3.1 Ricerca del gene responsabile della Sindrome di Mungan	<i>pag</i>	64
3.2 Trombocitopenia autosomica dominante THC2: esclusione del gene <i>MASTL</i> in 5 famiglie ed individuazione di mutazioni in un gene potenzialmente coinvolto nella patogenesi della sindrome	<i>pag</i>	80
3.3 Conclusioni	<i>pag</i>	110
Bibliografia	<i>pag</i>	112
Indice	<i>pag</i>	123