

**UNIVERSITA' DEGLI STUDI DI NAPOLI
" FEDERICO II "**



**Dottorato di Ricerca in Biologia Avanzata
Indirizzo
SISTEMATICA MOLECOLARE**

XX ciclo

L'ESPANSIONE ARXdup24 NEL GENE ARISTALESS HOMEobox

CAUSA RITARDO MENTALE X-LINKED:

ANALISI GENETICA E CONSEGUENZE FUNZIONALI

**Coordinatore
Prof.ssa Silvana Filosa**

**Tutore
Prof. Giuseppe Saccone**

**Candidata
Dott.ssa Carmela Laperuta**

**Co-Tutori
Dott.ssa Matilde Valeria Ursini**

Dott.ssa Maria Giuseppina Miano

Anno Accademico 2007-2008

INDICE

	INDICE	2
1	INTRODUZIONE	4
1.1	Malattie da espansione di triplette	5
	<i>Geni e malattie associati ad espansioni di poliglutammine</i>	8
	<i>Geni e malattie associati ad espansioni di polialanine</i>	11
1.2	Fattori di trascrizione ed i tratti di polialanine	15
1.3	<i>ARX</i> , un gene "hot spot" per mutazioni MRX	18
1.4	Eterogeneità fenotipica ed effetto pleiotropico delle mutazioni di ARX	20
1.5	Il modello murino transgenico ARX KO	23
1.6	Il cromosoma X nell'evoluzione dei vertebrati: diversificazione delle capacità cognitive dell'uomo	26
1.7	I Geni implicati nello sviluppo e nell'evoluzione del cervello: metodo del Ka/Ks	28
1.8	Linkage Genetico e Microsatelliti	30
2	<i>SCOPO DEL LAVORO</i>	36
3	<i>MATERIALI & METODI</i>	40
3.1	Estrazione di DNA	41
3.2	Analisi dei microsatelliti	41
3.3	Sequenziatore di DNA	43
3.4	Organizzazione esone/introne di geni candidati	45

3.5	Purificazione e sequenziamento automatico diretto di frammenti di PCR	45
3.6	Analisi dello stato di inattivazione del cromosoma X (Test HUMARA)	45
3.7	Programmi utilizzati	46
4	<i>RISULTATI</i>	49
4.1	Descrizione della famiglia MRX87	50
4.2	Identificazione del locus MRX87	50
4.3	Identificazione del difetto molecolare responsabile di MRX87	53
4.4	Ruolo patogenetico e conseguenze funzionali di ARXdup24 in ARX	56
4.5	Analisi Genomico-Comparativa dei domini di poli-Alanine di ARX	57
5	<i>DISCUSSIONE</i>	61
6	<i>FIGURE</i>	67
7	<i>BIBLIOGRAFIA</i>	101

1. INTRODUZIONE

1.1 Malattie da espansione di triplette

Le ripetizioni di triplette nel DNA sono presenti in numerosi geni e codificano per lunghe catene composte da un numero ripetuto di residui aminoacidici uguali. Attualmente il numero delle patologie neurologiche e neuromuscolari causate da espansioni in triplette instabili è in crescente aumento (Pearson CE et al. 2005).

La posizione della tripletta espansa rispetto alla regione codificante è certamente variabile nelle diverse patologie identificate. Espansioni di triplette sono state identificate sia nelle regioni 5' non tradotte di geni, all'interno della regione codificante (in esoni), in regioni introniche oppure nelle regioni 3' non tradotte (Figura 1) (Lutz RE 2007; Sinden RR et al. 2002). Triplette ripetute possono trovarsi in sequenze di geni che codificano per enzimi, per recettori ormonali, per fattori di trascrizione ed in sequenze regolatrici della trascrizione (Sinden RR et al. 2002).

La struttura che la tripletta ripetuta va a formare, è probabilmente implicata nella formazione della struttura tridimensionale della molecola, nella formazione, ad esempio, di DNA flessibile o curvo oppure di specifiche strutture che facilitano l'interazione del DNA stesso con altri elementi. La formazione dei nucleosomi, l'interazione del DNA con gli istoni, l'interazione di complessi repressori o attivatori o con le DNA e/o RNA polimerasi sono tutti processi che dipendono fortemente da un'unica struttura tridimensionale che si forma grazie alla presenza di triplette di DNA ripetute (Sinden RR et al. 2002).

Espansioni di trinucleotidi instabili sono responsabili di numerose patologie neurologiche, che vanno dal ritardo mentale sindromico e non sindromico a disordini neurodegenerativi come la Corea di Huntington e le atassie (Orr TH and Zoghbi HY 2007).

Le patologie causate da espansioni di triplette rappresentano un gruppo di disordini davvero eterogeneo ed variabile, sia dal punto di vista della regione in cui è presente la tripletta ripetuta, sia per le conseguenze che l'espansione causa a livello della struttura del DNA. La lunghezza dell'espansione è, in questo contesto, determinante. E' il caso specifico della sindrome dell'X fragile (FRAXA, OMIM 309550), la prima malattia neuromuscolare associata ad una tripletta di DNA espansa (Figura 1). L'analisi di sequenza ha rivelato la presenza di una tripletta ripetuta (CGG)_n presente nella regione 5' non tradotta del gene *FMRI* (Fu YH et al. 1991). Alleli normali contengono da 6 a 55 triplette ripetute, intervallate da una o più sequenze AGG che stabilizzano i tratti CGG (Verkerk AJMH et al. 1993). La perdita di queste triplette AGG e l'espansione del tratto CGG fino ad arrivare a 230 ripetizioni causa la malattia accompagnata da una ipermetilazione della tripletta (CGG)_n nel promotore di *FMRI* (Sutcliffe JS et al. 1992). Questa ipermetilazione causa lo spegnimento della trascrizione del gene *FMRI*, ne consegue una perdita di funzione della proteina FMRP implicata nella formazione di complessi per il trasporto dell'RNA messaggero (Ashley CT jr et al. 1993) e un successivo anormale metabolismo dell'RNA. Maschi con la premutazione sviluppano la neurodegenerazione in ritardo (Hagerman PJ & Hagerman RJ 2004a,

Hagerman PJ & Hagerman RJ 2004b). FRAXA ha una incidenza di 1:4000 ed è responsabile del 30% dei ritardi mentali legati al cromosoma X (Sherman SL et al. 1996; Turner G et al. 1996).

La malattia causa ritardo mentale che può essere da moderato a severo, macrorchidismo, orecchie grandi, iperattività. Le femmine sintomatiche mostrano un fenotipo meno grave grazie alla compensazione del dosaggio dovuta alla inattivazione del cromosoma X che porta l'allele mutato (Abrams MT et al 1994; Rousseau F et al. 1991). Nella FRAXA si riscontra anche premutazione; inoltre femmine che portano l'allele premutato sono a rischio di POF (Premature Ovarian Failure). Nel laboratorio dove ho svolto il mio progetto di dottorato è stato dimostrato che il gene FMR1 agisce come fattore di rischio per la POF, con un rischio relativo stimato del 20% (Miano MG et al 2007). In realtà, ancora non è noto quale sia la natura del difetto molecolare delle forme di POF associate ad anomalie del braccio lungo del cromosoma X (Fimiani G et al. 2005). Accanto alla sindrome dell'X fragile ci sono altre patologie causate da espansione di triplette di DNA. Si tratta dell'Atassia di Friedreich (X25), della Distrofia Miotonica (DM1), dell' Atassia Spino-Cerebellare di tipo 8 e 10 (SCA8 e SCA 10) (Lutz RE 2007) (figura 1).

L'Atassia di Friedreich (FRDA, OMIM 229300) è un disordine autosomico recessivo causato dall'espansione della tripletta instabile GAA. La tripletta si trova nel primo introne del gene X25, al centro di una sequenza *Al*. X25 codifica per Frataxina, una proteina implicata con la biosintesi del ferro (Campuzano V et al. 1996). L'allele normale contiene 7-

34 ripetizioni, l'allele mutato contiene un numero di ripetizioni ≥ 100 (Lutz RE 2007) (Figura 1).

Espansioni della tripletta (CTG)_n nel gene *DMPK* sono responsabili della Distrofia Miotonica (DM1, OMIM 160900). Si tratta di un disordine dominante con manifestazioni cliniche variabili e forte anticipazione genica (Harper PS 1989; Howeler CJ et al. 1989). I sintomi sono miotonia, progressivo indebolimento muscolare, atrofia testicolare, difetti renali, difetti cardiaci e di regolazione dell'insulina. La mutazione causativa della DM è una espansione della tripletta CTG nella regione 3' non tradotta del gene codificante per la proteina chinasi DMPK (Brook JD et al 1992; Mahadevan M et al. 1992). I cromosomi normali contengono la tripletta CTG ripetuta 5-37 volte, laddove gli individui affetti mostrano la tripletta ripetuta da 50 a 200 volte (Anvert M et al. 1993; Tsilfids C et al. 1992) (Figura 1).

L'Atassia Spino-Cerebellare tipo 8 e 10 (SCA8, OMIM 608768; SCA10, OMIM 603516) sono una patologia neurologica caratterizzata da una progressiva degenerazione del cervelletto e del tratto spinale (Greenfield JG 1954). Esse si distinguono dalle altre forme di SCA, perché determinate non da una espansione di poliglutammine bensì dal fatto che la tripletta ripetuta è trascritta sia in direzione 5'→3' che nella direzione 3'→5' (Koob MD et al. 1999). Ciò causa una alterazione nella regolazione dell'RNA di altri geni che contengono quella sequenza specifica (Figura 1).

* * *

Geni e malattie associati ad espansioni di poliglutammine.

Le triplette di DNA $(CAG)_n/(GAA)_n/(CTG)_n/(CGG)_n$ codificano per la glutammina. Espansioni di poliglutammine (poliQ) sono coinvolte in moltissime patologie neurodegenerative. Si tratta di malattie come la Corea di Huntington (HD), l'Atrofia Muscolare Spinobulbare (SBMA), l'Atrofia Dentarorubro-Pallidolusiana (DRPLA), l'Atassia Spino-Cerebellare SCA 1, 2, 3, 6, 7, 12, 17 (Figura 1).

Il meccanismo causativo delle patologia, in questo tipo di alterazioni è l'errore di scivolamento della DNA polimerasi. D'altra parte le ripetizioni imperfette di trinucleotidi sono loci che ben si prestano a questo tipo di mutazioni (Amiel J et al. 2004). Caratteristica fondamentale delle espansioni di poliQ è che mostrano una grande variabilità nella lunghezza del tratto espanso attraverso le generazioni. Altra peculiarità delle malattie da poliQ è la presenza di una progressiva neurodegenerazione con l'avanzare della malattia (Albrecht A and Mundlos S 2005).

Gli effetti neurodegenerativi relati alle espansioni di poliQ sono sempre associati a meccanismi che alterano la sopravvivenza cellulare quali la formazione di aggregati nucleari e citoplasmatici tossici che alterano il trasporto cellulare nei neuroni (Lutz RE 2007). Infatti nella maggior parte dei disordini dovuti a espansione di triplette, si osserva una acquisizione di funzione della proteina mutante che spesso porta alla formazione di inclusioni intracitoplasmatiche e/o intranucleari. La perdita del corretto assetto della proteina può interferire con la trascrizione del DNA ed il processamento dell'RNA, oppure causare l'attivazione del processo apoptotico (Lutz RE 2007).

Espansioni della tripletta (CAG)_n nel gene *HD* sono responsabili della Corea di Huntington (HD, OMIM 143100), una grave e progressiva degenerazione del SNC che porta alla demenza. Il numero di ripetute della tripletta (CAG)_n che si trova nel primo esone del gene HD passa da un valore normale di 6-35 (normale) ad un valore che può variare da 36 a 131 nell'individuo affetto (Cummings CJ and Zoghbi HY 2000). Nella Corea di Huntington si riscontra il fenomeno dell'anticipazione: l'aumento del numero delle triplette ripetute causa una neurodegenerazione sempre progressiva. Si osserva inoltre la presenza di aggregati intracellulari tossici della proteina mutata (Lutz RE 2007) (Figura 1). La proteina HD normale ha funzioni di mediazione ed interazione con numerose proteine nel SNC (Orr TH and Zoghbi HY 2007).

La tripletta polimorfica (CAG)_n che si trova nel primo esone del gene *AR* del recettore degli androgeni è responsabile della Atrofia Muscolare Spinobulbare (SBMA, Malattia di Kennedy, OMIM 313200). La proteina-ormone AR presenta un recettore nucleare con regioni di legame al DNA molto conservate capaci di attivare la trascrizione di geni responsivi all'ormone. La funzione del tratto di poliglutammine nella proteina-ormone AR è sconosciuta, inoltre l'espansione non altera il legame dell'ormone al DNA (Chamberlain NL et al. 1994; Mhatre AN et al 1993; Nakajima H et al 1996), ma causa solo una parziale perdita della funzione del recettore che determinano la sitomatologia della malattia come l'insensibilità agli androgeni che causa la femminilizzazione testicolare. Si osservano

inclusioni nucleari contenenti la proteina mutante nei motoneuroni del ponte, della medulla e della corda spinale (Li M et al. 1998) (Figura 1).

Mutazioni da espansioni di poliQ nella proteina Atrofina 1 causano invece l'Atrofia Dentarorubro-Pallidolusiana (DRPLA, OMIM 125370), una rara malattia dominante, associata ad atassia cerebrale, demenza e degenerazione cognitiva (Lutz RE 2007). In questo caso la tripletta (CAG)_n è localizzata nella porzione C-terminale del gene. Il normale "range" di ripetizione della tripletta varia da 6 a 35 copie, l'allele espanso porta invece da 49 a 88 ripetute (Koide R et al. 1994). La funzione della proteina Atrofina 1 è ad oggi sconosciuta ma essa è probabilmente implicata nella regolazione della trascrizione (Figura 1).

I geni SCA sono responsabili di tutte le diverse forme di Atassia Spino-Cerebellare (SCA1, OMIM 164400; SCA2, OMIM 183190; SCA3, OMIM 109150; SCA6, OMIM 183086; SCA7, OMIM 164500; SCA12, OMIM 604326; SCA17, OMIM 607136). La tripletta (CAG)_n si trova nella regione codificante di ciascun gene. Mutazioni da espansioni sono responsabili di forme diverse di atassia cronica, che presentano caratteri distintivi a seconda del gene mutato (Lutz RE 2007) (Figura 1).

* * *

Geni e malattie associati ad espansioni di polialanine.

Sempre più spesso negli ultimi anni, sono state identificate malattie dovute all'espansione di un codone imperfetto che codifica per l'alanina: (GCC)_n/(GCG)_n/(GCA)_n/(GCT)_n (Amiel J et al. 2004) (Figura 1). Espansioni di triplette (GCC)_n/(GCG)_n/(GCA)_n/(GCT)_n che codificano per

lunghi tratti di alanina (poliA), causano patologie come la Sinpolidattilia di tipo II (*HOXD13*), la Displasia dento-facciale (*RUNX2*), la Sindrome di Haddad (*PHOX2B*), l' Oloprosencefalia (*ZIC2*), la Distrofia Muscolare Oculofaringea (*PABPN1*), una Sindrome con malformazione a mani, piedi e genitali (*HOXA13*), la Blefarofimosi (*FOXL2*), il Ritardo Mentale con difetti della crescita (*SOX3*), Ritardo Mentale ed Epilessia (*ARX*), Sindrome di West (*ARX*), Sindrome di Partington (*ARX*), Ritardo Mentale X-linked e difetti minori (*ARX*) (Brown LY and Brown SA, 2004) (figura 2 e 3).

Le attuali conoscenze relative alle espansioni di tratti di polialanine indicano che il meccanismo causativo e le conseguenze fenotipiche sono differenti rispetto a quelli relativi alle espansioni da poliQ (Brown LY, Brown SA, 2004). Il meccanismo mutazionale causativo dell'espansione del tratto poliA e quindi della patologia, è una ricombinazione allelica ineguale durante la meiosi, il cui sito e la lunghezza sono generalmente stabili (Albrecht A, Mundlos S 2005). Inoltre, a differenza di un normale tratto di poliQ, che contiene 40 triplette ripetute, un normale tratto di poliA può contenerne al massimo 15 (Brown LY, Brown SA 2004). Contrariamente ai tratti di poliglutammine, inoltre, le catene di polialanine sono presenti in molti fattori di trascrizione. In letteratura sono infatti descritti nove geni malattia associati a espansioni di alanina. Di questi, otto codificano per fattori di trascrizione omeotici (Albrecht A, Mundlos S 2005). Analogamente alle malattie da espansione di poliglutammine (poliQ), anche in alcuni casi di patologie da espansione di polialanine (poliA), è stata mostrata la formazione di aggregati nucleari e l'aumento di mortalità

cellulare. Tale effetto è probabilmente dovuto, come nel caso delle malattie da poliQ, ad una acquisizione di funzione della proteina mutata (Nasrallah IM et al. 2004).

Mutazioni nel gene *HOXD13* causano la Sinpolidattilia di tipo II (SPDII, OMIM186000), una rara malattia dominante, nella quale gli individui affetti mostrano un eccesso di dita che sono fuse tra loro. A causare la malattia è una espansione di un tratto di alanine all' estremità N-terminale del fattore di trascrizione *HOXD13*. Il tratto di alanine passa da un valore di 15 della proteina normale a 22-29 dell'individuo affetto da SPD (Figura 2). Nella SPDII, la lunghezza del tratto espanso correla con la gravità del fenotipo (Brown LY, Brown SA 2004).

RUNX2 è il gene responsabile della Displasia Cleidocranica (CCD, OMIM 119600). La malattia è causata da una espansione in *RUNX2*, un fattore di trascrizione implicato nella differenziazione degli osteoblasti. La malattia da infatti origine a gravi difetti all'apparato scheletrico e bassa statura, ed è dovuta all' espansione di un tratto, che da 17 alanine, in individui normali, passa ad un valore di 27 alanine negli affetti (Brown LY, Brown SA 2004) (Figura 2).

Mutazioni nel fattore di trascrizione *PHOX2B* causano una patologia del sistema nervoso autonomo, denominata sindrome Congenita del centro di ipoventilazione (CCHS, OMIM 209880), un disordine autosomico dominante che causa una risposta ventilatoria anormale. La proteina *PHOX2B* contiene 2 domini di polialanine e l'espansione di uno di questi determina la malattia. Il tratto normale consta di 20 alanine, mentre il tratto

malattia può arrivare ad un numero di 25 e fino a 29 residui di alanine (figura 2) (Brown LY, Brown SA 2004).

Mutazioni nel gene *ZIC2* causano l' Oloprosencefalia (HPE, OMIM 603073.0003), una grave malformazione del cervello, con esito letale. *ZIC2* è un fattore di trascrizione che presenta all'estremità C terminale un tratto di 15 alanine. Tale tratto negli affetti da HPE si espande fino ad un valore di 25 residui di alanina (Brown LY, Brown SA 2004) (Figura 2).

Il gene responsabile della Distrofia Muscolare Oculofaringea (OPMD, 164300) è *PABPN1*, l'unico a non essere un fattore di trascrizione bensì una proteina che lega le poli-adenine nucleari. La proteina PABPN1 possiede normalmente un tratto di 10 alanine in pazienti OPMD, tale tratto raggiunge una lunghezza che varia da 12 a 17 alanine (figura 2). Nella OPMD è stata osservata la formazione di inclusioni nucleari dovute ad aggregazione della proteina (Brown LY, Brown SA 2004).

Espansioni nel gene *FOXL2* causano invece la Blefarofimosi ed epicanto inverso (BPES, OMIM 110100). È' una condizione autosomica dominante, che presenta anomalie degli occhi associate a menopausa precoce (POF). In *FOXL2* sono state descritte mutazioni a carico di ognuno dei suoi 4 diversi tratti di poliA presenti nella proteina (Albrecht A, Mundlos S 2005) (Figura 3).

Una espansione di un tratto di alanine da 15 a 26 residui in *SOX3* consiste (Figura 3) determina ritardo mentale e difetti della crescita (OMIM 300123). In associazione a questo fenotipo è stata anche descritta una mutazione che provoca una contrazione del numero di alanine (Albrecht A,

Mundlos S 2005).

Il gene *HOXA13* è stato trovato mutato in pazienti con una rara sindrome che causa anomalie delle mani, dei piedi e dei genitali (HFGS, OMIM 140000). Si tratta di una condizione autosomica dominante causata da espansioni di alanine che possono riguardare due diversi domini della proteina. Nel primo caso il tratto che si duplica (aa 74-84) passa da un valore di 12 della proteina normale ad un valore di 18 alanine nell'affetto; nel secondo caso, il tratto (aa 116-133) passa da un valore di 18 alanine della proteina normale a 27 nell'affetto (Figura 3) (Brown LY, Brown SA 2004).

Mutazioni nel gene *ARX* sono responsabili di patologie neurologiche differenti. Si evidenzia infatti una grande eterogeneità di fenotipi causati da espansioni nei primi due domini di poliA dei quattro che la proteina *ARX* possiede. Si tratta della Sindrome di West (ISSX, OMIM 308350), che causa spasmi infantili e severo Ritardo Mentale (Kato M. et al. 2003; Stromme P. et al 2002b). L'ISSX è causata da espansioni che cadono nel primo dominio di poliA della proteina *ARX* (aa 100-115). In questo caso il tratto espanso passa da un normale valore di 16 alanine a 23 alanine, come è evidente negli individui affetti da ISSX. Espansioni invece del secondo dominio di poliA di *ARX* (aa 144-155) causano Ritardo Mentale Sindromico e Non Sindromico (SMR e NSMR, OMIM 300382) (Bienvenu T et al. 2002) (Figura 3).

1.2 Fattori di trascrizione e tratti di polialanine.

Espansioni di tratti di polialanine sono state identificate in ben nove patologie umane congenite, otto di queste sono associate a fattori di trascrizione (figure 2 e 3): *HOXD13* (*Synpolydactyly type II*), *RUNX2* (*Cleidocranial displasia*), *ZIC2* (*Holoprosencephaly*), *HOXA13* (*Hand-foot-geniyal syndrome*), *FOXL2* (*Blepharophimosis, ptosis and epicanthus inversus*), *SOX3* (*X-linked Mental retardation with isolated growth hormone deficiency*), *PHOX2B* (*Congenital central hypoventilation syndrome/Ondine curse*) e *ARX* (*Infantile spasms syndrome, X-linked; Partington syndrome, lissencephaly with ambiguous genitalia X-linked; mental retardation, X-linked* (Brown LY, Brown SA 2004).

Quando un'espansione si genera in un fattore di trascrizione, può accadere sia una perdita di funzione oppure l'acquisizione di una funzione anormale (Brown LY, Brown SA 2004). Il meccanismo responsabile dell'espansione di un tratto di polialanine è una ricombinazione omologa ineguale durante la meiosi. Tale meccanismo è molto frequente in tratti omopolimerici ricchi in GC con codoni imperfetti [GCN] (Amiel J et al. 2004).

Dalla letteratura si apprende che espansioni nella lunghezza dei tratti di alanine del gene *ZIC2* (*Holoprosencephaly*), influenzano la capacità e la forza del legame di *ZIC2* al DNA, determinando, conseguentemente, una alterazione nell'attività trascrizionale di *ZIC2* (Brown L. et al 2005).

Riguardo al gene *HOXD13* (*Synpolydactyly type II*), espansioni dei tratti di alanine in *HOXD13* causano un'acquisizione di funzione dei mutanti alterando la regolazione trascrizionale (Goodman FR et al. 1997).

L'espansione di un tratto di polialanine nel fattore di trascrizione *SOX3* (*X-linked Mental retardation with isolated growth hormone deficiency*) causa riduzione dell'attività trascrizionale e perdita del segnale di localizzazione nucleare delle proteine mutanti (Woods KS et al. 2005).

In *PHOX2B* (*Congenital central hypoventilation syndrome/Ondine curse*) è stata addirittura osservata una correlazione tra la lunghezza dei tratti di polialanine espansi ed il decremento nella transattivazione del promotore di *PHOX2B*, in particolare la lunghezza del tratto di alanine espanso influenza direttamente il grado di transattivazione (Trochet D et al. 2005).

Riguardo al ruolo funzionale dei domini di polialanine Karlin et al., suggerisce che essi potrebbero agire come flessibili tratti spaziatori tra i diversi domini funzionali di un proteina (Karlin S and Burge C 1996). L'allungamento di questi tratti di polialanine, in proteine mutanti, andrebbe rapidamente a perturbare l'interazione di questi domini con le regolari zone di interazione attraverso mutamenti dell'orientamento e dell'assetto sterico della proteina (Goodman FR et al. 1997). Proteine che non possiedono un corretto assetto sono rapidamente ed efficacemente riconosciute e quindi degradate. In alcuni casi, le proteine che non possiedono un assetto corretto formano aggregati, gli aggregati a loro volta causano un funzionamento eccessivo del macchinario di degradazione proteica, ciò crea un profondo scompenso nella cellula e, conseguentemente, il fenotipo patologico. Dalla letteratura apprendiamo infatti che lunghi tratti di alanine inducono la formazione di α -eliche ed inoltre, l'inserzione di alanine induce l'aumento

di elicITÀ della proteina causando la transizione a β -foglietti. Tale transizione è influenzata dal pH, dalla temperatura e dai residui amminoacidici fiancheggiati (Shinchuk LM et al 2005).

Come mostrato da Albrecht AN. et al., l'overespressione della forma espansa di HOXD13 (*Synpolydactyly type II*) causa infatti, la formazione di aggregati citoplasmatici della proteina mutata, mentre la forma normale è sempre localizzata nel nucleo. Risultati simili sono stati ottenuti per i geni *RUNX2*, *SOX3* ed *HOXA13*, suggerendo un meccanismo patogenico simile (Albrecht A, Mundlos S 2005). Recentemente, Nasrallah et al. hanno dimostrato che l'espressione "in vitro" della proteina mutata ARX PolyA_I_{16->23} (MR con ISS) genera la formazione di aggregati intranucleari citotossici in cellule COS e aumento della morte cellulare (Figura 4) (Nasrallah IE et al. 2004). Tuttavia anche se la funzione dei domini di polialanine presenti nella proteina ARX rimane al momento sconosciuta, la loro conservazione durante l'evoluzione suggerisce l'esistenza di un loro potente ruolo funzionale e/o strutturale (Albrecht A, Mundlos S, 2005).

1.3 ARX, un gene "hot spot" per mutazioni MRX

Il gene *ARX* (*Aristaless-related homeobox*) (OMIM: 300382) è dopo *FMRI* (*Fragile Mental Retardation 1*, OMIM: *309550) il gene che più frequentemente si rinviene mutato nelle forme di ritardo mentale legato X-linked. Infatti si ritiene che mutazioni in *ARX* sono responsabili di circa il 10% di tutti i casi di diagnosi di XLMR. Per questo *ARX* è considerato un sito "hot spot" per mutazioni MR (Mandel JI, Chelly J 2004; Poirier K et al.

2006; Poirier K et al. 2005; Gecz J et al. 2006; Nawara M. et al. 2006; Gestinari-Duarte Rde S et al. 2006).

Il gene *ARX* (Figura 5) mappa nella banda citogenetica Xp22.13 e codifica per un fattore trascrizionale omeotico e bi-funzionale, essenziale per lo sviluppo neuronale. In particolare, la proteina *ARX* regola lo sviluppo della testa e dei territori cerebrali ed è necessario allo sviluppo ed al mantenimento dei neuroni GABAergici (Colombo E et al. 2004). Il gene *Aristaless* appartiene alla famiglia genica dei fattori omeotici *Q₅₀ Paired-like (Prd-like)*, una famiglia molto antica di fattori di trascrizione con un ruolo chiave nelle prime fasi dell'embriogenesi, ed in particolare dello sviluppo della testa e del sistema nervoso centrale e periferico (Galliot B, Miller D. 2000; Friocourt G et al. 2006).

Il gene *ARX* copre una regione genomica di 11,7 kb nella banda citogenetica Xp22.13; è costituito da 5 esoni e codifica per una proteina di 562 amminoacidi (Figura 5). Essa è caratterizzata da un dominio molto conservato di repressione trascrizionale, l'Octapeptide (Otp) in posizione 27-34 (Mailhos C et al 1998). Segue un dominio NLS di localizzazione nucleare (nelle posizioni 82-89). Riguardo ai domini di legame al DNA, essi sono 2: il dominio "*Homeobox*", noto anche come dominio "*Paired like*" (*Prd-like*), in posizione 328-387, un dominio molto antico e conservato; e il dominio *Aristaless*, dominio *OAR*, in posizione 530-543. Il dominio *Aristaless* è molto conservato in un gruppo di proteine, orthopedia (Otp), con funzione fondamentale nella regionalizzazione del diencefalo ventrale del topo. Delezioni del dominio *Aristaless*, nelle proteine *Otp* riducono

l'attività di transattivazione (Simeone A et al. 1994). Peculiarità della proteina ARX è la presenza inoltre di quattro domini di polialanine. Si tratta di regioni composte da ripetizioni di residui di alanine (poliA). Questi domini sono distribuiti lungo la proteina tra le posizioni 100-115 (poliAI, 16 residui), 144-155 (poliAII, 12 residui), 275-281 (poliAIII, 7 residui) ed infine 432-440 (poliAIV, 9 residui) (Figura 5). Tali regioni risultano composte da un numero ripetuto di diversi residui di alanine (Nawara M. et al. 2006). La funzione di questi domini di polialanine è sconosciuta. Secondo alcuni autori tali domini sarebbero probabilmente implicati nella repressione trascrizionale (McKenzie O. et al 2007; Han K., Manley JL, 1993), altri autori invece suggeriscono un ruolo nell'interazione proteina-proteina o proteina-DNA (Brown LY, Brown SA., 2004). È interessante sottolineare che la percentuale di conservazione di questi domini di polialanine non è del 100% tra uomo e topo (Friocourt G. et al. 2006).

1.4 Eterogeneità fenotipica ed effetto pleiotropico delle mutazioni di ARX.

Mutazioni nel gene ARX (Aristaless Related Homeobox), (OMIM #3000382) ricorrono nel 10% dei casi di Ritardo Mentale *X-linked* analizzati (Gecz J et al. 2006) e sono responsabili di patologie neurologiche differenti. Si evidenzia infatti una grande eterogeneità di fenotipi causati da alterazioni nel gene *Aristaless* (Figura 6):

- Lissencefalia con anomalie dei genitali (*XLAG*, *X-linked lissencephaly with abnormal genitalia*, OMIM #300215) (Kato M. et al. 2004);

- Sindrome di Proud (*Proud Syndrome*, OMIM #30004) ovvero Ritardo Mentale con agenesia del corpo calloso e microcefalia (Kato M. et al. 2004);
- Sindrome dell'epilessia mioclonica (*XMESID*, OMIM #300432) (Scheffer I.E. et al. 2002);
- Sindrome di West (*ISSX, X-linked infantile spasms*, OMIM #308350), che causa spasmi infantili e severo Ritardo Mentale (Kato M. et al. 2003; Stromme P. et al 2002b);
- Sindrome di Partington, (*PRTS, Partington sindrome*, OMIM #309510) con Ritardo Mentale e movimenti distonici delle mani (Stromme P. et al. 2002);
- Ritardo Mentale Non Sindromico (*NSMR*, OMIM #300382) (Bienvenu T. et al. 2002);

Le prime tre patologie sopra elencate la Lissencefalia con anomalie dei genitali, la Sindrome di Proud e la Sindrome dell'epilessia mioclonica, sono causate da mutazioni che cadono all'interno del dominio omeotico di legame al DNA oppure determinano la sintesi di proteine tronche. Tali mutazioni causano fenotipi molto severi, in cui è possibile una perdita di funzione che renda nulla la funzione di ARX, E' il caso delle mutazioni p.R332H e p.L343Q, che causano lissencefalia con anomalie dei genitali (Kato M. et al. 2004). Oppure della mutazione p.T333N che causa la sindrome di Proud, ovvero Ritardo Mentale con agenesia del corpo calloso e microcefalia (Kato M. et al. 2004); ed ancora la mutazione p.P353L che causa, invece, la sindrome dell'epilessia mioclonica (Scheffer IE et al. 2002)

(Figura 6). Le rimanenti tre patologie quali la Sindrome di West, la Sindrome di Partington ed Ritardo Mentale Non Sindromico sono causate da altri tipi di mutazioni, come le espansioni di triplette codificanti per i tratti di polialanine, mutazioni associate a patologie decisamente moderate (Friocuort G. et al. 2006). Di questo secondo gruppo di mutazioni, la più frequente è certamente la c428_451dup24 (Figura 6), denominata anche come ARXdup24. Essa è responsabile di circa il 6.6% dei casi di XLMR ed è stata identificata nel 41% delle famiglie con mutazioni di ARX (Mandel JI, Chelly J 2004; Poirier K et al. 2006; Getz J et al. 2006). Questa mutazione è dovuta ad una duplicazione di 24 bp nell'esone 2 del gene ARX. A livello proteico, essa causa l'espansione del secondo tratto di polialanine (posizione amminoacidica 144-155, poliA_II), che da un numero normale di 12 residui di alanine passa a 20 residui (Friocuort G. et al. 2006). Generalmente ARXdup24 non è mai stata trovata in associazione a fenotipi con malformazioni cerebrali e spesso determina una variabilità fenotipica elevata all'interno della stessa famiglia (Figura 6) (Turner G et al. 2002; Gronskov K et al. 2004). L'espansione di un tratto di polialanine può causare un fenotipo variabile, con alterazioni cognitive che si distribuiscono in un gradiente di gravità fenotipica che può variare da moderato a severo (Figura 6) (Friocuort G. et al. 2006).

Tale osservazione rafforza il concetto che ARX è un gene pleiotropico, che, in contesti genetici differenti e/o sotto il controllo di opportuni geni modificatori, può controllare differenti aspetti della morfogenesi e della funzioni del SNC (Laperuta C. et al. 2007). In aggiunta

alla duplicazione ARXdup24, è nota in letteratura, anche l'espansione nel primo tratto di polialanine [c.(GCG)₁₀₊₇] (Figura 6) che causa la Sindrome di West o *ISSX (X-linked infantile spasms)* (OMIM #308350), caratterizzata da spasmi infantili e severo Ritardo Mentale (Kato M. et al. 2003; Stromme P. et al 2002b). Tale mutazione è il risultato di una inserzione di 21 bp che causa l'espansione del primo tratto di polialanine (poliA_I), che da un normale numero di 16 alanine (posizione amminoacidica 100-115) passa a 23 residui (Figura 6) (Stromme P. et al 2002a; Stromme P. et al 2002b).

1.5 Il modello murino transgenico ARX KO.

Nel topo normale l'espressione di *ARX* (Figura 7) è molto precoce, *ARX* è infatti già presente allo stadio di 3 somiti durante l'embriogenesi murina. A 10 somiti, l'espressione di *ARX* è localizzata specificamente nella piastra anteriore. Successivamente, da E8.5 ad E11.5, *ARX* si distribuisce secondo un gradiente di intensità dorso-ventrale nella vescicola telencefalica (figura 7.1). In E12.5, *ARX* è localizzato nella regione telencefalica, nella zona ipotalamica anteriore, nell'area preottica (figura 7.2 A, C). Più tardi, in E14.5, la sua espressione diventa più ristretta alle zone dell'eminenza ganglionica laterale e mediale (figura 7.2 G, H, M). *ARX* partecipa infatti alla differenziazione ed alla migrazione dei neuroni GABAergici che, dalle eminenze ganglioniche laterale e mediale devono migrare attraverso lo striato verso la corteccia cerebrale. In E18.5 infatti, quando ormai gli eventi di migrazione tangenziale sono conclusi, *ARX* è presente in molti neuroni della corteccia (figura 7.2 A, B, C). Nel topo

adulto *ARX* è espresso in molti neuroni della corteccia e nei bulbi olfattori, mentre risulta poco espresso nel tessuto striato adulto (figura 7.3, A, E). *ARX* è inoltre molto espresso in una zona di migrazione vicina al ventricolo corticale che contiene principalmente cellule neurali staminali (figura 7.3, D frecce in bianco, Colombo E et al. 2004). Nel topo adulto, infine, *ARX* è inoltre espresso in ippocampo, amigdala, nucleo caudato, tutte regioni fondamentali per apprendimento e memoria (Colombo E et al. 2004).

Recentemente è stato generato un modello murino transgenico di *Arx*, al fine di indagare in dettaglio sulla funzione di *Arx* durante la formazione del cervello (Kitamura K et al. 2002).

Il topo transgenico è stato ottenuto distruggendo l'esone 2 del gene *Arx*, mediante l'inserzione del gene *LacZ* (figura 8A, a-b). Alla nascita i topi Arx^{X*Y} mostrano un cervello di dimensioni ridotte (figura 8A, e), bulbi olfattori e testicoli ridotti (figura 8A f, g, h). I topi Arx^{X*Y} muoiono dopo circa 12 ore dalla nascita (Kitamura K et al. 2002).

Il topo Arx^{X*Y} esibisce una aberrante morfologia di tutte le regioni in cui *Arx* è espresso. La piastra corticale in E19.5, è più spessa in Arx^{X*Y} (figura 8B, e) rispetto al topo normale (figura 8B, d). Studi condotti con per analizzare la proliferazione del neuroepitelio del ventricolo, mediante iniezioni di 5-bromo-2-deossiuridina (BrdU), nei topi allo stadio E12.5, hanno mostrato che nel topo Arx^{X*Y} (figura 8B, g, i) lo strato di neocortex è più sottile rispetto al topo normale (figura 8B f, h); ciò avviene però non a causa di eventi apoptotici, ma a causa di una ridotta proliferazione del neuroepitelio in tutta la zona ventricolare della neocortex del mutante

(Kitamura K et al. 2002).

Allo stadio E15.5, mediante l'utilizzo di marcatori specifici per l'eminenza ganglionica mediale (*Titf1*), (figura 8C, a, b), per l'eminenza ganglionica laterale (*Lhx6*), (figura 8C c, d) e per l'eminenza ganglionica mediale (*Dlx*), (figura 8C e, f). E' stato possibile osservare che nei topi Arx^{X*Y} si evidenzia una aberrante localizzazione degli interneuroni GABA-ergici. Si osserva inoltre una evidente perdita della migrazione degli interneuroni GABA-ergici che dall'eminenza ganglionica mediale devono migrare alla cortex. Le sole cellule che migravano dall'eminenza ganglionica alla cortex si trovavano nella zona cortico-striatale (figura 8D, g-i). Di conseguenza l'assenza di Arx da origine ad un difetto della migrazione degli interneuroni dall'eminenza ganglionica mediale devono migrare alla regione intermedia della cortex ma non altera la migrazione dall'eminenza ganglionica mediale verso le regioni corticali subventricolare attraverso l'eminenza ganglionica laterale (figura 8D, k, l). Si evidenzia inoltre, nel topo Arx^{X*Y} , una la perdita di interneuroni GABA-ergici positivi per Npy. Si tratta di neuroni specifici dell'eminenza ganglionica mediale, che devono poi differenziare nello striato (Kitamura K et al. 2002).

Tali dati relativi al modello murino transgenico di Arx dimostrano che non solo Arx/ARX è implicato nello sviluppo del cervello nei mammiferi, ma ha un ruolo fondamentale nella proliferazione, nella migrazione neuronale e nel differenziamento.

Del resto va tenuto conto del fatto che anche in organismi meno evoluti come *C.Elegans*, l'ortologo di ARX, *alr-1* è coinvolto nella funzione

chemiotattica mediante i processi di adesione cellulare (Tucker M et al. 2005). In *X.Laevis* l' ortologo di ARX, *xARX* è espresso nel telencefalo mediale e ventrale e sembrerebbe agire come un regolatore trascrizionale bi-funzionale durante lo sviluppo del cervello, capace di agire sia come attivatore che come repressore (Seufert DW et al. 2005).

1.6 Il cromosoma X nell'evoluzione dei vertebrati: diversificazione delle capacità cognitive dell'uomo.

Comparando il cromosoma X con gli autonomi, esso sembra racchiudere un numero davvero elevato di geni MRX che, quando mutati, causano difetti cognitivi (Figura 9). Sul cromosoma X sono attualmente mappati 931 geni, circa il 3.75% dell'intero genoma umano (Zechner U et al 2001). E' noto che i geni MRX sono coinvolti in meccanismi molto conservati come la regolazione della trascrizione, la stabilizzazione dell'mRNA, l'organizzazione del citoscheletro (Lahn BT, Page DC 1999). Va sottolineato in questa sede anche il fatto che i cromosomi sessuali dei vertebrati si sarebbero originati 300 milioni di anni fa da una coppia di cromosomi che non era molto diversa dagli altri cromosomi (Ohno S, 1967). Successivamente, durante l'evoluzione e la speciazione, secondo Zechner sarebbe avvenuto uno speciale processo di selezione che avrebbe forzato determinati geni presenti sull'X a sviluppare speciali funzioni implicate con l'intelligenza (Zechner U et al 2001). Il cromosoma X sarebbe quindi stato impegnato nello sviluppo dei caratteri sessuali per circa 300 milioni di anni e la selezione naturale potrebbe avere selezionato quei geni che hanno

portato allo sviluppo di capacità superiori (Zechner U et al 2001; Skuse DH 2005).

Va detto inoltre che molte patologie con *handicap* cognitivi sono spesso associate ad un decremento della fertilità (Figura 9). D'altra parte spesso si osserva l'espressione di tali geni sia nel cervello che negli organi genitali (Skuse DH 2005).

È proprio il caso del gene *ARX*, le cui mutazioni nel dominio di legame al DNA, causano la lissencefalia con anomalie dei genitali (*XLAG*, OMIM #300215) (Kato M et al. 2004). D'altra parte anche il topo $Arx^{-/Y}$, infatti, oltre a presentare aberrazioni nella migrazione e differenziazione degli interneuroni GABAergici in neocortex, mostrano anche alterazioni negli organi genitali, in particolare nei testicoli e nei tubuli semiliferi. (Figura 8A) (Kitamura K et al. 2002).

Un aspetto che è spesso possibile osservare in talune patologie legate all' X è la presenza di una inattivazione preferenziale del cromosoma X che porta la mutazione. Tale meccanismo serve a compensare l'azione patogenica della proteina mutata. Ciò accade perché opera una selezione post- inattivazione capace di generare fenotipi clinici diversi per gravità ed importanza (Miller AP, Willard HF 1998). Tale evento è selettivo dal momento che il cromosoma mutato è selezionato negativamente.

L'inattivazione del cromosoma X che porta la mutazione è un fenomeno osservato sia in associazione a riarrangiamenti cromosomici che a mutazioni puntiformi (Plenge RM et al. 1999). In particolare si osserva una

inattivazione preferenziale sbilanciata (90:10) in femmine portatrici di aberrazioni cromosomiche di vaste proporzioni, come la sindrome di Klinefelter (45,XXY) (Herzing B et al. 1997) o della trisomia del cromosoma X (47,XXX) (Ratcliffe SG et al 1986). Uno sbilanciamento preferenziale è stato osservato anche in associazione a mutazioni in geni necessari alla sopravvivenza cellulare. È quanto accade in disordini associati al cromosoma X quali la sindrome di Lesch Nyhan (Marcus S et al. 1992), l'Incontinentia Pigmenti (Parrish JE et al. 1996), la discheratosi congenita (Vulliamy TJ et al 1997), gravi casi di deficienza di G6PD (Filosa S. et al 1996). Riguardo al Ritardo Mentale associato al cromosoma X, solo in pochissime forme è stata verificata l'esistenza di una associazione tra fenotipo malattia, stato di inattivazione dell'X e mutazione (Belmont JW et al.1996, Vulliamy TJ et al. 1997, Orstavik KH et al 1998, Plenge RM et al.1999).

1.7 I Geni implicati nello sviluppo e nell'evoluzione del cervello: il metodo Ka/Ks.

La caratteristica distintiva dell'*Homo sapiens* è la dimensione del suo cervello che si accompagna ad un forte aumento delle capacità cognitive (figura 10). Recenti studi hanno messo in evidenza che i geni che controllano lo sviluppo del cervello, implicati in patologie come la microcefalia (un difetto congenito caratterizzato da una severa riduzione della dimensione del cervello), sono bersagli preferenziali della selezione naturale durante l'evoluzione (Gilbert SL et al. 2005).

Un recente studio effettuato attraverso un approccio di sistematico, ha esaminato 214 geni implicati in diversi aspetti del SNC, sono stati presi in considerazione geni come MCPH1 (Microcephalin), BDNF (Brain Derived Neurotrphic Factor), TMOD2 (Tropomodulin2) (Gilbert S.L. et al. 2005; Dorus S et al 2004).

Le sequenze codificanti di questi geni sono state comparate in quattro *taxa* di mammiferi: uomo-macaca (una scimmia del vecchio mondo) e ratto-topo. Per ognuno dei 214 geni è stato calcolato, separatamente per i primati (comparando uomo-macaca) e per i roditori (comprando topo-ratto), il rapporto tra numero di sostituzioni non sinonime e numero di sostituzioni sinonime, ovvero il Ka/Ks. (Gilbert S.L. et al. 2005).

Il metodo del il Ka/Ks è un parametro che consente di calcolare l'accelerazione evolutiva analizzando il numero di sostituzioni non sinonime ovvero di cambiamenti nucleotidici che avvengono nelle regioni codificanti di un gene e quindi modificano la sequenza aminoacidica. Le sostituzioni non sinonime (Ka) alterano le caratteristiche del prodotto proteico e pertanto sono oggetto, rispetto alle sostituzioni sinonime di selezione. Le sostituzioni sinonime (Ks), invece, avvengono in posizione degenerata del codone e pertanto non alterano il codice aminoacidico. Dal momento che quest' ultimo tipo di sostituzione non altera la proteina, esse sono considerate funzionalmente neutre. Il rapporto Ka/Ks è pertanto un parametro molto utilizzato negli studi di evoluzione molecolare e fornisce una misura del "passo" (inteso come velocità) attraverso il quale il prodotto

proteico di un gene si è evoluto. Un valore del rapporto K_a/K_s elevato indica che un gene si è evoluto rapidamente (Gilbert S.L. et al 2005).

Il risultato di questa analisi condotta applicando questo metodo a 214 geni implicati nello sviluppo del SNC, ha evidenziato un valore di K_a/K_s del 37% più elevato, per i geni del SNC, nei primati rispetto ai roditori. L'analisi è stata condotta in parallelo su geni ad espressione ubiquitaria che hanno mostrato un andamento identico tra primati e roditori. Il risultato di questa analisi suggerisce quindi un grado di evoluzione proteica, relativamente ai soli geni implicati nello sviluppo del SNC, più veloce nei primati rispetto ai roditori. Il dato è suggestivo di una evoluzione adattativa parallela allo sviluppo delle dimensioni e della complessità del SNC nei primati. Il maggiore grado di evoluzione mostrato dai geni analizzati potrebbe correlare con un aumento della complessità e della grandezza del cervello dei primati (Posthuma D et al. 2002; Gilbert S.L. et al. 2005). Geni come MCPH1 (Microcephalin, $K_a/K_s=0.833$) oppure TMOD2 (Tropomodulin, $K_a/K_s=0.333$) mostrano infatti un grado di evoluzione molecolare più rapido nei primati rispetto ai roditori: Microcephalin, $K_a/K_s=0.479$; Tropomodulin, $K_a/K_s=0.040$ (Dorus S et al 2004).

1.8 Linkage Genetico e Microsatelliti.

Nello studio dei ritardi mentali associati al cromosoma X, la strategia del linkage genetico e l'applicazione dei microsatelliti, ha permesso di identificare numerosi geni e loci malattia quali il locus MRX81 associato ad una forma di Ritardo Mentale non sindromico (Annunziata I et al. 2003).

Il linkage genetico è una strategia di mappaggio il cui principio risiede nella co-segregazione mendeliana di una regione genomica e di un tratto fenotipico qualitativo. Questo permette di stabilire il locus genico associato e di identificare il gene responsabile.

Nell'insieme la strategia da applicare deve soddisfare alcune condizioni: - sezionare ogni singolo cromosoma in intervalli di piccola dimensione; - marcare le regioni con elementi di DNA polimorfici per stabilire le fasi genetiche dei singoli cromosomi; - analizzare un numero sufficienti di meiosi per evidenziare gli eventi di ricombinazione; e infine - stabilire attraverso un test statistico la significatività delle osservazioni (Ott J, Lucek P, 1998).

La corretta applicazione della strategia implica un'analisi accurata della famiglia da analizzare, espressa come numero di meiosi, numero di individui affetti, segregazione del fenotipo, e penetranza del tratto (Miano MG et al. 2000).

Punto di partenza nel disegnare uno studio di linkage genetico è la scelta e l'impiego dei marcatori di DNA (Sheffield VC et al. 1995).

Un marcatore di DNA è una sequenza polimorfica, stabilmente presente nel DNA e trasmessa in modo mendeliano. Esso rappresenta la diversità genotipica intra-specifica degli individui per un determinato locus genico, dato che la sequenza di questo segmento genico può variare in modo stabile da un individuo all'altro all'interno di una stessa specie. Il motivo ammesso generalmente per spiegare la grande variabilità individuale delle ripetizioni all'interno dei microsatelliti è *l'errore per scivolamento*

intracromatidico (strand slippage) al momento della replicazione. Questo fenomeno è la riproduzione *in vivo* di ciò che accade talvolta *in vitro* quando l'enzima Taq polimerasi trova difficoltà a riprodurre fedelmente il numero di ripetizioni durante l'elongazione (Weber J. 1990).

Generalmente, la loro applicabilità è stabilita dall'indice di Eterozigosità, ovvero dalla *potenza* di ciascun marcatore di segregare in due loci con alleli diversi; e dalla densità di mappa ovvero quanto essi distano l'uno dall'altro.

Il grado di polimorfismo è stabilito attraverso il calcolo dell'indice di eterozigosità (H), secondo l'equazione

$$H = 1 - \sum p_i^2$$

dove H rappresenta la probabilità che un individuo, preso a caso in una popolazione di riferimento, risulti eterozigote in un dato locus genico, con frequenza allelica $[p_i]$ dell'allele "iesimo" (Ott J. 1998).

La distanza di mappa, invece, può essere espressione della frequenza di ricombinazione meiotica ed essere fornita in cM dalle mappe genetiche marcatore-marcatore ottenute dalle famiglie CEPH (Centre d'ETUDE du POLYMORPHISME HUMAIN); oppure essere una misura fisica ed espressa in Mb come annotato nelle mappe ottenute con il sequenziamento del Genoma Umano (Ott J. 1998).

I polimorfismi di DNA impiegati più frequentemente in studi di linkage genetico sono i microsatelliti o STRs (*short tandem repeats*). Essi sono caratterizzati da un motivo di base (2-6 nucleotidi) ripetuto "n" volte ($n \geq 8$) che può dar luogo ad alleli di diversa lunghezza. Tra questi i più

abbondanti nel Genoma Umano sono i dinucleotidi, distribuiti ogni 30.000 basi. I microsatelliti di tipo $(CA)_n/(GT)_n$ sono uniformemente distribuiti e particolarmente abbondanti (da 35.000 a 130.000 nell'intero genoma). Le ripetizioni di tre o quattro (trinucleotidi e tetranucleotidi) sono invece più rare, ma sono anche altamente polimorfiche, e pertanto di grande interesse e molto utili come marcatori di linkage. Per l'alta informatività e per l'ampia distribuzione che caratterizza gli STRs, essi costituiscono un sistema di indagine molecolare molto potente in studi di mappaggio genetico (Ott J. 1998). D'altra parte la loro tipizzazione è alquanto semplice e rapida, grazie soprattutto all'automatizzazione dei sistemi di rilevazione degli alleli in fluorescenza (Lindqvist A.K. et al. 1996). Infatti utilizzando oligonucleotidi nelle regioni fiancheggianti di ciascun microsatellite è possibile amplificarlo in reazioni di PCR ed analizzare i prodotti di reazione con un sistema ad alta risoluzione che permetta di discriminare variazioni di 2 sole basi.

Un'altra classe di polimorfismi, molto utile sia nel linkage genetico che in studi di associazione, sono gli SNPs (*Single Nucleotide Polymorphisms*; Altshuler D et al. 2000).

Rispetto ai microsatelliti, gli SNPs hanno una bassa velocità di mutazione tale da essere impiegati quali indicatori di elezione della storia dell'umanità (*The International SNP Map Working Group* http://snp.cshl.org/about/2001_TSC_project_overview.shtml).

Sulla base dei dati di sequenziamento è stato stimato che nel Genoma Umano esistono circa 1.5 milioni di SNPs, circa 1 SNP ogni 100-200 basi

nucleotidiche (Matisse TC et al. 2003). Questa altissima densità di mappa non è tuttavia coniugata ad un valore di eterozigosità (H) elevato. Infatti, per la tipologia degli alleli, gli SNPs hanno valori di H inferiori o al massimo pari al 50%.

Attualmente sono in commercio pannelli di SNPs (1500) che possono essere genotipizzati mediante microarrays, tecnologia che permette di analizzare migliaia di polimorfismi in modo veloce e riproducibile.

Il Metodo del Lod Score

La verifica statistica dell'associazione tra un fenotipo ereditario e una regione cromosomica è essenziale nel mappaggio genetico.

Nello studio delle patologie ereditarie monogeniche l'analisi statistica è eseguita mediante un'applicazione del test del χ^2 che valuta il rapporto di verosimiglianza tra due ipotesi opposte, *ipotesi 1*, esiste associazione; *ipotesi 2*, non esiste associazione.

Tale calcolo è eseguito in funzione della distanza genetica, espressa come frequenza di ricombinazione θ in funzione dei gameti che segregano alla meiosi ($\theta = N^\circ$ di gameti ricombinanti/ N° di gameti trasmessi).

Per loci che segregano indipendentemente, perché molto distanti o posti su cromosomi diversi, θ corrisponde ad una frequenza di ricombinazione massima di 0.5. Invece, per loci geneticamente legati o in *linkage* che segregano come aplotipi stabili nella discendenza, il valore di θ può variare tra 0 (linkage stretto) e 0.4.

Il rapporto di verosimiglianza tra l'ipotesi 1 e l'ipotesi 2 è eseguito per tutti i possibili valori di θ , utilizzando il metodo del Lod Score (Rice JP et al. 2001):

$$Z(\theta) = \log_{10} L(\theta) / L(0.5)$$

dove $Z(\theta)$ esprime il punteggio per una probabile verosimiglianza di associazione in linkage; $L(\theta)$ è l'ipotesi (1) di Linkage per $\theta \leq 0.4$; e $L(0.5)$ è l'ipotesi (2) di no Linkage per $\theta = 0.5$.

Il punteggio migliore corrisponde al valore di Lod Score massimo (Z_{maz}). Per patologie autosomiche sono accettati quali significativi di linkage i valori di $Z_{\text{maz}}(\theta) \geq 3$; per patologie associate al cromosoma X valori di $Z_{\text{maz}}(\theta) \geq 2$.

Per un valore di Z_{maz} pari a 3, si intende che vi è un'affidabilità di 1000 a 1 che nell'intervallo di θ corrispondente sia presente il locus malattia. Ne deriva che il calcolo del Lod Score permette di stabilire la distanza genetica tra i marcatori e dunque di posizionare, mediante studi a due punti e multipunti, l'intervallo in cui cercare il gene malattia. Infatti dall'applicazione del metodo dei Lod Score alle diverse strategie di genetica inversa sviluppate in questi anni, è stato possibile identificare le basi molecolari di centinaia di malattie monofattoriali (Sheffield VC et al. 1995). Questo ha in breve tempo dato origine ad altri campi di ricerca per lo sviluppo di strumenti in uso alla diagnosi molecolare e alla ricerca di terapie genetiche e/o farmacologiche.

2. SCOPO DEL LAVORO

Le anomalie delle funzioni cognitive costituiscono un gruppo eterogeneo di neuropatologie congenite. Tra esse il Ritardo Mentale (MR) rappresenta una condizione clinica complessa caratterizzata da una alterazione più o meno ampia delle capacità di apprendimento e di memoria dell'individuo (Chiurazzi P et al 2000; Miano MG et al 2006). Il 10% delle forme genetiche di Ritardo Mentale (MR) sono causate da mutazioni in geni *X-linked*. Infatti la causa più comune di deficit intellettivi e relazionali negli individui di sesso maschile, sia in forme moderate che severe, è dovuta ad alterazioni del cromosoma X (Ropers HH 2006). Il cromosoma X contribuisce pertanto con un elevato numero di geni e/o loci MRX (20-50%) allo sviluppo delle facoltà intellettive (Gecz J et al. 2000), apportando uno straordinario contributo allo sviluppo delle funzioni cognitive (Figura 9).

Il ritardo mentale *X-linked* (XLMR) è una patologia grave ad alto impatto sociale, la sua incidenza è stimata infatti pari a 1.66/1000 maschi nati vivi (Gecz J, Mulley J 2006). Recentemente di 178 XLMR, 120 sono forme di ritardo mentale di natura sindromica (S-XLMR) e 58 sono ritardi non specifici (MRX). La maggior parte di essi codifica per proteine che cooperano nello sviluppo del network neuronale e nel mantenimento della plasticità cellulare (Ramakers GJ, 2002).

Lo scopo del lavoro eseguito e presentato in questa tesi di dottorato è quello di stabilire le basi molecolari responsabili di una forma familiare di ritardo mentale ad ereditarietà *X-linked*.

Il lavoro è stato eseguito presso l'Istituto di Genetica e Biofisica del CNR di Napoli, nel laboratorio della Dott.ssa Ursini, sotto la guida della dott.ssa Miano.

Il progetto di dottorato è stato eseguito combinando metodi di genetica, biologia molecolare e di bioinformatica. Prima tappa è stata l'identificazione, mediante "*X-Chromosome Wide Screening*", del locus genetico associato alla patologia. A tale scopo sono stati selezionati accuratamente marcatori polimorfici da genotipizzare nella famiglia in analisi. In tal modo è stato possibile identificare l'intervallo di linkage genetico in associazione con la patologia.

Stabiliti i confini del nuovo aplotipo MRX87, sono state applicati metodi per la ricerca del gene candidato e della mutazione causativa della patologia. Dopo l'identificazione della mutazione e la verifica della sua segregazione nella famiglia in esame si è passati ad una analisi predittiva della funzione alterata della proteina mutata attraverso l'impiego di strumenti di bioinformatica.

L'identificazione della mutazione in un gene omeotico di grande interesse per la neurobiologia ha reso inoltre applicabili una serie di analisi evolutive allo scopo di investigare sulla la pressione selettiva esercitata sul nostro gene e conseguentemente il sul grado di evoluzione della proteina durante la speciazione.

Geni implicati nello sviluppo del Sistema Nervoso mostrano un grado di evoluzione più veloce nei Primati rispetto ai Roditori (Dorus et al. 2004).

Ciò avviene in relazione all'aumento della complessità e della grandezza del cervello dei primati durante il processo di evoluzione.

3. Materiali & Metodi

3.1 Estrazione di DNA.

L'estrazione di DNA genomico è stata eseguita secondo il metodo del “*Salting out*”, da linfociti periferici di tutti gli individui della famiglia **MRX87** disponibili allo studio (Figura 11a).

A 13 individui disponibili allo studio è stato prelevato un campione di 10 ml di sangue periferico in presenza di EDTA. Gli eritrociti sono stati sottoposti a lisi in presenza di Tris-HCl 0,01 M pH 7,5; MgCl₂ 0,005 M; Saccarosio M e TRITON 100X ed eliminati per centrifugazione. Il sedimento di linfomonociti è stato risospeso in 3 ml di un tampone A (10mM Tris-HCl pH 8,0, 2mM EDTA pH 8,0, 400mM NaCl + 200 µl di SDS 10% + 500 µg di Proteinasi K); ed è stato incubato a 37°C per almeno 24 ore. Successivamente, ciascun campione è stato sottoposto a digestione con l'enzima RNAsi (5µl di RNAsi 10mg/ml) a 37°C per 1h. Dopo la digestione, è stato aggiunto 1 ml di 6M NaCl e i campioni sono stati centrifugati a 2500 rpm per 15 minuti a 4°C. A questo punto il DNA precipitato è stato raccolto e lavato in etanolo 70%, asciugato e risospeso in 500 µl di tampone TE (0,001 Tris-HCl pH8,0, 0,001M EDTA pH 8,0) e conservato a 4°C.

3.2 Analisi di microsatelliti.

L'analisi degli STRs è stata eseguita mediante reazione di PCR su DNA genomico. Sono stati utilizzati oligonucleotidi marcatore specifici di cui l'oligo senso è modificato al 5' con un cromoforo del tipo fosforamidite (6-FAM; TET, HEX, NED).

L'impiego di oligo così modificati, ha permesso di eseguire la separazione e l'analisi dei frammenti in automatico mediante sequenziatori capillari ABI PRISM 3100 Applied Biosystem. Ciascun cromoforo emette una fluorescenza di lunghezza d'onda diversa rendendo i frammenti visibili da un sistema ottico, in dotazione al sequenziatore.

Per la reazione di amplificazione è stata utilizzata la DNA polimerasi termostabile TaqGold (Applied Biosystem) che ha la peculiarità di essere attivata a 95°C per la presenza di un gruppo chimico che la rende inattiva a temperature inferiori. La reazione di PCR è stata eseguita dopo l'attivazione della Taq polimerasi a 95°C per 10 minuti, secondo questo protocollo: 35 cicli che includono 95°C per 30 secondi, 55-65 °C per 30 secondi, 72°C per 45 secondi; e infine un ciclo di estensione finale dei singoli prodotti per 20-40 minuti, necessario per limitare la presenza di "stutters" aspecifici (Moretti TR et al. 2001).

Per ciascun microsatellite è stato utilizzato il seguente protocollo per una reazione di PCR in 10µl di volume finale:

Tampone TaqGold [10X]	1X
dNTP [2mM]	200 µM
Oligo senso 5'M [8µM]	0.8 µM
Oligo anti-senso [8µM]	0.8 µM
MgCl ₂ [25mM]	1.0 mM-2.0 mM
TaqGold [5U/µl]	10 ng

Per la maggior parte delle coppie di oligo è stato necessario mettere a punto la concentrazione ottimale di MgCl₂, allo scopo di ottenere amplificati

di qualità adatta alla risoluzione del sequenziatore automatico. Gli amplificati sono stati successivamente analizzati mediante elettroforesi su gel di agarosio 1.8% in tampone TBE 1X (Tris-Borato-EDTA).

Una volta verificata la qualità degli amplificati, si è passati all'analisi dei frammenti su sequenziatore automatico. Per fare questo, ad un aliquota di PCR è stata analizzata su sequenziatore capillare ABI PRISM 3100 Applied Biosystem e successivamente i risultati sono stati processati mediante i programmi GENESCAN e GENOTYPER2.5. Una volta stabiliti i valori dei singoli alleli, essi sono stati assegnati a ciascun individuo. Sono stati quindi ricostruiti gli aplotipo ed è stata verificata della corretta segregazione meiotica.

3.3 Sequenziatore di DNA.

Tale strumento è provvisto di: - un sistema di 96 capillari, entro i quali un particolare polimero funziona da matrice per la separazione elettroforetica di singoli frammenti; - di un laser che eccita continuamente e per tutta la lunghezza del capillare i vari cromofori legati ai prodotti di PCR ed ai frammenti dello standard; - una camera CCD (*charge-coupled device*) che registra i segnali di emissione; - un fotomoltiplicatore che converte in segnali elettrici; e - un pacchetto di softwares che opera la raccolta, il processamento, l'estrazione e la conversione in elettroferogramma. Il dato in uscita è quindi rappresentato da una serie di picchi ciascuno dei quali corrisponde ad un frammento di lunghezza diversa.

3.4 Organizzazione esone/introne di geni candidati.

I geni candidati per i loci **MRX87** sono stati selezionati integrando le informazioni ottenute dalla consultazione di diverse banche dati pubbliche quali NCBI (<http://www.ncbi.nlm.nih.gov/genome/guide/human/>); UCSC (<http://genome.ucsc.edu/>) Ensembl (<http://genome.ucsc.edu/>). Di ciascun gene è stato eseguito uno studio *in silico* comparando la sequenza genomica con la sequenza del trascritto completo. Una volta stabilite le giunzioni di *splicing* AG/GT, è stato possibile identificare l'organizzazione strutturale di ciascun gene. A questo punto, avendo a disposizione la sequenza completa di esoni ed introni sono state selezionate coppie di oligonucleotidi da impiegare nella ricerca di alterazioni del DNA.

3.5 Purificazione e sequenziamento automatico diretto di frammenti di PCR.

I prodotti di PCR sono stati sequenziati in automatico previo purificazione mediante reazione EXOSAP.

A ciascun prodotto di PCR sono stati aggiunti direttamente gli enzimi esonucleasi I e fosfatasi alcalina per eliminare e fosforilare gli eventuali nucleotidi in eccesso. Dopo una incubazione di 15 minuti a 37°C, gli enzimi sono stati inattivati a 80°C per 15 minuti. Ciascun prodotto di PCR così purificato è stato sottoposto a sequenziamento in automatico. La reazione di sequenza è stata eseguita mediante il metodo di Sanger e l'utilizzo di dideossiribonucleotidi trifosfato marcati con opportuni cromofori. Essa è stata eseguita in 10 µl di volume finale, secondo il seguente protocollo:

Oligo d'innesco	7.5 picomoli
Miscela BigDye Terminator2.PE (dNTP-*ddNTP-TaqPol-Tacs5x)	2 µl
PCR-DNA	100 µg

Le condizioni di incubazione sono:

96°C per 2 minuti; 25 o 30 cicli a 96°C per 10 minuti; 45°C per 5 secondi; e 60°C per 4 minuti.

Al termine della reazione di sequenza, i campioni sono stati purificati su membrana Sephadex G50 mediante centrifugazione a 1700 rpm per 2 minuti. A 5 µl di prodotto purificato sono stati aggiunti 10 µl di formammide deionizzata. Il campione così preparato è stato successivamente denaturato per 3 minuti a 95°C e poi sottoposto ad elettroforesi capillare.

3.6 Analisi dello stato di inattivazione del cromosoma X (Test HUMARA).

Il DNA di 5 femmine MRX87 è stato sottoposto ad analisi dello stato di inattivazione del cromosoma X mediante il saggio HUMARA (Human Androgen Receptor Assay). Questo prevede un'analisi dello stato di inattivazione attraverso una digestione con un enzima sensibile allo stato di metilazione, *HpaII*; e successivamente una reazione di PCR per la genotipizzazione allelica del locus polimorfico AR. Riguardo all'enzima *HpaII*, questo è in grado di tagliare solo il cromosoma attivo lasciando integro il cromosoma inattivo. Pertanto la successiva amplificazione del

locus AR ha luogo solo sull'allele del cromosoma X inattivo, lasciato integro dal taglio enzimatico.

Il protocollo di digestione prevede la preparazione di due aliquote di DNA, ciascuna contenente 1.5 µg di DNA: in una sono aggiunte 10 U di enzima e lo specifico tampone, nell'altra è addizionato solo il tampone di digestione. Entrambe le miscele, portate a 10 µl con H₂O bidistillata, sono state incubate a 37° per 12 ore. Trascorso tale tempo, l'enzima è stato inattivato a 95°C per 10 minuti. A questo punto i campioni sono stati sottoposti ad amplificazione mediante PCR, utilizzando una coppia di oligonucleotidi specifici per il locus AR. Di questi l'oligo senso è stato marcato con il fluorocromo fosforamide (FAM), adatto per la rilevazione al sequenziatore automatico.

La condizioni di reazione di PCR sono quelle precedentemente descritte a proposito dei microsatelliti.

Il grado di inattivazione di ciascun cromosoma X è stato calcolato applicando l'espressione :

$$\% \text{ Inattivazione} = Bd1/Bu1 / (Bd1/Bu1) + (Bd2/Bu2)$$

dove *B* è la quantità di fluorescenza emessa da ciascun allele; *d* e *u* sono gli alleli del cromosoma X digerito e non digerito, rispettivamente; *1* è l'allele che presenta un valore di intensità di fluorescenza maggiore nel campione digerito; e *2* quello con intensità di fluorescenza minore.

3.7 Programmi utilizzati.

AUTOASSEMBLER

Questo programma è stato concepito per effettuare allineamenti multipli di sequenze nucleotidiche. E' in grado di leggere *file* di testo e provenienti dal sequenziamento automatico. Inoltre permette di correggere eventuali errori nella lettura della sequenza nucleotidica.

OLIGO 4.0

Questo è un programma per la progettazione di oliogonucleotidi da impiegare in reazioni di PCR e di sequenziamento. Il programma consente di predire eventuali strutture secondarie dovute ad appaiamento di nucleotidi con formazione di “*hairpins*”.

GENESCAN e GENOTYPER 2.5

Il programma GENESCAN e GENOTYPER 2.5 consentono il processamento, l'estrapolazione e l'analisi di campioni separati in automatico mediante sequenziatori ABI PRISM 3100 Applied Biosystem. I dati vengono visualizzati mediante un elettroferogramma in cui è possibile visualizzare come picchi i singoli alleli relativi ad un determinato marcatore. In questo modo è possibile attribuire gli alleli a ciascun individuo e visualizzarli in un formato adatto per l'analisi.

VALIDAZIONE STATISTICA DEL LOD SCORE

La validazione del dato di linkage di “LOD SCORE” è avvenuta mediante il programma MLINKprogram, version 5.1from LINKAGE package software

TopPred SOFTWARE

Il SOFTWARE TopPred disponibile in rete a <http://bioweb.pasteur.fr> consente di valutare il grado di idrofobicità di una proteina e effettuare un predizione sulla predisposizione della proteina a formare α -eliche.

CLUSTALW SOFTWARE

Il software CLUSTALW, disponibile in rete all'indirizzo <http://www.ebi.ac.uk/clustalw>, consente di effettuare alberi filogenetici di proteine al fine di valutarne il grado di conservazione tra le diverse specie. Il programma consente inoltre di effettuare allineamenti multipli delle sequenze ortologhe di interesse.

KA/KS, K-ESTIMATOR 6.1 SOFTWARE

Il programma, disponibile in rete all'indirizzo <http://xblast.tamu.edu>, consente di valutare il grado di evoluzione di una proteina e la pressione selettiva esercitata su di essa durante la speciazione. Il metodo si basa sulla valutazione del rapporto tra il numero di sostituzioni non sinonime (Ka) ed il numero di sostituzioni sinonime (Ks) avvenute all'interno della regione codificante durante l'evoluzione. Un rapporto $Ka/Ks > 1$ indica una selezione positiva, un valore di $Ka/Ks = 1$ indica un'evoluzione neutra, mentre $Ka/Ks < 1$ è indice di una avvenuta pressione selettiva sul gene.

4. RISULTATI

4.1 Descrizione della famiglia MRX87.

La famiglia MRX87 è riportata in Figura 11a. La famiglia è formata da 5 individui maschi affetti da Ritardo Mentale sindromico e tre madri portatrici obbligate di normale intelligenza e clinicamente non distinguibili dalle sorelle non portatrici distribuiti in tre generazioni. La patologia segrega secondo una classica ereditarietà *X-linked*. I maschi affetti mostrano una elevata eterogeneità intra-familiare con handicap mentale e piccole alterazioni neurologiche.

La famiglia MRX87 include tra gli individui affetti 2 fratelli: l'individuo IV:13, affetto da Ritardo Mentale Sindromico moderato, piede flesso ed incontinenza urinaria; e l'individuo IV:14, con Ritardo Mentale Sindromico moderato, sialorrea, piede flesso e deficit del linguaggio; inoltre, due cugini di primo grado: l'individuo III:10, che mostra Ritardo Mentale Sindromico di grado severo, piede flesso, tonsille cerebellari discese attraverso il forame magno; ed infine uno zio, l'individuo II:5 che oltre ad essere affetto da un Ritardo Mentale Sindromico di grado severo mostra ipertonìa piramidale, movimenti bilaterali tipo Babinski, deficit dell' VII nervo cranico ed infine incontinenza urinaria (Figura 11b).

4.2 Identificazione del locus MRX87

Durante il triennio del mio dottorato ho in prima analisi, identificato la regione malattia selezionando un pannello di 18 marcatori *STRs* distribuiti

lungo l'intero cromosoma X, a circa 15-20 cM di distanza l'uno dall'altro (tabella 1).

Ho effettuato una analisi di linkage genetico analizzando i 13 componenti della famiglia MRX87, (Figura 11a) ovvero gli individui: II:3, II:4, II:5, II:6, III:4, III:5, III:6, III:7, III:9, III:10, VI:12, VI:13, VI:14. Per questo sono stati genotipizzati mediante un pannello di marcatori polimorfici STRs.

La scelta e l'identificazione di sequenze ripetute polimorfiche (STRs), è di fondamentale importanza per convertire i dati di mappa genetica in mappa fisica. Questo è stato possibile attraverso l'interrogazione di banche dati pubbliche quali *GDB* (<http://www.gdb.org/>) e *Genatlas* (<http://www.dsi.univ-paris5.fr/genatlas/>) dove sono annotati numerosi marcatori polimorfici in mappe fisiche e genetiche.

La prima fase dell'analisi ha previsto l'utilizzo di un set standard di marcatori polimorfici distribuiti ogni 10 cM, lungo l'intero cromosoma X (ABI PRISM Linkage Mapping Sets vs2, Applied Biosystem in tabella 1). I 13 individui della famiglia MRX87 sono stati genotipizzati ed i prodotti di PCR analizzati al sequenziatore automatico (ABI PRISM 3100, Applied Biosystem). L'analisi dei nuovi marcatori è stata eseguita in modo analogo a quanto descritto in precedenza nei Materiali e Metodi. Tale analisi mi ha consentito di stabilire una associazione della patologia al braccio corto del cromosoma X. In particolare alla banda citogenetica Xp22.2-Xp21.1, compresa tra i due marcatori DXS7104 ed il marcatore DXS1214. Tale intervallo cade in corrispondenza ad una regione definita "hot spot" per

mutazioni MRX (Figura 12) (Ropers H.H. et al. 2003). Successivamente, in questo intervallo, sono stati selezionati accuratamente 8 nuovi marcatori polimorfici addizionali da genotipizzare nella famiglia MRX87 (tabella 2). I nuovi marcatori sono stati selezionati da banche dati pubbliche, *GDB* (<http://www.gdb.org/>) e di *Genatlas* (<http://www.dsi.univ-paris5.fr/genatlas/>) secondo criteri di mappaggio ed indice di eterozigosità ($H \geq 60$). Essi sono i marcatori DXS7104, DXS987, DXS207, DXS8019, DXS1226, DXS8099, DXS1061, DXS1214, distribuiti ogni 1-2 cM nella regione identificata (Figura 13 e tabella 2).

Dalla genotipizzazione dei marcatori sono stati ricostruiti gli aplotipi e verificata la segregazione (Figura 14). Gli individui maschi affetti condividono il seguente aplotipo malattia, mostrato in rosso nella figura 14): DXS7104: allele 1 (167 bp), DXS987: allele 3 (218 bp), DXS207: allele 2 (234 bp), DXS8019 allele 1 (148 bp), DXS1226: allele 3 (295 bp), DXS8099: allele 1 (102 bp), DXS1061: allele 1 (221 bp), DXS1214: allele 3 (291 bp). Nell' aplotipo IV:13 esiste un doppio evento di ricombinazione tra i marcatori DXS987 ed DXS207 (limite superiore) ed i marcatori DXS1061 e DXS1214 (limite inferiore) (Figura 14 e 15).

A questo punto è stata eseguita un'analisi statistica parametrica a due punti per verificare l'associazione tra il fenotipo malattia ed i marcatori mediante il metodo del LOD SCORE. In questo modo è stato attribuito un valore positivo per 6 marcatori degli 8 analizzati, con valore di LOD SCORE massimo di 2.43 per il marcatore DXS207 (Figura 15). Tenuto conto che sono accettati come significativi per patologie X-linked valori ≥ 2 ,

ciò vuol dire che esiste una associazione statisticamente significativa tra MRX87 e la regione compresa tra DXS987 e DXS1214 (Figura 16). Paragonando infatti gli aplotipi malattia rispetto agli aplotipi normali e basandoci sugli eventi di ricombinazione è stato possibile stabilire il territorio del gene malattia che si estende dal dal marcatore DXS987 fino al marcatore DXS1214. Pertanto la regione in cui cercare il gene responsabile del fenotipo MRX87 si estende dal marcatore DXS987 fino al marcatore DXS1214, tra le bande citogenetiche Xp22.2 ed Xp21.1 (Figura 16).

4.3 Identificazione del difetto molecolare responsabile di MRX87.

La regione malattia MRX87 corrisponde ad un territorio del cromosoma X associato ad un “hot spot” per mutazioni MRX (Figura 12). A questo punto sono passata alla selezione dei geni malattia candidati. Dall’integrazione dei databases dell’*NCBI* (<http://www.ncbi.nlm.nih.gov/mapview>), di *UCSC* (<http://genome.ucsc.edu/>) e di *ENSEMBL* (<http://ensembl.org/index.html>) è stata ricostruita la mappa fisica della regione corrispondente compresa tra DXS987 (14,3 Mb sul cromosoma X) e il marcatore DXS1214 (30,9 Mb sul cromosoma X) (Figura 17). Dalla mappa fisica ricostruita, ho potuto verificare la presenza di sei geni malattia per patologie a carico del SNC, candidati per il fenotipo MRX87 (Figura 17 e 18). Si tratta dei geni *NHS*, (*Nance-Horan Syndrome*, OMIM: 302350); *CDKL5* (*Cyclin-Dependent Kinase-Like 5*, OMIM: 300203); *RSK2* (*Ribosomal Protein S6 Kinase*, OMIM: 300075); *SMS* (*Spermine Synthase*, OMIM: 300105); *IL1RAPL1* (*Interleukin 1 Receptor Accessory Protein-Like*,

OMIM: 300206); ARX, (*Aristaless-Related Homeobox*, OMIM: 300382) (Figura 17 e 18).

In particolare, il gene NHS (*Nance-Horan Syndrome*, OMIM: 302350) è responsabile della sindrome cataratto-dentale con Ritardo Mentale, appartiene alla superfamiglia delle NHS. Il gene codifica per una proteina con funzione chiave nella regolazione dello sviluppo di occhi, denti, cervello (Brooks SP et al 2004); il gene CDKL5 (*Cyclin-Dependent Kinase-Like 5*, OMIM: 300203) è invece responsabile di due diverse sindromi, una sindrome con spasmi e Ritardo Mentale ed una sindrome di Rett atipica. Il gene codifica per una proteina con attività catalitica, appartiene infatti alla super famiglia delle CDC2/CDKX ed è coinvolta nello sviluppo del SNC (Archer HL et al. 2006); il gene RSK2 (*Ribosomal Protein S6 Kinase*, OMIM: 300075) è invece responsabile della sindrome di Coffin-Lowry, una patologia che presenta Ritardo Mentale severo con dimorfismo facciale ed inoltre è anche responsabile di forme di Ritardo Mentale Non Specifico. Il gene codifica per una proteina con attività catalitica fosforilasi, una Ser/Tre chinasi con un ruolo chiave di interazione con il fattore di trascrizione CREB (Abidi F et al. 1999); il gene SMS (*Spermine Synthase*, OMIM: 300105) è invece responsabile della sindrome di Snyder-Robinson, una patologia con Ritardo Mentale e disfunzioni cerebellari. *SMS* è coinvolto nella biosintesi della spermina, in particolare nella biosintesi di amine, poliamine e della spermina dalla speridina (Cason A.L. et al. 2003); ILIRAPL1 (*Interleukin 1 Receptor Accessory Protein-Like*, OMIM: 300206) è un gene malattia associato a Ritardo Mentale X-linked Non Specifico. Il

gene codica per una proteina di membrana ed è coinvolto nel rilascio di neurotrasmettitori (Carrie A et al. 1999). Infine ARX, (*Aristaless-Related Homeobox*, OMIM: 300382), un gene malattia le cui mutazioni sono responsabili di una eterogenea varietà di fenotipi neurologici a partire da sindromi gravi, quali la lissencefalia (XLAG), la sindrome di Proud (Ritardo Mentale con agenesia del corpo calloso e microcefalia), la sindrome di West (ISSX), l'epilessia Mioclonica con deficit intellettivo, fino a quelli meno gravi come la sindrome di Partington ed il Ritardo Mentale Non Sindromico (Figura 6 e 18) (Kato M. et al. 2004). Tutti i geni sopra elencati, tranne ARX, sono riportati mutati in una bassa percentuale di casi. ARX è infatti stato stimato essere responsabile del 10% dei casi con diagnosi di XLMR tanto che, dopo *FMRI*, si ritiene esso sia il gene MRX più frequentemente mutato (Figura 18). Sulla base di queste considerazioni, ho iniziato la ricerca della mutazione responsabile del fenotipo MRX87 da questo gene. Primo punto di questa analisi è stato ricostruire la struttura esone/introne di ARX. Per questo ho analizzato i dati di sequenza disponibili in *GeneBank* (<http://www.ncbi.nlm.nih.gov/entrez>). Dopo avere ottenuto la sua struttura esoni/introni, ho disegnato un pannello di oligonucleotidi specifici per ciascun esone (figura 19a) da utilizzare nella ricerca di alterazioni del DNA. Dopo la messa a punto degli amplificati corrispondenti (Figura 19a, b), è stata eseguita una ricerca di mutazioni puntiformi. Dall'analisi degli elettroferogrammi ottenuti per ciascun frammento di PCR, è stata identificata la presenza di una duplicazione di 24 bp nell'esone 2 di ARX (ARXdup24) (Figura 20a). Si tratta di una duplicazione “*in frame*” che non

altera il codice di lettura della proteina e determina l'aggiunta di 8 residui di alanine al secondo tratto di polialanine (polyA_II) di ARX. Il polyA_II passa da una lunghezza di 12 residui a 20 residui (Figura 21). Successivamente è stata verificata la segregazione della mutazione negli individui disponibili. E' stato quindi stabilito che la mutazione *ARXdup24* è presente in tutti gli individui maschi affetti (II:5, III:5, III:10, IV:13, IV:14), nelle femmine portatrici obbligate (II:4, II:6, III:7) ed è invece assente negli individui maschi sani (Figura 20b).

Successivamente ho valutato lo stato di inattivazione preferenziale del cromosoma relativamente alla famiglia MRX87, allo scopo di verificare se la mutazione *ARXdup24* induca una selezione preferenziale del cromosoma X mutato nelle femmine portatrici obbligate MRX87, ho eseguito un TEST HUMARA sul DNA delle 6 femmine appartenenti alla famiglia MRX87 (Figura 11a). Si tratta delle femmine: II:3, II:4, II:6, III:6, III:7 e III:12. L'analisi dei dati ottenuti ha mostrato valori della percentuale di inattivazione casuali, tali da far escludere l'esistenza di una inattivazione preferenziale del cromosoma X che porta la mutazione *ARXdup24*.

4.4 Ruolo patogenetico e conseguenze funzionali di *ARXdup24* in ARX.

Per poter investigare sulle conseguenze funzionali della mutazione *ARXdup24*, ho effettuato una analisi "in silico", utilizzando il software bioinformatico *TopPred*. L'analisi è stata effettuata comparando la normale struttura secondaria della proteina ARX normale e della sua forma mutata *ARXdup24* (Figura 22). Il software, disponibile in rete a

<http://bioweb.pasteur.fr>, ha evidenziato come la forma normale contiene 4 α -eliche in corrispondenza dei 4 domini di polialanine, poliA_I, poliA_II, poliA_III e poliA_IV (Figura 22, frecce in giallo). Comparando la sequenza proteica normale con le proteine che portano le due mutazioni MRX note, ARXpoliA_16->24 nel poliA_I (tabella 3 e figura 22, frecce in rosa) ed ARXpoliA_12->20 (ARXdup24) nel poliA_II (tabella 3 e figura 22, frecce in blu), ho osservato un incremento di idrofobicità nella prima espansione con un valore di idrofobicità che passa da 1.569 ad 1.600 e nella seconda espansione con un valore che passa da 0.985 ad 1.600 (tabella 3).

Tale incremento di idrofobicità, soprattutto quello osservato nella struttura della seconda espansione potrebbe alterare la struttura secondaria della proteina che da una struttura ad α -elica potrebbe tendere ad una a β -foglietto (Figura 23). Conseguentemente, la proteina mutata potrebbe assumere una configurazione alterata, tale da comprometterne la sua funzione di fattore trascrizionale. A tale proposito infatti Nasrallah et al., nel 2004 hanno dimostrato che l'espressione *in vitro* della proteina espansa ARX PoliA_I 16->23 (MR con ISS) produce aggregati nucleari in cellule COS e aumento della morte cellulare *in vitro* (Figura 4).

4.5 Analisi Genomico-Comparativa dei domini di poli-Alanine di ARX nell'evoluzione di Primati /Non-primati.

Per investigare sul ruolo funzionale dei domini di polialanine di ARX, ho eseguito per prima cosa un allineamento multiplo (Figura 24) della proteina ARX per valutare il grado di conservazione attraverso le specie

mediante il programma bioinformatico CLUSTALW, disponibile in rete a <http://www.ebi.ac.uk/clustalw>. Ho per questo analizzato gli ortologi di ARX in *H.sapiens*, *M.mulatta*, *M.musculus*, *C.familiaris*, *R.norvegicus*, *M.domestica*, *X.tropicalis*, *F.rubripes*, *D.melanogaster* (Figura 24). Come mostrato in figura 24, i quattro tratti di polialanine, non sono presenti in tutte le specie. In particolare, il tratto poliA_I è presente in *H.sapiens*, *M.mulatta*, *M.musculus*. I tratti poliA_II e poliA_III sono presenti anche in *C.familiaris*, mentre il tratto poliA_IV è presente in tutti gli organismi tranne che in *D.melanogaster*. Dall'allineamento inoltre risulta che i domini poliA_II e poliA_III, entrambi siti per mutazioni MR, sono presenti solo a partire dai Mammiferi placentali, mentre sono assenti nei mammiferi meno evoluti. L'aggiunta di questi tratti di polialanine potrebbe essere avvenuta parallelamente all'acquisizione di funzioni cognitive superiori quali apprendimento e memoria, tipiche dei mammiferi più evoluti quali i primati.

Riguardo alla proteina ARX, come è noto, si tratta di una proteina molto conservata (Figura 25). Infatti i due domini di legame al DNA di ARX, il dominio *Aristaless* e il dominio *Homeobox* sono entrambi molto antichi e conservati nell'evoluzione (Figura 26). Per investigare ulteriormente sull'evoluzione dei tratti di polialanine durante la speciazione, ho ricostruito alberi filogenetici della proteina ARX completa e dei tre segmenti corrispondenti alle regioni contenenti i soli domini di polialanine, rispettivamente il segmento 1 contenente poliA_I, il segmento 2 contenente il poliA_II e poliA_III ed infine il segmento 3 contenente il poliA_IV (Figura 26) Riguardo all'analisi dei tratti di polialanine, da essa è

stato possibile dedurre che i quattro tratti di polialanine sono di età evolutive diverse, in particolare il segmento contenente PoliA_II/PoliA_III è più giovane rispetto al segmento contenente PoliA_I (Figura 27).

Successivamente, al fine di valutare inoltre il grado di evoluzione di *ARX*, ho valutato la pressione selettiva esercitata sul gene, applicando il metodo Ka/Ks. L'analisi è stata eseguita mediante il *SOFTWARE* disponibile in rete *K-ESTIMATOR 6.1* <http://xblast.tamu.edu>. Il metodo si basa sulla valutazione del rapporto tra il numero di sostituzioni non sinonime (Ka) ed il numero di sostituzioni sinonime (Ks) avvenute all'interno della regione codificante durante l'evoluzione. Un rapporto $Ka/Ks > 1$ indica una selezione positiva, un valore di $Ka/Ks = 1$ indica un'evoluzione neutra, mentre $Ka/Ks < 1$ è indice di una avvenuta pressione selettiva sul gene. Dalla letteratura apprendiamo che alcuni geni implicati nello sviluppo del Sistema Nervoso in particolare quelli coinvolti nel processo di encefalizzazione, mostrano un grado di evoluzione più veloce nei primati rispetto ai non primati (Dorus et al. 2004). Questo potrebbe correlare infatti con l'aumento della complessità e della grandezza del cervello dei primati durante il processo di evoluzione (Figura 10). L'analisi è stata eseguita comparando primati e non primati (roditori). I dati ottenuti suggeriscono che *ARX* ha subito una pressione selettiva maggiore nei roditori ($Ka/Ks = 0,167$) che nei primati ($Ka/Ks = 0,042$) (figura 28b). Geni invece, come *MCPHI* (*Microcephalin*), fortemente implicati nei processi di encefalizzazione, mostrano, come atteso, valori assoluti di Ka/Ks più elevati nei primati ($Ka/Ks = 0,833$) rispetto ai roditori ($Ka/Ks = 0,479$) (figura 28a). Tale

risultato di ARX è probabilmente da attribuire al precoce ruolo svolto da ARX durante lo sviluppo del SNC. Dopo avere valutato la pressione selettiva subita da ARX in valore assoluto ho voluto rappresentare graficamente l'andamento del Ka e Ks in Aristaless nei primati e roditori. A conferma del dato numerico infatti, in figura 29, si evidenziano fluttuazioni solo relativamente all'andamento del Ka e Ks nei roditori. Infatti una regione in cui è presente una maggiore pressione selettiva è presente nei soli roditori e non si evidenziano variazioni dell'andamento nei primati (Figura 29).

5. DISCUSSIONE

Durante il mio dottorato di ricerca ho avuto modo di applicare strategie di genetica, biologia molecolare e di bioinformatica al fine di identificare le basi molecolari del difetto responsabile di una forma di Ritardo Mentale X-linked e, successivamente indagare sulle conseguenze fisiologiche e funzionali della alterazione identificata. L'analisi di una famiglia affetta Ritardo Mentale X-linked è stata il punto di partenza per uno studio mirato ad approfondire il meccanismo di causa ed effetto del Ritardo Mentale MRX87. Si tratta di una famiglia italiana con Ritardo Mentale X-linked. I maschi affetti mostrano un'elevata eterogeneità intra-familiare con *handicap* mentale e piccole alterazioni neurologiche.

Prima tappa è stata l'identificazione, mediante "X-Chromosome Wide Screening", del locus genetico associato alla patologia. Si tratta del locus MRX87, identificato grazie a questo studio di linkage genetico. In una prima fase, ho effettuato un'analisi di *linkage* genetico sulla famiglia MRX87, in tal modo ho potuto ricostruire l'aplotipo malattia presente nei maschi affetti e nelle rispettive madri portatrici obbligate ed assente invece nei maschi sani. A partire dai dati di linkage ottenuti, ho potuto posizionare il locus MRX87 in Xp22, tra i marcatori DXS987 e DXS1214 (figura 16 e 17), una regione definita "*hot spot*" per mutazioni MRX (Figura 14).

In questa regione ho verificato l'esistenza di 6 geni malattia implicati in patologie del SNC. Tra essi il più frequentemente mutato in forme di XLMR è il gene *ARX* (*Aristaless-related homeobox*) (OMIM: 300382). È stato infatti stimato che mutazioni in *ARX* sono responsabili del 10% di

diagnosi di XLMR: *ARX* è, infatti, dopo *FMR1*, la causa più frequente di ritardo mentale legato al cromosoma X. La mutazione *ARXdup24* più frequentemente identificata in questo gene, può causare fenotipi molto diversi tra loro, ma tali fenotipi hanno come minimo comune denominatore il Ritardo Mentale a cui associato ad una sintomatologia sempre diversa (figura 7).

ARX possiede infatti una serie di caratteristiche che lo rendono un gene “straordinario” da diversi punti di vista. *ARX* infatti, è infatti frequentemente mutato nei ritardi mentali, è noto inoltre, che mutazioni in questo gene causano una ampia gamma di fenotipi, molto diversi tra loro soprattutto riguardo al grado di severità. Lo spettro dei fenotipi causati dalle mutazioni di *ARX* spazia da forme sindromiche e non sindromiche di MR fino ad arrivare a malattie molto gravi come la lissencefalia, l’epilessia mioclonica o l’agenesia del corpo calloso. *ARX* è inoltre coinvolto nello sviluppo della corteccia cerebrale e nel differenziamento degli interneuroni GABAergici della corteccia cerebrale (Stromme P. et al. 2002; Bienvenu T. et al. 2002).

Date le caratteristiche di *ARX*, dato il fatto che il gene è considerato un sito “*hot spot*” per mutazioni MRX, ho iniziato la mia analisi proprio da questo gene di grande interesse per lo studio delle patologie neurologiche.

Dall’analisi molecolare di *ARX* ho identificato la presenza di una duplicazione di 24 bp dell’esone 2 (*ARXdup24*) in tutti gli individui maschi affetti della famiglia MRX87 e nelle rispettive femmine portatrici obbligate (Figura 21). La duplicazione determina, a livello proteico, un’espansione nel

secondo tratto di polialanine (Figura 22). Per poter investigare sulle conseguenze funzionali di tale mutazione, ho quindi effettuato una analisi *in silico* della struttura secondaria della proteina ARX nella sua forma mutata ARXdup24 (Figura 23). La presenza dell'espansione incrementa l'idrofobicità del dominio polyA_II. Ciò potrebbe alterare la struttura secondaria di questa porzione della proteina che da una struttura ad α -elica potrebbe tendere ad una a β -foglietto. Conseguentemente, la proteina mutata potrebbe assumere una configurazione alterata tale da comprometterne la sua funzione di fattore trascrizionale. Ciò avviene in numerose patologie associate ad espansioni di triplette di DNA, in cui si osserva spesso la formazione di aggregati nucleari e l'aumento di mortalità cellulare. Infatti analogamente alle malattie da espansione di poliglutammine (poliQ), anche in alcuni casi di patologie da espansione di polialanine (poliA) è stata mostrata la formazione di aggregati nucleari e l'aumento di mortalità cellulare. Tale effetto è probabilmente dovuto, come nel caso delle malattie da poliQ ad una acquisizione di funzione della proteina mutante (Nasrallah I.M. et al. 2006). Tuttavia nei pazienti affetti da MR con mutazioni di ARXdup24 non ci sono evidenze di degenerazione neuronale.

Per investigare poi sul ruolo funzionale dei domini di polialanine, ho eseguito uno studio di genomica comparativa per esaminarne la storia evolutiva. Dall'allineamento multiplo di ortologi di ARX (*H.sapiens*, *M.mulatta*, *M.musculus*, *C.familiaris*, *R.norvegicus*, *M.domestica*, *X.tropicalis*, *F.rubripes*, *D.melanogaster*), è stato possibile comprendere che i quattro tratti di polialanine, che caratterizzano ARX, si sono aggiunti in

tempi diversi durante la speciazione (Figura 24). In particolare, poliA_I è presente in *H.sapiens*, *M.mulatta*, *M.musculus*, poliA_II e poliA_III sono presenti anche in *C.familiaris*, poliA_IV è presente in tutti gli organismi tranne che in *D.melanogaster*. Pertanto i domini poliA_II e poliA_III, siti per mutazioni MR, vengono acquisiti solo a partire dai Mammiferi placentali e sono assenti invece negli organismi meno evoluti (Figura 24).

L'aggiunta di questi tratti di polialanine potrebbe essere avvenuta parallelamente all'acquisizione di funzioni cognitive superiori quali apprendimento e memoria, tipiche dei mammiferi più evoluti quali i primati.

Al fine di poter valutare le conseguenze funzionali delle poliespansioni, sono in corso ulteriori studi funzionali mediante esperimenti di overespressione in cellule *HNSC* (*Human Neural Stem Cell*), di costrutti della proteina espansa ARX PoliA_II_{12->20} per valutare la presenza alterazioni fenotipiche quali la presenza di aggregati intranucleari e/o citoplasmatici e l'aumento della mortalità cellulare. Ciò è stato in precedenza provato, ma solo relativamente alla prima espansione, in un lavoro di Nasrallah et al., in cui si dimostra come l'espressione "in vitro" della proteina espansa ARX PolyA_I_{16->23} (MR con ISS) genera la formazione di aggregati intranucleari citotossici in cellule COS e aumento della morte cellulare (Figura 4) (Nasrallah I. M. et al 2004).

Va poi sottolineato che le informazioni riguardo ai geni regolati a valle da ARX sono molto limitate. Per questa ragione sono attualmente in corso, nel laboratorio dove ho svolto la tesi di dottorato, studi di espressione genica mediante la tecnologia dei microarray al fine di comprendere meglio

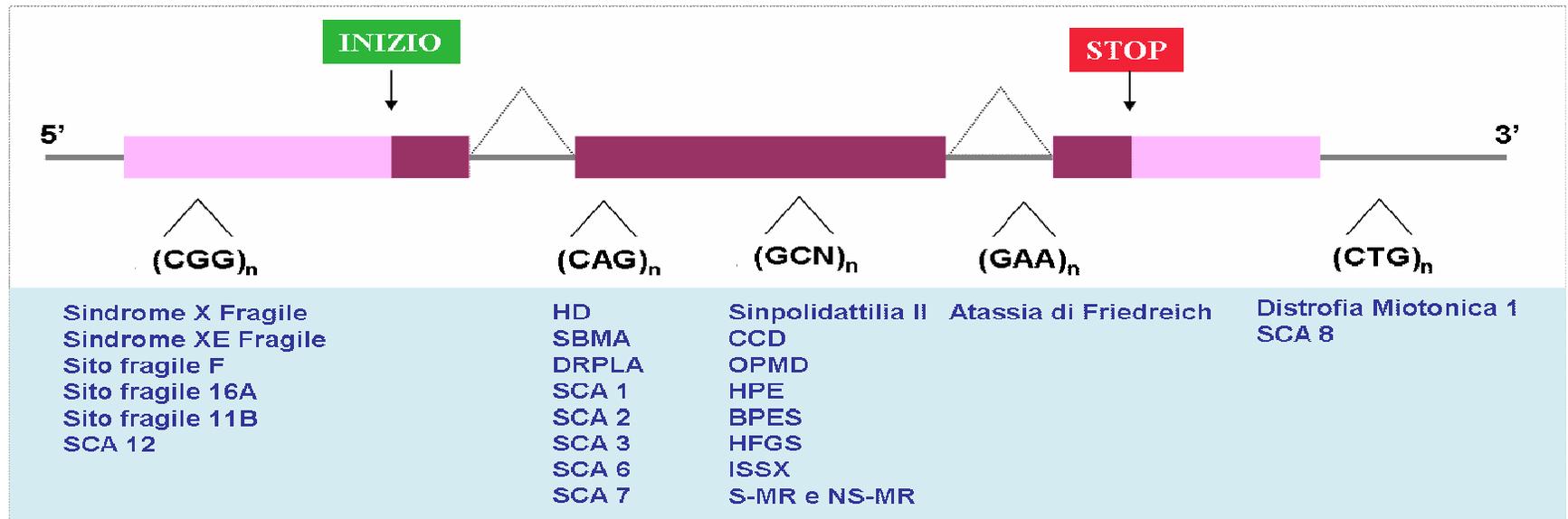
il ruolo e le interazioni di *ARX* durante la neurogenesi e la differenziazione neuronale.

Durante il processo di evoluzione si assiste ad un incremento enorme del volume, della grandezza e soprattutto della complessità del cervello soprattutto nei primati (Figura 10) ed inoltre differenze nel volume della corteccia sono associate a differenze nella complessità del circuito corticale.

Alla luce di queste considerazioni, un aspetto importante è stato capire se *ARX* è un gene importante per i processi di encefalizzazione, visto che è noto determinare lo sviluppo della corteccia cerebrale. Inoltre *Aristaless* appartiene alla famiglia genica delle *Q50 Paired-like (Prd-like)*, una famiglia molto primitiva di fattori di trascrizione con un ruolo chiave nelle prime fasi dell'embriogenesi dei vertebrati, nello sviluppo della testa dell'animale e del sistema nervoso centrale e periferico, infatti il dominio *Aristaless* è già presente in organismi poco evoluti come *C.elegans* (Figura 26) (Galliot B., Miller D. 2000; Friocourt G. et al. 2006).

Ho voluto, quindi, stimare il grado di evoluzione di *ARX*, valutando la pressione selettiva esercitata sul gene, applicando il metodo Ka/Ks. Geni implicati in processi come l'encefalizzazione, quali *Microcefalina*, mostrano un grado di evoluzione più veloce nei primati rispetto ai non primati (figura 28 e 29) (Dorus et al. 2004). L'analisi eseguita comparando nei primati e nei non primati (roditori) evidenzia che *ARX* ha subito una pressione selettiva maggiore nei roditori che nei primati. Ciò sarebbe in accordo con il precoce ruolo svolto da *ARX* durante lo sviluppo del SNC.

6. *FIGURE*



(CAG)_n/(GAA)_n/(CTG)_n/(CGG)_n = codoni per la glutammilina (Q).
 (GCC)_n/(GCG)_n/(GCA)_n/(GCT)_n e (GCN)_n = codoni per l'alanina (A).

Figura 1. Localizzazione nel gene dei tratti ripetuti nelle malattie da espansione di triplette.

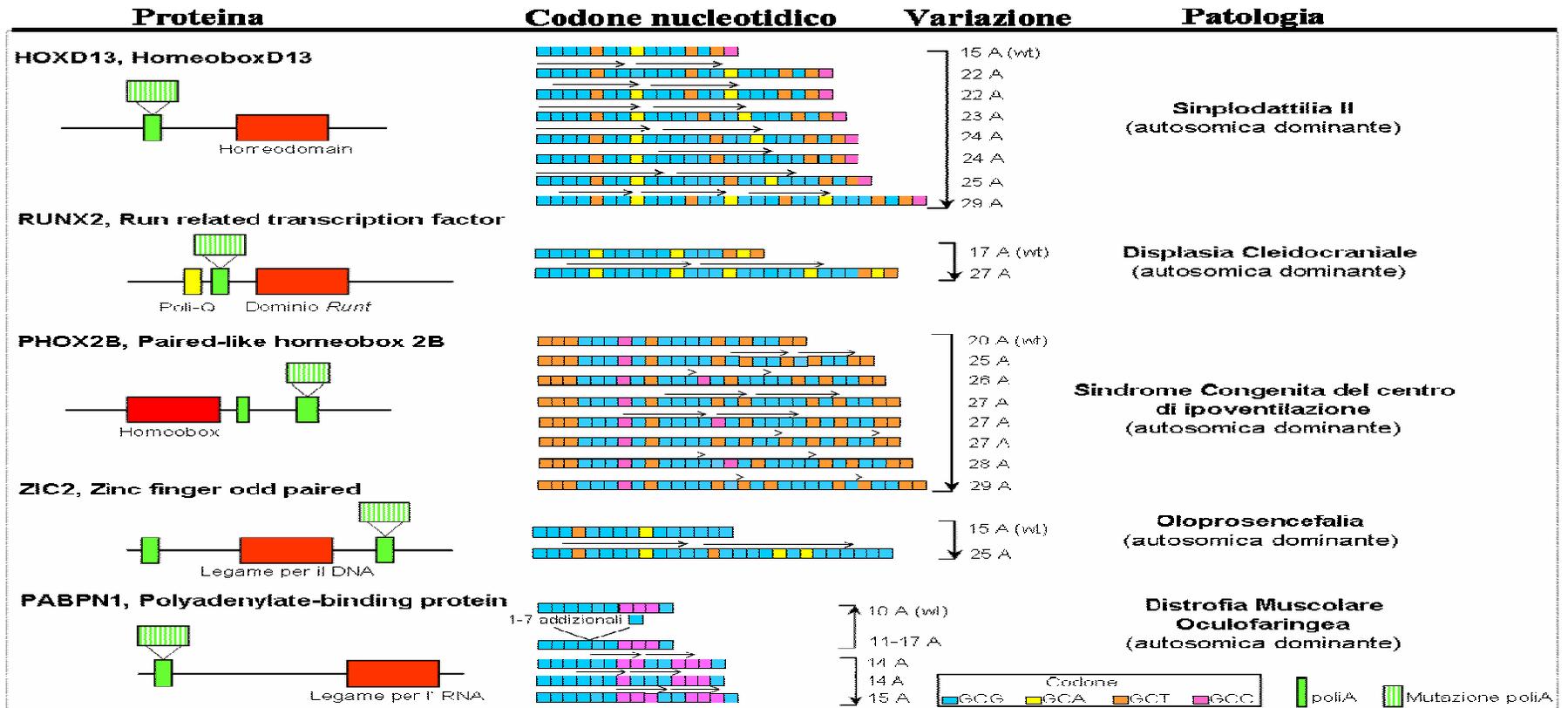


Figura 2. Geni e malattie associati a espansioni di polialanine.

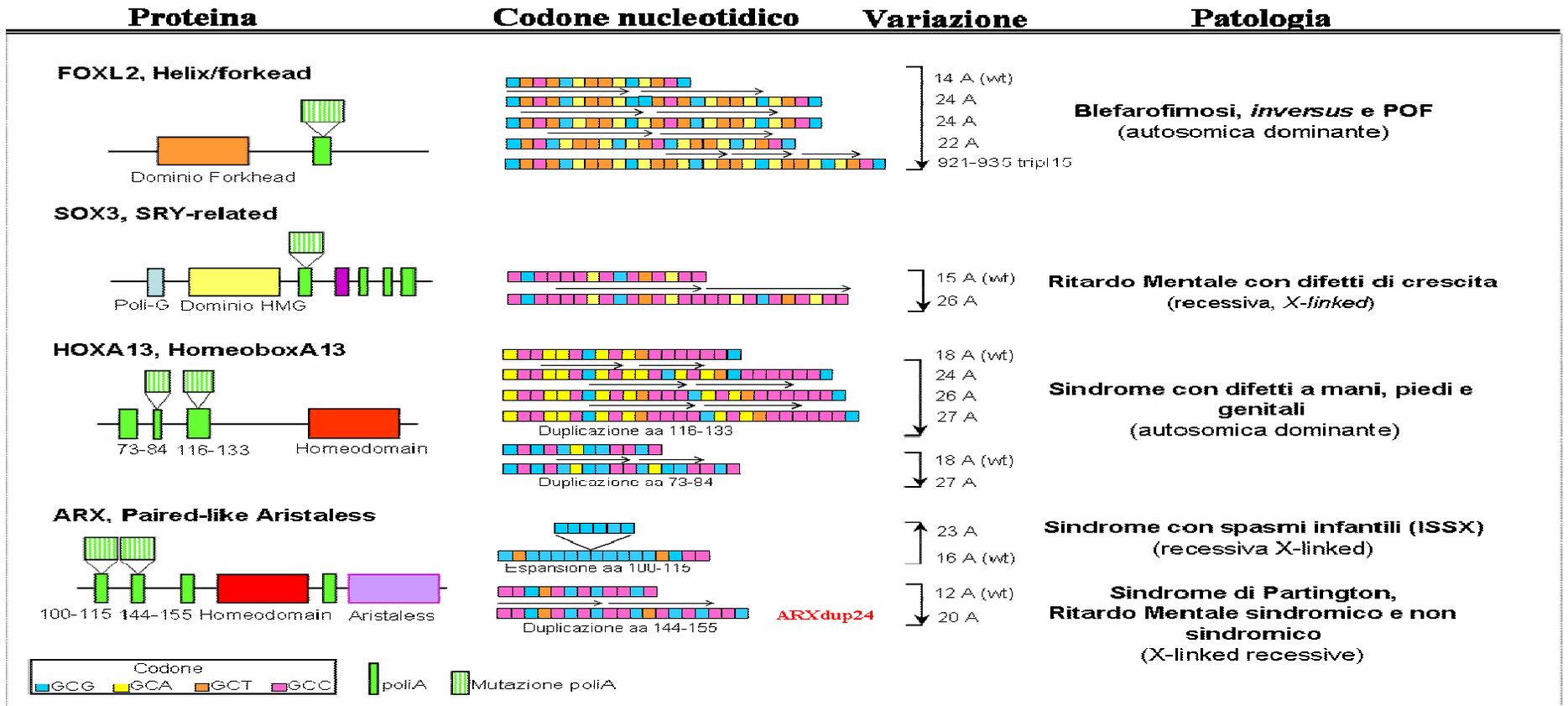
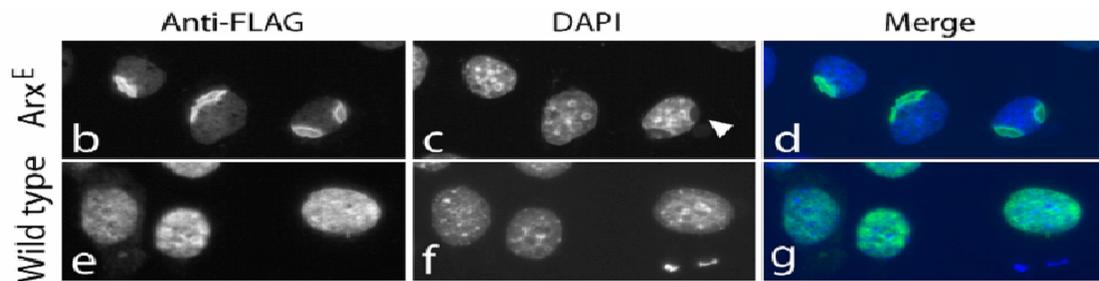


Figura 3. Geni e malattie associate a espansioni di polialanine



• **Espressione del costrutto poliA_I espanso di ARX** formazione di aggregati intranucleari

• **Espressione del wild type di ARX** diffusa espressione nucleare di ARX

L'espansione PoliA_I di Arx aumenta la morte cellulare in vitro

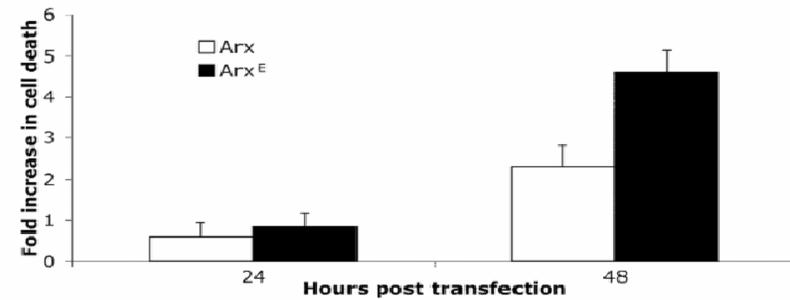


Figura 4. L'espressione in vitro della proteina espansa ARX PoliA_I 16->23 (MR con ISS) produce aggregati nucleari in cellule COS.

(Nasrallah I. M. et al. 2004)

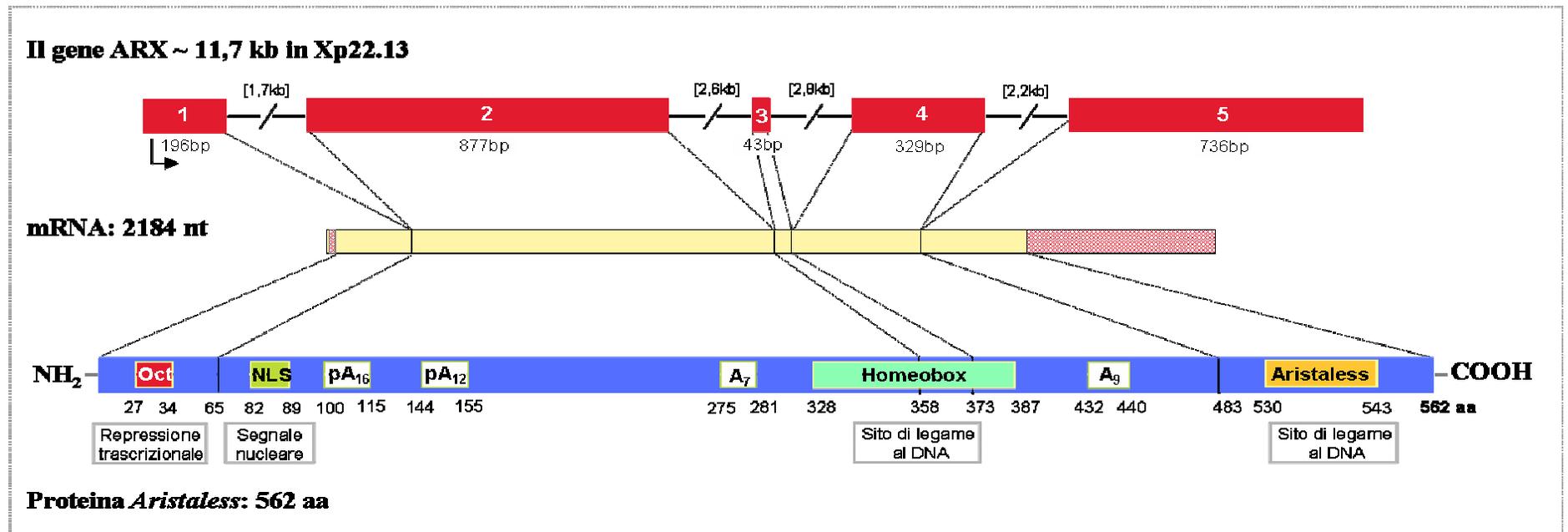


Figura 5. Il gene ARX: struttura genomica, trascritto e prodotto proteico.

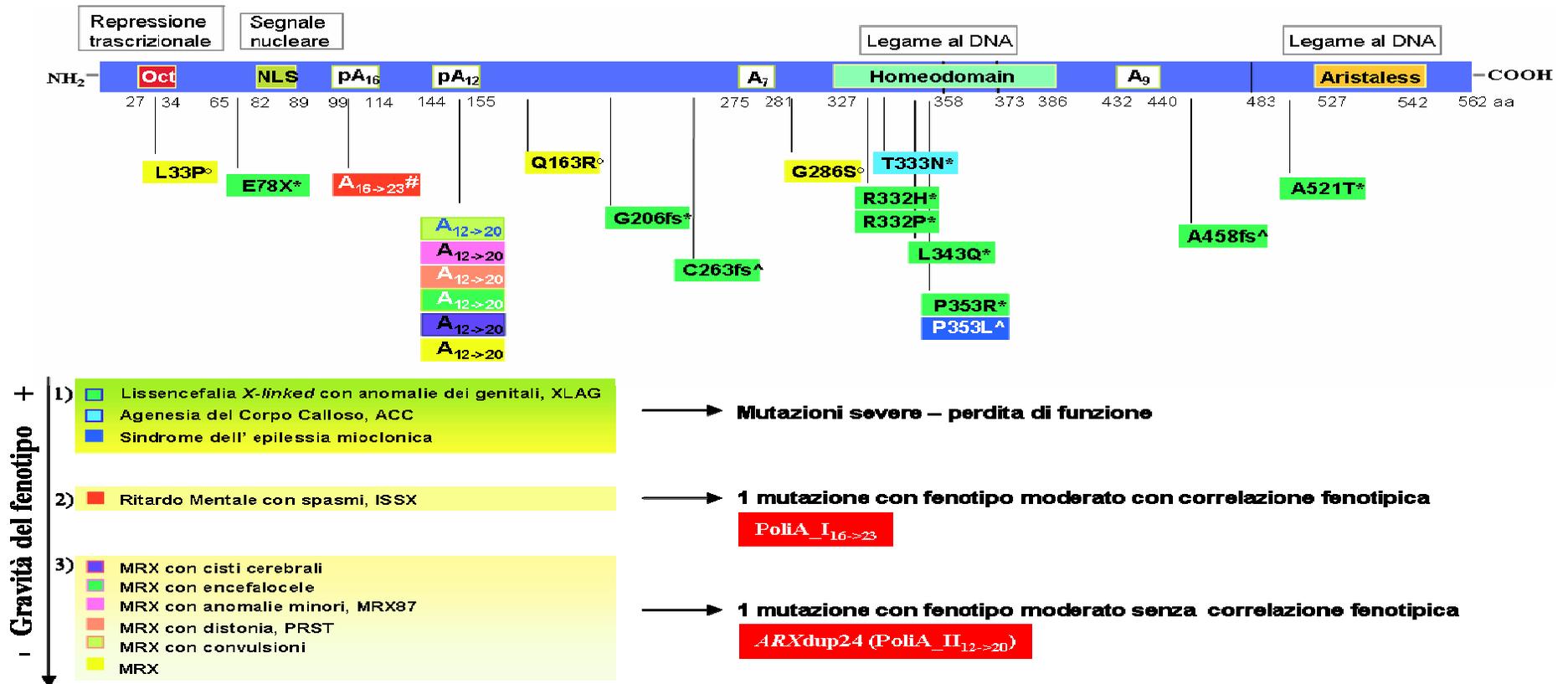


Figura 6. Spettro di mutazioni in ARX: eterogeneità di fenotipi, effetto pleiotropico e presenza di un gradiente di severità fenotipica.

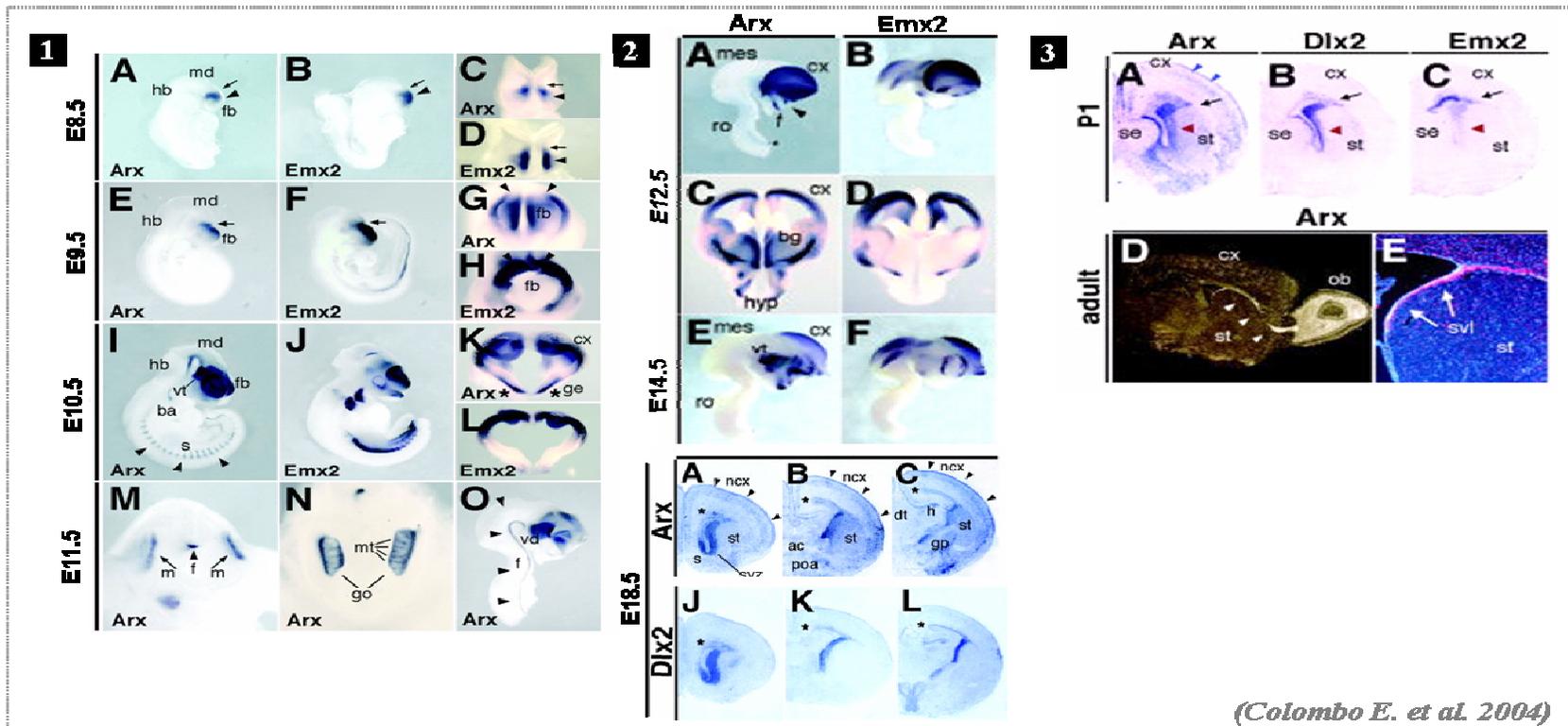


Figura 7. Espressione di Arx in topo negli stadi embrionali da E8.5 a E18.5 (1, 2) ed nel topo adulto (3). L'espressione di Arx (dorso-ventrale) è comparata con quella del marcatore caudo-dorsale Emx2 e con Dlx2 che mostra espressione simile a quella di Arx.

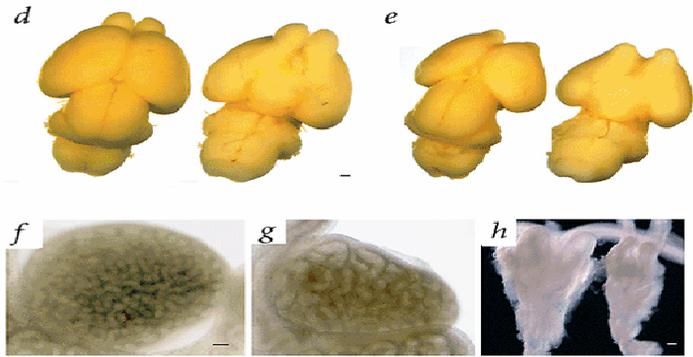
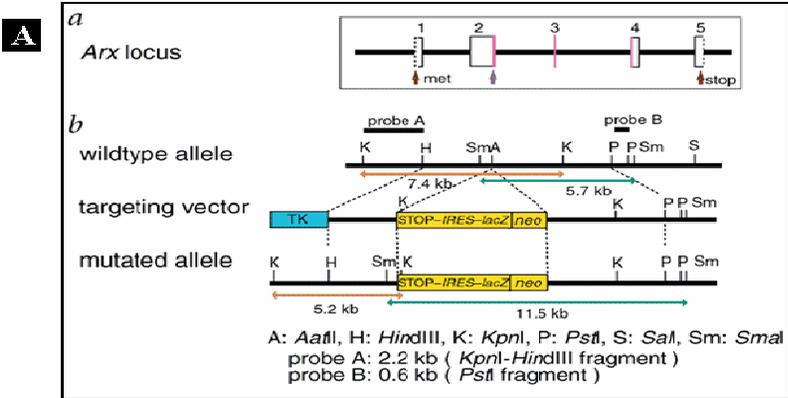


Figura 8A. In *a*, *b* struttura genomica e ablazione del gene *Arx* nel topo *Arx^{X^Y}*; *d*, *e*, visione dorsale e ventrale del cervello nel topo wt e di *Arx^{X^Y}*; *f-h*, i testicoli del wt (*f*) sono più grandi rispetto ad *Arx^{X^Y}* (*g*); le vescicole seminali sono ipoplastiche nel topo mutante rispetto al topo normale.*h*

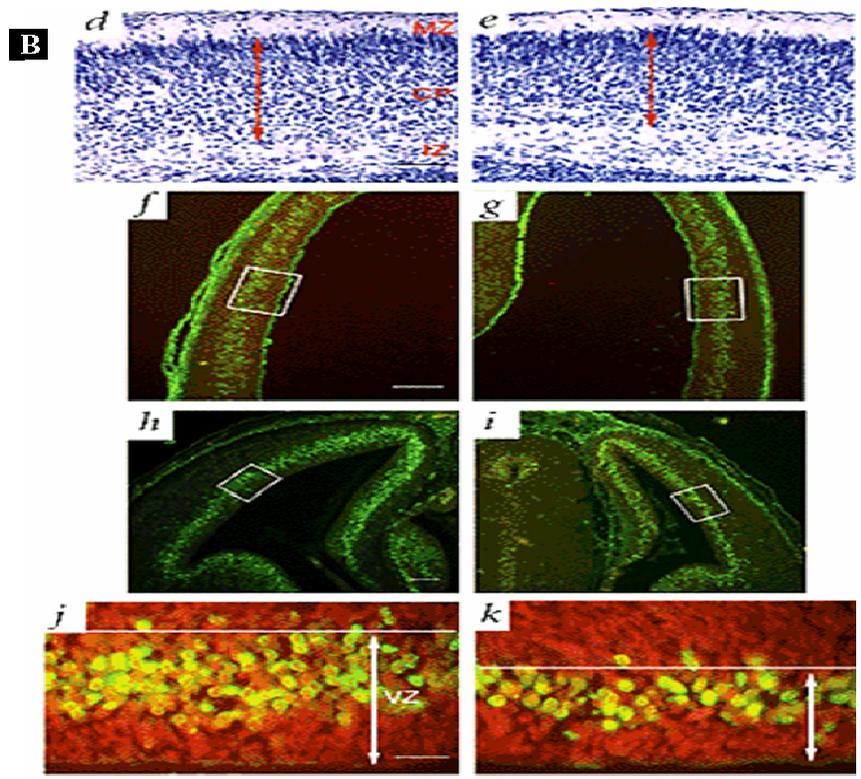


Figura 8B. Aberrazioni della proliferazione e del movimento delle cellule neuroepitettali in neocortex. *D*, *e*: lo spessore della neocortex del topo normale è più ampio rispetto al topo *Arx^{X^Y}*. *F-I*: cellule neuroepitettali trattate con BrdU (verde) del topo normale (*f*; *h*) e del mutante (*g*; *i*). *J*, *k*, cellule del ventricolo trattate con BrdU (cellule) e Propidio Ioduro (neuroepitello) del mutante (*k*) sono scarse rispetto al topo normale (*j*).

(Kitamura K et al. 2002)

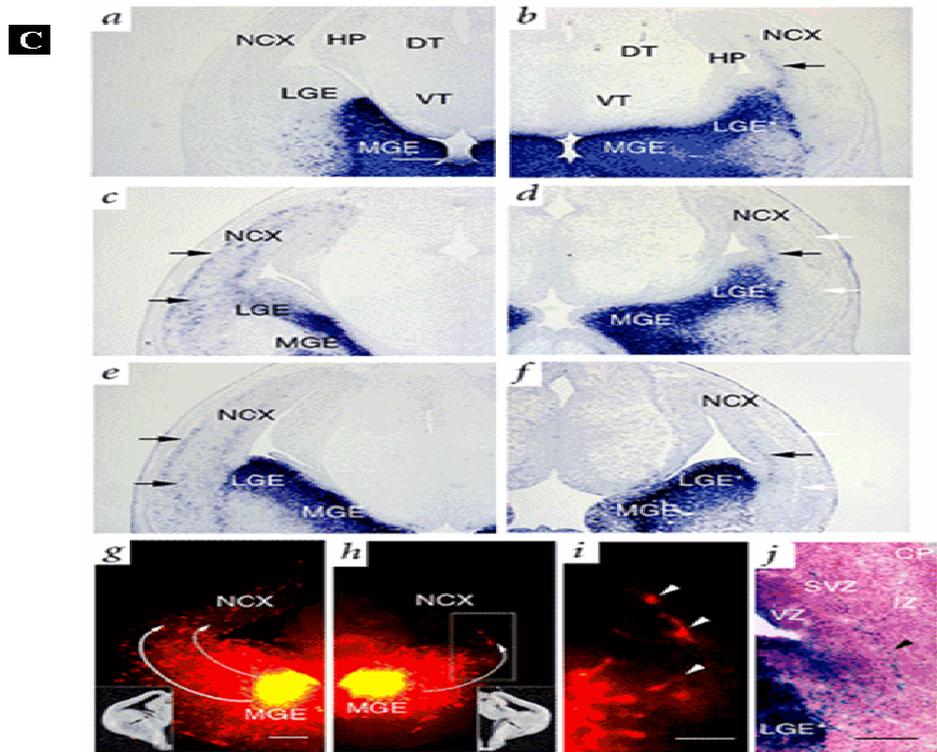


Figura 8C. Espansione dell'eminenza ganglionica mediale e laterale (MGE ed LGE) nel topo Arx^{XY} in b, d, f; rispetto al topo normale (a, c, e). Aberrante localizzazione degli interneuroni GABAergici in Arx^{XY} nell'eminenza ganglionica mediale (MGE).

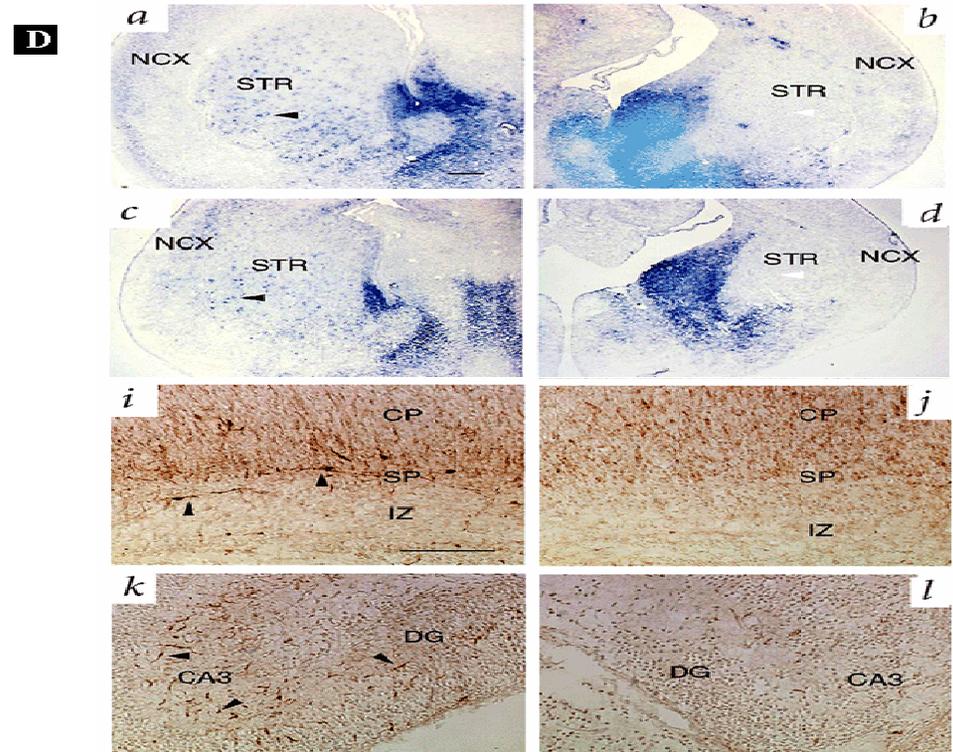
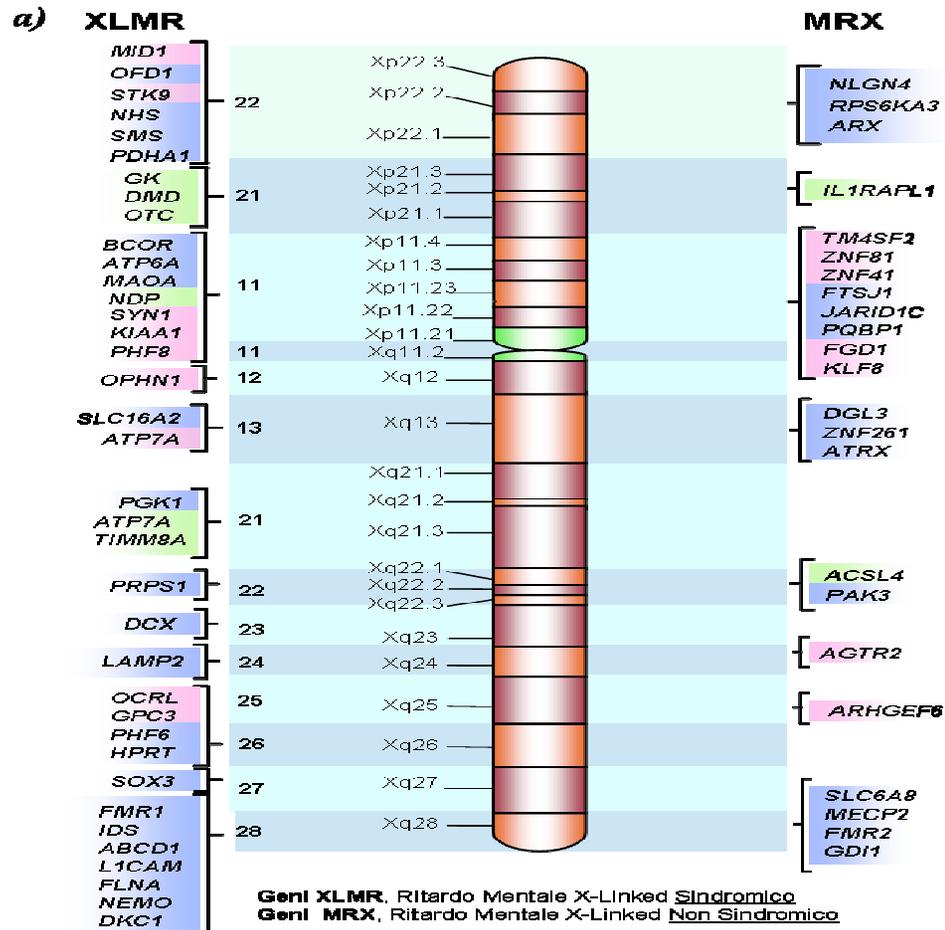


Figura 8D. Anomala espansione dell' MGE nel topo Arx^{XY} (b,d). Aberrante migrazione e localizzazione degli interneuroni GABAergici in Arx^{XY} nella piastra corticale.

(Kitamura K et al 2002)

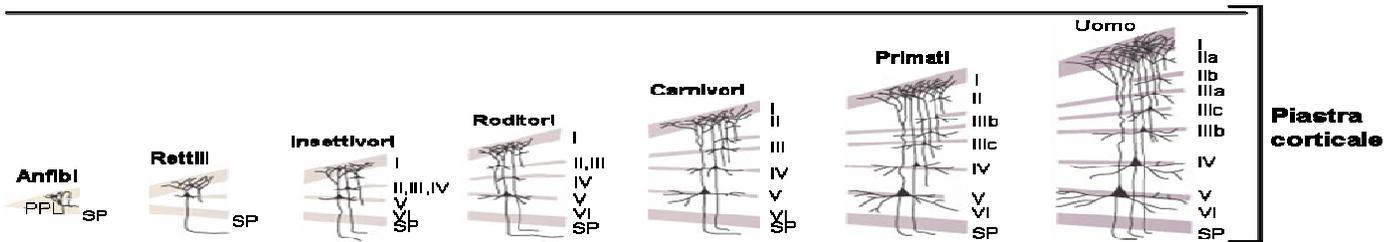
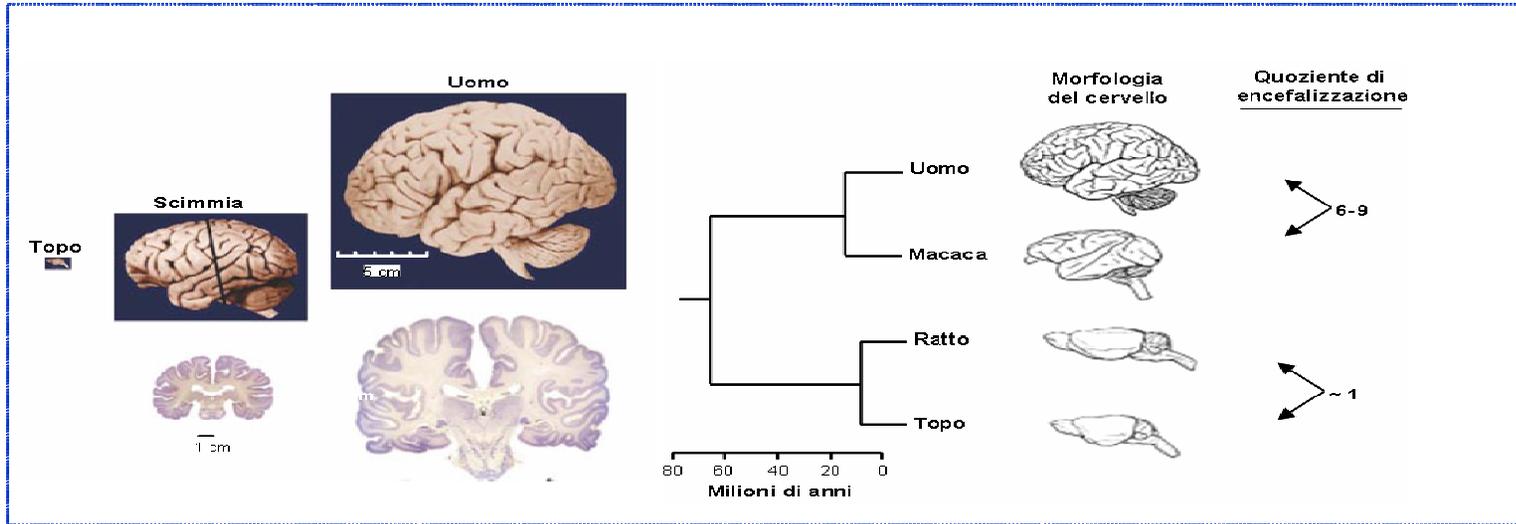


b)

Difetto Fenotipico	N° e F del fenotipi		% X loci : % A loci
	Cromosoma X (%)	Autosomi (%)	
Polidattilia	5 (1.21)	35 (0.55)	2.20
Palatoschisi	10 (2.34)	56 (0.88)	2.76
Displasia Facciale	7 (1.70)	45 (0.71)	2.39
Displasia Scheletrica	11 (2.67)	76 (1.19)	2.24
Ritardo della crescita	13 (3.16)	94 (1.48)	2.14
Ritardo Mentale	103 (25.0)	222 (3.49)	7.16
Ritardo Mentale e (Ipogonadismo/macroorchidismo/criptorchidismo)	17 (4.13)	27 (0.42)	9.83

Skuse, Hum. Mol. Gen. 2005

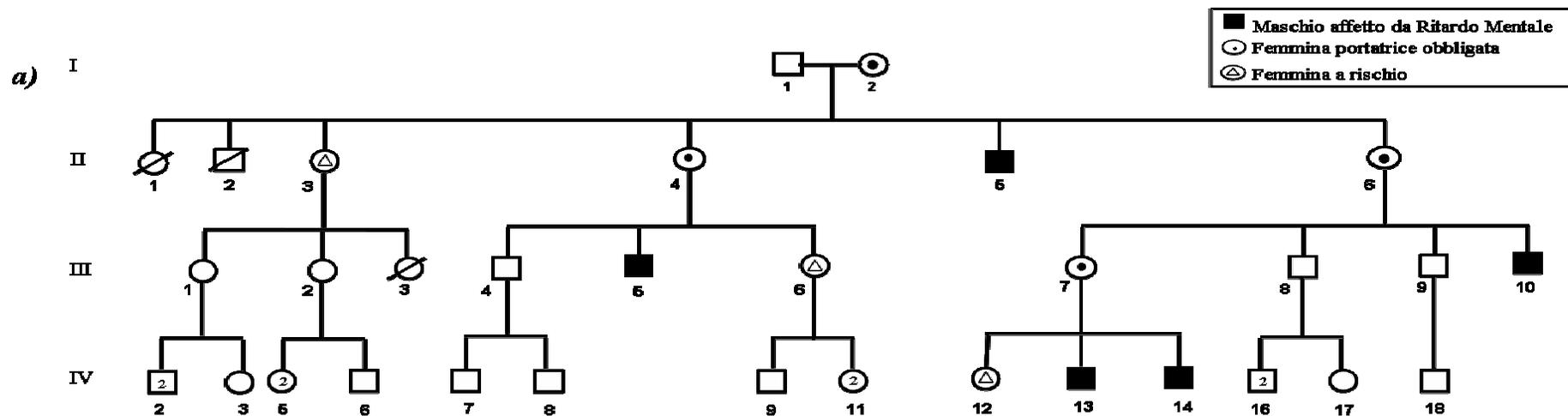
Figura 9. a) Elevata Densità sul cromosoma X di geni associati allo sviluppo delle capacità cognitive; b) numero e frequenza relativa a difetti associati al cromosoma X ed autosomi.



Le differenze nel volume della corteccia sono associate a differenze nella complessità del circuito corticale.

Da Sean Hill & Walsh. *Molecular insight into human brain evolution. Nature, 1: 64-67, 2005*

Figura 10. Evoluzione della corteccia nei mammiferi



b) Difetti neurologici associati:

- Ipertonia Piramidale; Movimenti Bilaterali tipo Babinski - Deficit dell' VII nervo cranico - **II:5**
- Ipoacusia Neurosensoriale - Demenza-
- Tonsille Cerebellari discese attraverso il forame magno - **III:10**
- Ampi spazi subaracnoidei -
- Incontinenza Urinaria - **II:5 and IV13**
- Difetto del Linguaggio **IV14**

Figura 11. a) La famiglia MRX87, Ritardo Mentale Sindromico legato al cromosoma X; b) I difetti neurologici associati.

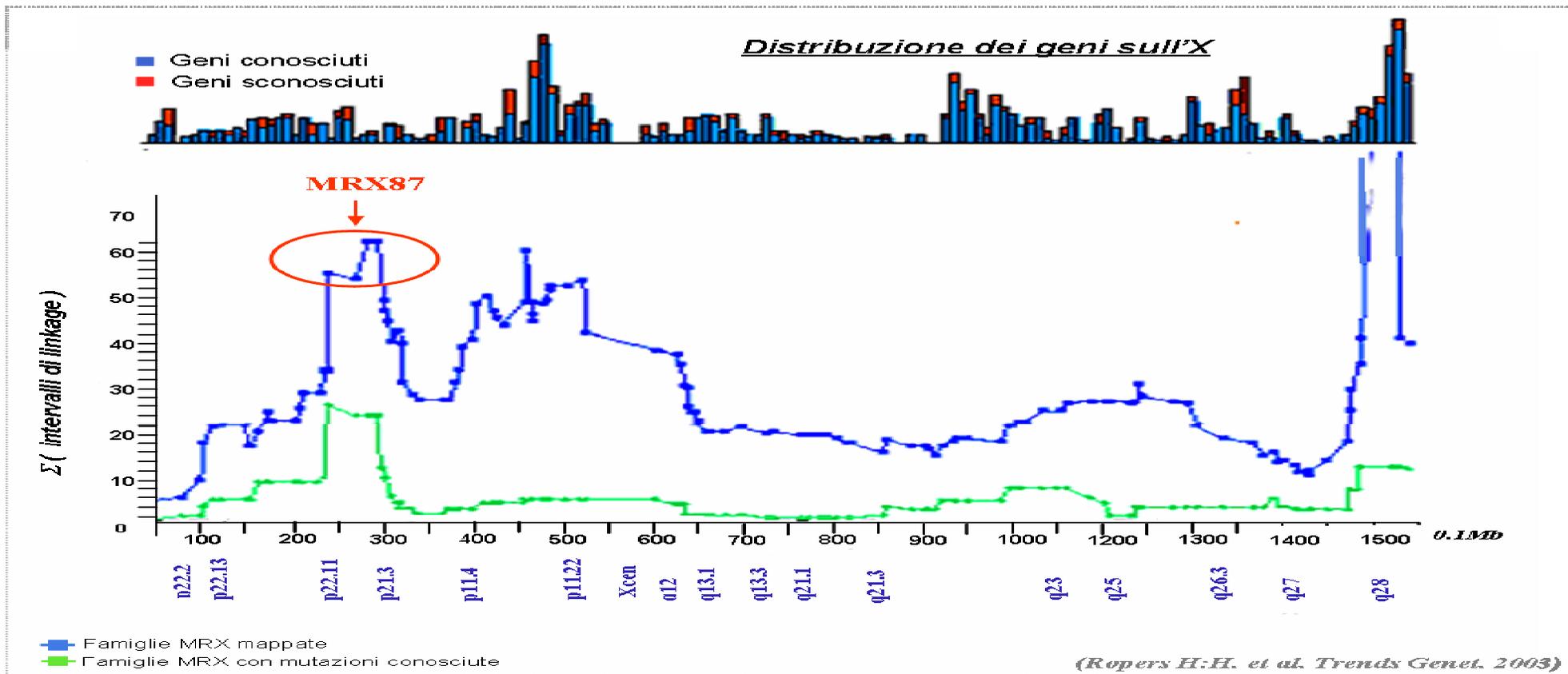


Figura 12. Ritardo Mentale X-linked: regioni "hot spots" - Xp22 - Xp11 - Xq28 -

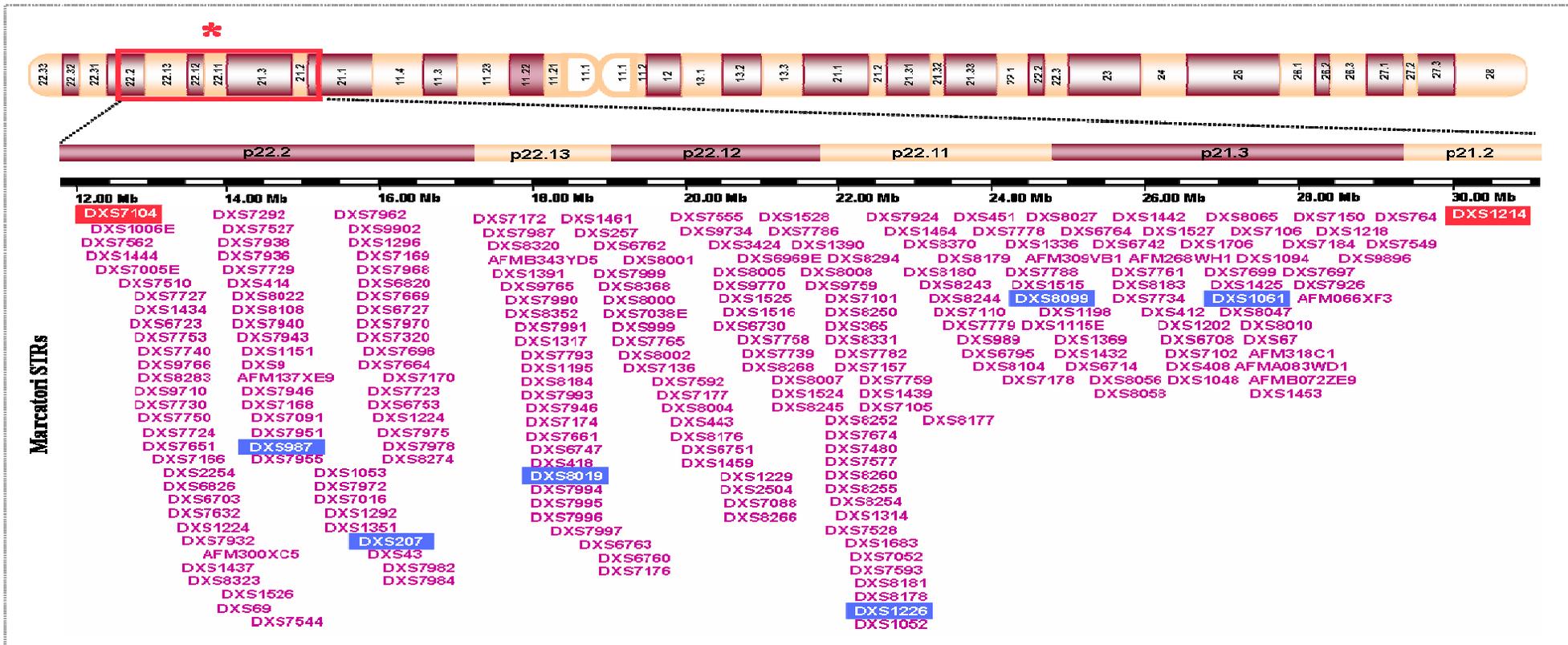


Figura 13. Identificazione della regione malattia in Xp22.2-Xp21.1
Genome-Browser Santa Cruz, Human Assembly Mar. 2006 (<http://genome.ucsc.edu>)

Marker	Sequenze Oligo
DXS1227	F 5'-AGAGGTCCGAGTCTTCCAC -3' R 5'-ATAAGGGTTTACTCCCCCAA -3'
DXS990	F 5'-AGCTATATGACCAGTACAAACATAC -3' R 5'-GACAGAAGGGACATCAACTC -3'
DXS986	F 5'-CCTAAGTGCTCATCATCCCA -3' R- 5'-AGCTCAATCCAAGTTGCTGA -3'
DXS987	F 5'GTTGAGATAATGAGGCCAGT-3' R 5' TAAAAGCCTGGTTCTTCTAAA-3'
DXS993	F 5'-GGATCCTGTTTACAGCCTGT -3' R 5'-CTACAGAGCAGTTCCTGGG -3'
DXS1073	F 5'-GGCTGACTCCAGAGGC -3' R 5'-CCGAGTTATTACAAAGAAGCAC-3
DXS8091	F 5'-CACATTCAGGTTCCACAGG -3' R 5'-CAAGATCCAGGCCAAAAGTC-3'
DXS1106	F 5'-TATGAGAACTCCCTAAACAAA -3' R 5'-TGATGCACCAAATACCA -3'
DXS1047	F 5'- CCGGCTACAAGTGATGTCTA 3' R 5'-CCTAGGTAACATAGTGAGACCTTG -3'

Marker	Sequenze Oligo
DXS1001	F 5'-CCGGCTACAAGTGATGTCTA -3' R 5'-CCTAGGTAACATAGTGAGACCTTG -3'
DXS1068	F 5'-TCCTCTAAAGCATAGGGTCCA -3' R 5'-CCCATCTGAGAACACGCTG -3'
DXS1214	F 5'-TAGAACCCAAATGACAACCA -3' R 5'-AAGATAGCAGGCAACAATAAGA -3'
DXS8055	F 5'-AGTGGGTCACTCTAGTCATCAT -3' R 5'-TCAGGTTTCTGTGTGGACAT -3'
DXS8051	F 5'-ACCAGAAATGAGCGATTATTG -3' R 5'-TTTTTGAAC TAAGAACCTGGAG -3'
DXS8043	F-5'-AGTTCTCAGAAACATTTGGTTAGGC -3' R-5'-AATTATTGGCAAAGAGTACAGGCAG -3'
DXS1060	F 5'-AAAATTAACATCTACTAGGCAGA -3' R 5'-TACTGCGAACGGTGAGAG -3'
DXS1226	F 5'-CTAAACCCATCTGNCCTC -3' R 5'-TTTCCAGCAACTACCTTTCAT -3'
DXS991	F 5'-ACTTCAACCACAGAAGCCTC -3' R 5'-ATCATTTGAGCCAATTCTCC -3'

**Tabella 1. I 18 marcatori polimorfici standard utilizzati nel 1° STEP di genotipizzazione della famiglia MRX87.
(ABI PRISM Linkage Mapping Sets vs2, Applied Biosystem)**

Marker	Sequenze Oligo
DXS7104	F 5'-TCAGACAAGTAGCCTTATGTG-3' R 5'-AGTCTCATCGCAACCTTT-3'
DXS987	F 5'GTTGAGATAATGAGGCCAGT-3' R 5' TAAAAGCCTGGTTCTTCTAAA-3'
DXS207	F 5'-CTCTTGTCAGATTTTCAGGTCC -3' R- 5'-GGAATACTTGTAAGTTCATTGGGA-3'
DXS8019	F- 5'-TTCATTAAAAGCCGTCCTTTG -3' R-5'-TGTTGAGTTTCCTCACAGC-3'
DXS1226	F-5'-CTAAACCCATCTGNCCTC-3' R-5'-TTTCCAGCAACTACCTTTCAT-3'
DXS8099	F-5'-AGCTGGTTTTGTGATTCTGC-3' R-5'-GGCCTTGAGATGTAGCCA-3
DXS1061	F-5'-ATTTGTAAACCCTTCTGATTG-3' R-5'-ACGATAGACCATTATGATGTGT-3'
DXS1214	F-5'-TAGAACCCAAATGACAACCA-3' F-5'-AAGATAGCAGGCAACAATAAGA-3'

Tabella 2. Marcatori polimorfici addizionali utilizzati nel 2° STEP di genotipizzazione della famiglia MRX87.
(<http://www.gdb.org/>)

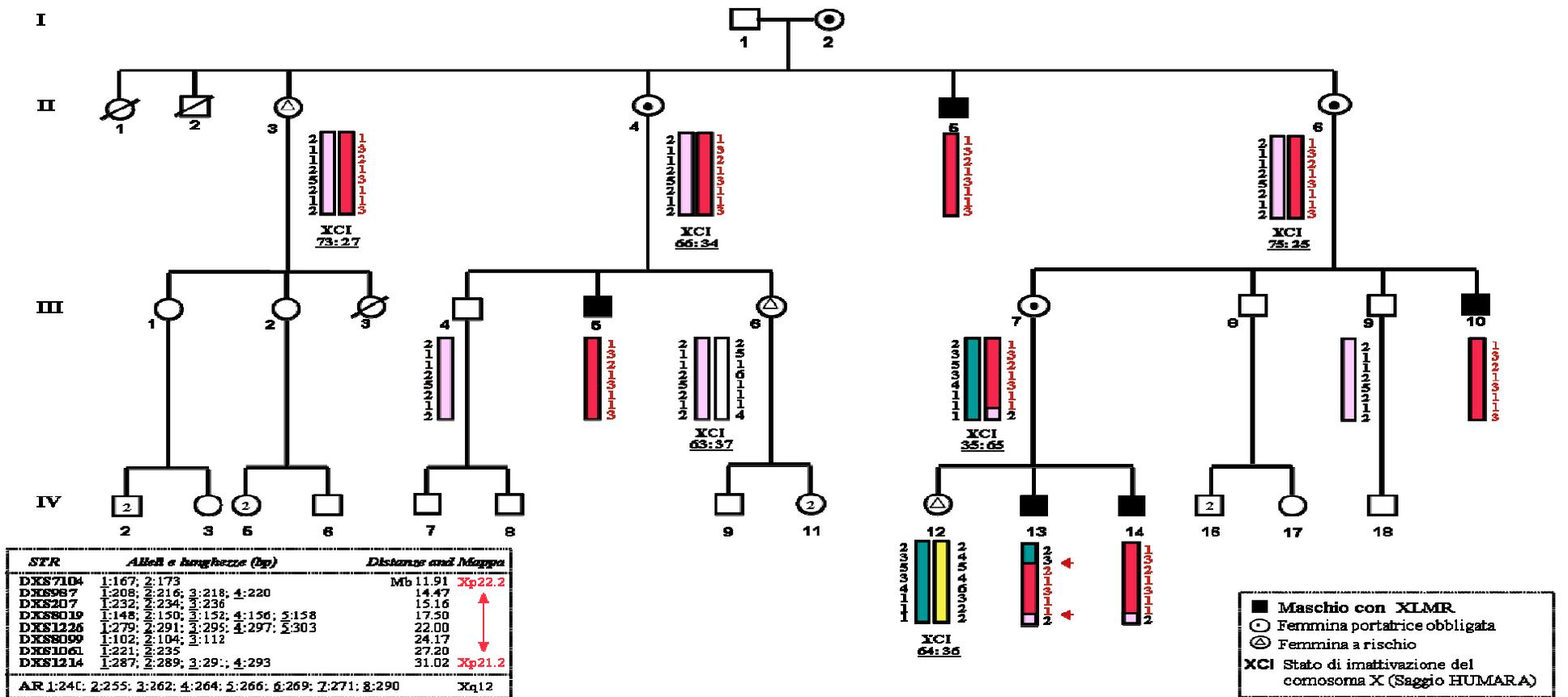


Figura 14. Ridefinizione del locus MERX87, Aplotipi e Segregazione.

(Laperuta C et al. 2007)

<i>Marker</i>	<i>0.0*</i>	<i>0.01</i>	<i>0.05</i>	<i>0.1</i>	<i>0.2</i>	<i>0.3</i>	<i>0.4</i>
DXS7104	-2.94	0.02	0.59	0.73	0.69	0.51	0.27
DXS987	1.75	1.72	1.59	1.43	1.09	0.73	0.36
DXS207	2.43	2.38	2.22	2.00	1.54	1.04	0.52
DXS8019	2.35	2.31	2.15	1.94	1.50	1.02	0.52
DXS1226	2.35	2.31	2.15	1.94	1.50	1.02	0.52
DXS8099	1.75	1.72	1.59	1.43	1.09	0.73	0.36
DXS1061	0.35	0.34	0.30	0.26	0.16	0.08	0.02
DXS1214	-7.31	0.31	0.87	0.99	0.90	0.66	0.35

* Frazione di Ricombinazione
MLINKprogram, version 5.1from LINKAGE package software

Figura 15. Validazione del dato di linkage mediante metodo del "LOD SCORE".

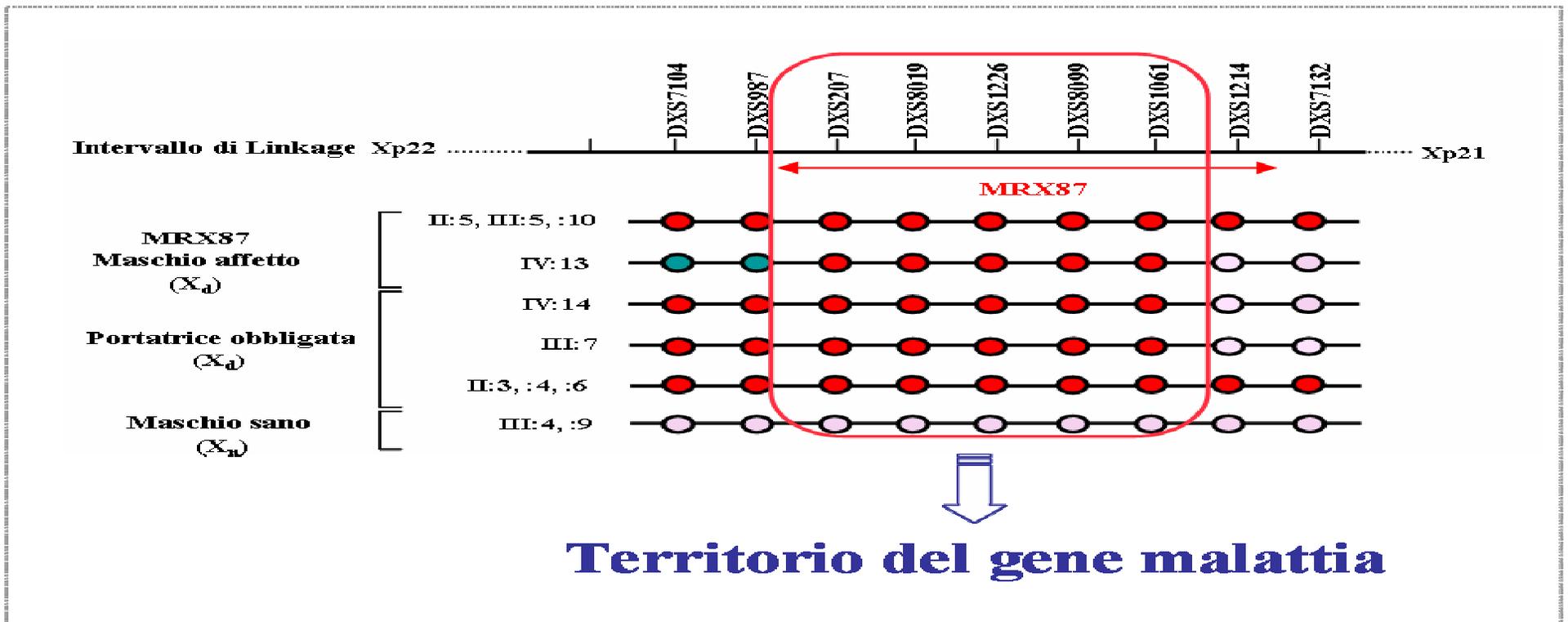
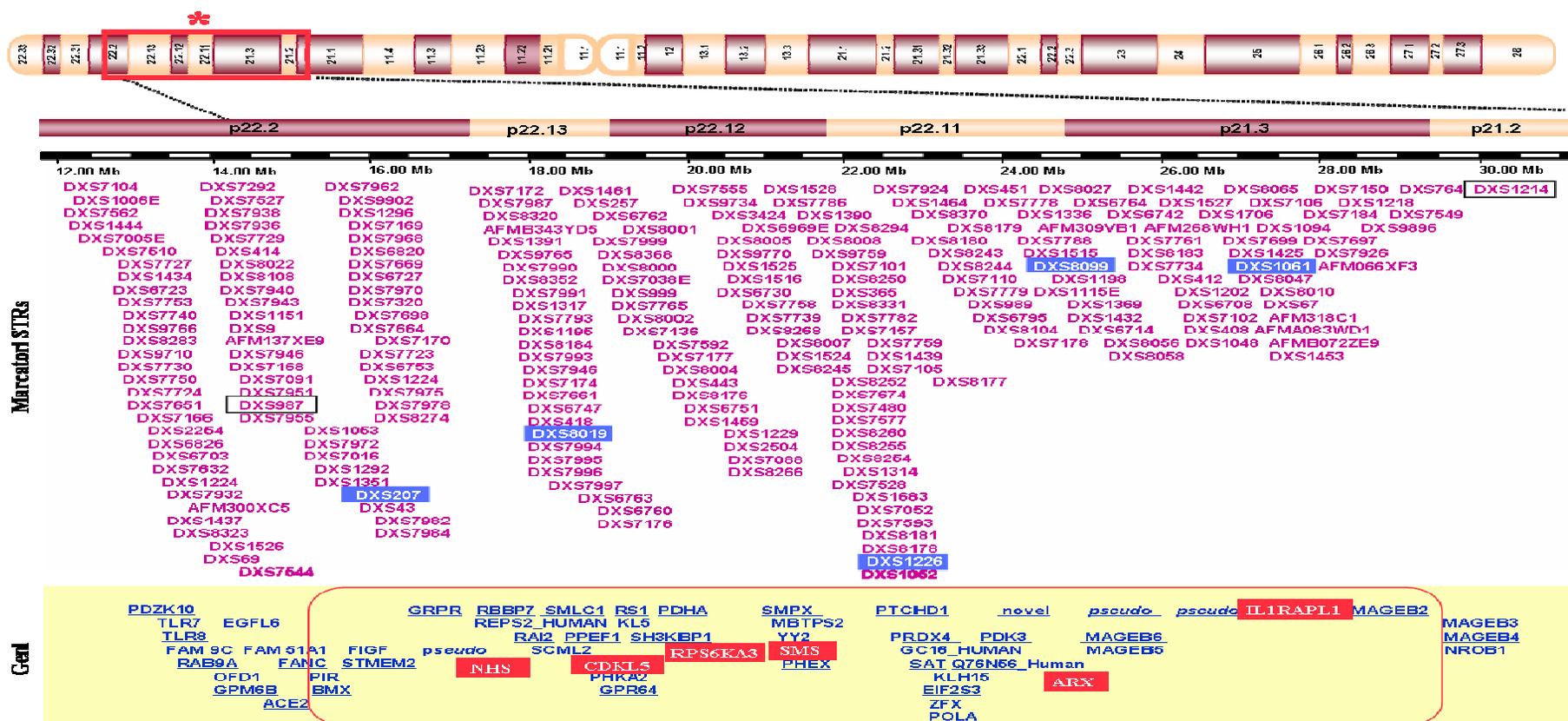


Figura 16. Aplotipo Malattia in MRX87.



NHS

Nance-Horan Syndrome (regolazione dello sviluppo di occhi, denti, cervello)

Patologia: Sindrome NHS (Sindrome cataratto-dentale con Ritardo Mentale)

Mutazioni rare : 5 casi

CDKL5

Cyclin-dependent kinase-like 5

Patologia: 1. Sindrome con spasmi infantili e Ritardo Mentale

2. Sindrome di Rett Atipica

Mutazioni Rare : 6 casi

RSK2

Ribosomal protein S6 kinase alpha 3

Patologia: 1. Sindrome di Coffin-Lowry (Ritardo Mentale severo e dismorfismo facciale)

2. Ritardo Mentale Non Specifico

Mutazioni rare : 19 casi

SMS

Spermine synthase

Patologia: Sindrome di Snyder-Robinson (Ritardo Mentale con disfunzione del circuito cerebellare)

Molto rara: 1 caso

ARX

Aristaless-related homeobox

Patologie: 1. Lissencefalia con anomalie dei genitali

2. Sindrome con spasmi Infantili e severo Ritardo Mentale

3. Epilessia Mioclonica con deficit intellettivo

4. Ritardo Mentale con movimenti distonici delle mani

5. Ritardo Mentale Non Specifico

Sito di mutazione *Hot spot*: 10% dei casi di XLMR

IL1RAPL1

Interleukin 1 receptor accessory protein-like 1

Patologia: Non-specific X-linked mental retardation

Molto rara: 1 caso

Figura 18. Geni coinvolti in disturbi cognitivi nel locus MRX87.

a)

Esone di ARX	Sequenze degli Oligo
1	F 5'-GCCACCCAATCTAAAGTTGC-3' R 5'-GGTGATGATGAGTGATAAGG-3'
2a	F 5CTCCTGTTCTCCTGGTATGTGA--3' R 5'AGAAGAATGCCTGTTNNGGGT-- 3'
2b	F 5'-CTTTTGTGCAAACCTGTTCC-3' R- 5'-TCTCTGCCTGGTACACG- 3'
3	F- 5'-TCTGAGAGCACCAACCAC-3' R-5'-TGCCAGTTTTCTTATGGAATC3'
4	F-5'-GGTGACTGACTCTGTGGC-3' R-5'-TGGAGTGAAATCAACATTTAA3'
5	F 5'-GAGTCAAGCCTTGATCCCAG-3' R 5-GTGTATGGATCGATTGCCTG-'3'

b)

Esone di ARX	Lunghezza dell'amplificato	Tm	MgCl2
1	300 bp	62°	1.5 mM
2a	280 bp	62°	1.5 mM
2b	250 bp	60°	1.5 mM
3	350 bp	62°	1.5 mM
4	345 bp	60°	1.3 mM
5	249 bp	62°	1.5 mM

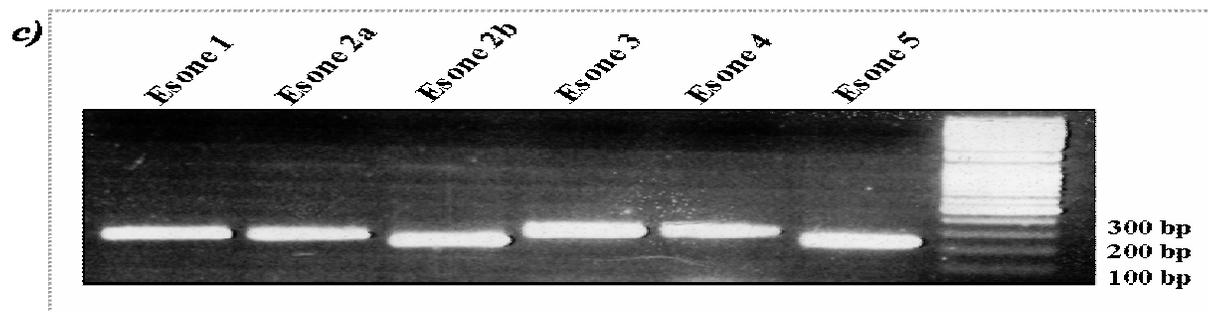


Figura 19. a) Condizioni di amplificazione dei singoli esoni del gene ARX. b) Sequenze nucleotidiche degli oligo gene specifici di ARX. c) Gel di agarosio che mostra i frammenti di PCR ottenuti con gli oligo ARX esone specifici.

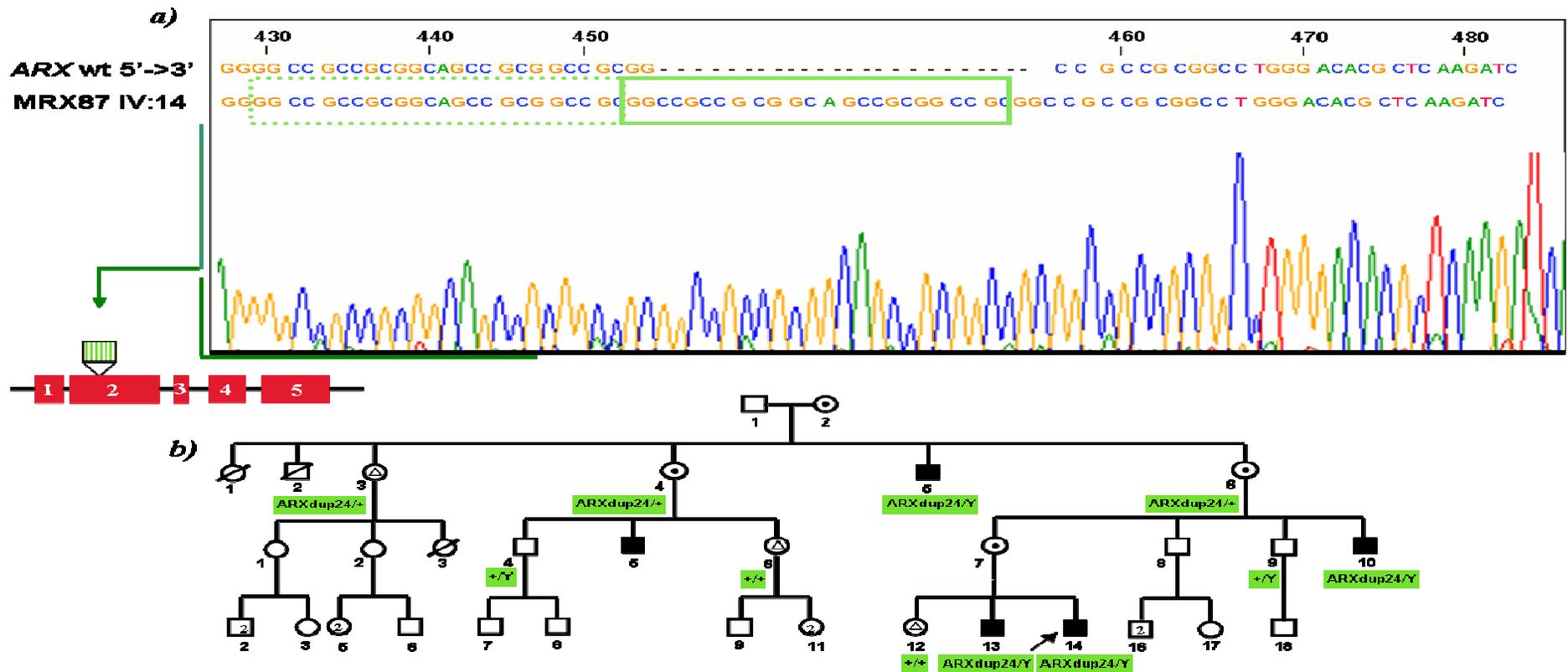


Figura 20. Identificazione e segregazione della mutazione ARX 429_452dup24 in MRX87.

(Laperuta C et al. 2007)

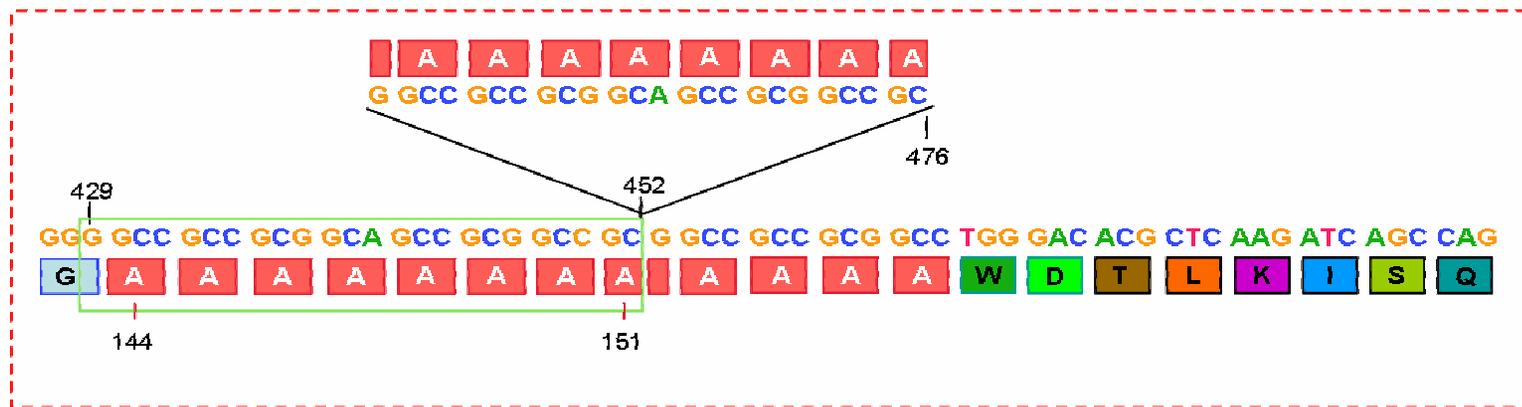
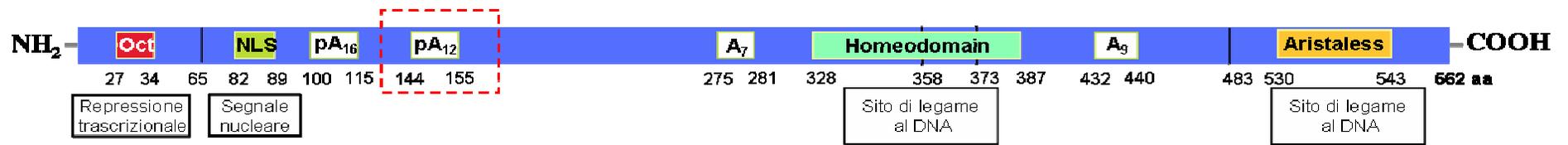


Figura 21. La mutazione ARXdup24 a livello proteico: inserzione di 8 alanine nel poliA_{II} di ARX.

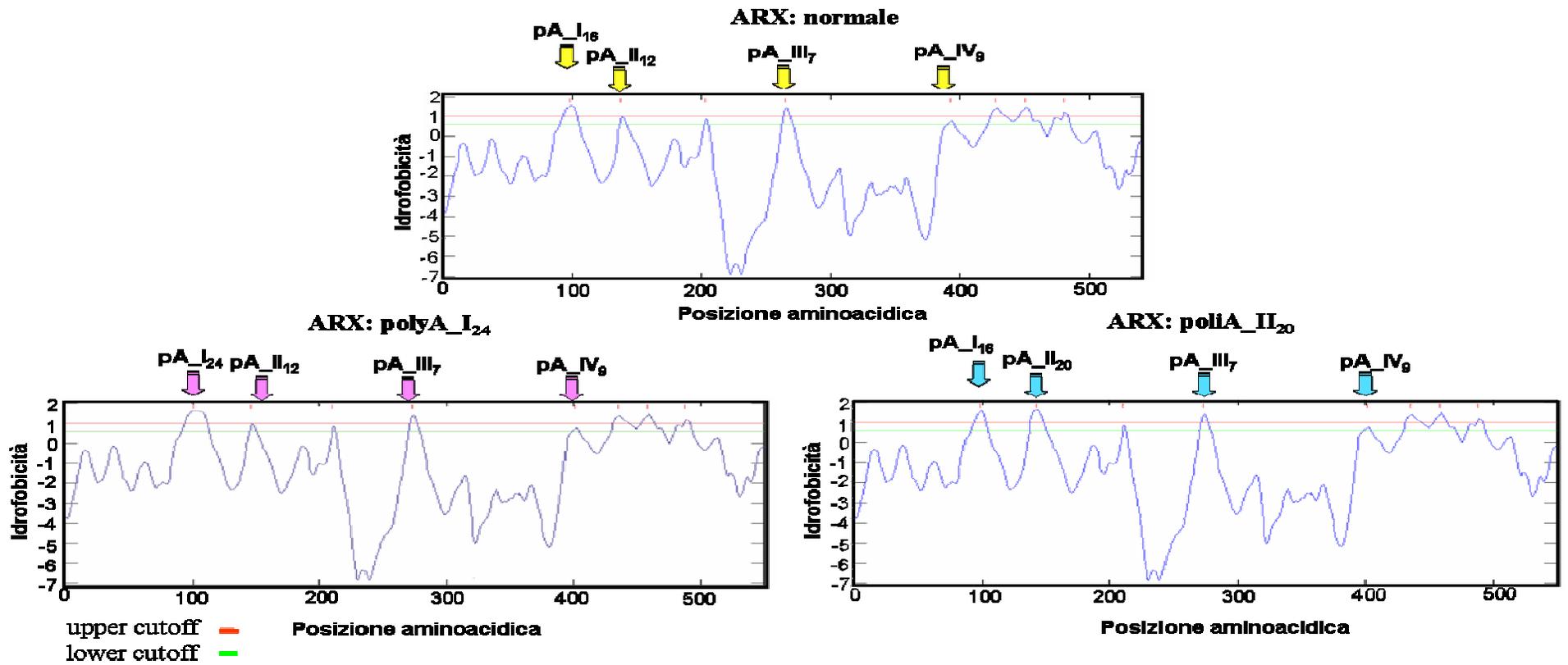


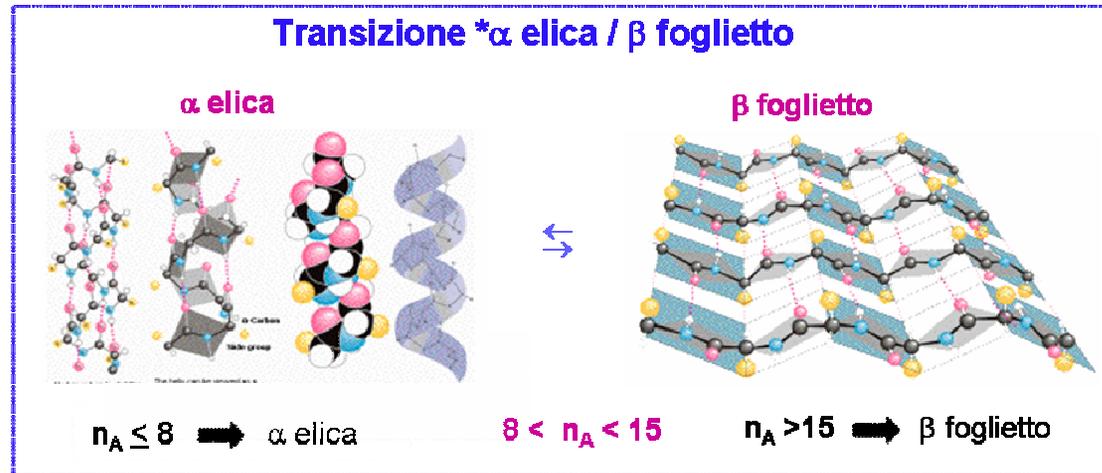
Figura 22. Analisi della idrofobicità della proteina ARX normale e della forma mutata.
TopPred software

Comparazione dei domini di polialanine di ARX

<i>ARX_{poliA}</i>	<i>Posizione aa</i>	<i>n di alanine</i>	<i>α eliche*</i>	<i>Idrofobicit�*</i>
pA_I₁₆ wt	100-116	16	99-119	1.569
pA_II₁₂ wt	144-155	12	139-159	0.985 ← WT
pA_III₇ wt	275-281	7	266-286	1.427
pA_IV₉ wt	431-439	9	428-448	1.414
pA_I₂₄ mut	100-123	24 (+8)	101-121	1.600
pA_II₂₀ mut	144-163	20 (+8)	144-164	1.600 ← II Exp

Tabella 3. Valori di idrofobicit  relativi ai quattro domini di polialanine di ARX normale(wt) e di ARX nelle sue forme espanse (pA_I24 e pA_II20).

* TopPred software <http://bioweb.pasteur.fr>



L'alanina è capace di stabilizzare l'elica:

- Lunghi tratti di A_n causano la formazione di α -eliche.
- L'inserzione di A accresce l'elicità e causa la transizione a β foglietti.

Figura 23. La transizione * α elica / β foglietto in tratti espansi di alanine (*La transizione è influenzata dal pH, dalla temperatura e dai residui amminoacidici fiancheggiati).

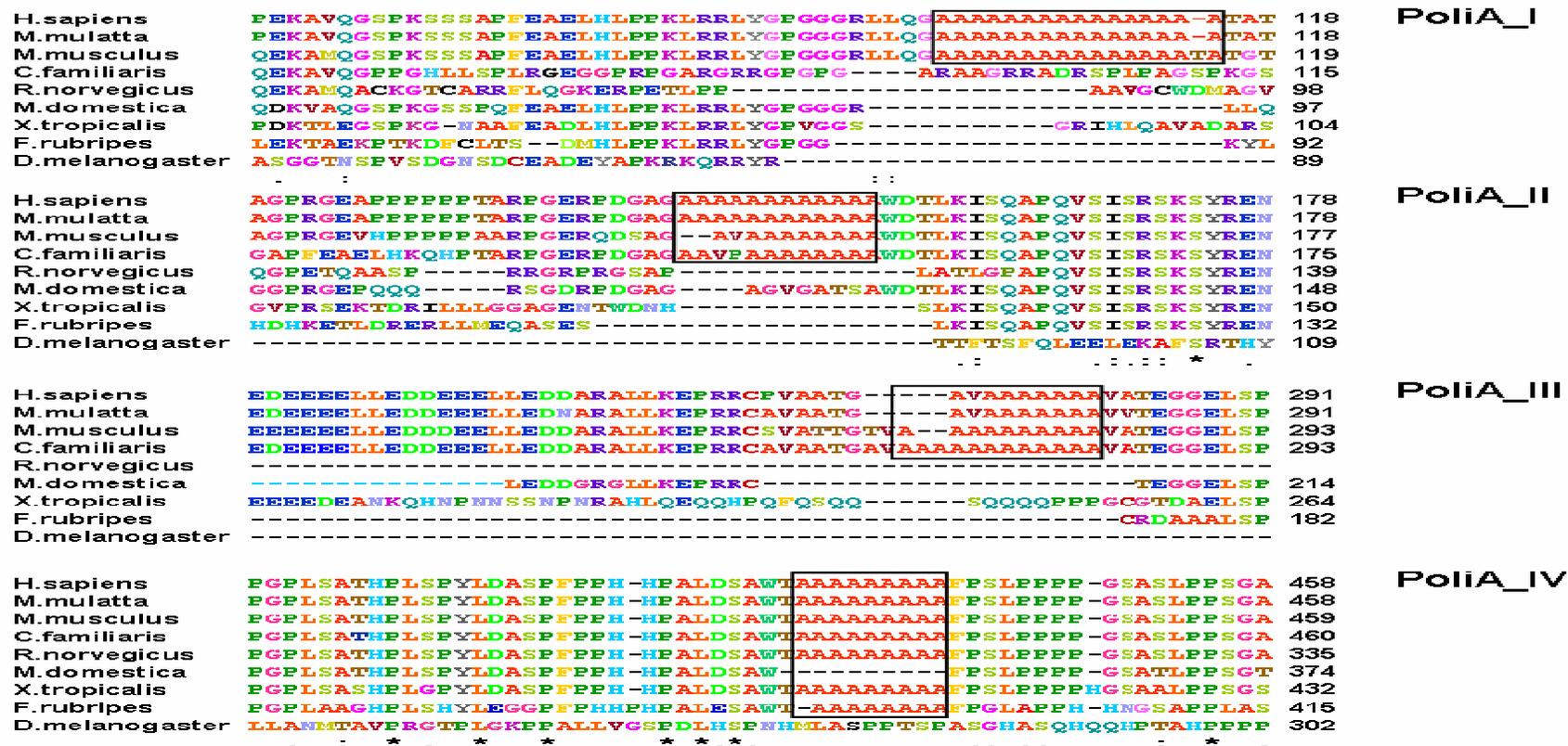


Figura 24. Allineamento multiplio e conservazione dei PolIA di ARX.
CLUSTALW software at <http://www.ebi.ac.uk/clustalw/>

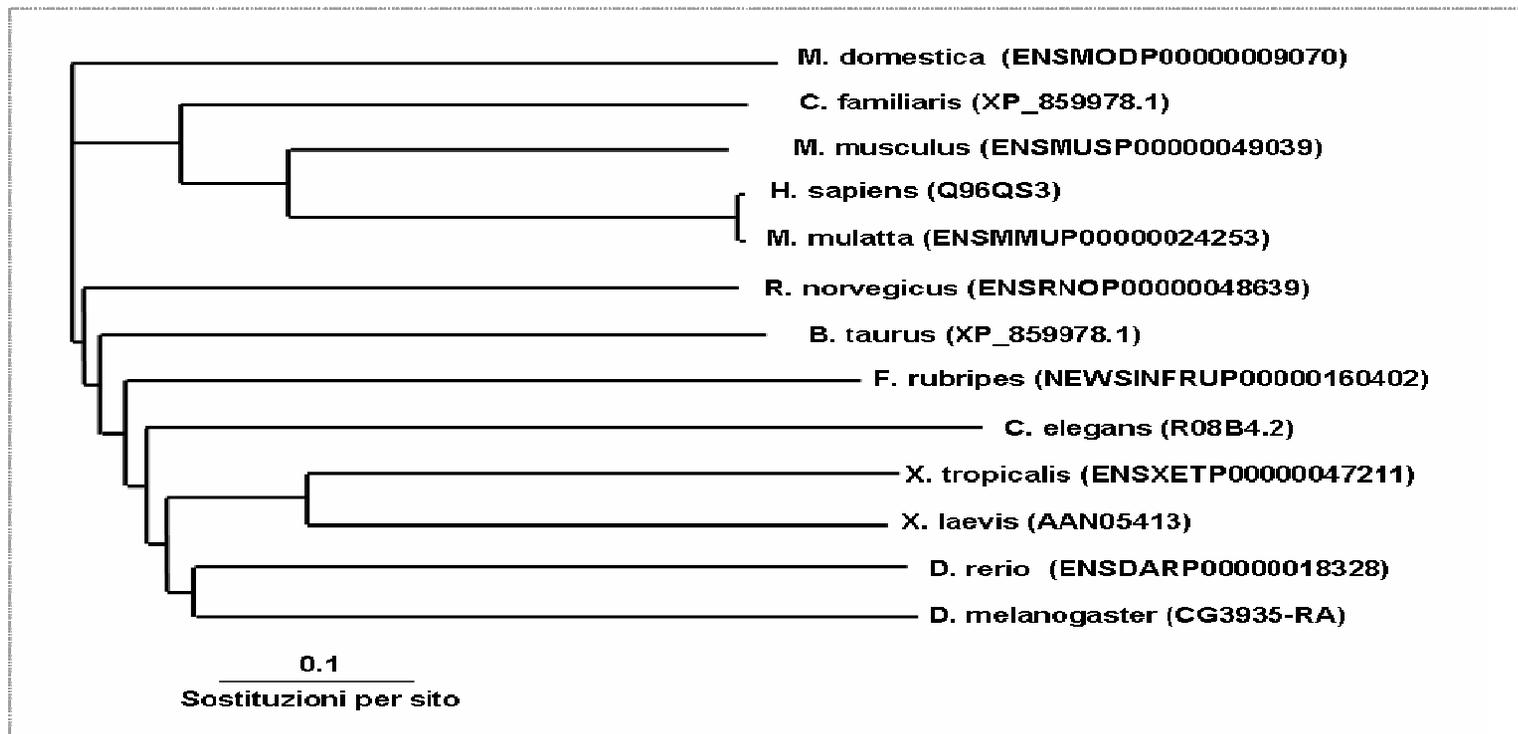


Figura 25. Albero filogenetico di ARX.
CLUSTALW software at <http://www.ebi.ac.uk/clustalw/>

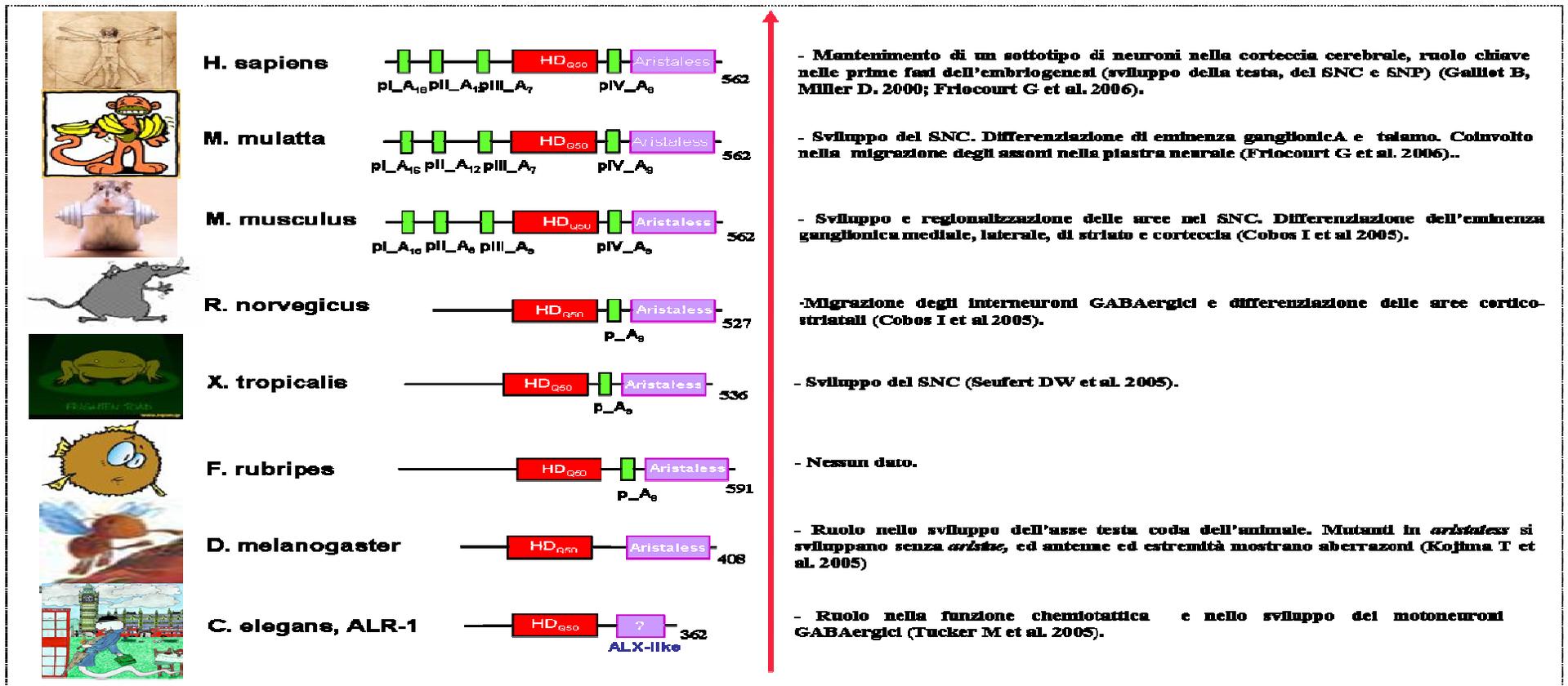


Figura 26. Co-evoluzione Strutturale e Funzionale di Aristaless e dei suoi ortologi.

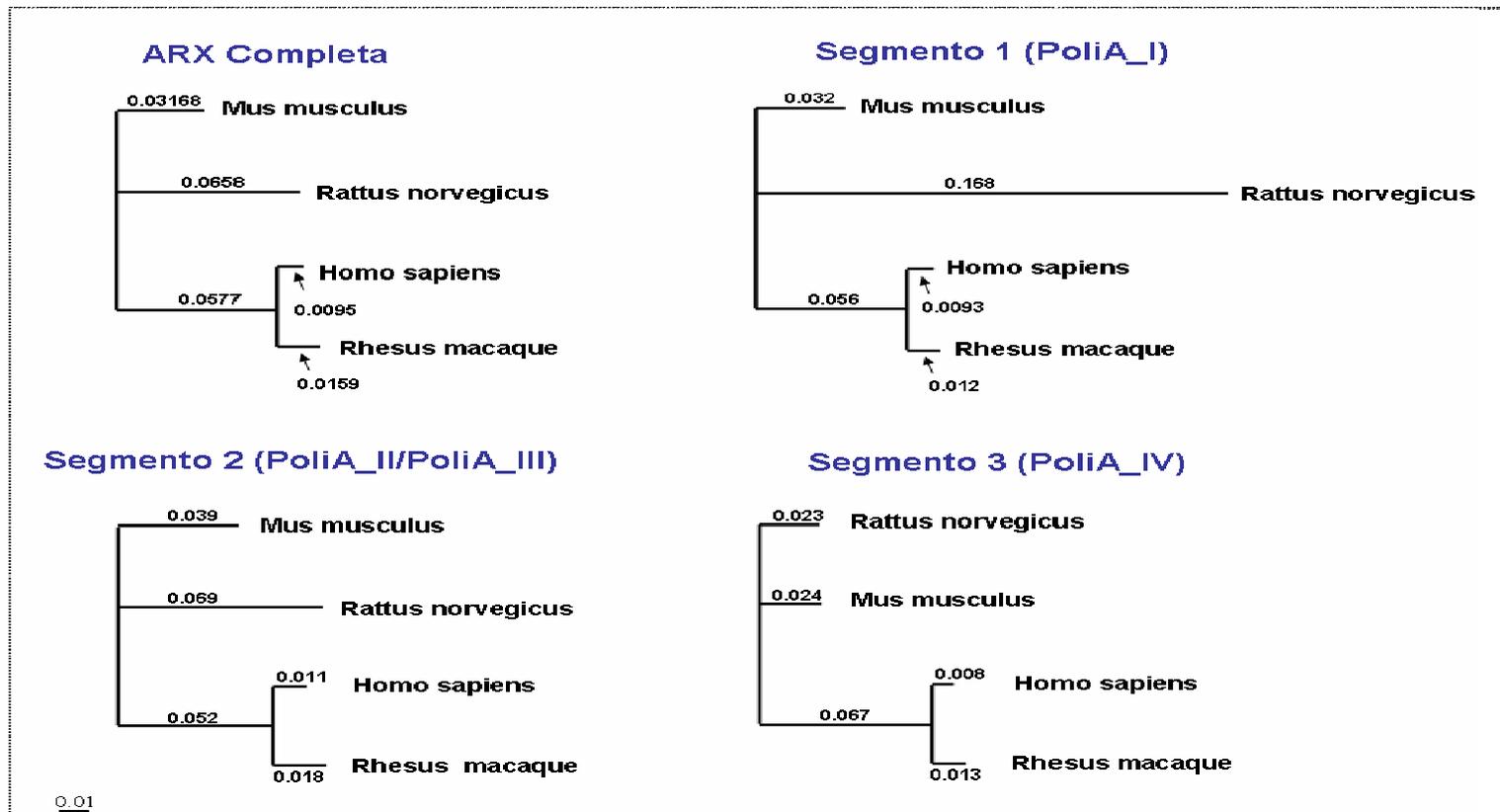


Figura 27: Alberi filogenetici di ARX completa e dei segmenti contenenti i tratti PolIA.
CLUSTALW software at <http://www.ebi.ac.uk/clustalw/>

a)

Gene	Nome del Gene	Primate			Roditori			*Lunghezza Comparata	Patologia
		Ka	Ks	Ka/Ks	Ka	Ks	Ka/Ks		
BDNF	Brain derived neurotrophic factor	0,006	0,026	0,231	0,005	0,100	0,050	558	Central hypoventilation syndrome
MCPHL	Microcephalin	0,040	0,048	0,833	0,070	0,146	0,479	2223	Microcephaly
DLX2	Distal-less homeobox2	0,009	0,137	0,066	0,002	0,057	0,026	834	----
ZIC5	Zic family member 5	0,007	0,036	0,194	0,006	0,113	0,027	423	----
TMOD2	Tropomodulin 2	0,011	0,033	0,333	0,006	0,124	0,040	333	----
POU3F2	POU domain, transcription factor	0,003	0,029	0,103	0,000	0,056	0,000	333	----
FGD1	Rho GEF protein	0,005	0,045	0,102	0,002	0,057	0,032	2670	X linked Mental retardation

b)

Gene	Nome del Gene	Primate			Roditori			*Lunghezza Comparata	Patologia
		Ka	Ks	Ka/Ks	Ka	Ks	Ka/Ks		
ARX	Aristalless related homeobox	0,003	0,083	0,042	0,032	0,127	0,167	1686	XLAG and Syndromic, non Syndromic MR
HOXB1	Homeobox B1	0,010	0,077	0,130	0,023	0,108	0,213	888	----
PHOX2B	Paired mesoderm homeobox 2B	0,000	0,058	0,000	0,127	0,230	0,505	543	Neuroblastoma
OTX1	Orthodenticle, Drosophila homologue ot, 1	0,005	0,107	0,047	0,009	0,167	0,054	648	----
IKBK2	Inhibitor of kappa light polypeptide gene	0,000	0,07	0,000	0,009	0,129	0,069	591	Incontinentia Pigmenti
IL1RAPL1	IL-1 receptor accessory protein like1	0,000	0,052	0,000	0,002	0,057	0,030	834	X-linked Mental Retardation
NCAM1	Cell adhesion molecule neural, 1	0,004	0,047	0,085	0,026	0,208	0,125	2067	----

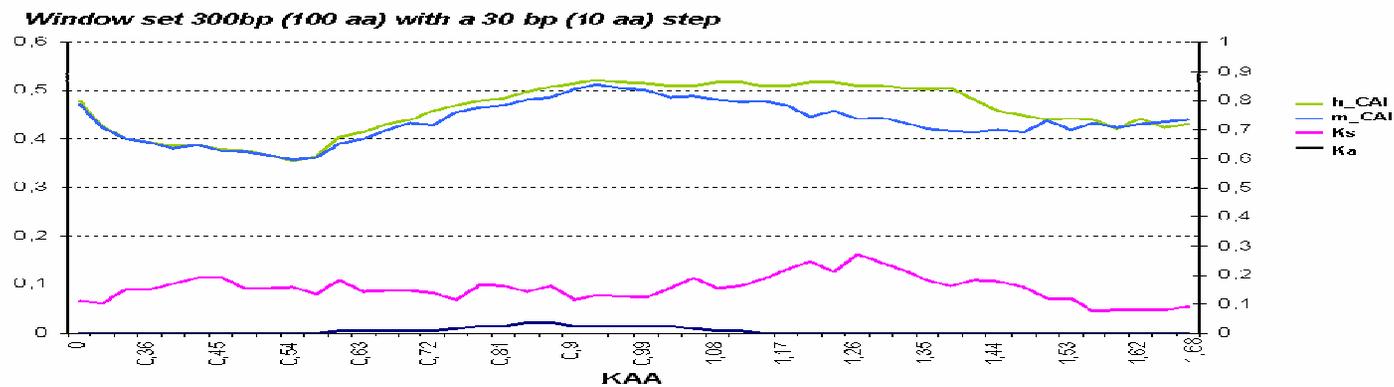
- in **nero**, geni analizzati in Dorus et al. Accelerated evolution, Cell 2004;

- in **rosso**, geni analizzati in questo studio Ka/Ks, mediante K-Estimator6.1 software* cds sequence

Figura 28. L'evoluzione molecolare di Aristalless è più rapida nei non-primati (roditori) che nei primati.

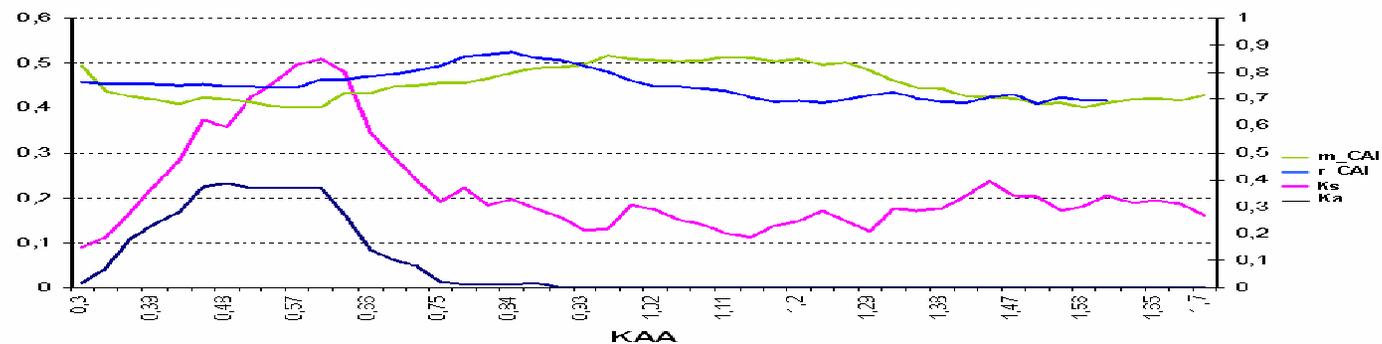
**Ka & Ks
Uomo/Macaca**

Ka = 0,0036
Ks = 0,083
Ka/Ks = 0,422



**Ka & Ks
Topo/Ratto**

Ka = 0,03254
Ks = 0,19507
Ka/Ks = 0,167



*Ka/Ks, K-Estimator6.1 software
 CAI, codon adaptation index
<http://xblast.tamu.edu>*

Figura 29. Andamento del *Ka* e *Ks* in *Aristaless* nei primati e roditori.

7. BIBLIOGRAFIA

Abidi F, Jacquot S, Lassiter C, Trivier E, Hanauer A, Schwartz CE.
Novel mutations in Rsk-2, the gene for Coffin-Lowry syndrome (CLS).
Eur J Hum Genet. 1999. 7(1):20-6.

Abrams MT, Reiss AL, Freund LS, Baumgardner TL, Chase GA, Denckla MB
Molecular-neurobehavioral associations in females with the fragile X full mutation.
Am J Med Genet. 1994. 51(4):317-27.

Albrecht A, Mundlos S
The other trinucleotide repeat: polyalanine expansion disorders.
Curr Opin Genet Dev. 2005. Jun;15(3):285-93.

Altshuler D, Pollara VJ, Cowles CR, Van Etten WJ, Baldwin J, Linton L, Lander ES
An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature.* 2000. 407(6803):513-6.

Amiel J, Trochet D, Clément-Ziza M, Munnich A, Lyonnet S
Polyalanine expansions in human.
Hum Mol Genet. 2004. 13 Spec No 2:R235-43.

Anderson S, Mione M, Yun K, Rubenstein JL.
Differential origins of neocortical projection and local circuit neurons: role of Dlx genes in neocortical interneuronogenesis.
Cereb Cortex. 1999. 9(6):646-54.

Annunziata I, Lanzara C, Conte I, Zullo A, Ventruto V, Rinaldi MM, D'Urso M, Casari G, Ciccodicola A, Miano MG
Mapping of MRX81 in Xp11.2-Xq12 suggests the presence of a new gene involved in nonspecific X-linked mental retardation.
Am J Med Genet A. 2003. 118(3):217-22.

Anvret M, Ahlberg G, Grandell U, Hedberg B, Johnson K, Edström L
Larger expansions of the CTG repeat in muscle compared to lymphocytes from patients with myotonic dystrophy.
Hum Mol Genet. 1993. 2(9):1397-400

Archer HL, Evans J, Edwards S, Colley J, Newbury-Ecob R, O'Callaghan F, Huyton M, O'Regan M, Tolmie J, Sampson J, Clarke A, Osborne J.

CDKL5 mutations cause infantile spasms, early onset seizures, and severe mental retardation in female patients.

J Med Genet. 2006. 43(9):729-34.

Ashley CT Jr, Wilkinson KD, Reines D, Warren ST

FMR1 protein: conserved RNP family domains and selective RNA binding.

Science. 1993. 262(5133):563-6.

Belmont JW.

Genetic control of X inactivation and processes leading to X-inactivation skewing.

Am J Hum Genet. 1996.58:1101-8.

Bienvenu T, Poirier K, Friocourt G, Bahi N, Beaumont D, Fauchereau F, Ben Jeema L, Zemni R, Vinet MC, Francis F, Couvert P, Gomot M, Moraine C, van Bokhoven H, et al.

ARX, a novel Prd-class-homeobox gene highly expressed in the telencephalon, is mutated in X-linked mental retardation.

Hum Mol Genet. 2002. 11(8):981-91.

Bonneau D, Toutain A, Laquerrière A, Marret S, Saugier-veber P, Barthez MA, Radi S, Biran-Mucignat V, Rodriguez D, Gélot A.

X-linked lissencephaly with absent corpus callosum and ambiguous genitalia (XLAG): clinical, magnetic resonance imaging, and neuropathological findings

Ann Neurol. 2002. 51(3):340-9.

Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, Hunter K, Stanton VP, Thirion JP, Hudson T, et al

Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member

Cell. 1992. 69(2):385.

Brooks SP, Ebenezer ND, Poopalasundaram S, Lehmann OJ, Moore AT, Hardcastle AJ

Identification of the gene for Nance-Horan syndrome (NHS).

J Med Genet. 2004. 41(10):768-71.

Brown L, Paraso M, Arkell R, Brown S

In vitro analysis of partial loss-of-function ZIC2 mutations in holoprosencephaly: alanine tract expansion modulates DNA binding and transactivation.

Hum Mol Genet. 2005. 14(3):411-20.

Brown LY, Brown SA.

Alanine tracts: the expanding story of human illness and trinucleotide repeats.

Trends Genet. 2004 . 20(1):51-8.

Campuzano V, Montermini L, Moltò MD, Pianese L, Cossée M, Cavalcanti F, Monros E, Rodius F, Duclos F, Monticelli A, Zara F, Cañizares J, Koutnikova H, et al.

Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion.

Science. 1996. 271(5254):1423-7.

Carrié A, Jun L, Bienvenu T, Vinet MC, McDonell N, Couvert P, Zemni R, Cardona A, Van Buggenhout G, et al.

A new member of the IL-1 receptor family highly expressed in hippocampus and involved in X-linked mental retardation.

Nat Genet. 1999. 23(1):25-31.

Cason AL, Ikeguchi Y, Skinner C, Wood TC, Holden KR, Lubs HA, Martinez F, Simensen RJ, Stevenson RE, Pegg AE, Schwartz CE

X-linked spermine synthase gene (SMS) defect: the first polyamine deficiency syndrome.

Eur J Hum Genet. 2003. 11(12):937-44.

Chamberlain NL, Driver ED, Miesfeld RL

The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function.

Nucleic Acids Res. 1994. 22(15):3181-6.

Chelly J, Mandel JL.

Monogenic causes of X-linked mental retardation.

Nat Rev Genet. 2001. 2(9):669-80.

Chiurazzi P, Hamel BC, Neri G

XLMR genes: update 2000.

Eur J Hum Genet. 200. 9(2):71-81.

Chiurazzi P, Tabolacci E, Neri G

X-linked mental retardation (XLMR): from clinical conditions to cloned genes.

Crit Rev Clin Lab Sci. 2004. 41(2):117-58.

Cobos I, Broccoli V, Rubenstein JL.

The vertebrate ortholog of *Aristaless* is regulated by *Dlx* genes in the developing forebrain.

J Comp Neurol. 2005. 483(3):292-303.

Colombo E, Galli R, Cossu G, Gécz J, Broccoli V.

Mouse orthologue of *ARX*, a gene mutated in several X-linked forms of mental retardation and epilepsy, is a marker of adult neural stem cells and forebrain GABAergic neurons.

Dev Dyn. 2004. 231(3):631-9

Cummings CJ, Zoghbi HY

Fourteen and counting: unraveling trinucleotide repeat diseases.

Hum Mol Genet. 2000. 12;9(6):909-16.

Dobyns WB, Berry-Kravis E, Havernick NJ, Holden KR, Viskochil D

X-linked lissencephaly with absent corpus callosum and ambiguous genitalia.

Am J Med Genet. 1999. 86(4):331-7.

Filosa S, Giacometti N, Wangwei C, De Mattia D, Pagnini D, Alfinito F, Schettini F, Luzzatto L, Martini G.

Somatic-cell selection is a major determinant of the blood-cell phenotype in heterozygotes for glucose-6-phosphate dehydrogenase mutations causing severe enzyme deficiency.

Am J Hum Genet. 1996. 59(4):887-95.

Fimiani G, Laperuta C, Falco G, Ventruto V, D'Urso M, Ursini MV, Miano MG.

Heterozygosity mapping by quantitative fluorescent PCR reveals an interstitial deletion in Xq26.2-q28 associated with ovarian dysfunction.

Hum Reprod. 2006. 21(2):529-35.

Friocourt G, Poirier K, Rakić S, Parnavelas JG, Chelly J.

The role of ARX in cortical development.

Eur J Neurosci. 2006. 23(4):869-76.

Fusco F, Bardaro T, Fimiani G, Mercadante V, Miano MG, Falco G, Israël A, Courtois G, D'Urso M, Ursini MV

Molecular analysis of the genetic defect in a large cohort of IP patients and identification of novel NEMO mutations interfering with NF-kappaB activation.

Hum Mol Genet. 2004. 13(16):1763-73.

Galliot B, Miller D

Origin of anterior patterning. How old is our head?

Trends Genet. 2000. 16(1):1-5.

Gécz J, Mulley J

Genes for cognitive function: developments on the X.

Genome Res. 2000. 10(2):157-63.

Gécz J, Cloosterman D, Partington M.

ARX: a gene for all seasons.

Curr Opin Genet Dev. 2006. 16(3):308-16.

Gestinari-Duarte Rde S, Santos-Rebouças CB, Boy RT, Pimentel MM.

ARX mutation c.428-451dup (24bp) in a Brazilian family with X-linked mental retardation.

Eur J Med Genet. 2006. 49(3):269-75.

Gilbert SL, Dobyns WB, Lahn BT

Genetic links between brain development and brain evolution.

Nat Rev Genet. 2005. 6(7):581-90.

Goodman FR, Mundlos S, Muragaki Y, Donnai D, Giovannucci-Uzielli ML, Lapi E, Majewski F, McGaughran J, McKeown C, Reardon W, Upton J, Winter RM, Olsen BR, Scambler PJ.

Synpolydactyly phenotypes correlate with size of expansions in HOXD13 polyalanine tract.

Proc Natl Acad Sci USA. 1997. 94(14):7458-63.

Greenfield JG

Spontaneous demyelinating diseases in man and animals

Acta Neurol Psychiatr Belg. 1954. 54(9):621-32.

Grønskov K, Hjalgrim H, Nielsen IM, Brøndum-Nielsen K

Screening of the ARX gene in 682 retarded males.

Eur J Hum Genet. 2004. 12(9):701-5.

Hagerman PJ, Hagerman RJ

Fragile X-associated tremor/ataxia syndrome (FXTAS).

Ment Retard Dev Disabil Res Rev. 2004b. 10(1):25-30.

Hagerman PJ, Hagerman RJ

The fragile-X premutation: a maturing perspective.

J Hum Genet. 2004a. 74(5):805-16.

Han K, Manley JL.

Transcriptional repression by the Drosophila even-skipped protein: definition of a minimal repression domain.

Genes Dev. 1993. 7(3):491-503.

Harper PS.

Postoperative complications in myotonic dystrophy.

Lancet. 1989. 2(8674):1269.

Hatten ME

Central nervous system neuronal migration.

Annu Rev Neurosci. 1999;22:511-39.

Herzing LB, Romer JT, Horn JM, Ashworth A.

Xist has properties of the X-chromosome inactivation centre.

Nature. 1997. 386: 272-5.

Höweler CJ, Busch HF, Geraedts JP, Niermeijer MF, Staal A

Anticipation in myotonic dystrophy: fact or fiction?

Brain. 1989 Jun;112 (Pt 3):779-97.

Karlin S, Burge C.

Trinucleotide repeats and long homopeptides in genes and proteins associated with nervous system disease and development.

Proc Natl Acad Sci USA. 1996. 93(4):1560-5.

Kato M, Das S, Petras K, Sawaishi Y, Dobyns WB

Polyalanine expansion of ARX associated with cryptogenic West syndrome.

Neurology. 2003. 61(2):267-76. Links

Kato M, Das S, Petras K, Kitamura K, Morohashi K, Abuelo DN, Barr M, Bonneau D, Brady AF, et al.

Mutations of ARX are associated with striking pleiotropy and consistent genotype-phenotype correlation.

Hum Mutat. 2004. 23(2):147-59.

Kaufmann WE, Moser HW

Dendritic anomalies in disorders associated with mental retardation.

Cereb Cortex. 2000 Oct;10(10):981-91.

Kitamura K, Yanazawa M, Sugiyama N, Miura H, Iizuka-Kogo A, Kusaka M, Omichi K, Suzuki R, Kato-Fukui Y, et al.

Mutation of ARX causes abnormal development of forebrain and testes in mice and X-linked lissencephaly with abnormal genitalia in humans.

Nat Genet. 2002. 32(3):359-69.

Koide R, Ikeuchi T, Onodera O, Tanaka H, Igarashi S, Endo K, Takahashi H, Kondo R, Ishikawa A, Hayashi T, et al

Unstable expansion of CAG repeat in hereditary dentatorubral-pallidolusian atrophy (DRPLA).

Nat Genet. 1994. 6(1):9-13.

Kojima T, Tsuji T, Saigo K

A concerted action of a paired-type homeobox gene, aristaless, and a homolog of Hox11/tlx homeobox gene, clawless, is essential for the distal tip development of the Drosophila leg.

Dev Biol. 2005. 279(2):434-45.

Koob MD, Moseley ML, Schut LJ, Benzow KA, Bird TD, Day JW, Ranum LP

An untranslated CTG expansion causes a novel form of spinocerebellar ataxia (SCA8).

Nat Genet. 1999. 21(4):379-84.

Lahn BT, Page DC

Four evolutionary strata on the human X chromosome.

Science. 1999. 286(5441):964-7

Lambert de Rouvroit C, Goffinet AM.

Neuronal migration.

Mech Dev. 2001. 105(1-2):47-56.

Laperuta C, Spizzichino L, D'Adamo P, Monfregola J, Maiorino A, D'Eustacchio A, Ventruto V, Neri G, D'Urso M, Chiurazzi P, Ursini MV, Miano MG

MRX87 family with Aristaless X dup24bp mutation and implication for polyAlanine expansions.

BMC Med Genet. 2007. 8:25.

Lehrke R.

Theory of X-linkage of major intellectual traits.

Am J Ment Defic. 1972. 76(6):611-9.

Leventer RJ, Mills PL, Dobyns WB

X-linked malformations of cortical development.

Am J Med Genet. 2000. 97(3):213-20.

Li M, Miwa S, Kobayashi Y, Merry DE, Yamamoto M, Tanaka F, Doyu M, Hashizume Y, Fischbeck KH, Sobue G

Nuclear inclusions of the androgen receptor protein in spinal and bulbar muscular atrophy.

Ann Neurol. 1998. 44(2):249-54

Lindqvist AK, Magnusson PK, Balciuniene J, Wadelius C, Lindholm E, Alarcón-Riquelme ME, Gyllenstein UB

Chromosome-specific panels of tri- and tetranucleotide microsatellite markers for multiplex fluorescent detection and automated genotyping: evaluation of their utility in pathology and forensics.

Genome Res. 1996. 6(12):1170-6.

Lutz RE

Trinucleotide repeat disorders.

Semin Pediatr Neurol. 2007. 14(1):26-33

Mahadevan M, Tsilfidis C, Sabourin L, Shutler G, Amemiya C, Jansen G, Neville C, Narang M, Barceló J, O'Hoy K, et al

Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene.

Science. 1992. 255(5049):1253-5.

Mandel JL, Chelly J

Monogenic X-linked mental retardation: is it as frequent as currently estimated? The paradox of the ARX (Aristaless X) mutations.

Eur J Hum Genet. 2004. 12(9):689-93

Martin, J. P. & Bell, J

A pedigree of mental defect showing sex-linkage.

J. Neurol. Psychiat. BMJ. 1943. 154-157.

Mailhos C, André S, Mollereau B, Goriely A, Hemmati-Brivanlou A, Desplan C.

Drosophila Goosecoid requires a conserved heptapeptide for repression of paired-class homeoprotein activators.

Development. 1998. 125(5):937-47.

Marcus S, Steen AM, Andersson B, Lambert B, Kristoffersson U, Francke U
Mutation analysis and prenatal diagnosis in a Lesch-Nyhan family showing non-random X-inactivation interfering with carrier detection tests.

Hum Genet. 1992. 89(4):395-400.

Matise TC, Sachidanandam R, Clark AG, Kruglyak L, Wijsman E, Kakol J, Buyske S, Chui B, Cohen P, et al.

A 3.9-centimorgan-resolution human single-nucleotide polymorphism linkage map and screening set.

Am J Hum Genet. 2003. 73(2):271-84.

McKenzie O, Ponte I, Mangelsdorf M, Finnis M, Colasante G, Shoubridge C, Stifani S, Gécz J, Broccoli V.

Aristaless-related homeobox gene, the gene responsible for West syndrome and related disorders, is a Groucho/transducin-like enhancer of split dependent transcriptional repressor.

Neuroscience. 2007. 146(1):236-47.

Mhatre AN, Trifiro MA, Kaufman M, Kazemi-Esfarjani P, Figlewicz D, Rouleau G, Pinsky L

Reduced transcriptional regulatory competence of the androgen receptor in X-linked spinal and bulbar muscular atrophy.

Nat Genet. 1993. 5(2):184-8.

Miano MG, Jacobson SG, Carothers A, Hanson I, Teague P, Lovell J, Cideciyan AV, Haider N, Stone EM, Sheffield VC, Wright AF.

Pitfalls in homozygosity mapping.

Am J Hum Genet. 2000. 67(5):1348-51.

Miano MG, Laperuta C, Chiurazzi P, D'Urso M, Ursini MV.

Ovarian dysfunction and FMR1 alleles in a large Italian family with POF and FRAXA disorders: case report.

BMC Med Genet. 2007. 8(1):18

Miano MG, Laperuta C, Ursini MV.

From Nonsyndromic X-linked Mental Retardation (MRX) diseases to discovery genes for cognitive circuitry in humans. Progress in Medical Genetics.

NovaScience Publisher, 2006.

Miller AP, Willard HF

Chromosomal basis of X chromosome inactivation: identification of a multigene domain in Xp11.21-p11.22 that escapes X inactivation.

Proc Natl Acad Sci U S A. 1998. 21;95(15):8709-14

Miura H, Yanazawa M, Kato K, Kitamura K.

Expression of a novel aristaless related homeobox gene 'Arx' in the vertebrate telencephalon, diencephalon and floor plate.

Mech Dev. 1997. 65(1-2):99-109.

Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM, Brown AL, Budowle B

Validation of STR typing by capillary electrophoresis.

J Forensic Sci. 2001. 46(3):661-76.

Nakajima H, Kimura F, Nakagawa T, Furutama D, Shinoda K, Shimizu A, Ohsawa N.

Transcriptional activation by the androgen receptor in X-linked spinal and bulbar muscular atrophy.

J Neurol Sci. 1996. 142(1-2):12-6.

Nasrallah IM, Minarcik JC, Golden JA

A polyalanine tract expansion in Arx forms intranuclear inclusions and results in increased cell death.

J Cell Biol. 2004. 167(3):411-6.

Nawara M, Szczaluba K, Poirier K, Chrzanowska K, Pilch J, Bal J, Chelly J, Mazurczak T.

The ARX mutations: a frequent cause of X-linked mental retardation.

Am J Med Genet A. 2006. 140(7):727-32.

Ogata T, Matsuo N, Hiraoka N, Hata JI

X-linked lissencephaly with ambiguous genitalia: delineation of further case.

Am J Med Genet. 2000. 94(2):174-6.

Ohno S, Klein J, Poole J, Harris C, Destree A, Morrison M.

Genetic control of lactate dehydrogenase formation in the hagfish *Eptatretus stoutii*.

Science. 1967. 156(771):96-8

Orr HT, Zoghbi HY

Reversing neurodegeneration: a promise unfolds.

Cell. 2000. 101(1):1-4.

Orstavik KH, Orstavik RE, Naumova AK, D'Adamo P, Gedeon A, Bolhuis PA, Barth PG, Toniolo D.

X chromosome inactivation in carriers of Barth syndrome.

Am J Hum Genet. 1998. 63: 1457-63.

Ott J, Lucek P.

Complex traits on the map

Recent Results Cancer Res. 1998. 154:285-91

Parnavelas JG

The origin and migration of cortical neurones: new vistas.

Trends Neurosci. 2000. 23(3):126-31.

Parrish JE, Scheuerle AE, Lewis RA, Levy ML, Nelson DL.

Selection against mutant alleles in blood leukocytes is a consistent feature in Incontinentia Pigmenti type 2.

Hum Mol Genet. 1996.5: 1777-83.

Pearlman AL, Faust PL, Hatten ME, Brunstrom JE

New directions for neuronal migration.

Curr Opin Neurobiol. 1998. 8(1):45-54.

Pearson CE, Nichol Edamura K, Cleary JD

Repeat instability: mechanisms of dynamic mutations.

Nat Rev Genet. 2005. 6(10):729-42.

Plenge RM, Tranebjaerg L, Jensen PK, Schwartz C, Willard HF

Evidence that mutations in the X-linked DDP gene cause incompletely penetrant and variable skewed X inactivation.

Am J Hum Genet. 1999. 64(3):759-67

Poirier K, Lacombe D, Gilbert-Dussardier B, Raynaud M, Desportes V, de Brouwer AP, Moraine C, Fryns JP, Ropers HH, Beldjord C, Chelly J, Bienvu T.

Screening of ARX in mental retardation families: Consequences for the strategy of molecular diagnosis.

Neurogenetics. 2006. 7(1):39-46

Poirier K, Abriol J, Souville I, Laroche-Raynaud C, Beldjord C, Gilbert B, Chelly J, Bienvenu T.

Maternal mosaicism for mutations in the ARX gene in a family with X linked mental retardation

Hum Genet. 2005. 118(1):45-8.

Posthuma D, De Geus EJ, Baaré WF, Hulshoff Pol HE, Kahn RS, Boomsma DI

The association between brain volume and intelligence is of genetic origin.

Nat Neurosci. 2002. 5(2):83-4.

Ramakers GJ

Rho proteins, mental retardation and the cellular basis of cognition.

Trends Neurosci. 2002. 25(4):191-9.

Ratcliffe SG, Murray L, Teague P.

Edinburgh study of growth and development of children with sex chromosome abnormalities.

Birth Defects Orig Artic Ser. 1986. 22:73-118.

Raymond FL

X linked mental retardation: a clinical guide.

J Med Genet. 2006. 43(3):193-200.

Ropers HH, Hamel BC

X-linked mental retardation.

Nat Rev Genet. 2005. 6(1):46-57.

Ropers HH, Hoeltzenbein M, Kalscheuer V, Yntema H, Hamel B, Fryns JP, Chelly J, Partington M, Gecz J, Moraine C

Nonsyndromic X-linked mental retardation: where are the missing mutations?

Trends Genet. 2003. 19(6):316-20.

Rousseau F, Heitz D, Biancalana V, Blumenfeld S, Kretz C, Boué J, Tommerup N, Van Der Hagen C, DeLozier-Blanchet C, Croquette MF, et al

Direct diagnosis by DNA analysis of the fragile X syndrome of mental retardation

N Engl J Med. 1991. 325(24):1673-81.

Rice JP, Saccone NL, Corbett J

The lod score method.

Adv Genet. 2001. 42:99-113.

Seufert DW, Prescott NL, El-Hodiri HM

Xenopus aristaless-related homeobox (xARX) gene product functions as both a transcriptional activator and repressor in forebrain development.

Dev Dyn. 2005. 232(2):313-24.

Sheffield VC, Weber JL, Buetow KH, Murray JC, Even DA, Wiles K, Gastier JM, Pulido JC, Yandava C, Sunden SL, et al

A collection of tri- and tetranucleotide repeat markers used to generate high quality, high resolution human genome-wide linkage maps.

Hum Mol Genet. 1995. 4(10):1837-44.

Sherman SL, Meadows KL, Ashley AE

Examination of factors that influence the expansion of the fragile X mutation in a sample of conceptuses from known carrier females.

Am J Med Genet. 1996. 64(2):256-60.

Shinchuk LM, Sharma D, Blondelle SE, Reixach N, Inouye H, Kirschner DA

Poly-(L-alanine) expansions form core beta-sheets that nucleate amyloid assembly.

Proteins. 2005. 61(3):579-89.

Simeone A, D'Apice MR, Nigro V, Casanova J, Graziani F, Acampora D, Avantaggiato V

Orthopedia, a novel homeobox-containing gene expressed in the developing CNS of both mouse and Drosophila.

Neuron. 1994 Jul;13(1):83-101.

Sinden RR, Potaman VN, Oussatcheva EA, Pearson CE, Lyubchenko YL, Shlyakhtenko LS

Triplet repeat DNA structures and human genetic disease: dynamic mutations from dynamic DNA.

J Biosci. 2002. 27:53-65

Skuse DH

X-linked genes and mental functioning.

Hum Mol Genet. 2005. 14 Spec No 1:R27-32.

Strømme P, Mangelsdorf ME, Scheffer IE, Gécz J.

Infantile spasms, dystonia, and other X-linked phenotypes caused by mutations in Aristaless related homeobox gene, ARX.

Brain Dev. 2002. 24(5):266-8.

Strømme P, Mangelsdorf ME, Shaw MA, Lower KM, Lewis SM, Bruyere H, Lütcherath V, Gedeon AK, Wallace RH, Scheffer IE, Turner G, Partington M, Frints SG, Fryns JP, Sutherland GR, Mulley JC, Gécz J

Mutations in the human ortholog of Aristaless cause X-linked mental retardation and epilepsy.

Nat Genet. 2002. 30(4):441-5.

Stevenson RE.

Splitting and lumping in the nosology of XLMR.

Am J Med Genet. 2000. 97(3):174-82

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. 1(6):397-400.

Toniolo D

In search of the MRX genes.

Am J Med Genet. 2000. 97(3):221-7.

Tsilfidis C, MacKenzie AE, Mettler G, Barceló J, Korneluk RG

Correlation between CTG trinucleotide repeat length and frequency of severe congenital myotonic dystrophy.

Nat Genet. 1992. 1(3):192-5.

Tucker M, Sieber M, Morphew M, Han M

The Caenorhabditis elegans aristaless orthologue, alr-1, is required for maintaining the functional and structural integrity of the amphid sensory organs.

Mol Biol Cell. 2005. 16(10):4695-704

Turner G, Partington M, Kerr B, Mangelsdorf M, Gecz J
Variable expression of mental retardation, autism, seizures, and dystonic hand movements in two families with an identical ARX gene mutation.
Am J Med Genet. 2002. 112(4):405-11.

Turner G, Webb T, Wake S, Robinson H
Prevalence of fragile X syndrome.
Am J Med Genet. 1996. 64(1):196-7.

Trochet D, Hong SJ, Lim JK, Brunet JF, Munnich A, Kim KS, Lyonnet S, Goridis C, Amiel J.
Molecular consequences of PHOX2B missense, frameshift and alanine expansion mutations leading to autonomic dysfunction.
Hum Mol Genet. 2005. 14(23):3697-708.

Verkerk AJ, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DP, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang FP, et al.
Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome.
Cell. 1991 May 31;65(5):905-14

Verkerk AJ, de Graaff E, De Boule K, Eichler EE, Konecki DS, Reyniers E, Manca A, Poustka A, Willems PJ, Nelson DL, et al
Alternative splicing in the fragile X gene FMR1.
Hum Mol Genet. 1993. 2(8):1348.

Vulliamy TJ, Knight SW, Dokal I, Mason PJ.
Skewed X-inactivation in carriers of X-linked dyskeratosis congenita.
Blood. 1997 Sep 15;90 (6):2213-6.

Weber JL.
Human DNA polymorphisms and methods of analysis.
Curr Opin Biotechnol. 1990. 1(2):166-71

Woods KS, Cundall M, Turton J, Rizotti K, Mehta A, Palmer R, Wong J, Chong WK, Al-Zyoud M, El-Ali M, Otonkoski T, Martinez-Barbera JP, Thomas PQ, Robinson IC, Lovell-Badge R, Woodward KJ, Dattani MT .

Over- and underdosage of SOX3 is associated with infundibular hypoplasia and hypopituitarism.

Am J Hum Genet. 2005. 76(5):833-49.

Zechner U, Wilda M, Kehrer-Sawatzki H, Vogel W, Fundele R, Hameister H
A high density of X-linked genes for general cognitive ability: a runaway process shaping human evolution?

Trends Genet. 2001. 17(12):697-701.

