



UNIVERSITÀ
DEGLI STUDI
FIRENZE

**DOTTORATO DI RICERCA IN
SCIENZE BIOMEDICHE**
curriculum Biotechnologie Endocrinologiche
Molecolari e Rigenerative

CICLO XXVIII

COORDINATORE Prof. Persio Dello Sbarba

**CROMOTRIPSI
UN NUOVO MECCANISMO ALLA BASE DI
RIARRANGIAMENTI GENOMICI COMPLESSI**

Settore Scientifico Disciplinare MED/03

Dottorando

Dott. Aldesia Provenzano

Tutor

Prof.ssa Paola Romagnani

Co-Tutor

Prof.ssa Sabrina Rita Giglio

Coordinatore

Prof. Persio Dello Sbarba

Anni 2012/2015

ABSTRACT

Alterazioni strutturali nei cromosomi sono una causa frequente di tumori e malattie congenite. Recentemente, il fenomeno di “crisi cromosomica”, costituito da un insieme di decine o centinaia di riarrangiamenti genomici, localizzati in uno o pochi cromosomi, è stata descritta in cellule tumorali con il termine di cromotripsia. Una migliore comprensione e il riconoscimento di questo evento catastrofico hanno portato alle seguenti definizioni: cromotripsia e cromosomopatia.

La complessità di questi riarrangiamenti e le iniziali descrizioni in cellule tumorali, hanno fatto pensare inizialmente che si trattasse di un'anomalia acquisita. In realtà, alcuni pazienti sono stati riportati con cromotripsia costituzionale. Mediante array-CGH abbiamo identificato complessi riarrangiamenti cromosomici in tre pazienti, sul cromosoma 17, 12 e 4, rispettivamente, senza alcuna anomalia visibile al cariotipo. Nel primo caso il paziente presentava un quadro clinico complesso in cui coesistevano 3 disordini noti: sindrome di Potocki-Lupski, malattia di Charcot-Marie-Tooth 1A e Neurofibromatosi tipo 1. Il secondo caso presentava disabilità intellettiva e tratti autistici. Il terzo caso presentava fenotipo è abbastanza ascrivibile alla sindrome 1p36.

È interessante notare che la successione di delezioni e duplicazioni identificate nei nostri casi non erano in cluster su di una singola regione cromosomica, ma presenti su tutta la lunghezza dei cromosomi. Abbiamo quindi deciso di effettuare un sequenziamento dell'intero genoma (WGS) di ciascun paziente e mostreremo i risultati che forniscono ulteriori informazioni per definire meglio i meccanismi molecolari e le origini di questi riarrangiamenti altamente complessi

INDICE

1. INTRODUZIONE	2
1.1. ANATOMIA DEL GENOMA UMANO	2
1.2. RIARRANGIAMENTI STRUTTURALI DEL GENOMA	4
1.3. SEQUENZE MEDIATRICI DEI RIARRANGIAMENTI	9
1.4. MECCANISMI ALLA BASE DEI RIARRANGIAMENTI	
CROMOSOMICI STRUTTURALI	12
1.4.1. Meccanismi di riparazione omologa	14
1.4.2. Meccanismi non replicativi di riparazione non omologa	17
1.4.3. Meccanismi replicativi di riparazione non omologa	19
1.5. RIARRANGIAMENTI CROMOSOMICI COMPLESSI	24
1.6. METODI DI RICONOSCIMENTO DI VARIAZIONI STRUTTURALI	
DEL GENOMA	32
2. SCOPO DELLA TESI	39
3. MATERIALI E METODI	40
3.1. MATERIALI	40
3.2. METODI	45
4. RISULTATI	59
4.1 CASO C/996	59
4.2 CASO E/1021	73
4.3 CASO C/1642	80
5. DISCUSSIONE	85
6. CONCLUSIONI	94
7. BIBLIOGRAFIA	97

1. INTRODUZIONE

1.1 ANATOMIA DEL GENOMA UMANO

DNA a sequenza unica

Circa il 46% del genoma umano può essere considerato a sequenza unica, ovvero presente in singola copia.

La gran parte di questo DNA è costituito da geni codificanti proteine (circa 23.000) e soltanto il 2% dell'informazione genetica è rappresentato da mRNA maturi (1,2% di esoni tradotti, 0,8% di sequenza 5' e 3' non tradotte). Le sequenze introniche ammontano quindi a circa il 25% del genoma.

La densità dei geni è estremamente variabile sia fra i vari cromosomi che tra le diverse porzioni dello stesso cromosoma; non di rado si incontrano famiglie di geni, ovvero cluster di geni simili (omologhi), vicini tra loro e derivati per duplicazione da un singolo gene ancestrale. All'interno del genoma si trovano numerose copie non funzionali di geni, dette pseudogeni, originatesi per duplicazione o per retrotrascrizione e successivo inserimento in altre porzioni del genoma: nel primo caso, generalmente, si forma una nuova copia del gene (pseudogene non processato) vicina al gene funzionale, completa di esoni e introni, con mutazioni inattivanti accumulate che ne impediscono la sintesi di una proteina funzionale. La maggioranza degli pseudogeni sono invece processati, originando dalla retrotrascrizione dell'mRNA della copia funzionale del gene, con successiva inserzione in uno o più altri siti cromosomici.

I geni codificanti RNA strutturali sono circa 1.000-2.000 e a fronte di circa 23.000 geni codificanti proteine, si stima che siano prodotti fino a 100.000 differenti *noncoding* RNA (ncRNA o trascritti non codificanti), aventi un ruolo in diversi processi cellulari quali lo *splicing* degli introni, l'assemblaggio dei ribosomi, il trasporto dell'mRNA dal nucleo al citoplasma e la regolazione trascrizionale e post-trascrizionale (ncRNA antisense e RNA *interference*)^{1,2}.

DNA ripetitivo intersperso

Circa il 45% del genoma umano è mediamente ripetitivo, costituito da numerosissime copie disperse di un numero piuttosto limitato di elementi trasponibili. La maggior parte di questi elementi si è moltiplicata attraverso un processo di retrotrascrizione del DNA (retroelementi), mentre meno del 3% del genoma è rappresentato da trasposoni a DNA. Le sequenze LINE (*Long INterspersed Element*) occupano circa il 20% del genoma totale, mentre le SINE (*Short INterspersed Element*) ne occupano circa il 13%; circa l'8% del genoma è invece rappresentato da 400.000 retroelementi con LTR (*Long Terminal Repeat*) e da retrovirus endogeni HERV (*Human Endogenous RetroVirus*). La presenza di numerosi elementi trasponibili nel genoma non costituisce soltanto un rischio di malfunzionamento, ma anche un'opportunità di generare nuovi geni funzionalmente attivi^{3,4}.

DNA ripetitivo in tandem

Circa il 9-10% del genoma umano è rappresentato da regioni di DNA altamente ripetitivo, composto da centinaia di unità ripetute in tandem di lunghezza variabile, noto con il nome di DNA satellite, in quanto si dispone in bande distinte da quella principale se il DNA viene centrifugato in un gradiente di densità. Circa 50 megabasi (Mb) del DNA sono occupate dai centromeri, composti prevalentemente da DNA α satellite, costituito da un'unità base di 171 paia di basi (pb) organizzata in blocchi lunghi, mediamente 2 Mb, e riconosciuto dalle proteine centromeriche per la corretta segregazione dei cromatidi fratelli durante l'anafase di mitosi e meiosi.

I telomeri, ovvero le estremità delle braccia cromosomiche, sono invece composti da un esanucleotide (TTAGGG) ripetuto migliaia di volte per una lunghezza totale di circa 5-15 kilobasi (Kb), inversamente proporzionale all'età delle cellule. Le braccia corte dei cromosomi acrocentrici (13, 14, 15, 21 e 22) contengono oltre 200 copie di geni codificanti per le subunità 18S, 5S, 8S e 28S dei ribosomi (rRNA).

Circa il 3% del genoma è rappresentato da microsatelliti (o *Short Tandem Repeat*, STR), ovvero sequenze di 1-6 pb ripetute decine o anche centinaia di volte, abbastanza omogeneamente distribuite lungo tutto il genoma. Un 1% del genoma è costituito da minisatelliti (o *Variable Number of Tandem Repeats*, VNTR), la cui unità ripetuta è compresa tra 10 e 100 pb⁵.

1.2 RIARRANGIAMENTI STRUTTURALI DEL GENOMA

I riarrangiamenti genomici strutturali sono variazioni della struttura fisica dei cromosomi, che in alcuni casi possono determinare difetti congeniti, ritardo mentale, aumentato rischio di infertilità o aborto spontaneo. Queste anomalie strutturali possono comparire *de novo*, possono essere ereditate direttamente da genitori che ne sono portatori o possono essere il risultato di ulteriori riarrangiamenti favoriti da un particolare background genomico presente in uno dei genitori⁶ (Figura 1).

Le anomalie strutturali cromosomiche possono essere causate dalla rottura di uno o più cromosomi, seguita da un'anomala riunione dei frammenti cromosomici, o da *crossing-over* ineguali⁵. Una singola rottura all'interno di un cromosoma produce una delezione terminale, mentre due rotture nello stesso cromosoma risultano in una delezione interstiziale, un cromosoma ad anello oppure un'inversione. Due rotture in due differenti cromosomi producono anomalie strutturali, come traslocazioni reciproche e Robertsoniane⁶.

Il *crossing-over* ineguale conduce a delezioni o duplicazioni⁷.

Le anomalie cromosomiche strutturali possono essere di due tipi, bilanciate (traslocazioni reciproche, Robertsoniane, inversioni) e sbilanciate (delezioni, duplicazioni): nel primo caso la cromatina rimane quantitativamente invariata, mentre nel secondo caso, il riarrangiamento comporta un eccesso o una perdita di cromatina rispetto alla norma⁶. Le anomalie cromosomiche bilanciate non implicano un fenotipo patologico nel soggetto portatore, a meno che non interrompano, inattivandolo, un gene specifico, portando perciò ad una malattia. Questi tipi di anomalie sono invece rilevanti nei confronti del rischio riproduttivo, poiché possono portare alla formazione di gameti sbilanciati, con conseguente concepimento di un individuo con sindrome cromosomica.

Le anomalie cromosomiche sbilanciate invece causano direttamente un fenotipo patologico⁶.

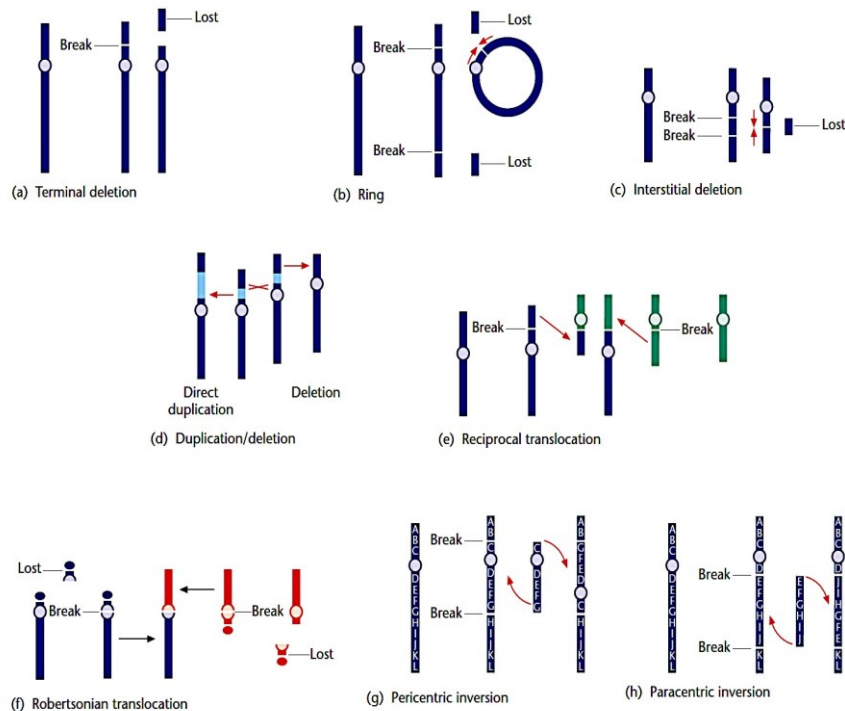


Figura 1. Riarrangiamenti strutturali

- (a) Delezione terminale: origina da una singola rottura del cromosoma, con perdita del segmento di cromatina terminale.
- (b) Cromosoma ad anello: origina da una rottura in ognuna della braccia del cromosoma, con perdita dei segmenti terminali e riunione di quello centromerico.
- (c) Delezione interstiziale: origina da due rotture sul medesimo braccio cromosomico, con perdita del segmento interstiziale e riunione dei due segmenti rimanenti.
- (d) Duplicazione/delezione: origine di una duplicazione diretta e di una delezione attraverso un evento di *crossing-over* ineguale.
- (e) Traslocazione reciproca: origina da una rottura in ognuno dei cromosomi e scambio dei segmenti non centromerici.
- (f) Traslocazione robertsoniana: origina da una rottura a livello del centromero di ognuno dei cromosomi, unione delle braccia lunghe e perdita delle braccia corte dei due cromosomi.
- (g) Inversione pericentrica: origina da una rottura in ognuna delle due braccia cromosomiche, con rotazione del segmento centromerico di 180° e riunione dello stesso ai due segmenti terminali.
- (h) Inversione paracentrica: origina da due rotture nello stesso braccio cromosomico, con rotazione del segmento interstiziale di 180° e riunione dello stesso ai segmenti terminali.

Delezioni

Si tratta di anomalie strutturali in cui una porzione di cromatina del cromosoma viene persa. Le delezioni, come accennato precedentemente, possono essere terminali (Figura 1a) o interstiziali (Figura 1c). Il fenotipo clinico associato alla delezione, varia a seconda dell'estensione del segmento deleto⁶. Delezioni

abbastanza estese da essere visibili con la microscopia ottica rappresentano la perdita di numerosi geni localizzati nella stessa banda o regione cromosomica, e risultano perciò in una monosomia per quella porzione del genoma. Il meccanismo patogenetico che determina le anomalie fenotipiche è l'aploinsufficienza (insufficiente quantità di prodotto genico che interessa geni per i quali il 50% di prodotto genico non è abbastanza per garantirne la funzione), per cui i trascritti dei geni inclusi nell'intervallo di delezione sono mediamente ridotti della metà rispetto alla norma⁸.

Cromosomi ad anello

Originano da due delezioni terminali: un evento di rottura in ognuna delle due braccia del cromosoma causa generalmente la perdita delle porzioni terminali del cromosoma e la fusione delle due estremità del segmento centromerico rimanente (Figura 1b). In alcuni casi l'origine del ring è il risultato di un processo più complesso di delezione/duplicazione, derivante da *crossing-over* ineguali⁹. I cromosomi ad anello rappresentano un riarrangiamento complesso, caratterizzato da instabilità mitotica.

Duplicazioni

Si tratta di riarrangiamenti cromosomici strutturali non bilanciati, che risultano in una trisomia parziale. Rispetto alle delezioni, le duplicazioni tendono a produrre un fenotipo clinico meno grave, nonostante condividano con le delezioni diverse caratteristiche cliniche. Le duplicazioni derivano principalmente da un *crossing-over* ineguale (Figura 1d), in particolare a livello delle regioni cromosomiche contenenti sequenza di DNA ripetute. Le duplicazioni possono essere distinte in dirette e invertite, nel caso il segmento duplicato mantenga o meno la direzione originaria rispetto al centromero^{6,7}.

Traslocazioni reciproche

Sono caratterizzate dallo scambio di cromatina tra due diversi cromosomi; un singolo evento di rottura su ognuno dei cromosomi si conclude con lo scambio dei

segmenti non centromerici tra i due cromosomi interessati, senza perdita visibile di cromatina (Figura 1e). Una traslocazione reciproca bilanciata è di solito innocua per il soggetto portatore, ma in alcune circostanze si associa ad un fenotipo patologico, che si configura come una malattia monogenica oppure come una patologia cromosomica. Nel primo caso, uno dei punti di rottura che hanno determinato la traslocazione avviene nel contesto di un gene trascritto che viene inattivato, mentre nel secondo caso la traslocazione è solo apparentemente bilanciata^{6,10}. Gli individui portatori di traslocazione reciproca bilanciata hanno un aumentato rischio riproduttivo, sia per abortività ricorrente sia per la possibilità di avere figli nati vivi affetti da sindrome cromosomica, e un aumentato rischio di infertilità.

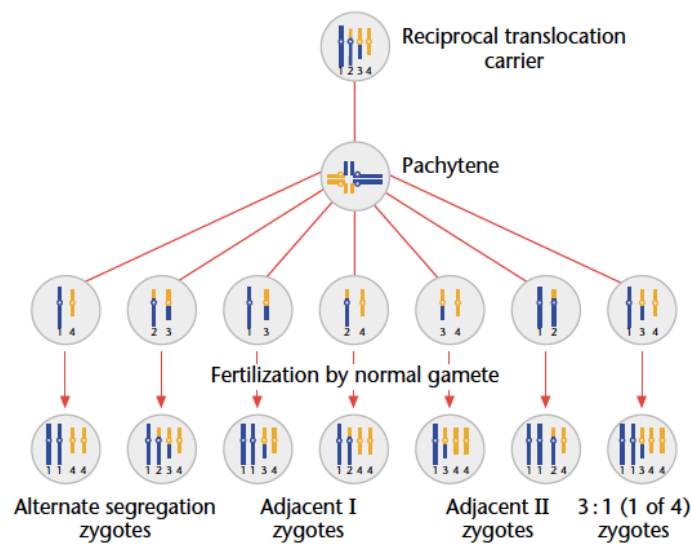
L'appaiamento dei cromosomi omologhi alla meiosi è alterato nei portatori di traslocazione reciproca bilanciata: anziché il normale appaiamento come bivalenti, i due cromosomi traslocati e i due cromosomi non traslocati si appaiano formando una tetrad quadrivalente, in cui ogni regione si appaia con la sua omologa¹⁰. Esistono quattro principali modalità di segregazione meiotica in caso di traslocazione reciproca bilanciata (Figura 2a): nella segregazione alternata, entrambi i cromosomi normali si muovono verso un polo e i due cromosomi traslocati si muovono entrambi verso il polo opposto, producendo quindi gameti bilanciati. Nella segregazione adiacente I, in cui i cromosomi non omologhi adiacenti si muovono verso lo stesso polo della cellula, avremo come risultato gameti sbilanciati, come anche nel caso della segregazione adiacente II, in cui sono i cromosomi omologhi adiacenti a muoversi verso il medesimo polo cellulare^{6,10}. Nella segregazione 3:1, tre cromosomi si muovono verso un polo della cellula e solo un cromosoma si muove verso il polo opposto¹¹.

Se lo zigote è portatore di uno sbilanciamento cromosomico, si verifica nella maggior parte dei casi un aborto spontaneo selettivo.

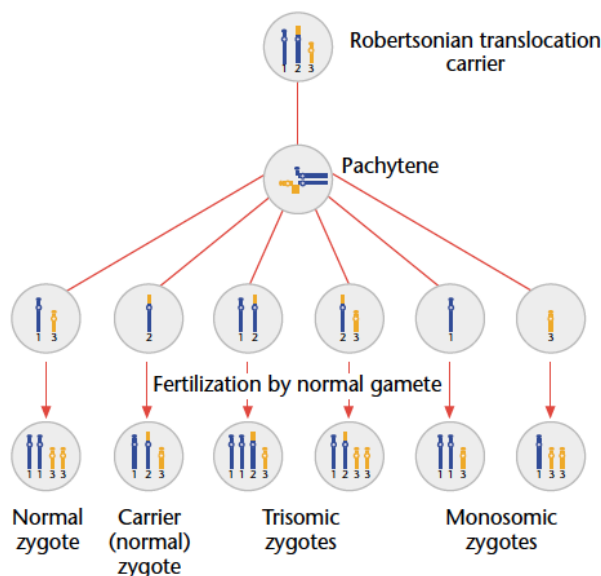
Traslocazioni Robertsoniane

Si originano dalla fusione di due cromosomi acrocentrici a livello del centromero, con formazione di un unico cromosoma costituito dalle braccia lunghe dei due cromosomi e perdita delle braccia corte di entrambi. In presenza di una traslocazione robertsoniana, il numero di cromosomi si riduce quindi

apparentemente di una unità^{6,12}. La traslocazione robertsoniana $t(13;14)(q10;q10)$ è la più frequente nella popolazione generale; sebbene innocua per il soggetto portatore, essa implica il rischio di concepire zigoti con trisomia 13¹². Come nel caso delle traslocazioni reciproche, l'appaiamento meiotico coinvolge sia i cromosomi omologhi normali che i cromosomi interessati dalla traslocazione, ma con la formazione di una triade e non di una tetrade (Figura 2b). Nella segregazione alternata si formano gameti bilanciati, normali o con la stessa traslocazione bilanciata, mentre nel caso della segregazione adiacente si formano gameti sbilanciati, per disomia o nullisomia di uno o l'altro dei cromosomi acrocentrici interessati dalla traslocazione^{6,12}.



(a)



(b)

Figura 2. Segregazione meiotica in caso di traslocazioni reciproche bilanciate (a) e di traslocazioni robertsoniane (b).

Inversioni

Le inversioni si formano da due rotture nello stesso cromosoma con conseguente scambio dei due segmenti terminali; le inversioni avvengono perciò essenzialmente nella stessa maniera con cui avvengono le traslocazioni, tranne che per il fatto che le rotture del DNA e lo scambio di materiale cromatinico si verificano all'interno di un solo cromosoma.

Le inversioni possono essere di due tipi, pericentriche (Figura 1g) e paracentriche (Figura 1h): le prime si formano per un doppio evento di rottura dello stesso cromosoma con inversione di 180° di un segmento contenente il centromero, mentre le seconde si formano per un doppio evento di rottura dello stesso cromosoma con inversione di 180° di un segmento che esclude il centromero^{6,13}. Nel caso dell'inversione pericentrica cambia il pattern di bande ed eventualmente anche la forma del cromosoma, mentre nel caso dell'inversione paracentrica la forma del cromosoma rimane invariata. I portatori di inversioni peri- o paracentriche in stato di eterozigosi non presentano un fenotipo clinico patologico, ma hanno un aumentato rischio riproduttivo dovuto ad anomalie durante l'appaiamento meiotico¹³.

1.3 SEQUENZE CHE MEDIANO RIARRANGIAMENTI

Circa il 45% del genoma umano consiste in sequenze di DNA ripetitivo intersperso, derivate principalmente da copie degenerate di elementi trasponibili⁵. Gli elementi trasponibili includono trasposoni a DNA e retrotrasposoni con e senza LTR (*Long Terminal Repeat*, lunghe circa 1 Kb). Sebbene la maggior parte degli elementi trasponibili sparsi nel genoma sia stata resa inattiva da mutazioni, la retrotrasposizione delle sequenze LINE-1 (*Long INterspersed Element-1*) continua a diversificare il genoma umano⁴ (Figura 3).

Trasposoni a DNA

I trasposoni a DNA costituiscono circa il 3% del genoma umano¹⁴. Nella loro forma più semplice, questi elementi trasponibili contengono sequenze terminali ripetute e invertite, che circondano una ORF (*Open Reading Frame*), che codifica per l'enzima trasposasi^{5,14}. L'mRNA risultante dalla trascrizione del trasposone

viene spostato nel citoplasma, dove viene tradotto, portando alla sintesi dell'enzima trasposasi; la trasposasi una volta trasferita nel nucleo, lega all'interno o nelle vicinanze del trasposone le sequenze terminali ripetute e invertite, "taglia" il trasposone e lo "incolla" in un'altra posizione nel genoma^{5,14}. Praticamente tutti i trasposoni a DNA del genoma umano sono mutati, incapaci di trasporre da una regione ad un'altra del genoma.

Retrotrasposoni LTR

I retroelementi contenenti LTR compongono circa l'8% del nostro genoma¹⁵. I retrovirus endogeni HERV fanno parte dei retrotrasposoni LTR e somigliano ai retrovirus nella struttura e nella mobilità; molti HERV contengono un gene ENV non funzionante, che perciò li relega ad un'esistenza intracellulare^{3,15}. Praticamente tutti gli HERV sono stati resi inattivi da eventi mutazionali e quindi non sono capaci di retrotrasposizione autonoma; soltanto la famiglia HERV-K, presente nell'uomo e nelle scimmie antropomorfe e quindi evolutivamente recente, conserva delle copie intere non mutate del genoma virale e rimane potenzialmente infettiva^{1,15}.

LINE-1

Le sequenze LINE-1 sono gli unici retrotrasposoni umani autonomamente attivi conosciuti, e costituiscono approssimativamente il 17% del nostro genoma¹⁶. I retroelementi LINE-1 capaci di retrotrasposizione (RC-L1) sono lunghi circa 6 Kb e contengono una regione 5'UTR (*UnTranslated Region*), due ORF e una regione 3'UTR che termina con una coda di poli-A.

Nella regione 5'UTR si trova un promotore interno per la RNA polimerasi II, che ha il compito di dirigere la trascrizione del retroelemento a partire dall'estremità 5'; la regione ORF1 codifica per la proteina ORF1p, mentre la regione ORF2 codifica per la proteina ORF2p con attività di endonucleasi e di trascrittasi inversa, necessarie per la retrasposizione di L1^{3,16,17}.

Retrotrasposoni non autonomi: elementi Alu e SVA

Il genoma umano contiene anche numerosi retrotrasposoni non autonomamente attivi, che si servono delle proteine codificate dai retroelementi L1 per poter acquisire la mobilità; questi retroelementi non autonomi consistono principalmente nelle sequenze Alu (SINE, *Short INterspersed Element*), che compongono circa il 10% del genoma umano¹⁸. Le Alu derivano da un piccolo RNA denominato 7SL e sono lunghe circa 300 pb. Le sequenze Alu hanno una struttura dimerica: il monomero di sinistra contiene un promotore interno per la RNA polimerasi III, ed è separato dal monomero di destra da una sequenza ricca di adenosine. Le Alu sono dotate di una coda di poli-A e non possiedono sequenze codificanti, per cui non sono capaci di retrotrasposizione autonoma, ma possono essere retrotrasposte da parte della trascrittasi inversa degli elementi L1¹⁸. Gli elementi SVA (SINE-R/VNTR/Alu) compongono circa lo 0,2% del genoma umano e sono lunghi approssimativamente 2,000 pb. La loro struttura è composta da una sequenza ripetuta esamerica CCCTCT, una sequenza ripetuta Alu-like invertita, una serie di VNTR (*Variable Number of Tandem Repeats*) ricchi di G+C, una sequenza SINE-R e una coda di poli-A¹⁹.

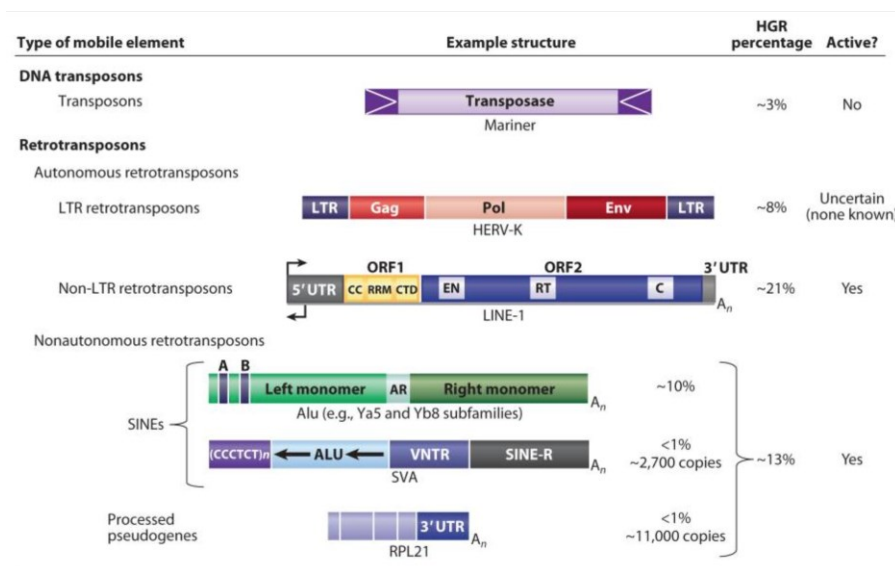


Figura 3. Elementi retrotrasponibili del genoma umano.

Elementi genetici “mobili” nel genoma umano. Abbreviazioni: HERV-K (retrovirus-K umano endogeno), LTR (*long terminal repeat*); Gag (antigene gruppo specifico), Pol (polimerasi), Env (proteina dell'involucro). **LINE-1**: UTR, regione non tradotta; CC, bobina a spirale; RRM, motivo riconoscimento RNA; CTD, dominio carbossi-terminale; EN, endonucleasi; RT, trascrittasi inversa; C, dominio cisteina. **Alu**: A e B, le sequenze che compongono promotore RNA polimerasi III; AR, il segmento ricco di adenosina separatore dei monomeri 7SL. **SINE-R / VNTR / Alu (SVA)**: VNTR, numero variabile di ripetizioni in tandem; SINE-R, dominio indicato da un HERV-K.

Duplicazioni segmentali

Altre sequenze di DNA ripetitivo sono caratteristicamente coinvolte in riarrangiamenti strutturali non casuali di particolari regioni cromosomiche, i quali, determinando uno squilibrio quantitativo di geni dose-sensibili, sono responsabili di disordini genomici²⁰. Queste sequenze sono note come dupliconi, o regioni LCR (*Low Copy Repeats*) o duplicazioni segmentali (SD). I dupliconi sono blocchi di poche sequenze ripetute con omologia molto elevata (94-99%) e sono lunghe centinaia di Kb (200-400 Kb).

Queste sequenze intersperse in tutto il genoma, rappresentano circa il 10%, ma sono localizzate in siti specifici, in particolare nelle regioni subtelomeriche e pericentromeriche. Vista la loro alta omologia reciproca, la presenza di due o più dupliconi nell'ambito della stessa regione cromosomica, può determinare errori di ricombinazione al *crossing-over* meiotico: il meccanismo di ricombinazione omologa non allelica mediato dai dupliconi può causare non solo delezioni e duplicazioni cromosomiche^{7,21}, ma anche inversioni e traslocazioni^{21,22}.

1.4 MECCANISMI ALLA BASE DEI RIARRANGIAMENTI CROMOSOMICI STRUTTURALI

I riarrangiamenti cromosomici di struttura sono il risultato di una o più rotture del DNA a cui fa seguito un ripristino erroneo della continuità dei filamenti nucleotidici coinvolti²³. Tali alterazioni possono derivare sia da un errore nel processo di riparazione che da un difetto del sistema di ricombinazione. Generalmente i riarrangiamenti genomici strutturali consistono in anomalie macroscopiche del DNA che vanno da migliaia a talvolta milioni di paia di basi: questi riarrangiamenti possono rappresentare polimorfismi senza alcun effetto fenotipico (*neutral polymorphism*) oppure possono condurre a fenotipi patologici attraverso diverse vie, tra cui la variazione del numero di copie (CNV- *Copy Number Variation*) di geni sensibili al dosaggio, la distruzione di geni o la creazione di geni di fusione. Le condizioni patologiche derivanti dai riarrangiamenti del genoma vengono definite collettivamente disordini genomici^{23,24,25}.

I riarrangiamenti genomici possono essere ricorrenti o non ricorrenti (Figura 4): i riarrangiamenti ricorrenti sono della stessa misura e presentano punti di rottura fissi (*breakpoints cluster*), mentre quelli non ricorrenti hanno misure differenti e diversi *breakpoint* (bp) per ogni singolo evento. I riarrangiamenti non ricorrenti condividono una comune regione di sovrapposizione, SRO (*Smallest Region of Overlap*), che contiene il locus associato al disordine genomico^{8,26,27}.

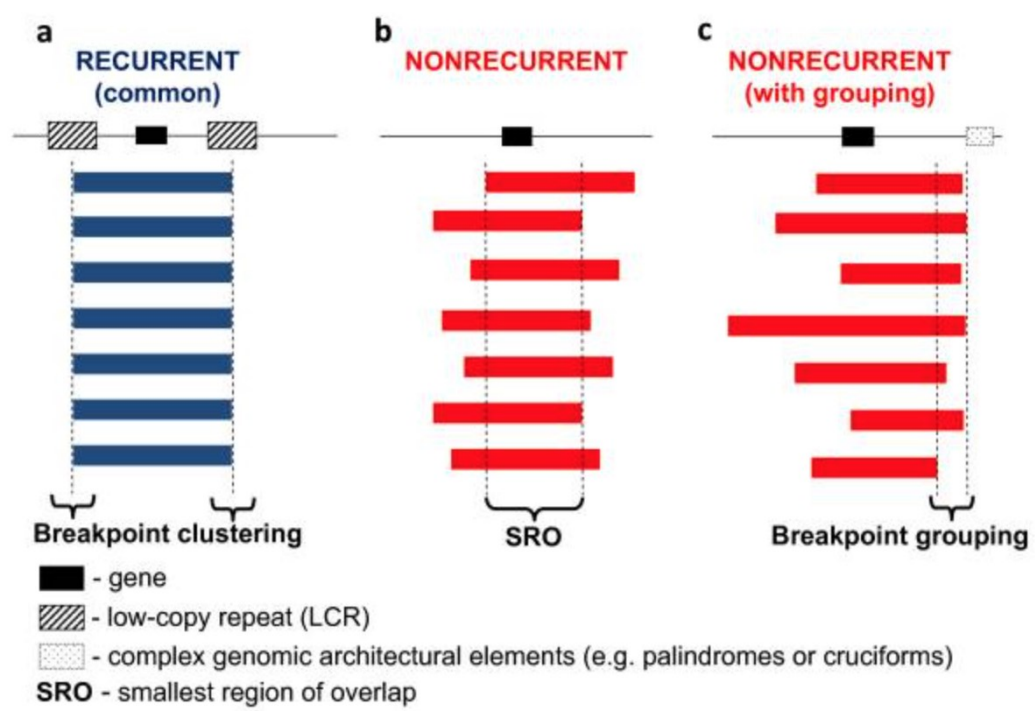


Figura 4. Meccanismo di formazione di riarrangiamenti strutturali ricorrenti e non ricorrenti.

(a). Riarrangiamenti ricorrenti con una dimensione comune e punti di rottura in *cluster*. La maggior parte dei riarrangiamenti ricorrenti risultano da una ricombinazione omologa non allelica (NAHR). I punti di rottura si trovano all'interno di LCR. (b). Riarrangiamenti non ricorrenti. I punti di rottura dei riarrangiamenti sono localizzati lungo la regione genomica. Tutti i riarrangiamenti condividono una comune regione genomica comune che comprende un gene probabilmente associato al disordine genomico. (c). Riarrangiamenti non ricorrenti con raggruppamento di un punto di rottura. Alcuni dei riarrangiamenti non ricorrenti hanno uno dei loro punti di rottura localizzati in una piccola regione genomica.

I riarrangiamenti cromosomici strutturali possono avvenire attraverso meccanismi di ricombinazione omologa (HR, *Homologous Recombination*) e meccanismi di ricombinazione e/o di riparazione del DNA non omologhe^{28,29,30}.

1.4.1 Meccanismi di ricombinazione omologa

Questi meccanismi richiedono lunghe sequenze di DNA identiche (più di 300 pb) e la proteina RecA/Rad51 per lo scambio del filamento. La ricombinazione omologa è alla base di numerosi accurati meccanismi di riparazione del DNA: anomalie strutturali cromosomiche possono essere generate da meccanismi di ricombinazione omologa, non perché questi risultano inaccurati, ma a causa della presenza di numerosi tratti del genoma contenenti LCR^{23,24}.

Ricombinazione omologa non allelica (NAHR, Non Allelic Homologous Recombination)

La ricombinazione omologa non allelica conduce a riarrangiamenti strutturali ricorrenti, mediati da un'architettura o una struttura genomica comune, in cui l'intervallo coinvolto nel riarrangiamento è fiancheggiato da sequenze ripetitive paraloghe o da LCR^{20,30-33}. A causa del loro alto grado di identità di sequenza, copie non alleliche di LCR possono allinearsi durante la meiosi e la mitosi, conducendo ad eventi di *crossing-over* ineguale. Eventi di ricombinazione omologa non allelica in cellule della linea germinale possono portare a riarrangiamenti genomici strutturali costituzionali, che si manifestano con disordini genomici³⁰. La ricombinazione omologa non allelica può verificarsi anche durante la mitosi, producendo popolazioni mosaico di cellule somatiche con riarrangiamenti genomici^{30,34}. La ricombinazione tra sequenze LCR dirette nella stessa direzione causa delezioni o duplicazioni, mentre la ricombinazione tra sequenze invertite causa inversioni^{20,23} (Figura 5a).

La ricombinazione tra cromosomi differenti può condurre a traslocazioni^{22,23} (Figura 5c). Eventi di ricombinazione tra sequenze LCR invertite localizzate su due cromosomi fratelli possono causare la formazione di un isocromosoma, ovvero un cromosoma strutturalmente anomalo, con un braccio parzialmente deleta e l'altro duplicato^{22,23,30} (Figura 5b).

Affinché la ricombinazione omologa non allelica si verifichi, sono necessari segmenti di DNA di lunghezza minima con elevato grado di omologia tra le LCR, denominati *minimal efficient processing segments* (MEPS).

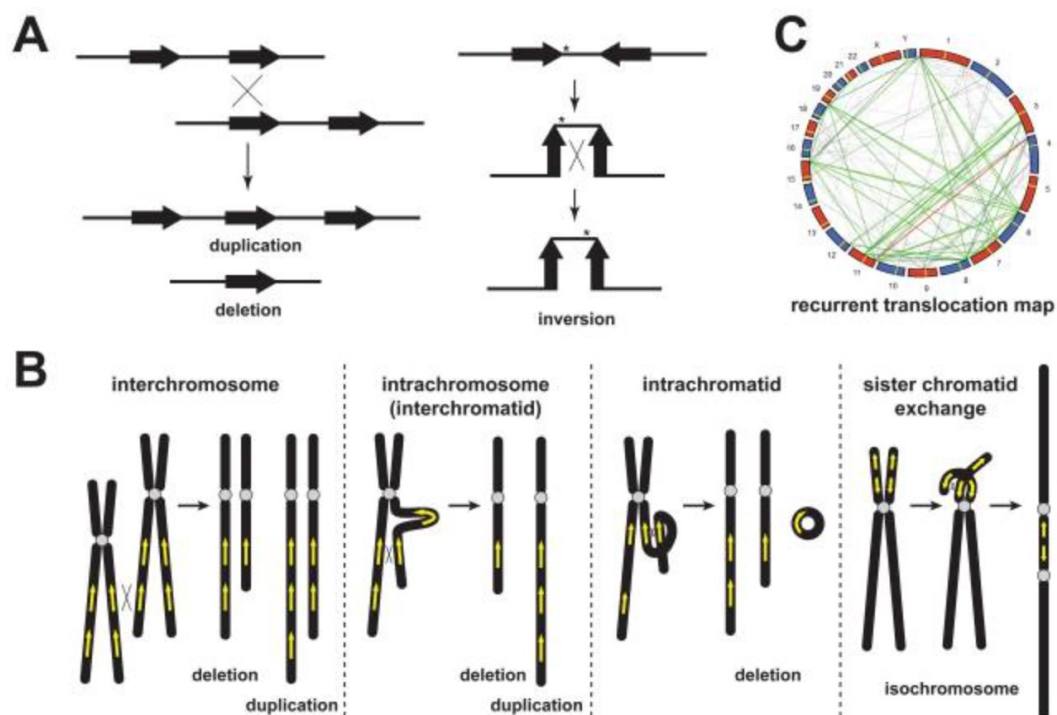


Figura 5. Riarrangiamenti cromosomici mediati da meccanismi di ricombinazione omologa non allelica (NAHR).

(a). *Crossing-over* LCR aventi la stessa direzione possono portare a delezioni o duplicazioni mentre la ricombinazione tra sequenze in direzione opposte possono portare ad inversione. (b) La ricombinazione omologa non allelica (NAHR) può essere intercromosomica o intracromosomica causando delezioni e duplicazioni, oppure intracromatidica causando esclusivamente delezioni. Se la ricombinazione avviene tra sequenze invertite su due cromosomi fratelli, si ha la formazione di un isocromosoma. (c) una mappa di una traslocazione ricorrente mediata da NAHR. La linea rossa indica le traslocazioni ricorrenti osservate e verificate sperimentalmente, mentre le linee grigio e verde (geni recettore olfattivo) denotano traslocazioni ricorrenti previste che potrebbero essere mediate da LCR paraloghe

Single-Strand Annealing (SSA)

Delezioni di piccole dimensioni possono verificarsi attraverso un meccanismo di riparazione di rottura del DNA, mediato di sequenze omologhe ripetute, definito *single-strand annealing* (SSA)^{30, 31, 35}. Nel caso in cui la resezione dell'estremità 5' di entrambe le estremità della rottura di DNA a doppio filamento non conduca all'invasione di sequenze omologhe da parte dell'estremità 3', la resezione continua, espone un tratto 3' a singolo filamento di notevole lunghezza: se vengono esposte sequenze complementari nei due tratti a singolo filamento, può avvenire un fenomeno di appaiamento. Concluso il processo di riparazione, avremo la delezione del tratto localizzato tra le sequenze ripetute e la delezione di una di queste (Figura 6)^{26,30}.

Nel genoma umano, il meccanismo di SSA indotto da rottura del DNA a doppio filamento è stato individuato tra sequenze Alu identiche ripetute, separate da poche centinaia di paia di basi; più è lunga la sequenza che separa le sequenze ripetitive mediatrici di riarrangiamento, meno è probabile che la rottura sia riparata usando questo meccanismo di ricombinazione omologa.

Difatti, è probabile che il meccanismo SSA sia un attore minore nella formazione dei riarrangiamenti genomici strutturali³⁶.

Single-strand annealing

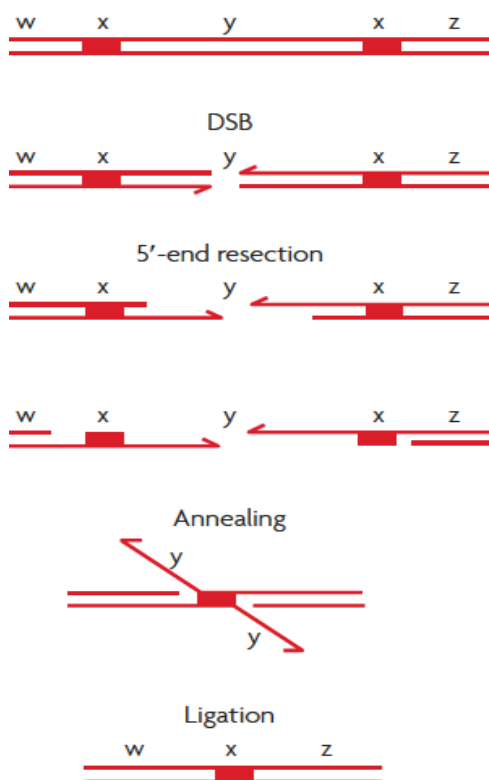


Figura 6. Schema del meccanismo Single-Strand Annealing (SSA).

Taglio all'estremità 5' di entrambe i filamenti della doppia elica del DNA, mancata invasione di sequenze omologhe che possano appaiarsi all'estremità 3', continua la resezione ed espone un tratto a singolo filamento di notevole lunghezza all'estremità 3'. Concluso il processo di riparazione, avremo la delezione del tratto localizzato tra le sequenze ripetute.

1.4.2 Meccanismi non replicativi di riparazione non omologa

Esistono dei meccanismi di riparazione di rotture del DNA che non si servono di sequenze mediatrici omologhe, o eventualmente sfruttano un'omologia limitata a poche paia di basi; quando non è richiesto un elevato grado di omologia, non è garantito che il meccanismo di riparazione conduca ad una riunione dei filamenti di DNA in posizione corretta, ed è perciò relativamente probabile che si verifichino anomalie cromosomiche strutturali³⁷.

Non-Homologous End Joining (NHEJ) e Microhomology-Mediated End Joining (MMEJ)

Il meccanismo di NHEJ può riparare una rottura a doppio filamento del DNA in modo corretto oppure può condurre a piccole delezioni (1-4 pb), e in alcuni casi anche a inserzioni di DNA libero, generalmente di origine mitocondriale o originato da un retrotrasposone³⁸. La riunione delle estremità della rottura a doppio filamento del DNA attraverso il meccanismo di MMEJ, è mediata da un'omologia di sequenza di poche paia di basi (5-25 pb), e conduce a delezione delle sequenze tra le regioni di micro-omologia³⁹. Il meccanismo di NHEJ procede attraverso 4 passi (Figura 7): per prima cosa si ha il riconoscimento del punto di rottura a doppio filamento del DNA, seguito dalla formazione di un ponte molecolare per tenere unite le due estremità libere del DNA; successivamente avviene la modificazione delle estremità per renderle compatibili e complementari, così da concludere il processo di riparazione con la riunione delle stesse^{8,20,21,24}.

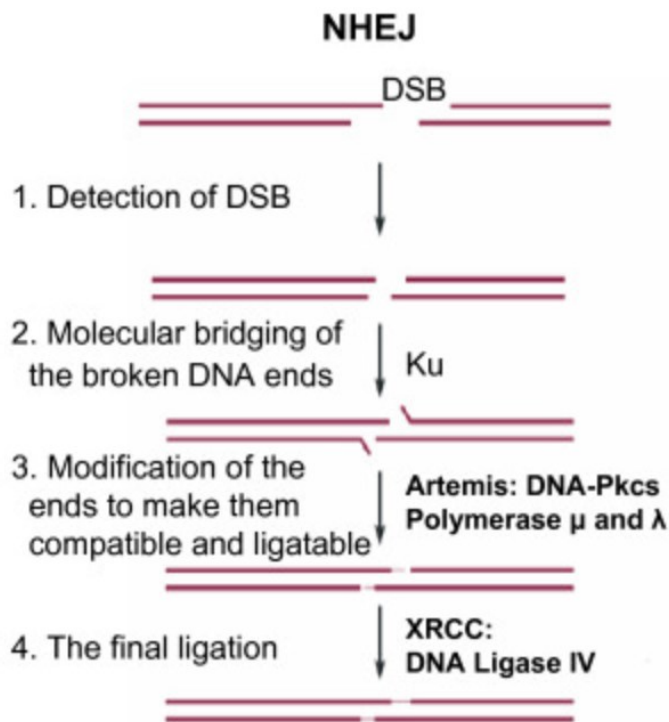


Figura 7. Meccanismo di non homologous end joining (NHEJ).

Quando una rottura del doppio filamento del DNA si verifica, viene riparata tramite il meccanismo NHEJ. Le due linee spesse rappresentano due filamenti di DNA con la rottura, i segmenti sottili in mezzo rappresentano le modificazioni a cui le estremità sono sottoposte per poter poi passare al processo di ligazione. La DNA ligasi IV provvede a ligare i frammenti per poi ricostituire il doppio filamento del DNA. E' possibile che durante il processo di riparazione delle estremità si possono avere delezioni o duplicazioni di basi nucleotidiche, lasciando la "firma" di un meccanismo NHEJ.

Breakage-Fusion-Bridge Cycle (BFB Cycle)

Secondo questo meccanismo proposto da Barbara McClintock⁴⁰, due cromatidi fratelli privi entrambi di telomero alle estremità, tendono a fondersi creando un cromosoma dicentrico. I cromatidi senza telomero derivano dalla replicazione di un cromosoma, che ha subito la perdita della sua estremità telomerica, per una rottura a doppio filamento del DNA.

Durante l'anafase, i due centromeri sono spinti verso poli opposti, causando una possibile rottura del cromosoma dicentrico; a causa della rottura e della susseguente replicazione del DNA, si generano nuove estremità prive di telomero, e così il ciclo riparte con la formazione di un altro cromosoma dicentrico^{41,42} (Figura 8).

Il ciclo cessa nel momento in cui il cromosoma acquista un telomero, che ne protegga l'estremità. Gli eventi di rottura casuali inducono grandi duplicazioni

invertite, e cicli ripetuti ne causano l'amplificazione: il meccanismo di *break-fusion-bridge cycle* fornisce una spiegazione della presenza di grandi sequenze ripetute invertite nel genoma di cellule tumorali umane⁴¹.

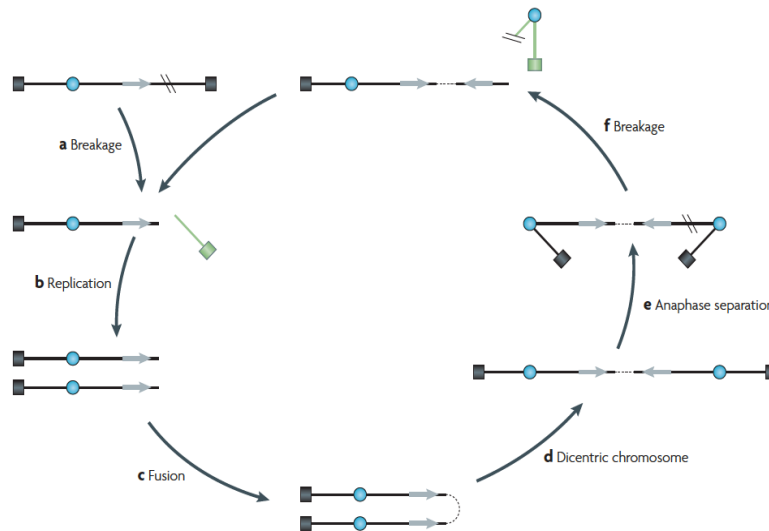


Figura 8. Breakage-Fusion-Bridge Cycle.

I centromeri sono indicati con un pallino blu, i telomeri con un quadratino nero e le sequenze genomiche da frecce grigie che indicano l'orientamento. I punti di rottura sono indicati dalle linee nere mentre i frammenti deleti in verde. (a) Una rottura del doppio filamento del DNA avviene a livello di un cromosoma non replicato, causando la perdita del telomero (b) Dopo la replicazione, tutti e due i cromatidi fratelli perdono i telomeri. (c) Le due estremità terminali si fondono insieme. (d) La fusione che avviene nella fase c forma un cromosoma dicentrico. (e) All'anafase, i due centromeri del cromosoma dicentrico sono separati, formando una struttura a ponte tra i due nuclei. (f) La struttura a ponte si rompe in modo random. Questo porta alla formazione di una grande duplicazione invertita. A questo punto il cromosoma ha una estremità non protetta e, dopo la replicazione, si formeranno due cromatidi fratelli che possono fondersi tra di loro e formare un nuovo cromosoma dicentrico. Il processo si ripete fino a quando l'estremità non acquista un telomero da una nuova fonte.

1.4.3 Meccanismi replicativi di riparazione non omologa

Il meccanismo FoSTeS (*Fork Stalling and Template Switching*) è stato proposto da Lee et al come causa dei riarrangiamenti strutturali non ricorrenti complessi alla base della malattia di Pelizaeus-Merzbacher (PMD), ed eventualmente come causa anche di numerosi altri riarrangiamenti non ricorrenti del genoma⁴³.

Secondo il modello FoSTeS (Figura 9), nel caso in cui la forca di replicazione attiva si blocchi, l'estremità 3' del filamento lento può slegarsi dal suo filamento stampo e appaiarsi con un altro filamento stampo, con cui condivide micro-omologie, all'interno di un'altra forca di replicazione. Si possono generare così

delezioni, duplicazioni, inversioni o traslocazioni, a seconda della posizione della nuova forca di replicazione rispetto alla precedente^{8,44}.

Secondo questo modello, sono due le caratteristiche principali dei riarrangiamenti FoSTeS-mediati, ovvero la complessità di sequenza derivante dal susseguirsi di giustapposizioni di sequenze di DNA da regioni diverse del genoma, e le microomologie a livello dei punti di unione⁴⁵.

Attraverso l'analisi dei riarrangiamenti complessi associati a disordini genomici nell'uomo, non è stato possibile determinare se una rottura del DNA a singolo o a doppio filamento sia implicata nell'invasione del nuovo filamento stampo, né è stato possibile determinare se il meccanismo sia innescato da un blocco della forca di replicazione oppure da una rottura della stessa⁴⁴.

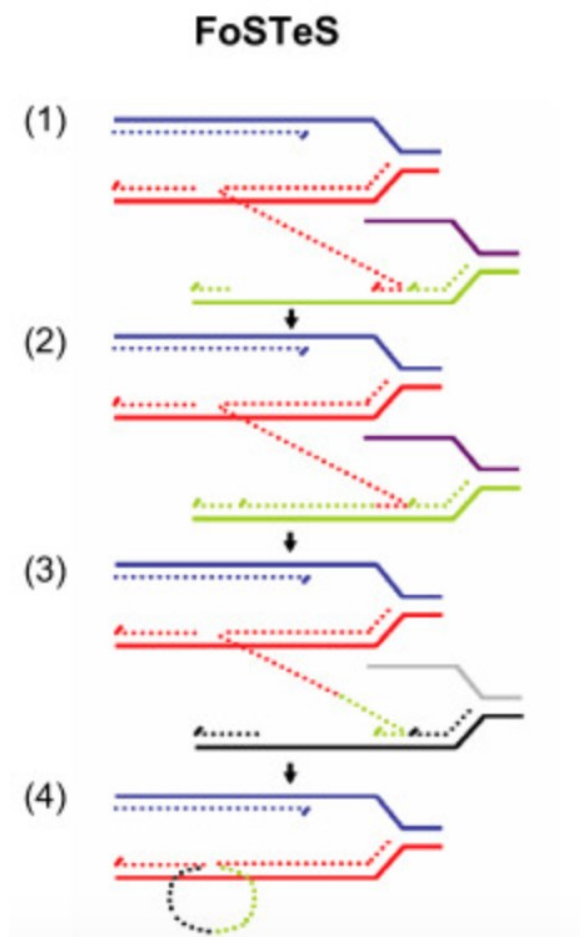


Figura 9. Il meccanismo FoSTeS (Fork Stalling and Template Switching)

Dopo lo stallo originario della forcella originaria di replicazione (linee continue, blu e rosso scuro), il filamento in ritardo (rosso, linea tratteggiata) si libera e si collega ad una seconda forcella (linee continue, viola e verde) per microomologia (1), seguito da (2) estensione della seconda forca e sintesi del DNA (verde, linea tratteggiata). Dopo lo sgancio della forcella (3), la forcella originale (linee continue blu e rosso scuro) legata con il suo filamento di ritardo (rosso e verde, le linee tratteggiate) potrebbero invadere una terza forca (linee continue, grigio e nero). Le linee tratteggiate rappresentano DNA appena sintetizzato. La replicazione seriale della forcella, che disimpegna e ritarda l'invasione del filamento potrebbe verificarsi più volte (ad esempio FoSTeS x 2, FoSTeS x 3, ecc) prima (4) della ripresa della replica sul modello originale.

Basato sull'evidenza a riguardo dei riarrangiamenti genomici in umani, batteri e lieviti, è stato proposto, accanto al modello FoSTeS, un nuovo modello replicativo, il modello MMBIR (*Microhomology-mediated break-induced replication*)^{39,46}.

MMBIR è basato sul meccanismo BIR (*Break-induced replication*) di ripresa della replicazione dopo rottura della forca di replicazione^{39,47,48} (Figura 10).

Quando l'enzima elicasi incontra una discontinuità nel filamento stampo, un braccio della forca di replicazione si rompe; non c'è una seconda estremità che possa essere coinvolta in uno dei meccanismi di riparazione della rottura a doppio filamento del DNA, in cui si procede alla riunione di due estremità³⁹.

L'estremità 5' del braccio rotto è reseccata da un'esonucleasi, in modo da lasciare un tratto sporgente dell'estremità 3', che va ad invadere una sequenza omologa, normalmente sul cromatidio fratello da cui proviene; l'invasione è mediata dalla proteina RecA/Rad51⁴⁹. L'estremità 3' avvia la sintesi del DNA e istituisce una forca di replicazione, a bassa processività, con la sintesi sia del filamento lento che di quello guida; successivamente, il braccio esteso si separa dal cromatidio fratello a cui si era appaiato. A questo punto, l'estremità 3' invade di nuovo il cromatidio fratello, appaiandosi con sequenze omologhe, e così il processo si ripete. Dopo alcuni cicli ripetuti di invasione, estensione e separazione, la forca di replicazione diventa a più elevata processività, probabilmente per uno scambio delle DNA polimerasi coinvolte^{39,47}.

BIR è strettamente RecA/Rad51-dipendente e omologia-dipendente⁴⁹.

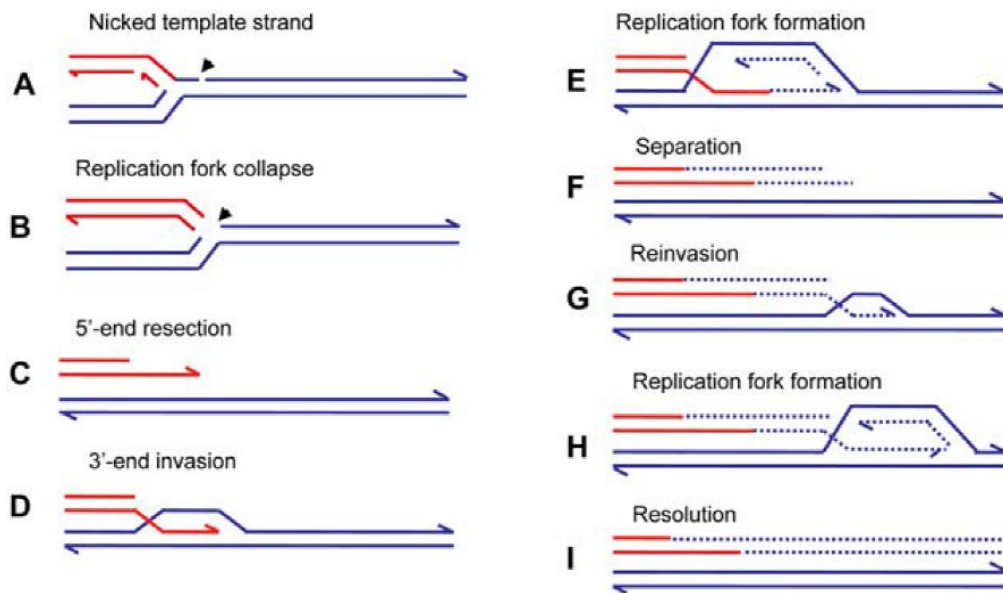


Figura 10. Riparazione di una forza di replicazione dopo il suo collasso, mediante Break-induced replication (BIR).

Quando una forcella di replicazione incontra un'interruzione nel filamento stampo (freccia nera) (A), un braccio della forcella si interrompe (rosso), producendo una forcella collassata (B). L'estremità 5' viene tagliata lasciando sporgere l'estremità 3' (C). L'estremità 3' a singolo filamento *End single-strand* 3' invade la sequenza omologa sul cromatidio fratello (blu), formando un D-loop (D), che diventa una forcella di replicazione (E). A questo punto si ha la separazione dal cromatidio fratello a cui si era appaiato. (F). L'estremità separata invade nuovamente il cromatidio fratello, e forma una nuova forcella di replicazione (G). Alla fine, la forcella di replicazione continua la replicazione fino alla fine dei cromosomi (H e I). Ogni linea rappresenta un filamento di DNA (*strand*). Il nuovo filamento di DNA sintetizzato è indicato da linee tratteggiate.

Quando una forza di replicazione collassa, generando una rottura del DNA a singola estremità, in una cellula sottoposta a stress, il classico meccanismo di riparazione BIR non può attuarsi perché la trascrizione del gene *RecA/Rad51* è sottoregolata, in risposta appunto allo stress⁴⁹.

Il meccanismo MMBIR si fonda sul presupposto che l'invasione da parte dell'estremità 3' del filamento del cromatidio fratello sia limitata quando la trascrizione di *RecA/Rad51* è sottoregolata: quindi, l'estremità 3' formatasi dopo la rottura della forza di replicazione si appaia con qualunque stampo a singolo filamento nelle vicinanze, con cui condivide una micro-omologia. Viene successivamente innescata la sintesi di DNA e la formazione di una forza di replicazione a bassa processività⁵⁰ (Figura 11).

L'appaiamento, che non necessita della proteina *RecA/Rad51*, ma solo di un piccolo tratto di omologia (micro-omologia appunto), può avvenire tra cromatidi fratelli sia a monte che a valle della posizione della forza di replicazione, nella

stessa direzione o in direzione invertite, conducendo a possibili duplicazioni, delezioni e inversioni^{39, 50}.

L'appaiamento mediato da micro-omologie con un cromosoma differente può condurre anche a traslocazioni³⁹. Il ripetersi del ciclo di appaiamento, estensione e separazione secondo MMBIR, si ritiene che possa essere responsabile di riarrangiamenti cromosomici non ricorrenti complessi^{10, 39}.

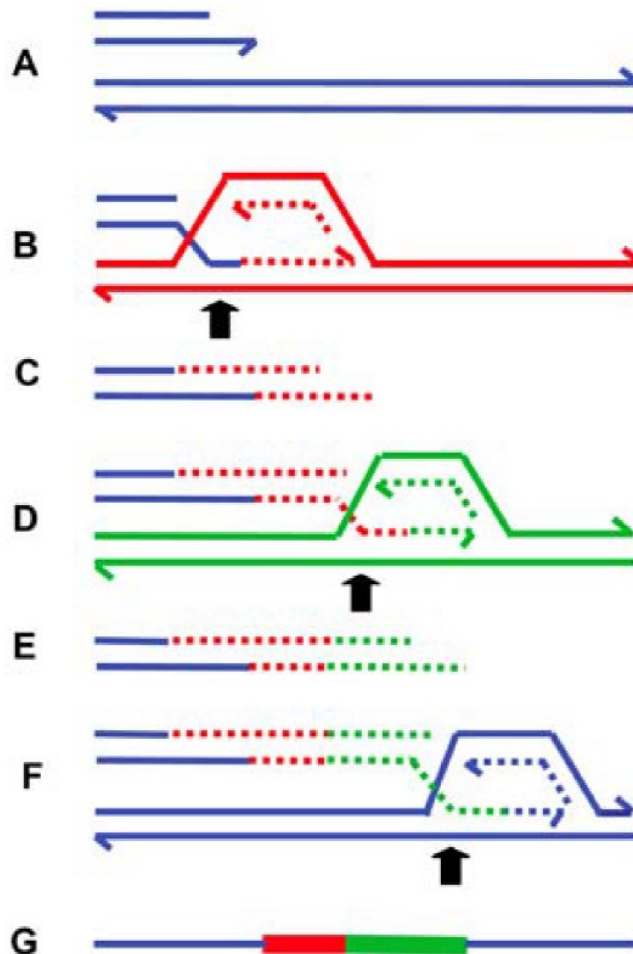


Figura 11. Schema del meccanismo replicativo Microhomology-mediated break-induced replication (MMBIR)

La figura mostra successivi passaggi verso diverse posizioni genomiche (con colori diversi), attraverso la formazione di giunzioni caratterizzate da micro-omologia (freccie nere). (A) mostra il braccio rotto di una forcella di replicazione collassata, che forma una nuova forcella di replicazione a bassa processività, come si vede in (B). L'estremità del filamento estesa si dissocia ripetutamente (C ed E) e forma una nuova forcella di replicazione con altri filamenti stampo (D e F). In (F) il passaggio avviene con il cromatidio fratello originario (in blu), formando una forcella che completa la replicazione. (G) mostra il prodotto finale del riarrangiamento, contenente sequenze di diverse regioni genomiche. Ogni linea rappresenta un filamento di DNA; la polarità del filamento è indicata dalla mezza freccia, che corrisponde all'estremità 3'.

La Tabella 1 riassume alcuni dei meccanismi coinvolti nella formazione dei riarrangiamenti genomici complessi.

Mechanism	Cause	DSB	Process	Breakpoint feature
FoSTeS	Stalled fork	Not required	Template switching between forks	Microhomology
MMBIR	Collapsed or broken fork	One-ended break	Template switching between forks	Microhomology
BIR	Broken fork	One-ended break	Homologous recombination-based template invasion between forks	Homology

Tabella 1. Alcuni dei meccanismi potenzialmente coinvolti nella formazione di riarrangiamenti complessi.

1.5 RIARRANGIAMENTI CROMOSOMICI COMPLESSI

Si definiscono complessi quei riarrangiamenti cromosomici che presentano tre o più punti di rottura del DNA. Recentemente, l'analisi dei punti di rottura tramite Next Generation Sequencing (NGS) ha permesso di definire la struttura genomica di questi riarrangiamenti, e di individuarne i meccanismi causativi.

Duplicazioni invertite adiacenti a delezioni terminali

Si tratta di riarrangiamenti di grandi dimensioni comunemente descritti sia in patologie tumorali che in disordini congeniti, principalmente disturbi del neurosviluppo.

Hermetz et al., hanno proposto un modello *fold-back* alla base della formazione di queste duplicazioni invertite: un iniziale rottura a doppio filamento del DNA determina una delezione terminale, a cui consegue la resezione dell'estremità 5' del cromosoma troncato; il tratto sporgente all'estremità 3', si ripiega, appaiandosi con se stesso, frequentemente in una regione di omologia di sequenza invertita⁵¹. A questo punto la sintesi di DNA colma l'intervallo resecato e si viene a creare un cromosoma dicentrico, che subisce un'altra rottura a doppio filamento del DNA, questa volta tra i due centromeri, riparata successivamente dall'aggiunta di un telomero. Questo modello di riarrangiamento è stato ipotizzato dopo il rilevamento, grazie all'analisi di sequenza di 34 punti di rottura, di sequenze

distanziatrici in normale numero di copie tra le duplicazioni invertite e di corte sequenze omologhe invertite all'estremità dei segmenti invertiti^{30,51} (Figura 12a).

Duplication-normal-duplication (DUP-NML-DUP)

Si tratta di duplicazioni adiacenti con una sequenza di DNA, in normale numero di copie, disposta tra loro, che presentano all'array-CGH una particolare pattern di ibridazione "DUP-NML-DUP" (Figura 12b). Le duplicazioni possono essere orientate nella stessa direzione, in tandem, oppure possono essere invertite. Queste duplicazioni interstiziali derivano da regioni dello stesso braccio cromosomico distanti da centinaia a migliaia di Kb; i riarrangiamenti DUP-NML-DUP non sono associati ad una particolare sindrome cromosomica, poiché derivano da diverse regioni del genoma e coinvolgono numerosi geni ³⁰.

Triplicazioni

Esistono due tipi di triplicazioni: le triplicazioni di tipo I sono orientate testa-coda, non sono fiancheggiate da duplicazioni e si formano per ricombinazione omologa non allelica (NAHR) tra duplicazioni segmentali (SD o LCR)^{23,30}; le triplicazioni di tipo II si trovano all'interno di grandi duplicazioni, e molto spesso il segmento triplicato è invertito rispetto alla duplicazione all'interno della quale si trova (DUP-TRP/INV-DUP)^{30,52,53} (Figura 12c).

Traslocazioni inserzionali

Si tratti di segmenti traslocati inseriti all'interno di cromosomi non omologhi, che frequentemente appaiono come semplici duplicazioni interstiziali dalla studio delle CNV (Figura 12d)³⁰; con le tecniche FISH e con l'analisi di sequenza dei punti di rottura è stato dimostrato che circa il 2% dei segnali di ibridazione in eccesso rilevati all'array-CGH si riferiscono in effetti a segmenti inseriti in un altro cromosoma^{30,54,55}.

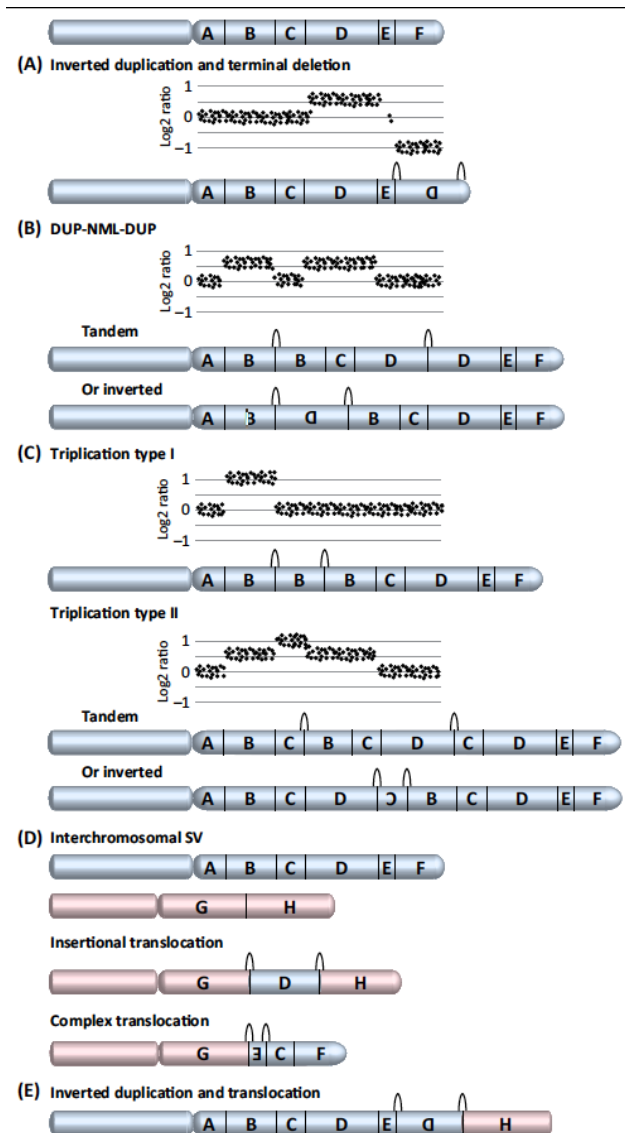


Figura 12. Riarrangiamenti cromosomici complessi.

I riarrangiamenti complessi e i loro pattern all'analisi array-CGH sono messi a confronto con il cromosoma di riferimento blu (in alto), diviso nei segmenti A-F. (a) Le duplicazioni invertite adiacenti a delezioni terminali presentano brevi sequenze distanziatrici disomiche (segmenti E) tra le duplicazioni invertite. (b) Una DUP-NML-DUP appare all'array come due incrementi nel numero di copie (segmenti B e D); le duplicazioni possono essere nella stessa direzione oppure un segmento duplicato (D) può essere invertito tra le due copie dell'altro segmento (B). (c) La triplicazione di tipo I ha tre copie di B orientate nella stessa direzione. Nel tipo II, la triplicazione è incorporata (C) all'interno della regione duplicata (B-D); il segmento triplicato potrà essere orientato correttamente o invertito. (d) Riarrangiamenti complessi intercromosomici avvengono tra il cromosoma blu e quella rosa. Una traslocazione inserzionale implica l'inserzione interstiziale di un segmento cromosomico (D) in un altro cromosoma. Alcune traslocazioni complesse presentano più segmenti cromosomici e/o inversioni ai punti di rottura. (e) Una duplicazione invertita con delezione terminale può finire con la porzione terminale traslocata di un cromosoma non omologo.

Cromoanagenesi: cromotrips e cromoanasintesi

Le tecniche di Whole Genome Sequencing (WGS) hanno permesso di scoprire un processo di riarrangiamento di struttura dei cromosomi, definito cromoanagenesi

(*Chromoanagenesis*, “*chromosome rebirth*”), alla base di alcune patologie tumorali e di alcune anomalie costituzionali dello sviluppo, in cui un grande numero di riarrangiamenti complessi avviene in uno o pochi cromosomi in un singolo catastrofico evento⁵⁶.

Sono stati riconosciuti due principali meccanismi alla base di questi riarrangiamenti catastrofici, ovvero la cromotripsi (*Chromothripsis*) e la cromoanasintesi (*Chromoanasyntesis*)^{57,58}. La cromotripsi (Figura 13) è stata individuata per la prima volta in una donna di 62 anni affetta da leucemia linfatica cronica, che presentava 42 riarrangiamenti somatici acquisiti localizzati sul braccio lungo del cromosoma 4⁵⁷.

Simili fenomeni di riarrangiamenti cromosomici complessi e localizzati, sono stati documentati in molte tipologie di tumore, come medulloblastoma, neuroblastoma, mieloma, melanoma, cancro del colon-retto, cancro della prostata e leucemia⁵⁹⁻⁶⁴: sembra che il 2-3% dei tumori nell'uomo abbia alla base un fenomeno di cromotripsi, in particolar modo per i tumori ossei e i glioblastomi^{59,62,65,66}. Kloosterman et al. sono stati i primi a descrivere il meccanismo di cromotripsi come causa di riarrangiamenti costituzionali complessi, eseguendo il sequenziamento del genoma di un paziente con ritardo dello sviluppo psicomotorio e del linguaggio, e di quello dei suoi genitori. Sono stati dimostrati 12 punti di rottura *de novo* confinati a livello dei cromosomi 1, 4 e 10, e dopo una loro dettagliata analisi è stato proposto che una serie di rotture del DNA a doppio filamento abbiano frantumato i cromosomi interessati e che i frammenti siano stati successivamente riuniti tra loro attraverso un meccanismo di NHEJ⁶⁷.

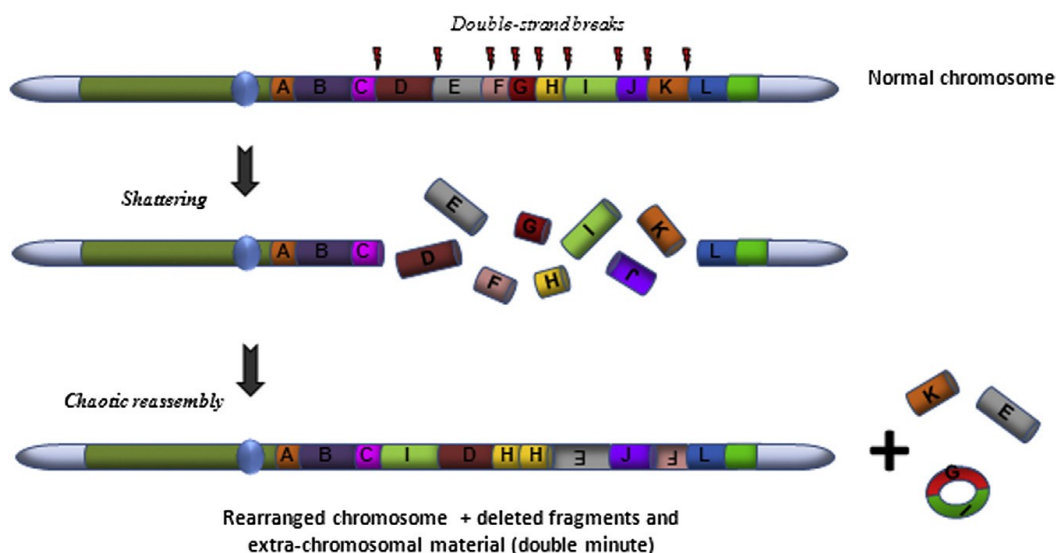


Figura 13. Rappresentazione grafica del meccanismo di cromotripsa

Evento cromosomico catastrofico in cui si verificano più rotture del doppio filamento, limitato ad una semplice segmento cromosomico o ad alcune regioni cromosomiche. La cromotripsa consiste nella polverizzazione di frammenti di cromosomi che vengono riassemblati insieme in maniera caotica.

Basandosi sui dati provenienti dagli studi su cellule tumorali e su pazienti affetti da disordini congeniti, è stato possibile identificare alcuni criteri di riconoscimento della cromotripsa: il fenomeno avviene in un unico catastrofico evento, che genera decine o centinaia di riarrangiamenti localizzati su uno o pochi cromosomi, con una tendenza dei punti di rottura e di giunzione a raggrupparsi in piccoli segmenti cromosomici (fino a 10 punti di rottura nell'arco di 50 Kb); la disgregazione dei cromosomi interessati causa una diffusa perdita di frammenti di sequenze nucleotidiche^{68,69}.

I frammenti cromosomici derivano dalla frantumazione di un solo cromosoma parentale e nella maggior parte dei casi vengono riassemblati senza alcun orientamento preferenziale o ordine; a livello dei punti di rottura non si riscontra omologia, o al massimo è possibile notare micro-omologie di pochi nucleotidi⁷⁰.

A livello delle regioni riarrangiate del genoma, il numero di copie oscilla tra 1 e 2, per cui si ha un regolare alternanza tra perdita e mantenimento dell'eterozigotà^{57,59,66,70} (Figura 14).

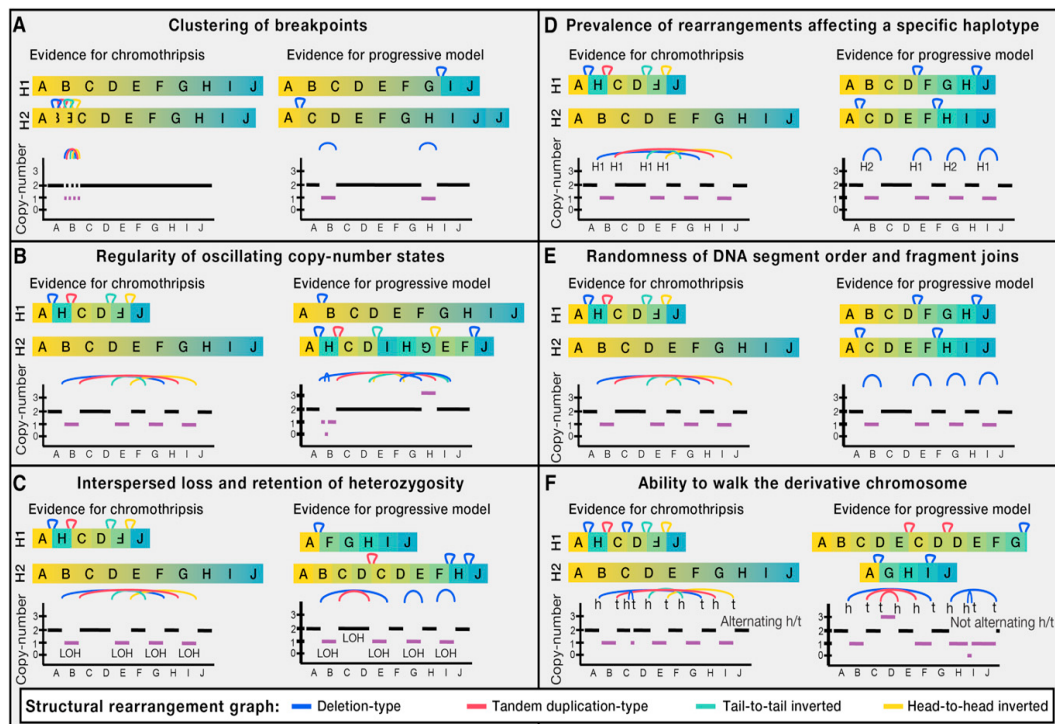


Figura 14. Criteri per il riconoscimento della cromotripsis.

- Punti di rottura raggruppati sono tipici del fenomeno della cromotripsis (a sinistra), mentre un così alto grado di raggruppamento non è rilevabile nei riarrangiamenti gradualmente (a destra).
- Nella cromotripsis si evidenzia un'oscillazione del numero di copie tra due stati (1 e 2), mentre nei riarrangiamenti gradualmente il numero di copie varia tra più stati diversi.
- La cromotripsis si caratterizza per un alternarsi regolare di regioni che hanno perso eterozigotà e altre che l'hanno mantenuta; i riarrangiamenti gradualmente non presentano un pattern altrettanto regolare.
- La cromotripsis avviene a livello di un singolo aploide e di un singolo cromosoma parentale, mentre i riarrangiamenti gradualmente non mostrano preferenze.
- Nella cromotripsis, i frammenti derivati dalla frantumazione del DNA sono uniti tra loro in modo casuale, mentre nei riarrangiamenti gradualmente si nota una preferenza per alcune forme di riarrangiamento, che risulta in un ordine non casuale e non uniforme dei frammenti riuniti.
- Nella cromotripsis, confrontando il segmento di DNA riarrangiato con il genoma di riferimento, vedremo che ogni segmento distinto inizia con un riarrangiamento alla "testa" e finisce con un riarrangiamento alla "coda".

Non è ancora chiaro quale sia il meccanismo che inneschi la cromotripsis, e quindi la disgregazione in frammenti di uno o pochi cromosomi.

Sono stati proposti diversi meccanismi capaci di indurre il fenomeno della cromotripsis, senza però che si sia raggiunta l'unanimità. Crasta et al. hanno dimostrato che la frammentazione del cromosoma può avvenire come risultato di una mis-segregazione durante la mitosi: errori nella segregazione cromosomica durante la mitosi, possono dar origine a cromosomi "lenti" che finiscono in un micronucleo all'interno della cellula. A questo punto la disgregazione del

cromosoma mis-segregato è innescata da un'asincronia della replicazione tra nucleo e micronucleo, ovvero da una prematura compattazione del cromosoma (PCC) nel micronucleo; il DNA può quindi essere polverizzato e successivamente integrato nel genoma⁷⁰. Altre ipotesi considerano che l'esposizione a stimoli esogeni, come radicali liberi o radiazioni ionizzanti durante la mitosi, in cui i cromosomi sono altamente condensati, possa innescare un processo di frantumazione del DNA, risultante in multipli punti di rottura raggruppati⁷¹.

E' possibile che anche lo stress replicativo sia coinvolto nel rendere più probabile un fenomeno di riarrangiamento genomico complesso, attraverso l'induzione del blocco delle forcelle di replicazione e del successivo ciclico scambio del filamento stampo, anche se alcuni autori ipotizzano sia più probabile che abbia un ruolo nella genesi della cromotripsia, piuttosto che della cromotripsia⁷². Alcuni esempi di cromotripsia nel cancro sono caratterizzati da riarrangiamenti in regioni che si estendono fino ai telomeri, e questo potrebbe suggerire un meccanismo di *breakage-fusion-bridge* (BFB) *cycle* alla base degli stessi riarrangiamenti. Tuttavia, una serie di cicli di BFB produrrebbe l'amplificazione di alcuni segmenti cromosomici, anche se tipicamente i cromosomi derivativi della cromotripsia non presentano ampliconi^{60,68,70}.

Rausch et al. hanno suggerito un legame tra lo stato della proteina TP53 e la cromotripsia: l'analisi con WGS di un medulloblastoma SHH (Sonic-Hedgehog) in un paziente affetto da sindrome di Li-Fraumeni (e quindi con TP53 mutata), ha rivelato un fenomeno di riarrangiamento cromosomico complesso, analogo alla cromotripsia⁶⁴. Sequenziando il genoma di altri pazienti è stata dimostrata una stretta correlazione tra TP53 mutata e cromotripsia, sia in medulloblastomi SHH che in casi di leucemia mieloide acuta. Una TP53 mutata potrebbe facilitare l'insorgenza della cromotripsia attraverso più vie: rendendo i cromosomi più suscettibili a rotture del DNA da esposizione a radiazioni ionizzanti, riparate poi preferenzialmente con NHEJ, oppure facilitando l'accorciamento dei telomeri e la fusione dei cromosomi alle estremità, o ancora contribuendo ad una prematura compattazione cromosomica (PCC), per alterazione del checkpoint nella transizione da fase G2 a M della mitosi⁶⁴. Inoltre, una mutazione di TP53 potrebbe avere un ruolo nel favorire la sopravvivenza della cellula all'evento catastrofico della cromotripsia^{70,71}.

Indipendentemente dal meccanismo causa della disgregazione cromosomica, le idee riguardo le procedure di assemblaggio dei frammenti sono più chiare: si ritiene che nella cromotripsi associata a cancro, poiché la maggior parte dei punti di rottura non presenta alcuna omologia di sequenza o solo piccole aree di micro-omologia, la riunione dei frammenti avvenga principalmente tramite NHEJ^{59,60}.

Per quanto riguarda la cromotripsi costituzionale, si ritiene che oltre al meccanismo di NHEJ siano coinvolti anche meccanismi di riparazione replicativi, come FoSTeS/MMBIR⁷³. La cromotripsi costituzionale esibisce, rispetto al fenomeno riscontrato nelle patologie tumorali, un minor numero di rotture del DNA, e i riarrangiamenti sembrerebbero tipicamente bilanciati, nonostante colpiscano multipli cromosomi; i connotati della cromotripsi costituzionale potrebbero essere frutto di un meccanismo molecolare diverso alla base del fenomeno, oppure potrebbero essere l'effetto della selezione⁷⁴. L'incidenza della cromotripsi nei pazienti con disordini congeniti è meno chiara che nel cancro: in uno studio, analizzando il genoma di 10 pazienti con riarrangiamenti genomici complessi, Kloosterman et al. hanno identificato 8 casi di cromotripsi costituzionale⁷³, mentre in un altro studio con 52 pazienti con riarrangiamenti bilanciati al cariotipo, Chiang et al. hanno rilevato 2 casi di cromotripsi *de novo*⁷⁴.

Riguardo l'origine della cromotripsi costituzionale, si pensa che il fenomeno di riarrangiamento catastrofico avvenga dopo la meiosi paterna, durante la spermiogenesi^{67,75}; non è stato neppure escluso che l'evento della cromotripsi possa manifestarsi negli stadi precoci dello sviluppo embrionale, generando poi un mosaico cromotriptico⁶⁷.

Liu et al., dallo studio di 17 pazienti con vari disturbi congeniti dello sviluppo, hanno identificato un nuovo pattern di riarrangiamento genomico costituzionale complesso, che hanno definito cromoanasintesi (sintesi di un nuovo cromosoma)⁷⁶. La cromoanasintesi è limitata ad un solo cromosoma o spesso alla porzione distale di un braccio cromosomico, e dà origine ad un cromosoma anomalo, con segmenti duplicati dovuti ad una sintesi extra del DNA.

Si osserva un aumento del numero di copie, con duplicazioni e triplicazioni, e l'analisi di sequenza a livello dei punti di giunzione ha rilevato micro-omologie e inserzioni di sequenze stampo (100-1000 pb), che suggeriscono un meccanismo replicativo di riparazione del DNA, come MMBIR, alla base dei riarrangiamenti^{23,27,44,71,75,76}. Nella cromoanasintesi, a differenza della cromotripsi,

non si ha perdita di eterozigotità (Figura 15). Nella cromosomasintesi i punti di rottura in serie non sono adiacenti, ma separati da lunghe distanze (0,01-1 Mb), e ciò sembra derivare da eventi di scambio del filamento stampo⁷².

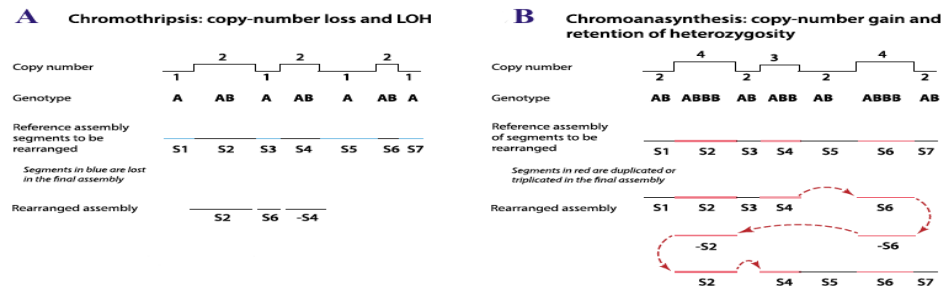


Figura 15. Confronto tra cromotrips e cromosomasintesi.

(a) Lo schema mostra il fenomeno di cromotrips con perdita dei frammenti genomici e riarrangiamento in modo casuale (b) cromosomasintesi che porta all'aumento del numero di copie dei frammenti riarrangiati che risultano duplicati e con punti di rottura non adiacenti tra loro.

1.6 METODI DI RICONOSCIMENTO DI VARIAZIONI STRUTTURALI DEL GENOMA

Bandeggio cromosomico

L'insieme completo di tutti i cromosomi metafasici di una cellula è definito cariotipo. L'identificazione inequivocabile di ogni cromosoma del cariotipo umano è stata possibile grazie a tecniche di trattamento e colorazione dei cromosomi, definite bandeggi: esistono diverse tecniche di bandeggio per le diverse regioni cromosomiche, delle quali riflettono la costituzione in basi e sequenze (uniche e ripetute) e lo stato trascrizionale. Il bandeggio G (Giemsa) e il bandeggio R (Reverse) sono specifici per l'eucromatina, ovvero la cromatina costituita da sequenze in singolo copia trascrizionalmente attive, e ne determinano una scomposizione in bande chiare e scure; il bandeggio Q è specifico per l'eucromatina, fornendo un pattern di bandeggio sovrapponibile al bandeggio G, e colora selettivamente anche il braccio lungo del cromosoma Y, nella sua porzione eterocromatica. I bandeggi G o R vengono usati routinariamente in ogni analisi cromosomica individuale, mentre il bandeggio Q è riservato principalmente allo studio del braccio lungo del cromosoma Y. Il bandeggio C colora selettivamente

l'eterocromatina centrometrica e pericentromerica, mentre l'eucromatina delle braccia corte e lunghe rimane pressoché incolore; la colorazione NOR è specifica per i satelliti dei cromosomi acrocentrici 13, 14, 15, 21 e 22. Il limite di risoluzione dell'esame di bandeggio cromosomico è circa 4-6 Mb; in caso di anomalie del pattern di bandeggio o in caso di alterazioni nella struttura dei cromosomi, sarà possibile individuare grandi delezioni, duplicazioni o traslocazioni, mentre non sarà comunque possibile rilevare anomalie cromosomiche di piccole dimensioni³⁰.

Ibridazione in situ fluorescente (FISH)

La tecnica FISH si basa sul principio di complementarietà dei due filamenti di DNA; il DNA di una sonda nota subisce un processo di marcatura, ovvero viene tagliato da una DNAasi e successivamente viene "ricucito" da una DNA polimerasi, in presenza di un nucleotide direttamente fluorescente oppure di uno specifico nucleotide contenente una molecola reporter, come la biotina, che sarà poi riconosciuta da un anticorpo fluorescente. Si procede poi alla denaturazione sia del DNA della sonda nota che del DNA oggetto di studio, e così i due DNA denaturati vengono ibridati. Se la sonda nota trova il suo complementare sul cromosoma in esame, l'ibridazione avviene e sarà visibile con il microscopio a fluorescenza; se l'ibridazione non ha luogo, non avremo alcun segnale al microscopio a fluorescenza. Le sonde molecolari utilizzate nella FISH possono essere di vario genere: sonde α satelliti centromeriche, sonde painting cromosoma-specifiche, che illuminano la cromatina di un intero cromosoma, e sonde locus-specifiche. L'analisi FISH permette di individuare anche piccole anomalie cromosomiche strutturali, fino a 40 kb di lunghezza, ed è capace di individuare sia variazioni strutturali quantitative, come delezioni o duplicazioni, sia qualitative, come inversione o traslocazioni bilanciate³⁰.

Array comparative genome hybridization (Array-CGH)

La tecnica prevede l'ibridazione del DNA del soggetto in esame e di un DNA di controllo, marcati ciascuno con un diverso fluorocromo, su piccole sequenze di DNA disposte su un vetrino, a costituire un array; la risoluzione di un esame di array-CGH può variare da 1 Mb con i BAC-array a 3,5-100 kb con gli

oligonucleotide-array. Se il DNA del soggetto in esame e il DNA di controllo sono equimolari, non si apprezzeranno differenze di ibridazione, in eccesso o in difetto, tra i due DNA; invece nel caso di delezioni o duplicazioni, il segnale di ibridazione sarà, rispettivamente, a favore del DNA di controllo o a favore del DNA testato. Tramite la tecnica array-CGH è possibile individuare le anomalie quantitative del genoma, mentre non è possibile rilevare quelle soltanto qualitative, che non si esprimono in varianti del numero di copie (CNV)⁷⁷.

Next Generation Sequencing (NGS)

Per sequenziamento si intende la lettura della sequenza esatta di nucleotidi di un acido nucleico, RNA o DNA. L'analisi della sequenza nucleotidica di un filamento di DNA o di RNA è possibile solo per frammenti contenenti circa un centinaio di nucleotidi, che vengono poi allineati e assemblati per ricostruire la sequenza che si vuole conoscere. Il primo metodo di sequenziamento del DNA è il metodo Sanger, definito anche metodo a terminazione di catena, con un *throughput* basso, costi e tempi elevati; i metodi di sequenziamento di nuova generazione vengono definiti ad *high throughput*, garantendo l'analisi di sequenza di numerosissimi frammenti di DNA contemporaneamente^{78,79} (Figura 16).

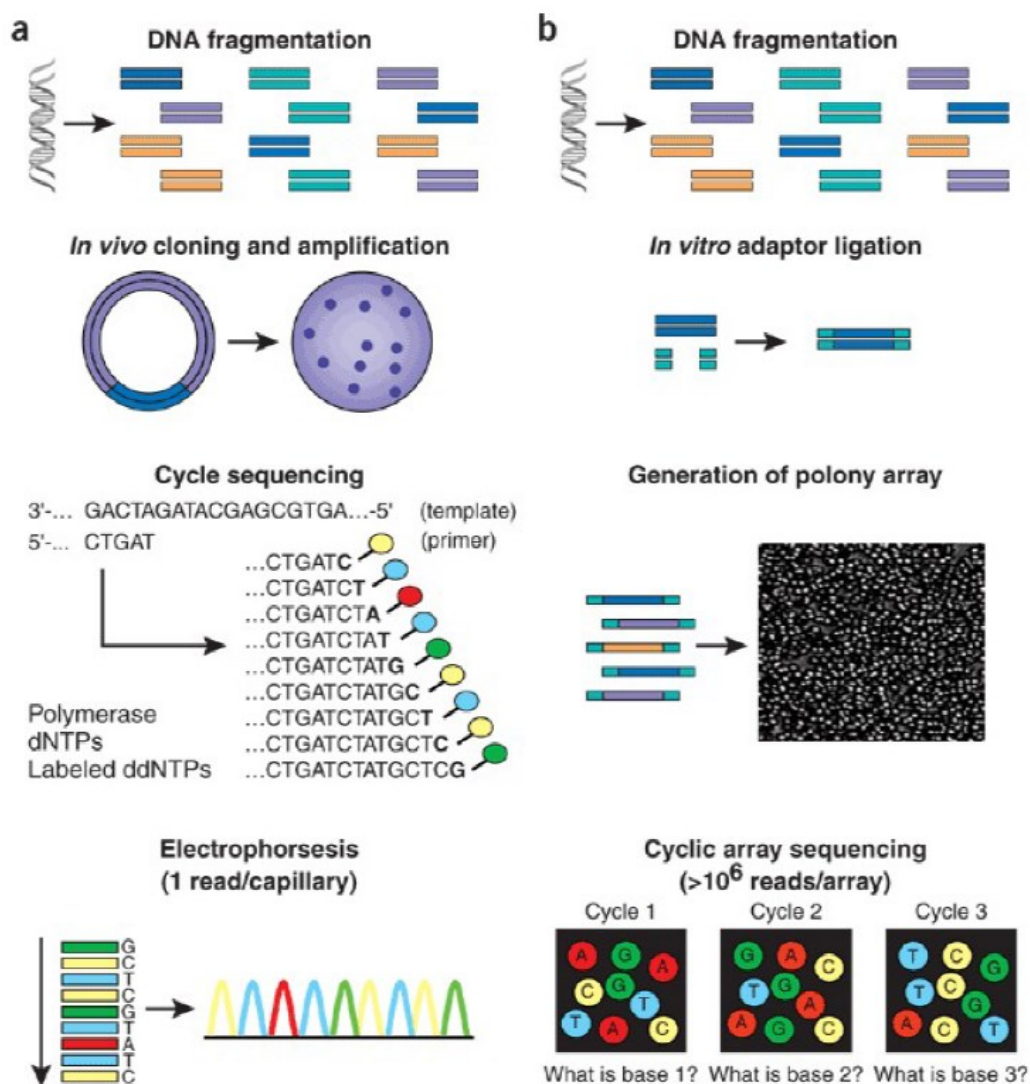


Figura 16. Flusso di lavoro del metodo Sanger (a) e dei metodi NGS (b).

(a) Per un sequenziamento Sanger ad alta processività, il DNA genomico è frammentato, poi clonato in vettori e trasformato utilizzando *E.coli*. Per ogni reazione di sequenziamento, viene prelevata una singola colonia ed estratto il DNA. Ogni reazione di sequenziamento genera dei frammenti marcati con nucleotidi terminatori che vengono sottoposti ad elettroforesi capillare con 96 o 384 capillari in una singola corsa di sequenziamento. Essendo frammenti marcati con molecole fluorescenti, saranno riconosciute da un laser a 4 canali di emissioni in modo da poter generare una traccia. (b) Nel sequenziamento di *shotgun* con metodiche di nuova generazione, adattatori universali sono legati ai frammenti genomici di DNA, che sono sottoposti a diversi protocolli al fine di creare milioni di polonie che saranno immobilizzate su un array. Ogni polonia consiste in molte copie di una singola libreria *shotgun*. Visto che le polonie sono immobilizzate su un array possono essere sequenziate contemporaneamente. L'incorporazione di molecole marcate con fluorescenza sono incorporate ed ogni estensione può essere acquisita in parallelo generando immagini che poi saranno convertite in *reads* per ogni array.

Esistono diverse piattaforme di NGS che si basano su diversi processi biochimici, ma con una medesima catena di passaggi fondamentali: per prima cosa è necessaria la preparazione di librerie di DNA che verranno amplificate e successivamente, inizia la vera fase di sequenziamento con conseguente

immagazzinamento dei dati. Le principali piattaforme di NGS sono Illumina e *Ion Torrent* di Life Technologies. Il sequenziatore Illumina utilizza una *flow cell* consistente in una “slide” otticamente trasparente costituita, sulla superficie, da 8 comparti a cui sono legati degli oligonucleotidi di ancoraggio. Il DNA stampo è frammentato in segmenti lunghi alcune centinaia di basi e modificato per generare estremità troncate 5'-fosforilate; una singola base di adenina viene aggiunta all'estremità 3' del frammento di DNA stampo fosforilato, per preparare i frammenti di DNA all'aggiunta di oligonucleotidi adattatori. Gli oligonucleotidi adattatori sono complementari agli oligonucleotidi ancorati alla *flow cell*; il DNA stampo a singolo filamento legato agli adattatori è aggiunto alla *flow cell* e immobilizzato mediante ibridazione agli oligonucleotidi di ancoraggio. I frammenti di DNA vengono amplificati mediante un meccanismo di amplificazione “a ponte”, che si basa sulla cattura dei filamenti di DNA ripiegati ad arco che si ibridano ad un oligonucleotide di ancoraggio adiacente; cicli multipli di amplificazione convertono la singola molecola di DNA stampo in un “cluster” di frammenti ripiegati, amplificati clonalmente, ciascuno composto approssimativamente da 1000 ampliconi clonali. Per il sequenziamento, i “cluster” sono denaturati, e un successivo lavaggio consente di ottenere solo i filamenti “senso”. Il sequenziamento dei filamenti “senso” avviene mediante l'ibridazione di un “primer” complementare alla sequenza dell'oligonucleotide adattatore; successivamente si aggiungono una DNA polimerasi e una miscela di 4 nucleotidi terminatori “reversibili” marcati con fluorofori diversi: i nucleotidi terminatori sono incorporati in base alla complementarietà della sequenza di ciascun filamento, all'interno di ogni “cluster” clonale. Dopo incorporazione, i reagenti in eccesso vengono rimossi con un lavaggio e la fluorescenza relativa ad ogni “cluster” viene rilevata e registrata; con alcuni passaggi, viene rimosso il blocco dei nucleotidi terminatori e così anche la marcatura fluorescente, consentendo di effettuare un nuovo ciclo di sequenziamento⁸⁰ (Figura 17).

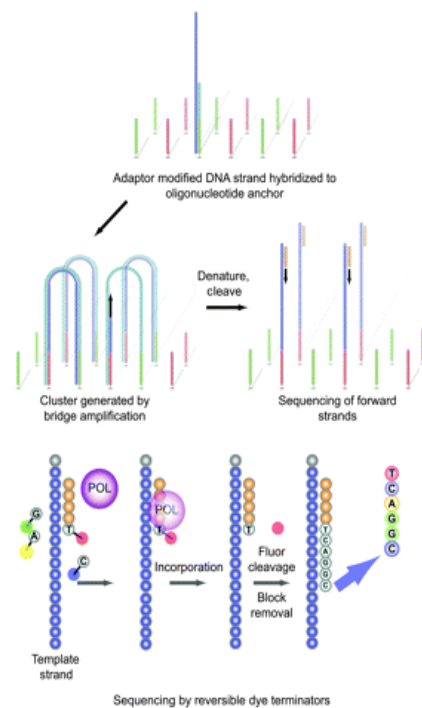


Figura 17. Funzionamento del sequenziatore Illumina.

Frammenti di DNA a cui sono legati degli adattori universali. Questi frammenti sono depositati su un array su cui sono *spottate* sequenze complementari agli adattatori in modo da formare delle strutture a ponte che sono amplificate in modo clonale. Si formano dei *cluster* di frammenti di DNA, polonie, che sono singolarmente sequenziate con l'incorporazione di nucleotidi terminatori marcati con fluorescenze diverse. I nucleotidi saranno colpiti da un laser e l'emissione di luce registrata da una CCD camera.

Il sequenziamento *paired-end* (Figura 18), permette di sequenziare un dato frammento di DNA in direzioni opposte partendo da entrambe le sue estremità: le *reads* ottenute con sequenziamento *paired-end* forniscono un allineamento più accurato di regioni di DNA contenenti sequenze ripetute, permettono una ricerca più sensibile di riarrangiamenti genomici, come inserzioni, delezioni o inversioni; inoltre producono sequenze più lunghe utili per il processo di assemblaggio de novo.

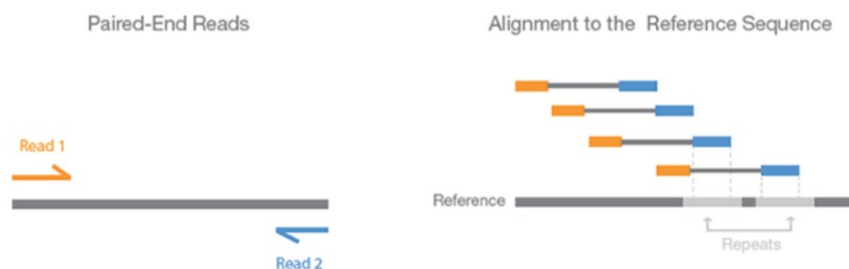


Figura 18. Sequenziamento *paired-end*.

Il sequenziamento *paired-end* è capace di sequenziare un frammento di DNA a partire da entrambe le estremità; dato che la distanza tra le *paired reads* è nota, l'allineamento delle *reads* sarà più accurato, e quest'informazione sarà utile soprattutto nel caso di regione del genoma altamente ripetitive.

Il sequenziamento *mate pair* (Figura 19), prevede la realizzazione di librerie di *paired-end reads* di lunghezza maggiore (rispetto al metodo *paired-end sequencing*), utili per il sequenziamento del DNA *de novo*, per la ricerca di varianti strutturali, riarrangiamenti genomici complessi e assemblaggio del genoma⁸⁰.

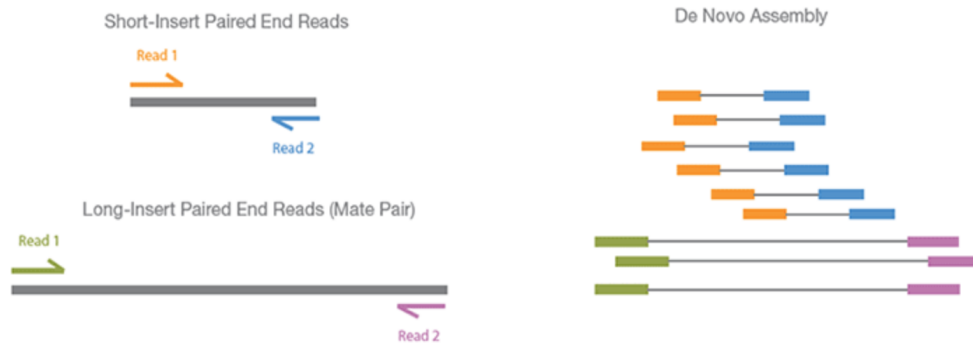


Figura 19. Sequenziamento *mate pair*.

La NGS ha molteplici campi di applicazione, tra cui il sequenziamento dell'intero genoma (WGS, Whole Generation Sequencing), con lo scopo di ottenere la migliore analisi delle variazioni strutturali del genoma; nonostante la possibilità di sequenziare tutto l'esoma e tutto il genoma, l'analisi di sequenziamento tende a concentrarsi su specifici geni o regioni genomiche selezionate sulla base del sospetto di malattia (*Targeted Resequencing*). Il sequenziamento del trascrittoma (RNA-seq) permette di individuare le differenze di espressione genica nei vari tessuti, o in diverse fasi dello sviluppo cellulare e dell'organismo^{79,80}

2. SCOPO DELLA TESI

In questo lavoro di tesi noi abbiamo voluto descrivere tre pazienti affetti da disabilità intellettiva e ritardo psicomotorio che presentano all'interno del loro genoma differenti complessi riarrangiamenti coinvolgenti il cromosoma 17,12 e 4, nei quali abbiamo potuto osservare appunto un meccanismo di cromotripsi.

Questi riarrangiamenti sono stati definiti con diverse metodologie di analisi e per la prima volta, almeno per dei casi costituzionali, con un sequenziamento massivo di tutto il genoma.

Il nostro scopo è stato quello di cercare di definire esattamente il tipo di complessità dell'anomalia cromosomica, di trovare gli esatti punti di giuntura dei riarrangiamenti e di cercare di definire la correlazione genotipo/fenotipo per ciascun caso.

Oggi grazie all'utilizzo della metodica di Next Generation Sequencing che può essere usata di routine, possiamo arrivare alla diagnosi e all'ulteriore caratterizzazione di questa nuova classe di riarrangiamenti che probabilmente sono più frequenti di quanto abbiamo ad oggi potuto ipotizzare.

3. MATERIALI E METODI

Nel nostro studio sono stati valutati da un punto di vista clinico tre pazienti in età pediatrica giunti presso l'ambulatorio di Genetica Medica dell'AOU "Anna Meyer" di Firenze per ritardo psicomotorio e disabilità intellettiva.

Successivamente sono stati analizzati i loro campioni di DNA.

Tutte le indagini sono state eseguite previa raccolta del consenso informato, sottoscritto dai genitori dei bambini. I campioni di DNA sono stati resi anonimi mediante l'apposizione di un codice alfa-numerico.

3.1 MATERIALI

CASO C/996

Il bambino di 3 anni e 5 mesi è secondogenito di genitori sani, che riferiscono di non essere consanguinei; ha una sorella maggiore, nata nel 2005, in apparente stato di buona salute. All'anamnesi familiare si segnala la presenza di ritardo psicomotorio in un fratello del nonno materno. Negata familiarità per difetti congeniti e poliabortività, in entrambi i rami parentali (Figura 20).

Il bambino nasce con taglio cesareo (TC, per precedente TC) alla 39° settimana di seconda gravidanza, caratterizzata da decorso regolare, salvo minacce d'aborto nelle prime settimane. I controlli ecografici sono nella norma e i MAF (movimenti attivi fetali) riferiti validi. Alla nascita il bambino presenta i seguenti parametri antropometrici: peso di 3140 g (circa 25° centile per età), lunghezza di 50 cm (circa 50° centile per età) e OFC di 36 cm (> 50° centile per età). Subito dopo la nascita il bambino mostra ipotonia con suzione ipovalida.

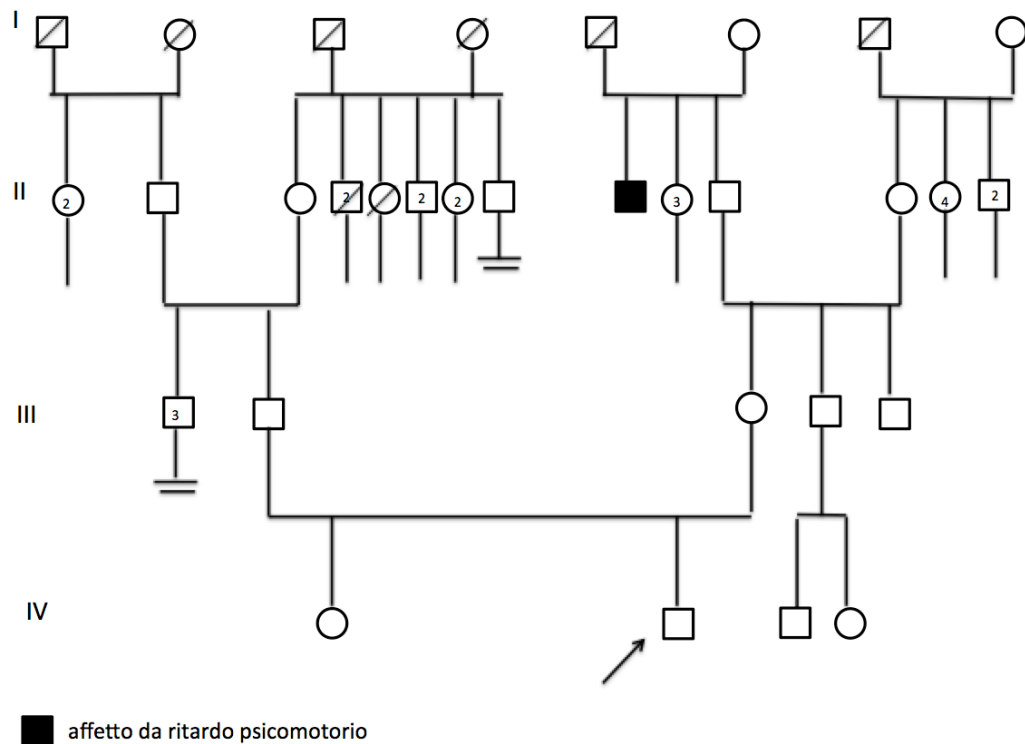


Figura 20. Albero genealogico del paziente C/996.

Ha presentato un ritardo nell'acquisizione delle prime tappe dello sviluppo psicomotorio: acquisizione della posizione seduta a circa 14-15 mesi e tuttora deambulazione con sostegno, e produzione verbale limitata a poche parole. Per la comparsa di spasmi affettivi (8-9 mesi) e di clonie, il bambino viene sottoposto ad accertamenti clinico-strumentali, tra cui una RMN, che evidenzia la chiusura dei canali auricolari. Alla nostra osservazione, all'età di 3 anni e 5 mesi, si rilevano peso di circa 14 kg (10° centile per età), lunghezza di circa 102 cm (circa 75° centile per età) e CC di 51.5 cm (circa 75° centile per età). Il fenotipo clinico si caratterizza per:

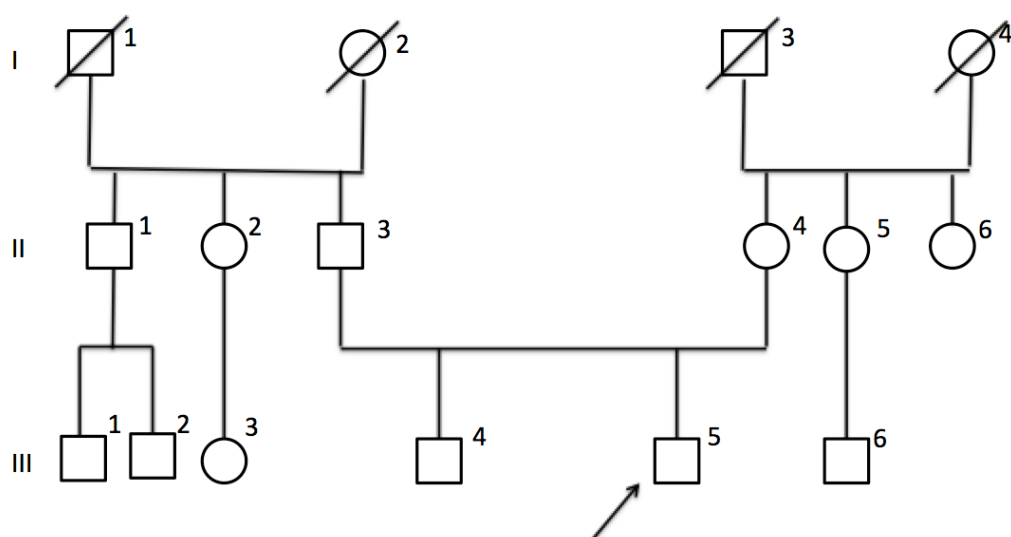
- lineamenti grossolani, fronte alta e bombata, capelli di colore castano chiaro, un po' ruvidi e arricciati sulla fronte,
- sopracciglia rade nel terzo mediale, rime palpebrali oblique verso il basso, ptosi palpebrale prevalentemente a destra, ipertelorismo,
- radice del naso rilevata,
- filtro nasolabiale corto e piatto,
- bocca tendenzialmente aperta, malocclusione, morso inverso,
- viso triangolare, mento piccolo a punta,

- padiglioni auricolari sporgenti, retrorotati, lievemente a basso impianto, di lunghezza 5.7 cm bilateralmente,
- mani in atteggiamento in chiusura con pollice in opposizione,
- cute: lentiggini diffuse sul tronco, alcune lentiggini nella regione ascellare destra, CALS (macchie caffè-latte) di circa 3x2 cm di diametro a destra (regione scapolare, fianco, regione inguinale), 3 CALS (3x1cm, 2x1cm, 1.5 x1cm) sulla parte anteriore della coscia destra, 1 CALS (3x1cm) sulla parte posteriore della coscia sinistra, alcune CALS di dimensioni 1x1cm, 1x 0.5 cm a livello del ginocchio destro,
- addome piatto, trattabile.

Alla visita, mostra ipotonia e apparente ipotrofia muscolare degli arti; le punte dei piedi sono tenute in estensione ma si possono ridurre passivamente. Il bambino non ha acquisito controllo sfinterico; il sonno viene riferito regolare. Al primo impatto risulta socievole, ma si preoccupa molto, e durante la visita piange.

CASO E/1021

Il bambino di 9 anni e 8 mesi è secondogenito di genitori non consanguinei e in apparente stato di buona salute; ha un fratello maggiore di circa 11 anni, riferito in buona salute. Nel ramo paterno della famiglia si segnala che una cugina di primo grado, figlia di una zia, è deceduta a circa due anni e mezzo a seguito di ricorrenti episodi di natura epilettica non meglio specificati; è stata inoltre riferita la presenza di patologia ematologica non meglio precisata in due cugini paterni di primo grado, tra loro fratelli solo da parte di padre (Figura 21). Il bambino nasce con parto spontaneo a 38 settimane, da seconda gravidanza normodecorsa. Indice di Apgar 9I-10V. Il decorso neonatale è regolare: il bambino è alimentato con latte materno, con suzione riferita valida; divezzamento normocondotto a partire dai 6 mesi. L'accrescimento staturo-ponderale è regolare, con parametri attuali di peso e statura, rispettivamente maggiori e uguali al 75° centile per età (altezza ai limiti superiori del target genetico); CC di 54 cm (> 50° centile per età, in accordo con la media parentale).



III3, deceduta a 30 mesi per episodi critici non meglio specificati
 III1 e III2, affetti da patologia ematologica non meglio specificata

Figura 21. Albero genealogico del paziente E/1021.

Il bambino raggiunge con ritardo variabile le prime tappe dello sviluppo psicomotorio: deambulazione autonoma dai 16 mesi, e linguaggio caratterizzato da primi bisillabi significativi a 12 mesi, e successiva lenta evoluzione con anomalie fonologiche e morfosintattiche, che richiedono un trattamento logopedico, tuttora in corso. I test psicometrici, effettuati intorno ai 7 anni di età, evidenziano un quadro caratterizzato da abilità cognitive in fascia di funzionamento borderline (Scala WISC IV: QIT 70), disturbo del linguaggio con importante compromissione in produzione e elementi di disprassia motoria. Sono riferiti progressivi miglioramenti sia in ambito linguistico che motorio, con riduzione dell'impaccio motorio e delle difficoltà di coordinazione che hanno caratterizzato i primi anni di vita del bambino. Vengono rilevati buon tono muscolare, ROT normoeccitabili ai quattro arti e discreta lassità legamentosa. Alle valutazioni oculistiche si apprezza distrofia corneale. All'esame obiettivo non si notano anomalie fisiche minori significative, da segnalare solo una macchia ipercromica ovalare in sede nucale e la presenza di fetal pads alle mani. Alla visita il bambino si presenta collaborante e in buone condizioni generali, non vengono riferiti episodi patologici significativi negli ultimi due anni (riferite pregresse otiti ricorrenti con perforazione timpanica sinistra, trattata con miringoplastica)

La bambina è giunta per la prima volta alla nostra osservazione per la presenza di ritardo psicomotorio associato a scarso accrescimento e tratti peculiari dello sviluppo fisico.

E' nata da seconda gravidanza, da genitori non consanguinei, riferita normodecorsa, salvo la presenza, nel primo trimestre, di minaccia di aborto, che hanno richiesto riposo a letto per circa 1 mese (Figura22). Le biometrie fetali vengono riferite nella norma, come anche i movimenti fetali attivi (MAF). E' stata anche eseguita amniocentesi (per ansia materna) che ha evidenziato la presenza di un cariotipo fetale femminile normale (46,XX).

Il parto è avvenuto, tramite taglio cesareo alla 38+5 settimana di gestazione dopo rottura spontanea delle membrane. La bambina presentava microcefalia e basso peso alla nascita (peso 2.800 g (3° percentile); lunghezza 50 cm (25° percentile); circonferenza cranica (CC) 32 cm (<3° centile). Il follow-up a 2, 5 e 8 mesi, rispettivamente, ha mostrato sempre microcefalia, basso peso, radice nasale ampia, ipertelorismo, ciglia lunghe, orecchie carnose e bocca larga.

A 2 mesi, si è chiusa la fontanella anteriore. Una risonanza magnetica a 5 mesi ha mostrato agenesia del commessura anteriore e del corpo calloso rostrale e agenesia parziale del setto pellucido. Non c'erano altre malformazioni o disfunzioni d'organo, quali malformazioni del sistema visivo, cisti interemisferiche o indicazioni di disfunzione ipofisaria. Non è stata in grado di mantenere il capo sollevato prima dei 7 mesi di vita. A 8 mesi ha mostrato un chiaro ritardo dello sviluppo psicomotorio corrispondente ad una età evolutiva di 4-5 mesi. A 7 mesi di vita mostrava anche segni di cardiomiopatia dilatativa. Successivamente comparivano ipertricosi, ipotiroidismo convulsioni, all'inizio dopo un episodio febbrile e successivamente anche in apiressia.

Un cariotipo convenzionale in bande G risultava nella norma (46,XX).

A 10 mesi di vita ha sviluppato ripetuti episodi di apnea in assenza di crisi che a volte hanno richiesto rianimazione. Ha iniziato a stare seduta da sola a 21 mesi. Gli esami metabolici erano nella norma. All'età di 1 anno (alla nostra osservazione) il peso era sempre al 3° centile, la lunghezza al 25° e la CC inferiore al 3° centile. Erano sempre presenti segni di cardiomiopatia e ipotonia.

All'ultima osservazione clinica a circa 3 anni di vita mostrava occhi infossati, ipoplasia medio-facciale, con le sopracciglia orizzontalizzate, grosso naso con radice appiattita mento a punta, basso impianto delle orecchie e orecchie a basso impianto e retrorotate. Inoltre era presente brachi-camptodattilia e piedi piccoli. RMN di controllo mostrava anche polimicrogiria.

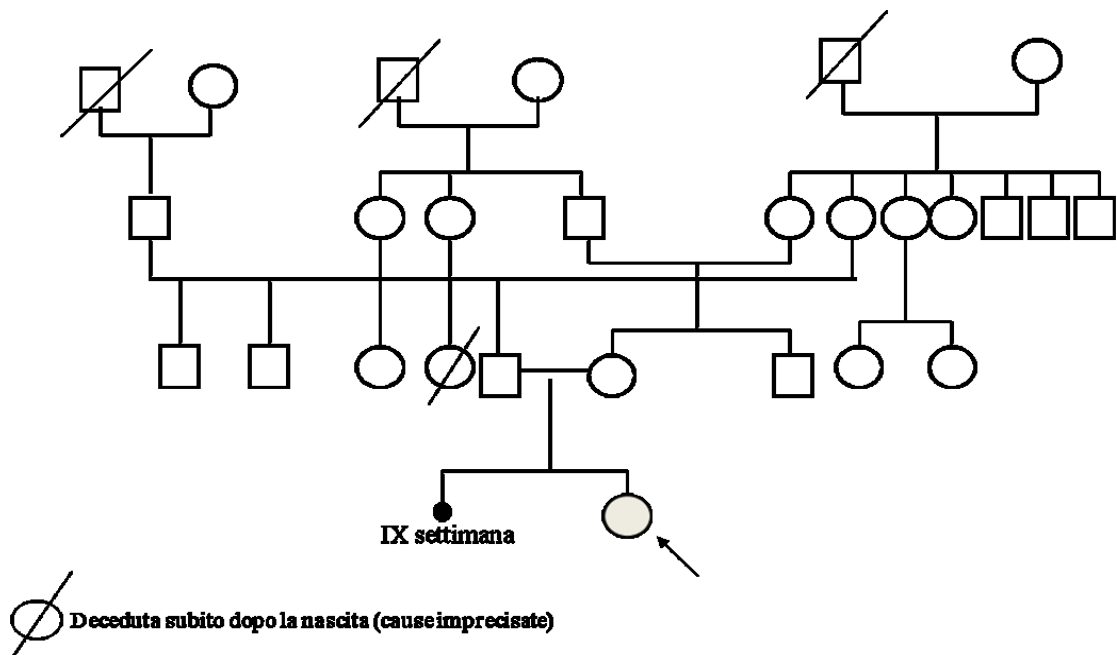


Figura 22. Albero genealogico del paziente.

3.2 METODI

Protocollo di estrazione del DNA genomico

Il DNA dei tre pazienti dello studio è stato estratto da prelievi di sangue periferico in EDTA. L'estrazione del DNA genomico è stata eseguita seguendo il protocollo del Kit QiAMP DNA Minikit (Qiagen). La concentrazione e la qualità del DNA sono state misurate con uno spettrofotometro NanoDrop ND-1000 UV sul quale viene caricato 1µl di DNA che ci fornisce sia la concentrazione del campione, espressa in ng/µl, sia i rapporti di assorbanza 260/230 e 260/280 nm, i quali permettono di stimare la purezza del DNA da vari contaminanti quali proteine, polisaccaridi e fenoli. Se il campione risulta soddisfacente (indice di purezza intorno a 2), il campione verrà impiegato per l'analisi con array-CGH e sequenziamento genomico.

Identificazione di CNVs

L'array-CGH è stato eseguito utilizzando il 4x180K (SurePrint G3 Human CGH Microarray Kit). Agilent Human Genome CGH microarray è un array a doppio colore contenente 60-mer di oligonucleotidi sintetizzati in situ utilizzando la tecnologia a getto d'inchiostro SurePrint di Agilent. I microarray SurePrint G3 sono caratterizzati da un'alta sensibilità e un'alta specificità che aumentano la risoluzione per mappare CNVs, punti di interruzione nelle varie aberrazioni e identificare i guadagni o le perdite di sequenze molto piccole fornendo allo stesso tempo la copertura completa dell'intero genoma. Il protocollo è costituito da vari step:

Step 1. Digestione del DNA

E' necessaria una concentrazione di DNA di 50ng/μl, dopodiché ad ogni campione sono aggiunti 2,9 μl di Digestion Master Mix e incubati per 2 ore a 37°C, per permettere agli enzimi di tagliare il DNA in frammenti di 400bp, e per 10 minuti a 65°C, per inattivare gli enzimi stessi (Tabella 2).

Mix di reazione

Nucl Free Water	1 μl
10x buffer C	1,3 μl
Acetylated BS	0,1 μl
AIU I	0,25 μl
RSA I	0,25 μl

Tabella 2. Componenti della Digestion Master Mix.

Step 2. Marcatura del DNA

Ai campioni e ai controlli digeriti sono stati aggiunti 2,5μl di Random Primers e incubati a 95°C per 5 minuti per permettere l'apertura della doppia elica. Dopodiché è stata aggiunta la Labeling Master Mix con cianine differenti per il campione (Cy5) e per il controllo (Cy3). I campioni e i controlli sono stati incubati

per 2 ore a 37°C e a 65°C per 10 minuti (Tabella 3). La marcatura deve avvenire rigorosamente in assenza di luce diretta sulle provette in modo da non interferire con il funzionamento delle cianine, che sono sensibili alla luce.

Mix di reazione

5X buffer	5 µl
10X dNTP	2,5 µl
Cy3 (controlli)	1,5 µl
Cy5 (campioni)	1,5 µl
Exo-Klenow	0,5 µl

Tabella 3. Componenti Labeling Master Mix.

Step 3. *Purificazione del DNA marcato*

Si tratta di purificare tutti i campioni per eliminare l'eccesso di marcatura e i frammenti non marcati. Questa procedura va fatta al buio. Viene utilizzato il SureTag labeling kit Purification column (Agilent Technologies). Il contenuto marcato viene riposto nella colonnina di purificazione con il buffer TE 1X e centrifugato. Per controllare l'attività specifica della cianina, 1 µl di DNA marcato e purificato viene dosato al Nanodrop.

Step 4. *Ibridazione*

Per calcolare l'attività specifica di ogni campione, si utilizza questa formula: pmol/µl Cy3 o Cy5/ug/µl DNA. Questo valore deve essere compreso tra 25 e 40 per Cy3 e tra 20 e 35 per Cy5. Si passa poi a calcolare la resa del campione ($[DNA] \times 8$), in questo modo verranno accoppiati quei campioni che presenteranno una resa ed una attività specifica simile tra loro. Si prepara la Hybridization Master Mix e se ne aggiungono 29 µl (Tabella 4). Per permettere al COT-DNA di ibridarsi con le regioni altamente ripetute, il campione viene incubato per 5 minuti a 95°C e per 30 minuti a 37°C (preannealing). Si procede poi con l'ibridazione sul vetrino aggiungendo 42 µl di campione. Tale vetrino viene poi incubato in rotazione (20rpm) in un fornetto a 65°C per 24 ore.

COT2	2 μ l
10X Block. Sol.	4,5 μ l
2XHRPM	22,5 μ l

Tabella 4. Componenti per la Hybridization Master Mix.

Questo passaggio consente di eliminare l'eccesso di DNA non ibridato. A questo scopo si utilizzano due diverse soluzioni di lavaggio: Oligo array-CGH Wash Buffer 1, a temperatura ambiente per 5 minuti, e Oligo array CGH Wash buffer 2, scaldata a 37°C per 1 minuto nelle quali verranno introdotti i vetrini contenenti i DNA ibridati.

Step 6 e 7. Scansione degli array e analisi dei risultati

La scansione degli array è stata eseguita con uno scanner Agilent (SureScan DX Microarray Scanner) che è in grado di rilevare i livelli di fluorescenza (Cy5 e Cy3) di ogni spot e fornisce l'immagine su cui poi si effettuerà l'elaborazione dei dati grazie all'utilizzo del software "feature extraction" (Agilent). Esso elabora per ogni spot il logaritmo in base 2 del rapporto tra le intensità di fluorescenza emesse da Cy5 e Cy3 ($\log_2 \text{Cy5/Cy3}$) inoltre sottopone l'informazione ottenuta dalla scansione del vetrino ad un processo di normalizzazione globale (sottrazione del background e correzione mediana) in grado di rimuovere le distorsioni sistematiche presenti. Il rapporto delle intensità di fluorescenza tra i due fluorofori è proporzionale al rapporto di quantità di materiale genomico tra campione e controllo. Se le intensità dei fluorofori sono uguali la regione analizzata non presenta alterazioni: avrò 2 copie del DNA da testare e 2 copie del DNA di controllo, per cui il rapporto tra i due fluorocromi è pari a 1 ($\log_2 2/2 = 0$). Se invece il rapporto Cy5/Cy3 è alterato si è in presenza di amplificazioni o delezioni: in caso di trisomia il rapporto sarà pari a 1,5 cioè $3/2$ ($\log_2 3/2 = 0,5$), viceversa in caso di monosomia il rapporto sarà pari a 0,5 e cioè $1/2$ ($\log_2 1/2 = -1$). Durante il processo di estrazione e analisi dei dati il software *feature extraction* elabora files .txt e QC reports grazie ai quali si possono visualizzare i dati elaborati e verificare la qualità

del vetrino. L'analisi è stata effettuata mediante il software Genomic Workbench (Agilent Technologies).

Studio dell'origine parentale delle CNVs

L'array-CGH è stato eseguito utilizzando il 4x180K (SurePrint G3 Human CGH+SNP Microarray Kit Agilent Technologies). Agilent Human Genome CGH + SNP microarray è un array a doppio colore contenente 60-mer di oligonucleotidi e sonde per SNP sintetizzati in situ utilizzando la tecnologia a getto d'inchiostro SurePrint di Agilent. Questi permettono l'identificazione simultaneamente di CNVs, LOH (Loss Of Heterozygosity) e disomie uniparentali (UPD) nello stesso array.

Step 1. Digestione del DNA

E' necessaria una concentrazione di DNA di 50ng/μl, dopodiché ad ogni campione sono aggiunti 2,9 μl di Digestion Master Mix e incubati per 2 ore a 37°C, per permettere agli enzimi di tagliare il DNA in frammenti di 400bp, e per 10 minuti a 65°C, per inattivare gli enzimi stessi (Tabella 5).

Mix di reazione

Nucl Free Water	1 μl
10x buffer C	1,3 μl
Acetylated BS	0,1 μl
AIU I	0,25 μl
RSA I	0,25 μl

Tabella 5. Componenti della Digestion Master Mix

Step 2. Marcatura del DNA

Ai campioni e ai controlli digeriti sono stati aggiunti 5 μl di Random Primers e incubati a 95°C per 5 minuti per permettere l'apertura della doppia elica. Dopodiché è stata aggiunta la Labeling Master Mix con cianine differenti per il

campione (Cy5) e per il controllo (Cy3). I campioni e i controlli sono stati incubati per 2 ore a 37°C e a 65°C per 10 minuti (Tabella 6). La marcatura deve avvenire rigorosamente in assenza di luce diretta sulle provette in modo da non interferire con il funzionamento delle cianine, che sono sensibili alla luce.

Mix di reazione

5X buffer	5 µl
10X dNTP	2,5 µl
Cy3 (controlli)	1,5 µl
Cy5 (campioni)	1,5 µl
Exo-Klenow	0,5 µl

Tabella 6. Componenti Labeling Master Mix

Step 3. Purificazione del DNA marcato

Si tratta di purificare tutti i campioni per eliminare l'eccesso di marcatura e i frammenti non marcati. Questa procedura va fatta al buio. Viene utilizzato il SureTag labeling kit Purification column (Agilent Technologies). Il contenuto marcato viene riposto nella colonnina di purificazione con il buffer TE 1X e centrifugato. Per controllare l'attività specifica della cianina, 1 µl di DNA marcato e purificato viene dosato al Nanodrop.

Step 4. Ibridazione

Per calcolare l'attività specifica di ogni campione, si utilizza questa formula: pmol/µl Cy3 o Cy5/ug/µl DNA. Questo valore deve essere compreso tra 25 e 40 per Cy3 e tra 20 e 35 per Cy5. Si passa poi a calcolare la resa del campione ($[DNA] \times 8$), in questo modo verranno accoppiati quei campioni che presenteranno una resa ed una attività specifica simile tra loro. Si prepara la Hybridization Master Mix e se ne aggiungono 29 µl (Tabella 7). Per permettere al COT-DNA di ibridarsi con le regioni altamente ripetute, il campione viene incubato per 5 minuti a 95°C e per 30 minuti a 37°C (preannealing). Si procede poi con l'ibridazione sul vetrino

aggiungendo 42 μ l di campione. Tale vetrino viene poi incubato in rotazione (20rpm) in un fornello a 65°C per 24 ore.

COT2	2 μ l
10X Block. Sol.	4,5 μ l
2XHIRPM	22,5 μ l

Tabella 7. Componenti per la Hybridization Master Mix.

Questo passaggio consente di eliminare l'eccesso di DNA non ibridato. A questo scopo si utilizzano due diverse soluzioni di lavaggio: Oligo array-CGH Wash Buffer 1, a temperatura ambiente per 5 minuti, e Oligo array CGH Wash buffer 2, scaldata a 37°C per 1 minuto nelle quali verranno introdotti i vetrini contenenti i DNA ibridati.

Step 6 e 7. Scansione degli array e analisi dei risultati

La scansione degli array è stata eseguita con uno scanner Agilent (SureScan DX Microarray Scanner) che è in grado di rilevare i livelli di fluorescenza (Cy5 e Cy3) di ogni spot e fornisce l'immagine su cui poi si effettuerà l'elaborazione dei dati grazie all'utilizzo del software "feature extraction" (Agilent). Contemporaneamente è possibile individuare l'alterazione nel numero di copie e le LOH in base al calcolo dei rapporti delle intensità di fluorescenza emesse da Cy5 e Cy3. L'analisi è stata effettuata con il software Genomic Workbench (Agilent Technologies).

Analisi del cariotipo

I campioni arrivati in eparina sono stati allestiti secondo la metodica di cariotipo costituzionale ad alta risoluzione e colorati con metodica QFQ (Q, Fluorescenza, Quinacrine).

Sono stati aggiunti 0.5 ml di sangue intero in eparina a 5 ml di terreno per colture cromosomiche standard (terreno RPMI 1640, siero fetale bovino FCS, antibiotici, fitoemoagglutinina). Dopo 72h di coltura in termostato a 37°C sono stati aggiunti

0.5 ml di bromodeossiuridina (BrdU) per permettere l'allungamento cromosomico con la sincronizzazione del ciclo cellulare in prometafase.

Dopo 17 ore le cellule sono state lavate due volte, dapprima con 5 ml di terreno RPMI 1640 addizionato con 1,5 ml di timidina e, successivamente, con 5 ml di terreno per colture cromosomiche non stimolate (RPMI 1640, FCS, antibiotici) contenente la stessa quantità di timidina per far ripartire il ciclo cellulare per ulteriori 5 ore. Successivamente, sono state aggiunte 20 µl di colchicina tenute in termostato a 37°C per 1 ora. La colchicina è una sostanza di origine vegetale che ha la funzione di bloccare le fibre del fuso mitotico, mantenendo i cromosomi in prometafase. Successivamente, le colture vengono sottoposte ad uno shock ipotonico con KCl alla concentrazione di 0.06% e poi fissate due volte in acido acetico e metanolo in proporzioni 1:3. Le cellule, in ultimo, sono state risospese in 5ml del fissativo (acido acetico e metanolo in proporzioni 1:3) utilizzato precedentemente e poi mantenute a 4°C per una notte.

Il giorno successivo, eliminato il surnatante, il pellet è stato risospeso in poche gocce di fissativo (acido acetico e metanolo in proporzioni 1:3) e sono stati allestiti i vetrini. I vetrini così allestiti possono essere utilizzati per cariotipo costituzionale dopo bandeggio e colorazione specifiche (Figura 23).

Colorazione dei vetrini per bandeggio ad alta risoluzione

E' stata eseguita la colorazione di bandeggio QFL (Q, Fluorescenza, Quinacrine) con la seguente metodica.

E' stata preparata la soluzione madre di Quinaccrina diluendo 100mg di Quinacrina mostarda in 100ml di acqua bidistillata. Successivamente è stato preparato il tampone Mc Ilvaine costituito da 21,01gr di acido citrico monoidrato, portato a 1000ml di acqua distillata, 25,6gr Na₂HPO₄ portato a 1000 ml con acqua distillata.

Successivamente è stata preparata la soluzione di lavoro prelevando 177 ml della soluzione 1 e aggiungendo 823 ml della soluzione 2 fino ad un pH 7. A questo punto i vetrini sono stati immersi per almeno 20 minuti a temperatura ambiente al buio nella soluzione di lavoro, costituita dall'unione di 5 ml di soluzione madre precedentemente preparata con 95 ml di tampone Mc Ilvaine e sciacquati in

tampone Mc Jivaine e coperti con un vetrino coprioggetto. I preparati si osservano al microscopio a fluorescenza.

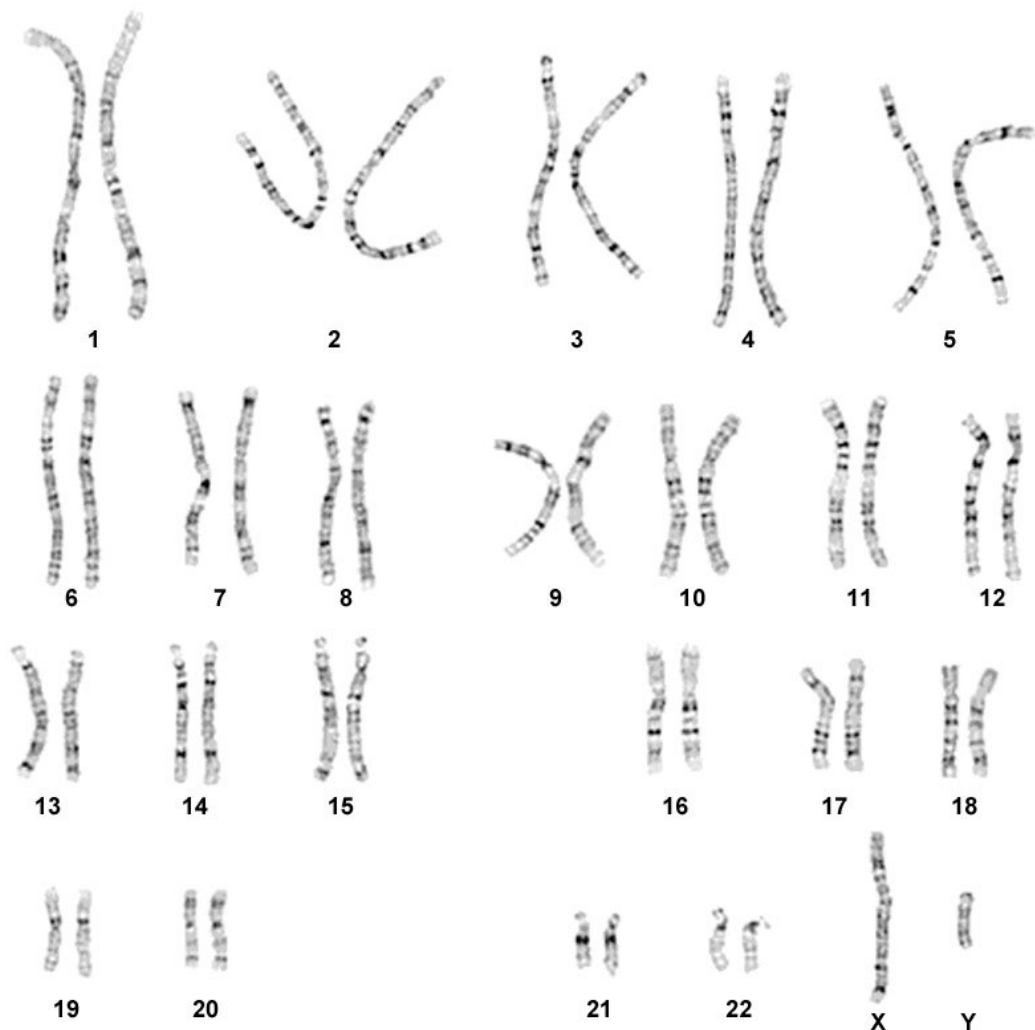


Figura 23. Esempio di cariotipo costituzionale maschile.

FISH

I vetrini precedentemente allestiti sono stabilizzati a 72°C in una soluzione 2XSSC per 5 minuti. Poi sono stati disidratati nella scala ascendente degli alcoli (70% 80% 95%) a temperatura ambiente per 2 minuti ciascuno. I vetrini sono stati immersi per 5 minuti a 72°C in una soluzione denaturante (costituita da 49 ml di formammide deionizzata, 14 ml di acqua distillata e 7ml 20XSSC portata a pH 7.5 con acido cloridrico). I preparati sono stati di nuovo immersi per 2 minuti ciascuno nella scala ascendente degli alcoli questa volta precedentemente refrigerati a -20°C in modo tale da mantenere le eliche del DNA denaturate. Tutte le sonde FISH necessitano di una precedente denaturazione a caldo 72°C per 5 minuti. Al termine

della denaturazione sono stati dispensati 10µl di sonda su ogni vetrino. I preparati sono stati posti in camera umida a 37°C per una notte per permettere l'ibridazione. Il giorno successivo si è proceduto con i lavaggi post ibridazione:

Lavaggio 1 per 5 minuti in 0,2XSSC a 72°C

Lavaggio 2 per 5 minuti in 2XSSC con l'aggiunta di TWEEN 20 a temperatura ambiente. In seguito ai lavaggi, sono stati aggiunti 10 µl di DAPI (4'-6' diamidino 2' phenylindole). I vetrini vengono poi osservati al microscopio ottico a fluorescenza con l'utilizzo dei filtri appropriati per il DAPI, per lo Spectrum Green (FITC) e per lo Spectrum Orange (rodamina). La fluoresceina e la rodamina sono eccitate dalla luce fluorescente, con emissione di luce verde e rossa, rispettivamente. Il DAPI che si colora in blu alla luce UV, permette un buon bandeggio dei cromosomi (Figura 24).

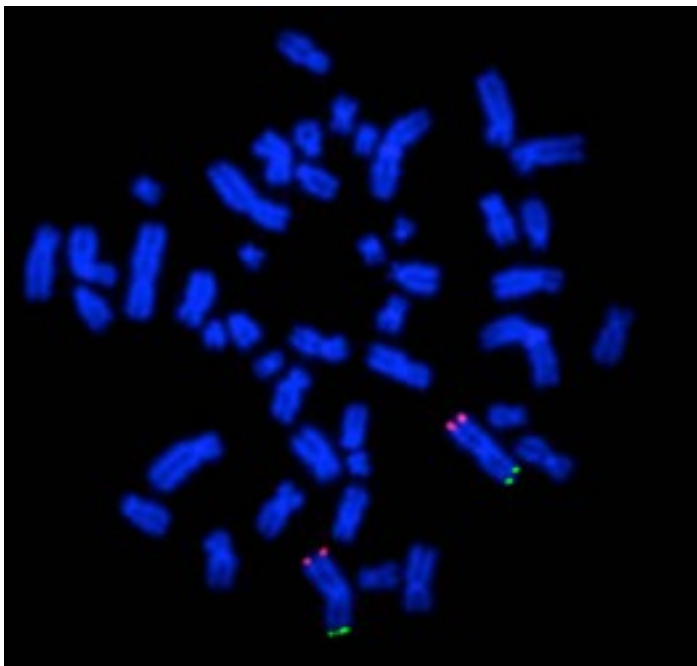


Figura 24. Esempio di ibridazione fluorescente in situ (FISH).

Una metafase con due diverse sonde, una rossa e una verde che si localizzano rispettivamente sulle braccia corte e sulle braccia lunghe dei cromosomi 3.

Validazione delle CNVs tramite Real Time PCR

Per la validazione dei risultati ottenuti con la metodica array-CGH, è stato utilizzato il protocollo ottimizzato per lo strumento Light-Cycler 480 della Roche.

La procedura prevede l'uso di un kit contenente SYBR Green I Master, che include il SYBR Green dye, che serve per l'amplificazione e la rilevazione del DNA ed è un marcatore specifico del DNA a doppio filamento. Durante ogni fase

della sintesi del DNA, il colorante SYBR Green I, incluso nella miscela di reazione, si lega ai prodotti PCR amplificati. L'amplicone può essere successivamente rilevato attraverso l'analisi della sua fluorescenza.

L'amplificazione avviene secondo i seguenti cicli:

95 °C per 5'	} per 30 volte
95 °C per 10''	
60 °C per 20''	
72 °C per 30''	
40 °C per 10'	

Il DNA è amplificato da reazioni a catena della DNA polimerasi; dopo ogni turno di amplificazione, il DNA viene quantificato. I metodi comuni di quantificazione includono l'utilizzo delle colorazioni fluorescenti che si intercalano al DNA. Per valutare in tempo reale la quantità totale di DNA a doppio filamento presente dopo ogni ciclo di sintesi, nella miscela di reazione è presente SYBR Green, un composto fluorescente intercalante del DNA. La fluorescenza del SYBR Green aumenta considerevolmente nel momento in cui la molecola intercala nel solco minore del DNA a doppio filamento, pertanto, la misura quantitativa della fluorescenza emessa è direttamente proporzionale alla quantità di DNA a doppio filamento presente in ogni campione (che è a sua volta proporzionale alla quantità di DNA presente nel campione iniziale), a patto che la molecola intercalante sia presente in eccesso. La lettura della fluorescenza avviene al termine di ogni ciclo di amplificazione. Dopo ogni rilevamento, i segnali di fluorescenza sono processati da un software e la cinetica di formazione dell'amplicato è visualizzata graficamente come incremento dei segnali di fluorescenza in ordinata per i cicli direzione in ascissa; viene poi effettuata una quantificazione relativa degli amplificati, che richiede la quantificazione di geni di controllo *housekeeping* per normalizzare l'espressione del gene studiato.

Sequenziamento del Genoma

Il sequenziamento del genoma dei tre bambini è stato effettuato in collaborazione con il Centro di Genomica Funzionale dell'Università di Verona diretto dal prof. Delledonne.

Le librerie sono state preparate utilizzando il protocollo TruSeq DNA PCR-free (350bp) Illumina. Questo protocollo permette di preparare di generare librerie per sequenziamento sia *single reads* che *paired end*.

Il DNA genomico è stato quantificato con metodo fluorimetrico e sono necessari 1mg di DNA genomico (gDNA) di partenza per generare librerie di 350bp di lunghezza. Quantità inferiori a queste aumentano la possibilità di generare duplicati e di avere un sequenziamento con un rendimento basso. I valori di assorbanza 260/280 devono essere compresi tra 1,8-2,0 per considerare il DNA relativamente puro

I campioni di DNA sono stati normalizzati con RSB fino ad ottenere un volume finale di 55µl ad una concentrazione di 20 ng/ml per ottenere una dimensione dell'inserto di 350bp. Successivamente il DNA viene frammentato con in maniera meccanica con il Covaris che genera frammenti di DNA a singolo filamento (dsDNA). A questo punto si prosegue con la reazione di End Repair per la quale è stata utilizzata una End Repair Mix nella quale si utilizza una polimerasi con attività esonucleasica 5'-3', seguito poi dalla selezione dei frammenti di giuste dimensioni tramite purificazione mediante biglie magnetiche. All'estremità 3' sono state aggiunte, tramite reazione di ligazione, delle Adenine in modo da impedire ai frammenti di legarsi l'uno con l'altro durante il legame con gli adattatori. La timina sull'estremità 3' dell'adattatore fa sì che si possa verificare un giusto legame tra frammento di DNA e adattatore. Questa strategia garantisce un basso tasso di formazione di chimera. Successivamente sono stati ligati gli adattatori per permettere poi l'ibridazione sulla *flow cell*. La fase conclusiva della preparazione delle librerie è la normalizzazione dei campioni a 2Nm per poter poi avere un volume uguale per tutti i campioni che saranno sequenziati (Figura 25).

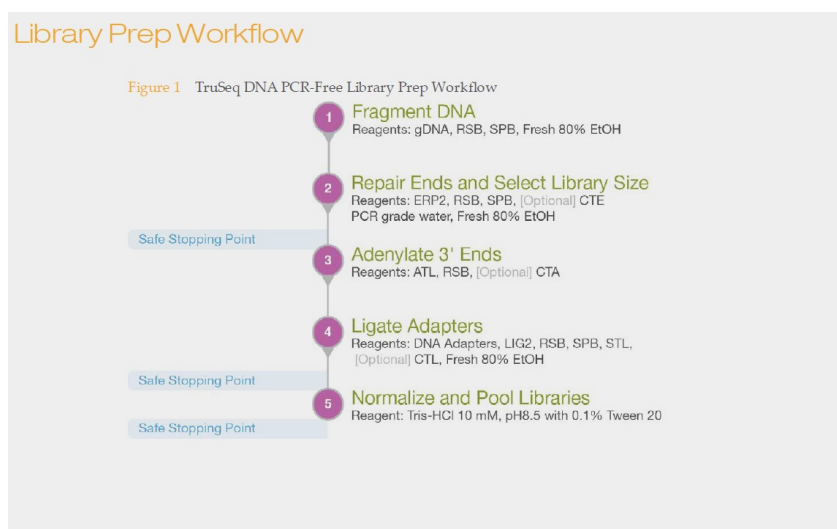


Figura 25. Workflow per la preparazione delle librerie Illumina.

Il sequenziamento delle librerie prodotte è stato effettuato su Hiseq X Ten con corsa 150 PE.

L'analisi dei dati è stata effettuata facendo un primo step di controllo di qualità delle *reads* (FastQ format) in modo tale da scartare le *reads* di bassa qualità. Le *reads* che al contrario passano il controllo di qualità sono state allineate con il genoma umano usando BWA. I file ottenuti in formato SAM vengono, quindi, convertiti in un formato binario di allineamento (BAM).

Prima di procedere a successive analisi, quindi, vengono effettuati dei processi di controllo e miglioramento della qualità delle letture.

I dati statistici relativi al *coverage* sono stati ottenuti con script sviluppati *in house*. Partendo dalle informazioni ottenute con l'analisi del *coverage*, abbiamo identificato i cluster di *reads* che mappavano in modo anomalo sul genoma di riferimento (UCSC build hg19): di queste *reads*, siamo andati a ricercare il *mate* corrispondente, per individuare la sua posizione sul cromosoma, e riuscire di conseguenza a ricostruire il riarrangiamento avvenuto (Figura 26).

Le *reads* prodotte sono state mappate sul genoma di riferimento e sono state analizzate con il software IGV (Integrative Genome Viewer, 2013 Broad Institute). Inoltre contemporaneamente è stata effettuata anche il *base calling* per l'identificazione di varianti nucleotidiche.

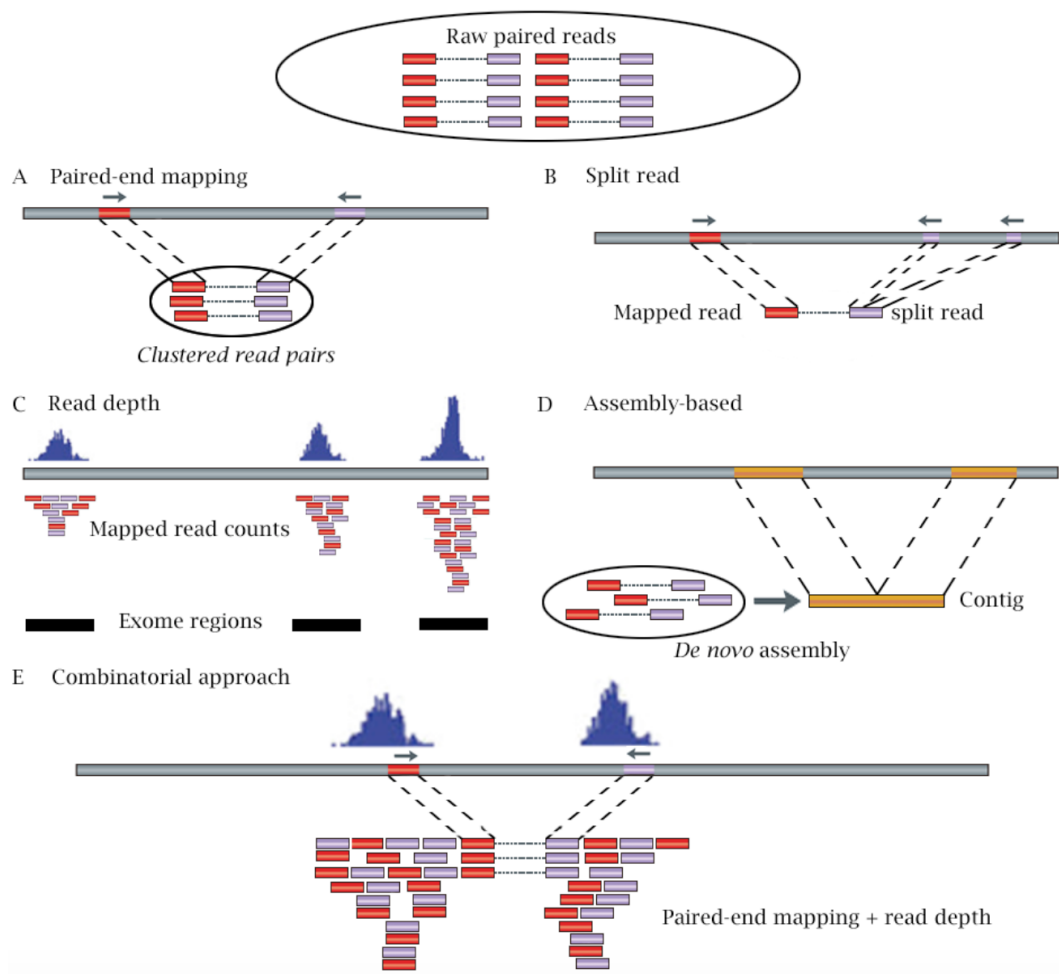


Figura 26. Esempio di ricerca di CNVs con NGS, tramite 5 diverse strategie.

- A. paired-end mapping
- B. ricerca *split reads*
- C. analisi *read depth*
- D. assemblaggio de novo del genoma
- E. una combinazione degli approcci precedenti

4. RISULTATI

4.1 CASO C/996

L'analisi del DNA, eseguita con metodica array-CGH (Figura 27 e 28), alla risoluzione di 100kb, ha rilevato la presenza di:

- una delezione interstiziale di circa 321kb nella regione 22 del braccio lungo del cromosoma 17 [17q22(51,360,166-51,626,927)x1]
- una delezione interstiziale di circa 18kb nella regione 11.2 del braccio lungo del cromosoma 17 [17q11.2(29,621,452-29,639,381)x1]
- una duplicazione interstiziale di circa 10Mb nella regione 12-11.2 del braccio corto del cromosoma 17 [17p12-p11.2(11,430,989-21,502,083)x3]

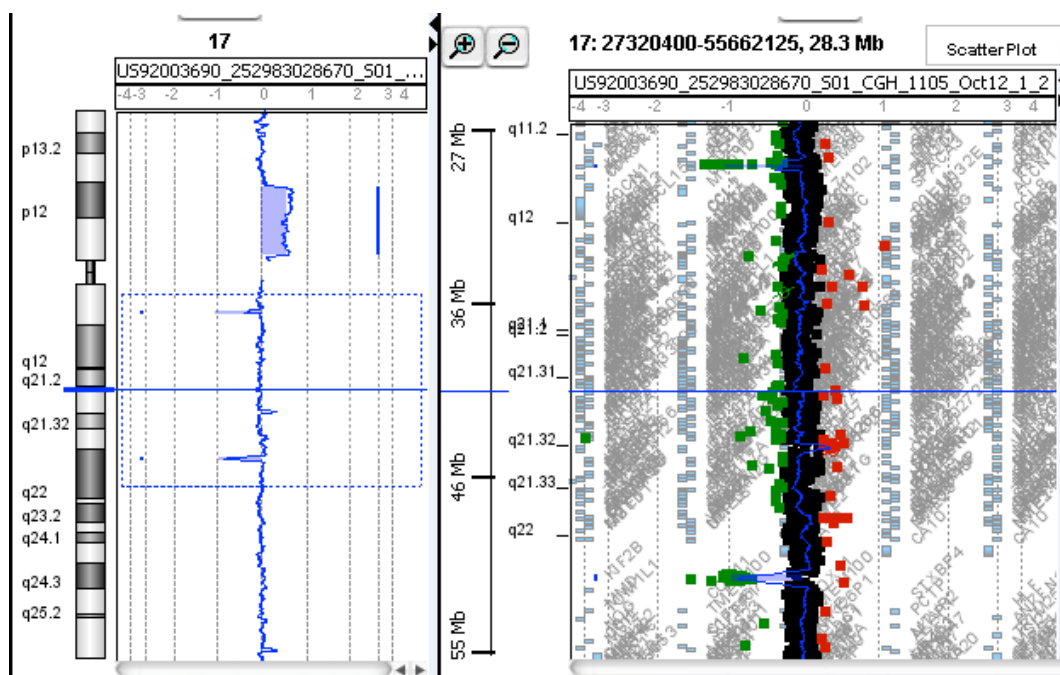


Figura 27. Ideogramma del cromosoma 17 del caso C/996.

L'analisi array evidenzia un'alterazione del numero di copie sul cromosoma 17, con due delezioni sul braccio lungo del cromosoma 17 e una duplicazione sul braccio corto.

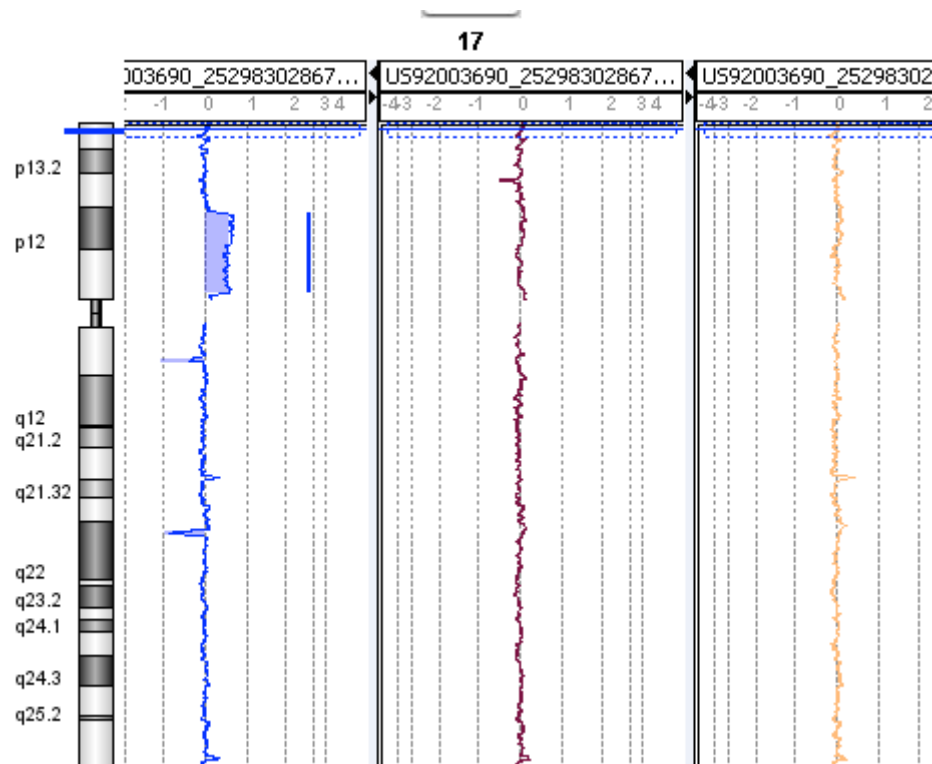


Figura 28. Ideogramma del cromosoma 17 del caso C/996 e dei suoi genitori. Analizzando le immagini ottenute con gli array-CGH non si rilevano alterazioni nei cromosomi parentali.

L'esame degli SNP array ha evidenziato che tutti i riarrangiamenti sono di origine materna (Figura 29).

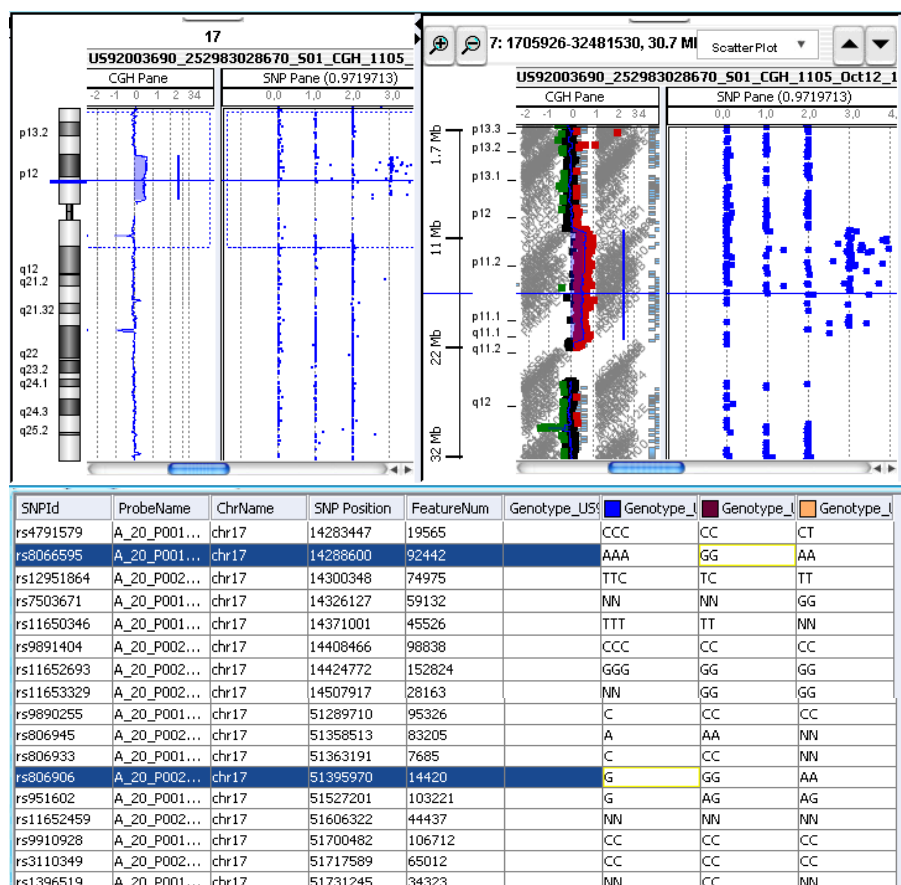


Figura 29. SNP array del cromosoma 17 del caso C/996 e dei suoi genitori.

L'esame del cariotipo non evidenziava nessuna delle anomalie e non si rilevava alcun riarrangiamento della regione del 17p duplicata. L'esame Real Time PCR confermava la presenza delle due delezioni sul braccio lungo del cromosoma 17. L'analisi FISH, effettuata con la sonda al locus LSI SMS (17p11.2) dimostrava che la regione duplicata si localizzava sul 17p (Figura 30).

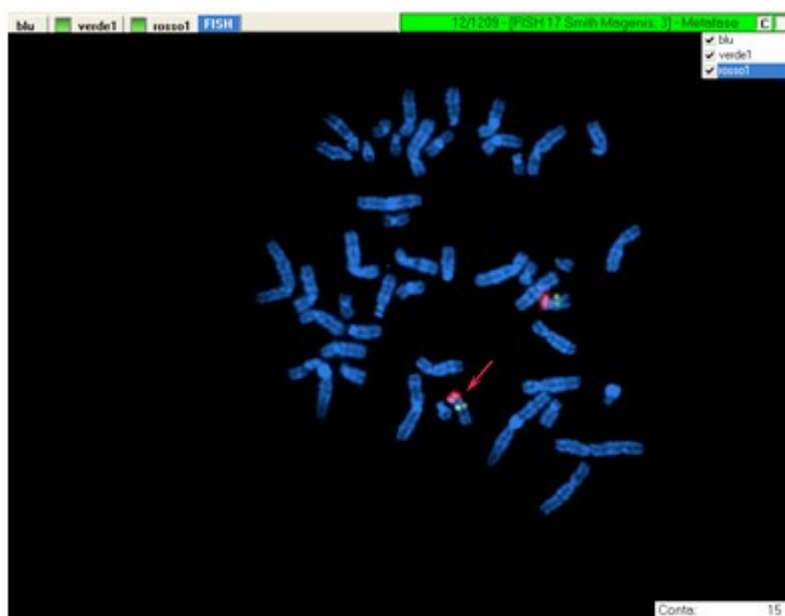


Figura 30. Analisi FISH cromosoma 17.

L'analisi FISH delle metafasi linfocitarie evidenzia segnali di ibridazione della regione 17p11.2. La freccia indica un segnale amplificato in uno dei due cromosomi 17.

Per studiare a fondo i riarrangiamenti cromosomici del paziente, è stato eseguito il sequenziamento del genoma (WGS), con le tecniche di Next Generation Sequencing: l'analisi del *coverage* ha confermato la presenza delle due delezioni e della duplicazione sul cromosoma 17, evidenziate all'array-CGH (Figura 31).

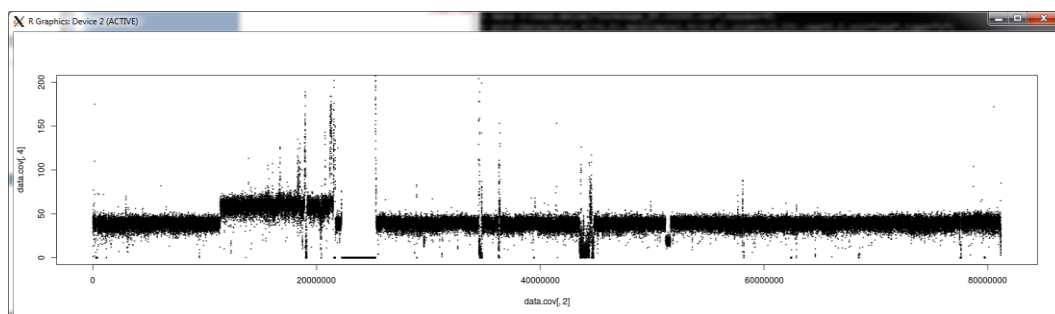


Figura 31. Analisi del coverage del cromosoma 17.

Duplicazioni e delezioni si presentano come alterazioni del *coverage*

L'analisi dei dati del sequenziamento è proseguita con lo studio delle regioni contenenti *reads* che mappavano in modo anomalo sul genoma di riferimento, indicando la presenza di presunte varianti strutturali (Figura 32).

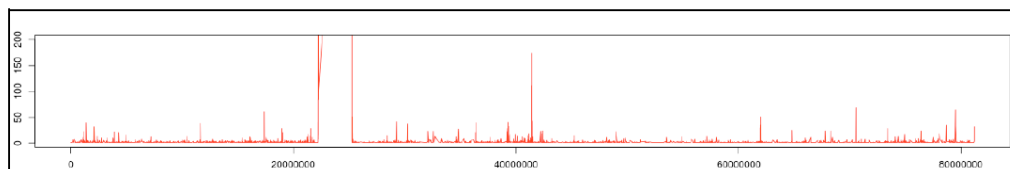


Figura 32. Analisi del coverage del cromosoma 17.

I *peaks* indicano cluster di *reads* che mappano in modo anomalo in una data regione (riportata sulle ascisse).

Delezione 17q11.2

Con le tecniche di sequenziamento NGS è stato possibile verificare con precisione il punto di rottura prossimale (29,590,099) e il punto di rottura distale (29,639,813) sul braccio lungo del cromosoma 17 (Figura 33).

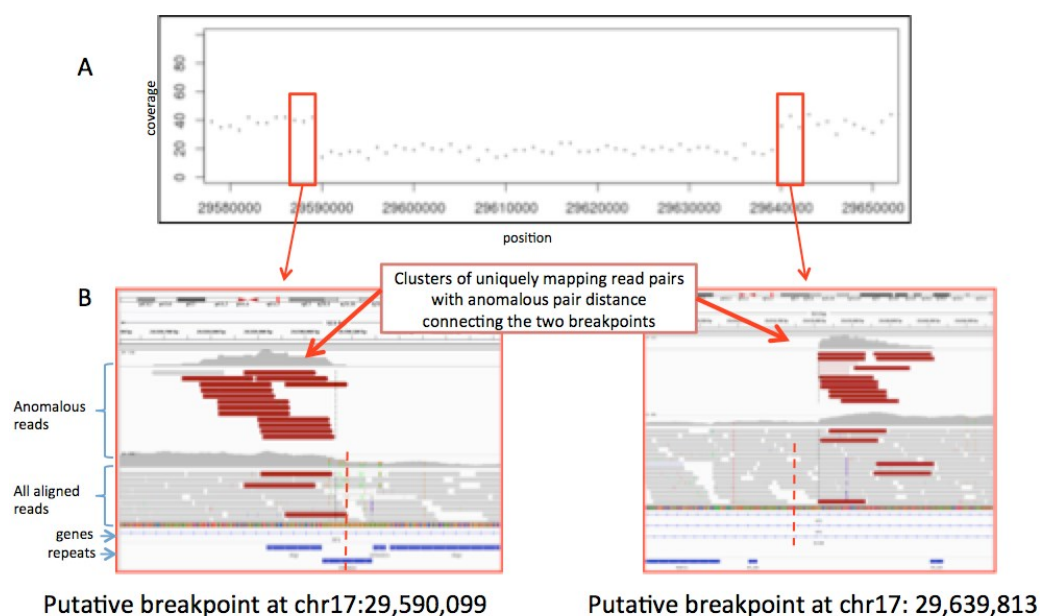


Figura 33. Analisi del *coverage* nella regione della prima delezione (A) e grafico delle *reads* che mappano in modo anomalo al bp prossimale e al bp distale della delezione (B).

Le *reads* che mappavano in modo anomalo a livello del punto di rottura prossimale, hanno dimostrato di avere il loro *mate* dove inizia il punto di rottura distale (Figura 28B): questo dato ha permesso di confermare che si tratta di una delezione, con perdita della sequenza tra i due punti di rottura. La verifica, tramite il programma *in silico*, del punto di giunzione tra i *mate*, ha permesso di identificare un'inserzione di 12 paia di basi a livello della giunzione: non è stato

ancora possibile validarla, poiché la sequenza inserita risulta altamente ripetitiva, ma sono in corso ulteriori analisi con PCR per la verifica della esatta sequenza di questa giunzione (Figura 34).



Figura 34. Verifica *in silico* del punto di giunzione.

Il programma *in silico* dimostra l'inserzione alla giunzione (A) di una breve sequenza di DNA altamente ripetuta (B).

La regione deleta contiene l'introne 35 del gene *NFI*, il gene *OMG* e il gene *EVI2B*, e un gran numero di sequenze ripetute come Alu, SINE e LINE (Figure 35 e 36).

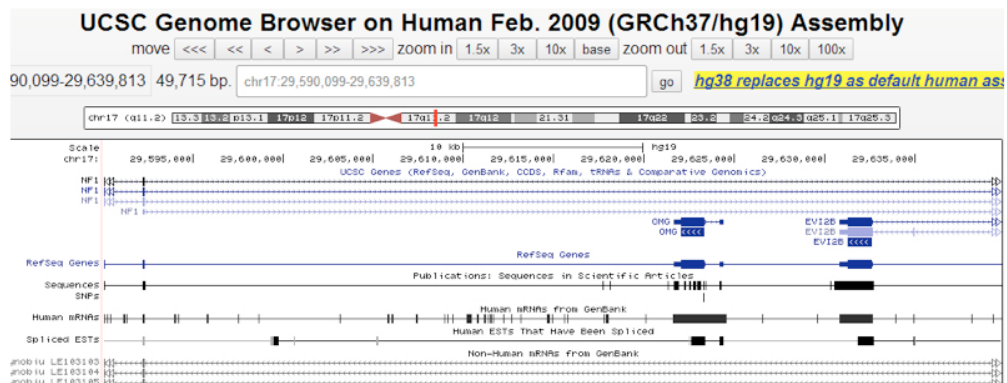


Figura 35. Schema del cromosoma 17

Dimensione e la sede della delezione interstiziale: 29,590,099 (*breakpoint* prossimale, ultima sonda presente prima della delezione) -29,639,813 (*breakpoint* distale, prima sonda presente dopo la regione deleta). Immagine ripresa da UCSC Genome Browser. Sono indicati i geni *NFI*, *OMG* e *EVI2*. Gli ultimi due sono inclusi all'interno dell'introne del gene *NFI*.

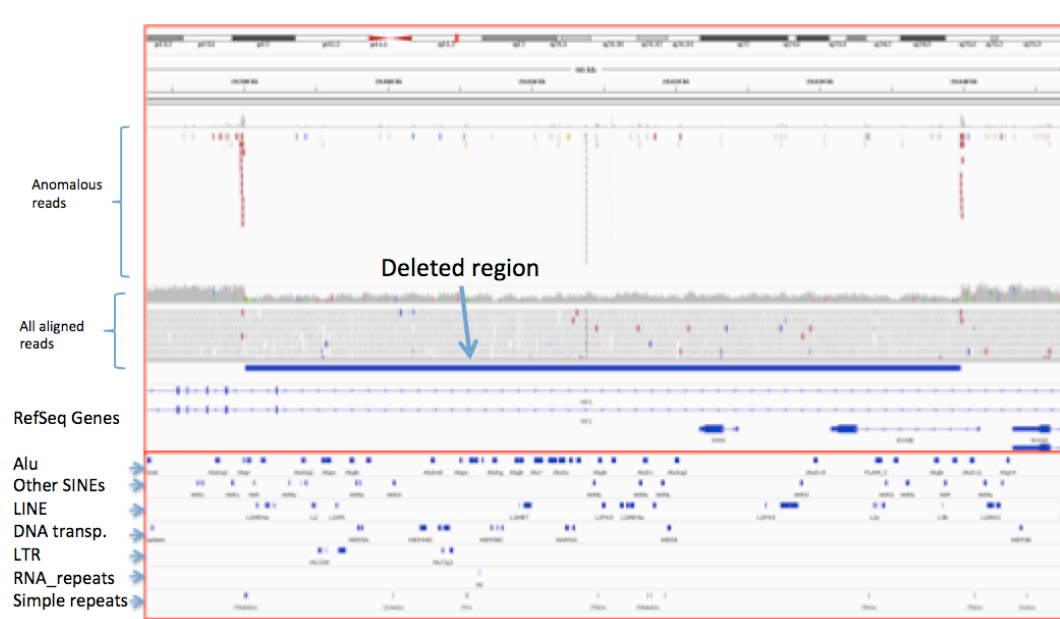


Figura 36. Geni e sequenze di DNA ripetitive localizzate nella regione deleta.

La regione contiene l'introne 35 del gene *NFI*, e i geni *OMG* e *EVI2B*, e numerose sequenze ripetute.

Delezione 17q22

Con l'analisi NGS sono stati identificati con precisione il punto di rottura prossimale e il punto di rottura distale della delezione, rispettivamente a 51,203,005 e a 51,637,727. Le *reads* che mappavano in modo anomalo a livello del punto di rottura prossimale, hanno dimostrato di non avere il loro *mate* a livello

del punto di rottura distale, ma in un'altra posizione nel cromosoma 17, ovvero a 49,887,817 (Figura 37).

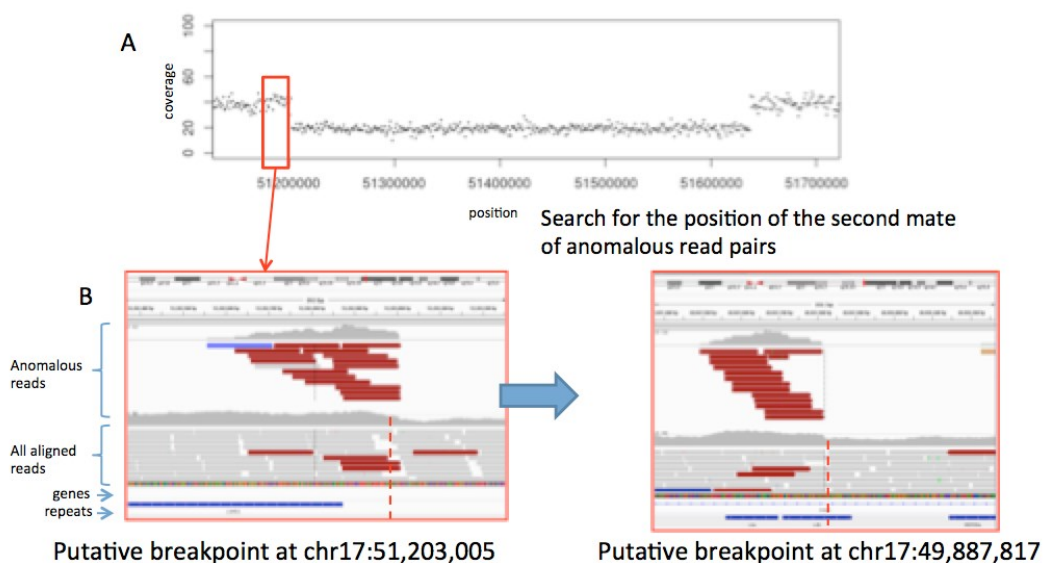


Figura 37. Analisi del *coverage* nella regione della seconda delezione (A) e grafico delle *reads* che mappano in modo anomalo al bp prossimale, che risultano avere il loro *mate* in una posizione diversa dal bp distale (B).

Utilizzando un programma *in silico*, è stato possibile validare la giunzione tra i due *mate* (Figura 38).

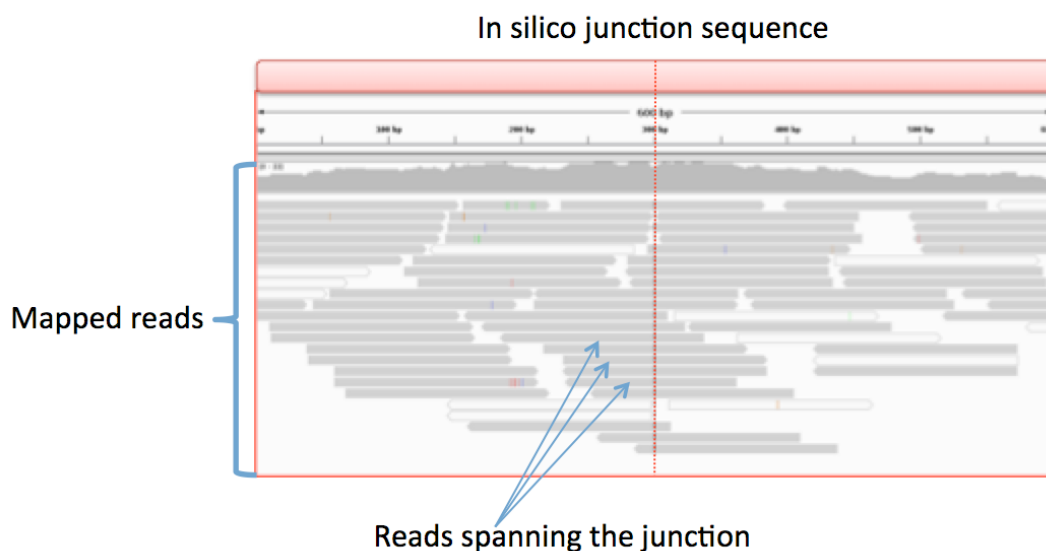


Figura 38. Verifica *in silico* del punto di giunzione.

Analogamente lo studio del punto di rottura distale, ha evidenziato che le *reads* che mappavano in modo anomalo in questa regione non avevano il loro *mate* a livello del punto di rottura prossimale, ma in posizione 49,881,360 sul cromosoma 17 (Figura 39).

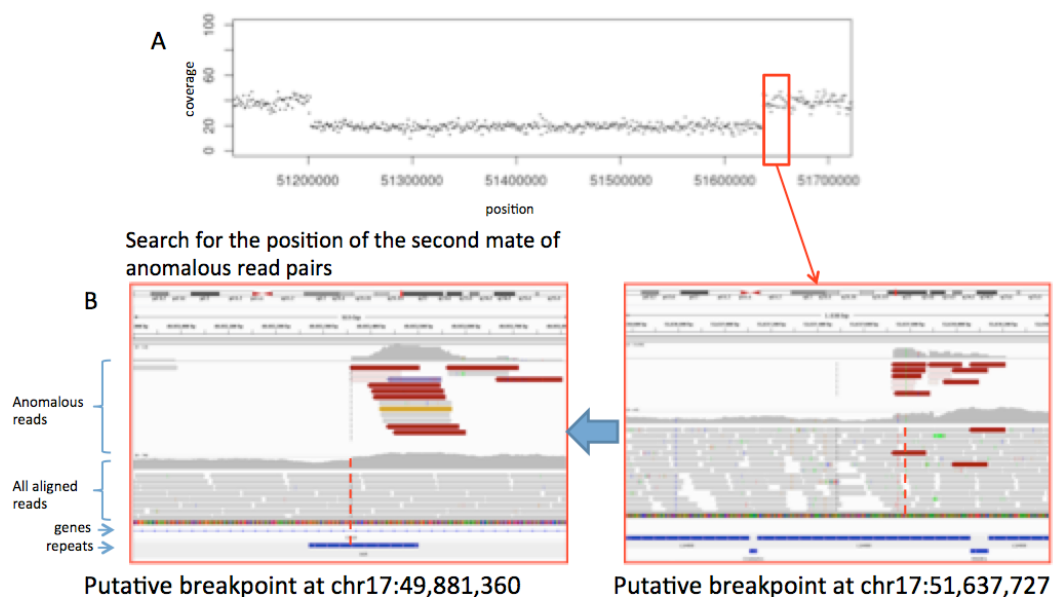


Figura 39. Analisi del *coverage* nella regione della seconda delezione (A) e grafico delle *reads* che mappano in modo anomalo al bp distale, che risultano avere il loro *mate* in una posizione diversa dal bp prossimale (B).

In quest'ultimo caso, il programma *in silico* non ha potuto validare la giunzione tra i due *mate*, a causa della presenza di un gap (Figura 40); per confermare tale presunta giunzione potrebbe essere utile eseguire un'analisi con PCR.



No read span the putative junction.

→ To confirm the putative junction a PCR assay should be performed.

Figura 40. Verifica *in silico* del punto di giunzione.

Sulla base dei risultati possiamo ipotizzare che all'interno della regione deleta (51,203,005-51,637,727), vi sia una inserzione della regione che va da 49,881,360 a 49,887,817. Tale regione si presenta quindi duplicata e invertita (Figura 41).

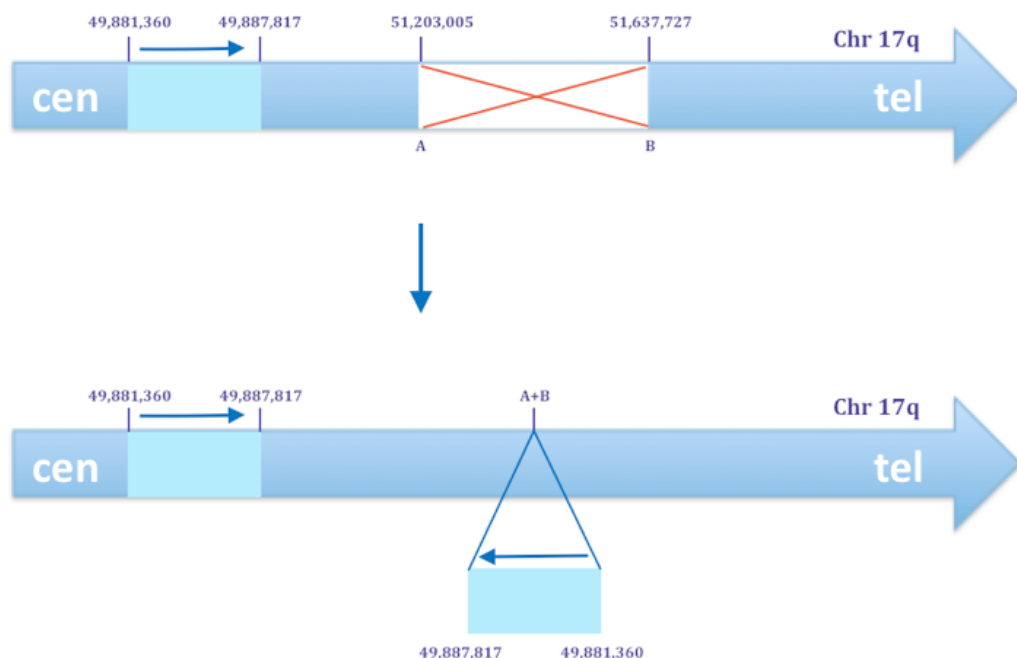


Figura 41. Grafico del riarrangiamento ipotizzato nella regione 17q22.

Il cromosoma 17 risulta deleta nella regione 51,203,005-51,637,727; la regione 49,881,360-49,887,817 (riquadro celeste chiaro), si presenta duplicata e invertita all'interno della regione deleta (A+B). La direzione delle sequenze di DNA sul cromosoma è indicata dall'orientamento delle frecce blu.

La conferma dell'avvenuta duplicazione della regione 49,881,360-49,887,817 è stata ottenuta dall'analisi del *coverage*, che quindi risulta aumentato, per la presenza di un numero maggiore di copie delle *reads* corrispondenti (Figura 42).

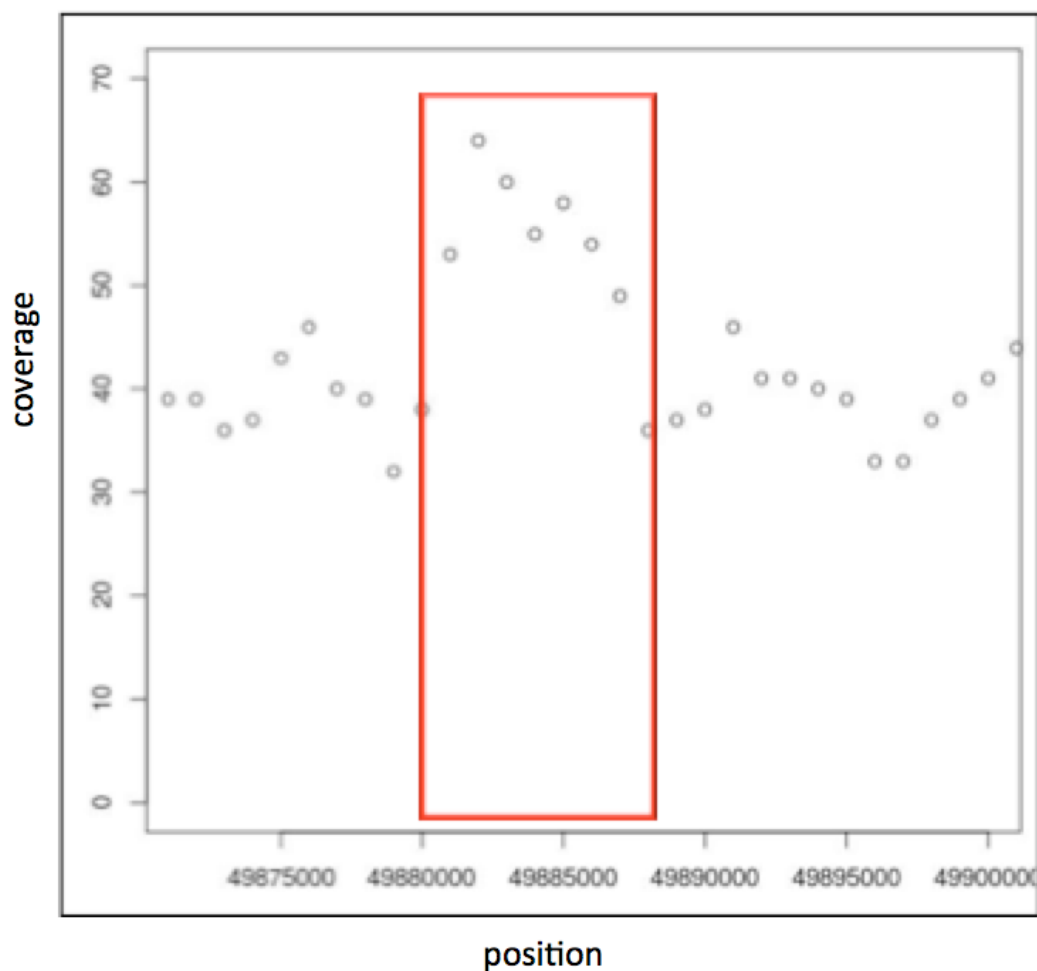


Figura 42. Analisi del *coverage* della regione 49,881,360-49,887,817, in accordo con la presunta duplicazione di questa regione.

La regione duplicata contiene al suo interno l'introne 4 del gene *CA10*, inserito in numerose sequenze di tipo SINE e LINE (Figure 43 e 44).

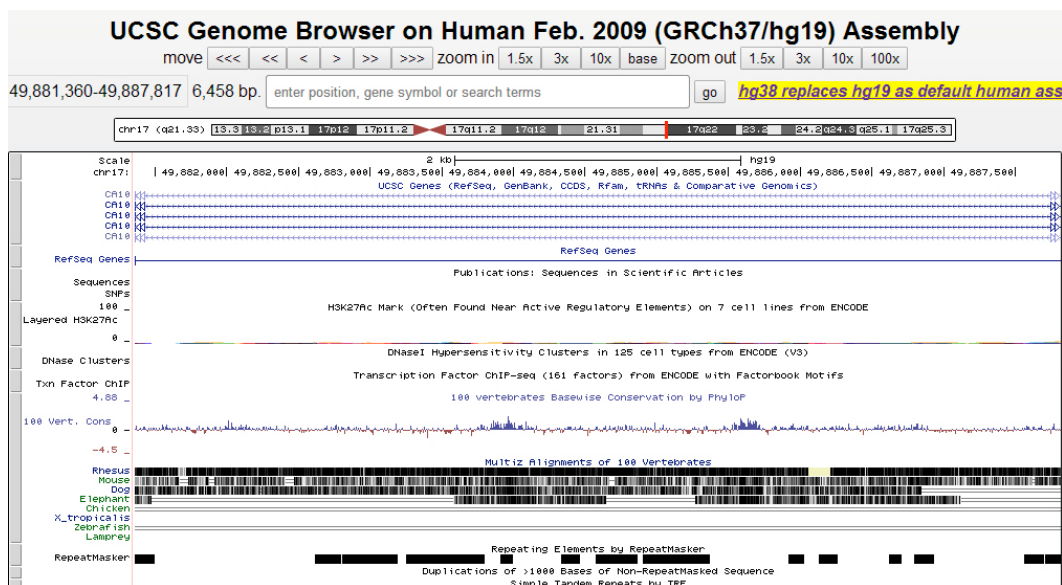


Figura 43. Schema del cromosoma 17

Dimensione e la sede della duplicazione interstiziale: 49,881,360 (*breakpoint* prossimale, ultima sonda presente prima della duplicazione) -49,887,817 (*breakpoint* distale, prima sonda presente dopo la regione duplicata). Immagine ripresa da UCSC Genome Browser. Sono indicati il gene *CA10* e i sottostanti *repeats*.

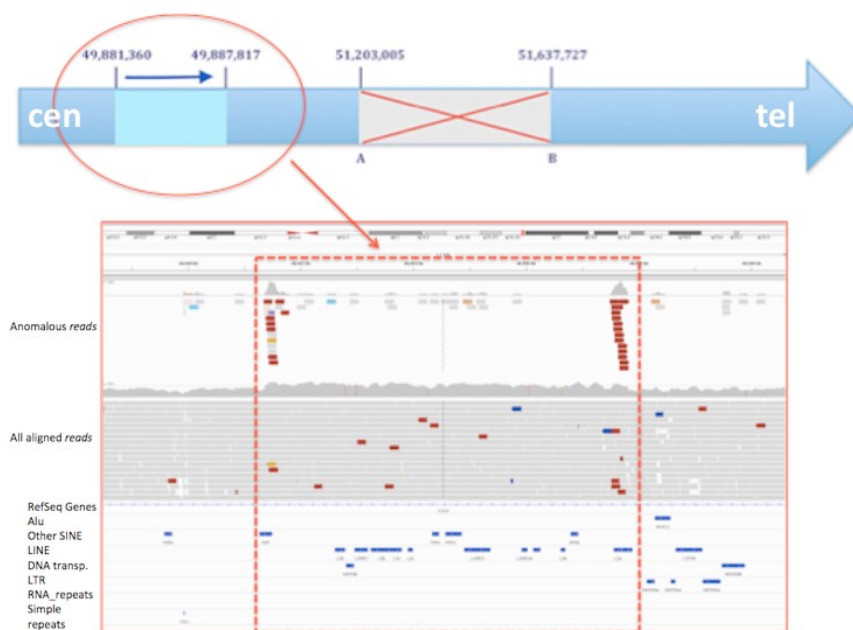


Figura 44. Geni e sequenze di DNA ripetitive localizzate nella regione duplicata.

La regione contiene l'introne 4 del gene *CA10*, inserito in numerose sequenze altamente ripetitive.

La regione deleta non contiene né sequenze esoniche né introniche, ma soltanto un grande numero di sequenze ripetute, SINE, LINE, Alu, trasposoni a DNA (Figura 45)



Figura 45. Geni e sequenze di DNA ripetitive localizzate nella regione deleta.

La regione deleta non contiene sequenze né esoniche né introniche, ma contiene numerose sequenze altamente ripetitive.

Duplicazione 17p12-p11.2

L'analisi NGS della regione duplicata, ha permesso l'identificazione precisa del punto di rottura prossimale, che si trova a 11,385,129 sul braccio corto del cromosoma 17, mentre rimane ancora da definire il *breakpoint* distale; le *reads* che mappavano in modo anomalo a livello di questo punto, hanno dimostrato di avere il loro *mate* a 14,514,149 sul braccio corto del cromosoma 17 (Figura 46).

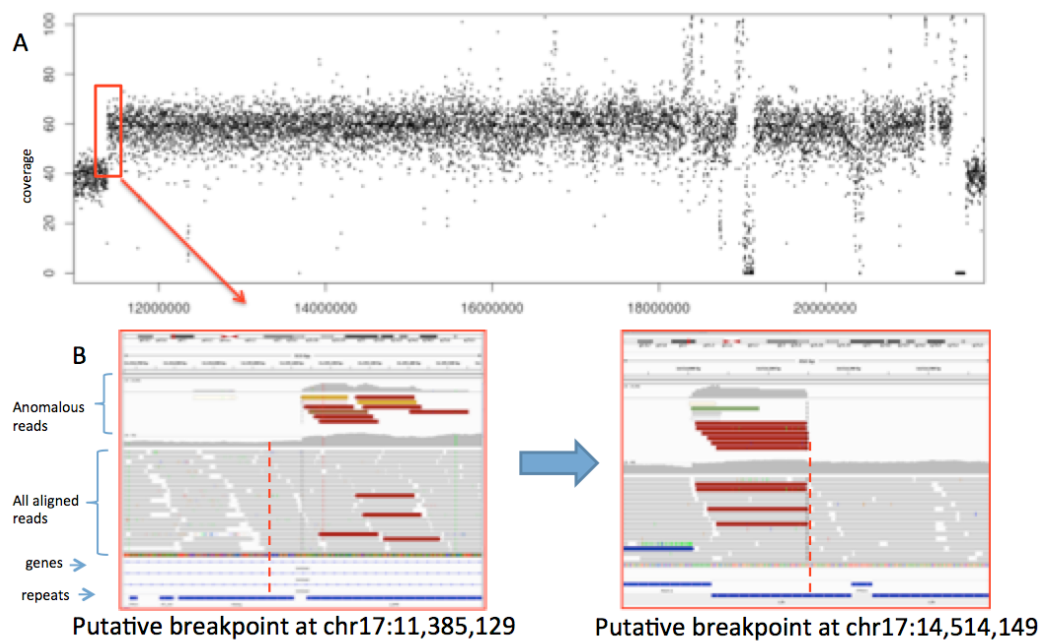


Figura 46. Analisi del *coverage* nella regione della duplicazione (A) e grafico delle *reads* che mappano in modo anomalo al bp prossimale, che risultano avere il loro *mate* in una posizione 14,514,149 (B).

Il punto di giunzione tra i due *mate* è stato validato con un programma *in silico* (Figura 47).

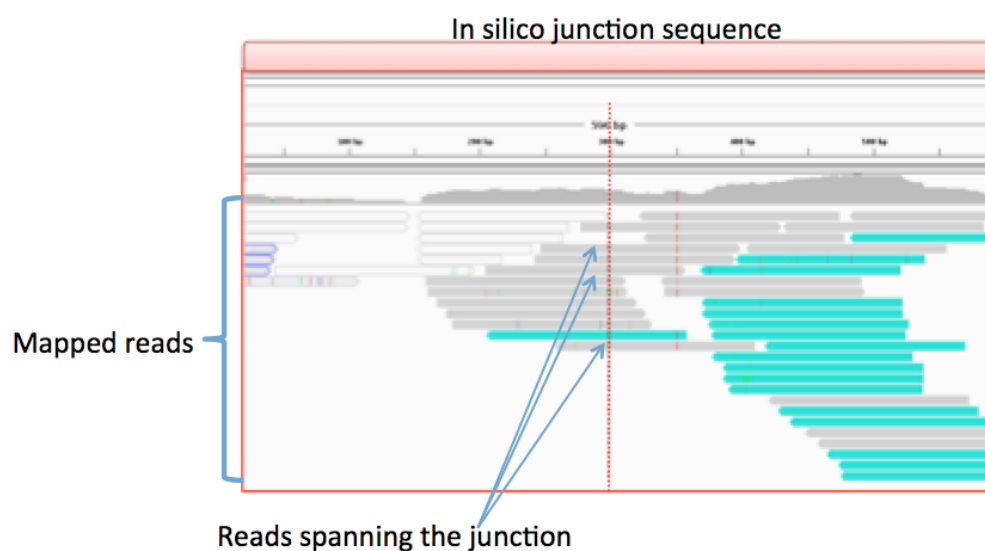


Figura 47. Verifica *in silico* del punto di giunzione.

Come già premesso, non è stato possibile individuare l'esatta posizione del punto di rottura distale (che agli array è risultato essere a 21,502,083): l'estremità distale della duplicazione è verosimile che si trovi all'interno di una regione ricca di

sequenze di DNA altamente ripetitive, non perfettamente definita nel genoma di riferimento usato nello studio (UCSC hg 19) (Figura 48);

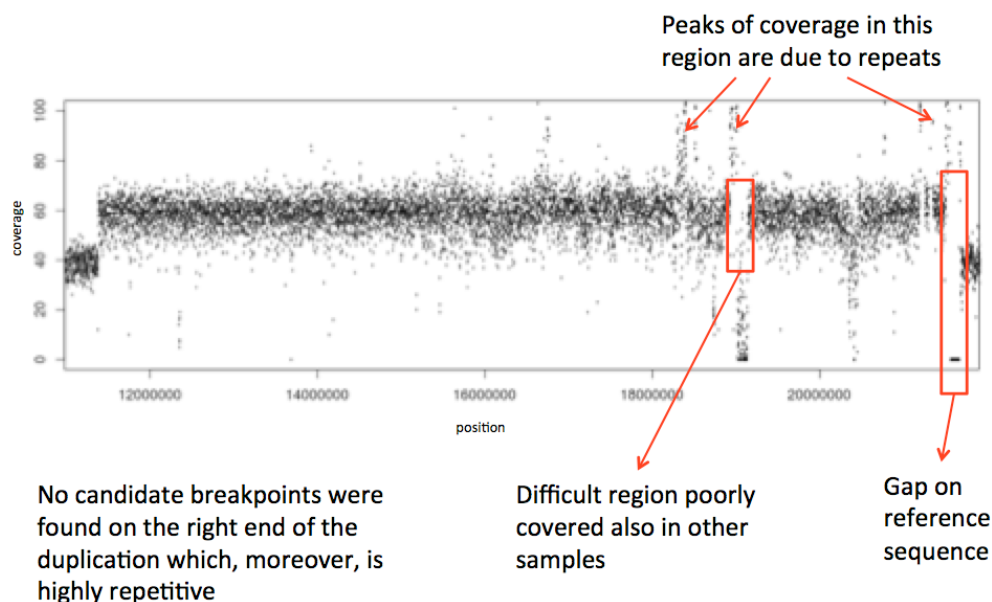


Figura 48. Analisi del *coverage* della regione duplicata sul 17p.

All'analisi del *coverage* questa regione risulta ricca di sequenze di DNA altamente ripetitive, con alcune sequenze non perfettamente definite nel genoma di riferimento UCSC hg19.

La regione duplicata contiene numerosi geni (Figura 49) tra cui *ELAC2*, *PMP22*, *COX10*, *PIGL*, *TTC19*, *RAI1*, *FLCN*, *TNFRSF13B*, *ATPAF2*, *B9D1*, *ALDH3A2*, *AKAP10*, *KCNJ18*, *MYO15A*

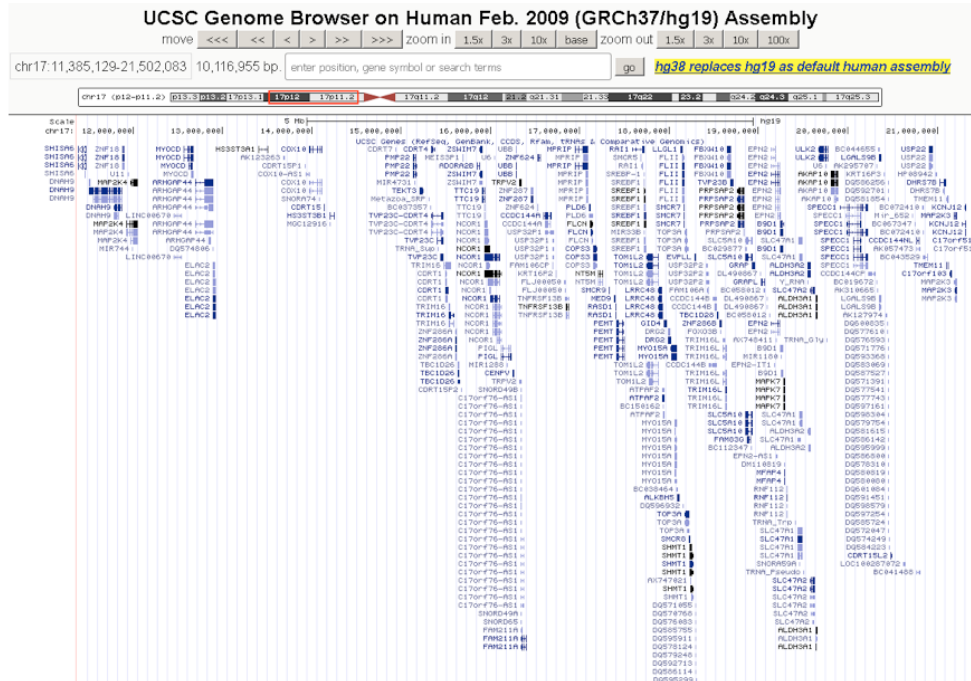


Figura 49. Schema del cromosoma 17

Dimensione e la sede della duplicazione interstiziale a livello del braccio corto: 11,385,129 (*breakpoint* prossimale, ultima sonda presente prima della duplicazione) -21,502,083 (*breakpoint* distale, prima sonda presente dopo la regione duplicata). Immagine ripresa da UCSC Genome Browser. Sono indicati i numerosi geni contenuti in questa regione, tra i quali prossimalmente il gene *PMP22*, la cui duplicazione è responsabile della neuropatia sensitivo-motoria CMT1A.

4.2 CASO E/1021

L'analisi del DNA del bambino, eseguita con metodica array-CGH (Figura 50 e 51) alla risoluzione di 100kb, ha rilevato la presenza di:

- una delezione interstiziale di circa 1Mb sul braccio lungo del cromosoma 12 [12q21.1(73,949,244-74,949,118)x1]
- una delezione interstiziale di circa 330kb sul braccio lungo del cromosoma 12 [12q21.33(89,625,405-89,955,801)x1]
- una delezione interstiziale di circa 255kb sul braccio lungo del cromosoma 12 [12q21.33(90,472,147-90,727,491)x1]
- una delezione interstiziale di circa 1,257Mb sul braccio lungo del cromosoma 12 [12q21.33(91,017,355-92,274,591)x1].

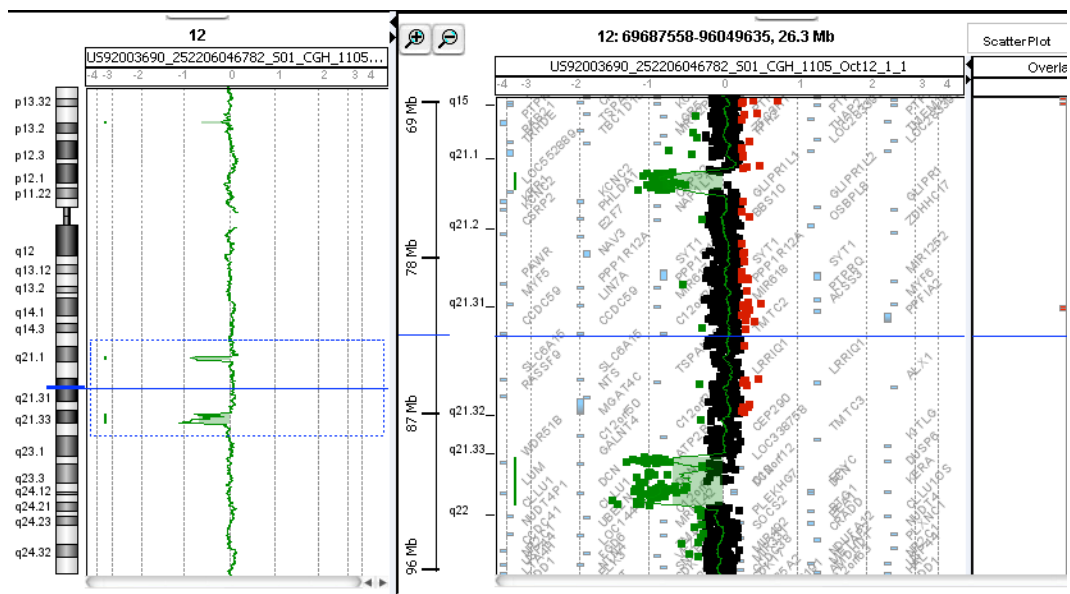


Figura 50. Ideogramma del cromosoma 12 del caso E/1021.

L'analisi array-CGH evidenzia un'alterazione del numero di copie sul cromosoma 12, con quattro delezioni sul braccio del cromosoma 12.

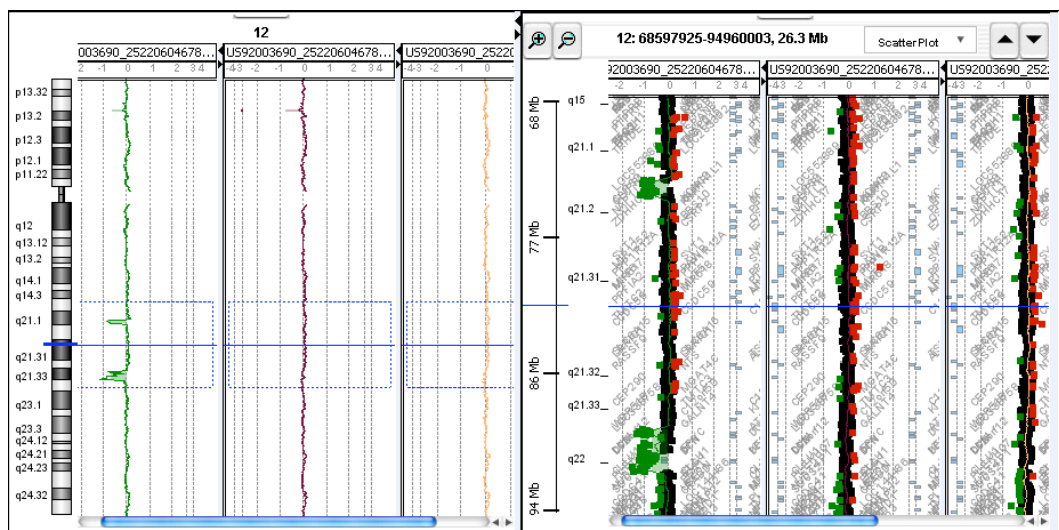


Figura 51. Ideogramma del cromosoma 12 del caso E/1021 e dei suoi genitori.

Analizzando le immagini ottenute con gli array non si rilevano alterazioni nei cromosomi parentali.

L'esame degli SNP array ha evidenziato che tutti i riarrangiamenti sono di origine paterna (Figura 52)

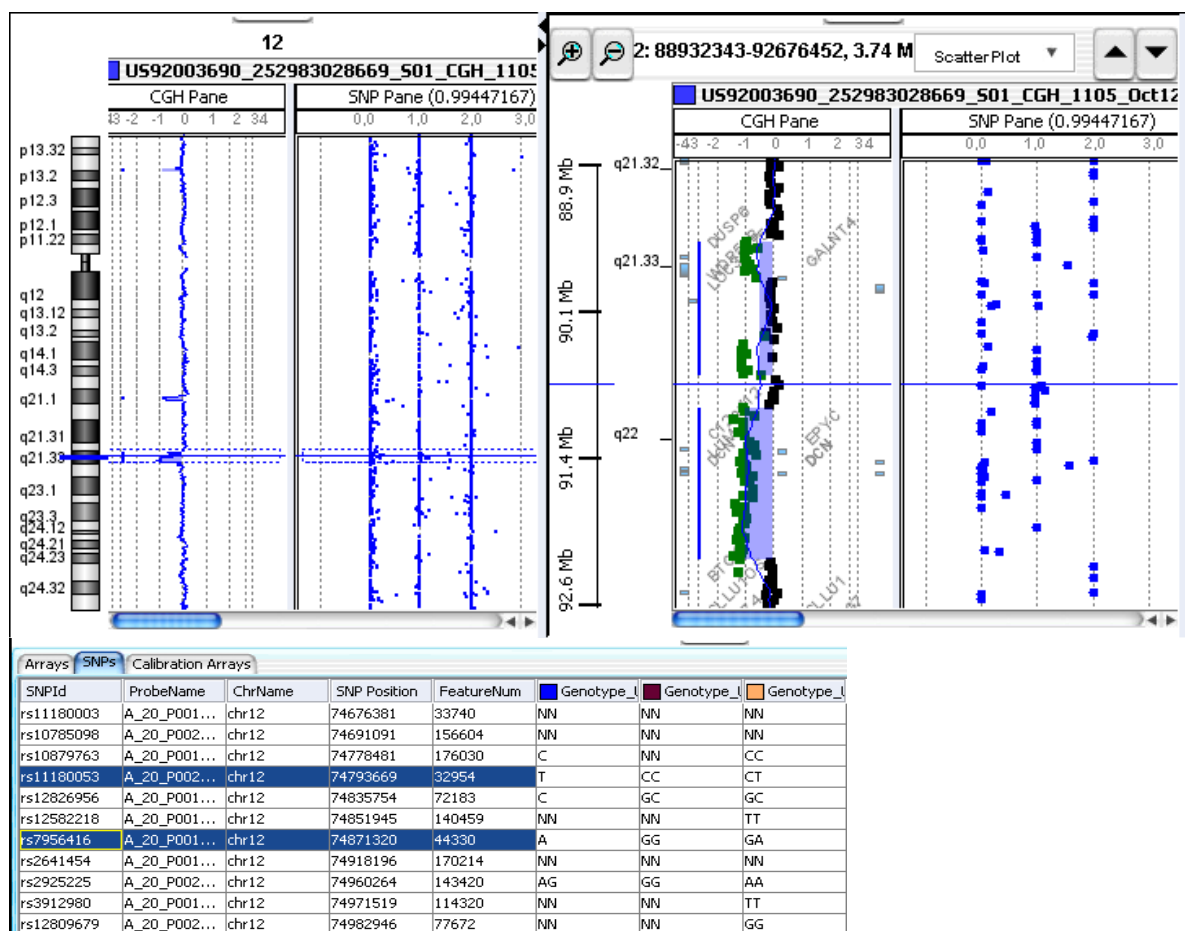


Figura 52. SNP array del cromosoma 12 del caso E/1021 e dei suoi genitori.

L'analisi del cariotipo ha rilevato una verosimile inversione paracentrica sul braccio lungo del cromosoma 12, con apparente bp prossimale in q21.2 e bp distale in q24.31 (Figura 53).

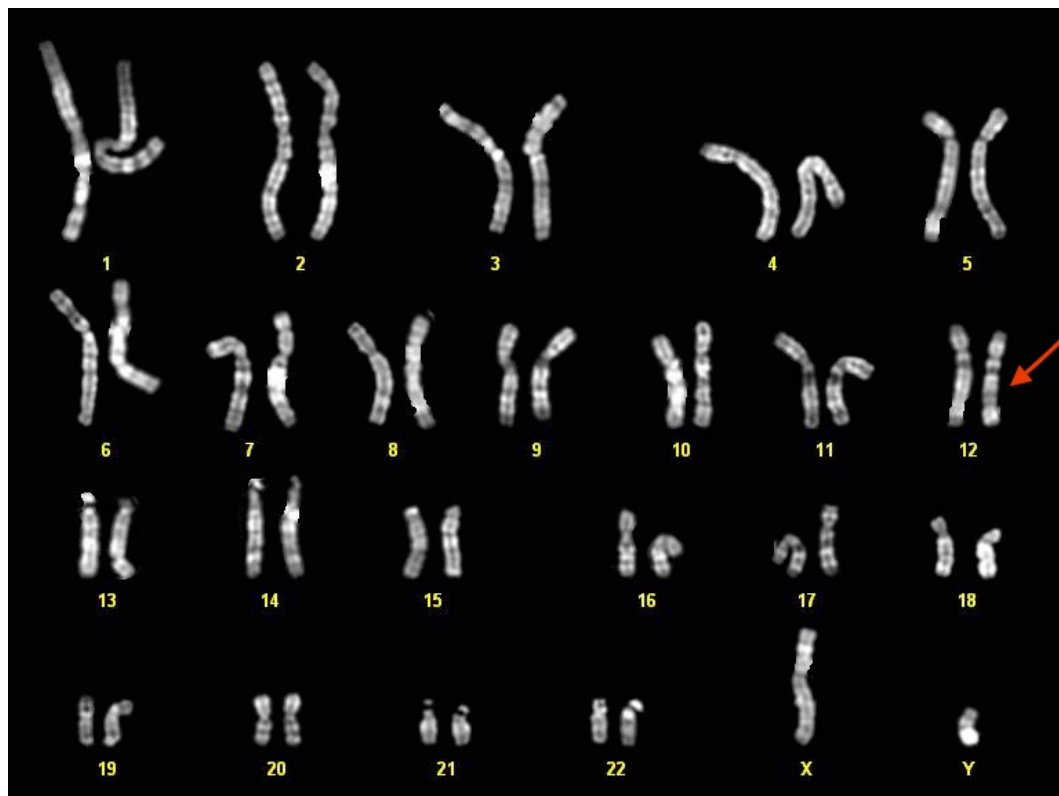


Figura 53. Cariotipo del caso E/1021.

La freccia rossa evidenzia l'inversione paracentrica a livello del braccio lungo del cromosoma 12.

E' stata effettuata anche un'analisi FISH utilizzando le due sonde subtelomeriche del braccio corto e del braccio lungo del 12 che hanno mostrato normale localizzazione dei segnali (Figura 54).

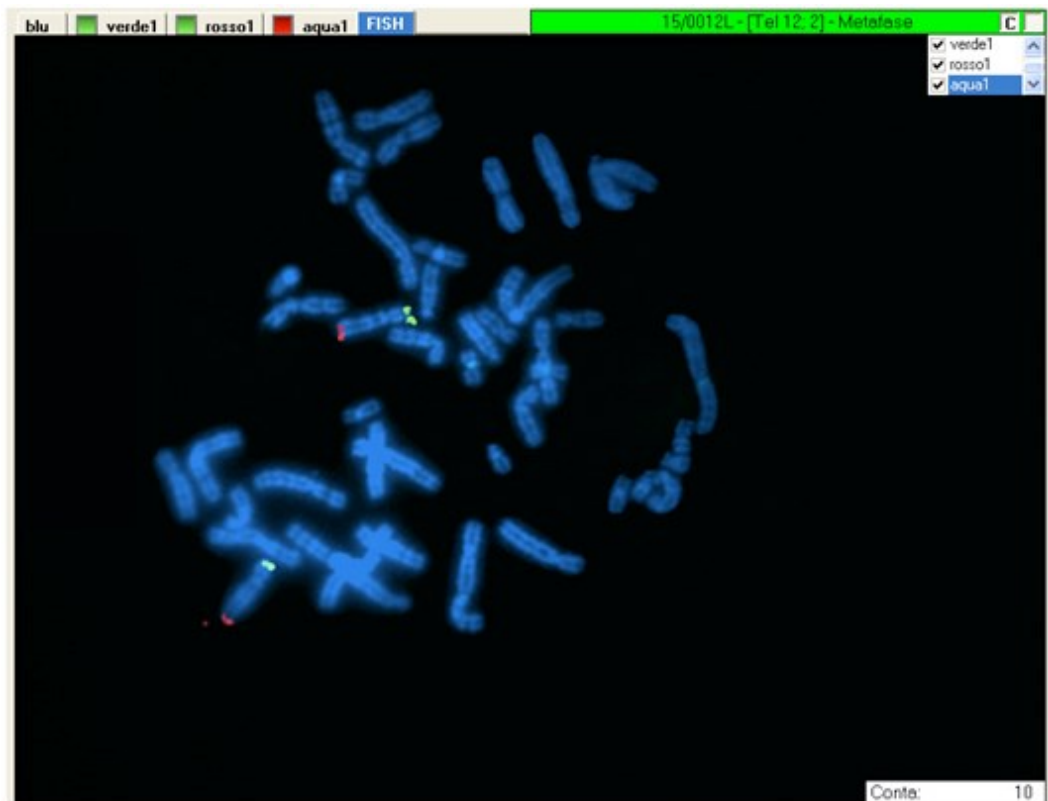


Figura 54. Analisi FISH.

Sonde subtelomeriche che mappano nel braccio corto (segnale verde) e nel braccio lungo (segnale rosso) cromosoma 12. Si rileva normale localizzazione dei segnali.

Per studiare a fondo i riarrangiamenti cromosomici del paziente, è stato eseguito il sequenziamento del genoma (WGS), con le tecniche di Next Generation Sequencing: l'analisi del *coverage* ha confermato la presenza delle quattro delezioni sul cromosoma 12, evidenziate all'array-CGH (Figura 55).

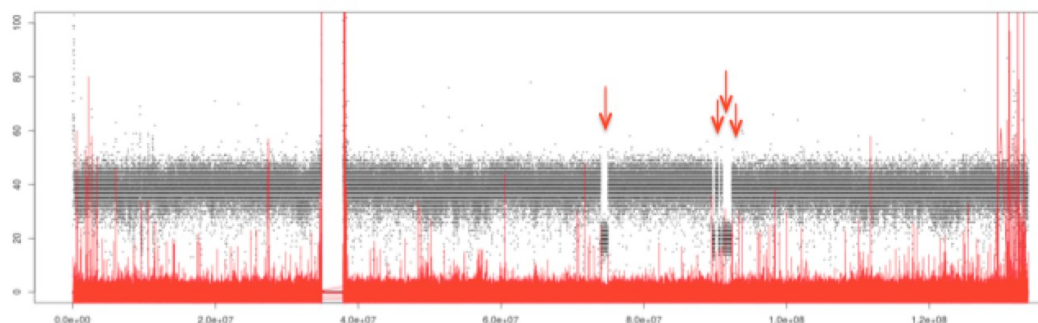


Figura 55. Analisi del *coverage* del cromosoma 12.

Le frecce rosse indicano le delezioni evidenziate dagli array; i *peaks* indicano cluster di *reads* che mappano in modo anomalo in una data regione (riportata sulle ascisse).

Il sequenziamento genomico ha permesso di identificare numerosi altri riarrangiamenti di struttura, che non sono stati rilevati dall'analisi array-CGH. Tutti questi riarrangiamenti sono localizzati a livello del braccio lungo del cromosoma 12; in particolare, rispetto all'analisi array, sono state individuate 15 delezioni prossimali al bp prossimale, e una delezione più distale rispetto al bp distale (892,274,591).

-12q12:43,977,440-43,979,554 (la regione non contiene geni)
-12q12:45,903,128-45,910,101 (la regione non contiene geni)
-12q13.13:52,867,871-52,886,730 (la regione contiene il gene *KRT6A*)
-12q13.13:53,375,650-53,378,987 (la regione non contiene geni)
-12q14.1:60,521,824-60,525,300 (la regione non contiene geni)
-12q14.1:61,600,344-61,617,534 (la regione non contiene geni)
-12q14.2:63,577,414-63,607,316 (la regione non contiene geni)
-12q15:69,255,738-69,256,563 (la regione contiene l'introne 7 del gene *CPM*)
-12q15:70,380,665-70,412,833 (la regione non contiene geni)
-12q15:70,594,863-70,597,489 (la regione non contiene geni)
-12q15:70,872,231-70,878,230 (la regione non contiene geni)
-12q21.1:71,709,544-71,709,916 (la regione contiene l'introne 2 del gene *TSPAN8*)
-12q21.1:72,189,159-72,190,574 (la regione non contiene geni)
-12q21.1:72,356,681-72,358,116 (la regione contiene un introne del gene *TPH2*)
-12q21.1:72,397,807-72,398,309 (la regione contiene un introne del gene *TPH2*)
-12q24.33:133,667,210-133,672,287 (la regione contiene un introne del gene *ZNF140*).

L'analisi dei dati del sequenziamento genomico, ha permesso di riconoscere, all'interno della regione compresa tra 73,943,003 e 130,878,247 (regione che comprende le delezioni individuate agli array-CGH), numerose delezioni di piccole dimensioni (Figura 56a):

- 12q21.1:73,943,003-74,969,210
- 12q21.33:89,535,464-89,535,489
- 12q21.33:89,539,276-89,545,832

- 12q21.33:89,576,403-89,576,425
- 12q21.33:89,581,067-89,964,474
- 12q21.33:90,457,062-90,754,308
- 12q21.33:90,996,549-92,305,750
- 12q21.33:92,367,775-92,369,962
- 12q21.33:92,380,391-92,380,690
- 12q21.33:92,384,519-92,387,539
- 12q22:92,954,764-92,958,817
- 12q22:93,353,092-93,353,195
- 12q23.1:98,268,803-98,268,830
- 12q23.1:100,011,892-100,011,923
- 12q24.33:130,825,589-130,825,824
- 12q24.33:130,877,754-130,878,247

Il riassetto del cromosoma ha individuato riarrangiamenti complessi, con inversioni e inserzioni (Figura 56b); purtroppo non è stato possibile ricollocare in modo preciso alcune regioni genomiche che non si sono perse, ma si sono riarrangiate (Figura 56c).

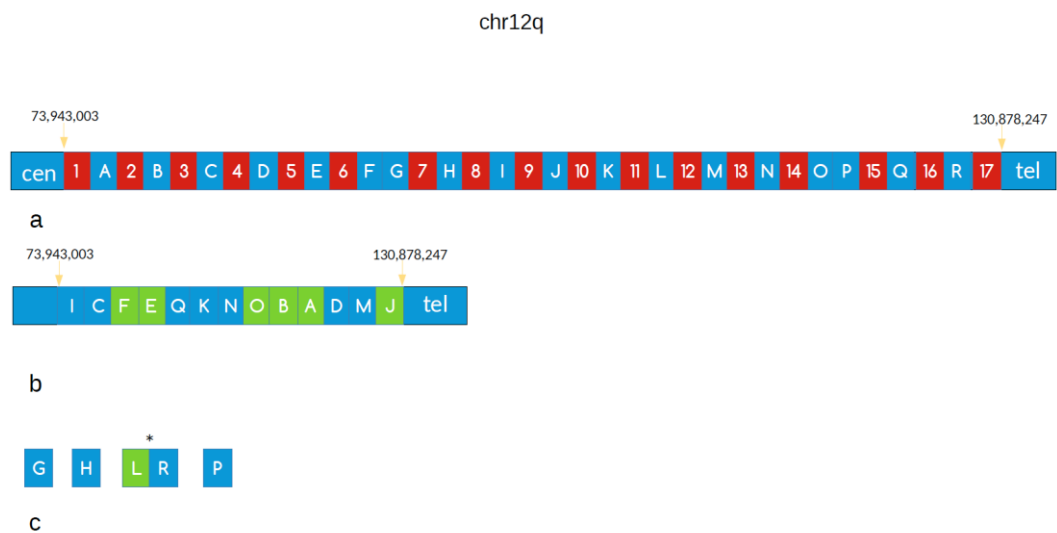


Figura 56. Schema del riarrangiamento complesso in 12q.

a) Nello schema, numeri e lettere si riferiscono al braccio lungo del cromosoma 12 da 73,943,003 a 130,878,247. I numeri negli intervalli rossi indicano le regioni che risultano delete nel cromosoma riarrangiato, e nel testo sono indicati gli intervalli di delezione. Le lettere (A-R) mostrano le regioni in 12q rimanenti. b) Schema del 12q dopo cromosomasintesi. In verde sono indicate le regioni riarrangiate e invertite, in base alla ricostruzione dopo mappaggio dei mate. c) Le regioni genomiche indicate dalle lettere G, H, L, R e P non si sono potute collocare nella giusta posizione del 12q riarrangiato per mancanza degli esatti punti di giunzione, che verosimilmente non si sono potuti localizzare per la presenza di elevate sequenze ripetute. L'asterisco tra L e R indica l'unico punto di giunzione che siamo stati in grado di mappare. La regione L è invertita.

4.3 CASO C/1642

L'analisi del DNA della bambina, eseguita con metodica array-CGH (Figura 57 e 58), alla risoluzione di 100kb, ha rilevato la presenza di:

- delezione interstiziale di circa 3 Mb sul braccio corto del cromosoma 1
[1p36.33-p36.32(1,860,742-4,843,523)x1]
- duplicazione interstiziale di circa 3,8 Mb sul braccio lungo del cromosoma 4
[4q32.1(155,627,657-159,502,724)x3]
- duplicazione interstiziale di circa 0,265Mb sul braccio lungo del cromosoma 4
[4q35.1(185,009,799-185,274,661)x3]
- duplicazione interstiziale di circa 1.5Mb sul braccio lungo del cromosoma 4
[4q35.1-q35.2(186,615,058-188,039,565)x3]
- duplicazione interstiziale di circa 0.294Mb sul braccio lungo del cromosoma 4
[4q35.2(190,175,556-190,469,478)x3]

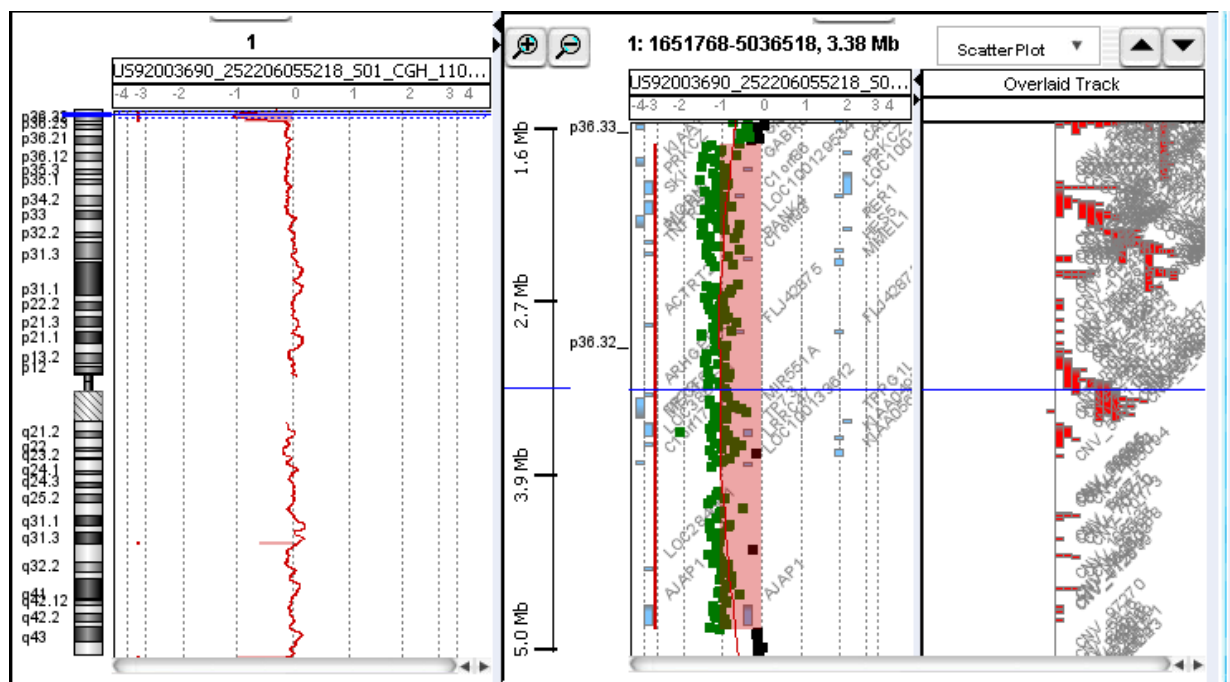


Figura 57. Ideogramma del cromosoma 1 del caso C/1642

L'analisi array-CGH evidenzia un'alterazione del numero di copie sul cromosoma 1, con una delezione sul braccio corto di un cromosoma 1.

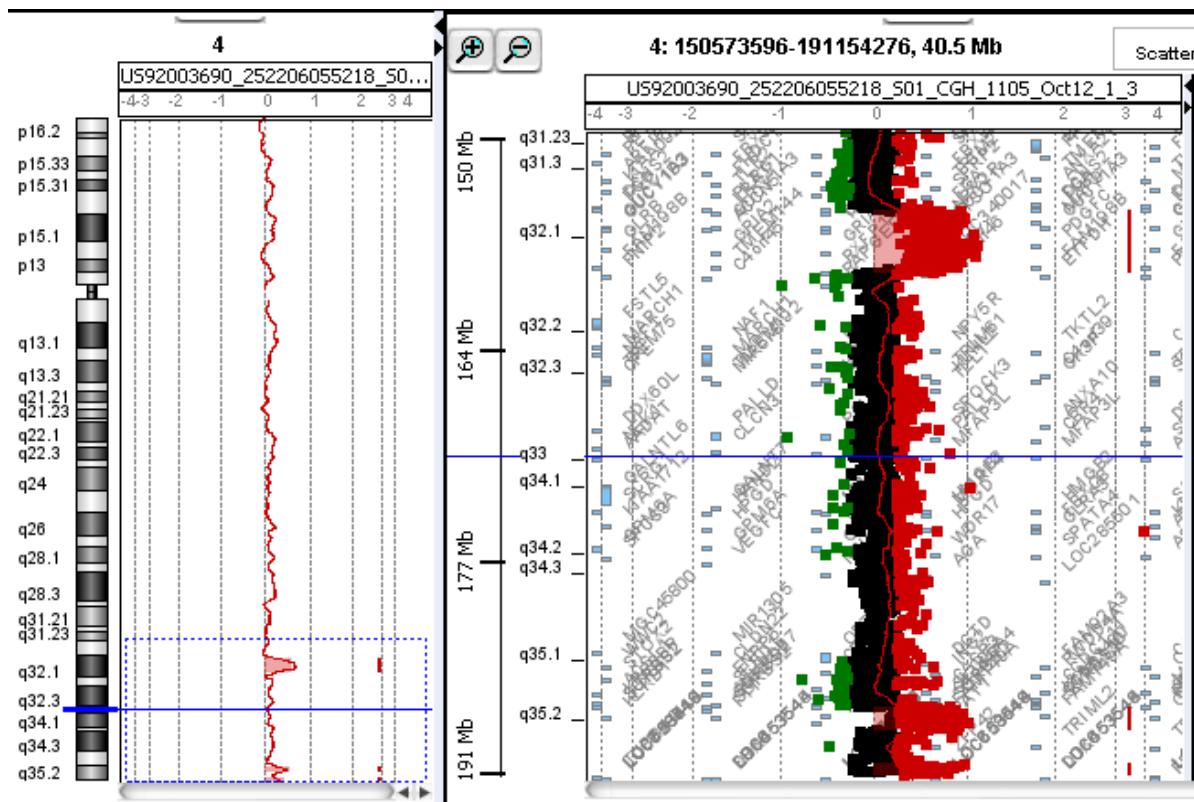


Figura 58. Ideogramma del cromosoma 4 del caso C/1642

L'analisi array-CGH evidenzia un'alterazione del numero di copie sul cromosoma 4, con quattro successive duplicazioni sul braccio lungo del cromosoma 4.

Delezione 1p36.33-p36.32

L'analisi NGS della regione deleta è in corso e i punti di rottura, da caratterizzare mediante sequenziamento, sono in via definizione.

La regione deleta nel cromosoma 1 coinvolge numerosi geni, fra cui i geni *GABRD*, *SKI*, *PEX10* e *PRDM16*. Tale delezione è causativa della sindrome da microdelezione 1p36 (Figura 59).

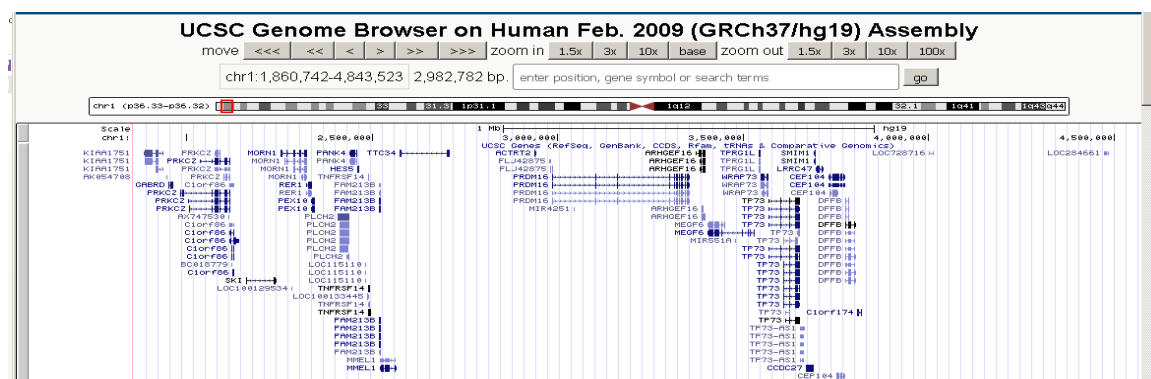


Figura 59. Schema del cromosoma 1

Dimensione e la sede della delezione interstiziale a livello del braccio corto: 1,860,742 (*breakpoint* prossimale, ultima sonda presente prima della delezione) -4,843,523 (*breakpoint* distale, prima sonda presente dopo la regione deleta). Immagine ripresa da UCSC Genome Browser. Sono indicati i numerosi geni contenuti in questa regione, tra i quali prossimalmente i geni malattia *GABRD*, *SKI*, *PEX10* e *PRDM16*.

Duplicazione 4q32.1

L'analisi NGS della regione duplicata è in corso e i punti di rottura, da caratterizzare mediante sequenziamento, sono in via definizione.

La regione duplicata nel cromosoma 4 coinvolge numerosi geni, tra cui i geni malattia *LRAT*, *GUCY1A3*, e *GLRB* (Figura 60).

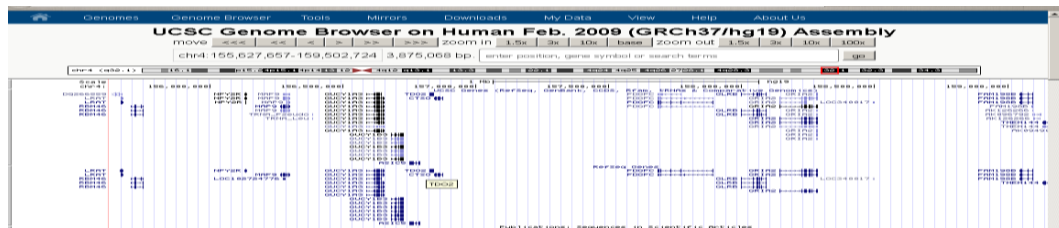


Figura 60. Schema del cromosoma 4

Dimensione e la sede della duplicazione interstiziale a livello del braccio lungo: 155,627,657 (*breakpoint* prossimale, ultima sonda presente prima della duplicazione) -159,502,724 (*breakpoint* distale, prima sonda presente dopo la regione duplicata). Immagine ripresa da UCSC Genome Browser. Sono indicati i numerosi geni contenuti in questa regione, tra i quali i geni malattia *LRAT*, *GUCY1A3* e *GLRB*.

Duplicazione 4q35.1

L'analisi NGS della regione duplicata è in corso e i punti di rottura, da caratterizzare mediante sequenziamento, sono in via definizione.

La regione duplicata nel cromosoma 4 coinvolge il gene *ENPP6* (Figura 61).

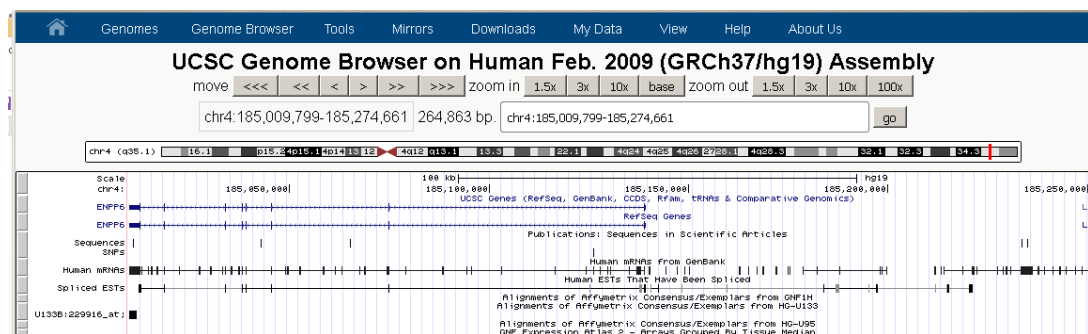


Figura 61. Schema del cromosoma 4

Dimensione e la sede della duplicazione interstiziale a livello del braccio lungo: 185,009,799 (*breakpoint* prossimale, ultima sonda presente prima della duplicazione) -185,274,661 (*breakpoint* distale, prima sonda presente dopo la regione duplicata). Immagine ripresa da UCSC Genome Browser. Il gene contenuto in questa regione è *ENPP6*.

Duplicazione 4q35.1-q35.2

L'analisi NGS della regione duplicata è in corso e i punti di rottura, da caratterizzare mediante sequenziamento, sono in via definizione.

La regione duplicata in 4q35.1-q35.2 coinvolge vari geni, fra cui i geni malattia *TLR3*, *CYP4V2*, *KLKB1*, *F11* (Figura 62).

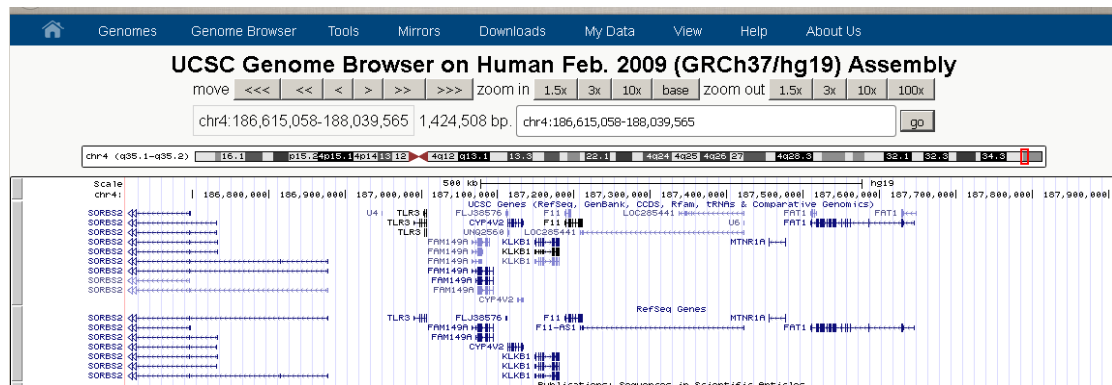


Figura 62. Schema del cromosoma 4

Dimensione e la sede della duplicazione interstiziale a livello del braccio lungo: 186,615,058 (*breakpoint* prossimale, ultima sonda presente prima della duplicazione) -188,039,565 (*breakpoint* distale, prima sonda presente dopo la regione duplicata). Immagine ripresa da UCSC Genome Browser. La regione duplicata in 4q35.1-q35.2 coinvolge vari geni, fra cui i geni malattia *TLR3*, *CYP4V2*, *KLKB1*, *F11*

Duplicazione 4q35.2

L'analisi NGS della regione duplicata è in corso e i punti di rottura, da caratterizzare mediante sequenziamento, sono in via definizione.

La regione duplicata in 4q35.2 non comprende geni malattia codificanti (Figura 63)

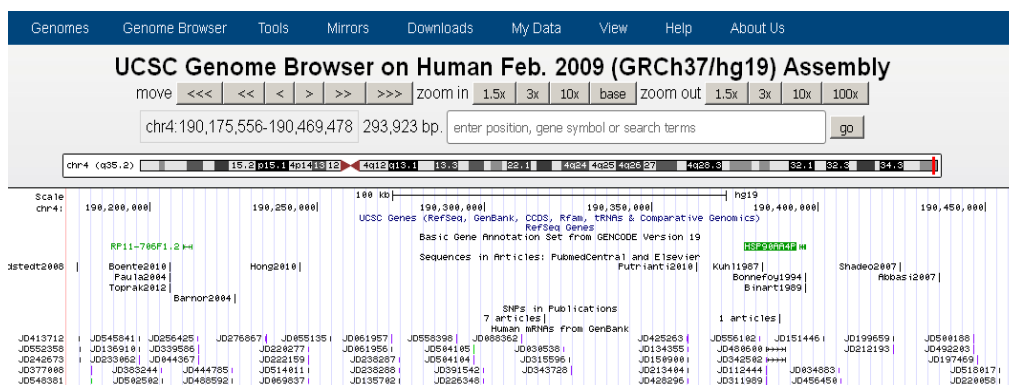


Figura 63. Schema del cromosoma 4.

Dimensione e la sede della duplicazione interstiziale a livello del braccio lungo: 190,175,556 (*breakpoint* prossimale, ultima sonda presente prima della duplicazione) -190,469,478 (*breakpoint* distale, prima sonda presente dopo la regione duplicata). Immagine ripresa da UCSC Genome Browser. La regione duplicata in 4q35.2 non coinvolge geni.

L'esame degli SNP array ha evidenziato che tutti i riarrangiamenti sono di origine materna (Figura 64).

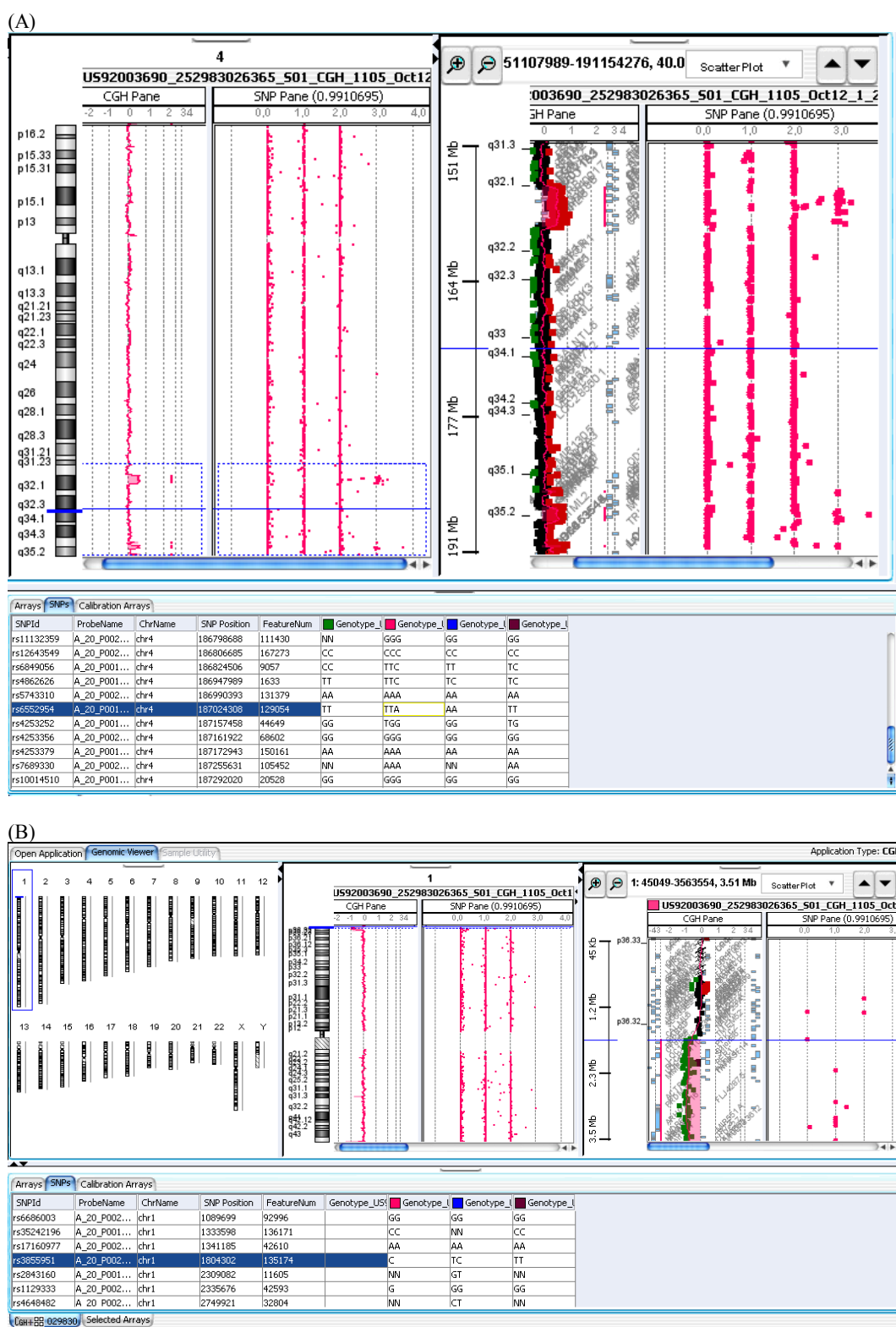


Figura 64. SNP array del cromosoma 4 (A) e del cromosoma 1 (B) del caso C/1642 e dei suoi genitori.

5. DISCUSSIONE

Il sequenziamento del genoma (WGS) ha messo in luce un nuovo fenomeno mutazionale sia nei tumori che nei disordini congeniti, chiamato cromotripsì.

La cromotripsì si verifica quando un cromosoma si frammenta in molti pezzi, tutti in una volta, e viene poi “ricucito” insieme in un ordine casuale, come descritto in origine dopo sequenziamento del genoma in un caso di leucemia linfatica cronica⁶⁰. Successivamente questo meccanismo è stato quindi identificato in diversi tipi di tumori e in disordini congeniti⁶⁹.

Le analisi dei breakpoints hanno identificato diverse caratteristiche che contraddistinguono la cromotripsì da altri processi di riarrangiamenti genomici e che sono alla base degli attuali modelli di spiegazione della cromotripsì.

In uno studio recente, Zhang et al. hanno proposto e identificato un meccanismo cellulare che potrebbe causare questo fenomeno nelle cellule tumorali⁸¹ partendo da un precedente lavoro di Pellman e al., che aveva dimostrato che i cromosomi intrappolati in micronuclei possono frammentarsi e quindi determinare un primo passo verso la cromotripsì⁵⁶.

Utilizzando una combinazione di *imaging* cellulare dal vivo e sequenziamento del genoma da singola cellula, il gruppo di Zhang ha dimostrato che la formazione di micronuclei può effettivamente generare uno spettro di riarrangiamenti genomici, alcuni dei quali ricapitolano tutte le caratteristiche note per la cromotripsì. Questi eventi sono limitati ad una cattiva segregazione cromosomica e si verificano all'interno di una divisione cellulare. Ha dimostrato che il meccanismo di cromotripsì può comportare la frammentazione e successivo rimontaggio di un singolo cromatide da un micronucleo. Messi insieme, questi esperimenti stabiliscono un nuovo processo di mutazione, del quale la cromotripsì rappresenta un esito estremo⁸¹.

All'inizio era stato il gruppo di Crasta ad aver proposto che l'isolamento fisico dei cromosomi in strutture nucleari aberranti, chiamate micronuclei, poteva spiegare la localizzazione delle rotture del DNA evidenziate nelle cromotripsì⁷⁰.

Un alterato attaccamento del centromero dei cromosomi al microtubulo può causare cosiddetti “cromosomi in ritardo” (*lagging*) durante l'anafase (1), che sono un segno distintivo di cellule cromosomicamente instabili (2), e può portare alla

formazione di micronuclei (MN), cioè di piccole formazioni simili al nucleo “principale”, che si formano se un cromosoma rimane separato dal gruppo principale di cromosomi al momento della formazione della membrana nucleare. I MN sono stati a lungo usati come marcatori di “fedeltà genetica”. Ad esempio, possono essere predittivi di tumorigenicità e possono anche essere utilizzati per lo screening di sostanze chimiche (genotossicità)⁸⁴⁻⁸⁶.

Crasta e Zhang hanno inoltre dimostrato che la suddivisione dei cromosomi intatti in micronuclei di nuova formazione porta ad una evidenza citologica di danni al DNA, in particolare sulla mis-segregazione del cromosoma. Dopo la mitosi, i cromosomi possono essere reincorporati dal micronucleo nel nuclei figli, potenzialmente integrando le mutazioni dal micronucleo nel genoma^{70,81}.

I micronuclei sono il risultato comune di molti difetti della divisione cellulare, inclusi gli errori mitotici in cui mis-segregano cromosomi intatti, e di errori nella replicazione o riparazione del DNA che generano frammenti residui di cromosomi acentrici^{82,83,84}.

Partendo da queste osservazioni e dal fatto che i cromosomi riarrangiati possono essere ripresi nel nucleo principale anche durante l’embriogenesi e non solo nelle cellule tumorali, in questo lavoro di tesi abbiamo esaminato i genomi di tre bambini giunti alla nostra osservazione per una serie di anomalie congenite e disabilità intellettiva. Abbiamo ottenuto il consenso informato da parte dei soggetti coinvolti per analizzare i loro genomi e discutere i risultati. Mediante sequenziamento Illumina, abbiamo individuato differenti riarrangiamenti sia nel cromosoma 17 (caso C/996) che nel cromosoma 12 (caso E/1021).

I nostri tre casi erano stati precedentemente analizzati mediante whole-genome array-CGH che aveva mostrato in tutti e tre i casi una serie di riarrangiemnti complessi che abbiamo considerato derivare da un meccanismo di cromotrips o cromoanasintesi.

Il primo bambino mostrava un fenotipo complesso, in cui coesistevano tre differenti disordini genetici: una neuropatia sensitivo-motoria, la malattia di Charcot-Marie-Tooth di tipo 1A, la sindrome di Potocki.Lupski (PLS) e la neurofibromatosi di tipo 1.

La malattia di Charcot-Marie-Tooth è una malattia genetica che colpisce il sistema nervoso periferico. È caratterizzata da debolezza e atrofia dei muscoli, che

appaiono come “smagriti”, e da ridotta sensibilità. I sintomi partono in genere dai piedi e si diffondono progressivamente verso l’alto a gambe, parte delle cosce e mani. Sono frequenti anche deformazioni dello scheletro, in particolare dei piedi, che si presentano cavi. La malattia esordisce di solito prima dei 20 anni e può manifestarsi in varie forme, più o meno invalidanti⁸⁷.

Il nostro bambino presentava già “piede cavo” bilateralmente e difficoltà alla deambulazione, anche se questa poteva essere associata ad ipotonia, caratteristica della PLS. E’ causata da una duplicazione del gene *PMP22*, che causa sovraespressione della proteina omologa, la quale determina l’instabilità della mielina periferica e la secondaria degenerazione assonale. La mutazione si trasmette in modo autosomico dominante.

La PLS è una condizione genetica caratterizzata dalla presenza di ritardo dello sviluppo, ipotonia, ritardo di crescita, ritardo mentale, disturbi pervasivi dello sviluppo, e malformazioni congenite agli organi interni. Tutti i casi riportati in letteratura si verificano *de novo*, senza la presenza di riarrangiamenti cromosomici nei genitori. La maggior parte delle duplicazioni ha una dimensione di circa 3,7 Mb in termini di dimensioni ed è quindi identificabile tramite tecniche di cariotipizzazione molecolare, quali array-CGH. Circa il 60% dei pazienti affetti da sindrome di Potocki-Lupski ha una microduplicazione del cromosoma 17p11.2 reciproca alla delezione ricorrente comune di 3,7 Mb presente nella sindrome di Smith-Magenis (SMS). La PLS è una sindrome da geni contigui e considerato che la regione duplicata è fiancheggiata da duplicazioni segmentali, è probabile che il meccanismo patogenetico sia quello di una ricombinazione omologa non allelica (NAHR). Spesso però il meccanismo è di tipo NHEJ, con un bp non ricorrente⁸⁸. Quindi la duplicazione nel braccio corto del 17, con punto di rottura proximale a 11,385,129 e distale 21,502,083, fa sì che questi due disordini genetici sussistano nello stesso paziente.

Proprio per le caratteristiche di questa regione, in cui coesistono tre regioni ad alta omologia, ha reso difficile la caratterizzazione del bp distale (vedi Risultati). Infatti è stato possibile caratterizzare il bp distale in una regione più proximale rispetto al reale punto di rottura evidenziato con gli array. Questo è un risultato che poteva essere atteso in quanto questa regione è ricca di sequenze altamente

ripetute, come ampiamente dimostrato da Lupski e coll⁸⁹ (Figura 65).

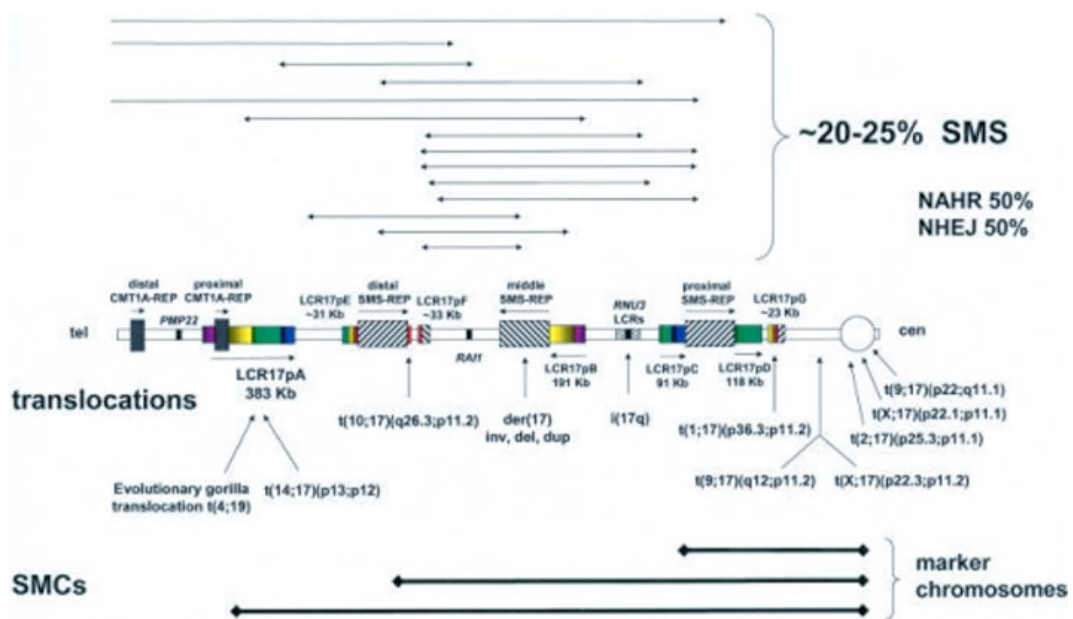


Figura 65. Riarrangiamenti non ricorrenti in 17p prossimale con la sua complessa architettura del genoma e molte LCR.

Il centromero (cen) è a destra, il telomero (tel) a sinistra. I rettangoli colorati raffigurano regioni LCR con identità di sequenza maggiore del 97%, e le frecce orizzontali raffigurano l'orientamento. Sono indicate le posizioni del gene *RAI1* (la cui aploinsufficienza determina la SMS) e i breakpoint dell'isocromosoma 17q. Sopra è mostrata la regione deleta nei pazienti con SMS, con delezioni non ricorrenti nei bp comuni. I punti di rottura sono indicati da frecce. Di seguito sono indicate le regioni contenute nei cromosomi marcatore soprannumerari (SMC). Inoltre, di seguito sono riportati i punti di rottura in 17p11.2 evidenziati nelle traslocazioni.

Il nostro caso presentava inoltre macchie caffè-latte che sono una delle caratteristiche della neurofibromatosi di tipo 1 (NF1). La NF1 è la più comune forma di neurofibromatosi con un'incidenza di 1 su 2.500-3.500 nati. E' una condizione multisistemica che coinvolge una varietà di tessuti e tipi cellulari; i segni clinici diagnostici sono ben definiti e sufficienti per porre diagnosi clinica.

E' una malattia a trasmissione autosomica dominante a penetranza età-dipendente completa (i segni clinici compaiono progressivamente fino a diventare del 100% all'età di 6-8 anni); l'espressività è anch'essa età dipendente ed è estremamente variabile, anche all'interno della stessa famiglia⁹⁰.

La NF1 è causata da alterazioni del gene *NF1* che mappa in 17q11.2. Il trascritto codifica per una proteina chiamata neurofibromina, che presenta una regione di 360 amminoacidi, denominata NF1-GRD (GAP related domain), che mostra un'omologia con il dominio catalitico delle "GTPase activating protein" (GAP) dei mammiferi, ed è in grado di svolgere una regolazione negativa sull'attività dell'oncogene p21ras. Questa attività della neurofibromina consente di inquadrare

il gene *NF1* nella categoria dei geni oncosoppressori: sia la perdita della copia *wild type* del gene che porta a LOH (*loss of heterozygosity*), sia la presenza di mutazioni somatiche puntiformi all'interno del gene che ne causano l'inattività, si sono riscontrate in uno svariato numero di neuro fibromi⁹².

Il nostro caso presenta una delezione che rompe il gene a livello dell'introne 35, determinando una verosimile aploinsufficienza della proteina. In questa regione deleta inoltre risulta inserita e duplicata in direzione invertita una regione prossimale di 17q, a dimostrazione della complessità del riarrangiamento.

Il secondo caso (E/1021) presentava, dopo analisi array-CGH, quattro delezioni interstiziali a livello del braccio lungo del cromosoma 12 e una verosimile inversione paracentrica (vedi cariotipo). Il riarrangiamento in realtà si presentava molto più complesso, con una serie di piccole delezioni e inserzioni di regioni probabilmente duplicate e invertite, messe in luce dal WGS. La complessità del riarrangiamento ha reso molto difficile la ricostruzione del cromosoma 12 e dell'esatta giunzione dei differenti bp.

Il fenotipo del paziente è caratterizzato da disabilità intellettiva, e soprattutto anomalie del comportamento in via di definizione e caratterizzazione. Il paziente inoltre presenta distrofia corneale, associabile alla aploinsufficienza del gene *DCN* (Figura 66), con opacizzazione corneale costituita da un gran numero di macchie in tutti gli strati della stroma. Essa si sta traducendo in una progressiva perdita della vista. Erosioni corneali e fotofobia sono assenti

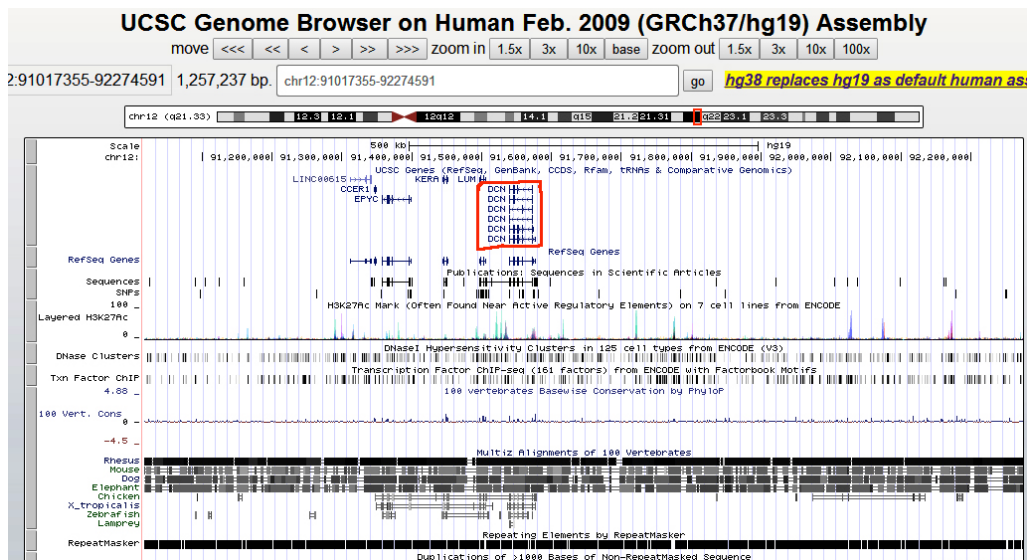


Figura 66. Schema del cromosoma 12 che indica la dimensione e la sede della delezione interstiziale: 91,017, 355 (*breakpoint* prossimale, ultima sonda presente prima della delezione) - 274,591 (*breakpoint* distale, prima sonda presente dopo la regione deleta). Immagine ripresa da UCSC Genome Browser. In rosso indicato il gene *DCN* le cui mutazioni determinano distrofia corneale.

Al momento sono davvero poche le famiglie riportate in letteratura con mutazioni nel gene della decorina (*DCN*), in cui sia presente una distrofia stromale della cornea che cosegrega con un meccanismo di tipo autosomico dominante⁹². Sino ad oggi, visti i lavori della letteratura, questo è il primo caso in cui è presente una delezione completa del gene e una forma di questa patologia ad insorgenza pediatrica.

Il terzo caso (C/1642) presenta una delezione della porzione distale del braccio corto di un cromosoma 1 (1p36.33-p36.32) di circa 3 Mb, che si associa alla sindrome da microdelezione 1p36, e quattro differenti duplicazioni a livello del braccio lungo di un cromosoma 4.

Delezioni del cromosoma 1p36 colpiscono circa 1 su 5.000 neonati e sono le delezioni terminali più comuni nei nati vivi. Il quadro clinico causato comunemente da delezioni terminali di 1p36 includono ritardo dello sviluppo psicomotorio, disabilità intellettiva, epilessia, problemi di visione, sordità, bassa statura, alcune caratteristiche facciali, anomalie cerebrali, labiopalatoschisi, cardiopatie congenite, cardiomiopatia, e anomalie renali. Sebbene la sindrome da delezione 1p36 sia considerata clinicamente riconoscibile, vi è una significativa variazione fenotipica tra gli individui affetti. Questo è dovuto, almeno in parte,

all'eterogeneità genetica di questo tipo di riarrangiamento che comprende delezioni terminali ed interstiziali di lunghezza variabile da poche kb fino a 30 Mb di DNA. L'analisi array-CGH può facilmente identificare delezioni genomiche in 1p36, tuttavia prevedere il fenotipo di un individuo unicamente sull'entità della delezione rimane una sfida, dal momento che la maggior parte dei geni che contribuiscono al fenotipo devono ancora essere identificati (Figura 67).

Inoltre, l'aploinsufficienza di più di un gene può contribuire ad alcuni fenotipi. In particolare, in base ai dati della letteratura, e ai diversi fenotipi da delezione descritti, i geni coinvolti in questa sindrome sono i seguenti: *MMP23B*, *GABRD*, *SKI*, *PRDM16*, *KCNAB2*, *RERE*, *UBE4B*, *CASZ1*, *PDPN*, *SPEN*, *ECE1*, *HSPG2*, e *LUZP1*⁹³.

La nostra paziente presentava una delezione di circa 3 Mb che coinvolge numerosi geni, fra cui i geni malattia *GABRD*, *SKI*, *PEX10* e *PRDM16*.

L'aploinsufficienza dell'acido gamma-aminobutirrico (GABA), recettore del gene delta (*GABRD*) è stata suggerita, nei soggetti affetti da questa sindrome, come possibile causa per le anomalie dello sviluppo neurologico, problemi neuropsichiatrici, e convulsioni⁹⁴. *GABRD* codifica per una subunità di un canale del cloro pentamerico ligando-dipendente, che viene attivato dal GABA⁹⁵, il principale neurotrasmettitore inibitorio nel cervello dei mammiferi, e si trova nella regione critica per l'epilessia definita da Rosenfeld et al.⁹⁶ Topi *Gabrd*-null esibivano comportamenti simil-depressivi e ansia. Tuttavia, il loro apprendimento e la memoria erano normali quando valutati utilizzando il condizionamento con la paura^{96,97}.

In effetti anche la nostra paziente mostrava un fenotipo caratterizzato da disabilità intellettiva, ritardo psicomotorio ed epilessia.

Il proto-oncogene *SKI* codifica invece per una proteina che agisce come un co-regolatore trascrizionale^{98,99}. Aploinsufficienza di *SKI* si pensa contribuisca al ritardo dello sviluppo psicomotorio, al ritardo mentale, alle crisi epilettiche, alla labiopalatoschisi, e alla cardiopatia congenita Colmenares et al. hanno dimostrato che i topi *Ski*-null hanno difetti del tubo neurale, anomalie del prosencefalo e della linea mediana, *clefting* facciale, e ridotta massa muscolare scheletrica¹⁰⁰. Mutazioni puntiformi e piccole delezioni in-frame di *SKI* causano la sindrome di

Shprintzen-Goldberg, che presenta alcune caratteristiche presenti anche nei soggetti con delezione di 1p36¹⁰¹.

Il gene *PEX10* codifica per una proteina che si localizza nella membrana perossisomale. Le mutazioni in questo gene producono fenotipi all'interno dello spettro della malattia di Zellweger, che vanno da adrenoleucodistrofia neonatale a sindrome di Zellweger. Queste patologie si trasmettono con meccanismo autosomico recessivo, quindi non basta la sola aploinsufficienza di questo gene, come nel caso della nostra paziente, per la determinazione di questo fenotipo.

Il dominio PR contenente il gene 16 (*PRDM16*) codifica per un fattore di trascrizione zinc finger. Molte evidenze suggeriscono che aploinsufficienza di *PRDM16* contribuisce a cardiomiopatia dilatativa e non-compattazione ventricolare sinistra. *PRDM16* si trova nella regione critica per la cardiomiopatia e la proteina si esprime nei nuclei dei cardiomiociti murini e umani sia nel periodo dello sviluppo che nell'età adulta¹⁰².

La cardiopatia è presente anche nella nostra paziente, con quindi una diretta correlazione genotipo/fenotipo.

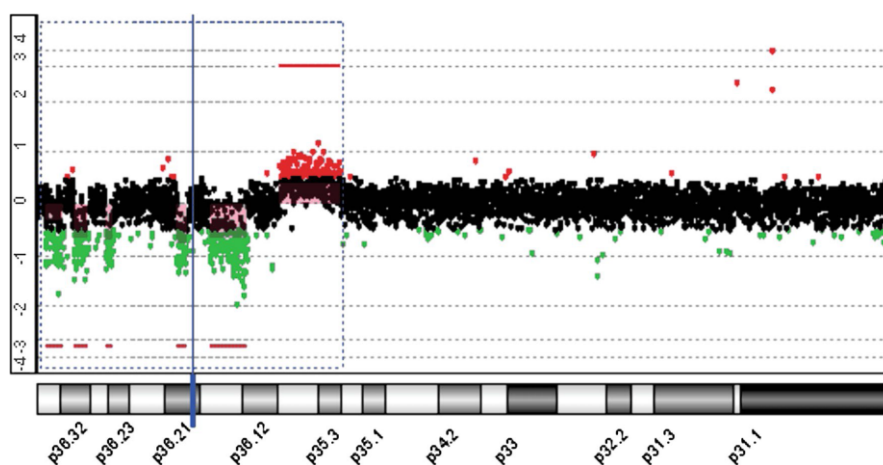


Figura 67. Profilo Array-CGH della regione terminale del braccio corto del cromosoma 1 che mostra cinque delezioni e una duplicazione in 1p36.33-1p35.3.

Un meccanismo di cromotripsia è stato anche osservato recentemente in un soggetto portatore di delezione in 1p36¹⁰³.

In questo caso è stato osservato un fenotipo atipico da delezione 1p36. Grazie alla sua combinazione unica di squilibri genomici, è stato difficile correlare il genotipo di questo caso confrontandolo con quello di altri pazienti con lo stesso riarrangiamento genomico.

Nel caso della nostra paziente invece il fenotipo è abbastanza ascrivibile alla sindrome che ha un quadro ben riconoscibile, definito e abbastanza grave, ed è più difficile riconoscere alcuni segni clinici che potrebbero essere determinati anche dalle differenti duplicazioni presenti a livello del braccio lungo del cromosoma 4, risultati da uno stesso meccanismo di cromotripsia.

La prima duplicazione, in 4q32.1, di circa 4 Mb, coinvolge numerosi geni, tra cui i geni malattia *LRAT*, *GUCY1A3*, e *GLRB* che si associano a malattia quando sono presenti mutazioni in entrambi gli alleli. Le malattie quindi vengono trasmesse con meccanismo autosomico recessivo (AR) e al momento non abbiamo indicazioni in letteratura di patologie legate a mutazioni *gain-of-function*, come potrebbe essere il caso determinato da duplicazione degli stessi. I casi riportati nel database Decipher sono solo in parte sovrapponibili al nostro caso e non abbiamo abbastanza dati clinici per una correlazione genotipo/fenotipo.

La seconda duplicazione in 4q35.1, contiene il gene *ENPP6*, al momento non associato a nessun quadro patologico.

La terza duplicazione, in 4q35.1-q35.2, di circa 1,5 Mb, coinvolge vari geni, fra cui i geni malattia *TLR3*, *CYP4V2*, *KLKB1*, *F11*. I geni *CYP4V2* e *KLKB1* sono stati descritti in associazione a disordini trasmessi sempre con meccanismo AR, e quindi anche in questo caso non ci aspettiamo di ritrovare segni di questi disordini nella nostra paziente. Solo il gene *F11*, che codifica per il fattore XI della coagulazione, può essere associato a forme dominanti del disordine, anche se sono soprattutto descritte mutazioni da perdita di funzione.

L'ultima duplicazione, in 4q35.2, di circa 300bp, non contiene geni codificanti.

Al momento, attraverso il sequenziamento genomico, è in corso la definizione sia dei *breakpoints* delle duplicazioni che l'estensione delle stesse. Inoltre sarà importante definire se siano presenti anche altri riarrangiamenti criptici non evidenziabili attraverso la metodica array-CGH.

6. CONCLUSIONI

Le alterazioni strutturali dei cromosomi sono stati a lungo implicate sia nelle anomalie dello sviluppo che nei tumori. Questi riarrangiamenti cromosomici sono definiti "complessi" quando il riassetto coinvolge più di due punti di rottura, come ampiamente discusso. Recentemente, utilizzando SNP/CGH-array e sequenziamento ad alta processività in una vasta gamma di tipi di cellule tumorali, si sono potuti evidenziare in queste cellule tanti tipi di rotture e tentativi di riparazioni, costituiti da un insieme di decine/centinaia di riarrangiamenti genomici in cluster, localizzati su uno o pochi cromosomi.

Questi riarrangiamenti complessi, simili ad una "polverizzazione" cromosomica, sono stati definiti cromotripsi, letteralmente "cromo" per il cromosoma e "*thripsis*" per "frantumare in mille pezzi". Sorprendentemente, questo fenomeno è risultato essere relativamente comune nei tumori, che rappresentano il 2-3% di tutti i campioni di cancro e fino al 25% dei tumori ossei primari. Tale meccanismo è una delle cause della cancerizzazione cellulare. Anche se le cause e il meccanismo alla base di questi eventi rimangono ancora poco chiari, molti studi hanno ipotizzato che siano coinvolti i meccanismi di riparazione in questi tipi di riarrangiamenti e, in particolare, la NHEJ. Prendendo in considerazione le caratteristiche del sistema di riparazione NHEJ, possiamo immaginare una situazione stressante in cui più rotture del doppio filamento di DNA appaiono simultaneamente in un'unica regione cromosomica. La regione cromosomica verrebbe così ad essere frantumata in pezzi, con la maggior parte dei pezzi che cercano di riunirsi in qualche modo. Mentre alcuni pezzi vengono persi (che potrebbe quindi formare cromosomi marcatori o double-minute), i frammenti conservati potrebbero risistemarsi in modo casuale attraverso il sistema di riparazione NHEJ. Anche il meccanismo di MMBIR sembra essere un buon candidato al ruolo di sistema di riparazione alla base di questi riarrangiamenti complessi. Infatti, questo è un sistema replicativo a bassa processività che utilizza, per riparare il DNA, regioni seriali e di micro-omologia che fungono da stampo, fornendo una spiegazione sia per le sequenze di micro-omologia ai bp che per gli inserimenti di segmenti brevi fiancheggiati, intorno alle giunzioni, di sequenze di micro-omologia. Il termine di cromoanasintesi è stato proposto proprio per illustrare questo secondo

meccanismo, letteralmente "cromo" per il cromosoma e "*anasyntesis*" per la ricostituzione.

Il sequenziamento genomico di tre pazienti analizzati per questo lavoro di tesi, ci ha permesso di individuare molteplici riarrangiamenti a carico dei cromosomi 1 e 4, 12 e 17, così da ipotizzare un meccanismo di cromosomasintesi. Al momento siamo stati in grado, mediante sequenziamento di identificare più di 40 punti di rottura nei pazienti con riarrangiamenti sui cromosomi 12 e 17. E' in corso l'analisi dei riarrangiamenti a carico del cromosoma 1 e 4 del terzo caso.

I riarrangiamenti contengono una successione di segmenti duplicati e deleti in tutto il cromosoma 17 e nel braccio lungo del cromosoma 12. Queste rare anomalie cromosomiche identificate nei nostri pazienti sono quindi fortemente evocative di cromosomasintesi o cromotripsie. Sorprendentemente, i bp non sono stati raggruppati in una specifica regione cromosomica ma sparsi in tutta la lunghezza di entrambi i cromosomi, vedi soprattutto il coinvolgimento del cromosoma 17. La maggiore complessità è stata individuata sul braccio lungo del cromosoma 12 in un paziente che presenta anche un fenotipo clinico apparentemente meno grave. Il sequenziamento mirato del punto di rottura ha consentito di identificare molte regioni di micro-omologia, ma anche di dimostrare come ancora siano presenti dei "buchi" nel nostro genoma, di difficile determinazione. Inoltre, i dati citogenetici ottenuti dal nostro paziente con anomalie del cromosoma 17 sono compatibili con un parziale trisomia 17p associata alla sindrome di Potocki-Lupski e alla CMT1A e con due sindromi da microdelezione, tra cui la NF1. Il fenotipo atteso è proprio compatibile con i riarrangiamenti individuati. Anche per il riarrangiamento del 12q è stato possibile determinare la causa della distrofia corneale osservata nel nostro bambino. Diventa più difficile la correlazione con il tipo di comportamento osservato, per gli scarsi dati della letteratura a proposito di questo fenotipo.

Resta da definire la completa correlazione genotipo/fenotipo per il terzo caso.

La cromosomasintesi è un modello interessante che può spiegare molti dei riarrangiamenti cromosomici complessi che sono stati identificati nei pazienti e particolarmente la cromosomasintesi/cromotripsie è probabilmente un meccanismo frequentemente coinvolto in malattie congenite.

Grazie alle tecnologie di Next Generation Sequencing, l'analisi dell'intero genoma umano ha semplificato sicuramente la possibilità di fare una diagnosi precisa e ci

ha consentito caratterizzare una nuova classe di riarrangiamenti cromosomici complessi alla base di malattia.

7. BIBLIOGRAFIA

1. Laurence D. Hurst, The evolution of genomic anatomy. *Trends in Ecology & Evolution*, 1999. 14(3):108–112
2. McKusick VA The Anatomy of the Human Genome. *JAMA*. 2001. 286(18):2289-95.
3. Richardson SR, Doucet AJ, Kopera HC et al., The Influence of LINE-1 and SINE Retrotransposons on Mammalian Genomes. *Microbiol Spectr*. 2015 Apr;3(2):MDNA3-0061-2014.
4. Beck CR, Garcia-Perez JL, Badge RM, et al. , LINE-1 elements in structural variation and disease. *Annu Rev Genomics Hum Genet*. 2011;12:187-215.
5. Smit AF, The origin of interspersed repeats in the human genome. *Curr Opin Genet Dev*. 1996 Dec;6 (6):743-8.
6. Moore CM and Best RG, Chromosomal Genetic Disease: Structural Aberrations. eLS, 2001.
7. Boerkoel CF, Inoue K, Reiter LT, et al., Molecular mechanisms for CMT1A duplication and HNPP deletion. *Ann N Y Acad Sci*. 1999. 14: 883:22-35. Review.
8. Gu W, Zhang F, Lupski JR, Mechanisms for human genomic rearrangements. *Pathogenetics*. 2008 Nov 3;1 (1):4.
9. Rossi E, Riegel M, Messa J, et al., Duplications in addition to terminal deletions are present in a proportion of ring chromosomes: clues to the mechanisms of formation. *J Med Genet*. 2008. 45(3):147-54.
10. Madan K, Nieuwint AW, van Bever Y. Recombination in a balanced complex translocation of a mother leading to a balanced reciprocal translocation in the child. Review of 60 cases of balanced complex translocations. *Hum Genet* 1997. 99(6):806-15.
11. Edelmann L, Spiteri E, Koren K, et al., BE.AT-rich palindromes mediate the constitutional t(11;22) translocation. *Am J Hum Genet*. 2001. 68 (1):1-13.
12. Page SL, Shin JC, Han JY, et al., Breakpoint diversity illustrates distinct mechanisms for Robertsonian translocation formation. *Hum Mol Genet*. 1996. 5(9):1279-88.
13. Kirkpatrick M. How and Why Chromosome Inversions Evolve. *PLoS Biol*. 2010 28;8(9)

14. Muñoz-López M, García-Pérez JL. DNA transposons: nature and applications in genomics. *Curr Genomics*. 2010. 11(2):115-28.
15. Havecker ER, Gao X, Voytas DF. The diversity of LTR retrotransposons. *Genome Biol*. 2004. 5(6):225.
16. Beck CR, Collier P, Macfarlane C, et al., LINE-1 retrotransposition activity in human genomes. *Cell*. 2010. 141(7):1159-70.
17. Kazazian HH Jr. Mobile elements: drivers of genome evolution. *Science* 2004. 303(5664):1626-32.
18. Bennett EA, Keller H, Mills RE, et al., Active Alu retrotransposons in the human genome. *Genome Res*. 2008. 18(12):1875-83.
19. Konkel MK, Batzer MA. A mobile threat to genome stability: The impact of non-LTR retrotransposons upon the human genome. *Semin Cancer Biol*. 2010. 20(4):211-21.
20. Sharp AJ, Locke DP, McGrath SD, et al., Segmental duplications and copy-number variation in the human genome. *Am J Hum Genet*. 2005. 77(1):78-88.
21. Goidts V, Cooper DN, Armengol L, et al., Complex patterns of copy number variation at sites of segmental duplications: an important category of structural variation in the human genome. *Hum Genet*. 2006. 120(2):270-84
22. Giglio S, Calvari V, Gregato et al., Heterozygous submicroscopic inversions involving olfactory receptor-gene clusters mediate the recurrent t(4;8)(p16;p23) translocation. *Am J Hum Genet*. 2002. 71(2):276-85.
23. Liu P, Carvalho CM, Hastings PJ, et al., Mechanisms for recurrent and complex human genomic rearrangements. *Curr Opin Genet Dev*. 2012 Jun;22(3):211-20.
24. Conrad DF, Pinto D, Redon R, et al., Origins and functional impact of copy number variation in the human genome. *Nature*. 2010. 464(7289):704-12.
25. Zhang F, Carvalho CM, Lupski JR, Complex human chromosomal and genomic rearrangements. *Trends Genet*. 2009 Jul;25(7):298-307.
26. Hasting PJ, Lupski JR, Rosenberg SM, et al., Mechanisms of change in gene copy number. *Nat Rev Genet*. 2009 Aug;10(8):551-64.
27. Lee AJ, Carvalho CM, Lupski JR, A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell*. 2007 Dec 28;131(7):1235-47.

28. Daley JM, Gaines WA, Kwon Y, et al., Regulation of DNA pairing in homologous recombination. *Cold Spring Harb Perspect Biol.* 2014. 6(11):a017954.
29. Mladenov E, Iliakis G. Induction and repair of DNA double strand breaks: the increasing spectrum of non-homologous end joining pathways. *Mutat Res.* 2011. 711(1-2):61-72.
30. Weckselblatt B, Rudd MK, Human structural variation: mechanisms of chromosome rearrangements. *Trends Genet.* 2015 Jul 22.
31. Lupski JR. Hotspots of homologous recombination in the human genome: not all homologous sequences are equal. *Genome Biol.* 2004;5(10):242.
32. Sasaki M, Lange J, Keeney S. Genome destabilization by homologous recombination in the germ line. *Nat Rev Mol Cell Biol.* 2010. 11(3):182-95.
33. Peng Z, Zhou W, Fu W, et al., Correlation between frequency of non-allelic homologous recombination and homology properties: evidence from homology-mediated CNV mutations in the human genome. *Hum Mol Genet.* 2015. 24(5):1225-33.
34. Koolen DA, Dupont J, de Leeuw N, et al., Two families with sibling recurrence of the 17q21.31 microdeletion syndrome due to low-grade mosaicism. *Eur J Hum Genet.* 2012. 20(7):729-33.
35. Ivanov EL, Sugawara N, Fishman-Lobell J, et al., Genetic requirements for the single-strand annealing pathway of double-strand break repair in *Saccharomyces cerevisiae*. *Genetics.* 1996. 142(3):693-704.
36. Youds JL, Boulton SJ. The choice in meiosis-defining the factors that influence crossover or non-crossover formation. *J Cell Sci.* 2011. 124(Pt 4):501-13.
37. McClintock B, Chromosome organization and genic expression. *Cold Spring Harb Symp Quant Biol.* 1951;16:13-47.
38. Burma S, Chen BP, Chen DJ Role o non-homologous end joining (NHEJ) in maintaining genomic integrity. *DNA Repair (Amst).* 2006. 5(9-10):1042-8.
39. Hastings PJ, Ira G, Lupski JR. A microhomology- mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet.* 2009 Jan;5(1):e1000327.
40. McClintock B. The Stability of broken ends of chromosomes in *Zea Mays*. *Genetics* 1941. 26(2):234–282.

41. Zakov S, Kinsella M, Bafna V. An algorithmic approach for breakage-fusion-bridge detection in tumor genomes. *Proc Natl Acad Sci U S A*. 2013. 110(14):5546-51.
42. Gisselsson D, Pettersson L, Höglund M, et al., Chromosomal breakage-fusion-bridge events cause genetic intratumor heterogeneity. *Proc Natl Acad Sci U S A*. 2000. 97(10):5357-62.
43. Lee JA, Inoue K, Cheung SW, et al., Role of genomic architecture in PLP1 duplication causing Pelizaeus-Merzbacher disease. *Hum Mol Genet*. 2006. 15(14):2250-65.
44. Hastings PJ, Lupski JR, Rosenberg SM, et al., Mechanisms of change in gene copy number. *Nat Rev Genet*. 2009. 10(8):551-64.
45. Cocquempot O, Brault V, Babinet C, et al., Fork stalling and template switching as a mechanism for polyalanine tract expansion affecting the DYX1C1 mutant of HOXD13, a new murine model of synpolydactyly. *Genetics*. 2009. 183(1):23-30.
46. Ottaviani D, LeCain M, Sheer D. The role of microhomology in genomic structural variation. *Trends Genet*. 2014. 30(3):85-94.
47. Malkova A, Ira G. Break-induced replication: functions and molecular mechanism. *Curr Opin Genet Dev*. 2013. 23(3):271-9.
48. Donnianni RA, Symington LS. Break-induced replication occurs by conservative DNA synthesis. *Proc Natl Acad Sci U S A*. 2013. 110(33):13475-80.
49. Aguilera A. Double-strand break repair: are Rad51/RecA-DNA joints barriers to DNA replication? *Trends Genet*. 2001. 17(6):318-21.
50. Haber JE. DNA recombination: the replication connection. *Trends Biochem Sci*. 1999. 24(7):271-5.
51. Hermetz KE, Newman S, Cooney NK et al., Large inverted duplications in the human genome form via fold-back mechanism. *PLoS Genet*. 2014. 10(1):e1004139.
52. Carvalho CM, Ramocki MB, Pehlivan D et al., Inverted genomic segments and complex triplication rearrangements are mediated by inverted repeats in the human genome. *Nat Genet*. 2011. 43(11):1074-81.
53. Beck CR, Carvalho CM, Baner L et al., Complex genomic rearrangements at the PLP1 locus include triplication and quadruplication. *PLoS Genet*. 2015. 11(3):e1005050.

54. Neill NJ, Ballif BC, Lamb AN et al., Recurrence, submicroscopic complexity, and potential clinical relevance of copy gains detected by array CGH that are shown to be unbalanced insertions by FISH. *Genome Res.* 2011.21(4):535-44.
55. Kang SH, Shaw C, Ou Z et al., Insertional translocation detected using FISH confirmation of array-comparative genomic hybridization (aCGH) results. *Am J Med Genet A.* 2010 152A(5):1111-26.
56. Zhang CZ, Spektor A, Cornils H, et al., Chromothripsis from DNA damage in micronuclei. *Nature.* 2015. 522(7555):179-84.
57. Zenz T, Mertens D, Döhner H, et al., Importance of genetics in chronic lymphocytic leukemia. *Blood Rev.* 2011. 25(3):131-7.
58. Pellestor F, Gatinois V, Puechberty J et al., Chromothripsis: potential origin in gametogenesis and preimplantation cell divisions. A review. *Fertil Steril.* 2014 Dec;102(6):1785-96
59. Holland AJ, Cleveland DW. Chromothripsis and cancer: mechanisms and consequences of localized, complex chromosomal rearrangements. *Nat Med.* 2012. 18(11):1630-8.
60. Stephens PJ, Greenman CD, Yang F et al., Massive genomic rearrangements acquired in a single catastrophic event during cancer development. *Cell.* 2011 Jan 7;144(1):27-40.
61. Magrangeas F, Avet-Loiseau H, Munshi NC, et al., Chromothripsis identifies a rare and aggressive entity among newly diagnosed multiple myeloma patients. *Blood.* 2011 Jul 21;118(3):675-8.
62. Molenaar JJ, Koster J, Zwiijnenburg DA et al., Sequencing of neuroblastoma identifies chromothripsis and defects in neuriteogenesis genes. *Nature.* 2012 Feb 22;483(7391):589-93.
63. Nagel S, Meyer C, Quentmeier H et al., Chromothripsis in Hodgkin lymphoma. *Genes Chromosomes Cancer.* 2013 Aug;52(8):741-7.
64. Rausch T, Jones DT, Zapatka M et al., Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations. *Cell.* 2012 Jan 20;148(1-2):59-71.
65. Wyatt AW, Collins CC, In brief: chromothripsis and cancer. *J Pathol.* 2013 Sep;231(1):1-3.

66. Malhotra A, Lindberg M, Faust GG et al., Breakpoint profiling of 64 cancer genomes reveals numerous complex rearrangements spawned by homology-independent mechanisms. *Genome Res.* 2013 May;23(5):762-76.
67. Kloosterman WP, Guryev V, van Roosmalen M et al., Chromothripsis as a mechanism driving complex de novo structural rearrangements in the germline. *Hum Mol Genet.* 2011 May 15;20(10):1916-24.
68. Korbel JO, Campbell PJ, Criteria for inference of chromothripsis in cancer genomes. *Cell.* 2013. 152(6):1226-36.
69. Kloosterman WP, Cuppen E, Chromothripsis in congenital disorders and cancer: similarities and differences. *Curr Opin Cell Biol.* 2013. 25(3):341-8.
70. Crasta K, Ganem NJ, Dagher R et al., DNA breaks and chromosome pulverization from error in mitosis. *Nature.* 2012.482(7383):53-8.
71. Maher CA, Wilson RK, Chromothripsis and human disease: piecing together the shattering process. *Cell.* 2012.148(1-2):29-32.
72. Zhang CZ, Leibowitz ML, Pellman D, Chromothripsis and beyond: rapid genome evolution from complex chromosomal rearrangements. *Genes Dev.* 2013. 27(23):2513-30.
73. Kloosterman WP, Tavakoli-Yaraki M, van Roosmalen ML et al., Constitutional chromothripsis rearrangements involve clustered double-stranded DNA breaks and non homologous repair mechanisms. *Cell Rep.* 2012. 1(6):648-55.
74. Chiang C, Jacobsen JC, Ernst C et al., Complex reorganization and predominant non homologous repair following chromosomal breakage in karyotypically balanced germline rearrangements and transgenic integration. *Nat Genet.* 2012. 44(4):390-7, S1.
75. Macera MJ, Sobrino A, Levy B, et al., Prenatal diagnosis of chromothripsis, with nine breaks characterized by karyotyping, FISH, microarray and whole-genome sequencing. *Prenat Diagn.*2015. 35(3):299-301.
76. Liu P, Erez A, Nagamani SC et al., Chromosome catastrophes involve replication mechanisms generating complex genomic rearrangements. *Cell.* 2011. 146(6):889-90
77. Shevell MI, Bejjani BA, Srour M, et al., Array comparative genomic hybridization in global developmental delay. *Am J Med Genet B Neuropsychiatr Genet.* 2008. 147B(7):1101-8

78. Shendure J, Ji H, Next generation DNA sequencing. *Nat Biotechnol.* 2008. (10):1135-45.
79. Voelkerding KV, Dames SA, Durtschi JD, Next-generation sequencing: from basic research to diagnostics. *Clin Chem.* 2009. 55(4):641-58.
80. Grada A, Weinbrecht K, Next-generation sequencing: methodology and application. *J Invest Dermatol.* 2013. 133(8):e11.
81. Zhang, C.-Z., Spektor, A, Cornils, H, et al., *Nature* 2015 522(7555):179-84.
82. Kinsella M, Patel A, Bafna V. The elusive evidence for chromothripsis. *Nucleic Acids Res.* 2014. 42(13):8231-42.
83. Terradas M, Martín M, Tusell L, et al., Genetic activities in micronuclei: is the DNA entrapped in micronuclei lost for the cell? *Mutat Res.* 2010. 705(1):60
84. Balmus G, et al. A high-throughput in vivo micronucleus assay for genome instability screening in mice. *Nat Protoc.* 2015 10(1):205–215.
85. Fenech M Cytokinesis-block micronucleus cytome assay. *Nat Protoc* 2007 2(5):1084–1104.
86. Fenech M, et al. Molecular mechanisms of micronucleus, nucleoplasmic bridge and nuclear bud formation in mammalian and human cells. *Mutagenesis* 2011 26(1):125–132.
87. van Paassen BW, van der Kooi AJ, van Spaendonck-Zwarts KY, et al., PMP22 related neuropathies: Charcot-Marie-Tooth disease type 1A and Hereditary Neuropathy with liability to Pressure Palsies. *Orphanet J Rare Dis.* 2014 Mar 19;9:38.
88. Magoulas PL, Liu P, Gelowani V, et al., Inherited dup(17)(p11.2p11.2): expanding the phenotype of the Potocki-Lupski syndrome. *Am J Med Genet A.* 2014 Feb;164A(2):500-4.
89. Lupski JR, Stankiewicz P. Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. *PLoS Genet.* 2005. 1(6):e49.
90. Ferner RE, Gutmann DH. Neurofibromatosis type 1 (NF1): diagnosis and management. *Handb Clin Neurol.* 2013;115:939-55.
91. Yap YS, McPherson JR, Ong CK et al., The NF1 gene revisited - from bench to bedside. *Oncotarget.* 2014 Aug 15;5(15):5873-92.
92. Bredrup C, Knappskog PM, Majewski J, et al., Congenital stromal dystrophy of the cornea caused by a mutation in the decorin gene. *Invest Ophthalmol Vis Sci.* 2005 46(2):420-6.

93. Jordan VK, Zaveri HP, Scott DA. 1p36 deletion syndrome: an update. *Appl Clin Genet*. 2015 Aug 27;8:189-200.
94. Windpassinger C, Kroisel PM, Wagner K, et al., The human gamma-aminobutyric acid A receptor delta (GABRD) gene: molecular characterisation and tissue-specific expression. *Gene*. 2002;292(1–2):25–31.
95. Mohler H. GABA(A) receptor diversity and pharmacology. *Cell Tissue Res*. 2006;326(2):505–516.
96. Rosenfeld JA, Crolla JA, Tomkins S, et al. Refinement of causative genes in monosomy 1p36 through clinical and molecular cytogenetic characterization of small interstitial deletions. *Am J Med Genet A*. 2010;152A(8):1951–1959.
97. Mihalek RM, Banerjee PK, Korpi ER, et al. Attenuated sensitivity to neuroactive steroids in gamma-aminobutyrate type A receptor delta subunit knockout mice. *Proc Natl Acad Sci U S A*. 1999;96(22):12905–12910.
98. Medrano EE. Repression of TGF-beta signaling by the oncogenic protein SKI in human melanomas: consequences for proliferation, survival, and metastasis. *Oncogene*. 2003;22(20):3123–3129.
99. Chen D, Xu W, Bales E, et al. SKI activates Wnt/beta-catenin signaling in human melanoma. *Cancer Res*. 2003;63(20):6626–6634.
100. Colmenares C, Heilstedt HA, Shaffer LG, et al. Loss of the SKI proto-oncogene in individuals affected with 1p36 deletion syndrome is predicted by strain-dependent defects in Ski^{-/-} mice. *Nat Genet*. 2002; 30(1):106–109.
101. Doyle AJ, Doyle JJ, Bessling SL, et al. Mutations in the TGF-beta repressor SKI cause Shprintzen-Goldberg syndrome with aortic aneurysm. *Nat Genet*. 2012;44(11):1249–1254.
102. Arndt AK, Schafer S, Drenckhahn JD, et al. Fine mapping of the 1p36 deletion syndrome identifies mutation of PRDM16 as a cause of cardiomyopathy. *Am J Hum Genet*. 2013;93(1):67–77.
103. Gamba BF, Richieri-Costa A, Costa S, et al., Chromothripsis with at least 12 breaks at 1p36.33-p35.3 in a boy with multiple congenital anomalies. *Mol Genet Genomics*. 2015 Dec;290(6):2213-6.