



ALMA MATER STUDIORUM
UNIVERSITA' DI BOLOGNA

Alma Mater Studiorum – Università di Bologna

Dottorato di Ricerca in Biologia e fisiologia cellulare

Ciclo XXI

Settore/i scientifico disciplinare di afferenza BIO/18 (Genetica)

***Genetica molecolare dell'autismo: studi di associazione ed
analisi di geni candidati***

Tesi presentata da:

Toma Claudio

Coordinatore Dottorato

Chiar.mo Prof. Rugolo Michela

Relatore

Chiar.mo Prof. Elena Maestrini

INDICE

CAPITOLO 1.....	pag.
Genetica di malattie complesse	1
1.1 Introduzione	2
1.2 Determinare le componenti genetiche di una malattia complessa	2
1.3 Strategie utilizzate per identificare le basi genetiche di malattie complesse	4
1.3.1 Geni candidati funzionali.....	5
1.3.2 Anomalie citogenetiche.....	5
1.3.3 Mappaggio genetico mediante analisi di linkage.....	6
1.3.4 Gli studi di associazione.....	8
1.3.4a Il progetto HapMap negli studi di associazione delle malattie complesse.....	9
1.3.5 Varianti strutturali nel genoma: Caratteristiche generali.....	11
1.3.5a Instabilità genomica e variabilità genetica.....	14
1.3.5b Formazione di CNVs tramite NAHR.....	14
1.3.5c CNVs e fenotipo.....	16
1.3.5d Metodi per l'identificazione di CNVs.....	18
CAPITOLO 2	
Autismo	21
2.1 Introduzione all'autismo	22
2.2 Aspetti clinici	22
2.3 Diagnosi	25
2.4 Anormalità neurologiche e fisiologiche correlate all'autismo	25
2.4.1 Anomalie cerebrali.....	25
2.4.2 Epilessia.....	26
2.4.3 Sistema dei neurotrasmettitori.....	26
2.5 Trattamenti farmacologici e terapia	27
2.6 Autismo: dalla sua scoperta al riconoscimento dell'importanza dei fattori genetici	27
2.7 Prevalenza dell'autismo	30
CAPITOLO 3	
Genetica molecolare dell'autismo	31
3.1 Contributi monogenici associati all'autismo	32
3.2 Studi di linkage per l'autismo estesi all'intero genoma	33

3.3 Geni candidati per l'autismo.....	39
---------------------------------------	----

CAPITOLO 4

<i>Il cromosoma X</i>	43
-----------------------------	----

4.1 <i>Il cromosoma X e l'autismo</i>	44
---	----

4.2 <i>Geni candidati per l'autismo che mappano sul cromosoma X</i>	45
---	----

4.2.1 <i>MECP2</i>	46
--------------------------	----

4.2.2 <i>NLGN3</i>	47
--------------------------	----

4.2.3 <i>FMRI</i>	49
-------------------------	----

4.3 <i>Inattivazione del cromosoma X (XCI)</i>	50
--	----

4.4 <i>Inattivazione sbilanciata del cromosoma X</i>).....	53
---	----

4.5 <i>XCI sbilanciata ed invecchiamento</i>).....	56
---	----

4.6 <i>Valutazione dello stato di inattivazione del cromosoma X</i>).....	57
--	----

CAPITOLO 5

<i>ASMT: gene candidato funzionale per l'autismo</i>	59
--	----

5.1 <i>La melatonina nell'autism</i>).....	60
---	----

5.2 <i>Studi di associazione e screening di mutazione del gene ASMT</i>	62
---	----

CAPITOLO 6

<i>Studi di associacizione ad alta densità ed analisi di CNVs in due loci di suscettibilità all'autismo: 2q24-32 e 7q21-32</i>	69
--	----

6.1 <i>I Loci di suscettibilità all'autismo AUTS1 e AUTS5</i>	70
---	----

6.2 <i>Genotyping su larga scala di SNPs tramite il sistema Illumina</i>	71
--	----

6.3 <i>Studi di associazione ad alta densità dei loci AUTS1 e AUTS5</i>	73
---	----

6.4 <i>Studio di replicazione in campioni di famiglie indipendenti e meta-analisi</i>	77
---	----

6.5 <i>Geni candidati per il disturbo autistico per i loci AUTS1 e AUTS5</i>	79
--	----

6.6 <i>Analisi di varianti strutturali nei loci AUTS1 e AUTS5</i>	80
---	----

CAPITOLO 7

<i>Scopo della ricerca</i>	83
----------------------------------	----

CAPITOLO 8

<i>Materiali e metodi</i>	87
---------------------------------	----

8.1 <i>Campioni di DNA</i>	88
----------------------------------	----

8.2. <i>Estrazione del DNA da sangue intero con kit Qiagen MAXI (per 5 ml di sangue)</i>	88
--	----

8.3. Disegno dei primers.....	89
8.4. Amplificazione tramite Polimerasi Chain Reaction (PCR)	89
8.5 Elettroforesi su gel di agarosio.....	95
8.6 Humara Assay.....	95
8.7 Reazione di sequenziamento ed analisi delle sequenze.....	96
8.8 Analisi di Linkage.....	97
8.9 Analisi di associazione: caso/controllo.....	97
8.10 qPCR mediante Real-Time.....	97
8.11 Analisi di varianti strutturali mediante MLPA(MultiplexLigationProbesAmplification) ...	98
8.12 Tecnologia Illumina: tecnologia array-based per il genotyping di SNPs su larga scala.....	100
8.13 Analisi di varianti strutturali mediante QMPSF(QuantitativeMultiplexPCRofofShortFluorescentFragments)	103

CAPITOLO 9

Risultati	105
9.1.1 Il Pattern di XCI.....	106
9.1.2 Profilo di XCI nel campione IMGSA.....	107
9.1.3 Profilo di XCI nel campione totale.....	109
9.1.4 Analisi di Linkage.....	112
9.1.5 Screening di mutazioni nei geni FMRI, Mecp2, NLGN3.....	113
9.1.6 Analisi di varianti strutturali nel gene Mecp2.....	113
9.2.1 Screening di mutazione del gene ASMT.....	117
9.2.2 Studio di associazione dei polimorfismi del promotore B del gene ASMT.....	120
9.3.1 Analisi dei trascritti alternativi del gene ZNF533.....	122
9.3.2 Analisi di mutazioni nel gene ZNF533 in un campione di individui autistici.....	125
9.3.3 Analisi di varianti strutturali nella regione IMMP2L-DOCK4 nella collezione di famiglie IMGSA.....	127

CAPITOLO 10

Discussione	131
10.1 Nessun gene soggetto a XCI è coinvolto nello spettro autistico.....	132
10.2 Il gene ASMT non sembra essere coinvolto nello spettro autistico.....	135
10.3 Studi di associazione ed analisi di varianti strutturali suggeriscono che la regione IMMP2L-DOCK4 del locus AUTS1 è coinvolta nella suscettibilità all'autismo.....	137

CAPITOLO 11

Bibliografia	141
---------------------------	-----

Abstract

Autism is a neurodevelopmental disorder characterized by impaired verbal communication, limited reciprocal social interaction, restricted interests and repetitive behaviours. Twin and family studies indicate a large genetic contribution to ASDs (Autism Spectrum Disorders). During my Ph.D. I have been involved in several projects in which I used different genetic approaches in order to identify susceptibility genes in autism on chromosomes 2, 7 and X:

High-density SNP association and CNV analysis of two Autism Susceptibility Loci

The International Molecular Genetic Study of Autism Consortium (IMGSAC) previously identified linkage loci on chromosomes 7 and 2, termed *AUTS1* and *AUTS5*, respectively. In this study, we evaluated the patterns of linkage disequilibrium (LD) and the distribution of haplotype blocks, utilising data from the HapMap project, across the two strongest peaks of linkage on chromosome 2 and 7. More than 3000 SNPs have been selected in each locus in all known genes, as well as SNPs in non-genic highly conserved sequences. All markers have been genotyped to perform a high-density association analysis and to explore copy number variation within these regions. The study sample consisted of 127 and 126 multiplex families, showing linkage to the *AUTS1* and *AUTS5* regions, respectively, and 188 gender-matched controls. Association and CNV analysis implicated several new genes, including *IMMP2L* and *DOCK4* on chromosome 7 and *ZNF533* and *NOSTRIN* on the chromosome 2. Particularly, my contribution to this project focused on the characterization of the best candidate gene in each locus:

On the *AUTS5* locus I carried out a transcript study of *ZNF533* in different human tissues to verify which isoforms and start exons were expressed. High transcript variability and a new exon, never described before, has been identified in this analysis. Furthermore, I selected 31 probands for the risk haplotype and performed a

mutation screen of all known exons in order to identify novel coding variants associated to autism.

On the *AUTS1* locus a duplication was detected in one multiplex family that was transmitted from father to an affected son. This duplication interrupts two genes: *IMMP2L* and *DOCK4* and warranted further analysis. Thus, I performed a screening of the cohort of IMGSAC collection (285 multiplex families), using a QMPSF assay (Quantitative Multiplex PCR of Short fluorescent Fragments) to analyse if CNVs in this genic region segregate with autism phenotype and compare their frequency with a sample of 475 UK controls.

Evidence for a role of *DOCK4* in autism susceptibility was supported by independent replication of association at rs2217262 and the finding of a deletion segregating in a sib-pair family.

Analysis of X chromosome inactivation

Skewed X chromosome inactivation (XCI) is observed in females carrying gene mutations involved in several X-linked syndromes. We aimed to estimate the role of X-linked genes in ASD susceptibility by ascertaining the XCI pattern in a sample of 543 informative mothers of children with ASD and in a sample of 164 affected girls. The study sample included families from different european consortia. I analysed the XCI inactivation pattern in a sample of italian mothers from singletons families with ASD and also a control groups (144 adult females and 40 young females). We observed no significant excess of skewed XCI in families with ASD. Interestingly, two mothers and one girl carrying known mutations in X-linked genes (*NLGN3*, *ATRX*, *MECP2*) showed highly skewed XCI, suggesting that ascertainment of XCI could reveal families with X-linked mutations. Linkage analysis was carried out in the subgroup of multiplex families with skewed XCI ($\geq 80:20$) and a modest increased allele sharing was obtained in the Xq27-Xq28 region, with a peak Z score of 1.75 close to rs719489. In this region *FMRI* and *MECP2* have been associated in some cases with autism and therefore represent candidates for the disorder. I performed a

mutation screen of *MECP2* in 33 unrelated probands from IMGSAC and Italian families, showing XCI skewness. Recently, Xq28 duplications including *MECP2*, have been identified in families with MR, with asymptomatic carrier females showing extreme (>85%) skewing of XCI. For these reasons I used the sample of probands from X-skewed families to perform CNV analysis by Real-time quantitative PCR. No duplications have been found in our sample. I have also confirmed all data using as an alternative method the MLPA assay (Multiplex Ligation dependent Probe Amplification).

ASMT as functional candidate gene for autism

Recently, a possible involvement of the *acetylserotonin O-methyltransferase (ASMT)* gene in susceptibility to ASDs has been reported: mutation screening of the *ASMT* gene in 250 individuals from the PARIS collection revealed several rare variants with a likely functional role; Moreover, significant association was reported for two SNPs (rs4446909 and rs5989681) located in one of the two alternative promoters of the gene. To further investigate these findings, I carried out a replication study using a sample of 263 affected individuals from the IMGSAC collection and 390 control individuals. Several rare mutations were identified, including the splice site mutation IVS5+2T>C and the L326F substitution previously reported by Melke et al (2007), but the same rare variants have been found also in control individuals in our study. Interestingly, a new R319X stop mutation was found in a single autism proband of Italian origin and is absent from the entire control sample. Furthermore, no replication has been found in our case-control study typing the SNPs on the *ASMT* promoter B.

1. Genetica di malattie complesse

1.1 Introduzione

Durante l'ultima decade, tecniche di clonaggio posizionale hanno permesso l'identificazione di centinaia di geni ad eredità mendeliana, in cui il fenotipo della malattia dipende dall'azione di un unico gene. Gli approcci tradizionali utilizzati per malattie monogeniche non sono sufficienti per individuare i geni che sono alla base delle cosiddette malattie complesse, in cui il fenotipo, a volte eterogeneo, dipende dall'azione concertata di un certo numero di loci, cui può sommarsi l'azione di fattori esterni o ambientali.

La difficoltà nello studio di queste patologie complesse consiste proprio nel correlare le varie componenti fenotipiche ad uno specifico fattore genetico e/o ambientale (Elston, 1995).

Le malattie complesse hanno spesso una elevata incidenza nella popolazione, per questo la sfida degli ultimi anni nel campo della genetica medica è lo sviluppo di nuove tecniche o approcci che permettano di individuarne i fattori genetici coinvolti.

Lo studio di malattie complesse comporta la necessità di disporre di ampie casistiche (famiglie o individui affetti) sia per ottenere un sufficiente potere statistico, sia per avere un campione rappresentativo dell'eterogeneità della malattia a livello di fattori genetici e ambientali (Lander 1994). Infatti, i fattori genetici implicati in malattie complesse hanno individualmente un modesto rischio relativo, che è difficilmente quantificabile *a-priori*.

1.2 Determinare le componenti genetiche di una malattia complessa

L'ereditarietà delle malattie complesse può essere descritto dal cosiddetto “modello multifattoriale a soglia”. Il fenotipo di una malattia complessa è il risultato dell'effetto di un numero variabile di geni di suscettibilità e di fattori ambientali. L'insieme di questi elementi costituisce la predisposizione individuale, che rappresenta una variabile continua con distribuzione gaussiana nella popolazione generale. Pertanto, ogni individuo presenta un certo grado di predisposizione, mentre solo una piccola

percentuale di individui si colloca agli estremi della curva, rispettivamente con una predisposizione molto bassa o molto alta di sviluppare la malattia. Gli individui la cui predisposizione supera un determinato livello, definito come "effetto soglia", manifestano la malattia (figura 1).

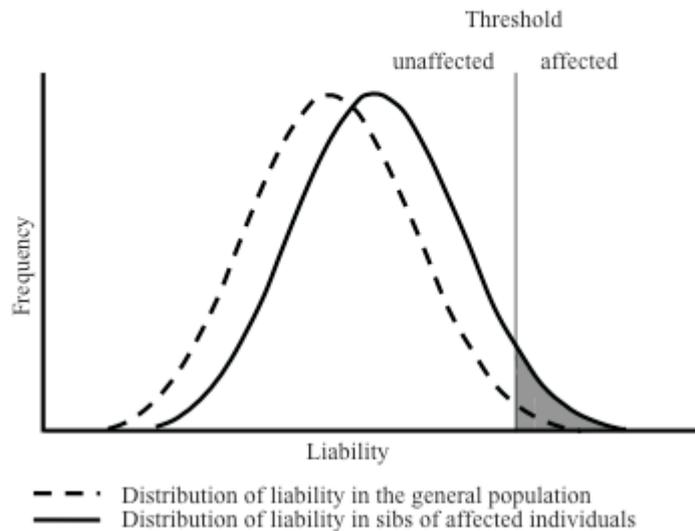


Figura 1: Modello multifattoriale a soglia: I fattori che predispongono all'insorgere della malattia sono distribuiti normalmente nella popolazione. Gli individui la cui predisposizione supera questa soglia manifestano la malattia. I fratelli di individui affetti mostrano un rischio maggiore rispetto alla media della popolazione e una maggiore proporzione di questi ultimi presenta una predisposizione che supera la soglia.

Gli studi di ricorrenza familiare sono spesso utilizzati per stimare l'importanza di fattori genetici nell'insorgenza di una malattia. Il grado di aggregazione familiare è misurato come rischio relativo (λ_r), ossia la probabilità di insorgenza della malattia in parenti di individui affetti rispetto alla popolazione generale.

$$\lambda_r = \frac{\text{frequenza della malattia in parenti di probandi affetti } (Kr)}{\text{frequenza della malattia nella popolazione generale } (K)}$$

La presenza di un elevato λ_r tuttavia non è sufficiente a stabilire la base genetica di una malattia, poichè individui tra loro imparentati condividono gli stessi geni, ma anche gli stessi fattori ambientali. Per questo motivo i classici studi di concordanza tra gemelli monozigotici (MZ) e dizigotici (DZ) sono utili a stabilire il contributo dei fattori genetici ed ambientali in una malattia. Maggiore è la differenza di concordanza tra gemelli monozigotici e dizigotici, maggiore sarà il peso della componente genetica, mentre una concordanza uguale indica una forte componente ambientale. I dati di questi studi, ove possibile, vengono utilizzati per effettuare una stima dell'ereditabilità e del numero di geni coinvolti nella malattia.

1.3 Strategie utilizzate per identificare le basi genetiche di malattie complesse.

Per l'identificazione di geni di suscettibilità di una patologia, sia essa ad eredità mendeliana o complessa, si utilizzano differenti strategie, che possono essere utilizzate in maniera combinata. Gli approcci che sono stati usati finora per l'identificazione dei geni di suscettibilità di malattie oligo- o poligeniche prevedono:

- 1) *Ipotesi del gene candidato funzionale*: si ipotizza il coinvolgimento di un gene in base al ruolo che ricopre nel processo fisiopatologico, se è noto.
- 2) *Studi sulle aberrazioni cromosomiche*: le anomalie citogenetiche possono risultare cruciali nell'individuare i geni responsabili di una malattia
- 3) *Studi di linkage*: permettono di localizzare la posizione cromosomica del locus malattia rispetto a quella di marcatori polimorfici, osservando la cosegregazione di loci in famiglie con almeno due figli affetti.
- 4) *Studi di associazione allelica*: mirano ad individuare associazione tra un particolare allele o aplotipo e la malattia a livello di popolazione.

5) *Studi sulle varianti strutturali*: recentemente è stata individuata la presenza nel genoma umano di numerosi varianti strutturali submicroscopiche, ossia tratti cromosomici di dimensioni comprese fra 1Kb e 3Mb, che variano in termini di numeri di copie. Tali varianti strutturali denominate *Copy Number Variants (CNVs)* in alcuni casi possono essere associate a malattie.

1.3.1 Geni candidati funzionali

I Loci di suscettibilità di una malattia possono essere identificati in base all'indicazione della funzione nota del gene e/o al profilo di espressione. Quando è conosciuto il ruolo di una data proteina nel processo fisiopatologico di una malattia è possibile ipotizzare il suo coinvolgimento diretto. Quindi è possibile risalire al gene ed analizzare possibili eventi mutazionali che interessano la sequenza genica e formulare quindi un'ipotesi sull'eziologia della malattia. Gli studi condotti su modelli animali sono di particolare interesse in questo campo. Per numerose patologie il modello murino è utilizzato in quanto mimica le caratteristiche cliniche della patologia in esame. Topi *Knock in* sono ad esempio utilizzati per analizzare l'acquisizione di una nuova funzione da parte di geni mutati (*gain of function*), mentre topi *knock out*, sono utili quando alla base di una patologia vi è l'abolizione della funzione di un gene (*loss of function*).

1.3.2 Anomalie citogenetiche

L'analisi di anomalie cromosomiche (delezioni, inserzioni, traslocazioni o inversioni) rappresenta un approccio alternativo e complementare che si è dimostrato molto utile nel mappaggio e nell'identificazione di geni malattia, in quanto può aiutare ad individuare direttamente il gene malattia, qualora il punto di rottura cada in corrispondenza del gene stesso, o a definire intervalli cromosomici critici per la localizzazione di un locus di suscettibilità. Le anomalie cromosomiche determinano cambiamenti genetici funzionali che si possono verificare in vario modo:

- “*effetto dosaggio*”: come conseguenza della variazione del numero di copie di un gene in seguito a delezioni o duplicazioni.
- “*effetto di posizione*”: come conseguenza di un riarrangiamento cromosomico uno o più geni possono risultare separati da sequenze regolatrici.
- i punti di rottura direttamente associati ad un riarrangiamento cromosomico possono distruggere il gene.
- delezioni in uno dei due alleli possono evidenziare una mutazione puntiforme localizzata nella regione corrispondente sul cromosoma omologo non deletato.

La sovrapposizione di regioni che presentano riarrangiamenti cromosomici con regioni individuate dai vari studi di *linkage* da una parte rappresenta una conferma di *linkage*, e dall'altra permette di localizzare in maniera più precisa la regione cromosomica di interesse.

1.3.3 Mappaggio genetico mediante analisi di linkage

L'analisi di *linkage* classica, di tipo **parametrico** (nota anche come **model-based analysis**), rappresenta una metodologia statistica classica dell'analisi genetica e si basa sull'identificazione di marcatori che segregano insieme alla malattia all'interno di famiglie. Il *linkage* si deve alla stretta vicinanza di due geni o marcatori lungo un cromosoma e si verifica qualora la cosegregazione tra un marcatore (per lo più microsatellite) ed un gene che influenza il fenotipo di interesse si manifesta con una frequenza maggiore di quella attesa per il puro effetto del caso. La base biologica è il concetto di ricombinazione o *crossing-over*, grazie al quale durante la prima fase meiotica avviene lo scambio fisico di materiale genetico tra i cromosomi omologhi paterno e materno. In pratica ciò che si osserva è il grado di indipendenza con cui avviene la trasmissione di un marcatore rispetto ad un altro, ciò riflette il grado di ricombinazione fra tali marcatori. Il parametro che esprime la frazione di ricombinazione è θ , intesa come la percentuale dei gameti che ci si attende siano ricombinanti.

Un *genome wide screen* consiste nel *genotyping* di 300-400 marcatori altamente polimorfici (microsatelliti), distribuiti in maniera uniforme ad una distanza di circa 10 cM l'uno dall'altro, in un elevato numero di famiglie contenenti coppie di fratelli affetti (*sib-pairs*) o altri tipi di parenti affetti (*relative pairs*). I dati di *linkage*, ottenuti in seguito all'analisi di tutti i marcatori, sono utilizzati per effettuare un conteggio diretto dei ricombinanti per ottenere il valore di θ . In alternativa si applica il *metodo del LOD Score*. Il LOD Score è il logaritmo in base 10 del rapporto tra la probabilità che due loci siano associati e quella che non lo siano:

$$\text{LOD score } Z(\theta) = \sum \log_{10} = \frac{\text{L (family se } \theta=X)}{\text{L (family se } \theta=1/2)}$$

La soglia di significatività di *linkage* è fissata per LODscore ≥ 3 (Morton, 1955). Se due loci si trovano lontani tra loro, allora i LOD score saranno per lo più negativi.

I metodi di *linkage* parametrici sono generalmente applicati ai tratti mendeliani e necessitano della preventiva conoscenza del modello di ereditarietà. Nelle malattie geneticamente complesse non si conosce il modello di ereditarietà, ed è raro avere a disposizione alberi genealogici estesi. Vengono pertanto utilizzati metodi di analisi di *linkage* non parametrici ("*model-free linkage methods*") nei quali non è necessario specificare il modello d'ereditarietà (Weeks DE, 1995; Lander ES, 1994). Tali metodi sono basati sull'analisi dello stato di condivisione degli alleli (*allele sharing*) fra due o più individui malati in famiglie nucleari. Vengono studiati gli alleli identici per discendenza (*IBD, identical by descent*), ossia quegli alleli che sono identici negli individui in cui è presente la patologia. Fra le metodologie di analisi di *linkage* non parametriche, l'analisi di coppie di fratelli affetti (*affected sib-pair method, ASP*) è quella più frequentemente utilizzata, e si può estendere anche a coppie di parenti che non siano fratelli (*affected Relative-Pair method*).

La presenza di un aumento significativo della condivisione di alleli nei fratelli affetti viene interpretata come prova di linkage genetico. In generale uno studio iniziale di *linkage* di una malattia complessa si conclude con l'individuazione di regioni candidate, generalmente molto ampie, all'interno delle quali dovrebbero risiedere i loci responsabili della malattia.

1.3.4 Gli studi di associazione

Gli studi di *linkage model-free* hanno un potere risolutivo relativamente basso, quindi, quando si inizia lo studio di una malattia complessa, di cui non si conosce l'eziologia e i geni coinvolti, l'approccio generalmente seguito prevede, inizialmente, un'analisi dell'intero genoma utilizzando metodi di *linkage* non-parametrici e, successivamente, un'analisi sempre più fine delle regioni di *linkage* individuate utilizzando ulteriori marcatori interni alle regioni stesse.

A tal fine, gli studi di associazione rappresentano un approccio particolarmente utile in quanto mirano all'individuazione, all'interno di una popolazione, di varianti che si trovano più frequentemente associate alla patologia. Tale associazione può essere dovuta a diversi fattori quali: un effetto diretto della variante genica sul fenotipo malattia, il *linkage disequilibrium* (LD), o associazioni spurie dovute alla stratificazione della popolazione. Una variante allelica può avere un effetto diretto sul fenotipo in quanto può alterare la sequenza aminoacidica di una proteina coinvolta nel processo fisiopatologico, o alterare la regione regolatoria del gene determinando anomalie a livello di espressione o, infine, interferire con la regolazione dello *splicing* e la stabilità del messaggero. Quando si ha una associazione non casuale tra la variante allelica analizzata e la vera variante eziologica siamo in presenza di *Linkage Disequilibrium*, in quanto i due loci sono vicini sul cromosoma e vengono ereditati insieme come unico blocco.

Esistono due tipi di studi di associazione:

- lo **studio caso-controllo**, condotto a livello di popolazione, che confronta la frequenza di un determinato allele tra un campione di individui affetti (casi) e un campione di individui non affetti (controlli) tra loro non imparentati. Il problema maggiore in questo tipo di studio è rappresentato dalla stratificazione della popolazione e, quindi, dalla presenza di sottopopolazioni genetiche, con diverse distribuzioni alleliche (Slager SL et al., 2001).
- il ***Disequilibrium Transmission Test*** (TDT), Questo test è condotto su famiglie composte da entrambi i genitori con uno o più figli affetti, e calcola se esiste differenza nella trasmissione di un allele rispetto ad un altro nella progenie affetta da genitori eterozigoti. (Spielman RS et al., 1993). Se la trasmissione dell'allele avviene con una frequenza maggiore del 50% si è in presenza di associazione allelica e *linkage* tra marcatore e malattia. A differenza del caso-controllo il TDT non è sensibile alla stratificazione nella popolazione.

1.3.4a Il progetto HapMap negli studi di associazione delle malattie complesse

Il progetto internazionale *HapMap* ha ampliato le prospettive degli studi di associazione in quanto ha portato alla formazione di un *database* delle varianti comuni del genoma umano. Grazie alle informazioni contenute in questo database è oggi possibile intraprendere degli studi di associazione molto più efficienti ed estesi all'intero genoma, un approccio praticamente irrealizzabile fino a pochi anni fa. Il progetto HapMap è nato nel 2002 con lo scopo di mappare la struttura di *linkage disequilibrium* lungo tutto il genoma umano. (The International HapMap Consortium: 2003, 2005).

Secondo alcune stime, nella popolazione mondiale, si trovano circa dieci milioni di SNPs per i quali entrambi gli alleli sono presenti con una frequenza superiore all'1% . Lo specifico insieme di marcatori che si trovano sullo stesso cromosoma o nella medesima regione sub-cromosomiale è detto *aplotipo*. Il *linkage disequilibrium* (LD) deriva dal fatto che alleli vicini vengono più facilmente co-ereditati in blocchi di

aplotipi risultando quindi associati all'interno di una popolazione. Siccome la probabilità di ricombinazione fra due SNPs aumenta con l'aumentare della distanza fisica fra i due, di norma il grado di associazione fra SNPs diminuisce progressivamente con la distanza (Reich et al. 2001). Uno dei risultati più interessanti scaturiti dal progetto HapMap è che il genoma umano risulta essere costituito da regioni cromosomiche, dette "blocchi di aplotipi", all'interno delle quali gli SNPs hanno un alto grado di associazione. Questo significa che all'interno di un "blocco" la diversità di una popolazione umana è rappresentata solo da pochi aplotipi (Patil et al. 2001). I "blocchi di aplotipi" sono separati da regioni con un basso livello di LD ed un'alta diversità, che corrispondono probabilmente a "hotspot" di ricombinazione. La conoscenza della struttura di LD e della struttura in blocchi del genoma ha dei risvolti pratici molto importanti (Johnson et al. 2001). Significa che analizzando solo alcuni SNPs opportunamente scelti (*tag SNPs*) è possibile conoscere le altre varianti comuni presenti nelle vicinanze in modo tale da rappresentare tutti gli aplotipi presenti nella regione cromosomica (figura 2).

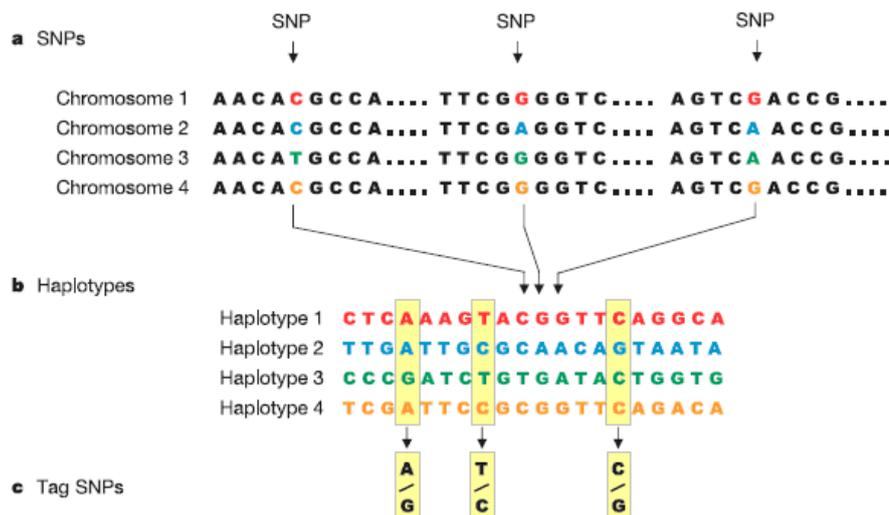


Figura 2: a) Sequenze genomiche caratterizzate dalla presenza di SNPs; b) Successioni di SNPs (aplotipi) che rappresentano una data sequenza genomica; c) Tag SNPs che rappresentano gli aplotipi di una data sequenza genomica.

1.3.5 Varianti strutturali nel genoma: Caratteristiche generali

Finora le differenze osservabili a livello genomico riguardavano rari casi di variazione del numero e della struttura dei cromosomi, spesso associati a gravi patologie (aneuploidie, riarrangiamenti, eteromorfismi, siti fragili, ecc.) osservabili tramite l'utilizzo del microscopio o con esami del cariotipo (Warburton D., 1991; Jacobs PA et al., 1992). Questo tipo di anomalie coinvolgono segmenti di DNA di lunghezza superiore a diverse Mb, abbastanza grandi da poter essere individuate a livello citogenetico. Attualmente, in seguito al diffondersi di tecnologie specifiche per gli studi a livello genomico è stato possibile individuare la presenza di un numero inaspettatamente grande di varianti genomiche di dimensioni intermedie, comprese fra 1Kb e 3Mb, complessivamente denominate *Copy-Number Variants* (CNVs) o *varianti strutturali* (Tabella 1) e (Figura 3).

Copy-number variant (CNV): Segmento di DNA di lunghezza superiore ad 1Kb presente in un numero variabile di copie rispetto ad un genoma di riferimento. Diverse classi di CNVs comprendono inserzioni, delezioni, duplicazioni. Delle CNVs fanno parte anche le <i>large-scale copy-number variants</i> che riguardano segmenti di DNA superiori a 50Kb.
Copy-number polymorphism (CNP): Un CNV presente nella popolazione con una frequenza maggiore dell'1%.
Segmental duplication o low-copy repeat (LCR): Segmento di DNA di lunghezza superiore ad 1Kb presente in due o più copie per genoma aploide tale per cui le copie hanno un'identità di sequenza del 90%. Le LCRs possono anche essere CNVs (fig.x).
Inversioni: Un segmento di DNA con orientamento invertito rispetto al resto del cromosoma.
Traslocazioni: Un cambiamento di posizione all'interno del genoma di un segmento di DNA senza perdita o guadagno di materiale genetico. Le traslocazioni possono essere <i>intra-</i> o <i>inter-</i> cromosomiche.
Segmental uniparental disomy: Si ha quando una porzione di cromosoma deriva da una coppia di cromosomi omologhi ereditata da un singolo genitore.

Tabella 1: Definizioni relative alle diverse tipologie di varianti strutturali.

Attualmente circa il 20% del genoma umano è stato visto associato a varianti strutturali che sono state annotate (Copper GM, 2007), benchè gli arrays finora utilizzati mancano di alta risoluzione e probabilmente una piccola percentuale rappresentano falsi positivi (Sharp AJ, 2009). Data la frequenza elevata, questa nuova classe di polimorfismi contribuisce in modo rilevante alla diversità umana, insieme ad altri tipi di polimorfismi come gli SNPs (Sharp et al. 2005), che costituiscono la classe di varianti più comuni nel genoma. Le varianti strutturali possono contenere interi geni o intere regioni regolatorie con potenziali effetti a livello genico quale il dosaggio di un gene o la sua regolazione (Sharp JA, 2009; Tuzun AJ et al., 2005; Korbel JO, 2007). Sebbene, possiamo considerarle come una nuova classe di polimorfismi che in alcune regioni del genoma non hanno alcuna conseguenza fenotipica (Sharp AJ et al. 2005; Sebat J et al., 2004; Iafrate AJ et al., 2004; Tuzun AJ et al., 2005)

Tuttavia si è ancora lontani dall'interpretazione esaustiva dell'effetto di determinate varianti sul fenotipo e della loro frequenza nella popolazione generale (Sharp AJ, 2009). Per questo motivo sono sorti dei consorzi come Il Wellcome Trust Case Control Consortium con l'obiettivo di analizzare un numero elevato di individui per meglio interpretare le varianti incontrate, ed in alcuni casi stabilire se siano varianti rare o varianti polimorfiche con una bassa frequenza nella popolazione (Itsara A., 2009). Negli ultimi anni sono stati condotti numerosi studi allo scopo di chiarire il significato evolutivo di questo tipo di varianti, il meccanismo in seguito al quale si formano, in che modo contribuiscono alla variabilità genetica e la loro possibile associazione con malattie ereditarie e complesse ad eziologia ignota quali, ad esempio, l'autismo (Buckley PG et al., 2005).

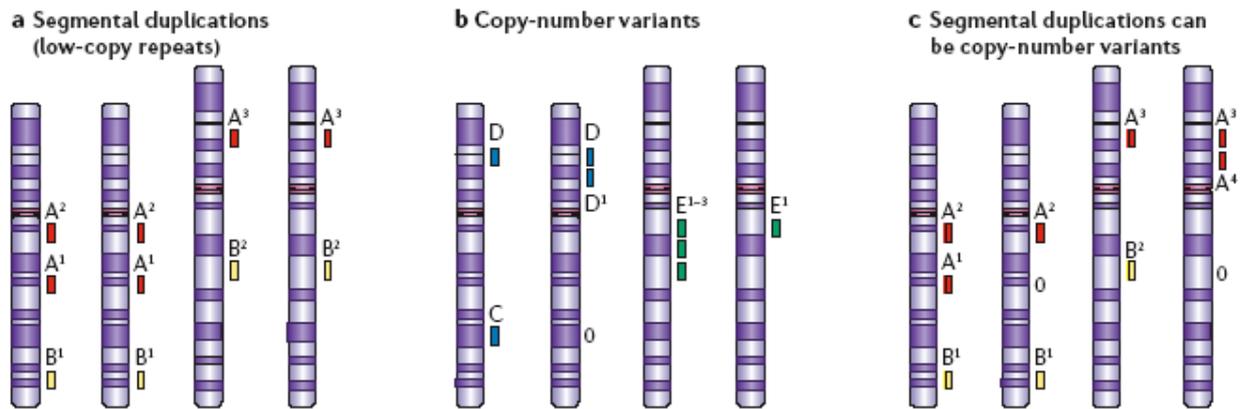


Figura 3: a) Segmental duplication: copie di un segmento genomico con omologia minima del 90% intersperse nel genoma aploide; b) Copy-number variants: quando un segmento genomico può essere presente un numero variabile di volte rispetto ad un genoma di riferimento; c) Esempio in cui alcune segmental duplication possono essere copynumber variant.

1.3.5a Instabilità genomica e variabilità genetica

Diversi studi hanno evidenziato la presenza nel genoma di alcuni segmenti particolarmente suscettibili a riarrangiamenti cromosomici che causano un alto grado di instabilità genomica a livello di specifiche regioni cromosomiche. Si ipotizza che ciò sia uno dei meccanismi responsabili dell'introduzione di variabilità genetica all'interno della specie. I meccanismi che generano tali riarrangiamenti e le conseguenze fenotipiche che le determinano, sono stati oggetto di numerosi studi negli ultimi anni. Analogamente, si ipotizza che la presenza nel genoma di varianti strutturali comuni quali le CNVs, sia responsabile di parte della comune variabilità fenotipica degli individui. Per molte malattie genomiche (CMT1A, *Charcot-Marie-Tooth1A*; HNPP, *Hereditary Neuropathy with liability to pressure palsies*; SMS, *Smith-Magenis syndrome*), così come per le varianti strutturali comuni, si è osservato che le alterazioni a livello genomico che ne sono la causa possono essere ereditate dai genitori oppure insorgere *de novo* sporadicamente.

1.3.5b Formazione di CNVs tramite NAHR

Spesso le CNVs sono delimitate da punti di rottura che mappano in corrispondenza di specifiche *segmental duplications* (o low copy repeats, LCRs) fiancheggianti la delezione o la duplicazione in questione. Si ipotizza che, nel 99% dei casi, tali varianti strutturali siano il risultato di un evento di ricombinazione omologa non allelica (*non-allelic homologous recombination*, NAHR) causato dall'appaiamento scorretto tra sequenze ripetute che le fianleggiano. La NAHR avviene esattamente con lo stesso meccanismo della ricombinazione omologa (appaiamento filamenti omologhi, formazione del chiasma, risoluzione del crossing-over) con l'eccezione che, nella NAHR, l'alta omologia di sequenza dei segmenti cromosomici che costituiscono ad esempio le LCRs, fa sì che il meccanismo cellulare di controllo della ricombinazione non sia in grado di riconoscere l'appaiamento di segmenti non perfettamente identici (Lupski JR et al., 2000).

L'appaiamento può avvenire fra LCRs appartenenti a cromosomi omologhi (*intercromosomico*), fra cromatidi fratelli (*intracromosomico*), all'interno di un singolo cromatidio (*intracromatide*) (Lupski JR et al., 2002). In base al meccanismo di appaiamento, dell'orientamento e della complessità delle LCRs, la NAHR può risolversi in delezioni, duplicazioni, inversioni o altri riarrangiamenti più complessi che possono anche coinvolgere cromosomi non omologhi (Figura 4).

Solo in pochi casi sono stati osservati riarrangiamenti originatisi da eventi di ricombinazione non omologa (*Non-Homologous End-Joining*, NHEJ), un tipo di ricombinazione che non richiede un grado minimo di omologia di sequenza e che si basa sulla ligazione fra due segmenti di DNA in cui si è verificata una rottura del doppio filamento (Gu W, 2008; Lupski JR, 2005).

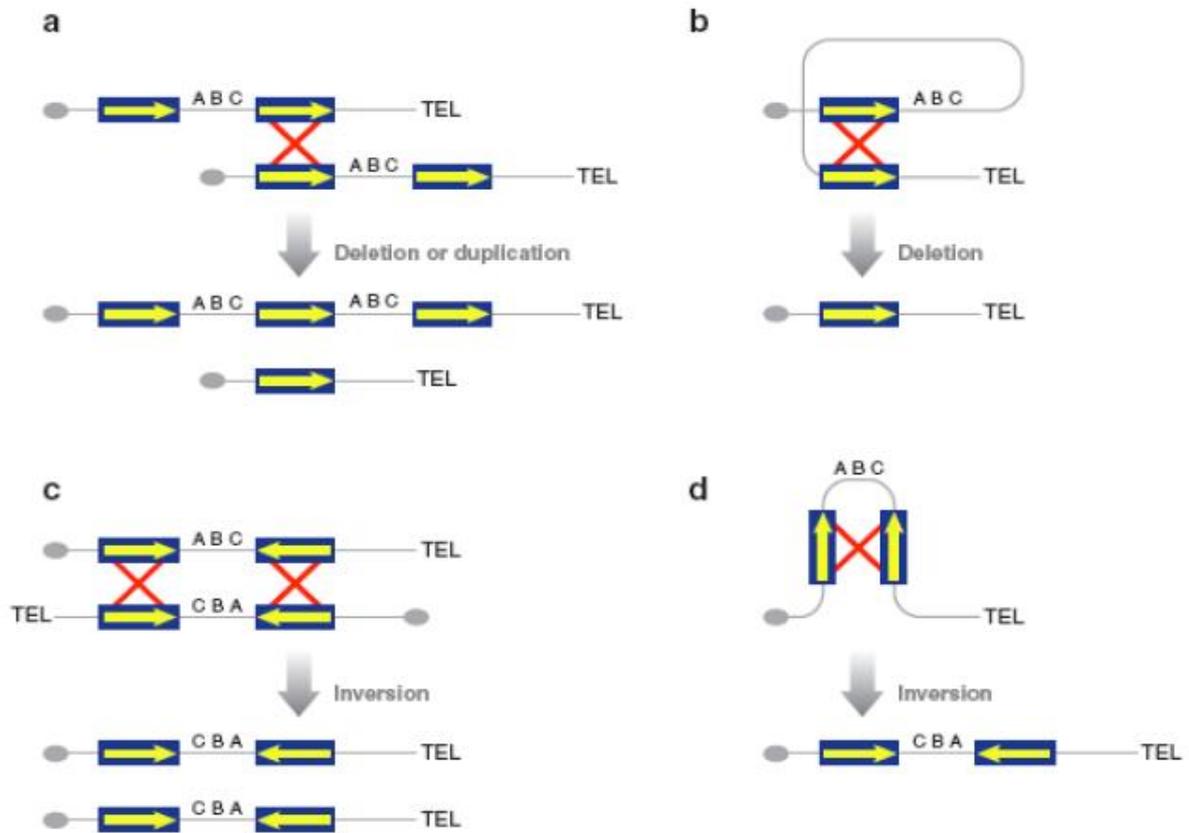


Figura 4: Esempi di ricombinazione omologa non allelica : Le *Segmental duplications* fungono da substrato per la NAHR. (*a e b*) NAHR intercromosomale, intracromosomale, o intracromatidica tra sequenze ripetute orientate nello stesso senso causa delezioni e/o duplicazioni. (*c e d*) NAHR intercromosomale, intracromosomale o intracromatidica tra sequenze ripetute orientate in senso opposto causa inversioni. Le sequenze ripetute sono rappresentate dalle box blu con l'orientamento indicato dalla freccia gialla. L'evento di ricombinazione è rappresentato da una croce rossa.

Oltre che fra LCRs, la NAHR può verificarsi in presenza di sequenze ripetute sub-telomeriche. Si è osservato che tali riarrangiamenti sono associati in modo significativo ad alcuni casi di ritardo mentale e/o a sindromi di malformazione (Flint J. et al., 1995; Knight et al., 1999). Altri studi hanno invece osservato il verificarsi della NAHR in corrispondenza di sequenze ripetute e intersperse che, complessivamente, rappresentano circa il 30% del genoma. In particolare, sono stati osservati eventi di NAHR fra sequenze *Alu* (classe principale di *Short Interspersed*

Elements, SINEs) e meno frequentemente fra sequenze L1 (classe principale di *Long Interspersed Elements*, LINEs) (Kolomietz E. et al., 2002; Burwinkel B. et al., 1998). E' probabile che molte varianti strutturali non rappresentino eventi rari, ma bensì il risultato di una predisposizione al riarrangiamento cromosomico determinata dall'architettura stessa del genoma che agisce come catalizzatore dell'instabilità cromosomica.

1.3.5c CNVs e fenotipo

Gli studi effettuati sulle varianti strutturali nel genoma hanno mostrato che individui sani differiscono per il numero di copie di ampi e numerosi segmenti cromosomici (De Stahl TD, 2008; Conrad D et al. , 2006; Hinds DA et al, 2006; Mc Carrol S et al., 2006). Ad oggi non è chiaro in che misura tali riarrangiamenti cromosomici possano influire su tratti fenotipici complessi o essere causa di suscettibilità per una patologia o essere semplici varianti polimorfiche.

In generale, le varianti strutturali possono essere causa di variabilità o dell'insorgere di una patologia, secondo molteplici meccanismi (Figura 5). Ad esempio, una CNV può influenzare direttamente il dosaggio di un gene oppure, indirettamente, alterarne l'espressione genica per effetto di posizione in quanto si trova in una regione regolatoria o potrebbe predisporre il locus in cui risiede ad ulteriori riarrangiamenti cromosomici.

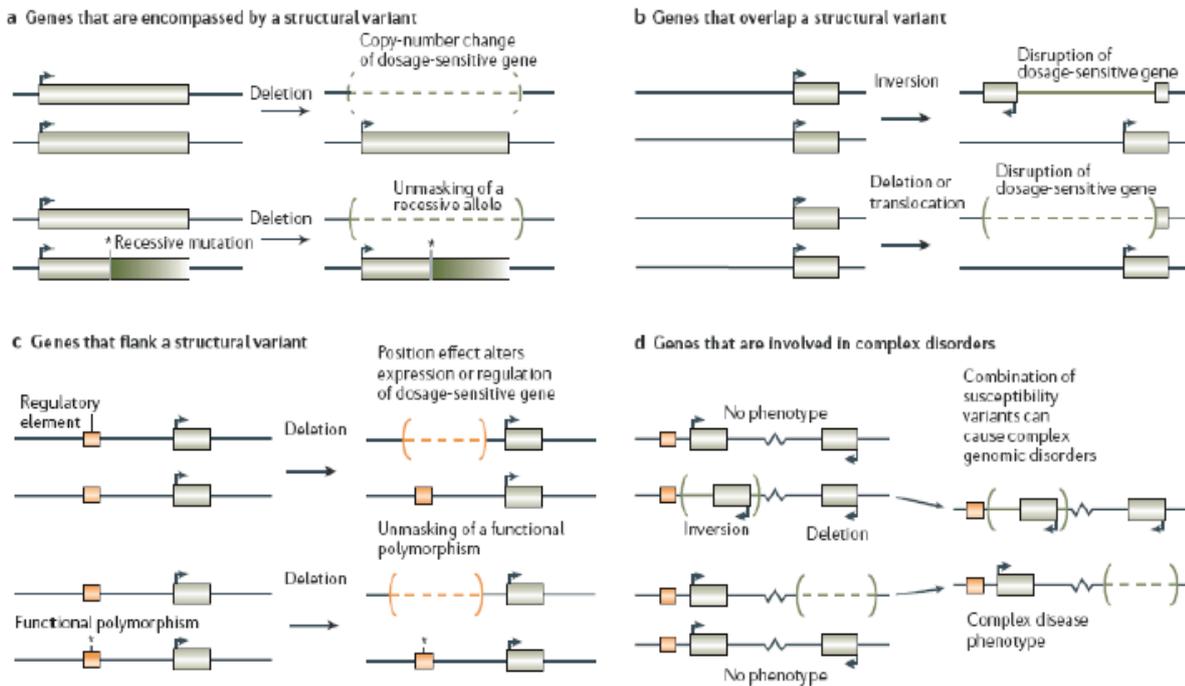


Figura 5: Casistica di effetti che la presenza di una CNV può produrre a livello fenotipico.

Le varianti strutturali rappresentano una buona parte della variabilità genetica umana, infatti si calcola che questa classe di polimorfismi è in grado di catturare circa il 20% del totale della variabilità genetica umana (Stranger BA et al., 2007). Una interessante sfida dei prossimi anni sarà quindi associare a questa ampia fonte di fattori genetici tratti fenotipici, anche se è di difficile interpretazione il contributo specifico al fenotipo malattia delle varianti strutturali, in quanto agiscono in combinazione con altri fattori genetici e/o ambientali, e distinguere tra quelle responsabili di tratti legati a malattie e quelle benigne e frequenti nella popolazione (Buchanan JA, 2008).

1.3.5d Metodi per l'identificazione di CNVs

Lo sviluppo di nuove strategie sperimentali e computazionali ha permesso di analizzare ad una risoluzione sempre maggiore le varianti strutturali umane (Speicher et al., 2005). L'identificazione di nuove varianti strutturali segue principalmente due approcci: uno basato sull'uso di array specificamente progettati per la loro identificazione, l'altro su metodi di PCR quantitativa.

Approcci basati sull'utilizzo di array: l'approccio attualmente più robusto è quello basato su esperimenti di genomica comparativa con array-CGH (*array-based Comparative Genome Hybridization*) (Solinas-Toldo et al., 1997; Pinkel D et al., 1998). Questo approccio sfrutta l'ibridazione competitiva di segmenti di un genoma di interesse e di un genoma di confronto diversamente marcati ad un *array* di target genomici (BACs, cDNAs, prodotti di PCR, oligonucleotidi) (Figura 6).

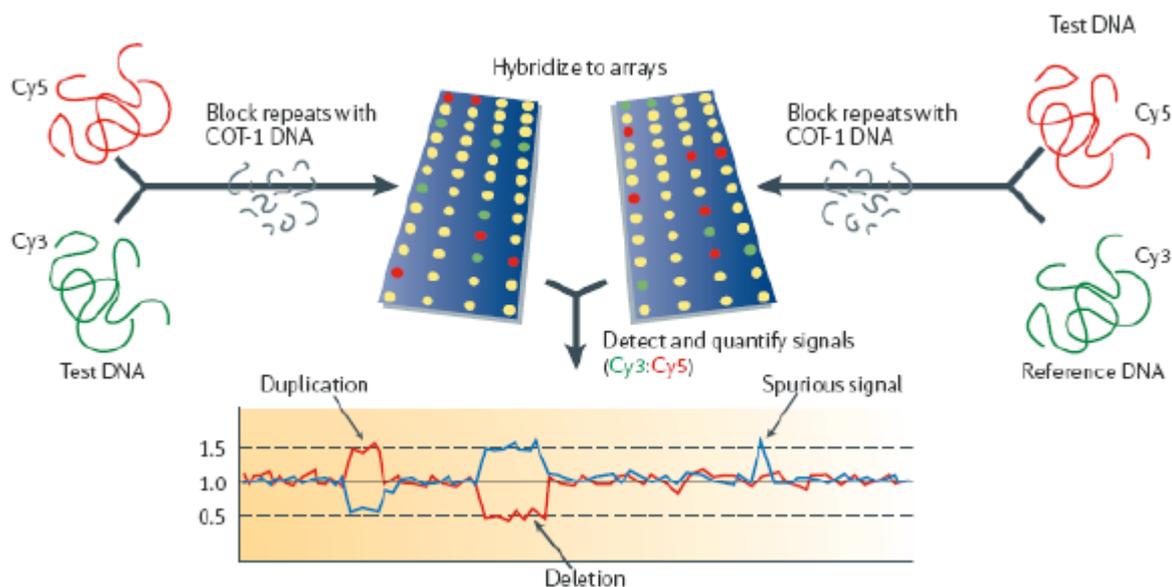


Figura 6: Schema di un esperimento con array-CGH.

In particolare, gli array-CGH con BACs (Iafate AJ et al., 2004; Tuzun E et al., 2005) sono molto diffusi perché provvedono ad un'ampia copertura del genoma (1 clone BAC ogni 1 Mb), perché sono disponibili mappe di BACs altamente affidabili,

perché i cloni BAC sono facilmente reperibili e perché il loro limite di risoluzione è abbastanza alto, (parte da 50Kb fino a raggiungere, teoricamente, poche Kb) (Feuk et al., 2006). Gli array-CGH con oligonucleotidi di 60-100 bp sono in grado di migliorare tale limite di risoluzione e vengono in genere utilizzati con un metodo detto ROMA (*Representational Oligonucleotide Microarray Analysis*) (Lucito R et al., 2003; Sebat et al., 2004). Il metodo ROMA consente di aumentare il rapporto “segnale/fondo” riducendo la complessità del DNA utilizzato per l’analisi (Kennedy G et al., 2003). Infatti, il DNA che deve essere ibridato all’array viene precedentemente trattato con enzimi di restrizione e poi ligato ad alcune sequenze adattatore”. Tali frammenti di sintesi vengono quindi amplificati ottenendo un pool di segmenti di lunghezza paragonabile che sono rappresentativi dell’intero genoma ma hanno una complessità inferiore (Figura 7).

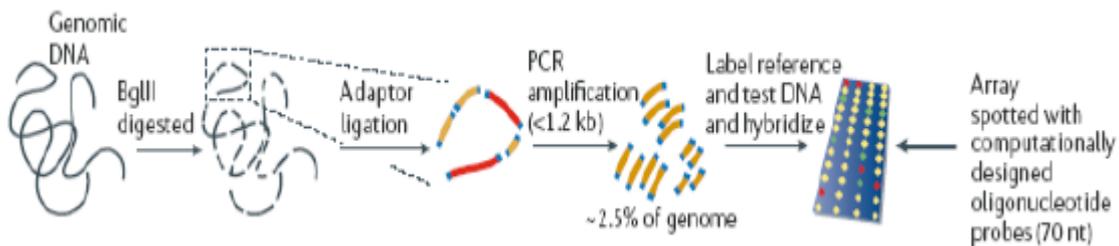


Figura 7: Metodo ROMA (*Representational Oligonucleotide Microarray Analysis*), una variante degli esperimenti con array-CGH.

Un altro metodo ultimamente molto utilizzato sfrutta gli **array per il genotyping di SNPs** allo scopo di rilevare anche CNVs (Slater SR et al. 2005). In questo caso vengono le intensità dei segnali di ibridazione di un campione di DNA genomico di interesse sono confrontati con il valore medio di un DNA di controllo. Una deviazione dell’intensità prodotta dal campione rispetto a quella prodotta dal controllo è indicativa di una variazione del numero di copie e può corrispondere ad esempio ad una delezione o presenza di disomia uniparentale. (Lindblad-toh K et al., 2000; Mei R et al., 2000; Altug- Teber O et al., 2005; Raghavan SK et al., 2005;

Bruce M et al., 2005). Un approccio simile è stato utilizzato nel lavoro presentato in questa tesi nel corso della tipizzazione ad alta densità di SNPs al cromosoma 2 e 7 per l'individuazione di CNVs che potessero essere eventualmente associate all'autismo (Cap. 9 § 9.3.3).

Approcci basati su PCR quantitativa: per identificare CNVs in specifiche regioni genomiche è preferibile un approccio che si basa su PCR. In particolare il metodo della *real-time quantitative* PCR (qPCR) è molto affidabile per l'individuazione di singole delezioni o duplicazioni, ma si adatta più difficilmente ad esperimenti che mirano alla rilevazione di diverse CNVs contemporaneamente in diverse regioni genomiche (Ponchel et al. 2003). A tale proposito sono stati messi a punto metodi alternativi quali la QMPSF (*Quantitative Multiplex PCR of Short Fluorescent Fragments*) (Charbonnier et al. 2000), la MAPH (*Multiplex Amplifiable Probe Hybridization*) (Armour et al. 2000) e la MLPA (*Multiplex Ligation-dependent Probe Amplification*) (Schouten et al. 2002) che sono in grado di rilevare in un solo esperimento la presenza di diverse CNVs fino a 40 regioni diverse. Ogni metodo per l'identificazione di varianti strutturali nel genoma risulta ottimale per alcune tipologie di varianti strutturali. Nel lavoro presentato in questa tesi sono state utilizzate differenti tecniche citate pocanzi quali: Real-time (Cap. 9 § 9.1.6), MLPA (Cap. 9 § 9.1.6) e QMPSF (Cap. 9 § 9.3.3).

2. Pulismo

2.1 Introduzione all'autismo

L'autismo (OMIM #209850) è una grave patologia neuropsichiatrica, la cui eziologia è nella maggior parte dei casi sconosciuta, caratterizzata principalmente dall'inabilità dei soggetti a sviluppare le normali capacità della sfera relazionale e comunicativa, oltre che da comportamenti stereotipati e ripetitivi (American Psychiatric Association, 2000).

La diagnosi di questo disturbo è complicata dalla mancanza di marcatori biologici e dall'elevata eterogeneità dei sintomi di tipo comportamentale. L'autismo rappresenta il prototipo di un gruppo più ampio di disturbi, detti "Disturbi Pervasivi dello sviluppo" (PDD, Pervasive Developmental Disorders) o "disturbi dello spettro autistico" (ASD, Autism Spectrum Disorders), che si distinguono sulla base di carenze in ambito linguistico o nella sfera cognitiva o in base alle relazioni sociali e comportamentali. La categoria PDD comprende, oltre all'autismo propriamente detto anche:

1. La Sindrome di Asperger, che include quei soggetti che pur presentando le classiche problematiche di interazione sociale, non evidenziano alcun disturbo nell'area cognitiva o nel ritardo dello sviluppo delle abilità linguistiche.
2. Il disturbo pervasivo dello sviluppo non altrimenti identificato (PDD-NOS) che include soggetti con caratteristiche simili a quelle osservate nell'autismo, ma non sufficienti per la diagnosi della forma classica del disturbo.
3. Il disturbo disintegrativo dell'infanzia (CDD), più raro delle altre forme, porta dopo un periodo di sviluppo apparentemente normale, ad una regressione con perdita delle capacità acquisite e ad una condizione di disabilità grave.

2.2 Aspetti clinici

Benché alcuni sintomi siano presenti fin dalla nascita, difficilmente è possibile effettuare una diagnosi di autismo prima di 24-36 mesi (Bailey *et al.*, 1996). L'autismo si manifesta con sintomi variabili da un individuo all'altro e secondo

molteplici livelli di gravità anche fra membri della stessa famiglia. Un individuo autistico è caratterizzato anzitutto da una grave difficoltà di gestione delle interazioni sociali. Spesso ha gravi problemi di comunicazione, anche a livello di linguaggio, e presenta comportamenti ripetitivi e stereotipati o interessi rigidi. Le principali caratteristiche cliniche (World Health Organization, WHO, 1992) sono descritte in tabella 2.

Compromissione dell'interazione sociale:

- Comprende la componente non verbale del linguaggio, come difficoltà a mantenere lo sguardo diretto, espressioni facciali, postura del corpo e gestualità nel contesto dell'interazione sociale.
- Incapacità a sviluppare legami ludici con i coetanei.
- Raramente gli individui affetti richiedono la presenza di altre persone per condividere momenti di gioia, difficoltà o interessi e difficilmente si prestano ad interpretare gli stati d'animo delle persone che li circondano.
- Mancanza della reciprocità socio-emozionale.

Compromissione della comunicazione:

- Ritardo lieve o grave nello sviluppo del linguaggio parlato, molto spesso caratterizzato da poca flessibilità espressiva, creatività e fantasia. Inoltre circa un terzo dei soggetti non sviluppa il senso del passato.
- Assenza nella comunicazione non verbale, della mimica, ma anche incapacità a svolgere giochi basati sull'imitazione e sulla gestualità.
- In soggetti capaci di adeguata capacità linguistica, mancanza di sincronia e reciprocità nelle relazioni interpersonali.
- Mancanza di risposte emozionali.
- Incapacità a modulare la voce e a conferirle giusta tonalità, per cui non è marcato alcun aspetto emotivo nel linguaggio verbale.
- Mancanza di enfasi o stupore nella comunicazione non verbale.

Comportamenti ristretti, stereotipati e ripetitivi:

- Attività motorie ristrette ad azioni stereotipate o ripetitive.
- Interesse per parti inusuali di oggetti, a volte dettagli non funzionali di giochi.
- Gli oggetti spesso rappresentano mezzi per compiere azioni compulsive e ripetitive.
- Stati d'angoscia per piccoli cambiamenti nel contorno familiare.
- Sottomissione a rituali o abitudini slegate dal contesto del vissuto e del tutto fini a se stesse.

Tabella 2: Manifestazioni cliniche che si possono osservare negli individui con disturbo autistico.

Ulteriori comportamenti osservati in soggetti autistici sono risposte anomale ai suoni, al tatto o ad altri stimoli sensoriali. Molti mostrano una ridotta sensibilità al dolore, ma possono anche essere straordinariamente sensibili ad altre sensazioni. Inoltre, sono frequenti disturbi del sonno, preferenze limitate per i cibi e problemi nel camminare e nello scrivere.

2.3 Diagnosi

Fondamentale per la ricerca genetica sull'autismo è partire da individui la cui diagnosi non sia erronea o ambigua, ecco perché è necessario avere mezzi che possano identificare lo spettro preciso dei disturbi. Recentemente complessi strumenti diagnostici sono stati messi a punto per consentire questo fine, come l'ADI-R (Autism Diagnostic Interview – Revised) (Lord, C. *et al* 1994) e l'ADOS-G (Autism Diagnostic Observational Schedule) (Lord, C. *et al* 2000), che, insieme, permettono una diagnosi con un'elevata specificità e attendibilità.

2.4 Anormalità neurologiche e fisiologiche correlate all'autismo

Al tipico quadro autistico spesso si possono sommare patologie non assimilabili alla sfera relazionale, come ad esempio anormalità fisiologiche a carico del sistema nervoso, del sistema dei neurotrasmettitori ed epilessia. E' da chiarire comunque, se si possa parlare di vera relazione o siano dei casi sporadici in cui le distinte patologie si manifestano in maniera non associativa.

2.4.1 Anomalie cerebrali

Nei soggetti autistici è possibile riscontrare caratteristiche cerebrali anomale, ed in particolar modo:

1) Un aumento del volume del cervello, evidenziato mediante studi condotti con MRI (magnetic resonance imaging), studi post-mortem, e misure della circonferenza cranica. Non è chiaro se ciò sia causato da un incremento della proliferazione cellulare non seguita da morte cellulare o se sia dovuto alla mancanza dello sfoltimento delle sinapsi. Comunque attualmente ci sono pochi studi per poter trarre conclusioni definitive.

2) Diminuzione delle cellule del Purkinje nel cervelletto, rilevato mediante studi post-mortem. Tale riduzione cellulare, avverrebbe nel tardo periodo prenatale o all'inizio dello sviluppo postatale (Whitney ER, 2008).

3) Ipometabolismo di alcune aree della corteccia cerebrale, evidenziate mediante studi di PET e fMRI. Ciò non vuol dire che tali aree del cervello siano direttamente coinvolte (Ring, H. *et al* 2000).

2.4.2 Epilessia

L'associazione tra epilessia ed ASDs è stata descritta fin dal 1960. La stima delle persone affette da epilessia varia, ma sicuramente è in percentuale molto più alta nei soggetti con autismo rispetto alla popolazione normale. Si calcola che almeno un terzo delle persone adulte autistiche manifesti disturbi epilettici, con schemi differenti in base all'età dell'insorgenza (Spence SJ, 2009).

2.4.3 Sistema dei neurotrasmettitori

Numerosi studi hanno indagato il coinvolgimento dei vari sistemi di neurotrasmettitori nell'autismo, con risultati poco significativi. L'unica eccezione è il sistema serotoninergico.

Alcuni tra questi studi hanno infatti confermato che una certa proporzione di individui autistici (un terzo circa) hanno aumentati livelli di serotonina (5-HT) nel sangue e/o nelle urine. Ciò sembra dovuto probabilmente ad un aumento dello storage nelle piastrine piuttosto che ad anomalie di sintesi (Croonenberghs, J. *et al.* 2000).

Questo potrebbe riflettere una bassa disponibilità di serotonina nel SNC, come evidenziato da studi di PET (Cook, E. H. & Leventhal, B. L 1996).

Inoltre, il trattamento dei pazienti con inibitori dei trasportatori o del re-uptake selettivo della serotonina, ha dimostrato di migliorare alcuni sintomi del disturbo autistico. Il gene che codifica per il trasportatore della serotonina (5-HTT) è stato oggetto di numerosi studi come possibile gene candidato per l'autismo, con risultati ancora poco chiari.

2.5 Trattamenti farmacologici e terapia

Non esistono cure psicofarmacologiche per i disturbi centrali dell'autismo, e fino a quando non saranno chiare le basi delle disfunzioni neurologiche che investono l'autismo nessuna terapia farmacologica potrà essere considerata specifica.

La terapia psicofarmacologica per l'autismo cerca di migliorare la sintomatologia associata al disturbo come: la scarsa attenzione e concentrazione, i fenomeni ossessivi e compulsivi, l'eccessiva ansietà, i problemi legati al sonno, la depressione e l'autolesionismo.

In particolare, i farmaci che regolano la via serotoninergica, come gli inibitori dei trasportatori o del re-uptake selettivo della serotonina (SSRIs), riducono i comportamenti ripetitivi ed aggressivi (Gordon,C., 1993; Sanchez,L.E. *et al.* 1996; Remington,G., 2001). Altri farmaci inibiscono la trasmissione dell'impulso nei neuroni dopaminergici, riducendo alcune sintomatologie comportamentali (Anderson,L.T. *et al* 1989; Naruse,H. *et al.* 1982). Altri ancora agiscono sul sistema adrenergico e noradrenergico riducendo l'iperattività (Gordon,C., 1992; Jaselskis,C.A., 1992; Lichter,D.G. 1996).

L'approccio farmacologico nella terapia deve essere affiancato con un valido programma psicopedagogico, in modo da migliorare gli aspetti cognitivi e comportamentali.

2.6 Autismo: dalla sua scoperta al riconoscimento dell'importanza dei fattori genetici

L'autismo fu descritto per la prima volta da Leo Kanner nel 1943, quando descrisse 11 bambini la cui condizione differiva da quella di ritardo mentale; infatti Kanner parlava di "inabilità innata per i contatti interpersonali", dovuta ad un disturbo dello sviluppo cerebrale, annotandolo come "autismo infantile" poiché mancava l'interazione sociale del bambino (Kanner L, 1943). Un anno dopo Hans Asperger in

Germania descrisse casi simili che chiamò soggetti a “psicopatia autistica” (Asperger H, 1944).

Kanner e Asperger ipotizzarono che questi individui fossero affetti da una malattia di origine neuropatologica. Questo primo tentativo di spiegare l’eziologia dell’autismo con una base biologica fu ben presto sostituito da altri autori che ne sostenevano la natura psicologica; infatti negli anni ’60-’70 la psichiatria tendeva a spiegare molte malattie come il risultato di esperienze traumatiche nei primi anni di vita. Ci sono molti autori quali: Mahler, Bettelheim, Winnecott, Klein e Tustin che incentrano i loro lavori sulla rottura troppo precoce del rapporto di dipendenza madre-bambino (le madri vengono definite “ frigorifero”). Si è successivamente dimostrato che i genitori di bambini autistici erano assolutamente nella norma, confrontati con genitori controllo.

Nel 1977 Rutter e Folstein furono i primi a condurre studi genetici con metodiche quantitative quali concordanza tra gemelli e studio di alberi genealogici che fornirono sufficienti prove sull’importanza della componente genetica nell’autismo. In uno degli studi più importanti compiuto finora (Bailey *et al.*, 1995), il tasso di concordanza per l’autismo in coppie di gemelli monozigotici (MZ) è risultato notevolmente più alto (circa 60%), rispetto al valore riscontrato nelle coppie di gemelli dizigotici dello stesso sesso (DZ) (vicino al 3%). Nello studio di Bailey *et al.* (1995), la concordanza dei gemelli MZ per lo spettro più ampio di anormalità socio/comunicative, in cui le caratteristiche fenotipiche sono presenti in forma meno grave (il cosiddetto “*broader phenotype*”) è del 92%, rispetto ad una concordanza del 10% nei gemelli DZ (figura 7).

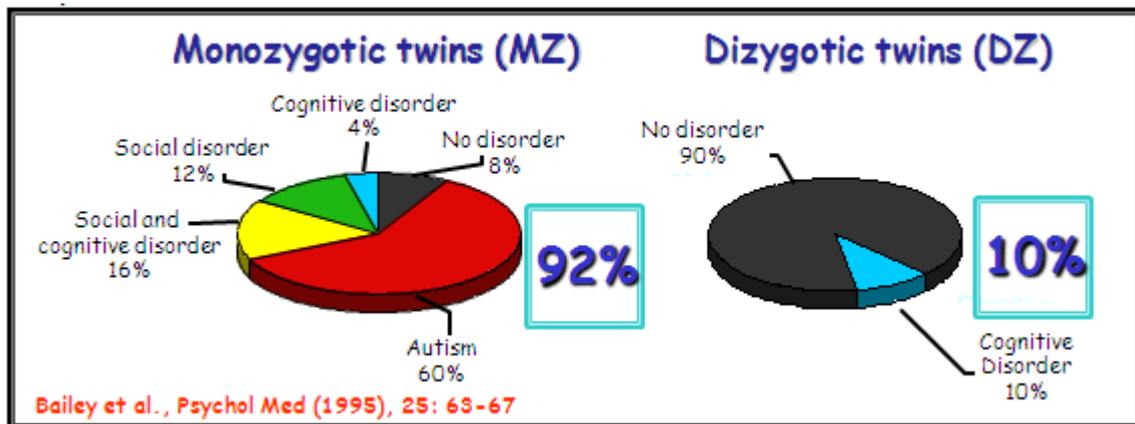


Figura 7: Concordanza nei gemelli monozigotici (MZ) e diziotici (DZ) dello spettro dei disturbi connessi all'autismo.

In studi di ricorrenza familiare (Szatmari *et al.*, 1998), la percentuale di rischio in fratelli di individui affetti è stimata attorno al 3%: valore quest'ultimo di 10-30 volte maggiore, rispetto al tasso di prevalenza nella popolazione generale (Cap. 2 § 2.7).

Gli studi di ricorrenza familiare e quelli sui gemelli sono stati fondamentali per dimostrare la forte influenza di fattori ereditari nell'autismo, tuttavia ne hanno anche indicato la complessità genetica. I pattern di ereditarietà infatti non sono compatibili con i modelli monogenici mendeliani, ma suggeriscono il probabile coinvolgimento di molteplici geni di suscettibilità (ereditarietà poligenica o oligogenica). Secondo questo modello di ereditarietà complessa, varianti funzionali nei geni di suscettibilità hanno verosimilmente un effetto debole o moderato: ciascun locus, da solo, contribuisce ad aumentare il rischio di sviluppare la malattia, ma non è in grado di determinare il completo fenotipo clinico. Inoltre diversi set di geni potrebbero essere coinvolti in diversi individui (eterogeneità genetica). Probabilmente i geni responsabili, agirebbero in maniera concertata con una combinazione di fattori epigenetici e ambientali diversi.

Quali siano gli specifici geni che contribuiscono alla patogenesi dell'autismo, rimane fino ad oggi sconosciuto. Solamente in una minoranza di casi (meno del 10%)

l'autismo è associato a patologie causate da un singolo gene, tra cui la sindrome dell'X fragile e la sclerosi tuberosa (TSC).

2.7 Prevalenza dell'autismo

Dagli anni '60 in poi sono stati portati a termine più di 30 studi epidemiologici sull'autismo (Volkmar et al., 2004; Baird et al., 2001; Wing et al., 2002; Fombonne et al., 2003; Croen et al., 2002). L'aspetto più rilevante di questi studi è un sostanziale aumento della prevalenza dell'autismo nella popolazione generale, la cui stima passa da 4.4 casi per 10.000 negli studi epidemiologici pubblicati dal 1966 al 1991, a 12.7 casi su 10.000 negli studi degli ultimi dieci anni. Tale incremento potrebbe essere dovuto, più che all'aumento effettivo della prevalenza, a più sensibili e precoci strumenti diagnostici.

Fombonne suggerisce, in base ai dati disponibili, che la stima di prevalenza sia di circa 10 casi di AD su 10.000 (Fombonne et al., 2003). Se si prendono in considerazione tutti i diversi sottotipi di PDD la prevalenza generale è nettamente più alta, raggiungendo il valore di 60 individui su 10.000 (Fombonne, 2003).

Tutti gli studi epidemiologici inoltre indicano una più elevata incidenza dell'autismo nei maschi rispetto le femmine, con un rapporto stimato di 3.5 o 4 : 1 (Fombonne, E., 2003; Lord, C., 1982; Volkmar, F., 1993).

Inoltre non vi sono evidenze di associazione tra autismo e particolari classi sociali o appartenenza etnica (Fombonne, E., 2003; Ritvo, E.R., 1989).

3. Genetica molecolare dell'autismo

3.1 contributi monogenici associati all'autismo

La presenza di eterogeneità genetica è anche indicata dal fatto che in una minoranza di casi (<10%) l'autismo è associato a patologie causate da un singolo gene, tra cui la sindrome dell'X fragile (FRAXA), la sclerosi tuberosa (TSC), la neurofibromatosi (Klauck, 2006) o a varie anomalie cromosomiche (Smalley et al. 1998; Dykens and Volkmar, 1997; Feinstein and Reiss, 1998).

TSC: La Sclerosi Tuberosa (Omim # 191100) è un disturbo autosomico dominante neurocutaneo caratterizzato dalla presenza di tubercoli nella regione corticale e cerebrale, lesioni renali, attacchi epilettici durante l'infanzia, resistenza a trattamenti antiepilettici (Bolton et al., 2002; Gutierrez et al., 1998; Jambaque et al., 2000). Questo disturbo è causato da mutazioni in due diversi geni: *TSC1* (Omim # 605284) e *TSC2* (Omim # 191092). Studi epidemiologici hanno mostrato che la prevalenza di Sclerosi Tuberosa in bambini autistici e di autismo in casi di Sclerosi Tuberosa è cento volte superiore a quella attesa (Harrison et al., 1997; Fombonne, 2003).

FRAXA: La Sindrome dell'X-fragile (Omim # 300624) è un disturbo caratterizzato da ritardo mentale con un'incidenza di 1 : 4000 negli individui di sesso maschile e 1 : 8000 negli individui di sesso femminile (Lombroso, 2003). La causa eziologica di questa malattia è l'ipermetilazione di una tripletta (CGG) altamente ripetuta (>200 *repeats*) presente nella regione non tradotta al 5' del gene *FMR1* (Omim # 309550) localizzato sul cromosoma Xq27 con conseguente riduzione del prodotto proteico (Sutcliffe et al., 1992). Il 2-5% di bambini con una diagnosi di autismo più lieve sono portatori di mutazioni associate a FRAXA (Fombonne, 2003; Bailey, 1993; Reddy, 2005; Wassink et al., 2001).

Neurologine: La classe dei geni delle neurologine è stata la prima ad essere associata all'autismo. Una descrizione più ampia degli studi finora condotti è riportata nel Cap. 9 § 9.3.6.

Questi studi mettono in risalto il ruolo del sistema glutamatergico nella manifestazione del fenotipo autistico.

SHANK3: Il gene *SHANK3* (*SH3 and Multiple Ankyrin Repeat Domains 3*) (Omim # 606230) mappa nella regione 22q13.3 e codifica per una proteina espressa preferenzialmente nella corteccia cerebrale e nel cervelletto (Bonaglia et al., 2001). *SHANK3* è responsabile della sindrome da delezione terminale del cromosoma 22q13.3. Gli individui che presentano questa anomalia cromosomica presentano ritardo mentale e/o disturbi del linguaggio e caratteristiche fenotipiche riconducibili a quelle dello spettro autistico (Bonaglia et al., 2001; Anderlid et al., 2002; Wilson et al., 2003; Durand et al., 2007)

I geni specifici che contribuiscono alla patogenesi nella restante maggioranza di forme idiopatiche di autismo rimangono ad oggi sconosciuti.

3.2 Studi di linkage per l'autismo estesi all'intero genoma

Essendo ampiamente sconosciute le basi neurobiologiche dell'autismo, negli ultimi anni diversi gruppi di ricerca hanno effettuato studi di *linkage* non parametrici estesi all'intero genoma con lo scopo di identificare quelle regioni che contengono potenziali loci di suscettibilità (*affected relative pair method*, descritto nel Cap 1) (Auranen et al. 2003; Barret et al., 1999; Buxbaum et al., 2001; Cantor et al., 2005; Coon et al., 2005; IMSCAC, 1998; IMGSA, 2001; Lamb et al., 2005; Lauritsen et al., 2006; Liu et al., 2001; McCauley et al., 2005; Philippe et al., 1999; Risch et al. 1999; Shao et al., 2002; Yonan et al., 2003).

Il primo *genome scan* sull'autismo è stato portato a termine, nel 1998, dall'***International Molecular Genetic Study of Autism Consortium*** (IMGSA), un consorzio composto da ricercatori e clinici di diversi paesi (Gran Bretagna, Germania, Francia, Olanda, Italia, Danimarca, Grecia e U.S.A.) e di cui fa parte anche il nostro gruppo di ricerca. La necessità di costruire un consorzio internazionale per questo tipo di studi deriva dalla relativa rarità di famiglie con più soggetti affetti e dalla necessità di analizzare un grande numero di famiglie diagnosticate con criteri omogenei.

In questo primo studio sono stati analizzati 354 marcatori in un campione di 83 coppie di fratelli affetti (*affected sibling pairs*, ASPs), ed è stato calcolato il MLS *multipoint* lungo i 23 cromosomi.

Sono state individuate 12 regioni, ma la regione con maggiore evidenza di *linkage* si è rivelata la regione 7q con un $MLS=2,53$ (IMGSAC, 1998). Nel 2001 il *genome scan* è stato esteso ad un campione di 152 ASPs, utilizzando 119 marcatori, all'interno delle regioni di *linkage* individuate nel primo *genome scan*. La regione con maggiore evidenza di *linkage* è risultata essere localizzata nella porzione cromosomica 2q ($MLS=3.74$) in corrispondenza del marcatore D2S2188, mentre se si considerano le coppie di fratelli affetti che soddisfano i criteri diagnostici più stringenti per l'autismo il valore di MLS sale a 4.80 (IMGSAC, 2001); tale valore rappresenta un risultato estremamente significativo ed il dato più importante finora ottenuto nei differenti *genome scans* effettuati per l'autismo. Altre regioni in cui è stato osservato un valore significativo di *linkage* sono la regione 7q ($MLS=3.20$), 16p ($MLS=2.93$), 16p ($MLS=2.93$) e 17q ($MLS=2.34$) (IMGSAC, 2001) (Figura 8).

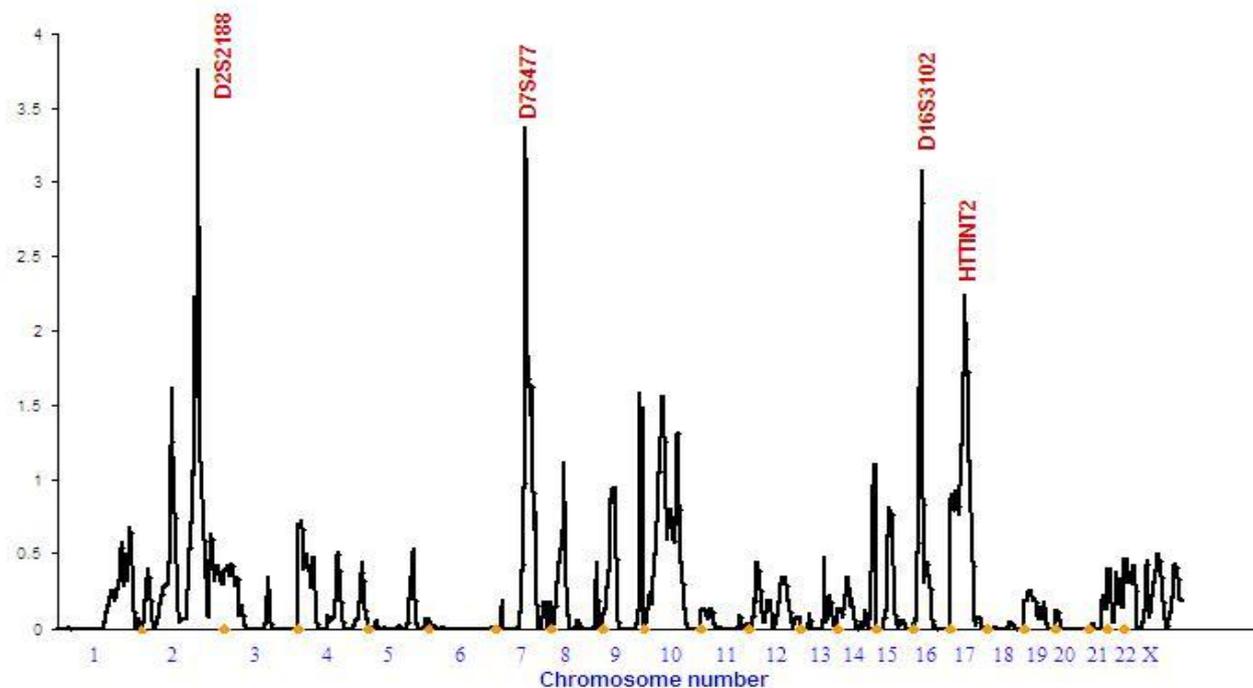


Figura 8: Risultati del primo *genome scan* effettuato dall'IMGSAC. In rosso sono indicati i marcatori in corrispondenza dei quali si ottiene il valore di LODscore più significativo rispettivamente sui cromosomi 2q, 7q, 6p, 17q.

Nel 2005 il consorzio ha esteso l'analisi a 219 ASPs analizzando 396 marcatori in corrispondenza dei loci precedentemente individuati sui cromosomi 2, 7, 9, 10, 16, 17 e includendo nell'analisi, sulla base dei risultati riportati in letteratura, anche i cromosomi 3, 13 e 15 (Lamb et al., 2005).

Valori di $MLS > 2$ sono stati trovati per i cromosomi 2q ($MLS = 2.54$) in corrispondenza del marcatore D2S2314, posto nelle vicinanze del marcatore D2S2218 individuato nel primo *genome scan*, 7q ($MLS = 2.31$) in corrispondenza del marcatore D7S530 e 9p ($MLS = 2.12$) in corrispondenza del marcatore D9S161.

I dati di *linkage* ottenuti dall'IMGSAC sono stati, in parte, replicati da altri 8 *genome scans*, condotti da altrettanti gruppi, utilizzando campioni di famiglie indipendenti, da cui sono emersi valori positivi ($MLS > 1$) per regioni a livello del cromosoma 6q (Philippe et al., 1999), 7q e 13q (Barrett et al., 1999), 1p (Rich et al., 1999), 2q (Buxbaum et al., 2001), 5q, 7q, 19q e X (Liu et al., 2001), 2q, 3p, 7q, 19q, Xq (Shao et al., 2002), 1p, 3q, 7q, 17q, Xq (Auranen et al., 2003), e a livello dei cromosomi 17q e 5p (Yonan et al., 2003) a dimostrazione dell'eterogeneità di questa patologia.

La figura qui sotto (Figura 9) mostra un quadro generale di tutte le regioni cromosomiche che presentano valori significativi di *linkage* individuate da dieci *genome-wide scans* effettuati da diversi gruppi di ricerca nel corso degli ultimi dieci anni.

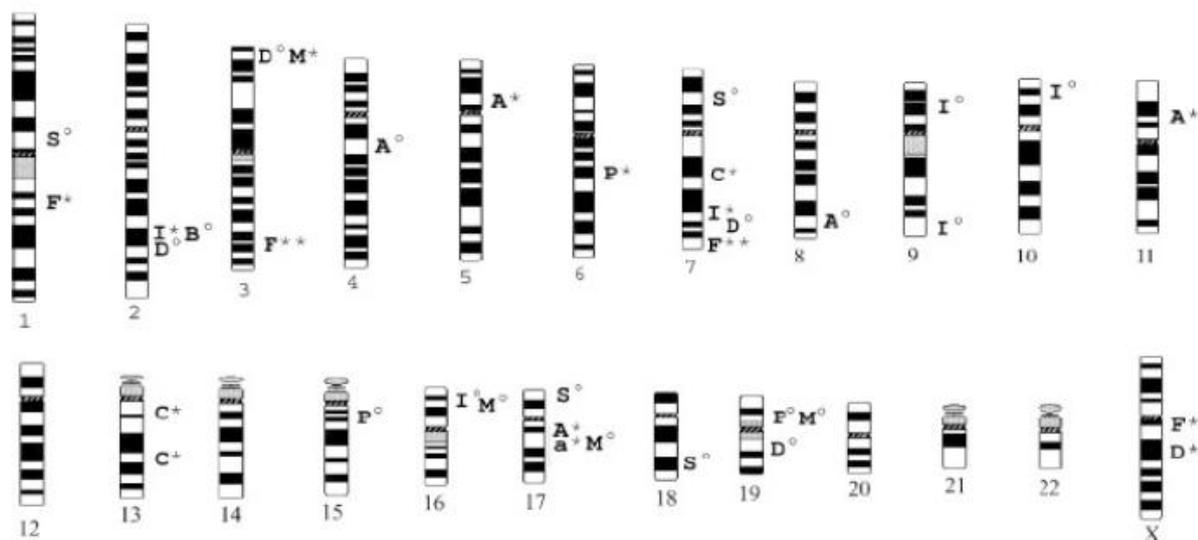


Figure 9: Regioni cromosomiche che presentano valori significativi di linkage individuate da 10 genome-wide scans effettuati per la ricerca di loci di suscettibilità per ASD.

I risultati di ogni studio sono indicati da una lettera seguita da ** se $MLS > 3.6$, da * se $MLS > 2.2$, e da ° se $MLS > 1$. **A:** AGRE (Yonan et al., 2003); **(a)** AGRE2 (Cantor et al., 2005); **(B)** (Buxbaum et al., 2001), **(C)** CLSA (Barrett et al., 1999); **(D)** Duke (Shao et al., 2002b); **(F)** Finland (Auranen et al., 2002); **(I)** IMGSAC (Lamb et al., 2005); **(M)** (McCauley et al., 2005); **(P)** PARIS (Philippe et al., 1999); **(S)** Stanford (Risch et al., 1999).

Le regioni sui cromosomi 2q, 7q, 3p, 17q, 19q sono quelle per le quali sono stati trovati risultati parzialmente sovrapponibili, confermati da almeno due diversi studi di *linkage* condotti da gruppi indipendenti.

La sovrapposizione dei risultati è incoraggiante e rappresenta un primo passo verso l'identificazione di loci di suscettibilità all'autismo. Tuttavia, l'interpretazione dei dati di *linkage* provenienti da diversi *genome scans* comporta alcune difficoltà insite in questo tipo di approccio applicato alle malattie complesse. Infatti, è tuttora poco chiaro quale valore di *linkage* sia da considerarsi significativo in uno studio esteso all'intero genoma, in quanto vengono effettuati molteplici test statistici e non si conosce il numero dei reali loci di suscettibilità.

Come già detto, pur utilizzando strumenti diagnostici internazionalmente standardizzati, la variabilità che si osserva tra i risultati ottenuti dai diversi studi indipendenti può dipendere da diversi fattori quali:

- criteri di inclusione adottati
- differenze nei metodi analitici
- differenze nella scelta dei marcatori tipizzati
- grandezza del campione

Per tutti questi motivi, anche condurre una meta-analisi dei dati, messi attualmente a disposizione dai vari gruppi di ricerca, risulta particolarmente complicato (Lamb et al., 2002; Klauck, 2006; Trikalinos et al., 2006).

Dato che l'autismo segue un modello di ereditarietà complesso in cui sono coinvolti molti loci di suscettibilità, ciascuno con un effetto moderato, la non replicazione di un risultato in uno studio indipendente non esclude completamente la possibilità che questo sia un risultato reale, ma suggerisce semplicemente la presenza di eterogeneità genica e l'esistenza di più geni che interagiscono in maniera epistatica.

Per limitare il problema rappresentato dall'eterogeneità genetica sono stati condotti studi di *linkage* o utilizzando criteri di inclusione più stringenti o suddividendo le varie famiglie in sottogruppi che presentano tratti caratteristici del fenotipo autistico (endofenocopie) quali problemi nel linguaggio, difficoltà nella lettura, comportamenti stereotipati e ripetitivi, sesso. Per esempio, stratificando le famiglie sulla base di tratti relativi al linguaggio si è osservato un aumento del potere di *linkage* in corrispondenza del cromosoma 7q (Folstein e Mankoski, 2000; Liu et al., 2001; Bradford et al., 2001) e del cromosoma 2q (Buxbaum et al., 2001; Shao et al., 2002), mentre utilizzando come endofenocopie famiglie con individui con comportamenti ossessivi-compulsivi è stato individuato un ulteriore locus di suscettibilità sul cromosoma 1q (Buxbaum et al., 2004; Silverman et al., 2002). Inoltre, l'osservazione di una prevalenza maggiore di casi di autismo in individui di sesso maschile rispetto

ad individui di sesso femminile ha suggerito il probabile coinvolgimento di fattori sesso-specifici e quindi la possibilità di stratificare le famiglie in base al sesso del probando (Cantor et al., 2005; Stone et al., 2004; IMGSAC, 2001). Anche in questo caso, la non riproducibilità dei risultati porta alla conclusione di uno scarso coinvolgimento del cromosoma X nella maggioranza dei casi di autismo.

Benché la stratificazione delle famiglie può essere un buon sistema per aumentare l'omogeneità genetica e, di conseguenza, la probabilità di identificare loci di suscettibilità, il problema fondamentale negli studi condotti fino ad ora rimane probabilmente la dimensione del campione analizzato che non è sufficientemente grande per identificare varianti con effetto moderato e per discriminare i veri dai falsi positivi (Bacchelli et al., 2006).

Per ovviare questo problema è stato creato un Consorzio Internazionale, l'*Autism Genome Consortium* (AGP), di cui il nostro gruppo di ricerca fa parte, che coinvolge 50 differenti gruppi di ricerca in tutto il mondo. Il consorzio ha raccolto un totale di 1168 famiglie *multiplex* provenienti da Stati Uniti, Canada ed Europa, creando un campione di dimensione tre volte superiore a quello utilizzato da qualsiasi altro studio genetico condotto fino ad oggi sull'autismo.

È stata condotta un'analisi di *linkage* estesa all'intero genoma analizzando più di 10.000 SNPs, utilizzando la piattaforma Affimetrix. Per ridurre l'eterogeneità presente nel campione, le famiglie sono state suddivise sulla base di tre diversi livelli diagnostici: famiglie con un fenotipo autistico stringente, famiglie con un fenotipo più lieve, tutte le famiglie. Il valore di *linkage* statisticamente più significativo è stato trovato in corrispondenza della regione 11p12-p13 se si

considerano tutte le famiglie, e in corrispondenza delle regioni 2q e 7q se si considerano solo le famiglie di origine europea (AGP, 2007). Tenendo conto della generale difficoltà nella replicazione dei risultati di *linkage* nello studio delle malattie complesse, è quindi molto interessante ed incoraggiante notare come le regioni sul

cromosoma 7q e cromosoma 2q, inizialmente identificate dall'IMGSAC, siano state successivamente confermate da ulteriori studi indipendenti.

3.3 Geni candidati per l'autismo

Gli studi di *linkage*, intrapresi negli ultimi dieci anni, hanno identificato diverse regioni cromosomiche di interesse per la suscettibilità all'autismo. Tuttavia, questo tipo di analisi non ha circoscritto abbastanza gli intervalli cromosomici e per questo motivo molti gruppi di ricerca hanno iniziato ad effettuare *screening* di geni candidati e studi di associazione nelle regioni definite dalle analisi di *linkage*.

La disponibilità della sequenza completa del genoma umano ed il suo costante aggiornamento ha facilitato l'efficienza nello *screening* di geni candidati che mappano in determinate regioni cromosomiche. Un'indagine della letteratura mostra l'enorme numero (circa un centinaio) di geni candidati analizzati e la quantità di lavori sull'argomento pubblicati negli ultimi dieci anni.

Di solito, i geni candidati sono scelti in base al loro profilo di espressione e al possibile ruolo nello sviluppo e funzionalità del SNC (candidati funzionali), alla loro localizzazione all'interno di intervalli individuati tramite un'analisi di *linkage* o alla presenza di riarrangiamenti cromosomici associati alla patologia (candidati posizionali).

Un'analisi accurata dei geni candidati viene condotta attraverso il sequenziamento del gene nelle regioni esoniche e nelle regioni regolatorie in individui affetti e in individui sani.

Nel caso di malattie complesse non è immediato riuscire ad identificare varianti funzionali in geni candidati anche perchè queste varianti possono essere localizzate all'interno di regioni introniche o lontane da elementi regolatori. Per questo motivo, la maggior parte degli studi di associazione partono da un approccio "indiretto" basato sul *linkage disequilibrium* (LD), dove per LD si intende un'associazione non-casuale di alleli vicini. Quindi, un'associazione di tipo indiretto viene individuata

solo se l'allele di un marcatore è in LD con la variante di suscettibilità, anche se il marcatore stesso non ha alcun effetto funzionale.

Negli ultimi dieci anni sono stati analizzati più di 100 geni candidati (posizionali e funzionali) alla ricerca di varianti implicate nell'eziologia del disturbo autistico. Tuttavia, per nessuno di questi geni è stato dimostrato un chiaro coinvolgimento nella patogenesi della malattia. Alcuni esempi sono descritti di seguito:

RELN: La Relina (*RELN*) (Omim # 600514) è un gene che mappa nella regione 7q22, una regione di *linkage* identificata da diversi *genome scan*. Questo gene codifica per una proteina coinvolta nella migrazione neuronale durante lo sviluppo cerebrale. Inoltre, si è visto che il topo *reeler*, un mutante naturale portatore di una larga delezione di questo gene, mostra alterazioni nella posizione dei neuroni nella corteccia celebrale e nell'ippocampo, le stesse regioni cerebrali in cui sono state riscontrate anomalie in individui autistici (D'Arcangelo et al., 1995; Bailey et al., 1998). In uno studio iniziale (Persico et al., 2001) era stata riportata l'esistenza di una trasmissione preferenziale dell'allele più lungo di una sequenza ripetuta (tripletta) presente nella regione non tradotta al 5' del gene. Gli autori suggeriscono che alcune varianti alleliche di questo polimorfismo sono associate all'espressione del gene (Persico et al., 2006). Inoltre, successivi studi evidenziano interazione genica tra il gene *RELN* ed il gene *APOE*, quest'ultimo è coinvolto nella corretta formazione della corteccia cerebrale insieme al gene *RELN* (Ashley-Koch AE, 2007). La replicazione di questi risultati in campioni indipendenti ha fornito poche evidenze di associazione tra il gene *RELN* e fenotipo autistico. (Krebs et al., 2002; Devlin et al., 2004; Li et al., 2004; Dutta S., 2008).

SLC6A4: Questo gene codifica per il trasportatore della serotonina (*SLC6A4/5-HTT*) (Omim # 182138) ed ha suscitato notevole interesse in seguito alle osservazioni sugli elevati livelli di serotonina nel sangue e nelle urine di individui autistici e alla risposta positiva a sintomi caratteristici del fenotipo autistico (comportamenti aggressivi e ripetitivi) dopo trattamento con inibitori del trasportatore stesso

(Hollander et al., 2005). Questo gene mappa nella regione cromosomica 17q11-12, precedentemente identificata tra tre distinti studi di *linkage* (Risch et al, 1999; Yonan et al., 2003; Cantor et al., 2005; McCauley et al, 2005).

Benchè rappresenti un ottimo gene candidato per l'autismo gli studi finora condotti mostrano risultati non significativi (Maestrini et al., 1999; Wu et al., 2005; Ramoz, 2006), o comunque contrastanti (Prasad HC, 2009; Guhathakurta S, 2008; Devlin et al., 2005; McCauley et al., 2005), evidenziando la presenza di eterogeneità tra gli studi condotti dai diversi gruppi di ricerca.

MECP2: Il gene *MECP2* (*methyl-CpG-binding protein-2*) (Omim # 300496) mappa nella regione cromosomica Xq28 e codifica per una proteina responsabile di un altro disturbo neurologico, la Sindrome di Rett (Amir et a., 1999) classificato tra i disturbi pervasivi dello sviluppo. *Screening* di mutazioni di questo gene in individui autistici hanno identificato alcune rare varianti presenti al 3'UTR del gene (Coutinho AM, 2007), altri studi invece non convergono su simili risultati (Xi C, 2007; Vourc'h et al., 2001). Una descrizione più dettagliata del gene *MECP2*, incluso nel lavoro di tesi, verrà trattato nel Cap. 4 § 4.2.1.

MET : Il gene *MET* (Omim # 164860) mappa nella regione 7q31 e codifica per un recettore tirosin chinasi. Nonostante il suo ruolo nella formazione di metastasi in diversi tipi di cancro, *MET* è stato considerato un buon candidato per l'autismo in quanto contribuisce allo sviluppo della corteccia cerebrale e del cervelletto, regioni altamente compromesse negli individui autistici. Due recenti studi suggeriscono il coinvolgimento di questo gene, in particolare nel primo studio condotto su 204 famiglie *multiplex*, e successivamente replicato su altre 539 famiglie, si ottiene associazione con un polimorfismo funzionale nella regione promotrice del gene *MET* (Campbell et al, 2006). Nel secondo studio condotto in 325 famiglie *multiplex* ed 82 trios del consorzio IMGSAC si ottiene associazione con un polimorfismo nell'introne 1 del gene *MET* (Sousa et al, 2009).

FOXP2: Il gene *FOXP2* (*Forkhead Box P2*) (Omim # 605317) è un gene che mappa sul cromosoma 7q ed è stato considerato un buon candidato in quanto mutazioni di questo gene erano state osservate in una famiglia con individui con problemi nel linguaggio (Lai et al., 2000). Tuttavia, con l'eccezione di due studi (Gong et al. 2004; Li et al., 2005) non è stata trovata associazione tra il fenotipo autistico e le varianti individuate in *FOXP2* (Gauthier et al., 2003; Marui et al., 2005; Wassink et al., 2002; Newbury et al., 2002). Il prodotto genico di *FOXP2* è un fattore di trascrizione che regola l'espressione del gene *CNTAP2* (Vernes SC, 2008). Recentemente, numerosi studi hanno evidenziato associazione di questo gene con l'autismo (Bakkaloglu B, 2008; Arking DE, 2008; Alarcón M, 2008). Inoltre è stato dimostrato che SNPs nel gene *CNTAP2* sono anche associati a disturbi specifici del linguaggio (SLI) (Vernes SC, 2008). Il pathway *FOXP2-CNTAP2* sembra quindi essere implicato in diversi disturbi: mutazioni di *FOXP2* causano forme monogeniche di disturbi del linguaggio, mentre varianti comuni in *CNTAP2* sono implicate nella predisposizione a ASDs ed a SLI. Il gene *CNTAP2* codifica per un membro della famiglia delle neurexine, a cui appartiene anche il gene *NRXN1*, anch'esso associato ad alcuni casi di autismo (Cap. 4 § 4.2.2).

Molti altri geni candidati, oltre agli esempi citati sopra, sono stati implicati nella predisposizione all'autismo, ma la difficile riproducibilità dei risultati ottenuti dai diversi studi non ne permette di evidenziare un chiaro coinvolgimento nell'eziologia del disturbo autistico. Questi studi riflettono l'elevata complessità ed eterogeneità genetica che caratterizza questa patologia.

4. Il cromosoma X

Il cromosoma X è un cromosoma subcentrico in cui sono stati identificati ad oggi 1098 geni. E' caratterizzato da regioni di omologia con il cromosoma Y localizzate a livello dei telomeri e definite regioni pseudoautosomiche (PAR1 e PAR2) che si comportano come regioni autosomiche in cui avvengono eventi di ricombinazione omologa.

Studi recenti suggeriscono che motivi funzionali possano aver influenzato il contenuto genico di questo cromosoma, con alcune classi di geni sovrarappresentate ed altre sottorappresentate: ad esempio i geni espressi nel muscolo e nel cervello sono particolarmente abbondanti sul cromosoma X rispetto ad altre classi di geni (Nguyen and Disteche, 2006).

Durante l'evoluzione il cromosoma X ha accumulato un numero elevato di geni coinvolti in funzioni cognitive. Infatti circa un terzo dei geni X-linked notiscono coinvolti in ritardo mentale (Skuse DH., 2005).

4.1 Il cromosoma X e l'autismo

L'autismo è caratterizzato da un rapporto tra maschi e femmine di 4:1, facendo ipotizzare un possibile coinvolgimento di geni del cromosoma X in questo disturbo.

I vari studi di linkage (*genome scans*) effettuati per l'autismo non hanno dimostrato la presenza di un locus predominante sul cromosoma X, tuttavia una modesta evidenza di linkage è stata riportata da Liu e collaboratori (Liu et al., 2001) con un MLS di 2.56 a livello del marcatore DXS1047, da Auranen e collaboratori (Auranen et al., 2002) con un MLS di 2.75 a livello del marcatore DXS7132 e da Yonan e collaboratori (Yonan *et al.*, 2003) con un MLS di 2.27 sempre nella regione distale del braccio lungo.

Oltre ai dati provenienti dalle analisi di *linkage*, a favore del coinvolgimento del cromosoma X nell'autismo, ci sono dei riarrangiamenti cromosomici riscontrati in alcuni pazienti autistici, ad esempio duplicazioni o delezioni nella regione Xp22 (Rao *et al.*, 1994; Thomas *et al.*, 1999); oppure una traslocazione X;8, in una femmina

autistica, con punto di rottura in Xp22.13, a livello del primo esone del gene GRPR (*gastrin-releasing peptide receptor*) (Ishikawa-Brush *et al.*, 1997).

4.2 Geni candidati per l'autismo che mappano sul cromosoma X

Il cromosoma X contiene un elevato numero di geni coinvolti nello sviluppo del cervello e nelle funzioni cognitive. Le differenze tra l'espressione del cromosoma X e l'espressione degli autosomi nei tessuti cerebrali sono significative (Nguyen and Disteché, 2006): nel complesso i geni dell'X hanno un'espressione nel tessuto cerebrale 2.8 volte maggiore rispetto agli autosomi (figura 10); probabilmente proprio per tale motivo mutazioni in un elevato numero di geni X-linked sono associate a ritardo mentale sindromico e non.

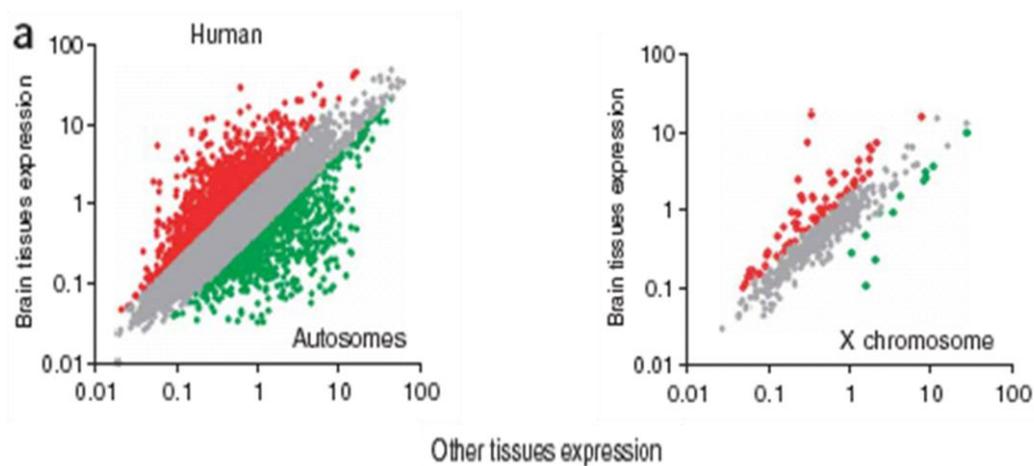


Figura 10. Espressione dei geni del cromosoma X e degli autosomi nel tessuto cerebrale rispetto agli altri tessuti. In rosso sono rappresentati i geni che hanno un'espressione due volte maggiore nel cervello, in verde quelli che hanno un'espressione due volte maggiore negli altri tessuti, in grigio i geni che hanno un'uguale espressione in tutti i tessuti. (da Nguyen and Disteché, 2006).

Tra le sindromi X-linked associate a ritardo mentale vi sono: la sindrome dell'X-fragile, dovuta a mutazioni nel gene FMR1; la sindrome di Rett, causata nella maggioranza dei casi da mutazioni nel gene MECP2. E' interessante notare che in una minoranza di casi la sindrome dell'X-fragile è associata ad autismo, mentre la

sindrome di Rett presenta alcune caratteristiche fenotipiche comuni ai disturbi dello spettro autistico.

Inoltre, mutazioni in due geni codificanti per le neuroligine, NLGN3 in Xq13 e NLGN4 in Xp22.3 (Jamain et al., 2003) sono state associati in alcuni casi di autismo.

4.2.1 MECP2

La sindrome di Rett (RTT, MIM 312750) è un disturbo neurodegenerativo progressivo che colpisce quasi esclusivamente le donne, circa 1 su 10.000-15.000, caratterizzata da ritardo mentale, movimenti stereotipati, microcefalia progressiva, aprassia e tratti autistici. La sindrome di Rett e l'autismo fanno parte entrambi dei disturbi pervasivi dello sviluppo (PDD), condividono infatti la perdita delle capacità linguistiche, sociali e cognitive e la comparsa di comportamenti ripetitivi e stereotipati, che seguono un periodo iniziale di sviluppo apparentemente normale. Mutazioni nel gene MECP2 sono coinvolte nella quasi totalità dei casi di sindrome di Rett (Amir *et al.*, 1999). Il gene mappa in Xq28 e codifica per una proteina di 486 aa, MeCP2 (methyl-CpG-binding protein 2), che funziona come repressore trascrizionale, legandosi ai dinucleotidi CpG metilati. La proteina ha due domini funzionali: uno di 85 aminoacidi, MBD (methyl-cytosine-binding domain), che si lega alle isole CpG metilate; l'altro di 104 aminoacidi, TRD (transcriptional repression domain), che interagisce con il corepressore trascrizionale SIN3A per reclutare delle deacetilasi, che agiscono a livello degli istoni. L'effetto finale è la repressione della trascrizione dei geni target (Nan *et al.*, 1998).

La proteina MeCP2 è altamente espressa a livello cerebrale, dove è importante per lo sviluppo delle cellule nervose e per la formazione delle sinapsi. Le mutazioni del gene alterano la struttura della proteina o ne riducono la quantità totale, con un conseguente effetto inibitorio sul processo di silenziamento genico.

Mutazioni nel gene MECP2 sono state riportate anche in rarissimi casi di autismo (Carney et al., 2003) e in casi di ritardo mentale non sindromico in individui di sesso

maschile (Meloni *et al.*, 2000; Zappella *et al.*, 2001). Uno studio recente ha inoltre mostrato una ridotta espressione di MeCP2 nel cervello di pazienti autistici, in particolare a livello della corteccia frontale e del giro fusiforme, in correlazione ad un'eccessiva metilazione di due isole CpG al 5' del gene (Nagarajan *et al.*, 2006).

MECP2, quindi, rappresenta un buon candidato per l'autismo per diversi motivi: svolge funzioni importanti per lo sviluppo cerebrale; è coinvolto nella sindrome di Rett e nel ritardo mentale (frequentemente associato anche all'autismo); mutazioni in questo gene sono state individuate in alcuni pazienti autistici (Lam *et al.*, 2000; Carney *et al.*, 2003). Infine è stata dimostrata la presenza di duplicazioni a carico del gene MECP2 in individui di sesso maschile caratterizzati da ritardo mentale. Le duplicazioni sono trasmesse da madri asintomatiche con XCI sbilanciato (Van Esch *et al.*, 2005; Del Gaudio *et al.*, 2006).

4.2.2 NLGN3

Il gene NLGN3 fa parte di una famiglia di geni codificanti le neuroligine, molecole di adesione cellulare presenti a livello postsinaptico. Oltre a NLGN3, localizzato in Xq13, appartengono alla famiglia i geni NLGN1 (3q26), NLGN2 (17p13), NLGN4X (Xp22.39) e NLGN4Y (Yq11.2). NLGN4X e NLGN4Y sono omologhi e sfuggono al processo di XCI.

Le neuroligine agiscono legandosi a proteine di membrana chiamate β -neurexine, e tale interazione è fondamentale per la formazione di sinapsi funzionali glutamatergiche e GABAergiche (Graf *et al.*, 2004; Chih *et al.*, 2005). Il ruolo fondamentale delle neuroligine nello sviluppo delle sinapsi e nella trasmissione di messaggi neuronali porta ad ipotizzare un loro possibile coinvolgimento nei processi di sviluppo cognitivo e negli ASD.

Recentemente sono state individuate mutazioni nei geni NLGN3 e NLGN4X in due famiglie con ASD (Jamain *et al.*, 2003). La mutazione nel gene NLGN3 causa la sostituzione di una cisteina con un'arginina (R451C) a livello del dominio esterasico,

con possibile alterazione del legame delle neuroligine con le neurexine. La mutazione nel gene *NLGL4X* determina un *frameshift* con produzione di una proteina tronca. Studi successivi su modello murino hanno dimostrato che la sostituzione R451C provoca una riduzione della proteina NLGN3 a livello neuronale del 90%, con una maggiore ritenzione all'interno del reticolo endoplasmico ed il knock out per NLGN4 nei topi ha mostrato delle alterazioni nei comportamenti sociali (Tabuchi *et al.*, 2007; Jamain *et al.*, 2008). Una ulteriore mutazione nel gene *NLGN4* è stata riscontrata in una famiglia multigenerazionale, con 13 individui maschi affetti da ritardo mentale non specifico (MRX), alcuni dei quali presentano anche autismo o ASD; la mutazione segrega con il fenotipo MR, e causa un *frameshift* con produzione di una proteina tronca (Laumonnier *et al.*, 2004). Recentemente, una delezione degli esoni 4-6 del gene *NLGN4X* è stata individuata in una famiglia con individui affetti da differenti disturbi neuropsichiatrici, facendo ipotizzare che mutazioni in *NLGN4X* siano in realtà associate a un ampio spettro fenotipico (Lawson-Yuen *et al.*, 2008). Lo *screening* di mutazioni di questi due geni, condotto su ulteriori individui affetti, ha comunque prodotto risultati contrastanti (Gauthier *et al.*, 2005; Vincent *et al.*, 2004; Yan *et al.*, 2005; Blasi *et al.*, 2006; Wermter *et al.*, 2008). L'ipotesi più plausibile, comunque, è che difetti funzionali di questi geni possano influire sullo sviluppo cognitivo e condurre ad un fenotipo autistico solo in una piccola percentuale di casi. Recentemente, numerose evidenze suggeriscono il coinvolgimento nell'autismo del gene *NRXN1*, che codifica per un partner molecolare delle neuroligine. Le neurexine rappresentano una famiglia di proteine, altamente espresse nel cervello, coinvolte nella formazione e nel mantenimento delle sinapsi attraverso il legame con le neuroligine (Graf *et al.*, 2004).

Il consorzio internazionale AGP (*Autism Genome Project Consortium*) effettuando un ampio studio su famiglie con autismo ha individuato una delezione submicroscopica, in eterozigosi, in una coppia di fratelli affetti in corrispondenza di esoni codificanti del gene *NRXN1*, delezione non presente in un campione di individui di controllo

(AGP,2007). Ulteriori lavori hanno evidenziato rare varianti strutturali e varianti di sequenza del gene *NRXN1* in individui autistici (Feng et al, 2006; Yan J. Et al., 2008; Kim HG, 2008). Di notevole interesse sarà comprendere il ruolo che ricopre questo gene nello scenario delle malattie neuropsichiatriche, infatti un recente studio evidenzia che benchè varianti strutturali siano presenti anche nella popolazione generale, microdelezioni che coinvolgono regioni esoniche in *NRXN1* sono altamente associate alla schizofrenia (Rujescu, 2009).

4.2.3 *FMR1*

Il silenziamento del gene *FMR1* rappresenta la causa genetica della sindrome dell'X-fragile (FXS, MIM 300624), una patologia monogenica X-linked che colpisce maggiormente gli individui di sesso maschile e responsabile di circa il 10% dei casi di ritardo mentale X-linked.

Normalmente il gene *FMR1* contiene tra 6 e 53 ripetizioni del codone CGG. Negli individui affetti dalla sindrome dell'X fragile si ha un'amplificazione del numero di queste triplette fino a più di 230 ripetizioni. Questo grado di espansione provoca la metilazione delle citosine nel promotore del gene *FMR1*, con conseguente silenziamento dell'espressione del gene *FMR1*. La metilazione del *locus* *FMR1*, che mappa in Xq27.3, provoca in questo punto la costrizione e la fragilità del cromosoma X, fenomeno che dà il nome alla sindrome.

La metilazione del gene *FMR1* porta all'abolizione della produzione della proteina per cui il gene *FMR1* codifica, detta FMRP (fragile X-mental retardation protein). FMRP è una proteina che lega gli RNA (RNA-binding protein) e viene espressa soprattutto nei testicoli e nel cervello. I targets della FMRP sono messaggeri codificanti importanti proteine neuronali, di cui regola alcuni aspetti essenziali, quali il trasporto lungo i dendriti verso le sinapsi, ma anche la stabilità dei trascritti per la traduzione in proteine.

I sintomi neurocomportamentali della sindrome presentano vari tratti fenotipici in comune con l'autismo, tra cui comportamenti ripetitivi e stereotipati, il ritardo nello sviluppo del linguaggio. Secondo le più recenti stime circa il 4% dei casi di autismo sono associati alla FXS, mentre la prevalenza dell'autismo nella popolazione con FXS varia tra il 18% e il 33% (Rogers *et al.*, 2001).

La presenza dell'espansione della tripletta nel gene FMR1 è stata esclusa negli individui affetti da autismo appartenenti al campione utilizzato nel nostro studio.

4.3 Inattivazione del cromosoma X (XCI)

Il processo di inattivazione del cromosoma X è uno dei migliori esempi di regolazione epigenetica, messo in atto dalle cellule per evitare che ci sia uno sbilanciamento dell'espressione dei geni del cromosoma X tra uomo (XY) e donna (XX). Questo processo esemplifica anche la plasticità degli eventi epigenetici, dato che durante l'oogenesi si verifica una riattivazione dell'X.

Si tratta di un meccanismo di *dosage compensation*, tramite il quale uno dei due cromosomi X va incontro a silenziamento genico, e inizia nelle prime fasi dello sviluppo embrionale (Lyon, 1961). La scelta del cromosoma X da inattivare in ciascuna cellula è casuale, ma una volta effettuato, il silenziamento è ereditato nelle successive divisioni cellulari (figura 11). Nel caso di triploidia o tetraploidia saranno inattivati tutti gli X tranne uno (Migeon *et al.*, 1979).

Ogni femmina adulta in teoria avrà un 50% di cellule somatiche con l'X paterno attivo e un 50% di cellule somatiche con l'X materno attivo; il rapporto tra Xp attivo e Xm attivo (*X inactivation pattern*) seguirà una distribuzione gaussiana in cui i valori estremi saranno rari.

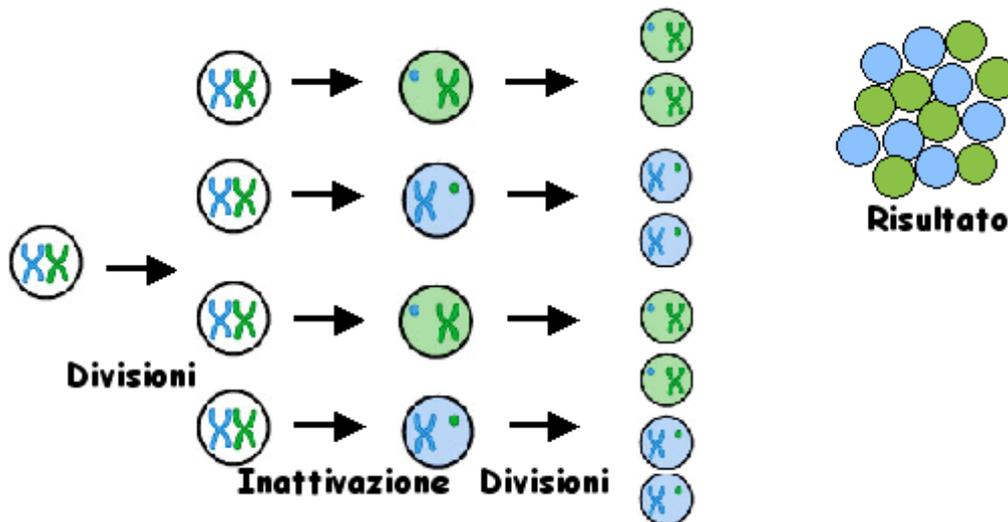


Figura 11. L'X-inactivation è un processo casuale che inizia nelle prime fasi dello sviluppo embrionale e rende la donna adulta un mosaico per quanto riguarda il cromosoma X.

Le principali differenze tra X attivo e X inattivo sono: l'Xi replica tardivamente durante la fase S della mitosi, presenta ipoacetilazione delle proteine istoniche, dà luogo ad un ammasso di eterocromatina detto Corpo di Barr e ha una localizzazione nucleare periferica.

Il processo di XCI si svolge in tre fasi: inizio, diffusione e mantenimento. Nella prima si ha la selezione del cromosoma X da inattivare, grazie alla presenza dell'*X Inactivation Centre* (XIC) che agisce in *cis* e mappa in Xq13. In questa regione il gene XIST (X Inactive Specific Transcript), viene espresso solo dall'X inattivo e codifica per un RNA non tradotto (Brown *et al.*, 1991). Nelle cellule staminali embrionali entrambi i cromosomi X esprimono XIST a bassi livelli; successivamente l'RNA XIST ricopre il cromosoma X che verrà inattivato, permettendo la diffusione dell'inattivazione a partire da XIC. L'accumulo di XIST in *cis* recluta specifici fattori proteici di silenziamento e porta alla riorganizzazione della cromatina in uno stato inattivo, caratterizzato da ipoacetilazione degli istoni e metilazione delle isole CpG. L'RNA attorno all'X attivo, invece, rimane visibile solo per un breve periodo di tempo, dopodiché il gene XIST sarà metilato su questo cromosoma, rimanendo in uno stato represso. Per il mantenimento dell'inattivazione è importante la continua

trascrizione di XIST nelle cellule adulte, ma anche la metilazione delle isole CpG sull'Xi e la sua replicazione tardiva sono importanti fattori di regolazione dell'inattivazione.

Non tutti i geni del cromosoma X vanno incontro ad inattivazione: alcuni infatti sono espressi da entrambi i cromosomi X e sono localizzati principalmente a livello delle regioni pseudoautosomiche PAR1 (2.7 Mb) e PAR2 (330 Kb) (Carrel *et al.*, 1999). Inoltre i geni che sfuggono l'inattivazione che non appartengono a queste due regioni hanno generalmente omologhi funzionali sul cromosoma Y, per cui probabilmente non è necessario il meccanismo di compensazione. Esistono anche altri geni, al di fuori di queste regioni, che sfuggono all'inattivazione (Carrel & Willard, 2005). Nello specifico, solo il 65% dei geni X-linked va incontro ad una completa inattivazione nelle femmine normali, mentre un 20% è inattivato in alcune ma non in tutte le cellule e un 15% sfugge costantemente all'inattivazione (figura 12).

Sono necessarie spiegazioni differenti per questa mancata inattivazione, come un'assenza di conseguenze per lo squilibrio nell'espressione di tali geni tra uomo e donna oppure un'utilità della doppia dose di essi per alcune specifiche funzioni femminili.

dell'organismo (Therman and Patau, 1974); al contrario, nel caso di traslocazioni bilanciate X-autosoma, sarà l'X normale ad essere inattivato, al fine di mantenere una euploidia funzionale (Schmidt and Du Sart, 1992).

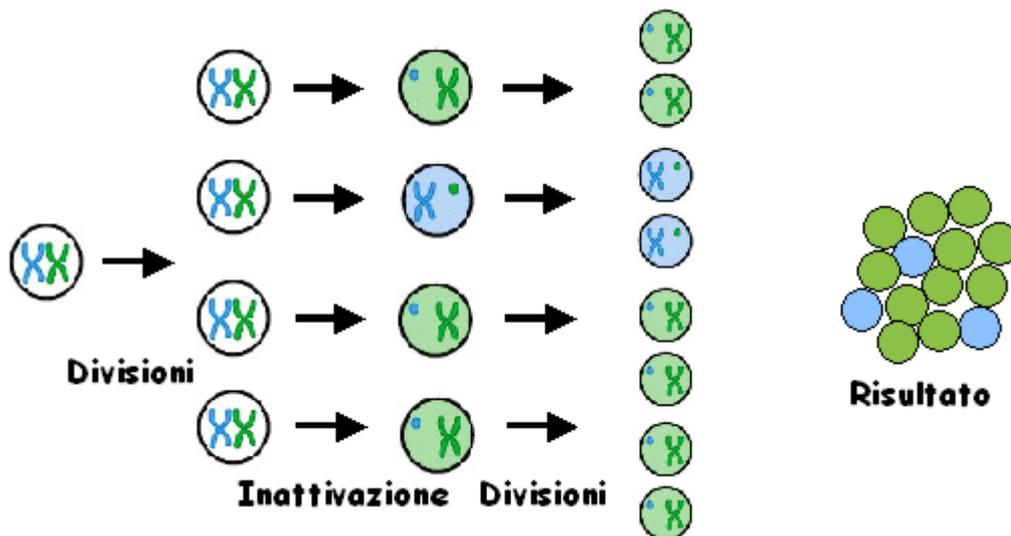


Figura 13. In alcuni casi l'X-inactivation non è più random ma è sbilanciata, provocando uno squilibrio tra cellule che esprimono l'Xp e cellule che esprimono l'Xm.

Processi selettivi simili agiscono anche in individui con mutazioni in geni X-linked, infatti sono state descritte varie sindromi (ritardo mentale legato all'X-XLMR, sindrome di Barth) (Plenge *et al.*, 2002; Orstavik *et al.*, 1998) in cui le femmine portatrici asintomatiche hanno un alto grado di sbilanciamento a favore dell'X normale, che viene espresso nella maggior parte delle cellule. Tale fenomeno può essere il risultato di una selezione contro le cellule che esprimono la mutazione deleteria, probabilmente dovuta alla ridotta vitalità o proliferazione delle stesse, o a meccanismi più complessi che agiscono durante lo sviluppo (Muers *et al.*, 2007).

In alternativa, l'inattivazione sbilanciata dell'X può essere un processo del tutto casuale, dovuto al basso numero di cellule presenti al momento dell'inattivazione (Fialkow, 1973).

Infine, a volte l'inattivazione sbilanciata a favore di un allele mutante può essere la causa del manifestarsi di una patologia X-linked recessiva in un sottogruppo di

individui di sesso femminile, come nel caso dell'anemia sideroblastica e della sindrome dell'X-fragile (Cazzola *et al.*, 2000).

Studi per definire la relazione tra l'XCI e rare mutazioni deleterie sono state effettuate da Plenge *et al.*, in cui utilizzando 24 famiglie con 20 diverse forme di XLMR (X-linked Mental retardation), è stata rilevata la presenza di X-skewed in una percentuale alta delle madri portatrici asintomatiche di mutazioni. Lo studio dimostra che l'inattivazione sbilanciata può essere riscontrata nelle cellule del sangue di donne portatrici di mutazioni X-linked e che il cromosoma inattivato contiene sempre la mutazione. Solitamente è presente un differente grado di associazione tra lo stato di carrier e l'inattivazione sbilanciata per le diverse forme di XLMR (Plenge *et al.*, 2002) (figura 14).

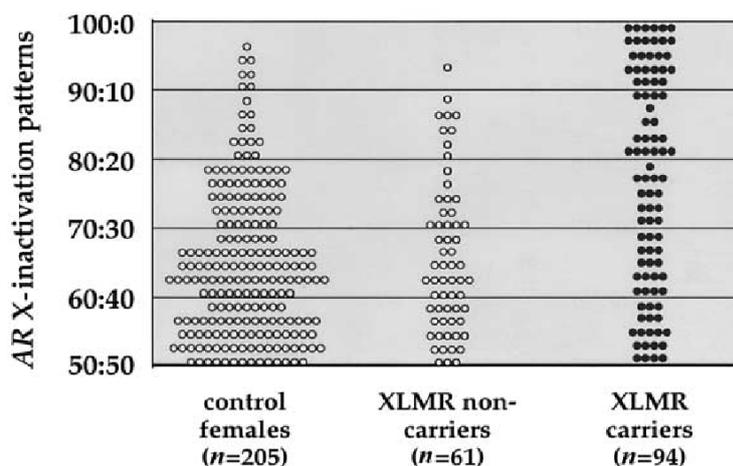


Figura 14. Profilo di XCI in madri appartenenti a famiglie con XLMR (carriers e non carriers) e nei relativi controlli (da Plenge *et al.*, 2002).

Ultimamente, lo studio di Talebizadeh, ha indagato la frequenza dello stato di X-skewing nell'autismo. Infatti sono presenti casi di ritardo mentale in individui autistici e sono state identificate alcune rare mutazioni in geni X-linked. Partendo da queste osservazioni, Talebizadeh e collaboratori (Talebizadeh *et al.*, 2005) hanno condotto uno studio su 30 ragazze con autismo e su 35 controlli correlati per età. I risultati hanno evidenziato che il campione di sesso femminile affetto da autismo

aveva una più alta percentuale di XCI sbilanciata rispetto ai controlli. Inoltre il 50% delle madri delle pazienti con XCI sbilanciata mostrava a sua volta XCI sbilanciata (figura 15).

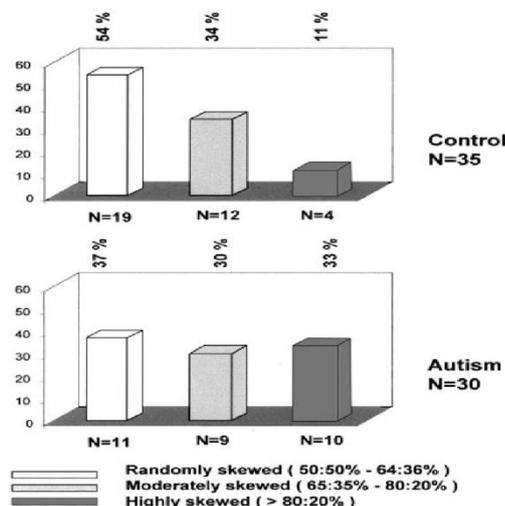


Figura 15. Profilo di XCI in un campione di probande con autismo e nei relativi controlli (da Talebizadeh et al., 2005).

4.5 XCI sbilanciata ed invecchiamento

Tutti i tessuti di uno stesso individuo dovrebbero avere lo stesso grado di XCI, poiché questa avviene prima che le cellule siano indirizzate verso particolari linee; in realtà è stata osservata una variazione intraindividuale nel pattern di XCI a seconda dei tessuti esaminati. Nelle cellule del sangue ad esempio, con l'aumentare dell'età della donna, può aversi una forte deviazione dal normale rapporto 50:50, in particolare dopo i 60 anni (Busque *et al.*, 1996). Le ragioni di tale fenomeno non sono ben note, ma un'ipotesi plausibile è che sia dovuto alla senescenza del pool di cellule staminali emopoietiche, che per fluttuazione casuale comporterebbe una diminuzione del numero di cellule progenitrici con l'uno o l'altro cromosoma X attivo; si è anche ipotizzato che i cambiamenti ormonali successivi alla menopausa possano contribuire a questo sbilanciamento (Kristiansen *et al.*, 2005). Una possibile conseguenza dello *skewing* correlato all'età è la comparsa, in età avanzata, di disturbi legati all'X in femmine portatrici. Un'altra spiegazione di questo aumento dello sbilanciamento con

l'età è una perdita della metilazione allele-specifica. Il saggio comunemente usato per valutare lo stato di inattivazione del cromosoma X si basa infatti sull'uso di enzimi di restrizione sensibili alla metilazione: la metilazione tende ad essere persa nelle cellule che invecchiano, sia *in vivo* che *in vitro*, e in un individuo con un'inattivazione del 50:50, la perdita a livello di uno solo dei due cromosomi X può portare ad un apparente sbilanciamento (Hatakeyama *et al.*, 2004).

4.6 Valutazione dello stato di inattivazione del cromosoma X

Per misurare lo stato di inattivazione del cromosoma X in passato si utilizzava l'espressione allele-specifica della glucosio 6-fosfato deidrogenasi (G6PD), ma tale metodo è limitato dal fatto che il polimorfismo in esame è raro nella popolazione caucasica (~ 1% di eterozigotità).

Il metodo più comunemente utilizzato per distinguere l'Xi dall'Xa consiste nella valutazione del grado di metilazione dei due cromosomi a livello di isole CpG adiacenti a siti polimorfici. A tal scopo si possono utilizzare i loci dell'ipoxantina-guanina fosforibosiltransferasi (HPRT) (Lock *et al.*, 1987) o della fosfoglicerato chinasi (PGK) (Singer-Sam *et al.*, 1990); gli RFLPs adiacenti a queste isole CpG sono infatti eterozigoti nel ~ 35%-40% delle donne. Un altro marker utilizzato per l'analisi dell'X-inactivation è il microsatellite DXS255, che risulta essere più polimorfico sia di HPRT che di PGK (Boyd and Fraser, 1990)

Il saggio attualmente più utilizzato è l'HUMARA assay, che sfrutta il locus del recettore degli androgeni (AR) in Xq11.2. Nel primo esone del gene è presente una ripetizione trinucleotidica (CAG) altamente polimorfica (~90%), nelle vicinanze dei siti di taglio di enzimi di restrizione sensibili alla metilazione (HpaII o HhaI), metilati solo sull'X inattivo (Allen *et al.*, 1992). Basandosi sulla digestione differenziale del DNA genomico di partenza in presenza ed in assenza dell'enzima sensibile alla metilazione, in seguito ad amplificazione del locus AR con primers appropriati è quindi possibile distinguere gli alleli metilati e non (Cap. 8 § 8.6).

5. *ASMI: gene candidato funzionale per l'autismo*

5.1 La melatonina nell'autismo

Numerosi studi hanno indicato un possibile ruolo della melatonina, un neurormone epifisario, nello spettro del disturbo autistico (ASD). E' evidente la presenza di familiarità per quanto riguarda i livelli di secrezione di questo ormone (Hallam KT), che regola la plasticità neuronale ed è coinvolto nel ritmo circadiano foto periodico (Liu C; El-sherif Y; Von Gall C). Ultimamente numerosi studi hanno dimostrato l'alta prevalenza di disturbi del sonno e ritmi circadiani alterati in individui autistici, suggerendo anomalie nella fisiologia della melatonina (Malow BA, 2004). Inoltre, sembra che la somministrazione orale di tale ormone ad affetti autistici possa migliorare i ritmi sonno-veglia (Hayashi E.,2000; Jan JE. et al., 1996). Ulteriori analisi hanno evidenziato una sostanziale riduzione dei livelli notturni di melatonina prodotta in individui autistici rispetto ai controlli di uguale età e sesso (Tordjam S. et al, 2005). Questi dati risultano di particolare interesse se si considera che la melatonina è prodotta a partire dalla serotonina, e numerosi sono gli studi che riportano alterazioni nel sistema serotoninergico centrale e periferico (Chroonenberghs J. et al., 2000; Cook E. et al, 1996).

ASMT (Acetyl Serotonin O-metyltransferase) codifica per l'ultimo enzima della via biosintetica della melatonina (figura 16). In questa via il triptofano è convertito in serotonina, che viene quindi N-acetilata, e la N-acetil serotonina subisce la O-metilazione da parte di ASMT, che determina la formazione della melatonina.

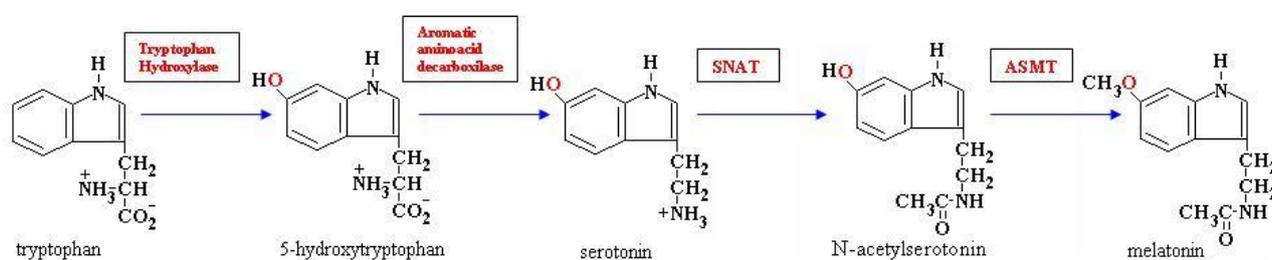


Figura 16. Schematizzazione della via biosintetica della melatonina.

Data l'importante funzione nella biosintesi della melatonina, il gene ASMT rappresenta un buon gene candidato funzionale per l'autismo. Il gene ASMT mappa all'estremità dei bracci corti dell'X e dell'Y (Xp22.3 o Yp11.3) nella regione pseudoautosomica PAR1 (figura 17). Questa è una regione di completa omologia dei cromosomi X ed Y, caratterizzata da eventi di ricombinazione meiotica.

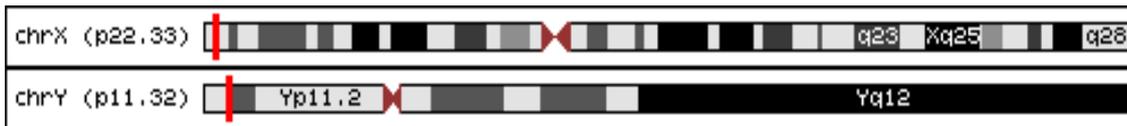
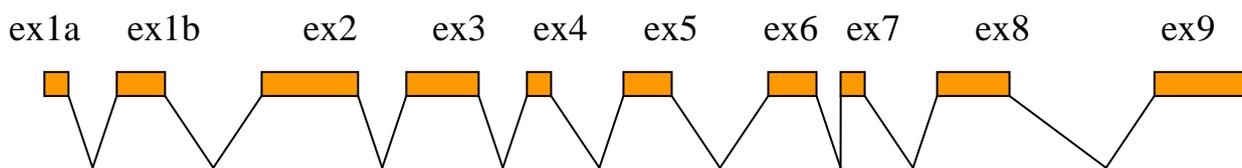


Figura 17: Il gene ASMT mappa nella regione pseudoautosomica PAR1 dei cromosomi X ed Y.

Il gene copre un intervallo di 35 kb ed è caratterizzato da due differenti promotori (figura 18). Contiene 9 o 10 esoni a seconda di quale promotore sia attivo. Inoltre, presenta prodotti di splicing alternativo degli esoni 6 e 7 che possono determinare tre peptidi differenti. Il prodotto proteico con l'isoforma più lunga (+ex6, +ex7) codifica per un prodotto di 373 aa, mentre il prodotto di splicing alternativo senza l'esone 6, codifica per un prodotto di 345 aa e rappresenta il trascritto maggiormente espresso. Infine, l'ultima isoforma manca dell'esone 6 e 7, e codifica per una proteina di 298 aa. Lo skipping di tali esoni non condiziona il modulo di lettura. Sembra che il promotore B sia responsabile della produzione di tutti e tre le varianti di splicing del gene ASMT a differenza del promotore A che sembra attivo esclusivamente nella retina. (Rodriguez et al. 1994).

Trascritto con il promotore A



Trascritto con il promotore B

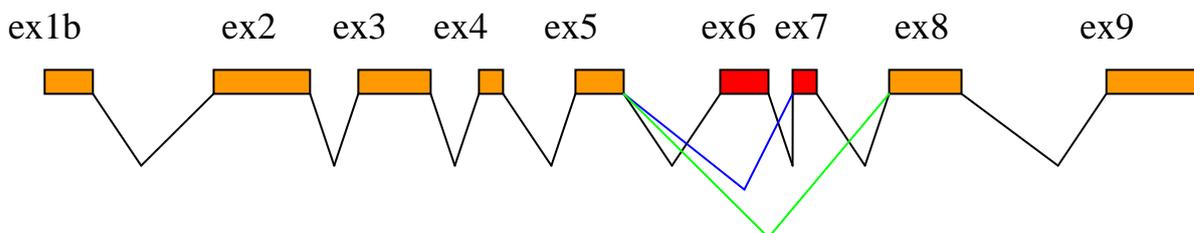


Figura 18. Organizzazione genomica del gene ASMT. Sono indicati i due possibili trascritti in base all'attivazione del promotore A o B; in rosso gli esoni coinvolti nello splicing alternativo.

5.2 Studi di associazione e screening di mutazione del gene ASMT

L'evidenza di un possibile coinvolgimento del gene ASMT proviene da uno studio in cui si è stato effettuato un'analisi di mutazioni e studi di associazione in un ampio campione di individui affetti da autismo (Melke et al.). In seguito allo screening di mutazioni condotto su tutti gli esoni di ASMT in 250 individui, sono state identificate diverse rare varianti genetiche :

1. Una variante di splicing (IVS5+2T>C) individuata in 3 famiglie con autismo
2. Quattro varianti non sinonime (N17K, K81E, G306A, L326F); di cui l'ultima in tre famiglie.
3. Due varianti sinonime (N167N, Q205Q)

I nove individui con le varianti identificate, presentano anche iperattività ed alcuni hanno disturbi del sonno. La segregazione di ogni variante per ciascuna famiglia è schematizzata in figura 19. Le varianti N17K (rs17149149) and L326F sono state individuate anche in un campione controllo, mentre la variante di splicing (IVS5+2T>C) è presente in due famiglie ed in nessuno dei 426 controlli analizzati.

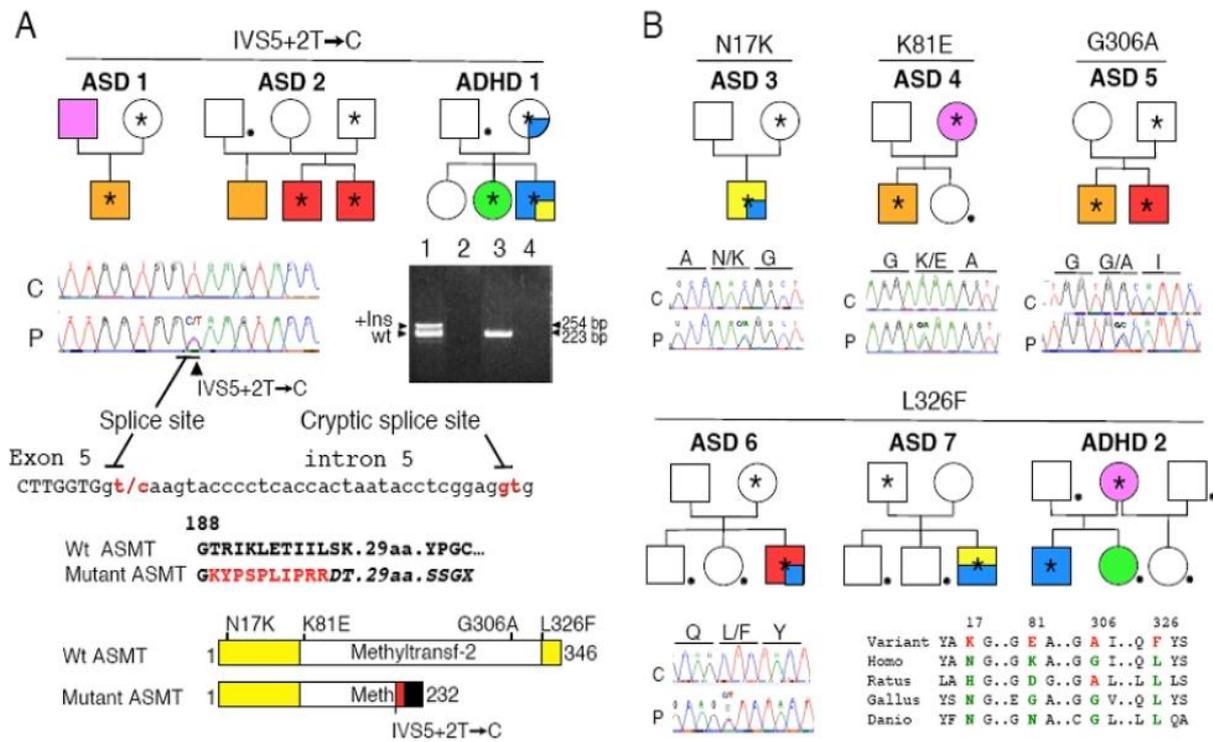


Figure 19. Varianti non sinonime in ASMT (A) Le tre famiglie con la variante di splicing IVS5+2T>C; gli studi di RT-PCR sono stati condotti amplificando il cDNA di BLCL tra l'esone 4 e 6 di ASMT del probando della famiglia ASD1. In figura, la lane 1 e 2 appartengono al probando (lane 1 +RT, lane 2 -RT). Si osserva, oltre al trascritto wt anche il trascritto in cui è presente un'inserzione di 31bp del trascritto originario dovuto alla variante di splicing IVS5+2T>C; si osserva il trascritto wt nella lane 3 (lane 3 +RT, lane 4 -RT). L'inserzione di 31bp dovuta alla creazione di un sito accettore di splicing alternativo, a valle dell'esone 5, porta alla formazione di una sequenza proteica addizionale indicata in rosso; il frame shift che ne deriva causa la terminazione prematura della proteina che manca del dominio metil-transferasico. (B) Sono indicate le famiglie in cui sono state individuate ulteriori varianti non sinonime e la loro conservazione tra le differenti specie. I colori nei pedigree delle distinte famiglie indicano aspetti relativi alla diagnosi: ritardo mentale moderato (arancione) o grave (rosso), sindrome di Asperger o autismo di tipo alto funzionamento (giallo), ADHD (blù) e depressione (rosa). Il probando ASD3 rientra nei criteri diagnostici sia per ADHD che per autismo ad alto funzionamento, quello ASD7 sia per ADHD che per la sindrome di Asperger. L'asterisco indica la presenza della mutazione mentre il punto l'assenza di DNA per l'analisi.

L'analisi biochimica ha mostrato che la variante IVS5+2T>C porta alla creazione, rispetto al sito canonico, di un sito accettore di splicing a valle dell'esone 5, con la conseguente formazione di un'inserzione di 31bp del trascritto. Il prodotto proteico che ne risulta è una proteina tronca che manca del dominio metil-transferasico. Inoltre, ulteriori dati dimostrano come le varianti IVS5+2T>C and L326F sono associate a bassi livelli dell'attività enzimatica di ASMT e della concentrazione di melatonina (figura 20)

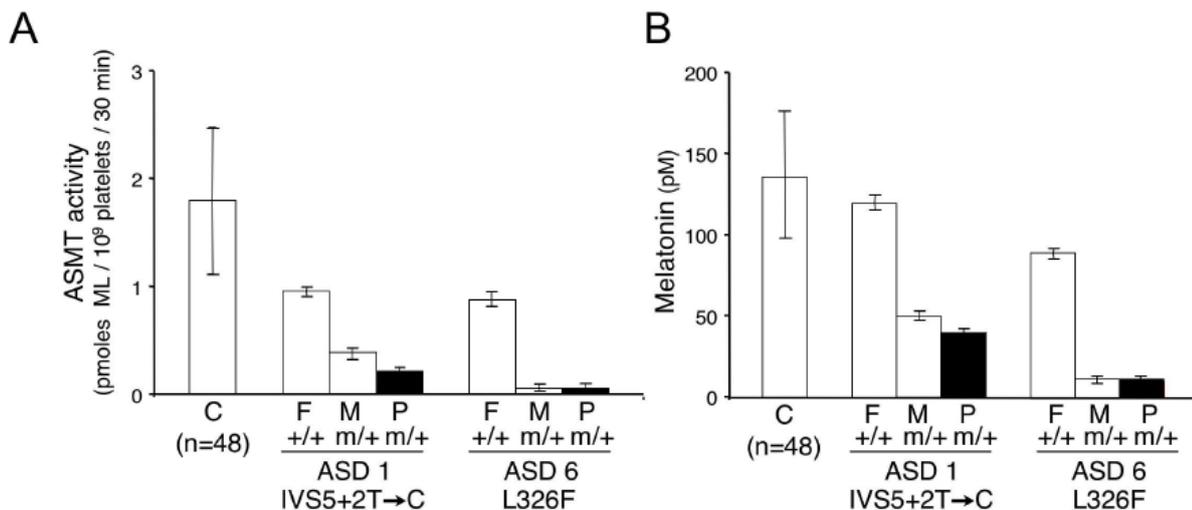


Figure 20. Attività enzimatica della proteina ASMT e concentrazione di melatonina nelle famiglie con le mutazioni individuate. (A) L'attività enzimatica della proteina ASMT è stata misurata sulle piastrine dei membri della famiglia ASD1 ed ASD6 che sono quelle che portano rispettivamente la variante di splicing IVS5+2T>C e la missenso L326F. (B) E' indicato negli stessi individui la concentrazione di melatonina nel sangue. C: controllo; F: padre; M: madre; P: individuo affetto.

Ulteriori evidenze sperimentali del contributo del gene ASMT al fenotipo autistico provengono dagli studi di associazioni di marcatori polimorfici lungo il gene. In particolar modo è stata effettuata un'analisi per stabilire i blocchi di *linkage disequilibrium* nella regione in cui mappa il gene ASMT, e sono stati selezionati 13 SNPs per catturare la maggior diversità aplo-tipica nei 3 blocchi di LD (figura 21) (Melke et al. 2007)

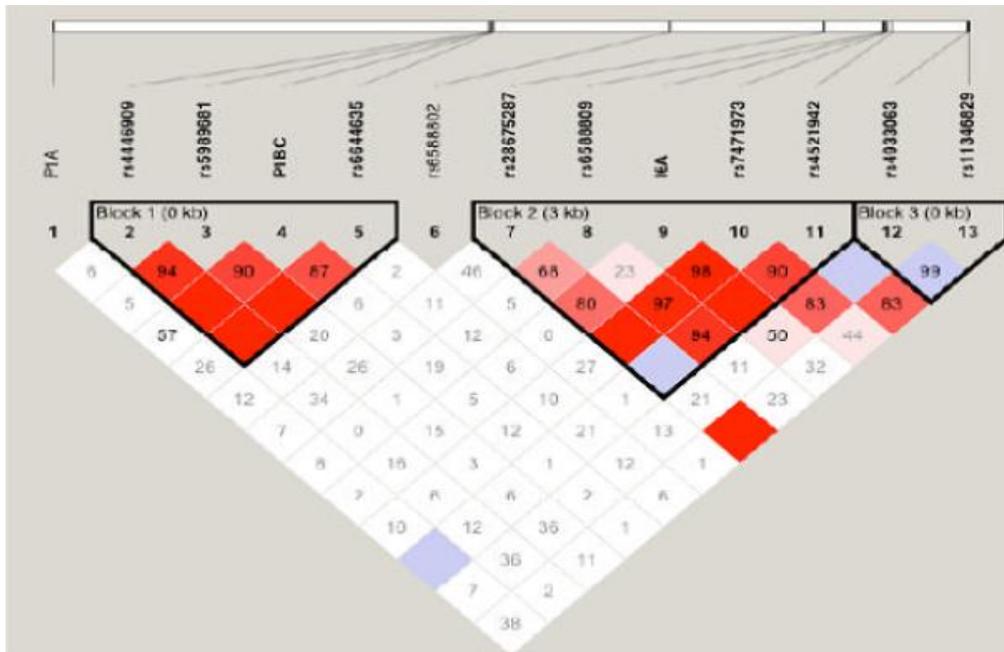


Figura 21: struttura dei blocchi di LD del gene ASMT

I risultati ottenuti evidenziano che le frequenze alleliche di due SNPs (rs4446909 e rs5989681) nel promotore B sono significativamente diverse tra il campione di 278 individui con autismo e quello di 255 controlli. In particolare modo, l'aplotipo H1 (GGGC) del primo blocco di LD è più frequente nel campione autistico rispetto al gruppo controllo, mentre l'aplotipo H3 (ACGC) è più frequente nei controlli (tabella 3).

Frequencies of the polymorphisms located in *ASMT* promoter B in ASD patients and controls

	ASD (n= 278)	Controls (n= 255)
<i>SNPs</i>		
rs4446909		
f(G)	0.77	0.70
P value (P_c value ^a)	0.006 (0.10)	
OR [95% CI] ^b	1.5 [1.1–2]	
rs5989681		
f(G)	0.73	0.65
P value (P_c value)	0.007 (0.12)	
OR [95% CI]	1.4 [1.1–2]	
<i>P1BC</i>		
f(G)	0.90	0.90
P value	0.78	
rs6644635		
f(C)	0.65	0.63
P value	0.66	
<i>Haplotypes^c</i>		
H1 GGGC	0.36	0.27
P value (P_c value)	0.002 (0.04)	
H2 GGGT	0.26	0.28
P value	0.54	
H3 ACGC	0.21	0.29
P value (P_c value)	0.005 (0.08)	
H4 GGAT	0.097	0.09
P value	0.78	
H5 GCGC	0.055	0.057
P value	0.88	

Abbreviations: ASD, autism spectrum disorders; CI, confidence interval; OR, odds ratio; SNP, single nucleotide polymorphism.

^a P_c value: Significance levels corrected for multiple comparisons using a step-down permutation procedure (comprising 100,000 permutations).

^bOdds ratio: major allele vs minor allele.

^cHaplotype using rs4446909, rs5989681, P1BC and rs6644635.

Results with significance < 0.05 are indicated in bold.

Tabella 3: frequenze alleliche dei polimorfismi del promotore B nei pazienti e nei controlli.

Il coinvolgimento dell'aplotipo H1 (GGGC) è emerso anche dall'analisi TDT (Transmission Disequilibrium Test), in cui è statisticamente significativo la trasmissione di questo aplotipo dai genitori ai figli affetti rispetto agli altri aplotipi (figura 21).

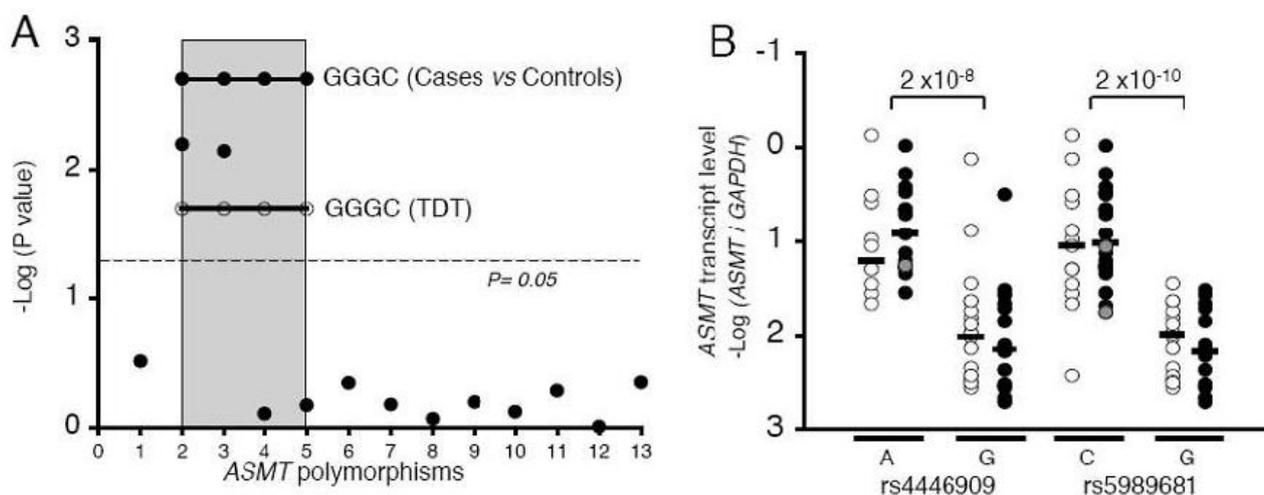


Figura 21: (A) Schema riassuntivo dei valori di P value ($-\log_{10}$) per tutti i marcatori analizzati utilizzando il caso/controlli o il TDT. I marcatori (in successione: rs4446909, rs5989681, P1BC, rs6644635) nella regione in grigio sono quelli del blocco di LD appartenente al promotore B. Gli SNPs che identificano l'aplotipo di rischio H1 (GGGC) sono indicati come una linea continua (cerchi neri nel caso del caso/controllo e cerchi vuoti nel caso del TDT). (B) Quantificazione del trascritto di ASMT in relazione ai genotipi dei marcatori rs4446909 e rs5989681 (A: genotipi A/A o A/G; C: genotipi C/C o C/G; G: genotipi G/G). I cerchi neri indicano gli individui autistici, mentre quelli bianchi indicano i controlli. Gli studi sono stati condotti su RT-PCR di linee linfoblastoidi di 38 individui autistici e 29 controlli.

L'analisi successiva dei livelli di espressione del gene ASMT in relazione agli aplotipi del promotore B, ha messo in risalto una stretta correlazione tra i bassi livelli del trascritto e l'allele G degli SNP rs4446909 e rs5989681 trovati precedentemente associati all'autismo (figura 21).

Questi dati suggeriscono che esistono fattori genetici del gene ASMT che contribuiscono a particolari aspetti del fenotipo autistico come disturbi del sonno e ritmi circadiani alterati. Le evidenze sperimentali riportate in questo studio risultano interessanti, benché le varianti genetiche incontrate sono in un numero limitato di casi e l'aplotipo di rischio non è il solo fattore che spiega la riduzione dell'attività enzimatica.

6. Studi di associazione ad alta densità ed analisi di CNVs in due loci di suscettibilità all'autismo: 2q24-32 e 7q21-32

6.1 I Loci di suscettibilità all'autismo AUTS1 e AUTS5

I risultati degli *screening* genomici condotti dal consorzio IMGSAC (IMGSAC, 1998; IMGSAC, 2001; Lamb et al., 2005) hanno portato ad individuare due principali loci di suscettibilità per l'autismo nella regione 2q21-33, locus AUTS5, e 7q21-32, locus AUTS1 (Cap. 3 § 3.2). In Figura 22, è riportato il profilo di MLS sul cromosoma 2q e 7q relativo alle 219 famiglie multiplex analizzate nell'ultimo *genoma scan* condotto dall'IMGSAC (Lamb et al., 2005).

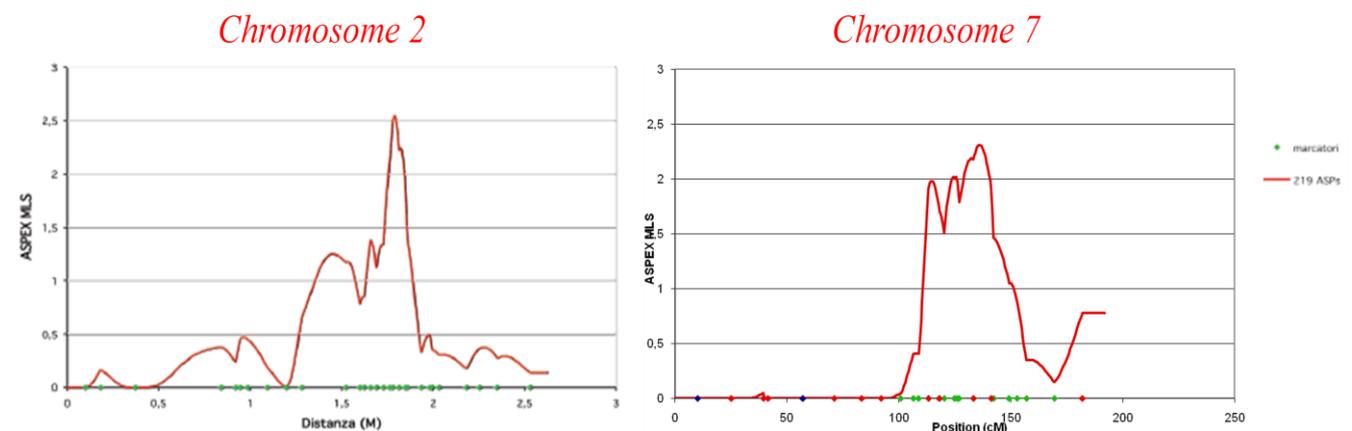


Figura 22: risultati di linkage per il cromosoma 2 e 7; è riportato per ciascun cromosoma il profilo di MLS relativo ai 219 ASPs (affected sib-pairs) (Lamb et al., 2005)

Risultati positivi che supportano il coinvolgimento del locus AUTS1 e AUTS5 si sono ottenuti anche in altri studi indipendenti (Shao et al., 2002; Buxbaum et al., 2001, Auranen et al., 2002). In seguito ai risultati ottenuti si è proceduto alla caratterizzazione dei due loci di suscettibilità all'autismo. In particolar modo il nostro gruppo all'interno del consorzio IMGSAC si è occupato del locus al cromosoma 2q. Benché i risultati di linkage abbiano dato buoni risultati, l'identificazione dei geni responsabili della malattia non risulta affatto semplice poiché le due regioni di linkage risultano essere molto ampie (~40 Mb ciascuna), ed ognuna contiene un numero elevato di geni: 190 geni per la regione di linkage al cromosoma 2q e 270 geni per la regione al cromosoma 7q. Infatti, la difficoltà che si incontra negli studi di

linkage per malattie multifattoriali è che la localizzazione dei risultati per loci ad effetto moderato non è precisa ed il locus reale è all'interno di un'ampia regione genomica. Per la caratterizzazione del locus AUTS1 e AUTS5 sono stati utilizzati due principali approcci:

1. Analisi di mutazioni nelle regioni codificanti, regolatorie e promotrici di geni che rappresentano dei buoni geni candidati per l'autismo.
2. *Genotyping* di SNPs allo scopo di effettuare *LD mapping* in geni candidati ed identificare per via indiretta varianti di rischio per l'autismo.

L'analisi condotta dal consorzio IMGSAC di numerosi geni candidati sulle due regioni di linkage non ha portato all'identificazione di nessuna mutazione chiaramente coinvolta nel disturbo autistico.

6.2 *Genotyping su larga scala di SNPs tramite il sistema Illumina*

Sulla base di questi presupposti il nostro gruppo di ricerca ha intrapreso un approccio sistematico che consiste in uno studio di SNPs ad alta densità all'interno della regione cromosomica 2q e 7q, allo scopo di effettuare studi di *LD mapping* e identificare per via indiretta varianti di rischio per l'autismo. Questo approccio è stato reso possibile dalla recente disponibilità di risorse pubbliche quale l'*International Haplotype Mapping Consortium* (HapMap), che sta caratterizzando i pattern più comuni di varianti di sequenza presenti nel genoma umano, e dalle nuove tecnologie che permettono una tipizzazione rapida ed efficiente di SNPs su larga scala. Inoltre, sfruttando la grande quantità di dati relativi all'analisi di SNPs è possibile mappare a livello genomico, su larga scala e a basso costo, la nuova fonte di variabilità rappresentata dalle CNVs.

Per ottimizzare la selezione degli SNPs, è stata effettuata un'approfondita analisi bioinformatica utilizzando i dati disponibili in HapMap per valutare il profilo di LD e la struttura a blocchi di aplotipi nella regione di interesse.

Abbiamo utilizzato due stage di tipizzazione dei marcatori basati sulla fase I (rel 13) e fase II (rel 22) di HapMap.

In totale, tra le due fasi sono stati tipizzati 3002 SNPs per ciascun cromosoma. Mediante questo approccio abbiamo analizzato 173 geni per il cromosoma 2 e 270 geni per il cromosoma 7. Il campione di pazienti autistici utilizzato per il nostro esperimento è stato di 126 famiglie multiplex per il cromosoma 2, e 127 famiglie multiplex per il cromosoma 7, scelte tra le 293 famiglie multiplex del nostro consorzio, in base alle famiglie che contribuivano maggiormente al picco di linkage sul cromosoma 2 e 7 rispettivamente. Solo uno dei due pazienti affetti per famiglia è stato utilizzato per il genotyping dei marcatori selezionati. Il campione controllo utilizzato è costituito da 188 individui donatori anonimi appartenente alla collezione ECACC (European Collection of Cell Culture).

Il sistema utilizzato per la tipizzazione degli SNPs è quello della tecnologia Illumina (*GoldenGate® assay*) (Cap. 8 § 8.12)

Una schematizzazione dei marcatori ed individui utilizzati nel presente studio è presente in tabella 4.

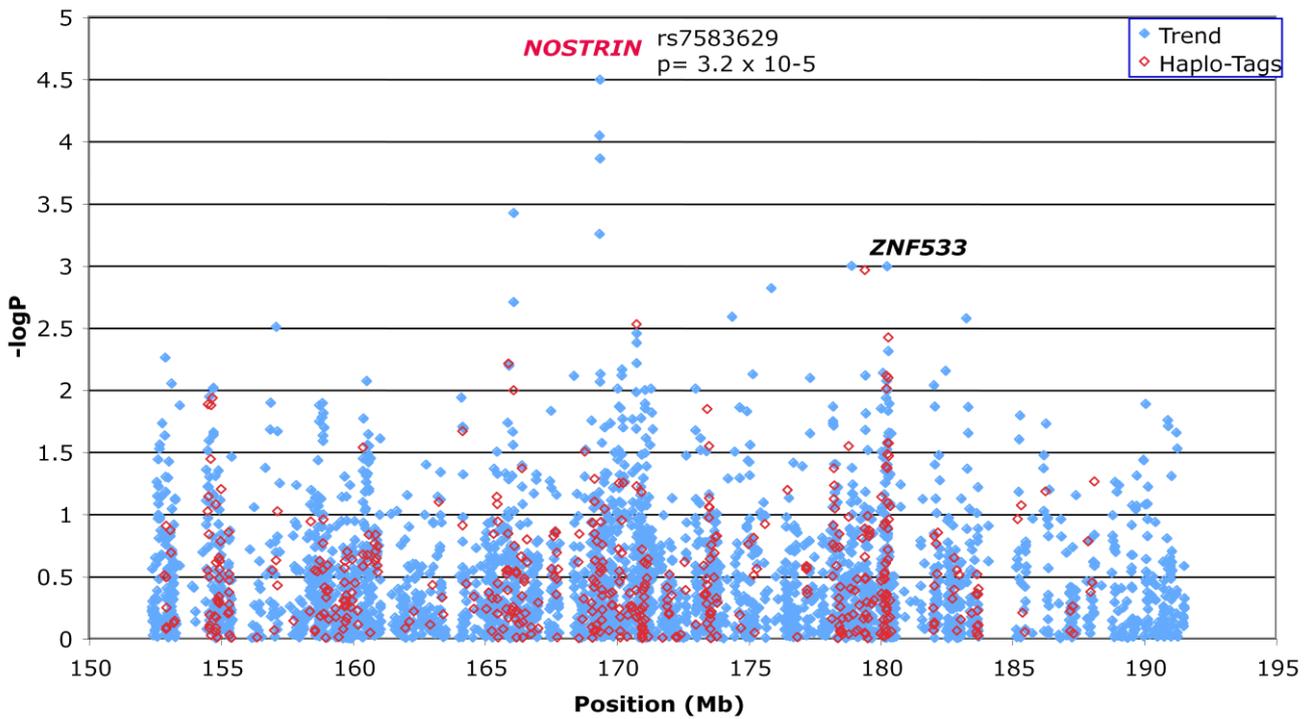
	<i>Cromosoma 2q</i>	<i>Cromosoma 7q</i>
SNPs	2.860	2.808
Individui affetti	126	127
genitori	245	244
controlli	188	188
Tot. individui	559	559

Tabella 4: sono schematizzati il numero finale di SNPs ed individui utilizzati nello studio di associazione e per l'analisi di CNVs. Sono stati eliminati tutti i marcatori: con una MAF (Minor Allele Frequency) < 0.05, con più di un errore mendeliano, un tasso di genotyping minore del 90%, problemi nella formazione dei clusters di genotipi, marcatori che nei controlli non rispettano l'equilibrio di Hardy-Weinberg ($p < 0.001$).

6.3 Studi di associazione ad alta densità dei loci AUTS1 e AUTS5

Gli studi di associazione sono stati condotti per ciascun cromosoma effettuando un'analisi di tipo caso/controllo e TDT per ciascun marcatore e aplotipo di tutti gli SNPs tipizzati nella fase I e fase II. I risultati per gli studi di associazione condotti sul cromosoma 2q sono rappresentati in figura 23.

A) Cromosoma 2 Caso/controllo



B) Cromosoma 2 TDT

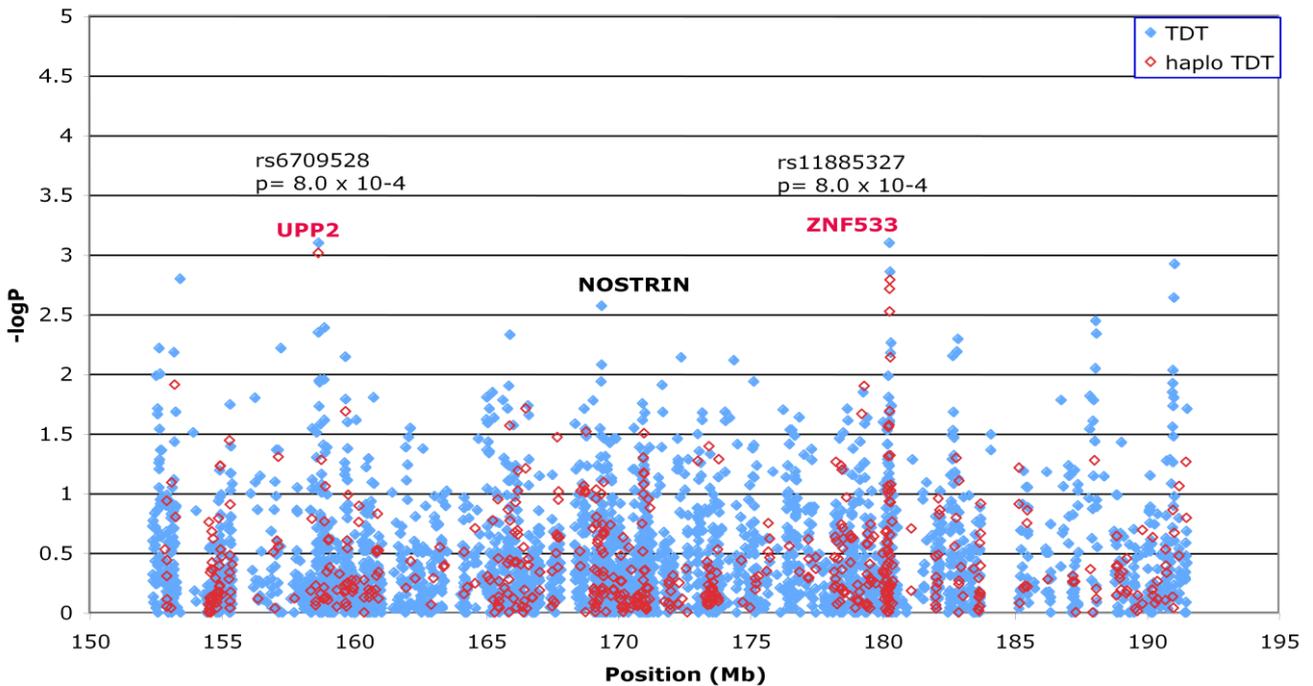


Figura 23: Studio di associazione caso/controllo e TDT per il locus al cromosoma 2q. In ordinata i valori di p-value in funzione del $-\log_{10}$; in ascissa le posizioni cromosomiche. I punti in azzurro indicano i valori di p-value ottenuti per i marcatori singoli, mentre i punti in rosso indicano i valori di p-value per l'analisi condotta per aplotipi. A) Studi di associazione di tipo caso/controllo. B) Studi di associazione di tipo TDT.

Nel locus al cromosoma 2q i risultati per quanto riguarda l'analisi di tipo caso/controllo mostrano i valori più significativi per 3 SNPs nel gene *NOSTRIN* (rs7583629, $P=3.2 \times 10^{-5}$; rs829957, $P=9.0 \times 10^{-5}$; rs482435, $P=1.4 \times 10^{-4}$), seguiti dal marcatore rs1020626 ($P=3.8 \times 10^{-4}$) nel gene *FAM130A2*. Per quanto riguarda lo studio di associazione di tipo TDT, i risultati più significativi appartengono al gene *ZNF533* (rs11885327, $P=8.0 \times 10^{-4}$; rs1964081, $P=1.4 \times 10^{-3}$), ed al gene *UPP2* (rs6709528, $P=8.0 \times 10^{-4}$).

Gli stessi test di associazione sono stati condotti sui marcatori del locus AUTS1 al cromosoma 7 ed i risultati sono schematizzati in figura 24.

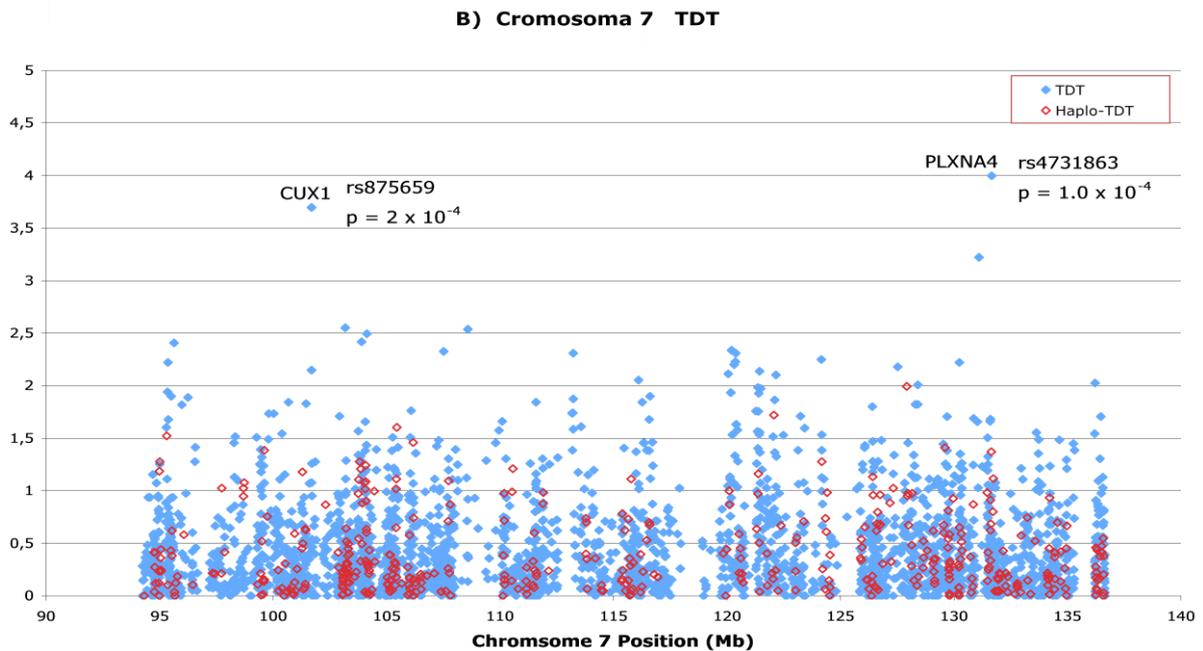
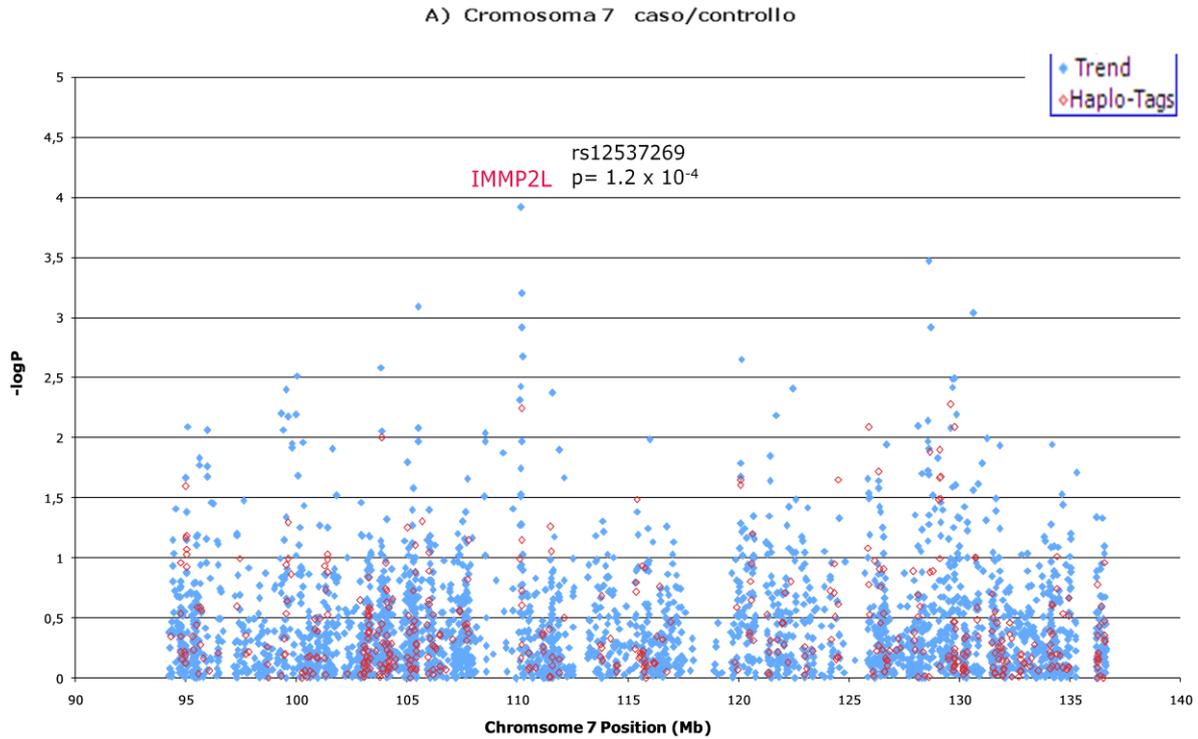


Figura 24: Studio di associazione caso/controllo e TDT per il locus al cromosoma 7q. In ordinata i valori di p-value in funzione del $-\log_{10}$; in ascissa le posizioni cromosomiche. I punti in azzurro indicano i valori di p-value ottenuti per i marcatori singoli, mentre i punti in rosso indicano i valori di p-value per l'analisi condotta per aplotipi. A) Studi di associazione di tipo caso/controllo. B) Studi di associazione di tipo TDT.

Nel locus al cromosoma 7q i risultati più significativi per quanto riguarda l'analisi caso/controllo risultano essere due marcatori nel gene *IMMP2L* (rs12537269, $P=1.2 \times 10^{-4}$; rs1528039, $P=6.3 \times 10^{-4}$) ed uno a monte del gene *SMO* (rs6962740, $P=3.4 \times 10^{-4}$). Mentre per l'analisi condotta mediante TDT i risultati più significativi coinvolgono il gene *PLXNA4* (*PLXNA4*; rs4731863, $P=1.0 \times 10^{-4}$) ed il gene *CUX1* (rs875659, $P=2.0 \times 10^{-4}$).

6.4 Studio di replicazione in campioni di famiglie indipendenti e meta-analisi

Lo studio di replicazione è stato condotto selezionando 28 SNPs più significativi di ciascun cromosoma su un campione di famiglie indipendenti costituito per lo più da famiglie singletons appartenenti al consorzio IMGSAC (IMGSAC-R) ed ad una collezione olandese (ND); la descrizione del campione di replicazione segue in tabella 5.

	<i>IMGSAC-R</i>	<i>ND</i>
PPC (Trios)	213	82
PC	47	14
C	34	-

Tabella 5: descrizione del campione utilizzato per la replicazione costituito da due campioni di famiglie indipendenti: IMGSAC-R ed ND (Northern Dutch). PPC, indica il numero di famiglie singletons costituite dai due genitori e figlio affetto; PC, indica il numero di famiglie costituito da un solo genitore ed il figlio affetto; C, indica il numero dei soli casi affetti.

Il risultato più significativo nel campione di replicazione si è ottenuto per il marcatore rs2217262 nel gene *DOCK4* ($P=9.2 \times 10^{-4}$) per il TDT. Il risultato rimane significativo anche in seguito alla correzione di Bonferroni per test multipli. Il resto degli SNPs non mostra alcun tipo di associazione nel campione di replicazione dopo la correzione per test multipli.

Tutti i genotipi dei 56 marcatori del campione primario e del campione di replicazione sono stati utilizzati per effettuare un'analisi complessiva con un numero totale di 515 famiglie con autismo. Nel locus AUTS1 il marcatore rs2217262, nel gene DOCK4, rimane il risultato più significativo nel caso dello studio di associazione per TDT. Mentre per il caso/controllo in un campione totale di 420 casi e 368 controlli, il risultato più significativo è ottenuto per il marcatore rs12537269 nel gene *IMMP2L*. Per quanto riguarda il cromosoma 2, il risultato più significativo si ottiene nel gene *ZNF533* per il marcatore rs7590028 nel caso/controllo. I risultati più significativi sono schematizzati in tabella 6.

Chr	SNP	Gene location	Risk Allele	ALL SAMPLES COMBINED (IMGSAC primary, IMGSAC-R, ND) 515 affected individuals Family-based analysis		IMGSAC SAMPLES COMBINED (IMGSAC primary, IMGSAC-R) 420 cases, 368 controls Case-control analysis	
				P-value (Ca-Co Freq)	OR (C.I.)	P-value (Ca-Co Freq)	OR (C.I.)
2	rs7590028	<i>ZNF533</i> intronic	T	0,3227 (0.52,0.50)		6,56E-04 (0.54,0.45)	1,41 (1.16-1.72)
7	rs2030781	<i>IMMP2L</i> intronic	C	0,08613 (0.25,0.22)		4,63E-04 (0.27,0.19)	1,53 (1.20-1.95)
7	rs12537269	<i>IMMP2L</i> intronic	A	0,01047 (0.27,0.22)		7,26E-05 (0.27,0.19)	1,62 (1.27-2.06)
7	rs2217262	<i>DOCK4</i> intronic	A	5,23E-05 (0.96,0.92)	2,37 (1.53-3.68)	1,75E-03 (0.96,0.92)	2,08 (1.31-3.32)
7	rs41620	3' of <i>TSPAN12</i>	A	0,08796 (0.77,0.74)		8,14E-04 (0.78,0.71)	1,48 (1.18-1.86)
7	rs538558	3' of <i>FEZF1</i>	A	0,4688 (0.36,0.34)		5,77E-04 (0.37,0.28)	1,45 (1.17-1.80)
7	rs11978485	3' of <i>SLC13A1</i>	G	0,04972 (0.82,0.79)		2,89E-04 (0.84,0.76)	1,59 (1.24-2.04)

Tabella 6: Sono elencati i 7 marcatori che mostrano i risultati più significativi per gli studi di meta-analisi utilizzando il campione primario ed il campione di replicazione nel caso del test TDT e caso/controllo. In rosso il marcatore che mostra il valore più elevato di p-value, in azzurro i marcatori con p-value significativi dopo la correzione per test multipli.

6.5 Geni candidati per il disturbo autistico per i loci AUTS1 e AUTS5

I risultati più significativi dello studio di associazione condotto per la regione di linkage nel cromosoma 2 e 7 mostrano evidenza di un possibile ruolo del gene *ZNF533*, *DOCK4* ed *IMMP2L*.

Il gene *ZNF533* copre un intervallo di circa 400 kb e mappa nella regione 2q31.2-q31.3, l'organizzazione genomica dell'isoforma principale è mostrata in figura 23. Il prodotto proteico codificato dall'isoforma più lunga è costituito da 471 a.a., contiene 4 motivi *zinc finger* (figura 25) e la localizzazione è potenzialmente nucleare, ma non vi sono dati sperimentali su questo né sulla sua funzione. L'analisi bioinformatica dei domini suggerisce un ruolo nella regolazione della trascrizione.

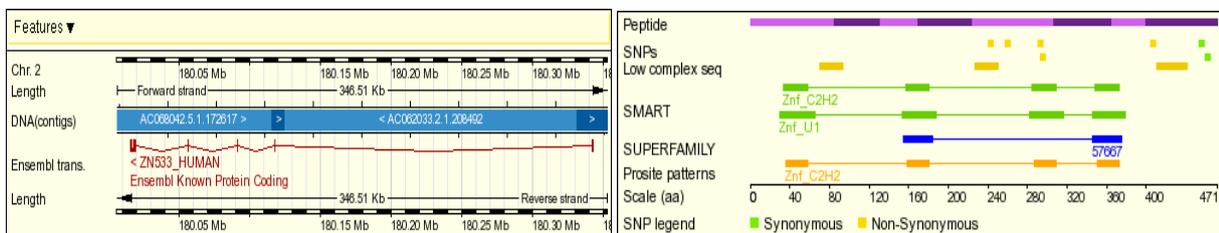


Figura 25: A sinistra è schematizzato l'organizzazione genomica dell'isoforma principale del gene *znf533*, a destra la descrizione dei domini presenti nel gene in seguito ad analisi bioinformatica

Sono indicate un numero elevato di isoforme alternative del gene (figura 26). L'espressione della principale isoforma avviene per lo più nei tessuti nervosi.

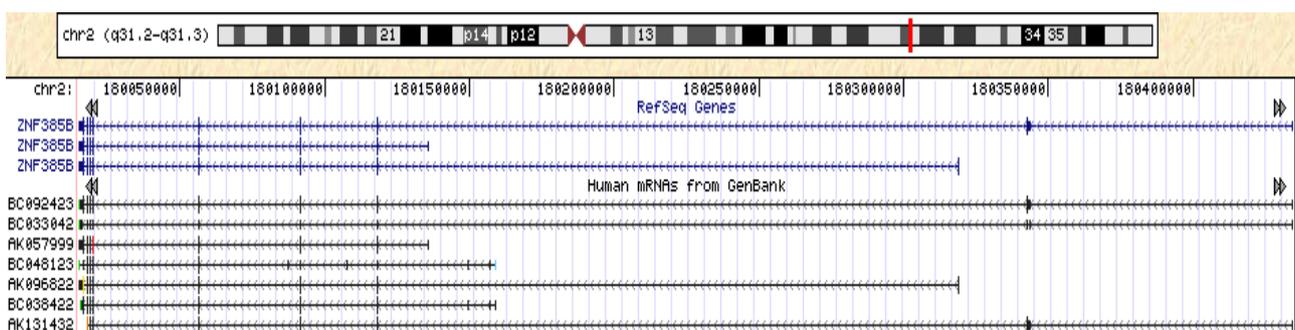


Figura 26: schematizzazione delle principali forme alternative del gene *znf533* indicate nel genome browser (build 36.1)

I risultati più significativi dello studio di associazione nella regione di linkage al cromosoma 7, nel campione primario e di replicazione, mostrano un possibile

coinvolgimento dei due geni adiacenti DOCK4 e IMMP2L (7q31.1), come evidenziato in figura 27, da cui è possibile notare la presenza di un ulteriore gene, LRRN3, nell'introne 3 di IMMP2L.

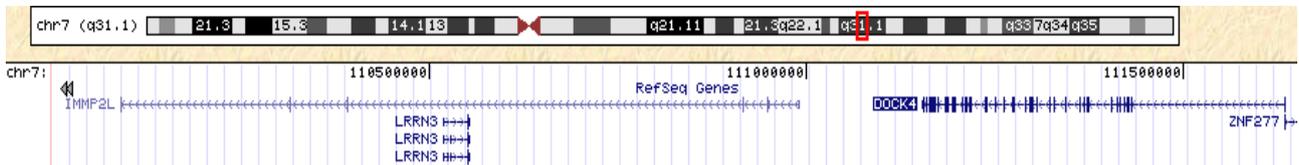


Figura 27: organizzazione genomica della regione 7q31.1 in corrispondenza dei geni candidati IMMP2L, LRRN3 e DOCK4

Il gene DOCK4, costituito da 52 esoni, codifica per una GEF (Guanine nucleotide Exchange Factor) in grado di attivare una Rac GTPase. Questo gene è coinvolto nella carcinogenesi quando è mutato (Yajnik, 2003), ma ultimamente alcuni studi hanno evidenziato il suo coinvolgimento anche a livello del SNC, in cui ha un'elevata espressione a livello dell'ippocampo e sarebbe responsabile della corretta formazione delle spine dendritiche (Ueda S, 2008).

Il gene IMMP2L, codifica per una delle due subunità catalitiche del complesso IMP (mitochondrial inner membrane peptide). Nell'introne 3 di IMMP2L mappa un interessante gene, LRRN3, che codifica per una proteina contenente domini *leucine rich repeats* altamente espresso nel cervello fetale.

6.6 Analisi di varianti strutturali nei loci AUTS1 e AUTS5

Gli arrays di SNPs come quello utilizzato nel nostro esperimento, possono essere utilizzati oltre che per condurre studi di associazione anche per identificare varianti strutturali. I dati dell'array della fase I e della fase II sono stati combinati per ottenere una sufficiente densità di SNPs in grado di poter condurre un'analisi di varianti strutturali utilizzando il programma QuantiSNP. Il software è in grado di inferire le potenziali varianti strutturali, siano esse delezioni o duplicazioni, analizzando la fluorescenza di marcatori contigui utilizzati nel genotyping e tenendo conto di due

parametri: 1) B-allele-frequency: rappresenta il rapporto del segnale di fluorescenza di una sonda allelica rispetto all'altra.

2) Log R ratio: rappresenta l'intensità del segnale di fluorescenza combinata che deriva da entrambe le sonde.

La maggior parte delle varianti strutturali incontrate nel cromosoma 2q e 7q sono ben rappresentate nel database di varianti genomiche o unicamente presenti nel nostro campione nei controlli o solo in un genitore. Il risultato più interessante di questo studio è la presenza di una duplicazione di circa 800 kb nella regione tra i due geni *IMMP2L-DOCK4*, trasmessa dal padre al figlio affetto (figura 28). Il dato è di particolare interesse se si considera che i due geni coinvolti hanno dato i migliori risultati negli studi di associazione condotti per la regione 7q.

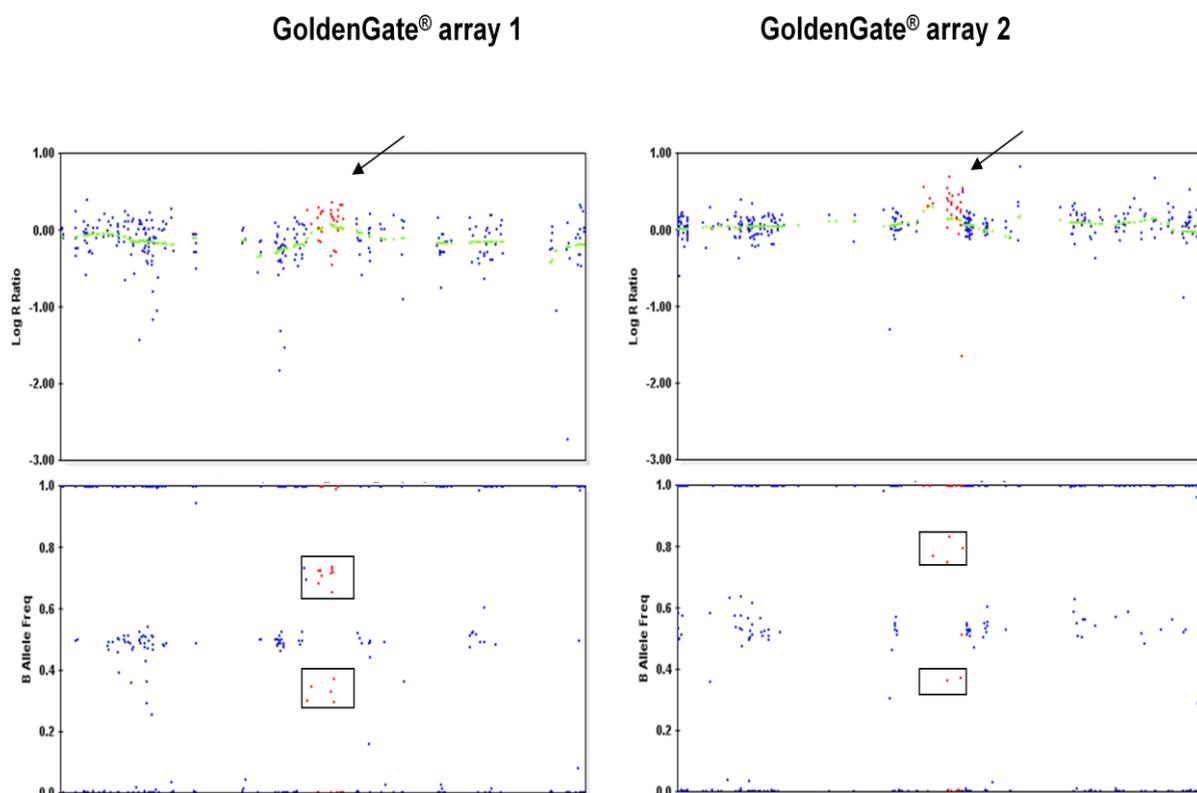


Figura 28: Sono indicati i due arrays in cui è possibile apprezzare la duplicazione della regione *IMMP2L-DOCK4* nella famiglia 148. In rosso i marcatori coinvolti nella variante strutturale in seguito all'analisi per "Log R Ratio" e "B allele frequency" con il software QuantiSNP.

7. Scopo della ricerca

L'autismo è un grave disturbo neuropsichiatrico che si manifesta durante l'infanzia e i cui sintomi permangono per tutta la vita del paziente. Gli studi epidemiologici hanno indicato la presenza di una forte componente genetica alla base di questo grave disturbo. L'identificazione delle basi genetiche dell'autismo rappresenta un obiettivo di primaria importanza per lo sviluppo di approcci diagnostici e terapeutici. Lo scopo di questo lavoro mira ad identificare i geni di suscettibilità all'autismo mediante differenti approcci quali: studi di linkage, studi di associazione, analisi di mutazioni ed analisi di varianti strutturali. Questi approcci sono stati utilizzati con l'obiettivo di identificare geni di suscettibilità all'autismo ed in modo particolare nel contesto di tre progetti:

- 1) Analisi del profilo di inattivazione del cromosoma X allo scopo di individuare geni X-linked coinvolti nella suscettibilità all'autismo.
- 2) Studio di replicazione riguardante la possibile associazione di varianti genetiche nel gene ASMT
- 3) Caratterizzazione di geni candidati per i loci di suscettibilità all'autismo AUTS1 e AUTS5.

1) Analisi di inattivazione del cromosoma X

L'ipotesi di un contributo da parte di geni sul cromosoma X nell'insorgenza dell'autismo è suggerita da diversi fattori: un maggior numero di maschi affetti rispetto a femmine affette (4:1); l'identificazione di mutazioni rare in geni X-linked in pazienti autistici; una debole evidenza di linkage riscontrata in 3 *genome scans* (Liu *et al.*, 2001; Auranen *et al.*, 2002; Yonan *et al.*, 2003). Mutazioni X-linked coinvolte in forme di ritardo mentale portano in madri portatrici asintomatiche all'inattivazione in maniera preferenziale del cromosoma X mutato. Per verificare l'ipotesi del coinvolgimento di geni legati al cromosoma X nell'autismo, abbiamo analizzato il profilo di inattivazione del cromosoma X (XCI) in un campione di 621 madri di pazienti autistici ed un campione di 182 pazienti autistiche, confrontandolo con quello di due gruppi di controllo correlati per età: 162 donne e 47 ragazze.

Finora l'unico studio condotto in questo senso, ha evidenziato una maggior frequenza del XCI sbilanciata (X-skewing) in madri di individui autistici benchè il campione utilizzato fosse ridotto (Talebizadeh et al., 2005).

L'obiettivo di questo studio è considerare l'ipotesi di coinvolgimento del cromosoma X nell'autismo verificando se vi è una differenza significativa del profilo di inattivazione tra il campione di casi e controllo; tale differenza potrebbe indicare la presenza di mutazioni X-linked. Questo approccio sarà inoltre utile per caratterizzare geni candidati per l'autismo considerando esclusivamente le famiglie X-skewing che avranno una probabilità più alta di avere mutazioni in questo cromosoma.

2) Studio di replicazione del gene ASMT

Nella regione pseudo-autosomica del cromosoma X ed Y (Xp22.3 o Yp11.3) mappa il gene ASMT, un interessante candidato funzionale per l'autismo coinvolto nella sintesi della melatonina. Ultimamente il consorzio per l'autismo PARIS ha evidenziato un possibile ruolo di questo gene nell'eziologia dell'autismo, in particolar modo associato a caratteristiche fenotipiche quali disturbi del sonno e ritmi circadiani alterati. Lo scopo del lavoro di replicazione, effettuato mediante uno studio di associazione di varianti nel promotore ed analisi di mutazione in un vasto campione europeo di individui affetti, mira a definire il ruolo di questo gene nell'autismo.

3) Caratterizzazione dei loci AUTS1 e AUTS5

Le regioni 2q24-33 e 7q21-32, definite rispettivamente come AUTS5 ed AUTS1 risultano le regioni più frequentemente replicate dagli studi di linkage per l'autismo (MGSAC, 1998; IMGSAC, 2001; Lamb et al., 2005; Shao et al., 2002; Buxbaum et al., 2001, Auranen et al., 2002). In queste due regioni, il consorzio IMGSAC a cui appartiene il mio laboratorio ha ottenuto i valori di linkage più consistenti, ed ha condotto recentemente una tipizzazione a larga scala di 6004 SNPs in 126 famiglie per il locus AUTS5 e 127 famiglie per il locus AUTS1. I genotipi sono stati utilizzati

per poter condurre ampi studi di associazione di tipo caso controllo e TDT. Inoltre sono stati utilizzati i dati dell'array ILLUMINA per poter ottenere una fitta copertura di marcatori contigui per condurre un'analisi di varianti strutturali. Lo scopo di questo lavoro è quindi caratterizzare i miglior geni candidati di ciascun locus e verificare:

- 1) L'espressione del gene candidato in tessuti cerebrali e caratterizzare le isoforme alternative nel caso non vi siano dati disponibili.
- 2) Effettuare uno screening di mutazione nelle regioni esoniche e nelle regioni regolatorie per individuare varianti con un potenziale ruolo eziologico per l'autismo.
- 3) Nel caso in cui sono state individuate varianti strutturali che coinvolgono geni di interesse e segregano con il fenotipo autistico, verificare la frequenza delle stesse in un più ampio campione di casi e di controlli.

8. Materiali e metodi

8.1 Campioni di DNA

I campioni di DNA genomico utilizzati in questo studio provengono dalla collezione di famiglie *multiplex e singletons* del Consorzio IMGSAC, di cui il nostro laboratorio fa parte. I criteri applicati per la diagnosi dell'autismo e per la raccolta delle famiglie, e le metodologie utilizzate per la preparazione dei campioni, sono stati precedentemente descritti (IMGSAC, 1998; 2001). Il DNA genomico dei pazienti è stato estratto da sangue o da saliva secondo le condizioni descritte in Maestrini et al. (1999).

Inoltre, nel corso di questi tre anni sono state raccolte, grazie ad una collaborazione del nostro laboratorio di Bologna con il Dott. Agatino Battaglia dell'Istituto Stella Maris di Pisa, circa 100 nuove famiglie *singleton* italiane che contribuiscono all'intera collezione IMGSAC.

Il DNA genomico degli individui appartenenti alle famiglie *singleton* italiane è stato estratto utilizzando il kit QIAGEN MAXI (§ 8.2)

Come popolazione di controllo sono stati utilizzati i DNA di individui caucasici di sesso misto appartenenti alla collezione ECACC (SIGMA) o una collezione di individui anonimi di origine italiana.

Gli stock di DNA sono dapprima diluiti ad una concentrazione finale di 80ng/μl in TE 1X e conservati a -20°C. Per effettuare le reazioni di amplificazione tramite PCR, si opera una ulteriore diluizione dei campioni in TE 0.1X, ottenendo una concentrazione finale di 8 ng/μl. I campioni così diluiti possono essere conservati a 4°C per alcuni mesi.

8.2. Estrazione del DNA da sangue intero con kit Qiagen MAXI (per 5 ml di sangue)

Il sangue, fresco o scongelato, è versato con attenzione in un tubo da 50 ml contenente 500 μl di Proteasi Qiagen, e lisato con 6 ml di buffer di lisi (AL).

I campioni sono agitati e vortexati 3 volte per 5 secondi, quindi incubati per almeno 10' a 70°C. Dopo raffreddamento a temperatura ambiente si aggiungono 5 ml di etanolo assoluto, si vortexano i campioni tre volte per 5 secondi, e il lisato è versato nel tubo contenente la colonna Qiamp. Dopo aver centrifugato per 3' a 3000 rpm (1850 g), si butta via il filtrato e si aggiungono 5 ml di buffer di lavaggio (AW1) e si centrifuga per 20' a 4500 rpm. Senza buttare via il filtrato, si aggiungono 5 ml di buffer di lavaggio (AW2), si centrifuga 20' a 4500 rpm, si sposta la colonna in un tubo da 50 ml pulito e si aggiungono successivamente 600 µl di buffer di eluizione (AE). Si lascia avvenire l'eluizione per 10' a temperatura ambiente, quindi si centrifuga 5' a 4500 rpm e si aggiungono altri 600 µl di buffer di eluizione (AE), ripetendo il procedimento precedente una seconda volta. Il DNA eluito viene quindi messo in un tubo sterile da 1.5 ml, e quantificato allo spettrofotometro, con determinazione dell'assorbanza a 260 nm. L'integrità del DNA genomico è verificata tramite elettroforesi su gel di agarosio.

8.3. Disegno dei primers

I primers per le reazioni di amplificazione sono stati disegnati utilizzando il programma Primer 3 (http://www-genome.wi.edu/cgi-bin/primer/primer3_www.cgi). Per lo *screening* di mutazioni, i primers sono stati disegnati in modo da coprire gli esoni e parte delle regioni introniche adiacenti, e almeno 300-400 bp del promotore.

8.4. Amplificazione tramite Polimerasi Chain Reaction (PCR)

Le reazioni di amplificazione sono effettuate in un volume finale variabile da 10 a 50 µl. Per ogni reazione tramite PCR è preparato un controllo negativo senza il DNA, per identificare eventuali contaminazioni. La composizione finale di una miscela di reazione standard è la seguente:

40 ng di DNA genomico in TE 0.1X (24 ng nelle reazioni in $V_f = 10 \mu\text{l}$, ; 40 ng nelle reazioni

in $V_f = 15\mu\text{l}$)

1X buffer per AmpliAq Gold DNA polimerasi (senza MgCl_2)

2-3 mM di MgCl_2

0.3 μM di primer Forward

0.3 μM di primer Reverse

0.2 mM di dNTPs

0.02 U/ μl di AmpliAq Gold DNA polimerasi (Applied Biosystems)

H_2O a volume

Per la PCR vengono utilizzati apparecchi a 96 posti della MJ Research e dell'Applied Biosystems. In generale si effettua un primo step di 15 minuti a 95°C per l'attivazione dell'enzima AmpliAq Gold, e per l'amplificazione si utilizza un protocollo definito *touchdown*: in questo tipo di protocollo lo step di *annealing* è dapprima impostato ad una temperatura più alta rispetto a quella calcolata sui primers utilizzati (T_{m1}), che viene diminuita di 0.5°C ad ogni successivo ciclo: in questo modo si conferisce una maggiore specificità iniziale alla reazione. Dopo 14 cicli si arriva alla temperatura di *annealing* (T_{m2}), che viene mantenuta costante per altri 25-30 cicli. Lo schema utilizzato per il protocollo di amplificazione *touch-down* è il seguente :

Pre-PCR step: 15' a 95°C

PCR 14 cicli:

95°C 30 secondi o 1 minuto

T_m 1 30 secondi con calo di 0.5°C ad ogni ciclo

72°C 30 secondi o 1 minuto

PCR 25-30 cicli:

95°C 30 secondi o 1 minuto

T_m 2 30 secondi o 1 minuto

72°C 30 secondi o 1 minuto

L'amplificazione di regioni molto ricche in G-C, nelle quali le due coppie di basi sono tenute

insieme da tre legami idrogeno, può risultare più difficoltosa; in tal caso la reazione è condotta con l'aggiunta di 0.1 mM finale di 7-deaza-2'-deoxiguanosina-5'-trifosfato (Deaza GTP) e/o 5% di DMSO. Questi reagenti aiutano la reazione della polimerasi prevenendo la formazione di strutture secondarie del DNA e facilitandone la denaturazione.

Al termine della reazione l'amplificazione è verificata tramite elettroforesi su gel di agarosio. Tutte le coppie di primers e le condizioni di amplificazione utilizzate in questo studio sono riportate in Tabella 7 (a-d)

Frammento esonico	Forward	Reverse	Condizione di amplificazione
ZNF533_promoter	TCTAGGCTTTGGGACCTCTG	CCCCTGAACTGTCCAACCTCT	Herculase, deAza, 5% DMSO,TD 61-54+ 32 cicli (1',30",1')
ZNF533_EX1a	AGTAAACCGTCACAGCTTCG	GACAAAGTGCAAACCTCTACCG	2mM Mg, de Aza, 5% DMSO,TD 63-56+ 32 cicli (1',30",1')
ZNF533_EX1b	CCCTTTCGAATCTTCCCTTA	GCATACACCTGGGCATATACAA	2mM Mg,TD 60-53+ 32 cicli (1',30",1')
ZNF533_EX1c	GGGTAGCTGGGCAACTTTTA	TTCGAGAGGTGACAAACAGC	2mM Mg,TD 63-56+32 cicli (1.30',30",1.30')
ZNF533_EX1.1	TCCTTCTTGTGGTTTGCTTG	TGCTCTTCAGCTTTGCTCCT	2mM Mg,TD 60-53+ 32 cicli (1',30",1')
ZNF533_EX1.2	TTCTCATTGATTTTTGTCTGAAAGA	TGGAAAAACAACATGTCACACC	2mM Mg,TD 62-55+ 30 cicli (1',30",1')
ZNF533_E 2.3	AGCTGCAACCTCATCTTTCA	AAAGCTGTCTATGGATTGAGAAAA	2.5mM Mg,TD 65-58+30 cicli (1',30",1')
ZNF533_EX 2	GCAAGTTGCAGCTCAACAGA	CACAACTGAGCAGCATAACCA	2.5mM Mg,TD 65-58+30 cicli (1',30",1')
ZNF533_EX 3	CTGTTATTTGCTCACACTGTTGTC	TGGGCTACCAATAGGAGTGG	2.5mM Mg,TD 59-51+30 cicli (1',30",1')
ZNF533_EX 4	TTTCATGAAGTGCCAAAAACA	TCCTTCCCGTGAAGAAAACA	2.5mM Mg,TD 65-58+30 cicli (1',30",1')
ZNF533_EX 5	TGCTGCAATAAAAACTTCACA	AGAGAAGCATGGAACAAACCT	2.5mM Mg,TD 65-58+30 cicli (1',30",1')
ZNF533_EX 6	GCCTGCAGGTAAACTTCTCAA	TGTGAGCACAGTAAAAACCATTC	2.5mM Mg,TD 65-58+30 cicli (1',30",1')
ZNF533_EX 7	CCTAGAGTTGGCAATGGAGAAA	AGACTCTATGCCCTCCTTGC	2.5mM Mg,TD 65-58+30 cicli (1',30",1')
ZNF533_EX 8	CTGAAACCAGCATTGTGAAAA	CTGCATTTTGTGGGTGAATG	2.5mM Mg,TD 65-58+30 cicli (1',30",1')
trascritti alternativi			
Frag. 1	GAGCGTGAGCAACAGCTTC	TCGTTGCGTCTAGTGCTTTG	2.5mM Mg,TD 63-56+32 cicli (1',30",1.30')
Frag. 2	CCAGGATGTGGAGTGGACTT	"	"
Frag. 3	AGTGCTCAAATGCAGACGTG	"	"
Frag. 4	ACAGGGATGGATTTGAGACG	"	"

Tabella 7a: Set di primers e relative condizioni di amplificazione utilizzate per il gene Znf533: sono presenti i primers utilizzati per amplificare le regioni esoniche ed i primer utilizzati per lo studio dei trascritti alternativi.

<u>Frammento esonico</u>	<u>Forward</u>	<u>Reverse</u>	<u>Condizione di amplificazione</u>
ASMT_ex1a	TGGTGTTCTCGCTGAGTCG	TGCGCACACTCCCAGGTG	2mM Mg,TD 61-54 + 32 cicli (1',30",1')
ASMT_ex1b	ACCTGCTCAATCCATAAGACG	CAATGGAACGTGAGTGTGATG	2mM Mg,TD 61-54 + 32 cicli (1',30",1')
ASMT_ex2	GCCATGGTATGGGGTGTTC	GGCTGCAGAGGGTGTCTATG	2mM Mg,TD 61-54 + 32 cicli (1',30",1')
ASMT_ex3	ATCACGATGTTGTCGAGCTG	GCGGTAGTCTGAGACTTGG	no condition
ASMT_ex3Falt	AAGTACATGGGCAGGACCAG		2mM Mg,TD 64-57 + 32 cicli (1', 30", 1'e20", Fast TAQ+GCrich buffer)
ASMT_x3bis	TGAAATCACGATGTTGTCGAGCTG	CCACCTCACGGCGTCTGC	2.5mM Mg, TD68-61+32 (1'-30"-1')
ASMT_ex4	CTGGGCTACAGAGCTGAAATG	TAAATGGCAGCTTCTCTCTG	2mM Mg,TD 61-54 + 32 cicli (1',30",1')
ASMT_ex5	CTGTGGGGTATAGCTCCGTTT	CGCACATGTCAAAGCATCAG	2mM Mg,TD 61-54 + 32 cicli (1',30",1')
ASMT_ex6	AGCTTGCACTGAGCGGAAATC	GCACCCATCGACTCGTCATTT	1.8mM Mg,TD 67-60 + 32 cicli (1',30",1')
ASMT_ex7	GGAAGGACCAAGTTCCTG	CGGGAGTGAGAGGAAAACC	2mM Mg,TD 61-54 + 32 cicli (1',30",1')
ASMT_ex8	GCAGGTGACTAGCCTGGAAG	CTGTGGGATGATTTCACTGC	2mM Mg,TD 61-54 + 32 cicli (1',30",1')
ASMT_ex9	TCCTTGAGGTCTGGCTGTC	GTCCACAGAAGCCTCCAGTC	2mM Mg,TD 61-54 + 32 cicli (1',30",1')

Tabella 7b: set di primers e relative condizioni di amplificazione utilizzate per il gene ASMT

<u>Frammento esonico</u>	<u>Forward</u>	<u>Reverse</u>	<u>Condizione di amplificazione</u>
NLGN3-X1	tggcacaaaaggagtaaatgg	agtggaggagcctgggatt	2.5mM Mg,TD 65-58 + 30 cicli (1',30",1')
NLGN3-X2	ctgggcatgtgaaacctctc	catgcaagccacaacacag	2.5mM Mg,TD 65-58 + 30 cicli (1',30",2')
NLGN3-X2 primers sequenza	cacctgtggttctctcagtt	tcactgggcagtggtactcg	
NLGN3-X2-exon alternat	agagcagctcagctcttgg	tgggctacttgaggaagg	2.5mM Mg,TD 65-58 + 30 cicli (1',30",1')
NLGN3-X3	ctgcagtcagctgttttga	ggaacagaaacaccaacg	2.5mM Mg,TD 65-58 + 30 cicli (1',30",1')
NLGN3-X4	cacctgggatagctttgctg	gcaccagctagagaagcaag	2.5mM Mg,TD 65-58 + 30 cicli (1',30",1')
NLGN3-X5	tattcccacctcccctgttg	aagagagctggccgattcc	2.5mM Mg,TD 65-58 + 30 cicli (1',30",1')
NLGN3-X6	cagcctcagtgacaaaggaa	ggaagggatgatggaagagg	2.5mM Mg,TD 65-58 + 30 cicli (1',30",2') + 5% DMSO
NLGN3-X6 primers sequenza	gacctgaggatggtgtctc	tctcaggggtgtcacgtct	
NLGN3-X7	aggtttggctgtcagaggaa	tgcatgtgttcttgatctg	2.5mM Mg,TD 65-58 + 30 cicli (1',30",2')
NLGN3-X7 primers sequenza	catctcacctgcctacagca		

Tabella 7c: set di primers e relative condizioni di amplificazione utilizzate per il gene NLGN3

Frammento esonico	Forward	Reverse	Condizione di amplificazione
mecp2-x1	cggccatcacagccaatg	cgccagccgtgtcgtc	2mM Mg +5% DMSO + deAZA GTP (1'-30''-1') 120 ng DNA in 50 v.f.
mecp2-x2	taggtaagctgggaaatagcc	acagatggccaaaccagga	2.5 mM Mg (30''-30''-1') 40ng DNA in 30 v.f.
mecp2-x3	aagcctgcctctgctcact	ccctgggcacatacatttc	2.5 mM Mg (30''-30''-1') 40ng DNA in 30 v.f.
mecp2-x4a	cgctctgcctatctctga	TTCTTGGGAATGGCCTGA	2.5 mM Mg (1'-30''-1') 40ng DNA in 30 v.f.
mecp2-x4aF_alt	ctgccctatctctgacattgc		
mecp2-x4b	cgtccttgcgTGCCTTT	ctttgggactctgagtgtt	2.5 mM Mg (30''-30''-1') 40ng DNA in 30 v.f.
mecp2-x4c	ggaaaggactgaagacctgtaag	ccctcccctcggtgtttg	2.5 mM Mg +5%DMSO + deAZA GTP (1'-30''-1') 64 ng DNA in 50 v.f.
mecp2-x4d	ggagaagatgccagaggag	tccccaatgctccaactact	2.5 mM Mg (30''-30''-1') 40ng DNA in 30 v.f.

Tabella 7d: set di primers e relative condizioni di amplificazione utilizzate per il gene Mecp2

8.5 Elettroforesi su gel di agarosio

Per la preparazione del gel si scioglie la quantità necessaria di agarosio in un volume di tampone TBE 1X che va dai 40 ml ai 400 ml a seconda delle dimensioni e dello spessore desiderato del gel. Dopo essere stato portato ad ebollizione il gel viene raffreddato a 60°C circa, successivamente si aggiungono 5 µg/ml di *Red Gel*. Il gel viene versato su di un supporto orizzontale e polimerizzato per 30 minuti.

Si trasferisce poi il supporto in un apparato per elettroforesi orizzontale, contenente il tampone di corsa (TBE 1X). Per il caricamento si aggiunge ai campioni il *loading buffer* per DNA (1X finale). La corsa è effettuata a 5 volts/cm per il tempo necessario. Al termine della corsa elettroforetica il DNA viene visualizzato su un transilluminatore a luce UV e fotografato. La percentuale di agarosio utilizzata varia dallo 0.8% al 2.5%, in funzione della grandezza dei frammenti da analizzare.

8.6 Humara Assay

Questo saggio permette di distinguere il cromosoma X attivo da quello inattivo e di calcolare il pattern di XCI. Si basa sulla presenza di una tripletta ripetuta (CAG) a livello dell'esone 1 del gene AR, caratterizzata da un'elevata eterozigosità (~90%) nella popolazione.

Il saggio consiste nella digestione differenziale del DNA genomico con l'enzima *RsaI*, in presenza e in assenza dell'enzima sensibile alla metilazione *HpaII*. *HpaII* presenta un sito di taglio nell'esone 1 di AR, in prossimità della *Short Tandem Repeat* (STR), e taglia il DNA solo se questo non è metilato a livello dei residui di citosina, che sono metilati sull'Xi. *RsaI* è un enzima indipendente dalla metilazione, che taglia entrambi i cromosomi X e viene utilizzato per ricreare condizioni analoghe di digestione per i campioni in esame.

Brevemente, 50ng di DNA genomico sono stati digeriti con 0.25U *RsaI* + 0.7U *HpaII*, e con 0.25U *RsaI* soltanto, a 37°C per 18 ore in un volume finale di 10µl. 5µl del DNA digerito sono stati amplificati in condizioni standard in un volume finale di

20µl, con il seguente programma: 15' a 95°C, 30 cicli 1' a 95°C.30" a 60°C-1' a 72°C, 5' a 72°C.

I primers utilizzati per l'amplificazione sono:

For: 5'-TCCAGAATCTGTTCCAGAGCGTGC-3' marcato con 6-FAM

Rev: 5'- GCTGTGAAGGTTGCTGTTCCCTCAT-3'

I prodotti di PRC sono stati corsi su un sequenziatore automatico ABI3730 e l'altezza dei picchi è stata misurata con il software Peak Scanner (Applied Biosystems).

L'XCI è stata calcolata per gli individui eterozigoti informativi come rapporto tra l'altezza dei picchi dei due alleli ottenuti per gli individui eterozigoti dopo digestione con *HpaII*+ *RsaI*, corretta per il rapporto tra l'altezza dei picchi dei due alleli ottenuta nel campione digerito con il solo *RsaI*, in modo da eliminare l'effetto dovuto all'amplificazione preferenziale degli alleli dovuto alle diverse dimensioni dei frammenti (Amos-Landgraf *et al.*, 2006).

Gli esperimenti sono stati condotti su piastre a 96 pozzetti, in duplicato, in presenza di un campione con XCI 95:5 e di un campione di sesso maschile, come controlli interni.

8.7 Reazione di sequenziamento ed analisi delle sequenze

Per le reazioni di sequenza è stato utilizzato il kit ABI PRISM Big Dye Terminator Cycle Sequencing (V2.1 Applied Biosystems) e i primers specifici di ciascun frammento. I prodotti di amplificazione sono stati purificati usando il sistema di filtrazione Multiscreen PCR della Millipore, seguendo le indicazioni della ditta fornitrice. Le reazioni di sequenza sono state effettuate in un volume finale di 10µl, secondo il protocollo fornito dalla ditta, e purificate mediante precipitazione con 1/10 di Na-acetato 3M pH5.2 e 2.5V di etanolo assoluto.

Le reazioni di sequenza sono state corse su un sequenziatore ABI PRISM 3730 (Applied Biosystem). Per l'analisi e la lettura dei file di sequenza sono stati utilizzati i programmi Sequence Navigator (Applied Biosystems) e Vector NTI (Invitrogen).

8.8 Analisi di Linkage

Per l'analisi di *linkage* del cromosoma X sono stati utilizzati i dati genotipici generati per le famiglie multiplex dell'IMGSAC dall'Autism Genome Project (AGP) utilizzando l'array di SNPs Affimetrix 10K v2 (Autism Genome Project consortium *et al.*, 2007). L'analisi di *linkage* non parametrica è stata svolta utilizzando la statistica NPL "all" del programma Merlin (Abecasis *et al.*, 2002). La significatività empirica è stata determinata attraverso 1000 simulazioni dei genotipi del cromosoma X.

8.9 Analisi di associazione: caso/controllo

Lo studio di associazione caso/controllo del promotore del gene ASMT è stato realizzato utilizzando il software Haploview (v4) (Barrett JC, 2005). Il software è stato impiegato per l'analisi dei singoli marcatori e dell'aplotipo. Si considerano significative le differenze alleliche per le quali il p-value è <0.05 .

8.10 qPCR mediante Real-Time

L'analisi di possibili varianti strutturali a carico del gene MECP2 è stata condotta mediante qReal-time PCR, usando il metodo del ddCt, come descritto da (Van Esch *et al.*, 2005). Due set di primers del gene MECP2 sono stati sottoposti a validazione, in modo da verificare che avessero un'efficienza di amplificazione analoga a quella della coppia di primers del gene NXF5 sul cromosoma Xq22, utilizzato come sequenza di riferimento per la normalizzazione. La reazione di amplificazione è stata condotta utilizzando la Power SYBRGreen PCR Master Mix (Applied Biosystems), con 500nM di ciascun primer e 10ng di DNA in un volume finale di 15 μ l. Le reazioni sono state effettuate in triplicato su un apparecchio 7500 Fast Real Time PCR System (Applied Biosystems), includendo i seguenti controlli: un campione di un individuo di sesso maschile normale e un campione di un individuo di sesso maschile con un'ampia duplicazione comprendente il gene MECP2 nelle piastre con individui di

Sesso maschile; un campione di un individuo di sesso femminile normale e un campione di un individuo di sesso femminile con un'ampia duplicazione comprendente il gene MECP2 nelle piastre con individui di sesso femminile.

8.11 Analisi di varianti strutturali mediante MLPA (Multiplex Ligation Probes Amplification)

Lo studio di duplicazioni nel gene Mecp2 è stato condotto utilizzando il kit P015 dell' MRC-holland. La sequenza ed il numero di sonde utilizzate nel nostro kit sono disponibili direttamente sul sito MRC-holland (<http://www.mlpa.com>). Lo schema utilizzato per il protocollo consiste di tre punti:

1. Denaturazione del DNA ed ibridizzazione delle sonde (SALSA MLPA probes):

2.5 ul DNA (DNA (20ng/ul) => 50 ng DNA)

Denaturare il DNA a 98 °C per 5 minuti

Quindi aggiungere al campione:

0.75 ul MLPA buffer

0.75 ul MLPA probes

A ciascun campione è stato aggiunto olio minerale per evitare evaporazione della soluzione di reazione. Reazione di ibridizzazione per a 60°C 16 ore

2. Ligazione:

Aggiungere la reazione di ligazione (Ligase-65) per ciascun campione:

1.5 ul ligase buffer A

1.5 ul ligase buffer B

12.5 ul H₂O

0.5 ul ligase

La reazione di ligazione prevede

- 54°C per 15 minuti

- 98°C per 5 minuti per disattivare l'enzima

3. Reazione di PCR:

5ul del prodotto di ligazione

13 ul H₂O

2 ul di buffer PCR SALSA

60°C per 10 minuti

Mix della polimerasi:

1 ul primers di PCR

1 ul di buffer enzima SALSA

2.75 ul H₂O

0.25 ul polimerasi

Mentre i tubi sono a 60°C si aggiunge la mix di polimerasi per permettere l'inizio della reazione di PCR.

Condizione di PCR:

35 cicli:

95°C per 30 secondi

60°C per 30 secondi

72°C per 30 secondi

Infine:

70°C per 20 minuti

Alla fine della PCR si caricano su gel 2 ul di prodotto di reazione ;in seguito alla corsa elettroforetica si può apprezzare una scala di prodotti di applicazione. in. Infine 1ul del prodotto di ciascun campione viene risospeso in una miscela di 9ul di formamide e size standard (8.8ul formamide e 0.2ul di LIZ600) e corso al

sequenziatore ABI PRISM 3730 (Applied Biosystem). Il protocollo è schematizzato in figura 29.

I dati sono stati quindi elaborati analizzando le corse al sequenziatore con GENEMAPPER (Applied Biosystem). L'analisi dei dati è stata condotta utilizzando il software "Coffalyser v5.4" disponibile in rete sul sito dell'MRC-holland.

MLPA, Multiplex Ligation-dependent Probe Amplification

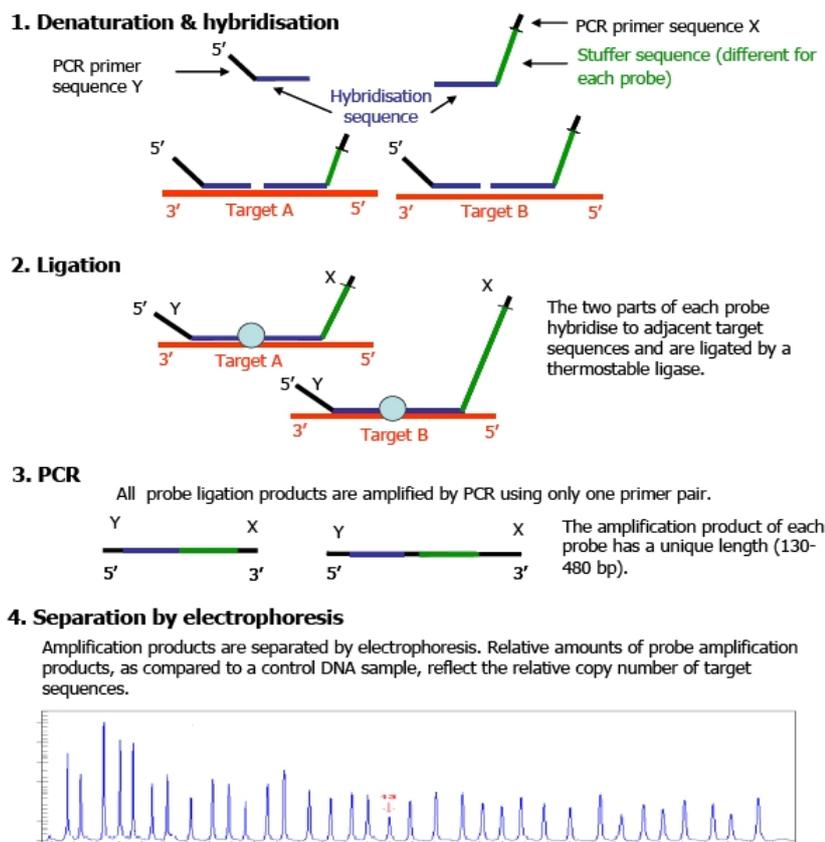


Figura 29: Schematizzazione del protocollo dell'MLPA eseguito per individuare duplicazioni nel gene Mecp2.

8.12 Tecnologia Illumina: tecnologia array-based per il genotyping di SNPs su larga scala

Il sistema Illumina (*BeadArray™*) si basa su 96 *arrays*, ognuno delle quali composto da circa 50.000 fibre ottiche, su ognuna delle quali è inciso un piccolo pozzetto che alloggia una *bead* di 3 micron, alla quale sono attaccate moltissimi

probes con la stessa sequenza nucleotidica. Ogni *arrays* contiene più di 1500 diverse sonde oligonucleotidiche, ognuna delle quali è presente in circa 30 copie, rendendo quindi il risultato di *genotyping* più affidabile perché basato su delle repliche.

Per ogni SNP da analizzare vengono disegnati due oligonucleotidi allele specifici (ASOs: *Allele-Specific Oligonucleotides*) e un oligonucleotide locus specifico (LSO: *Locus Specific Oligonucleotides*) circa 20 bp a valle dell'oligonucleotide allele specifico. Ogni ASO termina al 3' con una base che si ibridizza (in modo allele specifico) al DNA genomico in corrispondenza dello SNP, mentre al 5' contiene una regione complementare ad un primer universale per la PCR (P1 o P2 coniugati ad una cianina: Cy3 o Cy5 a seconda dell'allele). Il LSO, invece, è costituito da tre regioni: al 5' c'è una sequenza locus-specifica per lo SNP, nella porzione centrale una sequenza indirizzo complementare ad una delle 1500 sonde contenute in un *array* della piattaforma, e al 3' una sequenza complementare ad un primer universale (P3). Si utilizza quindi una polimerasi in grado di effettuare l'estensione di un oligonucleotide che si ibridizza perfettamente al 3' con il target e priva di attività esonucleasica e di *displacement* di sequenze a valle. In questo modo viene riempito il *gap* fra ogni ASO e LSO generando un substrato allele specifico per la reazione di PCR. Si introducono i tre primer universali (P1, P2, P3) e si effettua un classico programma di amplificazione. I prodotti di PCR, infine, si ibridizzano alle sonde negli *arrays* tramite la sequenza indirizzo. Come precedentemente spiegato, ad ognuno dei 1500 SNPs è assegnata una sequenza indirizzo contenuta nel LSO complementare ad una sola sonda presente nell'*array*. I rilevatori di fluorescenza utilizzati da Illumina scansionano infine la piastra contenente 96 *arrays* contemporaneamente a due diverse lunghezze d'onda adatte alla Cy3 e alla Cy5. Nel caso di un campione omozigote per un dato SNP, il rilevatore osserva un solo tipo di fluorescenza (o quella dipendente dal primer P1 coniugato alla Cy3 o dal primer P2 coniugato alla Cy5), invece in presenza di un campione eterozigote verranno rilevate entrambi i tipi di fluorescenza. Una schematizzazione del saggio è presente in figura

30. Quindi il rapporto dei segnali relativi alle due diverse fluorescenze identifica il genotipo come AA, AB, o BB.

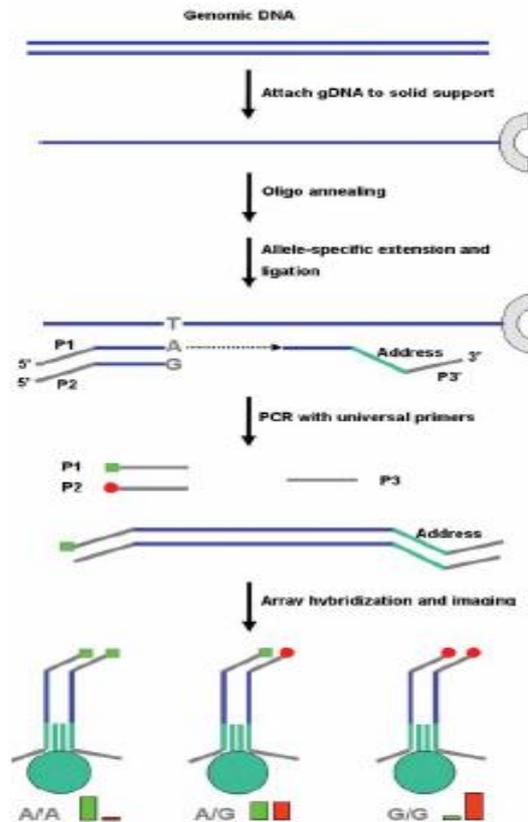


Figura 30: Schema del saggio Illumina (*GoldenGate™ genotyping assay*)

8.13 Analisi di varianti strutturali mediante QMPSF (*Quantitative Multiplex PCR of Short Fluorescent Fragments*)

Per analizzare la presenza di varianti strutturali nella regione IMMP2L-Dock4 abbiamo utilizzato una tecnica quantitativa che consiste nell'effettuare una multiplex con sonde specifiche della regione che si vuole studiare e sonde di riferimento.

I primers sono stati disegnati (tabella 8) e successivamente è stato utilizzato il software AutoDimer v.1 per testare l'assenza di prodotti aspecifici (<http://en.bio-soft.net/pcr/AutoDimer.html>).

La composizione finale della miscela di reazione è la seguente:

50 ng di DNA genomico in TE 0.1X in un Volume finale di 15µl

1X buffer per Amplitaq Gold DNA polimerasi (senza MgCl₂)

2 mM di MgCl₂

0.4-0.6 μM per ciascun primer

5% DMSO

0.2 mM di dNTPs

1 U di AmpliTaq Gold DNA polimerasi (Applied Biosystems)

H₂O a volume

L'analisi è stata condotta effettuando il rapporto del rapporto tra l'area della sonda locus specifica e la sonda di riferimento tra il paziente e la media di tutti i campioni (figura 31)

$$\text{Copy Number} = \frac{\text{Paziente} \quad \text{A sonda/A sonda reference}}{\text{Media tutti i campioni} \quad \text{A media sonda/A media sonda reference}} = \begin{cases} >1.3 \text{ DUPLICAZIONE} \\ <0.7 \text{ DELEZIONE} \end{cases}$$

Figura 31: Schematizzazione dell'analisi condotta per analizzare varianti strutturali mediante QMPSF nella collezione IMGSAC

<u>Frammento</u>	<u>Forward</u>	<u>Reverse</u>	<u>Condizioni di amplificazione</u>
Primo Assay Multiplex			
>IMMPL2_ex1	CCTCTCCCAGTCCTCACAAA	TTACCTTTCTCCCGCTCTT	60 + 20 (30"-30"-30"), 2 mM Mg 5uDNA Vf=15ul
>IMMPL2_ex2	GACAAGATTGCCCTTCGGTA	GACCCGATCCAAGAAAGTCA	"
>IMMPL2_ex3	ATCCCAGTGAGCCTTTTTGA	TTCCAGTGGTTCAAAAGCA	"
>DOCK4_lastEX	GCAGTCAAACAGCAAAAAGCA	AGGCAAGGTTTATCAGTGCAG	"
>HMBS	TTCTCACCAAATCCCACCTC	TGCCCCACAAACCAGTTAAT	"
>LRRN3_ex4	CCAAGCTGGAATCACTCAT	GGTCCATGAATCGAATGTT	"
>IMMPL2_ex7	CTTTATTGCTTTCTTTCTGTTTCTCA	GCAGTCATTCTCTTCTCTCTGT	"
>RNF20	GAGACAGCCGAATCACGAGT	CCAAGGTGAGTCTTCCTGA	"
Secondo Assay Multiplex			
Dock4_ex1	ggcgtgttcgcttagttcag	cttctccggctcacaacaat	60 + 20 (30"-30"-30"), 2 mM Mg 5uDNA Vf=15ul
Dock4_ex31	ttttatccctgcctgattg	tctgctcccagtcacata	"
FoxP2_intr2	tgctagaggagtgggacaagta	gaagcaggactctaagtgcaga	"
Dock_ex52	gcagtcaaacagcaaaaagca	aggcaaggtttatcagtgcag	"
Dock_ex25	ggatattctgctgaggacca	taacatcaagccacctgcaa	"
RNF20	gagacagccgaatcacgagt	ccaaggtgagtcttctga	"
Dock_ex8	tggcgtaggacatcacagtc	tctcccaggttggaaacacat	"
Dock_ex15	aacaacagtcccaggtggtc	cctccccttctggctaatac	"

Tabella 8: Set di primers utilizzati per l'analisi delle varianti nella regione IMMPL2-Dock4 mediante QMPSF

9. Risultati

9.1.1 Il Pattern di XCI

Il pattern di inattivazione del cromosoma X è stato valutato, nel nostro campione, attraverso l' HUMARA *assay*, un saggio che sfrutta la presenza di un STR a livello dell'esone 1 del gene AR, localizzato in Xq12. Il gene AR è inattivato durante il processo di XCI, tramite metilazione dei dinucleotidi CpG. Gli enzimi di restrizione utilizzati nell'HUMARA *assay* sono: *HpaII*, sensibile alla metilazione ed *RsaI*, che ha i siti di taglio al di fuori della ripetizione trinucleotidica e del frammento amplificato (figura 32).

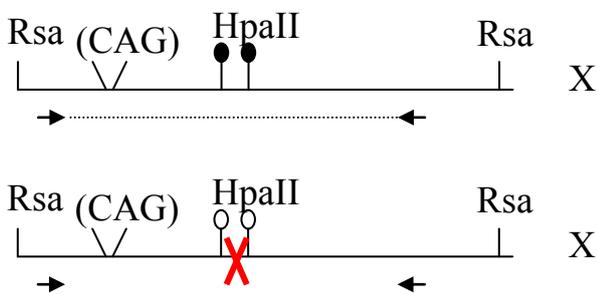


Figura 32: *HpaII* taglia il DNA solo se questo non è metilato a livello dei dinucleotidi CpG (X attivo). *RsaI* invece è indipendente dalla metilazione ed ha i siti di taglio all'estremità della ripetizione trinucleotidica, al di fuori del frammento amplificato.

In questo studio il pattern di inattivazione del cromosoma X è stato calcolato, per gli individui eterozigoti e quindi informativi, come il rapporto tra le altezze dei picchi dei due alleli della ripetizione (CAG) del gene AR dopo amplificazione del campione digerito con *HpaII*, corretto per un fattore ricavato dal rapporto tra le altezze degli stessi picchi dopo amplificazione del campione digerito con il solo *RsaI*, questo per controbilanciare l'amplificazione preferenziale degli alleli di dimensioni minori. Per ricavare le altezze dei picchi è stato utilizzato il software Peak Scanner v 1.0 o Genemapper. In figura 33, sono descritti 3 prototipi del pattern dei possibili risultati ottenuti dallo HUMARA *assay*: Human Male, femmina di controllo con pattern di X-Skewed, madre con un pattern di XCI bilanciato.

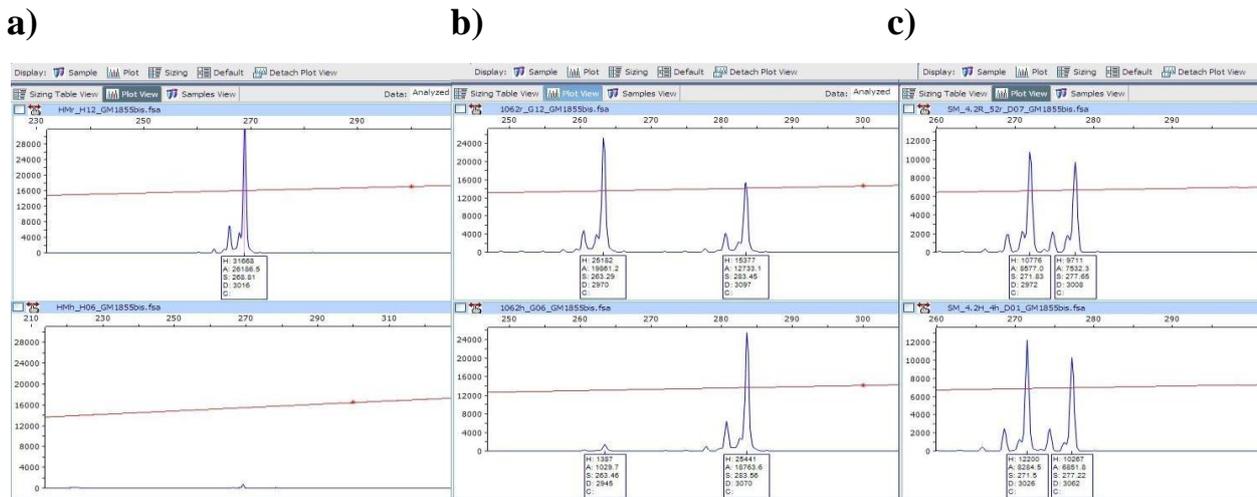


Figura 33. Risultati ottenuti dallo HUMARA assay per alcuni campioni. In alto vi sono i risultati ottenuti dalla digestione solo con RsaI, in basso quelli dalla digestione con RsaI e HpaII insieme. **a.** Controllo maschio sano. **b.** Controllo femmina con un pattern di XCI noto (5:95). **c.** Femmina con un pattern di XCI random.

Il pattern di inattivazione del cromosoma X per i campioni analizzati in questo studio è stato ricavato da esperimenti condotti in duplicato, applicando la formula riportata in figura 34.

$FC = \frac{\text{media allele 1 RsaI}}{\text{media allele 2 RsaI}}$	$XCI = \frac{\text{Allele 1 HpaII}}{\text{Allele 1 HpaII} + (\text{Allele 2 HpaII} \times FC)} \times 100$
--	--

Figura 34. Formule applicate per ricavare il pattern di X inattivazione. Allele 1: allele di dimensioni minori; allele 2: allele di dimensioni maggiori. FC: fattore di correzione. XCI: percentuale dello stato di inattivazione cromosoma X con l'allele HpaII più piccolo.

9.1.2 Profilo di XCI nel campione IMGSA C

L'analisi del profilo di inattivazione del cromosoma X è stata condotta mediante HUMARA assay, confrontando due campioni costituiti rispettivamente da madri di individui con autismo e probande, con altrettanti gruppi di controllo (tabella 9). I due

campioni con autismo dell'IMGSAC comprendono 275 madri di individui affetti da autismo (201 da famiglie multiplex, 74 da famiglie singletons) e 75 probande; tra queste, 249 madri e 69 probande sono risultate eterozigoti e quindi informative per l'XCI. Le famiglie sono state classificate in base al sesso degli individui affetti, in quanto è più probabile che ci siano mutazioni recessive X-linked in famiglie con soli maschi affetti (*males only*, MO) rispetto a quelle con una o più femmine affette (*female containing*, FC).

I due corrispondenti gruppi di controllo comprendono 162 donne adulte e 47 ragazze; tra queste, 144 donne adulte e 40 ragazze sono risultate informative per l'XCI (tabella 9). I gruppi di controllo sono stati selezionati in modo da avere un profilo di distribuzione dell'età simile a quello degli individui in analisi, in quanto è noto che all'aumentare dell'età della donna si osserva un maggiore grado di sbilanciamento dell'XCI su DNA ottenuto da cellule del sangue (Busque et al., 1996). Il pattern di XCI riscontrato nei gruppi di controllo è risultato simile a quello riportato in studi precedenti (Amos-Landgraf *et al.*, 2006).

I risultati ottenuti, utilizzando come valori soglia rispettivamente un $XCI \geq 80:20$ e un $XCI \geq 90:10$, sono riassunti in tabella 9.

CAMPIONI	N	INFORMATIVI	XCI \geq 80:20	P VALUE	XCI \geq 90:10	P VALUE
MADRI MO	197	181	28 (15.5%)	0.274	9 (5.0%)	0.353
MADRI FC	78	68	11 (16.2%)	0.299	3 (4.4%)	0.502
RAGAZZE AFFETTE	75	69	5 (7.2%)	0.093	1 (1.4%)	0.633
CONTROLLI ADULTI	162	144	18 (12.5%)		5 (3.5%)	
CONTROLLI GIOVANI	47	40	7 (17.5%)		0	

Tabella 9: Risultati ottenuti tramite HUMARA assay per l'XCI nei campioni in esame. N: totale individui per ciascun sottogruppo.

Come si può osservare, in nessun caso vi sono differenze statisticamente significative tra il campione di madri di individui con autismo e il corrispondente gruppo di controllo, né per le famiglie MO né per le famiglie FC. Anche tra il campione di probande e il relativo gruppo di controllo non sono state individuate differenze significative per entrambi i valori soglia di XCI considerati. Un analogo dato è ottenuto se si considera come valore soglia un valore di $XCI \geq 90:10$.

I dati sono stati ulteriormente analizzati suddividendo il campione di madri in base all'appartenenza a famiglie multiplex e singleton (figura 35a, b). Neanche in questo caso si osservano differenze significative tra i due gruppi.

a) Pattern XCI Famiglie multiplex b) Pattern XCI Famiglie singletons

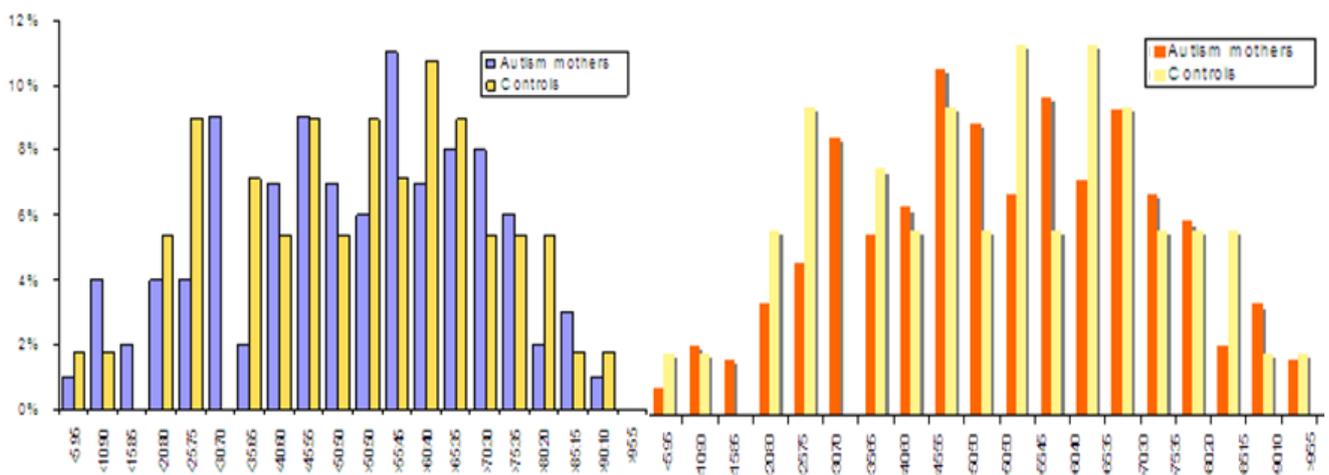


Figura 35: Distribuzione del pattern di XCI nel campione di madri provenienti da famiglie multiplex (a) e da famiglie singletons (b), confrontato con quello del gruppo di controllo.

9.1.3 Profilo di XCI nel campione totale

Il nostro studio è stato condotto in collaborazione con il consorzio PARIS e FINLAND. I due gruppi europei per lo studio dell'autismo, hanno a loro volta analizzato il profilo di inattivazione del cromosoma X in campioni di madri di individui con autismo e di probande. Nello studio sono state incluse anche 4 famiglie con individui affetti da disturbi dello spettro autistico, in cui erano presenti mutazioni

in geni X-linked già identificate in precedenza: tre famiglie con coppie di fratelli affetti con mutazioni rispettivamente nei geni ATRX, NLGN3 e NLGN4X (Jamain *et al.*, 2003); una famiglia singleton con una mutazione *de novo* R294X nel gene MECP2 in una ragazza affetta da autismo.

I risultati ottenuti per queste famiglie sono estremamente interessanti. Nella prima famiglia, la madre ha una XCI altamente sbilanciata (92:8) e trasmette la mutazione R451C nel gene NLGN3 ai suoi due figli affetti (Cap. 4 § 4.2.2). Nella seconda famiglia, la madre ha una XCI di 95:5 e trasmette la mutazione G1676A nel gene ATRX ai suoi due figli affetti. In entrambi i casi le mutazioni sono risultate associate all'allele AR inattivato; poiché sia ATRX che NLGN3 sono in stretto linkage con il locus AR, in cui un evento di ricombinazione è meno probabile.

Nella famiglia singleton, la probanda mostra una XCI estremamente sbilanciata (92:8) e una mutazione *de novo* R294X nel gene MECP2. Infine, come atteso, nella famiglia con la mutazione nel gene NLGN4X (Cap. 4 § 4.2.2) la madre portatrice ha un pattern di XCI bilanciato, poiché questo gene sfugge l'inattivazione (Carrel and Willard, 2005).

Per quanto riguarda l'analisi del profilo di XCI per l'intero campione analizzato (Paris, Finland e IMGSAC) non sono emerse differenze statisticamente significative dal confronto del pattern di XCI di tutte le madri informative incluse nello studio (543) con quello dei controlli (144): infatti la distribuzione dei valori di XCI ha un andamento simile nei due gruppi (figura 36).

In seguito all'analisi dell'intero campione per il profilo di XCI, 86 madri di individui affetti, senza mutazioni note, hanno mostrato un pattern di XCI altamente sbilanciato ($\geq 80:20$), 45 delle quali provenienti da famiglie multiplex (tabella 10).

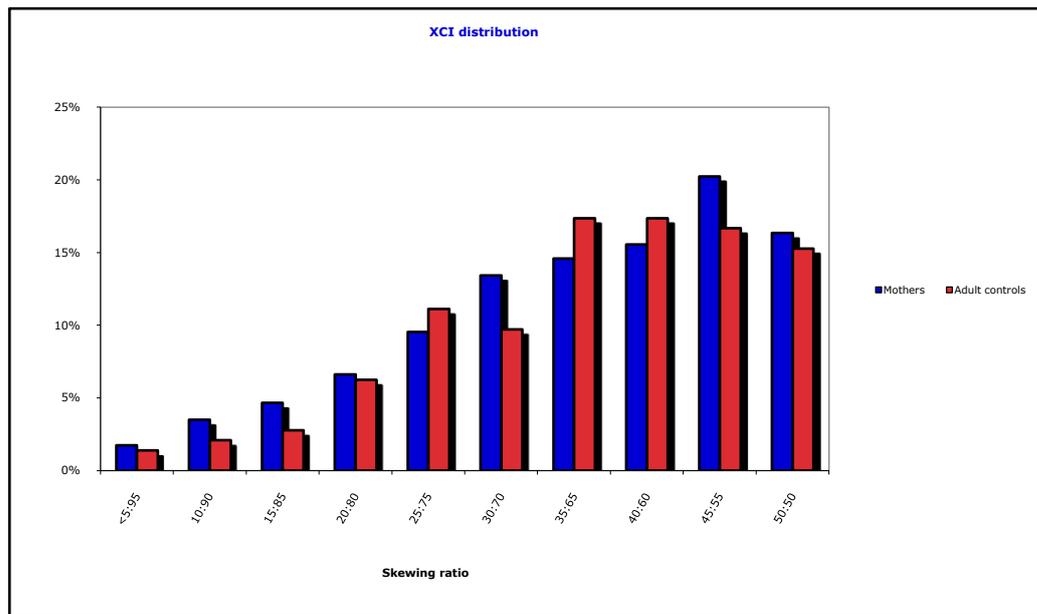


Figura 36: Andamento del pattern di XCI nel campione totale di madri di individui con autismo e nel gruppo di controllo

Samples	N	Informative ^a	XCI > 80:20	<i>P</i> -value ^b	XCI > 90:10	<i>P</i> -value ^b
PARIS						
Mother of MO	200	168	29 (17.3%)	0.155	7 (4.2%)	0.494
Mother of FC	89	78	9 (11.5%)	0.509	3 (3.8%)	0.578
Affected girls	85	75	7 (9.3%)	0.164	1 (1.3%)	0.652
Finland						
Mother of MO	37	31	5 (16.1%)	0.385	2 (6.5%)	0.36
Mother of FC	20	17	4 (23.5%)	0.184	2 (11.2%)	0.161
Affected girls	22	19	3 (15.8%)	0.593	1 (5.3%)	0.322
IMGSAC						
Mother of MO	197	181	28 (15.5%)	0.274	9 (5.0%)	0.353
Mother of FC	78	68	11 (16.2%)	0.299	3 (4.4%)	0.502
Affected girls	75	69	5 (7.2%)	0.093	1 (1.4%)	0.633
All samples						
Mother of MO	434	380	62 (16.3%)	0.172	18 (4.7%)	0.359
Mother of FC	187	163	24 (14.7%)	0.346	8 (4.9%)	0.37
All mothers	621	543	86 (15.8%)	0.195	26 (4.8%)	0.339
Affected girls	182	163	15 (9.2%)	0.112	3 (1.8%)	0.516
Controls						
Adult controls	162	144	18 (12.5%)		5 (3.5%)	
Girls controls	47	40	7 (17.5%)		0	

^aNumber of families without known mutations informative for XCI analysis.

^b*P*-value for one-sided Fisher's exact test.

Tabella 10. Risultati ottenuti per l'intero campione europeo (Gong et al., 2008)

9.1.4 Analisi di Linkage

Un totale di 185 famiglie multiplex informative per l'XCI e appartenenti ai gruppi IMGSAC e PARIS sono state analizzate per la presenza di linkage tra l'autismo e il cromosoma X mediante il metodo NPL (Cap. 1 § 1.3.3). Per tali famiglie erano disponibili i dati genotipici provenienti dallo studio collaborativo dell'Autism Genome Project, generati mediante l'Affymetrix 10K v2 SNP array (AGP, 2007). I risultati ottenuti sono mostrati in figura 20. Nessuna evidenza di linkage è stata riscontrata nel campione totale di 185 famiglie analizzato. Effettuando invece l'analisi di linkage per un sottogruppo di 28 famiglie con un pattern di XCI $\geq 80:20$, coerenti con un modello di ereditarietà legato all'X, si è ottenuto un incremento della condivisione allelica, con un valore di $Z = 1.75$ (LOD 0.85) in corrispondenza del marcatore rs719489, in Xq27-28 (figura 37). Il P value ottenuto mediante permutazioni non risulta però significativo ($P = 0.39$). Questo risultato può essere dovuto alle piccole dimensioni del campione, tuttavia sembra indicare l'assenza di un locus principale sul cromosoma X nel sottogruppo di famiglie con XCI sbilanciata.

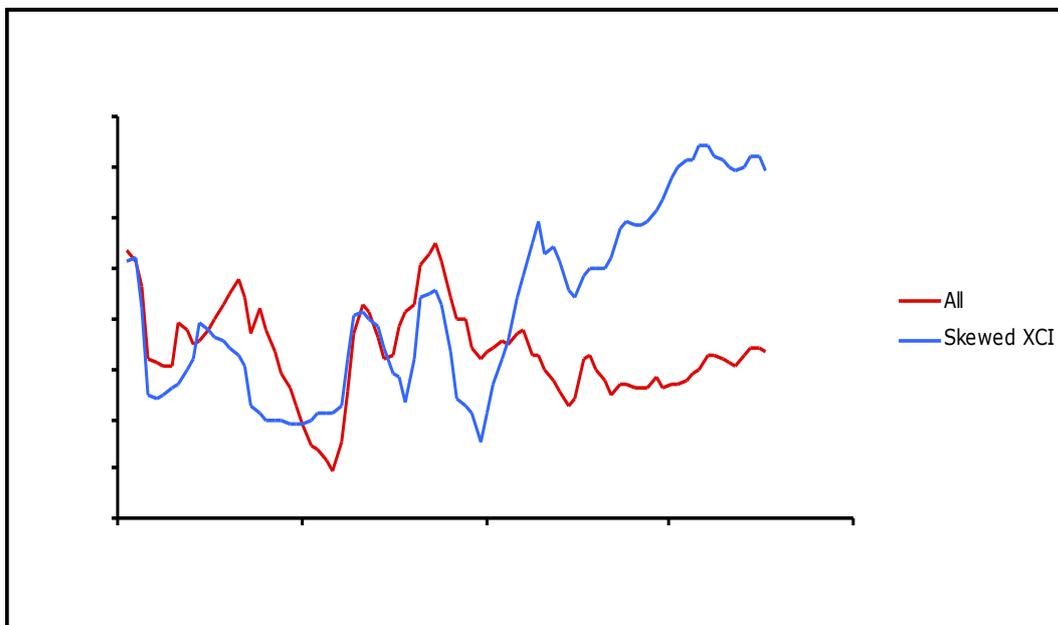


Figura 37. Analisi di linkage non parametrica effettuata su 185 famiglie multiplex e sul sottogruppo di 28 famiglie con un pattern di XCI $\geq 80:20$.

9.1.5 Screening di mutazioni nei geni *FMR1*, *Mecp2*, *NLGN3*

Nella regione Xq27-28 mappano i geni MECP2 e FMR1, entrambi candidati per il disturbo autistico (Cap. 4 § 4.2.1 e 4.2.3). Per identificare eventuali varianti di sequenza associate all'autismo, i geni sono stati sottoposti a screening di mutazioni mediante sequenziamento in un campione comprendente 22 individui appartenenti a famiglie multiplex, selezionati tra le famiglie con XCI \geq 80:20, coerente con un modello d'ereditarietà X-linked e con condivisione allelica nella regione Xq27-q28. Un campione aggiuntivo di 11 probandi, figli di madri con profilo XCI \geq 80:20, appartenenti a famiglie singletons del consorzio IMGSAC è stato analizzato per il gene *Mecp2*.

Lo screening, sugli stessi individui, è stato effettuato anche per il gene *NLGN3*, poiché precedentemente era stata trovata una mutazione (R451C) in questo gene in due fratelli affetti da disturbi dello spettro autistico (Jamain *et al.*, 2003) la cui madre portatrice è risultata avere un XCI sbilanciata (Cap. 4 § 4.2.2). L'analisi di mutazioni per il gene *NLGN3* è stato condotto su un campione, selezionato con gli stessi criteri descritti precedentemente, di 33 individui affetti, di cui 22 appartenenti a famiglie multiplex ed 11 a famiglie singleton.

Nessuna mutazione è stata riscontrata per i geni *FMR1*, *MECP2* e *NLGN3* negli individui sottoposti a screening. Le poche varianti identificate corrispondevano infatti a polimorfismi già noti, privi di significato funzionale.

9.1.6 Analisi di varianti strutturali nel gene *Mecp2*

Lo studio di varianti strutturali nel gene *MECP2* è stata condotta nei 22 individui affetti appartenenti a famiglie multiplex sottoposti a screening di mutazioni con profilo XCI \geq 80:20. L'analisi è stata condotta con PCR quantitativa, mediante un saggio di Real-Time. Nello specifico, sono state utilizzate due sonde specifiche che mappano nel gene *MECP2* (*mecp2-2a* e *mecp2-5'*) ed una sonda, utilizzata come controllo interno per la normalizzazione dei dati nel locus *NFX5* (figura 38).

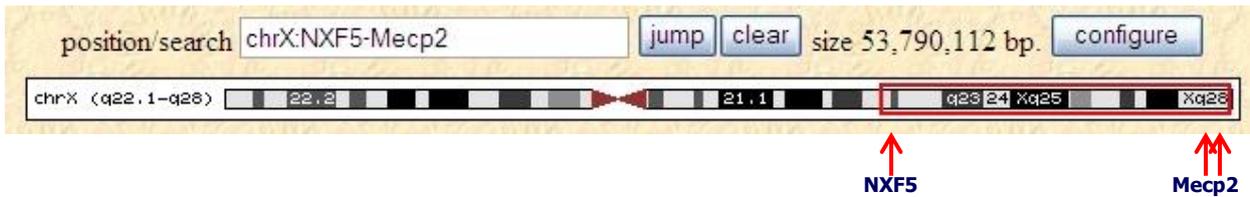


Figura 38. Localizzazione e numero delle sonde usate per la qPCR sul cromosoma nel gene Mecp2.

La quantificazione si basa sul calcolo del $\Delta(\Delta C_t)$ tra i campioni incogniti e il campione di controllo. Gli esperimenti sono stati condotti in triplicato per ciascun campione, utilizzando in ciascun esperimento un controllo normale maschio o femmina (a seconda del sesso dei campioni in analisi) ed un controllo positivo per la duplicazione nel locus Mecp2 (un maschio o una femmina a seconda del sesso dei campioni in analisi).

I risultati ottenuti per il campione di individui affetti appartenenti a famiglie multiplex sono mostrati in figura 39.

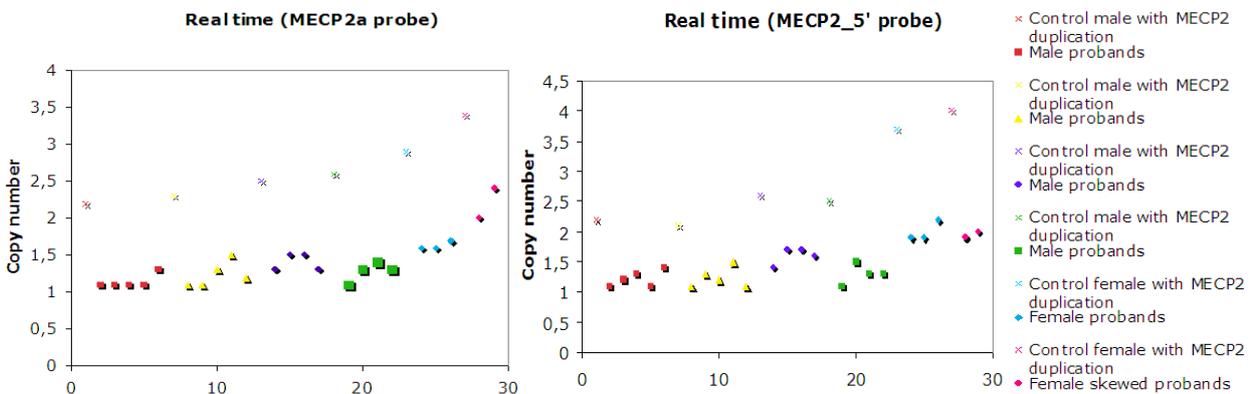


Figura 39: Risultati della qPCR ottenuti con entrambe le sonde locus specifiche sul campione di individui affetti appartenenti a famiglie multiplex. Per ogni gruppo di individui affetti analizzati è presente il controllo positivo per la duplicazione.

Come si può osservare, in nessun caso è stata identificata la presenza di duplicazioni comprendenti il gene MECP2 negli individui analizzati.

Un campione aggiuntivo di 11 individui affetti da autismo appartenenti al campione di famiglie singleton con profilo $XCI \geq 80:20$ è stato analizzato per varianti strutturali

nel gene *Mecp2* mediante MLPA (Multiplex Ligation Probe Amplification). In questo saggio vengono utilizzate un numero elevato di sonde specifiche per la regione genomica in esame ed un numero di sonde di controllo per la normalizzazione del segnale. Le sonde vengono fatte ibridare in maniera specifica al loro locus, in seguito viene effettuata una ligazione delle sonde adiacenti e quindi amplificate con dei primers fluorescenti. L'analisi dei dati è stata condotta mediante l'utilizzo del software Coffalyser v5 che si basa sul rapporto dei rapporti dell'area della sonda specifica e della sommatoria delle aree delle sonde di controllo tra individuo affetto e controllo (vedi figura 40).

$$\text{Copy Number} = \frac{\text{Caso}}{\text{Controllo}} = \frac{A\text{-sonda} / \sum A\text{-sonde controllo}}{A\text{-sonda} / \sum A\text{-sonde controllo}} = \begin{cases} > 1.3 = \text{duplicazione} \\ < 0.7 = \text{delezione} \end{cases}$$

Figura 40: Schema della formula generale che viene utilizzata per effettuare l'analisi condotta per MLPA.

Gli individui affetti analizzati non presentano alcuna variante strutturale in *Mecp2*, come evidenziato dalla figura 41.

P015C MLPA				sample	10.3	11.3	16.3	18.3	31.3	337	60.3	65.3	72.3	88.4	25.3	6.3^A	88.3	889
				Nr Probe	28	28	28	28	28	28	28	28	28	28	28	28	28	28
				Nr Control	17	17	17	17	17	17	17	17	17	17	17	17	17	17
				Ligation F	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
				Male / Fe	Male	Male	Male	Male	Male	Male	Male	Male	Male	Male	Male	Female	Female	Female
Gene	Chr pos	Length	MV36	Recomn	Ratio	Ratio	Ratio	Ratio	Ratio	Ratio	Ratio	Ratio	Ratio	Ratio	Ratio	Ratio	Ratio	Ratio
					controllo+										controllo+			
L1CAM probe 0819-L0337	Xq28	319	X-152.8	1	1.05	0.95	1.09	1.01	1.05	2.02	1.1	1.17	1.08	1	1.08	0.92	1.07	1.47
IRAK1 probe 1770-L1334	Xq28	418	X-152.9	2	1.01	0.87	0.86	0.99	1.03	1.72	0.9	0.93	0.99	0.9	1.01	0.91	0.98	1.69
MECP2 probe 2002-L1335	Xq28	292	3.0 exon 01A	3	0.94	0.86	0.99	0.92	0.98	1.6	1.06	0.94	1.01	1.01	0.92	0.91	1.03	1.95
MECP2 probe 1350-L0893	Xq28	391	3.0 exon 01B	4	0.85	0.74	0.74	0.81	0.83	1.28	0.8	0.83	0.86	0.82	0.86	0.77	0.89	2.06
MECP2 probe 1349-L0894	Xq28	373	53.0 exon 02	5	1.04	0.79	0.89	0.93	0.89	1.68	1.05	1.04	1.03	0.83	0.93	0.97	1.04	1.81
MECP2 probe 1348-L0895	Xq28	355	53.0 exon 03	6	1.11	0.95	1.06	1.09	0.96	1.95	1.05	1.11	1.16	1.1	0.89	0.98	1.02	1.47
MECP2 probe 1347-L0896	Xq28	337	3.0 exon 04A	7	1.03	0.84	1.15	1.07	1.02	1.94	1.17	1.08	0.98	1.05	0.91	0.95	1.07	1.53
MECP2 probe 3026-L2372	Xq28	229	3.0 exon 04B	8	0.75	0.91	0.91	0.98	1.06	1.21	0.95	1.02	0.74	0.95	0.88	0.83	0.92	1.2
MECP2 probe 1768-L1332	Xq28	274	3.0 exon 04C	9	1.15	0.99	1.11	1.06	1.04	1.87	0.98	1.13	0.98	0.97	0.9	0.92	1.02	1.33
MECP2 probe 1769-L1333	Xq28	256	3.0 exon 04D	10	1.12	1.08	1.08	1.04	1.09	2.04	1.24	1.12	1.08	1	1.03	0.97	0.96	1.35
SYBL1 probe 1094-L0659	Xq28	454	X-154.8	11	1.1	1	1.07	1.08	1.03	1.52	1	1.01	0.98	1.06	0.95	0.87	0.95	1.19
Control probe 1633-L1171	11q23	283	c	12	0.92	0.86	0.9	0.93	0.9	0.62	0.95	0.88	0.77	1	0.88	0.89	0.87	0.87
Control probe 0495-L0070	12p12.3	310	c	13	0.98	0.99	0.95	1	1	0.9	1.01	0.93	1.01	1.05	0.84	1.01	1.03	1.01
Control probe 1241-L0783	17p13	436	c	14	1.14	1.04	1.1	1.01	1.05	1	1.11	1.1	1.01	1.13	1.12	1.02	1.01	1.25
Control probe 0810-L0328	18q21.3	346	c	15	1	1.04	1.04	1.02	1	1.11	0.95	1.03	1.02	1.12	1.13	1.06	1.1	1.01
Control probe 1368-L1016	Xp21.2	265	c	16	1.06	0.93	0.99	0.97	0.87	0.99	0.93	0.98	0.84	1.05	0.96	0.99	1.03	1.03
Control probe 1375-L1023	Xp21.2	328	c	17	1.13	1	1.06	1.06	1.04	1.51	1.18	1.1	1.11	0.98	1.13	0.96	0.99	1.19
Control probe 1419-L1994	Xp21.2	364	c	18	1.01	1	1.01	0.99	0.93	1.11	0.87	0.97	0.96	1.04	1.15	1.02	0.98	0.97
Control probe 1716-L1284	Xp21.2	400	c	19	1.06	1.02	1.05	1.07	0.99	1.26	1.04	1.1	1	1.08	1.05	0.99	1.01	1
Control probe 1899-L1049	Xp21.2	210	c	20	0.95	0.91	0.99	0.96	0.96	1.05	0.89	0.89	0.9	1	1.06	1.02	1.07	0.91
Control probe 0880-L0465	Xp22.1	445	c	21	0.91	0.92	0.79	0.94	1.02	0.74	1	0.92	0.99	0.87	0.84	1.02	1.01	0.95
Control probe 1147-L0802	b22.32-PA	247	c	22	1.11	1.1	1.04	1.02	1.07	0.6	1.06	1.06	1.09	1.01	0.89	1	1	0.46
Control probe 1152-L0709	b22.32-PA	427	c	23	0.97	0.92	1.01	0.93	1.16	0.28	0.96	1.05	0.88	1.01	1	0.98	0.97	0.44
Control probe 1271-L0423	Xq11.2	238	c	24	1.07	1.06	0.97	1.06	1.01	1.02	0.99	1	1.02	0.99	1	0.97	1	0.83
Control probe 1279-L0963	Xq22	301	c	25	0.94	1.02	0.94	0.95	0.94	0.95	1	0.97	0.92	0.95	0.94	1	0.98	0.87
Control probe 0820-L0338	Xq25	409	c	26	0.97	1.04	1	1.07	1.24	0.92	1.19	1.14	1.07	0.93	0.9	1.01	1.06	1.12
Control probe 1881-L1450	Xq28.1	220	c	27	1.1	1.03	1.08	1.04	0.99	1.9	1.12	1.08	1.06	1	1	0.89	0.94	1.16
Control probe 1887-L1456	Xq28.1	382	c	28	0.94	0.73	0.86	0.79	0.88	1.46	0.93	0.81	0.84	0.78	0.94	0.86	0.95	2.17

Figura 41: Risultati delle varianti strutturali negli individui di famiglie singleton con profilo XCI \geq 80:20. A sinistra l'elenco delle sonde utilizzate in questo studio; in rosso sono evidenziate le sonde specifiche per il gene Mecn2, il restante sono sonde di controllo. Il campione è suddiviso in individui maschi e femmine con relativo controllo positivo maschio e femmina per la duplicazione. I valori maggiori di 1.3 indicano una duplicazione in quel locus e sono segnate in verde, i valori minori di 0.7 indicano una delezione e sono segnate in rosso. In nessun probando è presente un valore che suggerisce una duplicazione, mentre nel controllo positivo viene individuata la duplicazione nel locus di Mecn2 ed una delezione già nota della regione PAR1 (Xp22).

9.2.1 Screening di mutazione del gene ASMT

In seguito alle evidenze dello studio di Melke et al., che suggeriscono un coinvolgimento del gene ASMT alla suscettibilità nell'autismo, abbiamo condotto un'analisi di mutazioni su tutti gli esoni di questo gene utilizzando un ampio campione di famiglie con autismo. In particolar modo, l'analisi è stata condotta in 3 collezioni di famiglie indipendenti, appartenenti al consorzio di studio europeo per l'autismo:

- 1) 127 famiglie finlandesi, costituite da 83 famiglie singletons e 44 famiglie multiplex, e 100 controlli finlandesi
- 2) 69 famiglie singletons italiane e 200 controlli italiani
- 3) 194 famiglie multiplex appartenenti al consorzio IMGSAC e 190 controlli ECACC, per lo più inglesi.

Nel presente studio sono stati utilizzati due campioni di famiglie indipendenti, quello di famiglie singletons italiane e quello di famiglie multiplex IMGSAC con i relativi individui di controllo.

L'analisi di mutazione sui 9 esoni di ASMT e sulle regioni del promotore A e B è stata condotta su un primo campione di 47 individui autistici del campione italiano. Il risultato è schematizzato in figura 42.

47 individui	Nuove varianti identificate
EX 1a	1 intr. G/A-134 from start site (1758041)
prom_1b	/
EX 1b	1 intr. G/C -157 from start site (1777660)
	1 eson. C/T +62 (1777878)
EX 2	/
EX 3	/
EX 4	/
EX 5	1 IVS5+2T<C (1792661)
EX 6	1 G53A [Q205Q] (1795476)
EX 7	/
EX 8	/
	/
EX 9	1 C/T 45 [R319*] (1805567)
	1 C/T 66 [L326F] (1805588)

Figura 42: schematizzazione dello screening di mutazione sui 9 esoni e dei promotori A e B del gene ASMT nel campione italiano di 47 individui.

Dalla prima analisi di questo campione, seppur non ampio, sono emersi interessanti dati che provvedono a supportare il precedente studio. Infatti, vengono individuate le due varianti già descritte (Melke et al.): la variante di splicing (IVS5+2T>C) e la sostituzione non sinonima (L326F). La prima variante viene trasmessa dalla madre al figlio affetto, ma anche alla figlia non affetta (figura 43); la seconda viene trasmessa dal padre alla figlia affetta, ma anche alla figlia non affetta (figura 43). Inoltre, è stata identificata una variante di stop (R319*), mai descritta precedentemente, che viene trasmessa dalla madre al figlio affetto (figura 43)

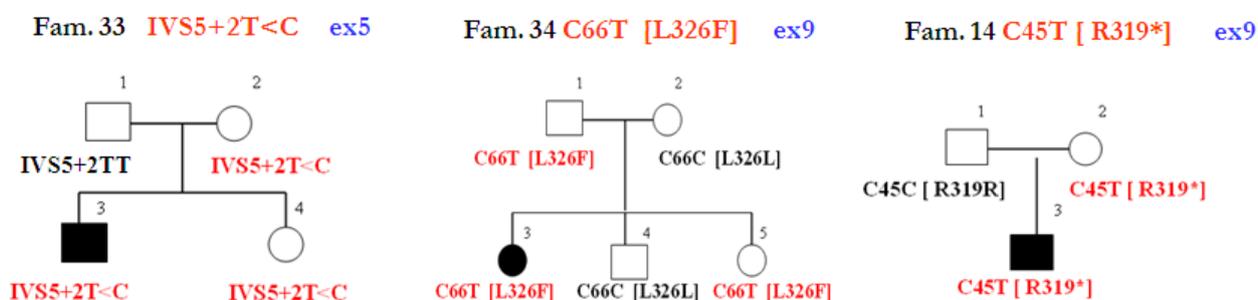


Figura 43: segregazione delle varianti individuate in seguito allo screening di mutazioni nelle tre famiglie italiane

La presenza nel campione di 47 individui autistici italiani delle varianti descritte da Melke et al., e della variante di stop, ci ha spinto ad indagare maggiormente la frequenza delle stesse, analizzando un vasto campione di individui autistici. La successiva analisi è stata condotta sul restante campione italiano di 22 individui autistici e 194 individui appartenenti alle famiglie multiplex IMGSAC. Per determinare la frequenza delle varianti di sequenza incontrate si è effettuato l'analisi di mutazione esclusivamente sull'esone 5 e sull'esone 9. Il risultato di questa analisi nel successivo campione preso in esame ha individuato la variante di splicing già descritta precedentemente (IVS5+2T>C) in una famiglia multiplex che viene trasmessa dal padre a solo uno dei due figli affetti (figura 44).

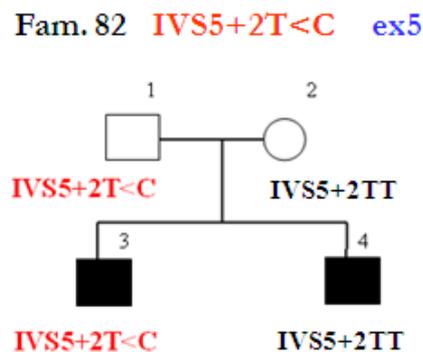


Figura 44: segregazione della variante (IVS5+2T>C) nella famiglia multiplex

Succesivamente abbiamo voluto stabilire se le varianti di sequenza incontrate nel campione di individui autistici analizzato erano presenti anche nella popolazione generale, quindi abbiamo condotto la stessa analisi degli esoni 5 e 9 in un vasto campione di individui di controllo. In particolar modo i due campioni di controllo sono stati selezionati in base all'origine dei casi: 200 individui italiani di controllo e 190 individui della collezione di controlli ECACC (European Collection of Cell Cultures) per lo più di origine inglese. L'analisi di questo campione ha portato ad individuare la variante di splicing (IVS5+2T>C) in un controllo italiano, la missenso (L326F) in 5 individui di controllo, mentre la mutazione di stop (R319*) non è stata individuata in nessuno dei 390 controlli analizzati in questo studio. In tabella 11 è

presente una schematizzazione generale delle varianti incontrate in questo studio con lo studio effettuato dai nostri collaboratori finlandesi (Toma C. et al, 2007). Complessivamente gli esoni 1b, 5 e 9 sono stati analizzati in un totale di 390 individui autistici nelle tre popolazioni (tabella 11), mentre il resto degli esoni sono stati analizzati in 194 casi (127 individui finlandesi, 47 individui di famiglie singleton italiani e 20 individui delle famiglie multiplex IMGSAC).

Varianti rare (NM_004043)	FINLANDESI		ITALIANI		IMGSAC		CAMPIONE TOTALE	
	Casi (n=127)	Contr. (n=100)	Casi (n=69)	Contr. n=(200)	Casi n=(194)	Contr. n=(190)	Casi (%) (n=390)	Contr. (%) (n=490)
IVS5+2T<C	1	0	1	1	1	0	3 (0.77)	1 (0.20)
E316D	0	0	0	0	1	1	1 (0.26)	1 (0.20)
R319*	0	0	1	0	0	0	1 (0.26)	0
L326F	4	2	1	2	0	3	5 (1.28)	7 (1.43)

Tabella 11: schematizzazione delle varianti rare in ASMT nei casi e nei controlli.

9.2.2 Studio di associazione dei polimorfismi del promotore B del gene ASMT

Ulteriore evidenza del coinvolgimento del gene ASMT nell'autismo è l'associazione con l'aplotipo H1 (GGGC) nel promotore B che potrebbe influenzarne l'espressione (Melke et al, 2008).

Per stabilire se questo aplotipo fosse un fattore di rischio per l'autismo abbiamo condotto uno studio di replicazione utilizzando un ampio campione totale di 390 casi e di 382 controlli. Lo studio condotto nel nostro laboratorio ha preso in esame il campione italiano e quello IMGSAC per un totale di 263 individui autistici ed un totale di 282 individui di controllo. Gli SNPs: rs4446909, rs5989681, rs56690322 e rs6644635 (figura 45) che definiscono gli aplotipi del promotore sono stati tipizzati mediante sequenziamento diretto.

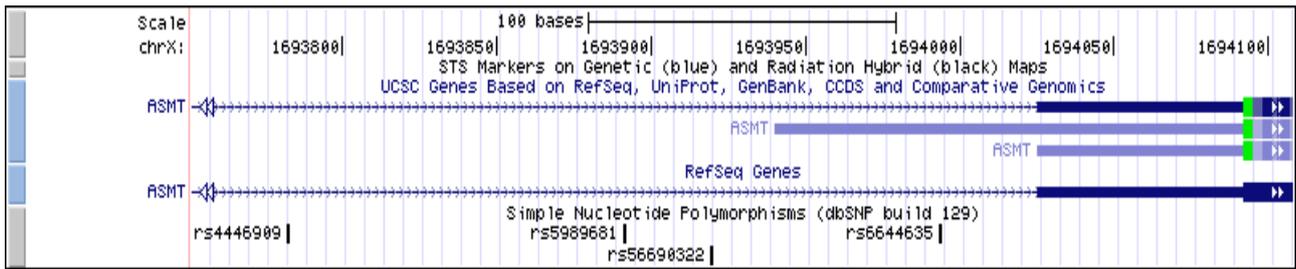


Figura 45: localizzazione degli SNPs del promotore B del gene ASMT secondo l'annotazione del database genomico UCSC.

Abbiamo condotto quindi uno studio di associazione caso/controllo effettuando l'analisi per singolo marcatore e per aplotipo utilizzando il software Haploview v4.0. I risultati sono riassunti in tabella 13.

	<i>Finnish</i>			<i>Italian</i>			<i>IMGSAC</i>		
	<i>Cases (127)</i>	<i>Control (100)</i>	<i>P (permP)^a</i>	<i>Cases (69)</i>	<i>Control (90)</i>	<i>P (permP)^a</i>	<i>Cases (194)</i>	<i>Control (192)</i>	<i>P (permP)^a</i>
<i>SNP (allele)</i>									
rs4446909 f(G)	0.89	0.87	0.64 (0.97)	0.76	0.74	0.74 (0.99)	0.76	0.74	0.43 (0.86)
rs5989681 f(G)	0.87	0.85	0.65 (0.97)	0.72	0.71	0.90 (1.00)	0.72	0.71	0.81 (1.00)
rs56690322 f(G)	0.90	0.90	0.98 (1.00)	0.88	0.87	0.75 (1.00)	0.88	0.91	0.18 (0.48)
rs6644635 f(C)	0.64	0.57	0.12 (0.33)	0.61	0.54	0.25 (0.57)	0.62	0.65	0.34 (0.76)
<i>Haplotypes</i>									
H1 GGGC	0.50	0.42	0.06 (0.21)	0.32	0.25	0.18 (0.59)	0.32	0.35	0.38 (0.88)
H2 GGGT	0.27	0.35	0.08 (0.25)	0.28	0.34	0.27 (0.72)	0.27	0.27	0.91 (1.00)
H3 ACGC	0.11	0.13	0.54 (1.00)	0.22	0.24	0.68 (1.00)	0.22	0.26	0.27 (0.79)
H4 GGAT	0.08	0.08	0.94 (1.00)	0.11	0.12	0.70 (1.00)	0.11	0.08	0.30 (0.83)
H5 GCGC	0.02	0.01	0.66 (1.00)	0.05	0.04	0.61 (1.00)	0.06	0.04	0.13 (0.44)
H6 GGAC	0.01	0.01	0.91 (1.00)	—	—	—	—	—	—

Tabella 13: Sono indicati i polimorfismi a singolo nucleotide (SNP) e la frequenza allelica per ciascun campione di casi e di controllo utilizzati nello studio di associazione. Infine il P value (P) per il test di χ^2 condotto per singolo marcatore o per aplotipo, in parentesi il valore di p-value effettuando 10.000 permutazioni per ottenere una correzione del valore per test multipli.

Come si può osservare nessuno dei marcatori singoli o la combinazione allelica dei 4 polimorfismi nei tre campioni analizzati in questo studio evidenzia associazione con l'autismo in seguito all'analisi di tipo caso/controllo.

9.3.1 Analisi dei trascritti alternativi del gene ZNF533

I risultati degli studi di associazione di tipo caso/controllo e TDT dell'esperimento ILLUMINA hanno mostrato risultati positivi per il gene *ZNF533* nel locus AUTS5 del cromosoma 2. La funzione del gene non è nota, e da dati bioinformatici risulta contenere 4 zinc fingers ed una potenziale localizzazione nucleare. L'organizzazione genomica del gene è alquanto complessa come si può apprezzare in figura X (Genome Browser 2004 Assembly). Per ottenere maggiori informazioni sulle reali isoforme o eventuali esoni alternativi, espressi a livello del cervello nel gene *ZNF533*, abbiamo condotto un'analisi dei trascritti tra tutti quelli descritti nel UCSC Genome Browser. In particolar modo abbiamo preso in esame gli esoni start delle 4 principali isoforme schematizzate in figura 46 per i tessuti di cervello fetale e placenta. Il disegno sperimentale è schematizzato in figura 47.

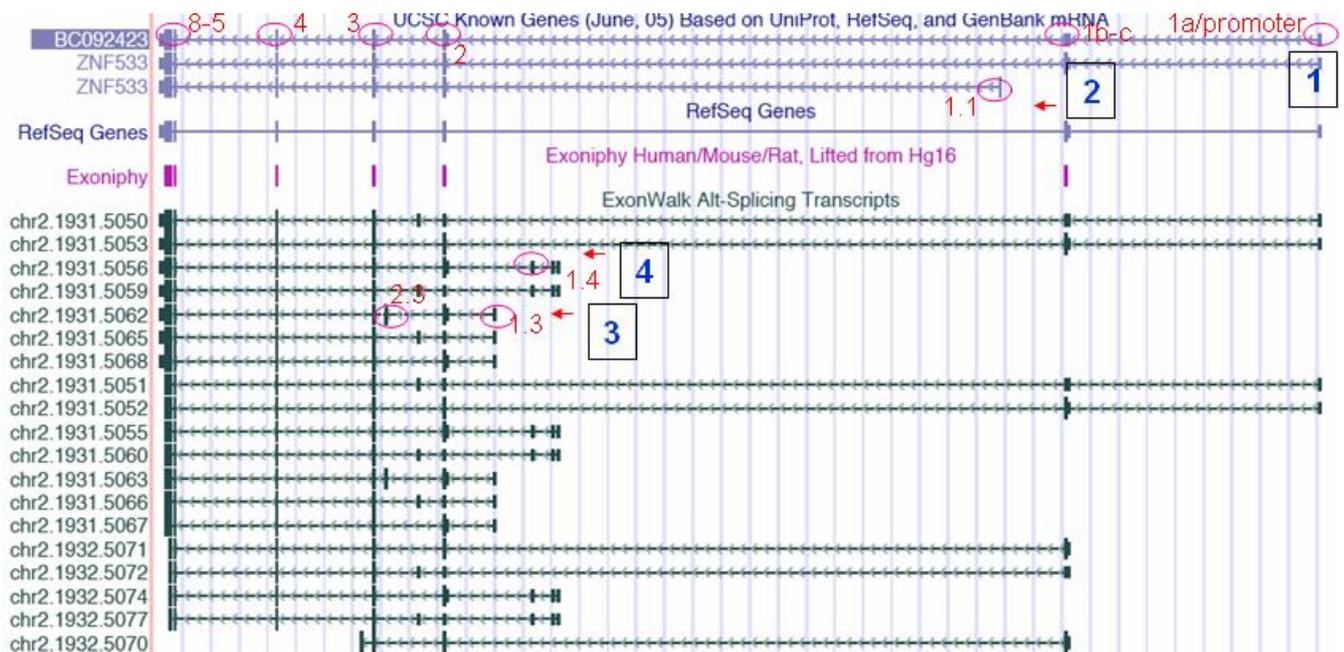


Figura 46: Organizzazione genomica del gene *Znf533*. Sono cerchiati in viola gli esoni delle distinte isoforme. I quadrati indicano le 4 isoforme che abbiamo considerato nel nostro studio di analisi di start alternativi del gene.

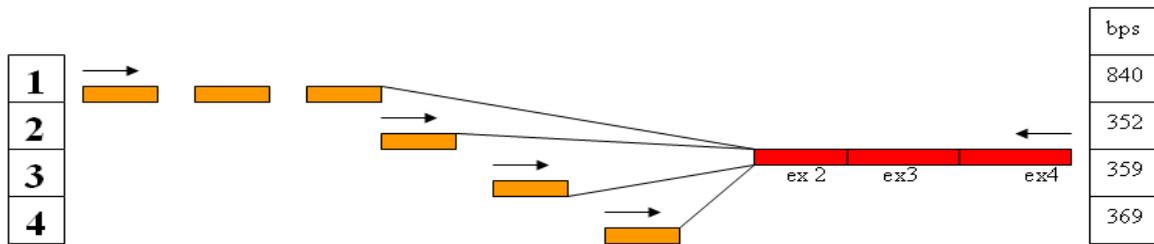


Figura 47: schematizzazione delle 4 isoforme del gene znf533 e della strategia utilizzata per l'analisi delle sue distinte isoforme in placenta e cervello fetale. Per ciascuna isoforma abbiamo utilizzato un primer Forward (freccia) disegnato in ciascun primo esone ed un primer Reverse comune nell'esone 4.

Ogni frammento è stato isolato e sequenziato. I risultati dimostrano la trascrizione a partire dai 4 start alternativi e suggeriscono un'espressione regolata nei differenti tessuti (figura 48); in particolar modo tutte le isoforme sono espresse in cervello. Inoltre come si può osservare in figura è presente una elevata variabilità delle distinte isoforme al 5' del gene. Questo studio ci ha portato anche ad individuare un nuovo esone in placenta non descritto nel Genome Browser (esone 1.2), supportato anche dall'elevato stato di conservazione di questa sequenza nei mammiferi (MCS) (figura 49)

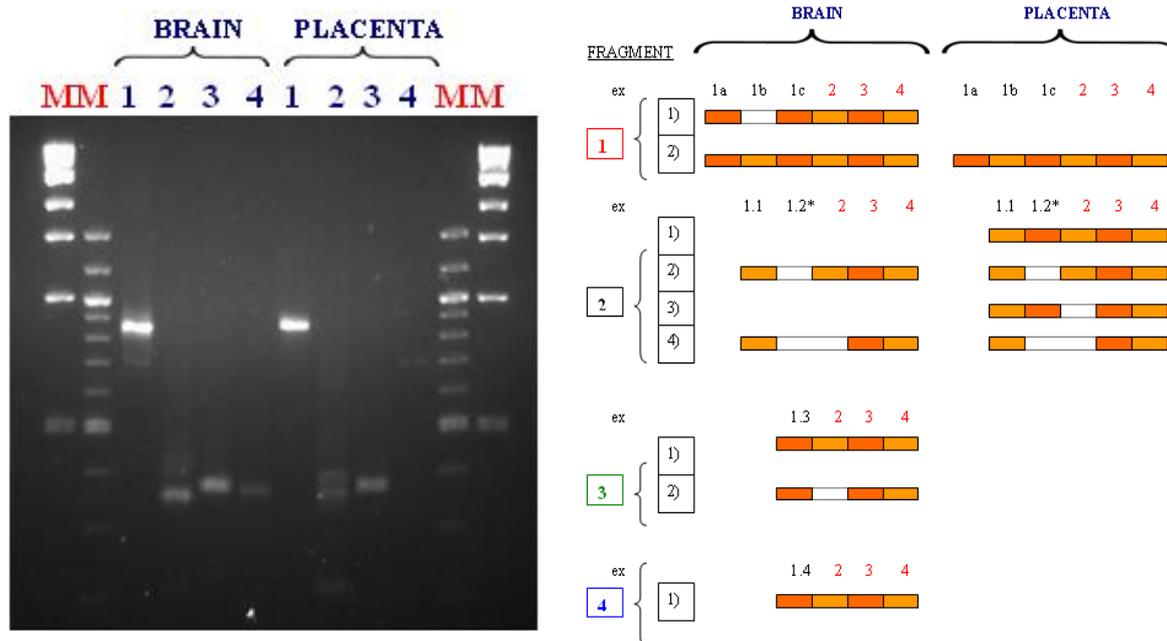


Figura 48: A sinistra la corsa elettroforetica in cui migrano gli amplificati dei 4 start alternativi nei tessuti di cervello fetale e placenta. A destra una schematizzazione dei risultati: sono indicati i 4 start alternativi e per ciascuno le isoforme identificate nei due tessuti presi in esame. In alto a ciascun start alternativo è indicato l'esone atteso: in rosso o arancio è indicato la presenza dell'esone nella corrispondente isoforma, in bianco l'assenza dell'esone.

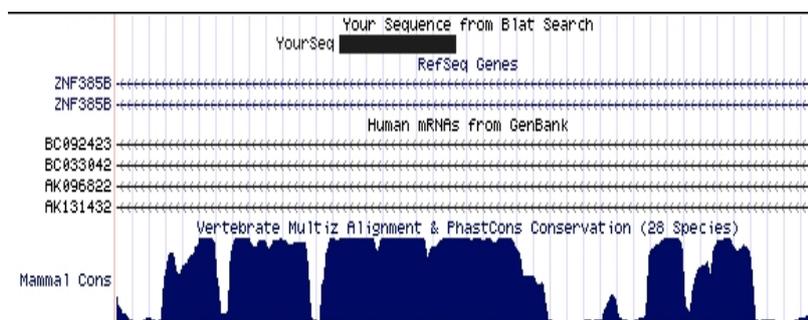


Figura 49: Stato di conservazione dell'esone 1.2 individuato nel nostro studio in placenta, ma non presente nel database genomico UCSC. La sequenza dell'esone corrisponde al rettangolo nero in alto, mentre in basso l'alto stato di conservazione di questa regione nei mammiferi.

Pertanto abbiamo incluso nell'analisi di mutazioni tutti gli esoni presenti nelle distinte isoforme espresse nei tessuti analizzati.

9.3.2 Analisi di mutazioni nel gene ZNF533 in un campione di individui autistici

Gli studi di associazione hanno dato evidenze di più aplotipi associati nel gene *ZNF533* per l'autismo. Per stabilire la presenza di eventuali varianti alleliche comuni o rare coinvolte nell'eziologia della malattia abbiamo condotto una analisi di mutazione sugli esoni del gene *ZNF533*, incluse le regioni regolatorie quali il promotore, tutti i siti di giunzione esone/introne ed i siti di ramificazione. Un campione totale di 31 pazienti autistici è stato selezionato tra le 126 famiglie iniziali dell'esperimento ILLUMINA per la presenza degli aplotipi di rischio nel gene *ZNF533*. Lo screening non ha individuato nessuna variante allelica con un potenziale ruolo eziologico per l'autismo, a parte quattro varianti introniche poco conservate nell'evoluzione (tabella 14).

<i>Znf533</i> (NM_152520)	<i>SNP</i> annotati	<i>MAF</i> (IMGSAC)	<i>Varianti rare</i>	<i>Posizione UCSC may2004</i>	<i>Nr eterozigoti in 31 famiglie</i>
introne 1a			A/T	180551248	1[A/T]
introne 1b	rs7574345	C(0.08)			5[G/A]
esone 1.1 5'UTR	rs2176060	T(0.016)			1[C/T]
introne 1.1	rs17823653	A(0.03)			2[T/A]
introne 1.2	rs7559882	G(0.016)			1[A/G]
promotore 1.3	rs3112931	C(0.125)			8[T/C]
esone 1.3 5'UTR	rs16866781	T(0.016)			1[C/T]
introne 1.2			G/A	180252881	1[G/A]
introne 2			G/T	180235002	1[G/T]
introne 2			C/T	180213290	4[C/T]
introne 3	rs2271759	G(0.26)			10[A/C]+3[C/C]
introne 5	rs2271760	A(0.05)			3[C/T]
ex8 (variante sinonima)	rs2367796 [G1802A] [G353G]	A(0.16)			10[G/A]
ex8 3'UTR	rs5836680	C(0.20)			8[del/C]+2[C/C]

Tabella 14: Schema dello screening di mutazione nel campione di 31 individui autistici selezionati per aplotipi di rischio del gene ZNF533. Sono indicati gli SNPs intronici (in nero) ed esonici (in arancio) per ciascun esone. MAF (Minor Allele Frequency). Eventuali varianti alleliche sono riportate per ciascun individuo.

9.3.3 Analisi di varianti strutturali nella regione *IMMP2L-DOCK4* nella collezione di famiglie *IMGSAC*

Per validare la duplicazione nella famiglia 148, e per stabilire inoltre il contributo di varianti strutturali nella regione *IMMP2L-DOCK4* in famiglie con autismo, abbiamo sviluppato un saggio di QMPSF, che consiste in una PCR multiplex di frammenti fluorescenti (QMPSF, quantitative multiplex PCR of short fluorescent fragments). Questa tecnica ci ha permesso di analizzare contemporaneamente gli esoni 2 e 3 di *IMMP2L*, l'esone 4 di *LRRN3* e l'ultimo esone di *DOCK4* (figura 50)

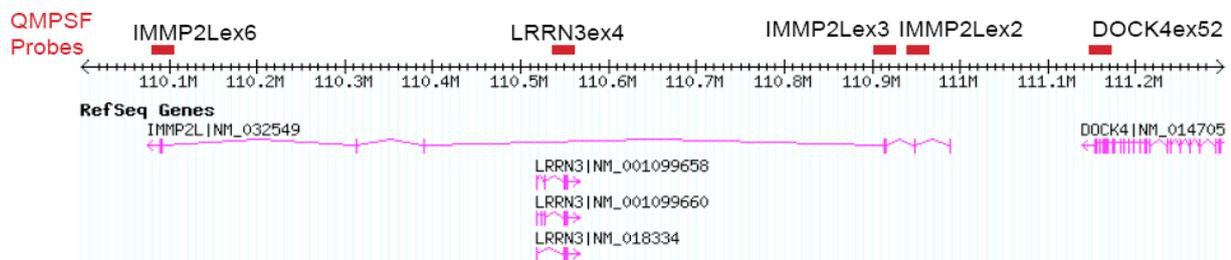


Figura 50: Schematizzazione dell'organizzazione genomica dei geni *IMMP2L*, *LRRN3* e *DOCK4*. In alto le 5 sonde utilizzate nel nostro saggio e la loro localizzazione

Il saggio consiste nella stessa reazione di coamplificazione di sonde marcate fluorescenti per regioni target e sonde per regioni di riferimento. Ogni sonda è caratterizzata da una lunghezza differente (figura 51).

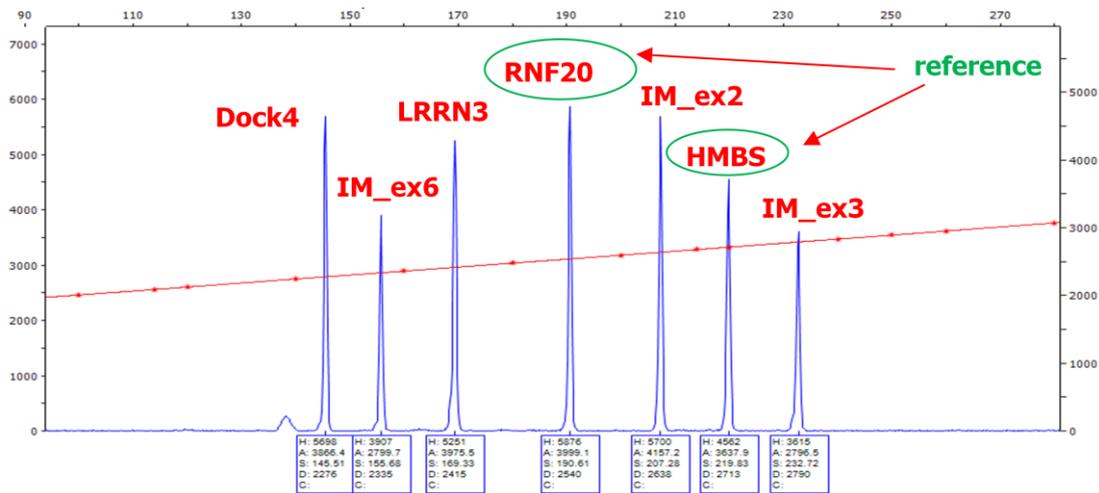


Figura 51: esempio di una reazione in un individuo normale in seguito alla corsa al sequenziatore; è possibile osservare gli amplificati di diversa lunghezza per le sonde specifiche e di riferimento.

L'analisi è stata condotta come rapporto dei rapporti tra l'area della sonda specifica e sonda di riferimento tra paziente e controllo. Per valori maggiori di 1.3 abbiamo un'indicazione di duplicazione nel locus dove mappa la sonda, per valori minori di 0.7 abbiamo indicazione di una delezione (figura 52)

$$\text{Calcolo numero di copie} = \frac{A \text{ Probe spec/ } A \text{ probe reference [paziente]}}{A \text{ Probe spec/ } A \text{ probe reference [controllo]}} = \begin{cases} >1.3 = \text{duplicazione} \\ <0.7 = \text{delezione} \end{cases}$$

Figura 52: Formula utilizzata per calcolare la presenza di CNVs con il saggio QMPSF

Complessivamente l'analisi è stata condotta su un totale di 285 famiglie multiplex del consorzio IMGSAC. I risultati sperimentali della QMPSF hanno validato la presenza della duplicazione nella regione IMMP2L-DOCK4 nella famiglia 148 identificata mediante QuantiSNP. La duplicazione non segrega con il fenotipo autistico in quanto è trasmessa dal padre ad una figlia affetta, ma non all'altro figlio affetto (figura 53). Abbiamo individuato 4 famiglie con delezioni nella regione analizzata, di cui 3 famiglie con delezioni che coinvolgono esclusivamente il gene IMMP2L : una famiglia con una delezione nell'esone 2 e due famiglie con una delezione nell'esone 2 ed esone 3 (figura 53). Le varianti strutturali in IMMP2L non segregano con il

fenotipo autistico (figura 53). L'ultima famiglia, la famiglia 419, presenta una delezione che coinvolge i geni *IMMP2L* e *DOCK4* che viene trasmessa dalla madre ai due figli maschi affetti, ma anche alla figlia non affetta da autismo (figura 53). Tuttavia, la madre e la figlia che hanno la delezione, presentano dislessia. Inoltre, per stabilire se vi fosse una frequenza di CNVs in questa regione maggiore nel nostro campione di casi abbiamo analizzato anche un campione di controllo di 475 individui per lo più di origine inglese (controlli ECACC). Nel campione di controllo sono stati individuati due delezioni nel gene *IMMP2L* in due individui, una delezione dell'esone 2 ed un'altra dell'esone 2 e 3. Nessuna variante strutturale è stata individuata nel campione di controllo a carico del gene *Dock4*.

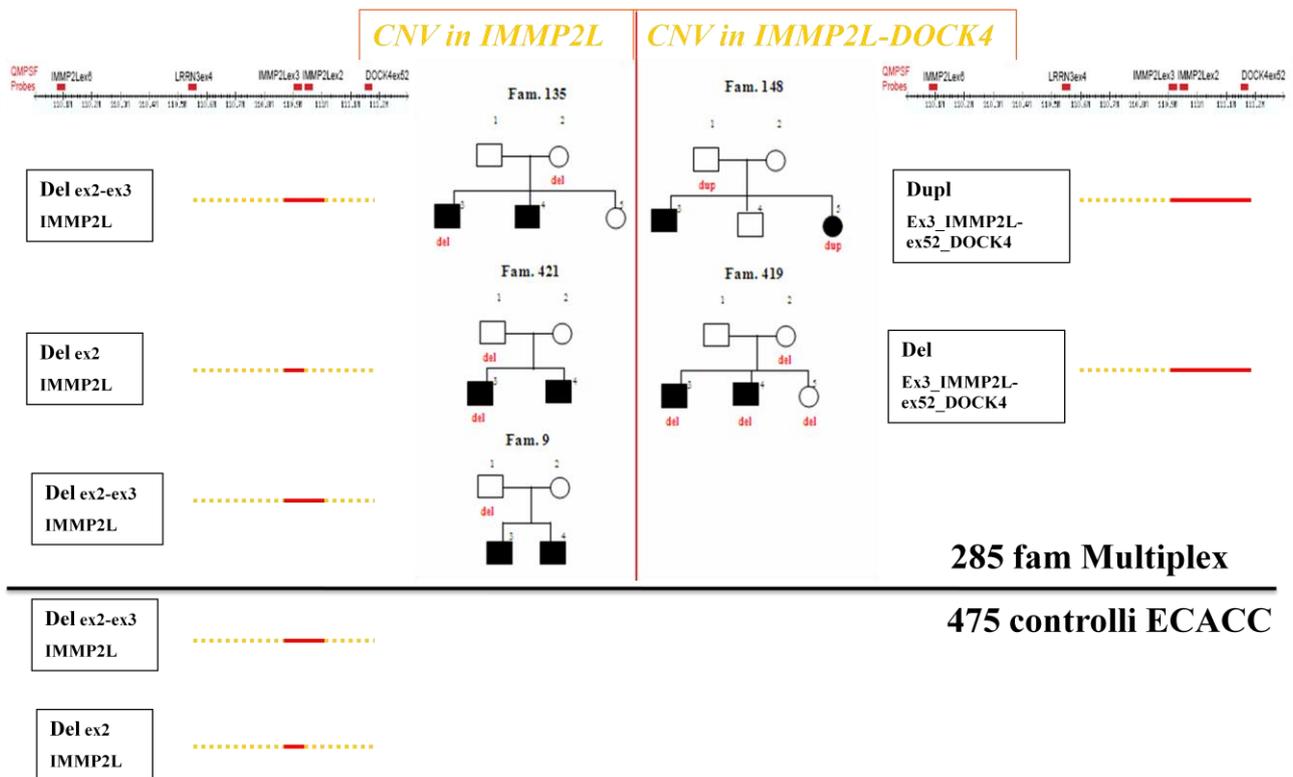


Figura 53: Schematizzazione dello studio di varianti strutturali nella regione *IMMP2L-DOCK4* nel gruppo di casi e di controlli. A sinistra sono indicate le CNVs individuate esclusivamente in *IMMP2L*, e la loro segregazione nelle tre famiglie multiplex, mentre sulla destra quelle individuate in *IMMP2L-DOCK4*, e la schematizzazione della segregazione nelle due famiglie. Le varianti strutturali identificate nei 475 controlli ECACC sono indicate in basso. Per ciascuna variante è indicato il tipo di CNV, in che esone è stato mappato.

Per mappare la delezione distale nella famiglia 419 abbiamo disegnato un nuovo set di sonde lungo il gene DOCK4 (figura 54)

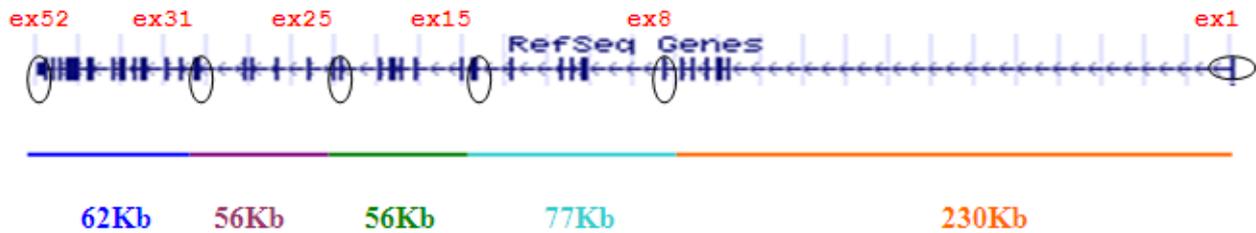


Figura 54: schematizzazione dell'organizzazione genomica del gene DOCK4; sono indicate in alto le sonde specifiche utilizzate nei diversi esoni; in basso le distanze in Kb tra una sonda e la successiva.

I risultati evidenziano che il punto di rottura distale della delezione nella famiglia 419 è tra l'esone 31 e l'esone 25 (tabella 15)

Probe	Campione	Rprobe/Sref	R_Control	CNV
DOCK_ex8	419.1	0,58	0,57	1,02
DOCK_ex15	419.1	0,73	0,77	0,96
DOCK_ex18_19	419.1	0,17	0,15	1,14
DOCK_ex25	419.1	0,32	0,27	1,16
DOCK_ex31	419.1	0,53	0,48	1,10
DOCK_ex52	419.1	0,75	0,78	0,96
DOCK_ex8	419.2	0,52	0,57	0,92
DOCK_ex15	419.2	0,65	0,77	0,84
DOCK_ex18_19	419.2	0,13	0,15	0,86
DOCK_ex25	419.2	0,29	0,27	1,06
DOCK_ex31	419.2	0,21	0,48	0,44
DOCK_ex52	419.2	0,36	0,78	0,47
DOCK_ex8	419.3	0,56	0,57	0,99
DOCK_ex15	419.3	0,77	0,77	1,00
DOCK_ex18_19	419.3	0,17	0,15	1,14
DOCK_ex25	419.3	0,33	0,27	1,21
DOCK_ex31	419.3	0,30	0,48	0,63
DOCK_ex52	419.3	0,41	0,78	0,53

Tabella 15: L'analisi condotta per rilevare CNVs mediante il saggio di QMPFSF. Nell'esempio l'esperimento condotto per mappare il sito distale della delezione nella famiglia 419. In azzurro i valori nell'intervallo 0.7-1.3 che indicano la presenza di due copie per ciascuna sonda locus specifica. In rosso i valori minori di 0.7 che indicano la presenza della delezione per quel locus. E' possibile osservare l'assenza della delezione nel padre (419.1) e la delezione, il cui punto di rottura distale è tra l'esone 31 e 25, trasmessa dalla madre (419.2) ai tre figli (dati del probando 419.3).

10. Discussione

L'autismo è un grave disturbo neuropsichiatrico, caratterizzato dalla compromissione della sfera relazionale e comunicativa, e da comportamenti stereotipati e ripetitivi. Gli studi epidemiologici condotti negli ultimi venti anni hanno dimostrato il coinvolgimento di fattori genetici nell'eziologia dell'autismo. Tali fattori tuttavia presentano un modello di ereditarietà complesso, con probabile presenza di un elevato grado di eterogeneità genetica ed effetti epistatici, che hanno reso l'identificazione delle cause genetiche del disturbo autistico un traguardo estremamente difficile.

Durante lo svolgimento del Dottorato, ho contribuito al progetto del consorzio internazionale per l'autismo IMGSAC, con l'obiettivo di individuare geni di suscettibilità all'autismo. In particolar modo il lavoro di ricerca si è concentrato su tre aspetti:

1. Stabilire il ruolo del cromosoma X nell'autismo.
2. Effettuare uno studio di replicazione del gene ASMT.
3. Caratterizzare geni candidati dei loci di suscettibilità all'autismo AUT5 ed AUTS1.

10.1 Nessun gene soggetto a XCI è coinvolto nello spettro autistico

Numerose sono le evidenze che suggeriscono un possibile ruolo nell'autismo di geni che mappano nel cromosoma X quali:

1. La maggior prevalenza dell'autismo nella popolazione maschile (rapporto maschi:femmine di 4 ad 1).
2. Numerosi geni di questo cromosoma hanno funzioni a livello cognitivo, di cui molti associati a ritardo mentale, e sono numerosi i casi di autismo accompagnati anche da questo tratto fenotipico.
3. Evidenza di una modesta condivisione allelica in tre studi di genome scan su campioni di famiglie indipendenti (Liu *et al.*, 2001; Auranen *et al.*, 2002; Yonan *et al.*, 2003).

L'analisi del profilo di inattivazione del cromosoma X può indicare la presenza di mutazioni in geni di questo cromosoma. Infatti, mutazioni in eterozigosi di geni X-linked comportano in femmine portatrici un'inattivazione in maniera preferenziale del cromosoma con la mutazione, fenomeno conosciuto come *skewed X chromosome inactivation* (SXCI).

Un recente studio ha evidenziato una proporzione maggiore di SXCI in ragazze affette da autismo e le rispettive madri, comparato a quella di un campione controllo (Talebizadeh *et al.*, 2005). Questo studio, benché condotto in un campione ridotto, potrebbe evidenziare il coinvolgimento di varianti associate all'autismo in uno o più geni soggetti a XCI.

Nonostante le evidenze citate pocanzi, non sono stati compiuti studi esaustivi per stabilire il ruolo di geni soggetti all'inattivazione del cromosoma X nell'autismo. Per questo motivo, il nostro consorzio IMGSAC ha intrapreso un'analisi del pattern di XCI in un vasto campione di famiglie, in collaborazione con altri due gruppi di ricerca, che hanno esteso l'analisi ad un gruppo di famiglie Francesi e Finlandesi.

Abbiamo analizzato il profilo di XCI in un campione di 275 madri di soggetti affetti ed un campione di 75 pazienti autistiche, confrontandolo con quello di due gruppi di controllo correlati per età: 144 femmine adulte e 40 ragazze.

L'analisi condotta nel nostro laboratorio del campione IMGSAC non ha evidenziato differenze significative nel profilo di XCI nel campione di probande ed in quello delle madri di individui affetti, rispetto ai rispettivi gruppi di controllo. Lo stesso risultato viene ottenuto quando si analizza l'intero campione europeo.

Tuttavia, il sottogruppo di famiglie con un'inattivazione altamente sbilanciata, costituisce un campione con una maggiore probabilità di portare mutazioni X-linked che contribuiscono all'eziologia dell'autismo. Ciò è dimostrato dal fatto che le tre famiglie del nostro campione europeo in cui erano già note mutazioni X-

linked nei geni ATRX, NLGN3 e MECP2, sono caratterizzate da elevato SXCI (90:10).

Quindi abbiamo selezionato 28 famiglie tra quelle che presentavano X-skewing ($\geq 80:20$) per condurre un'analisi di linkage non parametrica. Il risultato evidenzia una maggior condivisione allelica a livello del marcatore rs719489 in Xq27-28, dove mappano FMR1 e MECP2, due geni interessanti anche dal punto di vista funzionale. Nonostante il risultato dell'analisi di linkage coincide con alcune evidenze che hanno suscitato interesse in precedenti studi sull'autismo di questi due geni, il P-value non raggiunge la soglia di significatività in seguito a permutazioni.

Ultimamente sono state descritte duplicazioni nel gene MECP2 che vengono trasmesse da madri asintomatiche altamente X-skewed a figli maschi caratterizzati da ritardo mentale (Del Gaudio, 2006).

Abbiamo quindi selezionato 22 pazienti della nostra collezione, di cui 17 figli affetti di madri X-skewed e 5 pazienti autistiche X-skewed su cui abbiamo condotto un'analisi di mutazioni nelle regioni esoniche dei geni FMR1, MECP2 e NLGN3. Gli stessi individui sono stati analizzati per duplicazioni del gene MECP2.

Nessuna variante di sequenza con un potenziale ruolo eziologico è stata incontrata in seguito a *screening* di mutazione nei geni analizzati, nè duplicazioni sono state individuate nel gene MECP2.

I risultati di questo ampio studio, dimostrano che le famiglie con autismo non sono caratterizzate da un pattern di inattivazione del cromosoma X diverso da quello della popolazione generale. Inoltre, l'analisi di geni candidati posizionali e funzionali in individui autistici di famiglie X-skewed non ha rilevato nessuna nuova variante.

Queste evidenze suggeriscono l'assenza di un gene del cromosoma X soggetto ad inattivazione con un ruolo preponderante nella suscettibilità all'autismo. Tuttavia,

circa il 15-20% dei geni del cromosoma X sfugge l'inattivazione e non possiamo quindi escludere che proprio uno o più di questi geni sia coinvolto nell'autismo.

Proponiamo infine che l'analisi del profilo di XCI in famiglie con autismo possa costituire un criterio per la selezione di famiglie da sottoporre ad analisi di mutazioni nei geni candidati per gli ASD, localizzati sul cromosoma X.

10.2 Il gene ASMT non sembra essere coinvolto nello spettro autistico

Recentemente alcuni studi evidenziano il potenziale ruolo della melatonina, un neuroormone epifisario, in alcuni tratti del fenotipo autistico, quali disturbi del sonno e ritmi circadiani alterati (Malow BA, 2004). Questa ipotesi è supportata anche dalle ridotte concentrazioni notturne di questo ormone in pazienti autistici (Tordjam S. et al, 2005) ed un miglioramento dei sintomi legati al rapporto sonno/veglia in seguito alla somministrazione per via orale di melatonina (Hayashi E., 2000).

Lo studio condotto ultimamente dal consorzio PARIS sul gene ASMT potrebbe spiegare l'elevata prevalenza di ritmi circadiani alterati nell'autismo. Il gene ASMT, codifica per l'ultimo enzima del *pathway* che a partire dalla serotonina porta alla biosintesi di melatonina.

Il lavoro di Melke et al. (2008), consiste in un'analisi di mutazione ed uno studio di associazione del gene ASMT come gene candidato funzionale per l'autismo.

Lo screening di mutazione in 250 pazienti individua due rare varianti con un possibile ruolo eziologico: una missenso (L326F) ed una variante di splicing (IVS5+2T>C). Gli autori riportano un effetto di queste varianti a livello del prodotto genico. Inoltre, gli studi di associazione riportano due SNPs (rs4446909 e rs5989681) con frequenze alleliche significativamente diverse tra il gruppo di casi e quello dei controlli. L'aplotipo H1 (GGGC), identificato dai due SNPs associati, appartiene al blocco di LD localizzato al promotore B del gene ASMT. Inoltre, Gli

autori riportano che questo aplotipo è correlato ad una riduzione del trascritto genico in linee linfoblastoidi.

Benché le evidenze sperimentali di questo studio suggeriscono il coinvolgimento di ASMT nel fenotipo autistico sono necessari studi di replicazione in famiglie indipendenti. Abbiamo quindi condotto un'analisi di mutazione ed uno studio di associazione caso/controllo nella collezione IMGSAC di 263 famiglie multiplex e singleton. I risultati dello screening di mutazione svolto nel presente lavoro di tesi ha identificato tre rare varianti di cui due già identificate nel precedente lavoro:

1. La variante di splicing (IVS5+2T>C) è presente in due famiglie con autismo ed un solo individuo tra i 390 controlli utilizzati per l'analisi di mutazione. Nella famiglia multiplex la variante non segrega con il fenotipo autistico, infatti viene trasmessa dal padre solo ad uno dei due figli maschi affetti.

2. La missenso (L326F) è presente solo in un paziente e cinque controlli.

3. Una variante di stop (R319X) non descritta nel precedente lavoro viene identificata in una famiglia singleton e in nessuno tra gli individui controllo analizzati.

I risultati dello studio di associazione caso/controllo in seguito alla tipizzazione dei marcatori (rs4446909, rs5989681, rs56690322 e rs6644635) del blocco di LD al promotore B di ASMT non ha rilevato nessuna associazione né per l'analisi del singolo marcatore né per quello di aplotipi.

I risultati del nostro studio suggeriscono che le varianti identificate da Melke et al, non rappresentano varianti eziologiche per l'autismo, ma piuttosto rari polimorfismi, infatti nel nostro studio la missenso (L326F) è più frequente nel campione di casi e la variante di splicing (IVS5+2T>C) non segrega con il fenotipo autistico. Interessante risulta la presenza della variante di stop (R319X) in un individuo autistico di una famiglia singleton italiana, che tuttavia non presenta alcuna caratteristica fenotipica riconducibile a disturbi del sonno e ritmi circadiani alterati. Studi funzionali potrebbero suggerire se questa variante è realmente

associata all'abolizione della funzione enzimatica. Nonostante i dati presentati in questo studio non suggeriscono il coinvolgimento di questo gene nella suscettibilità all'autismo, rimane un ottimo gene candidato funzionale per altri disturbi neuropsichiatrici quali ad esempio ADHD, depressione, schizofrenia in cui sono frequenti disturbi del sonno e ritmi circadiani alterati.

10.3 Studi di associazione ed analisi di varianti strutturali suggeriscono che la regione IMMP2L-DOCK4 del locus AUTS1 è coinvolta nella suscettibilità all'autismo.

Negli ultimi dieci anni sono stati condotti numerosi genome scan per l'autismo, tra cui quello del nostro consorzio IMGSAC, che ha ottenuto i migliori risultati di linkage per le regioni 2q23.3-q32.3 (locus AUTS5) e 7q21.3-q33 (locus AUTS1), che rappresentano anche le regioni più replicate da parte degli altri consorzi per l'autismo. In queste due regioni mappano più di 400 geni ed il tentativo di individuare geni di suscettibilità in regioni così ampie non ha portato a risultati consistenti. Lo sviluppo in questi ultimi anni di tecniche di tipizzazione di marcatori ad alta densità ci ha permesso di effettuare il *genotyping* mediante il sistema ILLUMINA di circa 3000 SNPs per ciascuna regione, rispettivamente in 126 famiglie multiplex per il locus AUTS5 e 127 famiglie multiplex per il locus AUTS1. I dati sono stati utilizzati per effettuare studi di associazione (caso/controllo e TDT) ed analisi di varianti strutturali.

I migliori risultati per gli studi di associazione combinando i due test per il locus AUTS5 vengono ottenuti per i geni *NOSTRIN* e *ZNF533*, mentre per il locus AUTS1 vengono ottenuti per i geni *IMMP2L*, *PLXA4* e *DOCK4*. La replicazione dei marcatori più significativi per entrambi i loci è stata condotta in un campione di famiglie singleton indipendenti, in cui è stato effettuato uno studio di associazione caso/controllo e TDT.

Il risultato più significativo della meta-analisi per il locus al cromosoma 2q è rappresentato dal marcatore rs7590028 del gene ZNF533, mentre quello al cromosoma 7q è rappresentato dai marcatori rs12537269 e rs2217262 dei geni adiacenti IMMP2L e DOCK4. Il marcatore rs2217262 del gene DOCK4 rimane l'unico marcatore significativo in seguito a correzione di Bonferroni.

Il gene ZNF533 codifica per una proteina, altamente espressa nel cervello, caratterizzata da 4 zinc-fingers. Attualmente non sono noti studi funzionali, ma l'analisi bioinformatica suggerisce una localizzazione nucleare dove si pensa possa agire come fattore di repressione. Ultimamente alcuni studi hanno evidenziato il coinvolgimento di ZNF533 in attività cognitive e di sviluppo, infatti delezioni di questo gene sono state descritte in pazienti caratterizzati da ritardo mentale e problemi del linguaggio (Mencarelli MA, 2007; Monfort S, 2008).

Benché il gene *ZNF533* rappresenti un ottimo gene candidato per le evidenze citate pocanzi, i risultati dell'analisi di mutazione condotta nelle regioni regolatrici ed in tutti gli esoni non ha portato ad identificare nessuna variante chiaramente associata all'autismo. Tuttavia, non possiamo escludere la presenza di varianti eziologiche in regioni regolatorie distanti, o in regioni introniche che regolino ad esempio l'elevata variabilità dei trascritti, osservata dallo studio condotto sulle forme alternative del gene, o ulteriori varianti in nuovi esoni non presenti nel database e non rilevate nello studio dei trascritti che abbiamo effettuato.

L'analisi delle varianti strutturali per i due loci ha riportato come risultato più significativo una duplicazione che coinvolge i geni contigui IMMP2L-DOCK4. Il gene IMMP2L codifica per una proteina mitocondriale della membrana interna, e rappresenta un possibile candidato per il disturbo autistico in quanto anomalie cromosomiche di IMMP2L sono state associate alla sindrome di Tourette che mostra alcuni aspetti fenotipici in comune con l'autismo. Il gene DOCK4 codifica per una GEF (Guanine nucleotide Exchange Factor) in grado di attivare la RacGTPase. Benché il gene Dock4 sia stato associato a carcinogenesi (Yajnik V,

2003), ultimamente uno studio ha evidenziato la sua elevata espressione in ippocampo dove svolge un ruolo essenziale per la corretta formazione e morfologia dendritica (Ueda S, 2008).

L'evidenza di varianti strutturali nella regione IMMP2L-DOCK4 risulta essere di rilevante interesse alla luce dei risultati ottenuti dagli studi di associazione. Quindi abbiamo ritenuto necessario approfondire la presenza di varianti strutturali in questa regione ampliando il campione ad un numero finale di 285 famiglie multiplex. Complessivamente abbiamo identificato quattro famiglie con delezioni ed una famiglia con una duplicazione in questa regione. In particolar modo tre famiglie presentano una delezione che coinvolge esclusivamente il gene IMMP2L che non segrega con il fenotipo autistico. Inoltre questa variante è presente nei controlli analizzati ed è ben rappresentata nel database di varianti genomiche (figura 55).

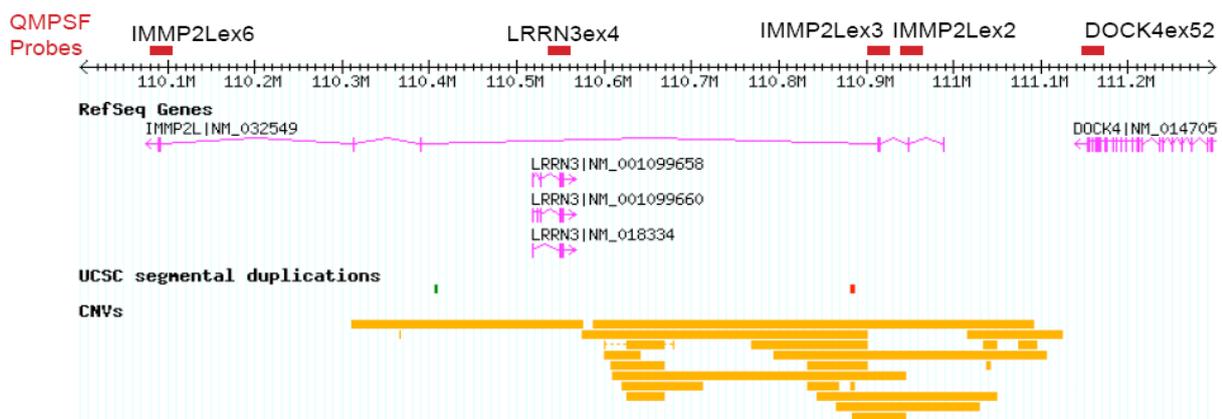


Figura 55: schematizzazione delle varianti strutturali in giallo presenti nel database di CNVs nella regione IMMP2L-Dock4

La duplicazione della regione IMMP2L-Dock4 sembra non essere associata al fenotipo autistico, in quanto viene trasmessa solo ad uno dei due figli affetti. Solo una famiglia presenta una delezione che si estende dal gene IMMP2L a DOCK4. La delezione segrega con il fenotipo autistico. Inoltre come si può osservare dal database di varianti genomiche non sono presenti CNVs nella regione di DOCK4, nè sono presenti nel campione di 475 individui di controllo analizzati

nel presente studio. La delezione è trasmessa dalla madre ai due probandi maschi affetti ed alla figlia non affetta. E' interessante notare che la madre e la figlia con la delezione in *IMMP2L-DOCK4*, benché non presentino un fenotipo autistico, sono affette da dislessia. Quindi è possibile che vi sia una penetranza differente di questa variante genetica in relazione al sesso. La delezione in *IMMP2L-DOCK4* potrebbe avere un effetto patologico comportando l'aploinsufficienza di *DOCK4* ed *IMMP2L*, oppure potrebbe avere un effetto dominante negativo, attraverso la formazione di un trascritto ibrido. Studi funzionali saranno utili per stabilire questo aspetto. Inoltre saranno necessari ulteriori studi su campioni di famiglie indipendenti per ottenere conferme che SNPs comuni nel gene *DOCK4*, *IMMP2L* e *ZNF533* siano coinvolti nella suscettibilità all'autismo

11. Bibliografia

- Abecasis, et al. Extent and distribution of linkage disequilibrium in three genomic regions. *Am. J. Hum. Genet.*, 2001; 68: 191–197.
- Alarcón M, Abrahams BS, Stone JL, Duvall JA, Perederiy JV, Bomar JM, Sebat J, Wigler M, Martin CL, Ledbetter DH, Nelson SF, Cantor RM, Geschwind DH. (2008) Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. (2008) *Am J Hum Genet.* 82(1):150-9.
- Allen et al. (1992). Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *Am J Hum Genet.* 51:1229-39.
- Altug-Teber O. et al. A rapid microarray based whole genome analysis for detection of uniparental disomy. *Hum. Mutat.* 2005; 26: 153–159
- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders*. Bailey, A., Le Counteur, A., Gottesman, I., Bolton P., Simonoff E., Yuzda E. and Rutter M., (1995). *Autism as a strongly genetic disorder: evidence from a British twin study*. *Psychol Med*, 25: 63-77.
- Amir et al. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genet.* 23:185-188.
- Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet.* 1999 Oct;23(2):185-8.
- Amos-Landgraf et al. (2006). X chromosome-inactivation patterns of 1,005 phenotypically unaffected females. *Am J Hum Genet.* 79:493-9.
- Anderlid BM, Schoumans J, Anneren G, Tapia-Paez I, Dumanski J, Blennow E, Nordenskjold M. FISHmapping of a 100-kb terminal 22q13 deletion. *Hum. Genet.* 2002; 110: 439-443.
- Anderson, L.T. et al. *The effects of haloperidol on discrimination learning and behavioral symptoms in autistic children*. *J. Autism Dev. Disord.* 19, 227-239 (1989).
- Arking DE, Cutler DJ, Brune CW, Teslovich TM, West K, Ikeda M, Rea A, Guy M, Lin S, Cook EH, Chakravarti A. (2008) A common genetic variant in the neurexin superfamily member CNTNAP2 increases familial risk of autism. *Am J Hum Genet.* 82(1):160-4.
- Ashley-Koch A, Wolpert CM, Menold MM, Zaeem L, Basu S, Donnelly SL, Ravan SA, Powell CM, Qumsiyeh MB, Aylsworth AS, Vance JM, Gilbert JR, Wright HH, Abramson RK, DeLong GR, Cuccaro ML, Pericak-Vance MA. Genetic studies of autistic disorder and chromosome 7. *Genomics.* 1999;61: 227-36.
- Ashley-Koch AE, Jaworski J, Ma de Q, Mei H, Ritchie MD, Skaar DA, Robert Delong G, Worley G, Abramson RK, Wright HH, Cuccaro ML, Gilbert JR, Martin ER, Pericak-Vance MA. (2007) Investigation of potential gene-gene interactions between APOE and RELN contributing to autism risk. *Psychiatr Genet.* 17:221-6
- Ashley-Koch AE, Jaworski J, Ma de Q, Mei H, Ritchie MD, Skaar DA, Robert Delong G, Worley G, Abramson RK, Wright HH, Cuccaro ML, Gilbert JR, Martin ER, Pericak-Vance MA (2007) Investigation of potential gene-gene interactions between APOE and RELN contributing to autism risk. *Psychiatr Genet.* 17:221-6
- Asperger, H. Die autistischen psychopathen im kindesalter. *Archiv fur Psychiatrie und Nervenkrankheiten* 117, 76–136 (1944).
- Auranen M, Varilo T, Alen R, Vanhala R, Ayers K, Kempas E, Ylisaukko-Oja T, Peltonen L, Jarvela I. Evidence for allelic association on chromosome 3q25-27 in families with autism spectrum disorders originating from a subisolate of Finland. *Mol Psychiatry.* 2003;8: 879-84.
- Autism Genome Project Consortium, Szatmari P, Paterson AD, Zwaigenbaum L, Roberts W, Brian J, Liu XQ, Vincent JB, Skaug JL, Thompson AP, Senman L, Feuk L, Qian C, Bryson SE, Jones MB, Marshall CR, Scherer SW, Vieland VJ, Bartlett C, Mangin LV, Goedken R, Segre A, Pericak-Vance MA, Cuccaro ML, Gilbert JR, Wright HH, Abramson RK, Betancur C, Bourgeron T, Gillberg C, Leboyer M, Buxbaum JD, Davis KL, Hollander E, Silverman JM, Hallmayer J, Lotspeich L, Sutcliffe JS, Haines JL, Folstein SE, Piven J, Wassink TH, Sheffield V, Geschwind DH, Bucan M, Brown WT, Cantor RM, Constantino JN, Gilliam TC, Herbert M, Lajonchere C, Ledbetter DH, Lese-Martin C, Miller J, Nelson S, Samango-Sprouse CA, Spence S, State M, Tanzi

- RE, Coon H, Dawson G, Devlin B, Estes A, Flodman P, Klei L, McMahon WM, Minshew N, Munson J, Korvatska E, Rodier PM, Schellenberg GD, Smith M, Spence MA, Stodgell C, Tepper PG, Wijsman EM, Yu CE, Rogé B, Mantoulan C, Wittmeyer K, Poustka A, Felder B, Klauck SM, Schuster C, Poustka F, Bölte S, Feineis-Matthews S, Herbrecht E, Schmötzer G, Tsiantis J, Papanikolaou K, Maestrini E, Bacchelli E, Blasi F, Carone S, Toma C, Van Engeland H, de Jonge M, Kemner C, Koop F, Langemeijer M, Hijmans C, Staal WG, Baird G, Bolton PF, Rutter ML, Weisblatt E, Green J, Aldred C, Wilkinson JA, Pickles A, Le Couteur A, Berney T, McConachie H, Bailey AJ, Francis K, Honeyman G, Hutchinson A, Parr JR, Wallace S, Monaco AP, Barnby G, Kobayashi K, Lamb JA, Sousa I, Sykes N, Cook EH, Guter SJ, Leventhal BL, Salt J, Lord C, Corsello C, Hus V, Weeks DE, Volkmar F, Tauber M, Fombonne E, Shih A, Meyer KJ. Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet.* 2007 Mar;39(3):319-28.
- Bacchelli E, Maestrini E. Autism spectrum disorders: molecular genetic advances. *Am J Med Genet C Semin Med Genet.* 2006 Feb 15;142C(1):13-23
 - Bailey, A., Palferman., Heavey L., Le Counter A. (1998). *Autism: the phenotype in relatives. J Aut Dev Dis, 28: 439-445.*
 - Bailey, A., Rutter M., Philips W. (1996). *Autism: the phenotype in relatives. J Aut Dev Dis, 28:439-445.*
 - Baird, G., Charman, T., Cox, A., Baron-Cohen, S., Swettenham, J., Wheelwright, S., (2001). *Current topic: screening and surveillance for autism and pervasive developmental disorders. Archives of disease in Childhood, 84, 468-475.*
 - Bakkaloglu B, O'Roak BJ, Louvi A, Gupta AR, Abelson JF, Morgan TM, Chawarska K, Klin A, Ercan-Sencicek AG, Stillman AA, Tanriover G, Abrahams BS, Duvall JA, Robbins EM, Geschwind DH, Biederer T, Gunel M, Lifton RP, State MW. (2008) Molecular cytogenetic analysis and resequencing of contactin associated protein-like 2 in autism spectrum disorders.*Am J Hum Genet.* 82(1):165-73.
 - Baron-Cohen,S.,Wheelwright,S., Hill,J., Raste,Y. & Plumb,I. *The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism.* *Journal of Child Psychology and Psychiatry* 42(2), 241-251. (2001).
 - Barrett JC, Fry B, Maller J,Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263-5.
 - Barrett S, Beck JC, Bernier R, Bisson E, Braun TA, Casavant TL et al. An autosomal genomic screen for autism. *Am J Med Genet* 1999; 88: 609–615.
 - Battye R, Stevens A, Perry RL,Jacobs JR. Repellent signaling by Slit requires the leucine-rich repeats. *J Neurosci* 2001; **21**: 4290-8.
 - Blasi et al. (2006). Absence of coding mutations in the X-linked genes neuroligin 3 and neuroligin 4 in individuals with autism from the IMGSAC collection. *Am J Med Genet B Neuropsychiatr Genet.* 141B:220-1.
 - Blasi F, Bacchelli E, Pesaresi G, Carone S, Bailey AJ, Maestrini E; International Molecular Genetic Study of Autism Consortium (IMGSAC) Absence of coding mutations in the X-linked genes neuroligin 3 and neuroligin 4 in individuals with autism from the IMGSAC collection. (2006) *Am J Med Genet B Neuropsychiatr Genet.*
 - Bolton PF, Park RJ, Higgins JN, Griffiths PD, Pickles A. *Neuro-epileptic determinants of autism spectrum disorders in tuberous sclerosis complex.* *Brain.* 2002 Jun;125(Pt 6):1247-55.
 - Bonaglia MC, Giorda R, Mani E, Aceti G, Anderlid BM, Baroncini A, Pramparo T, Zuffardi O. Identification of a recurrent breakpoint within the SHANK3 gene in the 22q13.3 deletion syndrome. *J.Med. Genet.* 2006; 43: 822-828.
 - Boyd Y, Fraser NJ. (1990). Methylation patterns at the hypervariable X-chromosome locus DXS255 (M27 beta): correlation with X-inactivation status. *Genomics.* 7:182-7.
 - Bradford Y, Haines J, Hutcheson H, Gardiner M, Braun T, Sheffield V, Cassavant T, Huang W, Wang K, Vieland V, Folstein S, Santangelo S, Piven J. Incorporating language phenotypes strengthens evidence of linkage to autism. *Am J Med Genet.* 2001 Aug 8;105(6):539-47.
 - Brown et al. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature.* 349:38-44.
 - Bruce SK et al. Global analysis of uniparental disomy using high-density genotyping arrays. *J. Med.Genet.*2005; 42: 847–851

- Buckley P. G., Mantripragada K. K., Piotrowski A., Diaz de Stahl T., Dumanski J.P., Copy-number polymorphisms: mining the tip of an iceberg. *Trends in Genetics* , 2005; vol.21, N°6 .
- Buoni et al., The syndrome of inv dup (15): clinical, electroencephalographic, and imaging findings. *J.Child. Neurol.* 2000; 15: 380-385.
- Burwinkel B., Kilimann M.W., Unequal homologous recombination between LINE-1 elements as a mutational mechanism in human genetic disease. *J. Mol. Biol.* 1998; 277: 513–517.
- Busque et al. (1996). Nonrandom X-inactivation patterns in normal females: lyonization ratios vary with age. *Blood.* 88:59-65.
- Buxbaum JD, Silverman J, Keddache M, Smith CJ, Hollander E, Ramoz N et al. Linkage analysis for autism in a subset families with obsessive-compulsive behaviors: evidence for an autism susceptibility gene on chromosome 1 and further support for susceptibility genes on chromosome 6 and 19. *Mol Psychiatry* 2004; 9: 144–150. 150–159.
- Buxbaum JD, Silverman JM, Smith CJ, Kilifarski M, Reichert J, Hollander E et al. Evidence for a susceptibility gene for autism on chromosome 2 and for genetic heterogeneity. *Am J Hum Genet* 2001;68: 1514–1520.
- Campbell DB, Sutcliffe JS, Ebert PJ, Militerni R, Bravaccio C, Trillo S, Elia M, Schneider C, Melmed R, Sacco R, Persico AM, Levitt P. A genetic variant that disrupts MET transcription is associated with autism. *Proc Natl Acad Sci U S A.* 2006; 103: 16834-9.
- Cantor RM, Kono N, Duvall JA, Alvarez-Retuerto A, Stone JL, Alarcón M, Nelson SF, Geschwind DH. Replication of autism linkage: fine-mapping peak at 17q21. *Am J Hum Genet.* 2005 Jun;76(6):1050-6. Epub 2005 Apr 1.
- Carlson, et al. Additional SNPs and linkage-disequilibrium analyses are necessary for wholegenome association studies in humans. *Nature Genet.*, 2003; 33: 518–520.
- Carney et al. (2003). Identification of MeCP2 mutations in a series of females with autistic disorder. *Pediatr Neurol.* 28:205-11.
- Carney et al. (2003). Identification of MeCP2 mutations in a series of females with autistic disorder. *Pediatr Neurol.* 28:205-11.
- Carrel et al. (1999). A first-generation X-inactivation profile of the human X chromosome. *Proc Natl Acad Sci U S A.* 96:14440-4.
- Carrel L, Willard HF. (2005). X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature.* 434:400-4.
- Casas et al., Four new patients with 2q terminal deletion syndrome: hemidiaphragmatic hernia as a novel findings and patterns of malformation among 47 reported cases. *Am. J. Hum. Genet.*, 2002; 71(4):A569.
- Cazzola et al. (2000). Familial-skewed X-chromosome inactivation as a predisposing factor for late-onset X-linked sideroblastic anemia in carrier females. *Blood.* 96:4363-5.
- Chakrabarti S., Fombonne E. Pervasive developmental disorders in preschool children, *Jama* 2001; 285:3093-9.
- Chih et al. (2005). Control of excitatory and inhibitory synapse formation by neuroligins. *Science.* 307:1324-8.
- Colella S, Yau C, Taylor JM, Mirza G, Butler H, Clouston P *et al.* QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data. *Nucleic Acids Res* 2007; **35**: 2013-25.
- Conrad D. F., Andrews D., Carter N., Hurler M., Pritchard J., A high-resolution survey of deletion polymorphism in the human genome. *Nature Genetics*, 2006; 38: 75-81.
- Conroy J, Meally E, Kearney G, Fitzgerald M, Gill M, Gallagher L. (2004) *Serotonin transporter gene and autism: a haplotype analysis in an Irish autistic population.* *Mol Psychiatry* 9: 587-93.
- Contemplating effects of genomic structural variation Buchanan JA, Scherer SW. (2008) *Genet Med* 10:639-47
- Cook EH Jr, Courchesne R, Lord C, Cox NJ, Yan S, Lincoln A, Haas R, Courchesne E, Leventhal BL. Evidence of linkage between the serotonin transporter and autistic disorder. *Mol Psychiatry.* 1997 May;2(3):247-50.
- Cook, E. H. & Leventhal, B. L. *The serotonin system in autism. A review of the evidence for involvement of the serotonin system in the aetiology of autism.* *Curr. Opin. Pediatr.* 8, 348–354 (1996).

- Coon H, Matsunami N, Stevens J, Miller J, Pingree C, Camp NJ, Thomas A, Krasny L, Lainhart J, Leppert MF, McMahon W. Evidence for linkage on chromosome 3q25-27 in a large autism extended pedigree. *Hum Hered.* 2005;60(4):220-6. Epub 2006 Jan 2.
- Cooper GM, Nickerson DA, Eichler EE. (2007) Mutational and selective effects on copy-number variants in the human genome. *Nat Genet* 39:S22–S29.
- Coutinho AM, Oliveira G, Katz C, Feng J, Yan J, Yang C, Marques C, Ataíde A, Miguel TS, Borges L, Almeida J, Correia C, Currais A, Bento C, Mota-Vieira L, Temudo T, Santos M, Maciel P, Sommer SS, Vicente AM. (2007) MECP2 coding sequence and 3'UTR variation in 172 unrelated autistic patients. *Am J Med Genet B Neuropsychiatr Genet.* 5;144B(4):475-83.
- Croen, I., Grether, J. K., Hoogstrate, J., Selvin, S. (2002). *The changing prevalence of autism in California.* *Journal of Autism and developmental disorders* 32, 207-215.
- Croonenberghs, J. et al. *Peripheral markers of serotonergic and noradrenergic function in post-pubertal, caucasian males with autistic disorder.* *Neuropsychopharmacology* 22, 275-283 (2000).
- D'Arcangelo G, Homayouni R, Keshvara L, Rice DS, Sheldon M, Curran T (1999). *Reelin is a ligand for lipoprotein receptors.* *Neuron* 24, 471-479.
- Dawson, et al. A first-generation linkage disequilibrium map of human chromosome 22. *Nature*, 2002; 418: 544–548.
- de Ståhl TD, Sandgren J, Piotrowski A, Nord H, Andersson R, Menzel U, Bogdan A, Thuresson AC, Poplawski A, von Tell D, Hansson CM, Elshafie AI, Elghazali G, Imreh S, Nordenskjöld M, Upadhyaya M, Komorowski J, Bruder CE, Dumanski JP (2008) Profiling of copy number variations (CNVs) in healthy individuals from three ethnic groups using a human genome 32 K BAC-clone-based array. *Hum Mutat.* 29:398-408.
- Del Gaudio et al. (2006). Increased MECP2 gene copy number as the result of genomic duplication in neurodevelopmentally delayed males. *Genet Med.* 8:784-92.
- Devlin B, Cook EH, Coon H, Dawson G, Grigorenko EL, McMahon W, Minshew N, Pauls D, Smith M, Spence MA (2005) *Autism and the serotonin transporter: the long and short of it.* *Mol Psychiatry.* 10: 1110-6.
- Devlin Bernie, Pamela Bennett, Geraldine Dawson, Denise A. Figlewicz, Elena Grigorenko, William McMahon, Nancy Minshew, David Pauls, Moyra Smith, M. Anne Spence, Patricia M. Rodier, Chris Stodgell, The CPEA Genetics Network, and Gerard D. Schellenberg. *Allele of a Reelin CGG repeat do not convey liability to autism in a sample from the CPEA network.* *American journal of medical genetics part B (neuropsychiatric Genetics)* 126B:46-50 (2004).
- Durand CM, Betancur C, Boeckers TM, Bockmann J, Chaste P, Fauchereau F, Nygren G, Rastam M, Gillberg IC, Anckarsater H, Sponheim E, Goubran-Botros H, and 11 others.
- Dutta S, Sinha S, Ghosh S, Chatterjee A, Ahmed S, Usha R. (2008) Genetic analysis of reelin gene (RELN) SNPs: no association with autism spectrum disorder in the Indian population *Neurosci Lett.* 441:56-6
- Dykens E. and Volkmar F. R., Medical conditions associated with autism, in *Handbook of Autism and Pervasive Developmental Disorders*, 1997; (Volkmar C., ed.), Wiley, New York, pp. 388–410.
- E Bonora, KS Beyer, JA Lamb, JR Parr, SM Klauck, A Benner, M Paolucci, A Abbott, I Ragoussis, A Poustka, AJ Bailey AP Monaco and the International Molecular Genetic Study of Autism Consortium (IMGSAC). *Analysis of reelin as a candidate gene for autism.* *Molecular Psychiatry* (2003) 8, 885–892
- El-Sherif Y, Tesoriero J, Hogan MV, Wieraszko A. Melatonin regulates neuronal plasticity in the hippocampus. *J Neurosci Res* 2003;72:454–460.
- Elston R. , The genetic dissection of multifactorial traits. *Clin Exp Allergy* , 1995; 25:103– 6.
- Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 2003; **164**: 1567-87.
- Feinstein C. and Reiss A. L., Autism: the point of view from fragile X studies. *J. Autism Dev. Disord.* 1998; 28: 393–405.
- Feuk L., Carson A. R., and Scherer S. W., Structural variation in the human genome. *Nature Genetics*, February 2006; 7: 85-97.
- Fialkow PJ. (1973). Primordial cell pool size and lineage relationships of five human cell types. *Ann Hum Genet.* 37:39-48.

- Flint J., Wilkie M.A.O., Bucke V.J., Winter R.M., Holland A.J., McDermid H.E., The detection of subtelomeric chromosomal rearrangements in idiopathic mental retardation. *Nat. Genet.*, 1995; 9: 132–140.
- Folstein SE, Mankoski RE. Chromosome 7q: where autism meets language disorder? *Am J Hum Genet.* 2000 Aug;67(2):278-81
- Folstein SE, Rosen-Sheidley B. Genetics of autism: complex aetiology for a heterogeneous disorder. *Nat Rev Genet* 2001; 2:943–55.
- Fombonne, E. *Epidemiological surveys of autism and other pervasive developmental disorders: An update.* *Journal of Autism and developmental disorders*, 33, 365-382. (2003).
- Fukumachi K, Matsuoka Y, Ohno H, Hamaguchi T, Tsuda H. Neuronal leucine-rich repeat protein-3 amplifies MAPK activation by epidermal growth factor through a carboxyl-terminal region containing endocytosis motifs. *J Biol Chem* 2002; **277**: 43549-52.
- Gabriel, S. B. et al. The structure of haplotype blocks in the human genome. *Science* 296, 2225–2229.
- Gauthier J, Bonnel A, St-Onge J, Karemera L, Laurent S, Mottron L, Fombonne E, Joober R, Rouleau GA. NLGN3/NLGN4 gene mutations are not responsible for autism in the Quebec population. *Am J Med Genet B Neuropsychiatr Genet.* 2005; 132: 74-5.
- Gauthier J, Joober R, Mottron L, Laurent S, Fuchs M, De Kimpe V, Rouleau GA. Mutation screening of FOXP2 in individuals diagnosed with autistic disorder. *Am J Med Genet A.* 2003 Apr 15;118A(2):172-5.
- Ghaziuddin et al., Deletion of chromosome 2q37 and autism: a distinct subtype. *J. Autism Dev. Disord.*, 1999; 29: 259-263.
- Goldstein, Ahmadi, Weale, Wood, Genome scans and candidate gene approaches in the study of common diseases and variable drug responses. *Trends Genet.*, 2003; 19: 615–622.
- Gong et al. (2008). Analysis of X chromosome inactivation in autism spectrum disorders. *Am J Med Genet B Neuropsychiatr Genet.* 147B:830-5.
- Gong X, Jia M, Ruan Y, Shuang M, Liu J, Wu S, Guo Y, Yang J, Ling Y, Yang X, Zhang D. Association between the FOXP2 gene and autistic disorder in Chinese population *Am J Med Genet B Neuropsychiatr Genet.* 2004 May 15;127B(1):113-6.
- Gordon, C., Rapoport, J., Hamburger, S., State, R. & Mannheim, G. *Differential response of seven subjects with autistic disorder to clomipramine and desipramine.* *American Journal of Psychiatry* 149(3), 363-366. (1992).
- Gordon, C., State, R., Nelson, J., Hamburger, S. & Rapoport, J. *A double-blind comparison of clomipramine, desipramine, and placebo in the treatment of autistic disorder.* *Archives of General Psychiatry* 50(6), 441-447. (1993).
- Graf ER, Zhang X, Jin SX, Linhoff MW & Craig AM. Neurexins induce differentiation of GABA and glutamate postsynaptic specializations via neuroligins. *Cell.* 2004; 119: 1013–1026.
- Graf et al. (2004). Neurexins induce differentiation of GABA and glutamate postsynaptic specializations via neuroligins. *Cell.* 119:1013-26.
- Gu W, Zhang F, Lupski JR. Mechanisms for human genomic rearrangements (2008). *Pathogenetics.* 3;1:4.
- Guhathakurta S, Sinha S, Ghosh S, Chatterjee A, Ahmed S, Gangopadhyay PK, Usha R. (2008) Population-based association study and contrasting linkage disequilibrium pattern reveal genetic association of SLC6A4 with autism in the Indian population from West Bengal. *Brain Res.* 13;1240:12-21.
- Gutierrez GC, Smalley SL, Tanguay PE. Autism in tuberous sclerosis complex. *J Autism Dev Disord.* 1998; 28: 97-103.
- Hallam KT, Olver JS, Chambers V, Begg DP, McGrath C, Norman TR. The heritability of melatonin secretion and sensitivity to bright nocturnal light in twins. *Psychoneuroendocrinology* 2006;31:867–875.
- Harrison JE, Bolton PF. *Annotation: tuberous sclerosis.* *J Child Psychol Psychiatry.* 1997 Sep;38(6):603-14
- Hatakeyama et al. (2004). The dynamics of X-inactivation skewing as women age. *Clin Genet.* 66:327-32.

- Hinds D.A., Kloek A. P., Jen M., Chen X., Frazer K.A., Common deletions and SNPs are in linkage disequilibrium in the human genome. *Nature Genetics*, February 200; vol. 38, num. 1: 82-85.
- Hollander E, Kaplan A, Schmeidler J, Yang H, Li D, Koran LM, Barbato LM. Neurological soft signs as predictors of treatment response to selective serotonin reuptake inhibitors in obsessive-compulsive disorder. *J Neuropsychiatry Clin Neurosci*. 2005 Fall;17(4):472-7.
- Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, Lee C. Detection of large-scale variation in the human genome. *Nat Genet* 2004; 36:949–951.
- IMGSAC (International Molecular Genetic Study of Autism Consortium), A full genome screen for autism with evidence for linkage to a region on chromosome 7q. *Hum. Mol. Genet.*, 1998; 7: 571–578.
- IMGSAC (International Molecular Genetic Study of Autism Consortium), A genomewide screen for autism: strong evidence for linkage to chromosomes 2q, 7q, and 16p. *Am J Hum Genet*. 2001; 69: 570-81.
- Inoue K, Lupski JR. Molecular mechanisms for genomic disorders. *Annu Rev Genomics Hum Genet*. 2002; 3:199-242. *J Autism Dev Disord*. 1989; 19: 349-51.
- ISHIKAWA-BRUSH Y, POWELL JF, BOLTON P, MILLER AP, FRANCIS F, WILLARD HF, LEHRACH H, MONACO AP. 1997. AUTISM and multiple exostoses associated with an X;8 translocation occurring within the GRPR gene and 3' to the SDC2 gene. *Human Molecular Genetics* 6: 1241-1250.
- Itsara A, Cooper GM, Baker C, Girirajan S, Li J, Absher D, Krauss RM, Myers RM, Ridker PM, Chasman DI, Mefford H, Ying P, Nickerson DA, Eichler EE (2009) Population analysis of large copy number variants and hotspots of human genetic disease *Am J Hum Genet*. 84(2):148-61
- Jacobs PA, Browne C, Gregson N, Joyce C, White H, Estimates of the frequency of chromosome abnormalities detectable in unselected newborns using moderate levels of banding. *J. Med. Genet*, 1992; 29: 103–108.
- Jamain et al. (2008). Reduced social interaction and ultrasonic communication in a mouse model of monogenic heritable autism. *Proc Natl Acad Sci U S A*. 105:1710-5.
- Jamain S, Quach H, Betancur C, Rastam M, Colineaux C, Gillberg IC, Soderstrom H, Giros B, Leboyer M, Gillberg C, Bourgeron T; Paris Autism Research International Sibpair Study. Mutations of the X linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nat Genet*. 2003; 34: 27-9.
- Jambaque I, Chiron C, Dumas C, Mumford J, Dulac O. Mental and behavioural outcome of infantile epilepsy treated by vigabatrin in tuberous sclerosis patients. *Epilepsy Res*. 2000; 38: 151-60.
- Jan JE, O'Donnell ME. Use of melatonin in the treatment of paediatric sleep disorders. 1996 *J Pineal Res*.;21(4):193-9.
- Jaselskis, C.A., Cook, E.H., Jr., Fletcher, K.E. & Leventhal, B.L. *Clonidine treatment of hyperactive and impulsive children with autistic disorder*. *J. Clin. Psychopharmacol*. 12, 322-327 (1992).
- Kanner, L. *Autistic disturbances of affective contact. The original description of autism, with excellent case examples*. *Nervous Child* 2, 217–250 (1943).
- Kemper T.L. and Bauman P., (1998). *Neuropathology of infantile autism*. *J. Neuropathol. Exp. Neurol*. 57: 645-662.
- Kennedy G. C., et al., Large-scale genotyping of complex DNA. *Nature Biotechnol*. 2003; 21: 1233–1237.
- Kim HG, Kishikawa S, Higgins AW, Seong IS, Donovan DJ, Shen Y, Lally E, Weiss LA, Najm J, Kutsche K, Descartes M, Holt L, Braddock S, Troxell R, Kaplan L, Volkmar F, Klin A, Tsatsanis K, Harris DJ, Noens I, Pauls DL, Daly MJ, MacDonald ME, Morton CC, Quade BJ, Gusella JF (2008) Disruption of neurexin 1 associated with autism spectrum disorder *Am J Hum Genet*. 82:199-207.
- Kim SJ, Cox N, Courchesne R, Lord C, Corsello C, Akshoomoff N, Guter S, Leventhal BL, Courchesne E, Cook EH Jr. Transmission disequilibrium mapping at the serotonin transporter gene (SLC6A4) region in autistic disorder. *Mol Psychiatry*. 2002;7(3):278-88.
- Klauck SM, Genetics of autism spectrum disorder, *Europ. Journ. of Hum. Genet*. 2006; 14: 714-720.

- Knight AJ, Regan R, Nicod A, Horsley SW, Kearny L, Homfray T, Winter RM, Bolton P, Flint J, Subtle chromosomal rearrangements in children with unexplained mental retardation. *Lancet* 1999; 354: 1676–1681.
- Kolomietz E, Meyn MS, Pandita A, Squire JA, The role of Alu repeat clusters as mediators of recurrent chromosomal aberrations in tumors. *Genes, Chromosomes Cancer* 2002; 35: 97–112.
- Korbel JO, Urban AE, Affourtit JP, Godwin B, Grubert F, Simons JF, Kim PM, Palejev, D, Carriero NJ, Du L, Taillon BE, Chen Z, Tanzer A, Saunders AC, Chi J, Yang F, Carter NP, Hurles ME, Weissman SM, Harkins TT, Gerstein MB, Egholm M, Snyder M. 2007. Paired-end mapping reveals extensive structural variation in the human genome. *Science* 318:420–426.
- Krebs MO., Betancur C., Leroy S., Bourdel MC., Gillberg C., Leboyer M. and Paris Autism Research International Sibpair (PARIS) Study. *Absence of association between a polymorphic GGC repeat in the 5' untranslated region of reelin gene and autism*. *Molecular Psychiatry* 7, 801-804 (2002).
- Kristiansen et al. (2005). Twin study of genetic and aging effects on X chromosome inactivation. *Eur J Hum Genet.* 13:599-606.
- Kruglyak L, Nickerson DA , Variation is the spice of life. *Nature Genet.*, 2001; 27: 234–236.
- Kruglyak, Nickerson, Variation is the spice of life. *Nature Genet.*, 2001; 27: 234–236.
- Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*. 2001 Oct 4;413(6855):519-23.
- Lam et al. (2000). Spectrum of mutations in the MECP2 gene in patients with infantile autism and Rett syndrome. *J Med Genet.* 37:E41.
- Lamb JA, Barnby G, Bonora E, Sykes N, Bacchelli E, Blasi F et al. Analysis of IMGSAC autism susceptibility loci: evidence for sex limited and parent of origin specific effects. *J Med Genet* 2005; 42: 132–137.
- Lamb JA, Parr JR, Bailey AJ, Monaco AP. Autism: in search of susceptibility genes. *Neuromolecular Med* 2002; 2:11–28.5.
- Lander ES, Schork NJ, Genetic dissection of complex traits. *Science*. 1994; 265: 2037-48.
- Lander ES, Schork NJ, Genetic dissection of complex traits. *Science*. 1994; 265: 2037-48.
- Laumonier et al. (2004). X-linked mental retardation and autism are associated with a mutation in the NLGN4 gene, a member of the neuroligin family. *Am J Hum Genet.* 74:552-7.
- Laumonier F, Bonnet-Brilhault F, Gomot M, Blanc R, David A, Moizard MP, Raynaud M, Ronce N, Lemonnier E, Calvas P, Laudier B, Chelly J, Fryns JP, Ropers HH, Hamel BC, Andres C, Barthelemy C, Moraine C, Briault S. X-linked mental retardation and autism are associated with a mutation in the NLGN4 gene, a member of the neuroligin family. *Am J Hum Genet.* 2004; 74: 552-7.
- Lauritsen MB, Als TD, Dahl HA, Flint TJ, Wang AG, Vang M, Kruse TA, Ewald H, Mors O. A genome-wide search for alleles and haplotypes associated with autism and related pervasive developmental disorders on the Faroe Islands. *Mol Psychiatry*. 2006 Jan;11(1):37-46.
- Lawson-Yuen A, Saldivar JS, Sommer S, Picker J Familial deletion within NLGN4 associated with autism and Tourette syndrome (2008) *Eur J Hum Genet.* 16:614-8
- Lawson-Yuen et al. (2008). Familial deletion within NLGN4 associated with autism and Tourette syndrome. *Eur J Hum Genet.* 16:614-8.
- Li H, Yamagata T, Mori M, Momoi MY. Absence of causative mutations and presence of autism-related allele in FOXP2 in Japanese autistic patients. *Brain Dev.* 2005 Apr;27(3):207-10.
- Li Jun, Loan Nguyen, Christopher Gleason, Linda Lotspeich, Donna Spiker, Neil Risch and Richard M. Myers. *Lack of evidence for an association between WNT2 and RELN polymorphisms and autism*. *American journal of medical genetics part B (neuropsychiatric Genetics)* 126B:51-57 (2004).
- Lichter,D.G. & Jackson,L.A. *Predictors of clonidine response in Tourette syndrome: implications and inferences*. *J. Child Neurol.* 11, 93-97 (1996).
- Lindblad-Toh K, Winchester E, Daly MJ, Wang DG, Hirschhorn JN, Laviolette JP, Ardlie K, Reich DE, Robinson E, Sklar P, Shah N, Thomas D, Fan JB, Gingeras T, Warrington J, Patil N, Hudson TJ, Lander ES. Large-scale discovery and genotyping of single-nucleotide polymorphisms in the mouse. *Nat Genet.*2000; 24: 381-6.
- Liu C, Weaver DR, Jin X, Shearman LP, Pieschl RL, Gribkoff VK, et al. Molecular dissection of two distinct actions of melatonin on the suprachiasmatic circadian clock. *Neuron* 1997;19:91–102.

- Liu JJ, Nyholt DR, Magnussen P, Parano E, Pavone P, Geschwind D et al. A genomewide screen for autism susceptibility loci. *Am J Hum Genet* 2001; 69: 327–340.
- Lock et al. (1987). Methylation of the Hprt gene on the inactive X occurs after chromosome inactivation. *Cell*. 48:39-46.
- Lombroso PJ. Genetics of childhood disorders: XLVIII. Learning and memory, Part 1: Fragile X syndrome update. *J Am Acad Child Adolesc Psychiatry*. 2003 Mar;42(3):372-5
- Lord, C., Schopler, E., Revicki, D. (1982). *Sex difference in autism. Journal of Autism and developmental disorders*. 12, 317-330.
- Lord,C. et al. *The autism diagnostic observation schedulegeneric: a standard measure of social and communication deficits associated with the spectrum of autism*. *J Autism Dev. Disord*. 30, 205-223 (2000).
- Lord,C., Rutter,M. & Le Couteur,A.*Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders*. *J. Autism Dev. Disord*. 24, 659-685 (1994).
- Lucito R., et al., Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation. *Genome Res*. 2003; 13: 2291–2305.
- Lupski JR, Genomic disorders: Structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet*. 1998; 14: 417–422.
- Lupski JR, Shaffer LG, Molecular mechanisms for constitutional chromosomal rearrangements in humans. *Ann. Rev. Genet*. 2000; 34: 297–329.
- Lupski JR, Stankiewicz P, Genome architecture, rearrangements and genomic disorders. *Trends Genet*. 2002; 18: 74–82.
- Lupski JR, Stankiewicz P, Genomic Disorders: Molecular Mechanisms for Rearrangements and Conveyed Phenotypes. *Plos Genetics* , 2005; 1, Issue 6 e 49.
- Lyon M. (1961). Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature*. 190:372-3.
- Maestrini E, Lai C, Marlow A, Matthews N, Wallace S, Bailey A, Cook EH, Weeks DE, Monaco AP, and IMGSAC. Serotonin transporter (5-HTT) and γ -aminobutyric acid receptor subunit β -3 (GABRB3) gene polymorphisms are not associated with autism in the IMGSAC families. *Am. J. Med. Genet.*, 1999; 88: 492-496.
- Malow BA. Sleep disorders, epilepsy, and autism.(2004) *Ment Retard Dev Disabil Res* ;10(2):122-5. Hayashi E. Effect of melatonin on sleep-wake rhythm: the sleep diary of an autistic male. *Psychiatry Clin Neurosci* 2000; **54**: 383-384.
- Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, Skaug J, Shago M, Moessner R, Pinto D, Ren Y, Thiruvahindrapuram B, Fiebig A, Schreiber S, Friedman J, Ketelaars CE, Vos YJ, Ficicioglu C, Kirkpatrick S, Nicolson R, Sloman L, Summers A, Gibbons CA, Teebi A, Chitayat D, Weksberg R, Thompson A, Vardy C, Crosbie V, Luscombe S, Baatjes R, Zwaigenbaum L, Roberts W, Fernandez B, Szatmari P, Scherer SW. (2008) Structural variation of chromosomes in autism spectrum disorder *Am J Hum Genet*. 82:477-88.
- Marui T, Koishi S, Funatogawa I, Yamamoto K, Matsumoto H, Hashimoto O, Nanba E, Kato C, Ishijima M, Watanabe K, Kasai K, Kato N, Sasaki T. No association of FOXP2 and PTPRZ1 on 7q31 with autism from the Japanese population. *Neurosci Res*. 2005 Sep;53(1):91-4.
- Marui T, Koishi S, Funatogawa I, Yamamoto K, Matsumoto H, Hashimoto O, Nanba E, Kato C, Ishijima M, Watanabe K, Kasai K, Kato N, Sasaki T. (2005) No association of FOXP2 and PTPRZ1 on 7q31 with autism from the Japanese population. *Neurosci Res.*;53(1):91-4.
- McCarrol S et al., Common deletion polymorphisms in the human genome. *Nature Genetics*, February 2006; vol. 38, num. 1: 86- 92.
- McCauley JL, Li C, Jiang L, Olson LM, Crockett G, Gainer K et al. Genome-wide and Ordered-Subset linkage analyses provide support for autism loci on 17q and 19p with evidence of phenotypic and interlocus genetic correlates. *BMC Med Genet* 2005; 6: 1.
- McCauley JL, Li C, Jiang L, Olson LM, Crockett G, Gainer K, Folstein SE, Haines JL, Sutcliffe JS. (2005) *Genome-wide and Ordered-Subset linkage analyses provide support for autism loci on 17q and 19p with evidence of phenotypic and interlocus genetic correlates*. *BMC Med Genet* 6:1.
- Mei R, et al. Genome-wide detection of allelic imbalance using human SNPs and high-density DNA arrays. *Genome Res*. 2000; 10, 1126–1137.

- Melke J, Goubran Botros H, Chaste P, Betancur C, Nygren G, Anckarsäter H, Rastam M, Ståhlberg O, Gillberg IC, Delorme R, Chabane N, Mouren-Simeoni MC, Fauchereau F, Durand CM, Chevalier F, Drouot X, Collet C, Launay JM, Leboyer M, Gillberg C, Bourgeron T. (2008) Abnormal melatonin synthesis in autism spectrum disorders. *Mol Psychiatry* ;13(1):90-8.
- Meloni et al. (2000). A mutation in the rett syndrome gene, MECP2, causes X-linked mental retardation and progressive spasticity in males. *Am J Hum Genet.* 67:982-5.
- Mencarelli MA, Caselli R, Pescucci C, Hayek G, Zappella M, Renieri A *et al.* Clinical and molecular characterization of a patient with a 2q31.2-32.3 deletion identified by array-CGH. *Am J Med Genet A* 2007; **143A**: 858-65.
- Migeon et al. (1979). Stability of the "two active X" phenotype in triploid somatic cells. *Cell.* 18:637-41.
- Monfort S, Rosello M, Orellana C, Oltra S, Blesa D, Kok K *et al.* Detection of known and novel genomic rearrangements by array based comparative genomic hybridisation: deletion of ZNF533 and duplication of CHARGE syndrome genes. *J Med Genet* 2008; **45**: 432-7.
- Morton, Sequential tests for the detection of linkage. *Am J Hum Genet* , 1955;7(3):277 –318.
- Muers et al. (2007). Defining the cause of skewed X-chromosome inactivation in X-linked mental retardation by use of a mouse model. *Am J Hum Genet.* 80:1138-49.
- Nagarajan et al. (2006). Reduced MeCP2 expression is frequent in autism frontal cortex and correlates with aberrant MECP2 promoter methylation. *Epigenetics.* 1:e1-11.
- Nan et al. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature.* 393:386-9.
- Naruse,H. *et al.* A multi-center double-blind trial of pimozide (Orap), haloperidol and placebo in children with behavioural disorders, using crossover design. *Acta Paedopsychiatr.* 48, 173-184 (1982).
- Newbury DF, Bonora E, Lamb JA, Fisher SE, Lai CS, Baird G, Jannoun L, Slonims V, Stott CM, Merricks MJ, Bolton PF, Bailey AJ, Monaco AP; International Molecular Genetic Study of Autism Consortium. FOXP2 is not a major susceptibility gene for autism or specific language impairment. *Am J Hum Genet.* 2002 May;70(5):1318-27. Epub 2002 Mar 13.
- Nguyen DK, Distèche CM. (2006). High expression of the mammalian X chromosome in brain. *Brain Res.* 1126:46-9.
- Orstavik et al. (1998). X chromosome inactivation in carriers of Barth syndrome. *Am J Hum Genet.* 63:1457-63.
- Ott J, Bhat A, Linkage analysis in heterogeneous and complex traits. *Eur Child Adolesc Psychiatry.* 1999; 3: 43-6.
- Pääbo, The mosaic that is our genome. *Nature*, 2003; 421: 409–412.
- Pardo CA, Eberhart CG The neurobiology of autism (2007) *Brain Pathol.* 45: 356-365.
- Patil, et al. Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science*, 2001 ; 294 : 1719–1723.
- Persico AM, L D'Agruma, N Maiorano, A Totaro, R Militerni, C Bravaccio, TH Wassink for the CLSA, C Schneider, R Melmed, S Trillo, F Montecchi, M Palermo, T Pascucci, S Puglisi-Allegra, K-L Reichelt, M Conciatori, R Marino, CC Quattrocchi, A Baldi, L Zelante, P Gasparini and F Keller.(2001) *Reelin gene alleles and haplotypes as a factor predisposing to autistic disorder.* *Molecular Psychiatry* (2001) 6, 150–159.
- Persico AM, Levitt P, Pimenta AF. (2006) Polymorphic GGC repeat differentially regulates human reelin gene expression levels. *J Neural Transm.* 113(10):1373-82
- Persico AM, Pascucci T, Puglisi-Allegra S, Militerni R, Bravaccio C, Schneider C, Melmed R, Trillo S, Montecchi F, Palermo M. (2002) *Serotonin transporter gene promoter variants do not explain the hyperserotonemia in autistic children.* *Mol Psychiatry.* 7: 795-800.
- Philippe A, Martinez M, Guilloudbataille M, Gillberg C, Rastam M, Sponheim E et al. Genome-wide scan for autism susceptibility genes. *Hum Mol Genet* 1999; 8: 805–812.
- Pickles A, Bolton P, Macdonald H, Bailey A, Le Couteur A, Sim CH, Rutter M. Latent-class analysis of recurrence risks for complex phenotypes with selection and measurement error: a twin and family history study of autism. *Am J Hum Genet.* 1995; 57: 717-26.
- Pinkel D. Et al., High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nature Genet.* 1998; 20: 207–211.

- Plenge et al. (2002). Skewed X-chromosome inactivation is a common feature of X-linked mental retardation disorders. *Am J Hum Genet.* 71:168-73.
- Prasad HC, Steiner JA, Sutcliffe JS, Blakely RD. (2009) Enhanced activity of human serotonin transporter variants associated with autism *Philos Trans R Soc Lond B Biol Sci.* 27;364:163-73.
- Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics* 2000; **155**: 945-59.
- Raghavan M. et al. Genome-wide single nucleotide polymorphism analysis reveals frequent partial uniparental disomy due to somatic recombination in acute myeloid leukemias. *Cancer Res.* 2005; 65: 375–378.
- Ramoz N, Reichert JG, Corwin TE, Smith CJ, Silverman JM, Hollander E, Buxbaum JD. (2006) *Biol Psychiatry* 15;60(2):186-91 Lack of evidence for association of the serotonin transporter gene SLC6A4 with autism.
- Reddy KS. Cytogenetic abnormalities and fragile-X syndrome in Autism Spectrum Disorder. *BMC Med Genet.* 2005; 6: 3.
- Reich, et al. Linkage disequilibrium in the human genome. *Nature*, 2001; 411: 199–204.
- Reich, Gabriel, Altshuler, Quality and completeness of SNP databases. *Nature Genet.*, 2003; 33: 457–458.
- Reik W, Walter J. *Imprinting mechanisms in mammals.* *Curr Opin Genet Dev.* 1998 Apr;8(2):154-64.
- Remington,G., Sloman,L., Konstantareas,M., Parker,K. & Gow,R. *Clomipramine versus haloperidol in the treatment of autistic disorder: a double-blind, placebo-controlled, crossover study.* *Journal of Clinical Psychopharmacology* 21(4), 440-444. (2001).
- Ring,H. et al. *Cerebral correlates of preserved cognitive skills in autism: a functional MRI study of embedded figures task performance.* *Brain* 122(7), 1305-1315. (1999). Press, Oxford (2000).
- Risch N, Spiker D, Lotspeich L, Nouri N, Hinds D, Hallmayer J et al. A genomic screen of autism: evidence for a multilocus etiology. *Am J Hum Genet* 1999; 65: 493–507
- Risch N, Spiker D, Lotspeich L, Nouri N, Hinds D, Hallmayer J, Kalaydjieva L, McCague P, Dimiceli S, Pitts T, Nguyen L, Yang J, Harper C, Thorpe D, Vermeer S, Young H, Hebert J, Lin A, Ferguson J, Chiotti C, Wiese-Slater S, Rogers T, Salmon B, Nicholas P, Myers RM, et al. *A genomic screen of autism: evidence for a multilocus etiology.* *Am J Hum Genet.* 1999 Aug;65(2):493-507.
- Ritvo, E.R., Freeman, B. J., Pingree, C., Mason-Brothers, A., Jorde, L., Jenson, W. R., McMahon, W.M., Petersen, P.B., Mo, A., Ritvo A. (1989). *The UCLA-University of Utah epidemiologic survey of autism: Prevalence.* *American Journal Of Psychiatry*, 146, 194-199.
- Rodriguez IR, Mazuruk K, Schoen TJ, Chader GJ. Structural analysis of the human hydroxyindole-O-methyltransferase gene. Presence of two distinct promoters. *J Biol Chem* 1994; **269**: 31969-31977.
- Rogers et al. (2001). The behavioral phenotype in fragile X: symptoms of autism in very young children with fragile X syndrome, idiopathic autism, and other developmental disorders. *J Dev Behav Pediatr.* 22:409-17.
- Rutter M. Genetic studies of autism: from the 1970s into the millennium. *J Abnorm Child Psychol.* 28(1):3-14 (2000)
- Sanchez,L.E. et al. *A pilot study of clomipramine in young autistic children.* *J. Am. Acad. Child Adolesc. Psychiatry* 35, 537-544 (1996).
- Sankaranarayanan K, Wassom JS, Ionizing radiation and genetic risks XIV. Potential research directions in the post-genome era based on knowledge of repair of radiation-induced DNA double-strand breaks in mammalian somatic cells and the origin of deletions associated with human genomic disorders. *Mutat Res.* 2005; 578: 3.
- Schmidt M, Du Sart D. (1992). Functional disomies of the X chromosome influence the cell selection and hence the X inactivation pattern in females with balanced X-autosome translocations: a review of 122 cases. *Am J Med Genet.* 42:161-9.
- Sebat J. et al., Large-scale copy number polymorphism in the human genome. *Science*, 2004; 305: 525–528.
- Shao YJ, Wolpert CM, Raiford KL, Menold MM, Donnelly SL, Ravan SA et al. Genomic screen and follow-up analysis for autistic disorder. *Am J Med Genet* 2002; 114: 99–105.

- Sharp AJ, Cheng Z, Eichler EE, Structural Variation of the Human Genome. *Annu. Rev: Human Genet.* 2006; 7:407-442.
- Sharp AJ, et al. Segmental duplications and copy number variation in the human genome. *Am. J. Hum. Genet.* 2005; 77: 78–88.
- Silverman JM, Smith CJ, Schmeidler J, Hollander E, Lawlor BA, Fitzgerald M, Buxbaum JD, Delaney K, Galvin P; Autism Genetic Research Exchange Consortium. Symptom domains in autism and related conditions: evidence for familiarity. *Am J Med Genet.* 2002 Jan 8;114(1):64-73
- Singer-Sam et al. (1990). Use of a HpaII-polymerase chain reaction assay to study DNA methylation in the Pcg-1 CpG island of mouse embryos at the time of X-chromosome inactivation. *Mol Cell Biol.* 10:4987-9.
- Skaar DA, Shao Y, Haines JL, Stenger JE, Jaworski J, Martin ER, DeLong GR, Moore JH, McCauley JL, Sutcliffe JS, Ashley-Koch AE, Cuccaro ML, Folstein SE, Gilbert JR, Pericak-Vance MA.(2005) Analysis of the RELN gene as a genetic risk factor for autism. *Mol Psychiatry.*10:563-71.
- Slager SL, Schaid DJ, Case-control studies of genetic markers: power and sample size approximations for Armitage’s test for trend. *Hum Hered* , 2001; 52: 149– 53.
- Slater SR et al. High-resolution identification of chromosomal abnormalities using oligonucleotide arrays containing 116,204 SNPs. *Am. J. Hum. Genet.* 2005; 77: 709–726.
- Smalley SL, Asarnow RF, Spence A. Autism and genetics. *Arch Gen Psychiatry.* 1988; 45: 953–961.
- Smith et al., Molecular genetic delineation of 2q37.3 deletion in autism and osteodystrophy: report of a case and of new markers for deletion screening by PCR. *Cytogenet. Cell. Genet.*, 2001; 94: 15-22.
- Solinas-Toldo, et al., Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* 1997; 20: 399–407.
- Sousa I, Clark TG, Toma C, Kobayashi K, Choma M, Holt R, Sykes NH, Lamb JA, Bailey AJ, Battaglia A, Maestrini E, Monaco AP. MET and autism susceptibility: family and case-control studies. (2008) *Eur J Hum Genet.*
- Speicher MR. Monitoring chromosome rearrangements. *Adv Exp Med Biol.* 2005;570:19-41.
- Spence SJ, Schneider MT. The Role of Epilepsy and Epileptiform EEGs in Autism Spectrum Disorders. (2009) *Pediatr Res.*
- Spielman RS, McGinnis RE, Ewens WJ, Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet*, 1993; 52: 506– 16.
- Stone JL, Merriman B, Cantor RM, Yonan AL, Gilliam TC, Geschwind DH et al. Evidence for sexspecific risk alleles in autism spectrum disorder. *Am J Hum Genet* 2004; 75: 1117–1123.
- Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, Redon R, Bird CP, de Grassi A, Lee C, Tyler-Smith C, Carter N, Scherer SW, Tavaré S, Deloukas P, Hurles ME, Dermitzakis ET. (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science.* 315:848-53.
- Suarez BK, Hampe CL, Linkage and association. *Am J Hum Genet.* 1994; 54: 554-9.
- Sutcliffe JS, Nelson DL, Zhang F, Pieretti M, Caskey CT, Saxe D, Warren ST.DNA methylation represses FMR-1 transcription in fragile X syndrome. *Hum Mol Genet.* 1992 Sep;1(6):397-400.
- Szatmari P, Jones MB, Zwaigenbaum L, MacLean JE. Genetics of autism: overview and new directions. *J Autism Dev Disord.* 1998 Oct;28(5):351-68
- Tabuchi et al. (2007). A neuroligin-3 mutation implicated in autism increases inhibitory synaptic transmission in mice. *Science.* 318:71-6.
- Talebizadeh et al. (2005). Brief report: non-random X chromosome inactivation in females with autism. *J Autism Dev Disord.* 35:675-81.
- The International HapMap Consortium, A haplotype map of the human genome. *Nature*, 2005; 437: 1299-1320.
- The International HapMap Consortium, The International HapMap Project. *Nature*, 2003; 426: 789-796.
- THERMAN E, PATAU K. 1974. Abnormal X chromosomes in man: origin, behaviour and effects. *Humangenetik* 25: 1-16.

- Tordjman S, Anderson GM, Pichard N, Charbuy H, Touitou Y. Nocturnal excretion of 6-sulphatoxymelatonin in children and adolescents with autistic disorder. *Biol Psychiatry* 2005; **57**: 134-138.
- Tordjman S, Gutknecht L, Carlier M, Spitz E, Antoine C, Slama F, Carsalade V, Cohen DJ, Ferrari P, Roubertoux PL et al. (2001) *Role of the serotonin transporter gene in the behavioral expression of autism*. *Mol Psychiatry* 6: 434-9.
- Trikalinos TA, Karvouni A, Zintzaras E, Ylisaukkoja T, Peltonen L, Jarvela I et al. A heterogeneity-based genome search metaanalysis for autism-spectrum disorders. *Mol Psychiatry* 2006; 11: 29–36.153
- Tuzun E, Sharp AJ, Bailey JA, Kaul R, Morrison VA, Pertz LM, Haugen E, Hayden H, Albertson D, Pinkel D, Olson MV, Eichler EE, Fine-scale structural variation of the human genome. *Nat Genet.* 2005; 37: 660-1.
- Ueda S, Fujimoto S, Hiramoto K, Negishi M, Katoh H. Dock4 regulates dendritic development in hippocampal neurons. *J Neurosci Res* 2008.
- Van Esch et al. (2005). Duplication of the MECP2 region is a frequent cause of severe mental retardation and progressive neurological symptoms in males. *Am J Hum Genet.* 77:442-53.
- Vernes SC, Newbury DF, Abrahams BS, Winchester L, Nicod J, Groszer M, Alarcón M, Oliver PL, Davies KE, Geschwind DH, Monaco AP, Fisher SE. (2008) A functional genetic link between distinct developmental language disorders. *N Engl J Med.* 27;359(22):2337-45. Epub 2008 Nov 5.
- Vernes SC, Newbury DF, Abrahams BS, Winchester L, Nicod J, Groszer M, Alarcón M, Oliver PL, Davies KE, Geschwind DH, Monaco AP, Fisher SE. (2008) A functional genetic link between distinct developmental language disorders. *N Engl J Med.*;359(22):2337-45
- Vincent JB, Herbrick JA, Gurling HM, Bolton PF, Roberts W, Scherer SW. Identification of a novel gene on chromosome 7q31 that is interrupted by a translocation breakpoint in an autistic individual. *Am J Hum Genet.* 2000 Aug;67(2):510-4. Epub 2000 Jul 7.
- Vincent JB, Kolozsvari D, Roberts WS, Bolton PF, Gurling HM, Scherer SW. Mutation screening of X-chromosomal neuroligin genes: no mutations in 196 autism probands. *Am J Med Genet B Neuropsychiatr Genet.* 2004;129: 82-4.
- Volkmar F., Lord C., Bailey A., Schultz R., Klin A. *Autism and pervasive developmental disorders*. *Journal of Child Psychology and Psychiatry* 45:1 (2004), pp 135-170.
- Volkmar, F., Szatmari, P., Sparrow, S.S. (1993). *Sex difference in pervasive developmental disorders*. *Journal of Autism and developmental disorders.* 23, 579-591.
- Volkmar, F.R. & Nelson, D.S. *Seizure disorders in autism*. *J. Am. Acad. Child Adolesc. Psychiatry* 29, 127-129 (1990).
- Von Gall C, Garabette ML, Kell CA, Frenzel S, Dehghani F, Schumm-Draeger PM, et al. Rhythmic gene expression in pituitary depends on heterologous sensitization by the neurohormone melatonin. *Nat Neurosci* 2002;5:234–238.
- Vorstman JA, Staal WG, Hochstenbach PF, Franke L, van Daalen E, van Engeland H. Overview of cytogenetic regions of interest (CROIs) associated with the autism phenotype across the human genome. *Mol Psychiatry* 2006; 11: 18–28.
- Vourc'h P, Bienvenu T, Beldjord C, Chelly J, Barthélémy C, Müh JP, Andres C. No mutations in the coding region of the Rett syndrome gene MECP2 in 59 autistic patients. *Eur J Hum Genet.* 2001 Jul;9(7):556-8.
- Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 2008; **320**: 539-43.
- Warburton D., De novo balanced chromosome rearrangements and extra marker chromosomes identified at prenatal diagnosis: clinical significance and distribution of breakpoints. *Am. J. Hum. Genet.* 1991; 49:995–1013.
- Warburton P, Baird G, Chen W, Morris K, Jacobs BW, Hodgson S, Docherty Z. Support for linkage of autism and specific language impairment to 7q3 from two chromosome rearrangements involving band 7q31. *Am J Med Genet.* 2000 Apr 3;96(2):228-34.

- Wassink TH, Brzustowicz LM, Bartlett CW, Szatmari P. The search for autism disease genes. *Mental Retard Dev Disabil Res Rev* 2004; 10: 272–283.
- Wassink TH, Piven J, Patil SR. Chromosomal abnormalities in a clinic sample of individuals with autistic disorder. *Psychiatr Genet* 2001; 11: 57–63.
- Wassink TH, Piven J, Vieland VJ, Pietila J, Goedken RJ, Folstein SE, Sheffield VC. Evaluation of FOXP2 as an autism susceptibility gene. *Am J Med Genet.* 2002 Jul 8;114(5):566-9.
- Weeks DE, Lathrop GM, Polygenic disease: methods for mapping complex disease traits. *Trends Genet.* 1995; 11: 513-9.
- Wermter et al. (2008). No evidence for involvement of genetic variants in the X-linked neuroligin genes NLGN3 and NLGN4X in probands with autism spectrum disorder on high functioning level. *Am J Med Genet B Neuropsychiatr Genet.* 147B:535-7.
- Whitney ER, Kemper TL, Bauman ML, Rosene DL, Blatt GJ. Cerebellar Purkinje cells are reduced in a subpopulation of autistic brains: a stereological experiment using calbindin-D28k. (2008) *Cerebellum.*;7(3):406-16.
- Wigginton JE, Abecasis GR. PEDSTATS: descriptive statistics, graphics and quality assessment for gene mapping data. *Bioinformatics* 2005; **21**: 3445-7.
- Wilson HL, Wong ACC, Shaw SR, Tse, WY, Stapleton GA, Phelan MC, Hu S, Marshall J, McDermid HE. Molecular characterisation of the 22q13 deletion syndrome supports the role of haploinsufficiency of SHANK3/PROSAP2 in the major neurological symptoms. *J. Med. Genet.* 2003; 40: 575-584.
- Wing, L., Potter, D (2002). *The epidemiology of autistic spectrum disorders; is the prevalence rising?* Mental Retardation and developmental disabilities Research Review, 8, 151-161
- World Health Organization (1992). *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines.* WHO, Geneva Switzerland.
- Wu S, Guo Y, Jia M, Ruan Y, Shuang M, Liu J, Gong X, Zhang Y, Yang J, Yang X (2005) *Lack of evidence for association between the serotonin transporter gene (SLC6A4) polymorphisms and autism in the Chinese trios.* *Neurosci Lett.* 381: 1-5.
- Xi CY, Ma HW, Lu Y, Zhao YJ, Hua TY, Zhao Y, Ji YH. MeCP2 gene mutation analysis in autistic boys with developmental regression. (2007) *Psychiatr Genet.* 17(2):113-6.
- Yajnik V, Paulding C, Sordella R, McClatchey AI, Saito M, Wahrer DC *et al.* DOCK4, a GTPase activator, is disrupted during tumorigenesis. *Cell* 2003; **112**: 673-84.
- Yan J, Noltner K, Feng J, Li W, Schroer R, Skinner C, Zeng W, Schwartz CE, Sommer SS. Neurexin 1alpha structural variants associated with autism (2008) *Neurosci Lett.*438:368-70.
- Yan J, Oliveira G, Coutinho A, Yang C, Feng J, Katz C, Sram J, Bockholt A, Jones IR, Craddock N, Cook EH Jr, Vicente A, Sommer SS Analysis of the neuroligin 3 and 4 genes in autism and other neuropsychiatric patients. *Mol Psychiatry.* 2005; 10: 329-32.
- Yirmiya N, Pilowsky T, Nemanov L, Arbelle S, Feinsilver T, Fried I, Ebstein RP. (2001) *Evidence for an association with the serotonin transporter promoter region polymorphism and autism.* *Am J Med Genet* 105: 381-6.
- Yonan AL, Alarcon M, Cheng R, Magnusson PKE, Spence SJ, Palmer AA *et al.* A genomewide screen of 345 families for autism susceptibility loci. *Am J Hum Genet* 2003; 73: 886–897.
- Yu CE, Dawson G, Munson J, D'Souza I, Osterling J, Estes A, Leutenegger AL, Flodman P, Smith M, Raskind WH, Spence MA, McMahon W, Wijsman EM, Schellenberg GD. Presence of large deletions in kindreds with autism. *Am J Hum Genet.* 2002 Jul;71(1):100-15.
- Zappella *et al.* (2001). Preserved speech variants of the Rett syndrome: molecular and clinical analysis. *Am J Med Genet.* 104:14-22.
- Zhang H, Liu X, Zhang C., Mundo E., Macciardi F., Grayson DR., Guidotti AR and Holden JJA. *Reelin gene alleles and susceptibility to autism spectrum disorders.* *Molecular Psychiatry,* 7, 1012-1017 (2002).
- Zhong N, Ye L, Ju W, Brown WT, Tsiouris J, Cohen I. (1999) *5-HTTLPR variants not associated with autistic spectrum disorders.* *Neurogenetics.* 2: 129-31.
- Zoghbi HY. Postnatal neurodevelopmental disorders: meeting at the synapse. *Science.* 2003; O302: 826-30.

