



Università degli Studi di Padova

Dipartimento di Pediatria

SCUOLA DI DOTTORATO DI RICERCA IN
MEDICINA DELLO SVILUPPO E SCIENZE DELLA PROGRAMMAZIONE

INDIRIZZO GENETICA, BIOCHIMICA E MOLECOLARE

XXIII CICLO

**VALUTAZIONE DEL PROFILO DI ESPRESSIONE GENICA CEREBRALE
NEL MODELLO MURINO
PER LA MUCOPOLISACCARIDOSI DI TIPO II (SINDROME DI HUNTER)
EFFETTUATA MEDIANTE TECNOLOGIA RNA-Seq**

Direttore della Scuola: Ch.mo Prof. Giuseppe Basso

Coordinatore d'indirizzo: Ch.mo Dott. Maurizio Scarpa

Supervisore: Ch.mo Dott. Maurizio Scarpa

Co-supervisore: Dott.ssa Rosella Tomanin

Dottoranda: Dott.ssa Marika Salvalaio

INDICE

RIASSUNTO	1
ABSTRACT	5
INTRODUZIONE	9
Patologie da accumulo lisosomiale	9
I glicosaminoglicani	11
Le mucopolisaccaridosi	13
La sindrome di Hunter (MPS II)	15
Aspetti clinici	16
<i>Coinvolgimento neurologico</i>	<i>16</i>
<i>Diagnosi</i>	<i>17</i>
Approcci terapeutici	18
<i>Terapia Enzimatica Sostitutiva</i>	<i>18</i>
<i>Genisteina</i>	<i>19</i>
<i>Trattamento del coinvolgimento neurologico</i>	<i>20</i>
Patofisiologia	21
L'animale modello per la MPS II	23
L'analisi del trascrittoma	25
SCOPO	31
MATERIALI E METODI	33
L'animale modello per la MPS II	33
Piano sperimentale	33
Genotipizzazione degli animali	33
Test comportamentali	34
<i>Open Field Test (OF)</i>	<i>34</i>
<i>Footprint test</i>	<i>34</i>
<i>Analisi statistica</i>	<i>35</i>
Preparazione campioni tessuto cerebrale	35
Estrazione dell'RNA totale	35

Analisi mediante tecnologia RNA-Seq	37
<i>Purificazione dell'mRNA</i>	37
<i>Preparazione della library ed emulsion PCR</i>	38
<i>Deposizione delle biglie</i>	39
<i>Sequenziamento per ligazione</i>	39
Analisi dei dati	42
<i>Allineamento</i>	42
<i>Identificazione dei geni differenzialmente espressi</i>	42
<i>Analisi funzionale</i>	43
RISULTATI E DISCUSSIONE	47
Test comportamentali	47
<i>Open Field test (OF)</i>	48
<i>Footprint test</i>	51
Analisi del trascrittoma	53
<i>RNA-Seq: Preparazione e analisi dei campioni</i>	53
Valutazione integrità RNA	53
Purificazione dell'mRNA	54
Frammentazione dell'RNA	55
Preparazione della library	56
<i>Analisi dei dati</i>	57
Allineamento	57
Identificazione dei geni differenzialmente espressi	59
Analisi funzionale	61
Gene Ontology	62
Categorie funzionali	67
Pathways	72
<i>Analisi delle vie di segnale alterate</i>	78
Eparan solfato binding protein	78
Omeostasi del calcio	79
Stress ossidativo	84
Autofagia	86

Axon guidance	90
Neuroinfiammazione	93
Comparazione con altre malattie neurodegenerative	95
La malattia di Alzheimer	95
La malattia di Parkinson	96
La malattia di Huntington	98
Ormone della crescita	100
CONCLUSIONI	103
APPENDICE	107
BIBLIOGRAFIA	113

RIASSUNTO

La Sindrome di Hunter (o Mucopolisaccaridosi di tipo II, MPS II) è una patologia ereditaria appartenente al più vasto gruppo delle malattie da accumulo lisosomiale (Lysosomal Storage Disorders, LSDs), comprendente quasi 50 diverse patologie. Tali patologie, sebbene individualmente molto rare, presentano un'incidenza complessiva che va da 1:4000 a 1:7000, a seconda della popolazione considerata. Le LSDs, che risultano per lo più da deficit singoli, e più raramente multipli, di enzimi lisosomiali deputati alla degradazione di molecole complesse, sono devastanti malattie multiorgano e multisistemiche che, in buona parte dei casi, comprendono un coinvolgimento neurologico grave.

Poco è noto tuttora sulla patofisiologia di queste sindromi e, ancor meno, sulle cause del loro deficit neurologico. Nel momento in cui alcune di queste patologie trovano finalmente beneficio dall'applicazione della terapia enzimatica sostitutiva, risalta maggiormente la problematica del trattamento della componente cognitiva e comportamentale. Essa infatti non trova beneficio da questi nuovi approcci terapeutici, poiché gli enzimi impiegati non sono in grado di attraversare la barriera emato-encefalica. Nasce da qui la necessità di comprendere la patogenesi neurologica di queste sindromi e, nel caso specifico di questo lavoro, della sindrome di Hunter. Tale comprensione, seppure molto complessa, potrebbe consentire, tra le altre cose, lo sviluppo di specifiche strategie terapeutiche mirate al cervello.

In questo lavoro di ricerca preclinica, il modello murino per la MPS II è stato impiegato per effettuare una valutazione molecolare complessa attraverso l'impiego di tecnologie high throughput. RNA derivati da 2 aree cerebrali, la corteccia e il mesencefalo, sono stati analizzati mediante next generation sequencing, impiegando la procedura SOLiD[®] (Sequencing by Oligo Ligation and Detection). Questa tecnologia, seppure estremamente indicata proprio per il sequenziamento dell'RNA, è stata finora poco utilizzata a questo scopo, a causa dei costi ancora piuttosto elevati, ma soprattutto, a causa della complessità dell'analisi che richiede notevoli competenze bioinformatiche e capacità di gestione dei software. Il sequenziamento dell'RNA è una

tecnologia estremamente potente in grado di evidenziare, a differenza della tecnica del microarray, tutti i trascritti cellulari in maniera indistinta.

Il lavoro qui presentato è un'analisi di tipo comparativo tra le aree cerebrali dell'animale knock-out per l'IDS e le corrispondenti aree dell'animale sano di controllo. I dati, dopo la fase di allineamento e di filtrazione, sono stati classificati secondo i domini della Gene Ontology e analizzati in base alle principali categorie funzionali.

L'analisi ha chiaramente evidenziato il coinvolgimento di molti geni e di parecchie vie specificamente implicati in processi neurologici. L'alterata struttura e funzione cellulare sono state evidenziate sia in modo generico, attraverso analisi di termini ugualmente alterati nelle due diverse aree cerebrali, sia in modo specifico, all'interno di ciascuna area cerebrale. Ciò consente di mettere in risalto i geni la cui alterata espressione è direttamente correlata allo stato di accumulo patologico della cellula e i geni con espressione differenziale che, invece, rappresentano pathways specifici per le funzioni svolte da quell'area cerebrale.

Molto interessante potrebbe risultare anche l'alterazione di geni implicati in alcune comuni patologie neurodegenerative croniche quali il morbo di Alzheimer e di Parkinson. Vie comuni potrebbero essere ipotizzate per l'instaurarsi delle malattie o come conseguenza dello stato patologico.

La parte finale dell'analisi ha preso in considerazione le vie di segnale e le correlazioni più interessanti, alcune delle quali già precedentemente considerate come potenziali candidati nella patogenesi delle malattie da accumulo lisosomiale: l'eparan solfato binding protein, l'omeostasi del calcio, lo stress ossidativo, il processo dell'autofagia, l'axon guidance, la neuroinfiammazione, la correlazione con altre malattie neurodegenerative e l'ormone della crescita.

Una forte alterazione del comparto endocellulare dovuta al progressivo, patologico aumento dei depositi primari di glicosaminoglicani, ma anche di quelli secondari quali i gangliosidi, potrebbe giustificare il coinvolgimento delle proteine implicate nel metabolismo del calcio cellulare, rilevato da questo lavoro. Essendo poi il calcio un messaggero ubiquitario coinvolto in differenti processi biologici non stupisce il ruolo di filo conduttore in molti pathways, evidenziato da questa analisi.

In conclusione, seppure fortemente complessa, l'analisi intrapresa in questo studio ha evidenziato le enormi potenzialità della procedura, dovute alla sua caratteristica capacità di mettere in luce, oltre a processi correlati in modo sospetto alla patologia, anche pathways non sospetti o la cui implicazione nella determinazione dello stato patologico non sia ancora stata definita.

ABSTRACT

Hunter Syndrome (mucopolysaccharidosis type II, MPS II) is an inherited metabolic disease belonging to the wide group of lysosomal storage disorders (LSDs), including nearly 50 different pathologies. Although individually rare, these pathologies have a collective incidence from 1:4000 to 1:7000 live births, depending on the analyzed population.

Most LSDs are due to single, or more rarely multiple, deficit of lysosomal enzymes, responsible of macromolecules degradation. Unmetabolized substrates result in multiorgan and multisystem diseases, which in the vast majority of the patients seriously affect also the central nervous system.

Very little is known on LSDs pathophysiology and even less on the determinants of their neurological impairment.

Since ERT (enzymatic replacement therapy) is getting promising results only in treatment of the LSDs without neurological involvement, because the enzyme cannot efficiently cross the blood brain barrier, more attention should be put in handling the cognitive and behavioral components of these pathologies.

In this scenario, the comprehension of the neurological pathogenesis of these diseases, and in this specific case, of Hunter syndrome, becomes mandatory. Such understanding, although quite complex, might allow, among others, the development of new specific therapeutic strategies targeted to the brain.

In this pre-clinic investigation, the murine MPS II model was used for a complex molecular analysis through high-throughput technology. RNAs from two different brain areas, cortex and midbrain, have been analyzed with next generation sequencing, by using SOLiD® (Sequencing by oligo ligation and detection) technology. Although this technology is considered the most specific for RNA sequencing, it has not been extensively used so far due to its very high costs and to the complexity of the data analysis requiring advanced bioinformatics competence and a good capacity of

software handling. RNA sequencing is a very powerful technology that, in contrast to microarray, can point out every single cellular transcript indistinctly.

This work is a comparative study between brain areas derived from the IDS-knock-out and the healthy control mice. Data obtained, after alignment and filtration, have been classified according to Gene Ontology Domains, and analyzed by functional categories. The analysis has clearly pointed out the involvement of a wide group of genes and pathways implicated in neurological processes.

Altered structures and cellular functions have been highlighted both through the analysis of terms that have the same alteration in the two cerebral areas and, in a more specific way, through the analysis of terms related to each area. This approach can underline genes whose altered expression is directly related to the cell pathological condition and, at the same time, genes with differential expression, representing instead specific pathways for the function of the brain area considered.

Also the detection of alterations in genes involved in some common neurodegenerative diseases (as Parkinson's and Alzheimer's) could be very interesting; common pathways could be hypothesized for the disease development or as a consequence of the pathological state.

The last part of this work has focused on some selected pathways that have been chosen as the most interesting candidates for the pathogenesis of the disease: heparan sulphate binding protein, calcium homeostasis, oxidative stress, autophagy, axon guidance, neuroinflammation, correlation with other neurodegenerative diseases and the growth hormone.

A significant endocellular alteration, due to progressive increase of glycosaminoglycan deposits and also of secondary deposits as gangliosides, could justify the involvement of proteins related to calcium metabolism, detected by the present analysis; since calcium plays an important ubiquitous messenger function in different biological processes, its role of leitmotiv in many pathways emerging from this analysis is not surprising.

In conclusion, although very complex, the analysis here presented has highlighted the great power of this technology, due to its ability to detect, not only pathways

obviously related to this disease, but also non-suspected pathways, whose role in the determination of the pathological condition has not been yet clarified.

INTRODUZIONE

PATOLOGIE DA ACCUMULO LISOSOMIALE

Le patologie da accumulo lisosomiale (LSDs) devono la loro classificazione alla scoperta dei lisosomi da parte di Christian de Duve nel 1955 e allo sviluppo del concetto di malattia lisosomiale dieci anni dopo grazie a Hers.

Hers si occupava di malattie da accumulo di glicogeno e notò in una di queste, la patologia di Pompe, accumuli non nel citoplasma, ma in organelli rivestiti di membrana, i lisosomi. Ciò comportò la classificazione delle già clinicamente riconosciute patologie da accumulo lisosomiale.

Tali patologie rappresentano un gruppo di circa 50 malattie genetiche causate dalla carenza di proteine solubili lisosomiali che risiedono nel lume dei lisosomi.

La degradazione di macromolecole biologiche come i glicosaminoglicani (GAGs), le glicoproteine e i glicolipidi avviene nei lisosomi presenti in ogni cellula dell'organismo.

Un difetto dei lisosomi, più comunemente dovuto a mutazioni nei geni che codificano per gli enzimi catalitici, comporta deposito di materiale che normalmente viene degradato e conseguente malattia da accumulo. Altre forme di disfunzioni lisosomiali sono dovute a deficit di proteine di membrana, errori di targeting e di attivazione enzimatica. In alcune di queste patologie gli accumuli coinvolgono diversi organi comportando sintomi gravi e morte prematura [Ballabio e Gieselmann, 2009; Beck, 2010].

Ad eccezione della malattia di Fabry, della Mucopolisaccaridosi di tipo II (sindrome di Hunter) e della malattia di Dannon, a ereditarietà di tipo X-linked recessiva, le altre LSDs sono trasmesse con modalità autosomica recessiva [Meikle et al., 1999].

Le LSDs sono monogeniche e per ogni gene sono state descritte numerose mutazioni che comprendono varianti missense, nonsense e di splicing, delezioni e inserzioni parziali. Alcune mutazioni comportano la perdita completa dell'attività enzimatica, mentre altre inducono una riduzione dell'attività. Inoltre, generalmente non si

riscontrano correlazioni genotipo-fenotipo, di conseguenza non sempre è possibile predire il decorso clinico della malattia sulla base della mutazione individuata.

Le LSDs si snodano su un ampio spettro di gravità della patologia e una buona parte di queste presenta una forma infantile, una giovanile e una adulta. Spesso la più severa, la forma infantile, si presenta con un acuto coinvolgimento neurologico e i pazienti muoiono nei primi anni di vita. Nella forma adulta i sintomi si sviluppano più lentamente e la disabilità spesso deriva da sintomi periferici. La forma giovanile è intermedia tra le precedenti. I sintomi neurologici possono includere apoplezia, demenza, disfunzioni del tronco encefalico e iperattività. Tra i sintomi periferici più comuni si registrano epatosplenomegalia, cardio- e nefropatie, disostosi multipla, rigidità articolari, atrofia muscolare e opacità corneale.

In alcune sindromi la forma severa può essere caratterizzata da un preponderante coinvolgimento neurologico e da un minimo coinvolgimento periferico (es. malattia di Sanfilippo, Mucopolisaccaridosi di tipo III), in altre da gravi disfunzioni periferiche e un raro coinvolgimento neurologico (es. malattia di Fabry, di Pompe e Mucopolisaccaridosi di tipo IV e VI).

Ogni LSD presenta un quadro clinico e patologico caratteristico che, presumibilmente, è collegato alla natura del substrato accumulato e ai tipi cellulari in cui esso si accumula. Ad esempio, tutte le forme della malattia di Pompe (o Glicogenosi di tipo II) sono caratterizzate da miopatia, come ci si aspetta per il ruolo importante che il glicogeno riveste nel muscolo e molte sfingolipidosi sono caratterizzate da un coinvolgimento neurologico, come si può supporre sulla base degli alti livelli di glicosfingolipidi che si trovano nel cervello. Tuttavia, per ragioni ancora del tutto non chiarite, solo una piccola percentuale di pazienti con la malattia di Gaucher (una sfingolipidosi) mostrano segni neurologici. In alcune LSDs la variabilità fenotipica può essere spiegata dai differenti livelli di attività enzimatica residua, ma in altre non c'è una chiara correlazione genotipo-fenotipo, e pazienti con un background genetico simile, a volte con la stessa mutazione, possono presentare differenti sintomi clinici o addirittura risultare sintomatici e asintomatici. Ciò è probabilmente imputabile al diverso background genetico e/o a fattori comportamentali che possono influenzare il decorso clinico.

Le LSDs possono essere raggruppate in diversi modi, tuttavia la classificazione più pratica è basata sulla caratterizzazione della proteina o dell'enzima difettivo o, piuttosto, sulla natura del substrato accumulato [Futerman e van Meer, 2004].

La prevalenza delle patologie da accumulo lisosomiale varia nelle diverse popolazioni e in letteratura si trovano dati relativi a queste malattie in Australia (12.9/100000 nati vivi), Olanda (14/100000 nati vivi), Italia (12.1/100000 nati vivi), Portogallo (25/100000 nati vivi) e Repubblica Ceca (12.25/100000) [Meikle et al., 1999; Poorthuis et al., 1999; Dionisi-Vici et al., 2002; Pinto et al., 2004; Poupetova, 2010].

Individualmente, pertanto, le LSDs sono considerate malattie genetiche rare, ma nel loro complesso sono relativamente comuni e rappresentano una causa importante di patologia.

I GLICOSAMINOGLICANI

Gli eteropolisaccaridi più abbondanti nel corpo sono i glicosaminoglicani (GAGs). Queste molecole sono lunghi polisaccaridi non ramificati contenenti unità disaccaridiche ripetute. Le unità disaccaridiche sono costituite da due zuccheri modificati, N-acetilgalattosamina (GalNAc) o N-acetilglucosamina (GlcNAc), e un acido uronico come il glucuronato o l'iduronato.

I GAGs sono molecole con una forte carica negativa, con conformazione distesa che conferisce una alta viscosità alla soluzione in cui si trovano. Sono localizzati principalmente sulla superficie delle cellule o nella matrice extracellulare (ECM).

Oltre all'alta viscosità i GAGs sono poco comprimibili e questa loro caratteristica li rende ideali come componenti del fluido lubrificante nelle giunture. Allo stesso tempo, la rigidità di queste molecole fornisce integrità alle cellule, permettendo la loro migrazione. I GAGs fisiologicamente significativi sono l'acido ialuronico, il dermatan solfato, il condroitin solfato, l'eparina, l'eparan solfato e il cheratan solfato.

La degradazione dei vari tipi di GAGs è catalizzata dall'azione, in sequenza, degli enzimi idrolitici lisosomiali, esoglicosidasi ed endoglicosidasi. Il catabolismo procede sempre a partire dall'estremità non ridotta, ad opera di glicosidasi e solfatasi. Nei tessuti adulti il ricambio di questi polimeri è relativamente lento, con tempi di semivita dell'ordine di

giorni o settimane. In **Figura I** sono riportate le vie degradative dell'eparan-, del cheratan- e del dermatan-solfato.

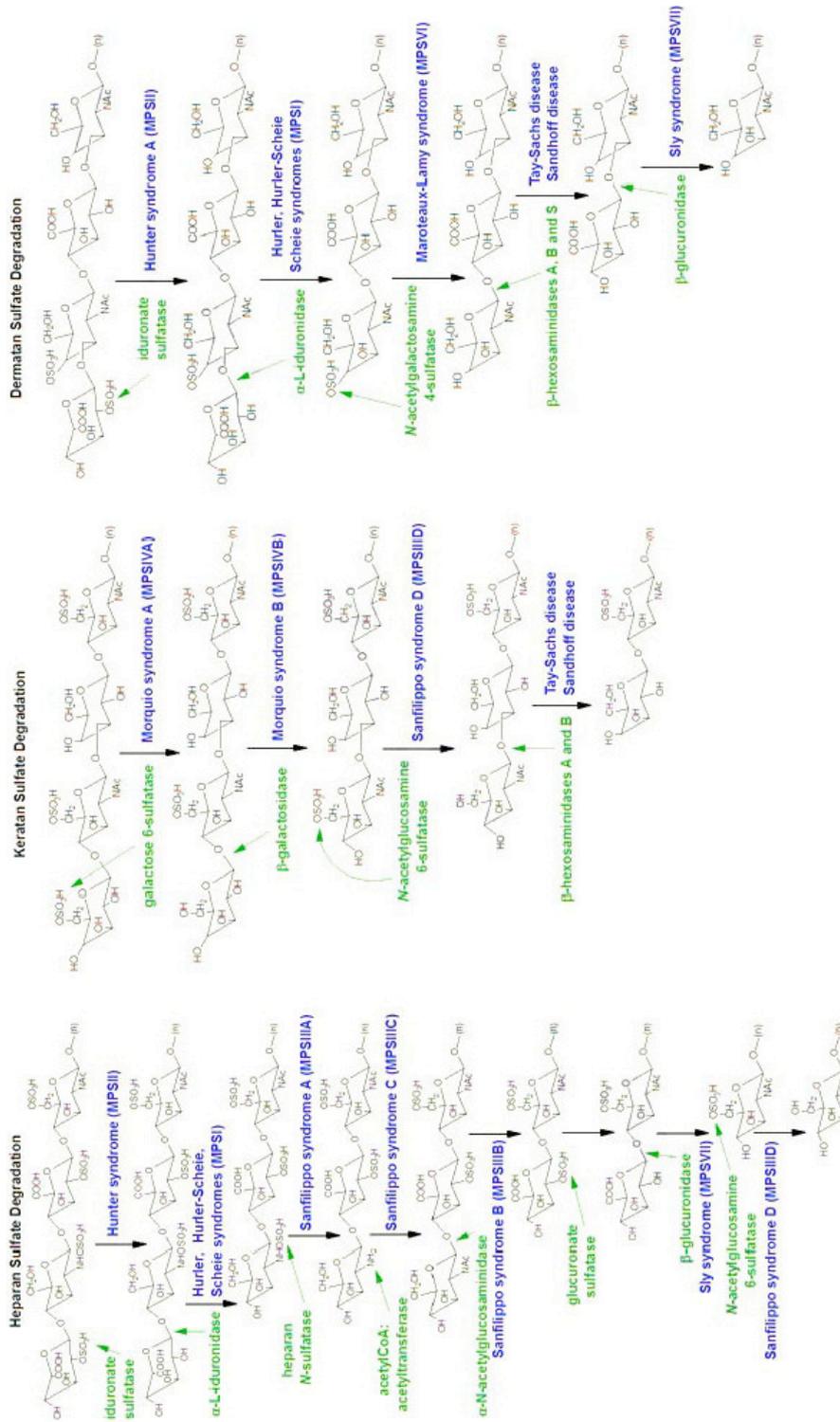


Figura I: Catabolismo di eparan, cheratan e dermatan solfato [M.W. King, 2008; www.themedicalbiochemistrypage.org].

LE MUCOPOLISSACARIDOSI

Le mucopolisaccaridosi (MPS) sono una famiglia di malattie ereditarie causate dalla mancanza di uno degli enzimi necessari alla degradazione dei glicosaminoglicani o mucopolisaccaridi (GAGs). I GAGs parzialmente o non degradati sono accumulati nei lisosomi ed eliminati attraverso le urine. Ogni MPS conosciuta coinvolge il deficit di uno degli 11 enzimi necessari nei vari livelli di degradazione del dermatan-, eparan-, cheratan-, o condroitin-solfato singolarmente o in combinazione.

Le patologie sono croniche, progressive e spesso presentano un ampio spettro di severità clinica. La classificazione delle maggiori caratteristiche delle MPS sono riassunte in **Tabella I** con i rispettivi enzimi deficitari, localizzazione del gene e substrato accumulato.

NOME	OMIM	ENZIMA DEFICITARIO	LOC. GENE	MOLECOLA ACCUMULATA
MPS I (Hurler, Hurler-Scheie, Schieie)	607014 607015 607016	α -L-iduronidasi	4p16.3	Dermatan solfato Eparan solfato
MPS II (Hunter)	309900	iduronato 2-solfatasi	Xq28	Dermatan solfato Eparan solfato
MPS III A (Sanfilippo A)	252900	Eparan-sulfamidasi	17q25.3	Eparan solfato
MPS III B (Sanfilippo B)	252920	N-acetil- α -glucosaminidasi	17q21	Eparan solfato
MPS III C (Sanfilippo C)	252930	AcCoA α -glucosamide-N-acetiltransferasi	8p11.1	Eparan solfato
MPS III D (Sanfilippo D)	252940	N-acetilglucosamina-6-solfatasi	12q14	Eparan solfato
MPS IV A (Morquio A)	253000	Solfatasi N-acetilgalattosamina-6-solfato	16q24.3	Cheratan solfato Condroitin solfato
MPS IV B (Morquio B)	253010	β -D-galattosidasi	3p21.33	Cheratan solfato Galattosio-Oligosaccaridi
MPS VI (Maroteaux-Lamy)	253200	Arilsolfatasi B	15q12	Dermatan solfato
MPS VII (Sly)	253220	β -D-glucuronidasi	7q21.11	Dermatan solfato Eparan solfato Condroitin solfato
MPS IX (Natowicz)	610492	Ialuronidasi	3p21.1	Acido ialuronico

Tabella I: Classificazione delle 11 mucopolisaccaridosi.

Le mucopolisaccaridosi sono malattie rare trasmesse con modalità autosomica recessiva, ad eccezione della MPS II che presenta un'ereditarietà di tipo X-linked recessiva. Le mutazioni alla base di queste malattie sono estremamente eterogenee, spesso private, ma uno o pochi alleli mutati possono essere predominanti in specifiche popolazioni. Per la maggior parte sono mutazioni puntiformi o piccole variazioni geniche, anche se nella MPS II si possono riscontrare grossi riarrangiamenti del gene e ampie delezioni. La correlazione della severità della malattia con il genotipo è a volte possibile, ma è generalmente difficile predire l'effetto di una mutazione sul fenotipo. Sono disponibili saggi enzimatici per la diagnosi di tutte le MPS sia postnatale, usando fibroblasti, leucociti o siero, sia prenatale attraverso amniocentesi o analisi dei villi coriali [Neufeld e Muenzer 2001].

Le mucopolisaccaridosi sono rare e sono riscontrabili in tutte le etnie, con una incidenza stimata tra i 3.4 e i 4.5 su 100000 nati vivi come si può vedere dalla **Tabella II** [Martin et al., 2008].

MPS	Australia	Australia occidentale	Germania	Irlanda del Nord	Paesi Bassi	British Columbia
I	0.90	0.93	0.69	1.66	1.19	0.79
II	0.62	0.31	0.64	0.71	0.67	0.30
III	1.37	1.71	1.57	0.35	1.89	0.29
IV	0.50 ^a	ND	0.38	1.30 ^a	0.36	0.49 ^a
VI	0.40	ND	0.23	ND	0.15	ND
VII	0.05	ND	0.00	ND	0.24	0.29
Totale	3.86	3.43	3.53	4.00	4.5	2.16

ND= dato non disponibile

^a il valore rappresenta solo MPS IV A

Tabella II: Incidenza delle MPS per 100.000 nati [Martin et al., 2008].

Oltre ai numerosi animali modello spontanei per le mucopolisaccaridosi, quali cani e gatti MPS I, capre MPS IIID, gatto MPS VI, cani e topi MPS VII, sono disponibili anche topi modello, ottenuti attraverso il knock-out del corrispondente gene murino.

Le caratteristiche biochimiche e patologiche di questi animali modello sono spesso simili a quelle delle loro controparti umane, ma la presentazione clinica può essere più attenuata. Gli animali modello per le MPS risultano estremamente utili per sviluppare e testare nuove terapie.

Le mucopolisaccaridosi sono state a lungo considerate potenzialmente adatte per la terapia con somministrazione di enzima esogeno. Il successo del trapianto di midollo osseo nell'alterare il decorso di alcune MPS ha dimostrato che l'enzima derivante dalle cellule ematopoietiche del donatore può ridurre l'accumulo di GAGs nei tessuti somatici del ricevente [Neufeld e Muenzer 2001]. Sfortunatamente, i lunghi tempi necessari per il ritrovamento di un donatore compatibile, la variabilità nei benefici neurologici e l'alto rischio associato al trapianto di midollo osseo limitano il valore di questo tipo di cura. La terapia enzimatica sostitutiva (Enzyme Replacement Therapy, ERT) è in attivo sviluppo ed è ad oggi disponibile per i pazienti di MPS I, MPS II, MPS VI; essa, nonostante i buoni risultati registrati nella riduzione degli accumuli sistemici, risulta del tutto inefficace sul comparto neurologico. La terapia genica, invece, è a stadi di studio meno avanzati [Clarke, 2008].

LA SINDROME DI HUNTER (MPS II)

La sindrome di Hunter (Mucopolisaccaridosi di tipo II; MPS II; OMIM 309900) è una patologia da accumulo lisosomiale a trasmissione X-linked recessiva causata dal deficit dell'enzima lisosomiale iduronato 2-solfatasi (IDS; EC3.1.6.13) coinvolto nel catabolismo dell'eparan- e dermatan-solfato. L'accumularsi dei metaboliti non degradati nei lisosomi colpisce quasi tutti i tipi di cellule, tessuti e diversi organi, inclusi quelli viscerali, lo scheletro, il tessuto connettivale e il sistema nervoso centrale (CNS). La sindrome di Hunter è una delle più comuni MPS e presenta un'incidenza di circa 1.3/100000 maschi nati vivi in Germania e Olanda [Poorthuis et al., 1999; Baehner et al., 2005]. Approssimativamente due terzi dei casi presentano coinvolgimento

neurologico, e rappresentano il fenotipo più severo della patologia [Wraith et al., 2008]. La malattia è dovuta a mutazioni nel gene dell'IDS localizzato in posizione Xq28. Sono state identificate più di 300 mutazioni, molte delle quali risultano private. La maggior parte delle mutazioni riscontrabili risultano puntiformi (70-80%), mentre mutazioni nelle regioni di splicing (10%) o riarrangiamenti del gene IDS (20%) sono meno comuni [Froissart et al., 2007; Manara et al., 2010].

ASPETTI CLINICI

La mucopolisaccaridosi di tipo II rappresenta una patologia multiorgano e multisistemica con un ampio spettro di caratteristiche cliniche.

Alla nascita i pazienti MPS II sono fenotipicamente normali, ma gradualmente con lo sviluppo manifestano cambiamenti morfologici, ostruzioni e infezioni ricorrenti delle vie respiratorie, irrigidimento articolare, organomegalia, displasia delle valvole cardiache, perdita di udito, degenerazione della cornea e, frequentemente, coinvolgimento neurologico.

Le varianti cliniche vanno da attenuate a severe e differiscono per età di insorgenza, presenza o assenza di peggioramento progressivo intellettivo e tempo di sopravvivenza. I pazienti affetti dalla forma severa risultano sintomatici tra 1 e 4 anni e la morte sopraggiunge frequentemente prima dei 15 anni, mentre i pazienti attenuati vengono, di solito, diagnosticati attorno ai 5-8 anni e possono sopravvivere fino all'età adulta. Il ritardo mentale e le difficoltà di deambulazione sono variabili e più frequentemente presenti nella forma severa [Neufeld e Muenzer, 2001; Martin et al., 2008; Burrow e Leslie, 2008; Manara et al., 2010].

Coinvolgimento neurologico

Il coinvolgimento neurologico nella sindrome di Hunter varia considerevolmente. Per definizione un paziente con fenotipo severo manifesta un coinvolgimento neurologico importante con compromissione delle abilità cognitive con conseguente ritardo e regressione dello sviluppo mentale. In tali pazienti sono comuni, nei primi anni di vita, problemi comportamentali, inclusi ostinazione, aggressività e iperattività

[Young et al., 1982]. Nei pazienti con fenotipo attenuato il coinvolgimento neurologico è limitato mentre presentano un progressivo deterioramento somatico.

I metodi clinici di valutazione della morfologia e della fisiologia del sistema nervoso centrale nella sindrome di Hunter non sempre forniscono una buona correlazione con lo status delle funzioni cognitive. In uno studio di risonanza magnetica (MRI) su 19 pazienti, la prevalenza dell'atrofia cerebrale, l'idrocefalo e le lesioni della materia bianca risultano significativamente maggiori nei pazienti severi con coinvolgimento cognitivo [Vedolin et al., 2007]. Nonostante ciò, la presenza di queste lesioni cerebrali non sono diagnostiche del danneggiamento delle abilità cognitive, in quanto, lesioni della materia bianca, allargamento dei ventricoli ed evidente atrofia cerebrale talvolta sono presenti anche in pazienti privi di coinvolgimento neurologico [Muenzer et al., 2009].

Si osserva anche compressione progressiva del midollo spinale con risultante mielopatia cervicale, che si manifesta inizialmente come attività ridotta e difficoltà ad alzarsi dalla posizione seduta, ma che qualora non trattata provoca un danno irreversibile alla corda. La compressione cervicale richiede un intervento tempestivo mediante decompressione.

Inoltre, nei pazienti affetti da MPS II è molto comune la sindrome del tunnel carpale con una rapida progressione dei sintomi. La decompressione chirurgica del nervo mediano nelle fasi iniziali permette un miglioramento parziale o totale dei sintomi nella maggior parte dei casi [Wraith et al., 2008].

Diagnosi

I sintomi precedentemente descritti non sono specifici della sola sindrome di Hunter e non tutti sono presenti in ogni paziente. Pertanto, in assenza di una familiarità della malattia, il riconoscimento diagnostico precoce è difficile. Solo pochi dei sintomi della sindrome di Hunter sono presenti alla nascita e, solo dopo alcuni anni, iniziano ad emergere in modo più consistente.

Nelle MPSs si riscontra elevata escrezione di glicosaminoglicani nelle urine e tale caratteristica viene sfruttata nella diagnosi biochimica. L'analisi dei livelli di GAGs urinari è in grado di confermare il sospetto di MPS; l'eccesso di dermatan- ed eparan-

solfato, tuttavia, può essere indice di MPS I, MPS II o MPS VII [Neufeld e Muenzer, 2001]. Per giungere alla diagnosi di sindrome di Hunter è necessario perciò passare al dosaggio dell'attività enzimatica dell'IDS.

L'IDS è presente in tutte le cellule (ad eccezione degli eritrociti maturi), pertanto i saggi comunemente utilizzati sono basati su colture di fibroblasti, leucociti, plasma o siero. Attività enzimatica ridotta o assente dell'enzima IDS nei maschi è considerata diagnosi di sindrome di Hunter, purché un'altra solfatasi misurata presenti un'attività enzimatica normale, in modo da escludere un deficit multiplo di solfatasi. Il dosaggio enzimatico, però, non può predire la severità del fenotipo né identificare le portatrici. All'analisi enzimatica fa seguito la ricerca delle mutazioni, usata per confermare e specificare la sindrome di Hunter. L'analisi genetica è, inoltre, l'unica via sicura per identificare le portatrici e può essere utilizzata per la diagnosi prenatale.

APPROCCI TERAPEUTICI

Fino a poco tempo fa, la cura della sindrome di Hunter era per lo più palliativa e focalizzata sul trattamento dei sintomi.

Il trapianto di midollo, o più recentemente di cellule staminali ematopoietiche (HSCT), non ha dimostrato efficacia significativa nella MPS II così come per la MPS I; inoltre, sono state riscontrate numerose problematiche legate alla difficoltà di trovare donatori istocompatibili e all'elevata morbilità della procedura.

Terapia Enzimatica Sostitutiva

Più recentemente è stata introdotta la terapia enzimatica sostitutiva (ERT) che sembra dare buoni risultati [Wraith et al., 2008].

Il trattamento dei pazienti affetti dalla sindrome di Hunter prevede la somministrazione settimanale per via endovenosa di Elaprase® (idursulfase; Shire Human Genetic Therapies Pharmaceuticals, Cambridge, MA, USA), forma purificata dell'enzima lisosomiale iduronato 2-solfatasi.

L'idursulfasi, peso molecolare di circa 76 kD, viene secreta come una glicoproteina costituita da 525 aminoacidi e contiene 8 siti di glicosilazione all'N, che sono occupati

da catene oligosaccaridiche di tipo complesso, ibrido e ad alto contenuto di mannosio. I residui di mannosio-6-fosfato (M6P) sulle catene oligosaccaridiche rendono possibile il legame specifico dell'enzima ai recettori dell'M6P sulla superficie cellulare e quindi l'internalizzazione dell'enzima nella cellula verso i lisosomi intracellulari ove avviene il catabolismo dei GAGs accumulati.

L'Elaprase è stato approvato a luglio del 2006 negli Stati Uniti e a gennaio del 2007 nell'Unione Europea per il trattamento della sindrome di Hunter, con una dose di 0.5 mg/kg/settimana. Questa approvazione si è basata sui risultati di un trial clinico randomizzato, placebo-controllo, effettuato in doppio cieco che ha dimostrato un beneficio clinico nei pazienti trattati con Elaprase rispetto a pazienti che hanno ricevuto il placebo [Muenzer et al., 2006; Burrow e Leslie, 2008].

Le sperimentazioni cliniche hanno mostrato miglioramenti significativi a livello epatico, splenico, della deambulazione e la stabilizzazione del quadro respiratorio e cardiaco. Tuttavia, un limite importante dell'ERT riguarda la cura dei pazienti con coinvolgimento neurologico, dato che l'enzima non attraversa la barriera emato-encefalica, risultando così inefficace per la cura della forte compromissione del sistema nervoso centrale.

Un altro problema, dal quale anche questa terapia non si discosta, è che l'efficacia non può essere seguita e quantificata con uno specifico marker biologico. Al momento è valutata esclusivamente tramite test clinici (diminuzione dell'organomegalia, incremento nella capacità motoria, ecc) e tramite il dosaggio dei GAGs urinari che, tuttavia, non riflettono sempre l'efficacia terapeutica.

Genisteina

Recentemente è stato proposto un nuovo trattamento, basato sull'uso della genisteina, per le mucopolisaccaridosi prive di terapia e, dove disponibile, in supporto all'ERT [Jakobkiewicz-Banecka et al. 2009].

La genisteina è un isoflavone derivato dalla soia con struttura molecolare simile a quella degli estrogeni endogeni e si lega, sebbene con minore affinità rispetto all'estradiolo, ai recettori alfa e beta per gli estrogeni. E' una molecola nota da anni proprio per il suo effetto estrogenico, per il quale è stata utilizzata per la terapia del cancro, nella profilassi e terapia dell'osteoporosi, delle patologie cardio-vascolari e

della menopausa [Gu et al., 2005; Wang et al., 2006; Vera et al., 2007; D'Anna et al., 2009].

L'effetto di riduzione sulla sintesi dei GAGs indotto dalla genisteina non sembra tuttavia legato al suo effetto estrogenico, quanto piuttosto alla sua capacità di ridurre l'espressione di geni coinvolti nella loro sintesi, effetto che la molecola determinerebbe tramite inibizione dell'attività proteinchinasica del recettore per l'epidermal growth factor (EGF).

Questo effetto è stato di recente valutato *in vitro* e *in vivo* con risultati mediamente positivi anche nel comparto neurologico [Piotrovskaya et al., 2006; Friso et al., 2010].

Per tali ragioni, nonostante questa molecola sia circondata da grandi controversie, la comunità scientifica e, soprattutto, le associazioni dei familiari dei pazienti MPS II hanno fatto pressione affinché venisse proposta la somministrazione della genisteina nei bambini malati, nella maggior parte dei casi affiancata alla terapia di sostituzione enzimatica.

Trattamento del coinvolgimento neurologico

Nonostante la buona efficacia dell'ERT nel trattamento dei compartimenti periferici, esso risulta del tutto inefficace nel miglioramento del coinvolgimento neurologico. Infatti, gli enzimi ricombinanti attualmente utilizzati sono disponibili con formulazioni, dosaggi e somministrazione non appropriati ad oltrepassare la barriera emato-encefalica. Devono pertanto essere studiati altri metodi di trattamento che comprendano, oltre a nuove posologie, anche via alternative di somministrazione dell'enzima.

Negli ultimi anni sono stati eseguiti alcuni studi basati sulla somministrazione intratecale (IT) degli enzimi ricombinanti negli animali modello per la MPS I e MPS VI con risultati positivi sia in termini di sicurezza sia di efficacia soprattutto nella riduzione degli accumuli a livello delle meningi [Kakkis et al., 2004; Dickson et al., 2010; Auclair et al., 2010].

In ambito clinico tale procedura è stata applicata solo in due pazienti, rispettivamente con sindrome di Hurler (MPS I) e di Maroteaux-Lamy (MPS VI), nei quali è stato riscontrato un miglioramento della compressione cerebro-spinale e conseguente

maggiore mobilità; tale somministrazione non ha portato a eventi avversi duraturi dovuti alla terapia [Munoz-Rojas et al., 2008 e 2010].

Tale approccio terapeutico non è stato ancora sperimentato nella MPS II, poichè al momento non sono ancora state definite le procedure e l'enzima ricombinante adeguati alla somministrazione diretta nel cervello, e infusioni intracerebrali ripetute non sono praticabili in quanto sottopongono il paziente a grandi stress e rischi pre- e post-operatori dovuti all'anestesia e all'eventualità di incorrere in infezioni.

Un trattamento più efficace del coinvolgimento neurologico potrebbe ottenersi modificando gli attuali farmaci in modo da favorirne il passaggio attraverso la barriera emato-encefalica, e/o abbinando la terapia enzimatica sostitutiva ad altri approcci, quali la terapia genica/cellulare.

Le mucopolisaccaridosi, compresa la sindrome di Hunter, sono buone candidate per la terapia genica in quanto monogeniche e gli studi sugli animali modello stanno dando risultati promettenti [Eto et al., 2004; Herati et al., 2008; Wang et al., 2009; Polito e Cosma, 2009; Polito et al., 2010].

La scelta del vettore e la via di somministrazione adeguati, la necessità di impiego degli immunosoppressori sono solo alcune tra le numerose problematiche che necessitano di essere ancora risolte prima che si possano effettuare i primi trial clinici umani di terapia genica per le MPS.

PATOFISIOLOGIA

La patofisiologia della sindrome di Hunter è tuttora in gran parte sconosciuta, infatti non è ancora noto l'esatto meccanismo che lega l'accumularsi di GAGs al conseguente decorso della malattia. Scoprire come questi depositi compromettano le funzioni cellulari risulta indispensabile per meglio comprendere la patofisiologia che è alla base delle mucopolisaccaridosi e progettare nuovi approcci terapeutici. In generale si ritiene che i depositi di mucopolisaccaridi possano interferire nel trafficking cellulare, nelle funzioni dei recettori delle cellule, nelle cascate di segnale e che possano comportare un'alterazione della matrice extracellulare.

Studi effettuati su animali modello MPS VI e VII hanno evidenziato l'infiammazione come effetto secondario ai depositi di GAGs. Nelle mucopolisaccaridosi gli accumuli extracellulari portano all'attivazione non fisiologica di recettori di segnali di traduzione come il Toll like receptor 4 (TLR4). Ciò determina la secrezione di una varietà di citochine proinfiammatorie (IL-1 β , TNF- α e TGF- β), che stimolano la risposta del sistema immunitario innato, la cui espressione aumenta con l'età degli animali e con lo sviluppo della patologia [Simonaro et al., 2008]. Valutazioni eseguite mediante microarray, Northern blot e Real-Time PCR in cervelli di modelli murini per la MPS I, MPS III B e MPS VII hanno confermato, anche in questo organo, l'attivazione delle vie di segnale implicate nell'infiammazione e nell'apoptosi. L'infiammazione cerebrale è caratterizzata da astrocitosi e microgliosi, dall'aumentata produzione di citochine e di chemochine, dalla down-regolazione di geni neurotrofici (brain-derived neurotrophic factor, BDNF) e componenti associati alla mielina (myelin basic protein, MBP; myelin-associated oligodendrocytic basic protein, MOBP) che suggeriscono neurodegenerazione e demielinizzazione [Ohmi et al., 2003; Villani et al., 2007, Richard et al., 2008; DiRosario et al., 2009].

In fibroblasti derivati da pazienti con sindrome di Hurler, invece, si è osservato che gli elevati livelli di eparan-solfato extracellulare interferiscono con il fibroblast growth factor 2 (FGF-2) impedendone il legame con il proprio recettore. Ne deriva l'indebolimento della via di segnale attivata dall'FGF-2, che permette la proliferazione e protezione di diversi tipi cellulari tra i quali i neuroni e i precursori cellulari neuronali, ciò potrebbe contribuire alla neurodegenerazione riscontrata nei pazienti MPS I e probabilmente anche in altre MPS [Ballabio e Gieselmann, 2009].

Inoltre, uno studio *in vitro* sulle MPS II, MPS III A e la malattia da Deficit Multiplo di Solfatasi (MSD) ha dimostrato l'implicazione del fattore trascrizionale EB (TFEB) nella biogenesi dei lisosomi e nell'aumento della degradazione di molecole complesse, il quale potrebbe diventare un potenziale target terapeutico per migliorare la condizione cellulare nelle malattie da accumulo lisosomiale [Sardiello et al., 2009].

Oggi la priorità dell'attività di ricerca è realizzare nuovi metodi per la cura delle LSDs, con particolare attenzione al sistema nervoso centrale, mediante la comprensione

della patofisiologia, del decorso del coinvolgimento neurologico e l'identificazione di marcatori di patologia.

L'ANIMALE MODELLO PER LA MPS II

Il primo animale modello MPS II caratterizzato nel 1998 era un cane Labrador Retriever spontaneo che presentava tutte le caratteristiche tipiche di tale patologia [Wilkerson et al., 1998].

I modelli animali spontanei, anche di altre MPS, sono stati inizialmente utilizzati per lo sviluppo della terapia enzimatica sostitutiva; tuttavia, le loro dimensioni e la durata di vita hanno limitato il loro utilizzo per gli studi di efficacia a lungo termine o per la valutazione di nuovi approcci terapeutici. Ben presto per questi motivi, per i lunghi tempi di riproduzione e per gli elevati costi di mantenimento sono stati sviluppati i più versatili modelli murini.

Il modello murino a tutt'oggi utilizzato per la MPS II è stato per la prima volta descritto nel 1999. Il topo knock-out (ceppo C57BL/6) per l'iduronato 2-solfatasi è stato ottenuto mediante la sostituzione, per ricombinazione omologa, dell'esone 4 e parte dell'esone 5 con il gene per la resistenza alla neomicina [Muenzer, 1999]. L'animale è stato successivamente descritto e caratterizzato in diverse pubblicazioni [Cardone et al., 2006; Garcia et al., 2007; Polito e Cosma, 2009].

Alla nascita i topi knock-out (KO) sono indistinguibili dai wild-type. A 4 settimane di età si ha già un aumento del peso del fegato e l'analisi radiografica mostra sclerosi e ingrossamento delle ossa del cranio. A partire dalle 12 settimane cominciano a essere visibili altri segni della patologia che progrediscono con il tempo e presentano un'elevata variabilità inter-animale. Si osserva alopecia, deformità delle articolazioni e ispessimento delle ossa degli arti e delle dita, che assumono spesso una forma incurvata simile al fenotipo umano. La deformità delle giunture tibio-tarsali degli arti inferiori porta ad anchilosi nel 6%, 25% e 37.5% dei topi KO rispettivamente a 3, 6 e 12 mesi [Garcia et al., 2007]. A 8 mesi di età si notano deformità del muso con prognatismo dovuto a dispostosi (**Figura II**).

Le anomalie scheletriche interessano tutte le altre ossa del corpo dando progressivi problemi di deambulazione con cambiamenti funzionali osservabili radiograficamente nel decorso clinico dell'MPS II. A 40 settimane i topi presentano declino della mobilità, e nonostante risultino più piccoli dei fratelli wild-type, il peso relativo degli organi appare significativamente aumentato [Cardone et al., 2006]. L'aspettativa di vita dei topi KO, circa un anno, è considerevolmente ridotta paragonata alle circa 120 settimane dei topi wild-type [Garcia et al., 2007].



Figura II: Fotografia del muso del topo wild-type (14 mesi) e del modello per la sindrome di Hunter (9 e 14 mesi).

L'attività enzimatica iduronato 2-solfatasi è nulla nel siero e in tutti gli organi analizzati: fegato, rene, milza, cuore e cervello. I GAGs urinari sono elevati già a 4 settimane e rimangono alti per tutta la durata della vita dell'animale. A 7 settimane la concentrazione dei GAGs è elevata in fegato, rene, milza e cuore. Gli accumuli nella pelle, nel polmone e nel cervello risultano il doppio rispetto l'animale wild-type mentre non si osservano differenze tra i due genotipi nel muscolo scheletrico. Già a 4 mesi il topo KO presenta neurodegenerazione con astrogliosi e vacuolizzazione cellulare in diverse regioni del cervello: corteccia, ippocampo, talamo, cervelletto e tronco encefalico, mentre gli accumuli di GAGs sono evidenti nei plessi coroidei e nelle meningi. A 60 settimane si riscontrano anche vacuolizzazione del tessuto delle valvole cardiache e necrosi neuronale [Cardone et al., 2006; Garcia et al., 2007; Polito e Cosma, 2009].

Recentemente un gruppo coreano ha sviluppato un altro modello murino, il quale assomiglia molto al precedente che ha la peculiarità di manifestare prima le caratteristiche morfologiche fenotipiche [Jung et al., 2010].

Le analogie di questi animali modello con il fenotipo e la progressione clinica dei pazienti con sindrome di Hunter li rendono molto utili per lo studio della patofisiologia della MPS II e per lo sviluppo di nuove terapie.

L'ANALISI DEL TRASCRITTOMA

Un enigma che avvolge la biologia molecolare è come lo stesso corredo genetico delle cellule possa portare a tipi cellulari diversi, ognuna delle quali con un proprio ruolo preciso e definito nel funzionamento degli organismi multicellulari. Questa diversità fenotipica è legata al fatto che differenti tipi cellulari esprimono set diversi di geni (trascrittoma) portando le cellule a diverso destino e funzioni. Per decenni l'interesse primario, da parte dei biologi, è stato quello di definire una correlazione di tale destino e funzioni cellulari con il profilo di espressione genica e quindi la valutazione del trascrittoma.

Più precisamente per trascrittoma si intende l'insieme del set completo di trascritti di una cellula e la loro quantità in una specifica condizione fisiologica o stadio dello sviluppo. Comprendere il trascrittoma permette di interpretare gli elementi funzionali del genoma e rilevare i costituenti molecolari delle cellule e dei tessuti. Inoltre, essa consente di comprendere la patogenesi e lo sviluppo delle malattie. Gli scopi principali della trascrittomica sono catalogare tutte le specie di trascritti presenti (compresi mRNA, RNA non codificanti e small RNAs), determinare la struttura trascrizionale dei geni (sito di inizio della trascrizione, pattern di splicing e modifiche post-trascrizionali) e quantificare la variazione dei livelli di espressione di ogni trascritto durante lo sviluppo e nelle diverse condizioni fisiologiche [Wang et al., 2009].

Nel 1977 fu sviluppato il primo metodo volto a individuare RNA specifici: il **Northern blot**. Tale tecnica risultò subito molto complessa in quanto necessita di sonde radioattive, una grande quantità di RNA e permette la distinzione solo di pochi geni conosciuti. Successivamente, lo sviluppo della PCR quantitativa con retrotrascrizione

(RT-qPCR) ha permesso di ridurre la quantità iniziale di RNA necessario e ha facilitato l'identificazione di un maggiore numero di geni, pur mantenendosi lontana dall'idea di una analisi ampia del trascrittoma.

Lo sviluppo dei **microarrays** ha soppiantato l'approccio basato sull'analisi del singolo gene per una caratterizzazione simultanea dei livelli di espressione di centinaia di trascritti conosciuti o putativi. Il loro utilizzo fu ampliato ad aspetti della trascrittomica come la determinazione degli RNAs non codificanti, dei polimorfismi di singolo nucleotide (SNPs) e degli eventi di splicing alternativo. Grazie ai costi relativamente bassi e all'elevata efficienza sono ancora utilizzati in approcci di trascrittomica in tutto il mondo. Nonostante la capacità di misurare simultaneamente i livelli di espressione di centinaia di geni, i microarrays presentano alcuni limiti. Innanzitutto, la tecnica non consente di identificare trascritti non noti, inoltre è un metodo indiretto nel quale la quantità del trascritto viene dedotta dall'intensità di ibridazione ottenendo risultati con un intenso background dovuto alle cross-ibridazioni, e ciò rende difficile la comparazione tra esperimenti diversi richiedendo complicati sistemi di normalizzazione.

L'analisi del trascrittoma mediante costruzione di librerie di cDNA attraverso cloni è diventata un metodo alternativo ai microarrays. L'elevato costo del metodo Sanger, usato in questa tecnica, la complessità di associare i cloni tra di loro e quindi l'incapacità di caratterizzare l'intero trascrittoma in specie multicellulari ha reso però inapplicabile, in modo routinario, il sequenziamento del cDNA a catena intera (**FLcDNA**). L'introduzione dell'analisi seriale dell'espressione genica (**SAGE**) ha rappresentato un passo avanti importante nelle tecniche di sequenziamento del trascrittoma facilitando e velocizzando l'uso della procedura di Sanger per la rilevazione del profilo di espressione genica. La SAGE permette di individuare nuovi trascritti e nuove isoforme alternative di splicing, quantificando i trascritti e rendendo più semplice il confronto tra i diversi campioni. Tuttavia, essa prevede una procedura di creazione dei cloni molto laboriosa e costosa con produzione di sequenze molto corte (14-21bp) difficili da discriminare nel caso di trascritti con sequenze codificanti simili.

Con il completamento del sequenziamento del genoma umano e dei maggiori organismi modello di importanza biomedica, il ri-sequenziamento è diventato procedura all'avanguardia, portando alla realizzazione di una serie di nuove tecnologie di sequenziamento definite “**sequenziatori di nuova generazione**”, con un rapporto costo/efficacia migliore del sequenziamento con metodo Sanger. Le tecnologie di nuova generazione attualmente disponibili in commercio sono quattro: Roche/454™, Illumina®, Applied Biosystems SOLiD™, e il più recente Helicos HeliScope™, i quali forniscono un elevatissimo numero di piccole sequenze (reads), di gran lunga superiore alla resa ottenibile con il sequenziatore Sanger (**Tabella III**).

Piattaforma	ABI PRISM™ 3730xl DNA Analyzer	Roche (454)™ FLX Titanium	Illumina® Genome Analyzer	ABI SOLiD™	HeliScope™
<i>Chimica di sequenziamento</i>	Sequenziamento Sanger automatico	Piro-sequenziamento su supporto solido	Sequenziamento mediante sintesi con terminatori reversibili	Sequenziamento mediante ligazione	Sequenziamento mediante sintesi con terminatori virtuali
<i>Metodo di amplificazione dello stampo dell'acido nucleico</i>	Amplificazione in vivo con uso di cloni	PCR per emulsione “Emulsion PCR”	PCR a ponte “Bridge PCR”	PCR per emulsione “Emulsion PCR”	Nessuna (molecola singola)
<i>Lunghezza delle reads</i>	700-900 bp	400 bp	35-150 bp	50-75 bp	25-35 bp
<i>Throughput</i>	0.7-2.1 Mb/run	0.4-0.6 Gb/run	4-38 Gb/run	60-300 Gb/run	21-35 Gb/run
<i>Sito web della compagnia fabbricante</i>	http://www.appliedbiosystems.com	http://www.roche-applied-science.com	http://www.illumina.com	http://www.appliedbiosystems.com	http://www.helicosbio.com

Tabella III: Sequenziatori in commercio usati per sequenziamento del trascrittoma e loro caratteristiche principali.

I sequenziatori di nuova generazione hanno permesso di abbassare enormemente i costi di sequenziamento e di ridurre la complessità sperimentale, così come di aumentare la copertura dei trascritti eseguendo più velocemente l'analisi del sequenziamento del trascrittoma rendendolo disponibile in breve tempo. L'applicazione può essere ampliata al ri-sequenziamento di diversi genomi e in

particolare di quello umano, in modo da meglio comprendere le variazioni genetiche tra individuo sano e patologico ed effettuare il sequenziamento dell'intero trascrittoma "RNA-Seq" [Morozova et al., 2009].

L'RNA-Seq è forse una delle più complesse applicazioni dei sequenziatori di nuova generazione. I livelli di espressione, gli splicing alternativi, l'espressione di alleli specifici e la fusione tra i trascritti forniscono informazioni molto importanti quando si comparano campioni nello studio di patologie. Sono molte le caratteristiche che differenziano l'RNA-Seq mediante le piattaforme di nuova generazione dalle altre tecnologie. In primo luogo esso non si basa su un metodo di ibridazione e non è vincolato dalla conoscenza a priori del trascritto, permettendo così l'identificazione, la caratterizzazione e la quantificazione di nuove isoforme. Non si avvale più della tecnica di clonaggio *in vivo* la quale viene sostituita dall'amplificazione mediante PCR, eliminando il rischio di sequenziare erroneamente parti del genoma dell'ospite. Inoltre, esso consente di determinare il gene corretto e i suoi confini, grazie alla risoluzione di ogni singolo nucleotide, e identificare gli SNPs. Gli altri vantaggi dell'RNA-Seq, confrontato con i microarrays, sono il basso background, l'assenza di un limite superiore per la quantificazione, e di conseguenza dei range più ampi di livelli di espressione nei quali i trascritti possono essere rilevati. I dati ottenuti con l'RNA-Seq mostrano alti livelli di riproducibilità sia nelle repliche tecniche sia biologiche. Diversi studi hanno chiaramente dimostrato i vantaggi dell'utilizzo dell'RNA-Seq. Molti gruppi di ricerca sono stati in grado di quantificare con precisione trascritti già conosciuti, di scoprire nuove regioni trascritte, di caratterizzare trascritti antisenso, di identificare splicing alternativi con nuove combinazioni di esoni noti o di nuovi esoni trascritti, di valutare l'espressione di elementi ripetuti, di analizzare un ampio numero di SNPs conosciuti e di altri possibili candidati, così come di identificare fusione di trascritti e altre nuove categorie di RNA [Costa et al., 2010].

Nell'ambito del sequenziamento del trascrittoma la maggior parte degli attuali studi sono stati effettuati utilizzando la tecnologia Illumina[®]. Tale scelta sembra essere dovuta all'iniziale ampia diffusione di questa piattaforma, più economica, rispetto alla pioniera Roche/454™, che tuttavia rimane l'eccellenza per il ri-sequenziamento. L'Applied Biosystems SOLiD™ è stato commercializzato solo successivamente

all'Illumina® e attualmente si sta imponendo soprattutto per quanto riguarda l'RNA-Seq, grazie alla superiore affidabilità di sequenza, nonché accuratezza e robustezza del sistema.

HeliScope e l'ultimo sviluppato, non ancora in commercio, Pacific Biosciences, basati sul sequenziamento di ogni singola molecola, potrebbero diventare il futuro del sequenziamento del DNA/RNA, ma richiedono ancora del tempo per convincere il mercato.

SCOPO

La mucopolisaccaridosi di tipo II è una malattia genetica rara, X-linked recessiva, causata dal deficit dell'enzima lisosomiale iduronato-2-solfatasi (IDS), coinvolto nella degradazione di dermatan- ed eparan-solfato. L'accumulo che ne deriva produce una grave disfunzione cellulare e tissutale in diversi organi. La patofisiologia della sindrome di Hunter è tuttora in gran parte sconosciuta, infatti non è ancora noto l'esatto meccanismo che lega l'accumularsi di glicosaminoglicani (GAGs) al conseguente decorso della malattia. Nonostante tali accumuli di GAGs risultino relativamente bassi nel sistema nervoso centrale, se paragonati agli altri organi, nelle forme più severe di MPS II si evidenzia un grave coinvolgimento neurologico e tale discordanza non è ancora del tutto compresa.

Il trattamento più efficace al momento disponibile per la cura della malattia, la terapia enzimatica sostitutiva, è in grado di determinare un miglioramento del fenotipo patologico in diversi distretti, ma risulta del tutto inefficace sul comparto neurologico, poiché l'enzima non può attraversare la barriera emato-encefalica. E' auspicabile che una comprensione della patologia a livello del sistema nervoso centrale possa consentire lo sviluppo di approcci terapeutici innovativi diretti al trattamento della compromissione neurologica.

Lo scopo di questa tesi è di analizzare l'espressione differenziale dei trascritti nella corteccia cerebrale e nel mesencefalo del modello murino per la MPS II mediante tecnologia RNA-seq con il sequenziatore SOLiD™ 3 Plus (Applied Biosystems). Tale comparazione tra l'animale sano e il modello patologico potrebbe consentire l'identificazione dei meccanismi patogenetici del coinvolgimento neurologico nella sindrome di Hunter, e permettere, eventualmente, di individuare efficaci marcatori di patologia.

L'applicazione della tecnologia SOLiD™ potrebbe così consentire di delucidare il decorso neurologico della malattia, parzialmente traslabile in ambito pediatrico, permettendo di identificare tempestivamente la gravità della patologia e successivamente di sviluppare nuovi approcci terapeutici più mirati.

MATERIALI E METODI

L'ANIMALE MODELLO PER LA MPS II

I topi impiegati per questi esperimenti appartengono al ceppo C57BL/6. I topi wild-type sono stati acquistati da Charles River Italia, mentre i knock-out (KO, IDS-KO) sono stati gentilmente forniti dal Prof. Muenzer (Univ. of North Carolina, Chapel Hill, USA). Le caratteristiche essenziali dell'animale sono già state fornite nella sezione introduttiva della presente tesi.

Gli animali sono stati mantenuti in allevamento con cicli di luce/buio, abbeveraggio e nutrizione in conformità con le leggi nazionali e internazionali vigenti. Anche le procedure applicate per il trattamento degli animali, sia nella fase di valutazione comportamentale che in corso di sacrificio, erano conformi all'attuale normativa.

PIANO SPERIMENTALE

Il piano sperimentale prevede la valutazione del profilo di espressione genica mediante tecnologia RNA-Seq di due macroaree del cervello, corteccia e mesencefalo, dell'animale modello per la mucopolissacaridosi di tipo II.

A tale scopo sono stati identificati e sacrificati, mediante dislocazione cervicale, 7 animali IDS-KO di 9 mesi di età e 7 rispettivi fratelli wild-type (WT).

GENOTIPIZZAZIONE DEGLI ANIMALI

L'estrazione del DNA genomico è stata effettuata utilizzando il kit commerciale Wizard SV Genomic DNA Purification System (Promega) a partire da porzioni di code di 0.5 – 0.7 cm.

L'analisi del genotipo è stata effettuata mediante amplificazione di 2 µl di DNA genomico con PCR multiplex. Sono stati utilizzati i seguenti 3 primer:

1. 5'- GGGGAGGAGCTACTTGCATAGTTG-3'
2. 5'- AGGTGGAAAAGACCAGCTATAT-3'
3. 5'-AAAAGAGGACTGCGTGTGGG-3'

Un primer si appaia sia alla sequenza wild-type del gene IDS che alla sequenza dell'allele knock-out (ottenuto mediante inserzione del gene per la resistenza alla neomicina); ciascuno degli altri due primer è specifico per uno dei due alleli. I primer sono disegnati in modo tale da poter ottenere due amplificati di lunghezza diversa, 152 bp e 181 bp per l'allele WT e KO rispettivamente, presenti entrambi nell'animale eterozigote; questi due amplificati possono pertanto essere separati mediante elettroforesi su gel di agarosio NuSieve (NuSieve GTG Agarose - Biospa) al 2.5%.

TEST COMPORAMENTALI

7 topi IDS-KO e 7 fratelli wild-type, di 9 mesi di età, sono stati sottoposti all'Open Field test e al Footprint test al fine di valutare diversi parametri che, a causa delle alterazioni scheletriche e del coinvolgimento neurologico della malattia, potrebbero essere alterati nei topi patologici.

Open Field Test (OF)

La capacità motoria e il comportamento esplorativo sono stati valutati sottoponendo gli animali all'Open Field test per 10 minuti in una arena di 60 cm X 60 cm X 40 cm con un fondo in plexiglass. Tutti gli animali sono stati filmati e le registrazioni sono state analizzate utilizzando il programma ANY-maze (Stoelting, IL, USA).

Per ciascun animale sono stati contati anche il numero di boli fecali prodotti, come indice di ansia [Gould et al., 2009].

Footprint test

Le anomalie nella camminata sono state determinate colorando le zampe degli animali con una pittura lavabile non tossica. Per le zampe anteriori è stato utilizzato il colore nero mentre per le posteriori il rosso, in modo da poter valutare l'overlapping

che ci si aspetta nei topi wild-type. Gli animali, quindi, sono stati fatti correre su un foglio di carta bianco attraverso un corridoio largo 10 cm e lungo 50 cm. Sono state valutate la lunghezza del passo, la sovrapposizione delle zampe anteriori e posteriori e la linearità della camminata.

Analisi statistica

In entrambi i test la barra di errore rappresenta la SEM (Standard Error of the Mean) e l'analisi statistica è stata effettuata applicando un test-t di Student. I confronti con $p \leq 0.05$ sono stati considerati statisticamente significativi.

PREPARAZIONE CAMPIONI TESSUTO CEREBRALE

L'animale è stato sacrificato mediante dislocazione cervicale. Il cervello una volta asportato dalla scatola cranica è stato posizionato sopra una piastra petri, mantenuta in ghiaccio secco, per effettuare la suddivisione delle diverse aree. Prelevati i bulbi olfattivi, la corteccia, il mesencefalo, i nuclei della base, il midollo allungato e il cervelletto (di questi saranno oggetto della presente indagine soltanto corteccia e mesencefalo), le macroaree sono state riposte in provette RNasi free contenenti 700 μ l di RNA later[®] Tissue Collection: RNA stabilization solution (Ambion[®], Applied Biosystems).

In seguito ad incubazione overnight a 4°C dei tessuti cerebrali, in modo da consentire l'infiltrazione dello stabilizzante nelle regioni più interne del tessuto, l'eccesso di RNA later[®] è stato rimosso e i campioni conservati a -80°C fino al momento dell'estrazione dell'RNA.

ESTRAZIONE DELL'RNA TOTALE

L'estrazione dell'RNA totale è stata eseguita utilizzando materiale RNasi-free e dedicato, soluzioni preparate con appositi reagenti, superfici e strumenti sono stati trattati con RNaseZAP[®] (Sigma[®]) per evitare contaminazioni con RNasi esogene.

L'estrazione è stata effettuata seguendo il protocollo standard del reagente TRIzol[®] Reagent (Invitrogen[™]).

I campioni, prima di essere congelati a -80°C, sono stati valutati per concentrazione allo spettrofotometro NanoDrop ND 1000 (Thermo Scientific) diluendo l'RNA totale 1:20 in H₂O RNasi-free.

L'RNA totale è stato sottoposto ad una preliminare valutazione qualitativa mediante corsa elettroforetica in gel di agarosio. È stato utilizzato un gel all'1% di agarosio SeaKem GTG[®] in TBE 0.5% ed Etidio Bromuro. 1µg di RNA è stato diluito 1:10 con il loading buffer (RNA Sample Loading Buffer, Sigma[®]) per un volume totale di 10 µl e fatto denaturare 10 minuti a 65°C. I campioni, in seguito, sono stati raffreddati per 1 minuto in ghiaccio prima di essere caricati nei pozzetti e fatti correre un'ora a 60 V.

Successivamente i campioni ritenuti idonei sono stati considerati per una determinazione più dettagliata dell'integrità dell'RNA totale mediante RNA 6000 Nano chip - 2100 Bioanalyzer (Agilent Biotechnologies), che fornisce, oltre alla bande della corsa elettroforetica e all'intensità dei picchi ribosomali 18S e 28S, anche un valore numerico, il RIN (RNA integrity number). In **Figura III** è riportato un esempio di dato ottenuto con tale procedura.

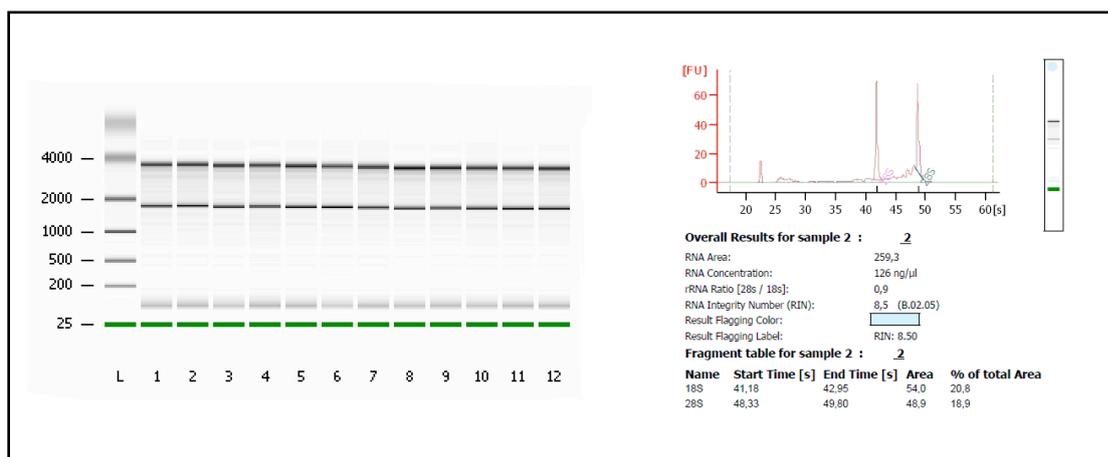


Figura III: Output dei dati ottenuti dall'analisi dell'RNA mediante il 2100 Bioanalyzer. A sinistra l'immagine virtuale del tipico gel di agarosio di alcuni campioni di corteccia e di mesencefalo e a destra l'elettroferogramma di uno di questi.

ANALISI MEDIANTE TECNOLOGIA RNA-Seq

I campioni da analizzare mediante tecnologia RNA-Seq sono stati preparati effettuando un pool dell'RNA come da tabella seguente (**Tabella IV**):

Numero animali per pool	Genotipo	Area cerebrale	ID campione	Concentrazione RNA totale per pool
7	knock-out (KO)	Corteccia (Cx)	Cx_KO	70 µg
		Mesencefalo (M)	M_KO	70 µg
7	wild-type (WT)	Corteccia (Cx)	Cx_WT	70 µg
		Mesencefalo (M)	M_WT	70 µg

Tabella IV: Pool di RNA delle aree cerebrali analizzate mediante RNA-Seq.

La procedura è stata eseguita dai ricercatori del Cribi dell'Università di Padova attraverso l'utilizzo della tecnologia di seconda generazione SOLiD™ 3 Plus System (Sequencing by Oligo Ligation and Detection) (Applied Biosystems).

Il SOLiD™ Whole Transcriptome Analysis (WTA) Kit consente la conversione delle molecole di RNA in una libreria di cDNA che può quindi essere sequenziata attraverso il SOLiD System, permettendo così di analizzare il completo set di trascritti della cellula, compresi gli RNA a funzione ignota.

Purificazione dell'mRNA

A partire dall'RNA totale, la purificazione dell'mRNA può essere effettuata mediante sottrazione dell'RNA ribosomale o mediante arricchimento in poli(A). In questo caso è stata scelta la seconda strategia, utilizzando il Dynabeads® mRNA Purification Kit (Invitrogen™). Questo protocollo si basa sul legame tra i residui di poli(A) all'estremità 3' della maggior parte delle molecole di mRNA e i residui di oligo (dT)25 legati covalentemente alla superficie delle Dynabeads; le molecole di RNA prive della coda di poli(A) non si ibridano con le biglie e quindi vengono lavate via. L'arricchimento in mRNA è stato valutato mediante l'RNA 6000 Pico chip - 2100 Bioanalyzer (Agilent Biotechnologies).

Preparazione della library ed emulsion PCR

Dopo lo step di purificazione, l'mRNA è stato frammentato mediante RNasi III in corte sequenze di circa 50-150 nucleotidi e purificato utilizzando l'Ambion flashPAGE™ Fractionation System per rimuovere i frammenti inferiori ai 50 nucleotidi. Ai frammenti selezionati sono stati ligati degli adattatori direzionati, in grado di conservare lo strand di lettura della molecola di partenza, utilizzati per la sintesi di molecole di cDNA a doppio filamento, mediante trascrittasi inversa, valendosi di un'emulsion PCR (**Figura IV**).

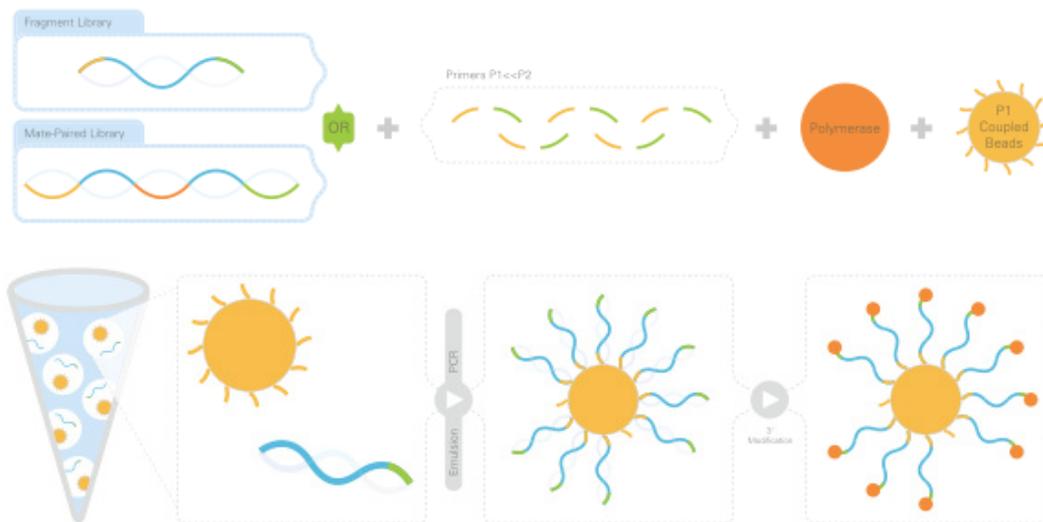


Figura IV: Preparazione della library ed emulsion PCR [<http://solid.appliedbiosystems.com>].

Questo tipo di PCR consiste in microreattori contenenti il template, i componenti della reazione di PCR, le biglie e i primers. Dopo la PCR, i template sono stati denaturati, tra le biglie arricchite sono state selezionate quelle in cui il template è stato amplificato (> 90%) e questi sono stati sottoposti ad una modificazione al 3' per consentire il successivo legame con il vetrino [<http://solid.appliedbiosystems.com>].

La preparazione della library è stata opportunamente controllata durante tutti gli step, qualitativamente mediante l'uso dello spettrofotometro Qubit® (Invitrogen™) e quantitativamente attraverso DNA 1000 Pico chip - 2100 Bioanalyzer (Agilent Biotechnologies).

Deposizione delle biglie

Le biglie con la modificazione al 3' sono state depositate sul vetrino (**Figura V**). La suddivisione in camere del vetrino ha consentito di processare contemporaneamente i 4 diversi campioni (Cx_KO, M_KO, Cx_WT, M_WT).

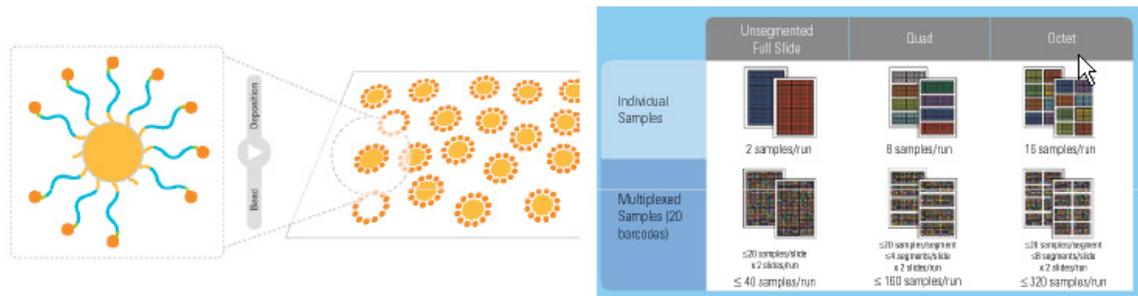


Figura V: Deposizione delle biglie [<http://solid.appliedbiosystems.com>].

Sequenziamento per ligazione

I primers sono stati ibridati alla sequenza dell'adattatore all'interno della libreria di templati. Un set di 4 sonde fluorescenti di-nucleotidiche compete con il primer per il legame alla sequenza di cDNA. La specificità della sonda di-nucleotidica è ottenuta esaminando tutte le prime e seconde basi in ogni reazione di ligazione. Sono stati effettuati cicli multipli di ligazione, rilevamento e taglio in numero tale da ottenere sequenze di 50 bp. A questo è seguita una serie di cicli di ligazione, il prodotto dell'estensione è stato rimosso e il template è stato resettato con un primer complementare alla posizione n-1 per un secondo round di cicli di ligazione (**Figura VI**) [[Metzker, 2010; http://solid.appliedbiosystems.com](http://solid.appliedbiosystems.com)].

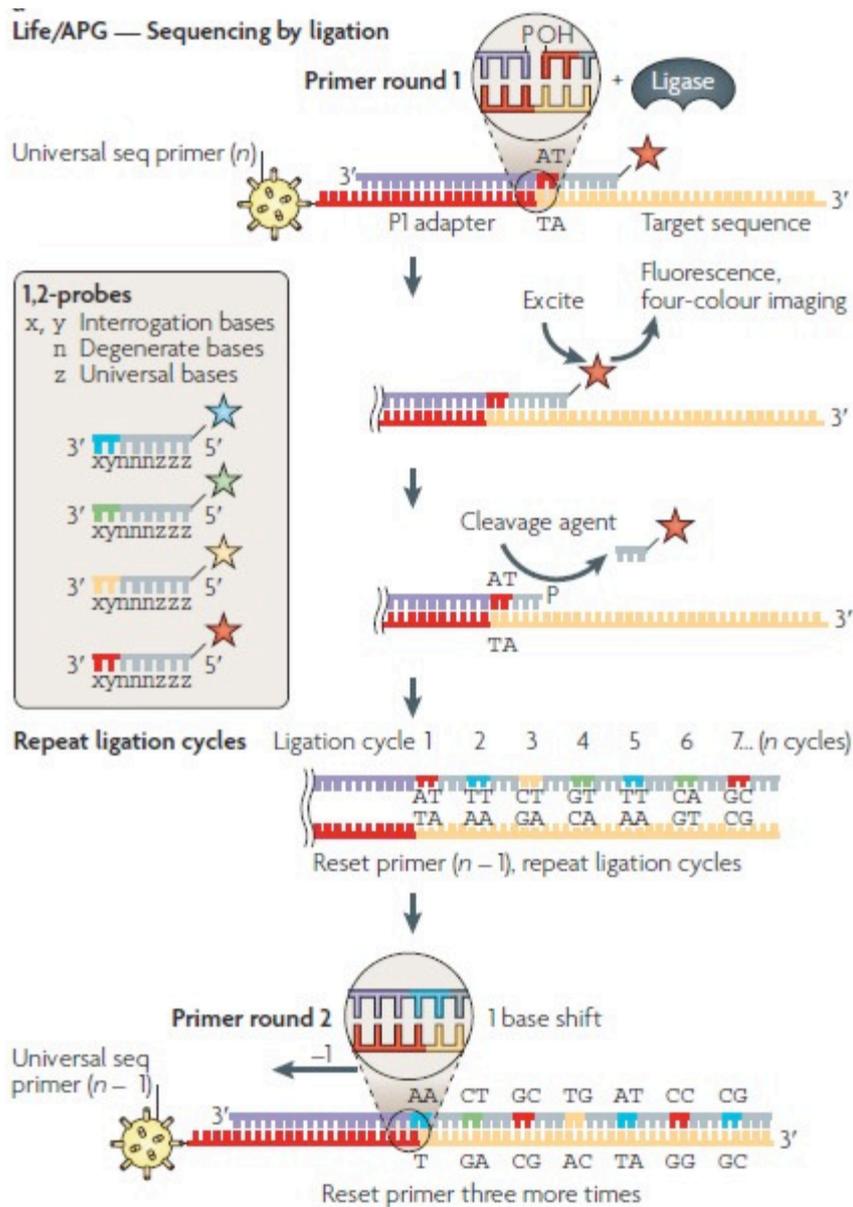


Figura VI: Sequenziamento per ligazione [Metzker, 2010].

Per ogni tag vengono completati 5 cicli di primer reset. Attraverso questo processo, ogni base viene sondata in due indipendenti reazioni di ligazione con due differenti primers (**Figura VII**).

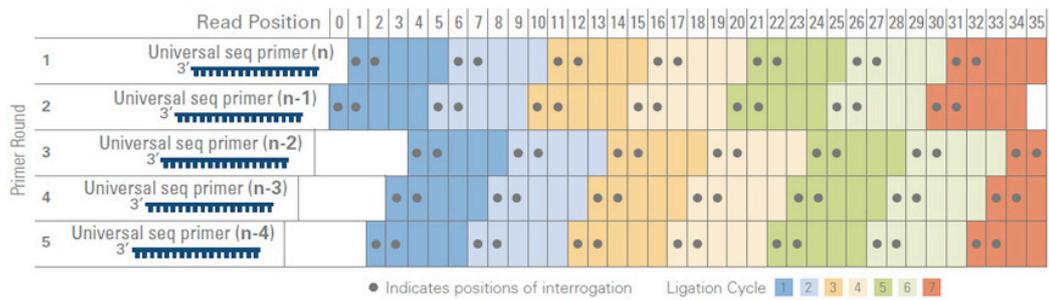


Figura VII: Primer reset [http://solid.appliedbiosystems.com].

Per esempio, la base in posizione 5 è analizzata con il primer numero 2 nel ciclo di ligazione 2 e con il primer numero 3 nel ciclo di ligazione 1. Questa doppia interrogazione è fondamentale per l'accuratezza e la tipologia di codifica a due basi che caratterizza il SOLiD System e che permette di identificare senza errore la presenza di SNPs nella sequenza (**Figura VIII**) [Metzker, 2010; http://solid.appliedbiosystems.com; Peckham, ASHG 2007].

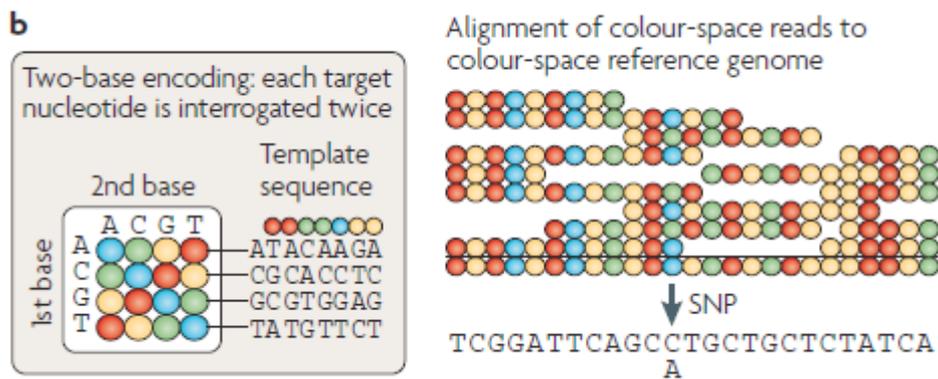


Figura VIII: Two-base encoding [Metzker, 2010].

Ogni corsa del SOLiD 3 Plus System produce 12.5-15 Gb di dati di sequenze (reads) di 50 bp di lunghezza.

ANALISI DEI DATI

I dati ottenuti dal sequenziamento dell'RNA con SOLiD™ 3 Plus System sono stati sottoposti ad: allineamento, identificazione dei geni differenzialmente espressi e analisi funzionale.

Allineamento

Le reads ottenute sono state sottoposte a controllo di qualità e allineate sul genoma murino dai bioinformatici del Cribi dell'Università di Padova utilizzando il programma PASS [Campagna et al., 2009; <http://pass.cribi.unipd.it>]. PASS (Program to Align Short Sequences) è un software ottimizzato per effettuare veloci allineamenti gapped e ungapped su una sequenza di riferimento e disegnato per trattare le grandi quantità di sequenze corte generate dalle tecnologie di sequenziamento di nuova generazione, compreso l'RNA sequencing mediante SOLiD™ System, usato in questo esperimento.

Come sequenza di riferimento è stata utilizzata la versione mm9 del genoma murino (aggiornata a luglio 2007) depositata nell'UCSC Genome Bioinformatics Site [<http://genome.ucsc.edu/>] e come opzioni sono state scelte best-hit, massimo 2 mismatch e nessun GAP. I trascritti allineati sono stati filtrati considerando solo gli UNIQUE, allineati su un solo gene, e quelli con copertura del gene superiore al 50 % della sua lunghezza.

Identificazione dei geni differenzialmente espressi

Per l'individuazione dei geni differenzialmente espressi evidenziati nei confronti Cx_KO vs Cx_WT e M_KO vs M_WT ci si è avvalsi di DEGseq, un pacchetto del software R per l'analisi statistica dell'espressione differenziale di dati derivanti da RNA sequencing, utilizzando come metodo statistico LTR (Likelihood Ratio Test) [Wang et al., 2010; www.bioconductor.org/packages/2.6/bioc/html/DEGseq.html].

In ciascuno dei due confronti sono stati considerati differenzialmente espressi i geni che soddisfano le due condizioni:

- p-value <0.05
- $|\log_2 (KO/WT)| \geq 0.7$

dove KO e WT rappresentano il numero di reads uniche che si allineano sul gene per il campione KO e WT, rispettivamente.

Inoltre, prima di procedere con l'analisi delle liste di geni differenzialmente espressi è stato necessario apportare una correzione. Alcuni geni allineati hanno un valore di espressione pari a zero in un campione e diverso da zero nell'altro. Tali geni potrebbero risultare molto interessanti ai fini dell'analisi ma, poiché il software DEGseq non è predisposto per trattare valori nulli, con la filtrazione questi geni andrebbero persi. Si è deciso pertanto di attribuire, in questi casi, un valore di espressione pari a 0.5 in sostituzione dello zero.

Analisi funzionale

Per individuare i processi e i componenti biologici maggiormente interessati dalla variazione di espressione genica nei campioni KO rispetto al controllo, le liste di geni differenzialmente espressi per ciascuna area sono state sottoposte ad un'analisi di arricchimento funzionale. L'arricchimento è una misura della probabilità che una lista di geni contenga più membri di una data categoria biologica rispetto a quanto sarebbe atteso dal caso, dato il numero totale di membri della categoria e il numero di geni della lista da analizzare. Come sistemi di annotazione sono state utilizzate la Gene Ontology e altri sistemi di classificazione, integrati in vari tools.

Gene Ontology

La Gene Ontology (GO) è un vocabolario controllato di termini che ha lo scopo di annotare i geni e i prodotti genici e descrivere le loro caratteristiche con l'intento di standardizzare la rappresentazione degli stessi tra le diverse specie e tra i diversi database. L'ontologia copre tre domini: *Biological Process*, *Molecular Function* e *Cellular Component*. Il dominio *Biological Process* è definito come un'operazione o un set di eventi molecolari con un inizio e una fine determinati, riguardanti il funzionamento di unità viventi integrate (cellule, tessuti, organi e organismi) in cui il gene è coinvolto; la *Molecular function* consiste nell'attività elementare del prodotto genico a livello molecolare (ad esempio legame o catalisi) mentre il dominio *Cellular*

Component si riferisce alla specifica parte della cellula o del suo ambiente extracellulare in cui il gene o prodotto genico è attivo.

PANTHER (Protein ANalysis THrough Evolutionary Relationships)

[<http://www.pantherdb.org/>] è un database pubblico che classifica i geni e i prodotti genici in base alla loro funzione, usando evidenze sperimentali pubblicate e relazioni evolutive per predire la funzione anche in assenza di evidenze sperimentali dirette. Le molecole sono classificate da biologi esperti in base all'associazione a famiglie e sottofamiglie geniche, alle categorie della Gene Ontology, ai termini delle classificazioni *PANTHER protein* e *Pathways*. PANTHER contiene l'informazione relativa a 165 pathways metabolici e regolatori. I tools di analisi di dati di espressione genica consentono di confrontare una lista di geni di interesse con una lista di riferimento per determinare la sovra- o sotto-rappresentazione delle diverse categorie. Per il confronto viene utilizzato un test binomiale che restituisce un valore di p-value per l'associazione ad ogni termine della classificazione di PANTHER.

ExPlain™

[<http://www.biobase-international.com/>] è un sistema di analisi che combina tools di classificazione funzionale, analisi promotoriale e analisi di pathways. Usando i dati relativi ai fattori di trascrizione presenti nel database TRANSFAC[®], ExPlain™ permette di identificare i fattori di trascrizione che potrebbero influenzare l'espressione genica di una lista di geni differenzialmente espressi e predire come questi fattori, in combinazione, possano indurre i patterns di espressione genica osservati. Il software permette di estendere l'analisi fino ad individuare i regolatori chiave del signaling cellulare che influenzano l'attività di questi fattori di trascrizione. Il database TRANSFAC[®] contiene informazioni sui fattori di trascrizione, sui loro siti di legame provati da dati sperimentali, e sui geni che regolano. Comprende 14490 fattori, 30118 siti di legame, 41859 interazioni fattore-sito di legame, 277337 sequenze promotoriali (di cui 57819 murine) e 24560 riferimenti bibliografici.

Il modulo Functional classification permette di identificare i termini statisticamente significativi usando diversi sistemi di annotazione e quelli usati sono elencati di seguito:

- *Expression, BKL manual curation* in cui le categorie di annotazione sono i differenti organi, tessuti, tumori e tipi cellulari dell'organismo umano. I geni di topo e ratto sono annotati per ortologia;
- *GO annotation, BKL manual curation* in cui le categorie di annotazione corrispondono ai termini della Gene Ontology e l'annotazione è curata manualmente nella Biobase Knowledge Library (BKL);
- *Proteome BKL Disease View* in cui i geni sono associati alle patologie umane annotate nel database BKL. I geni di topo e ratto sono annotati per ortologia;
- *SwissProt keywords* in cui le categorie di annotazione sono termini derivati dal database UniProt;
- *Transcription Factor Classification* in cui la classificazione è basata sui fattori di trascrizione annotati nel database TRANSFAC
- *TRANSPATH Molecule Classification* in cui si utilizza una classificazione di molecole annotate nel database TRANSPATH.

IPA® (Ingenuity® Pathways Analysis)

[Ingenuity Systems; [http:// www.ingenuity.com](http://www.ingenuity.com)] è un software web-based disegnato per organizzare l'informazione biologica in modo da permettere di analizzare e interpretare i dati derivati da esperimenti di espressione genica, proteomica e metabolomica.

Si basa sull'*Ingenuity Knowledge Base*, un database di annotazioni funzionali e interazioni biologiche comprendente 17400 geni umani, 13100 geni murini, 7500 geni di ratto, 13415 sostanze chimiche, 2.75 milioni di referenze bibliografiche e 1.6 milioni di interazioni chimiche e annotazioni funzionali con evidenza scientifica curate manualmente.

L'analisi funzionale di IPA è basata su tre categorie principali di funzioni: *Molecular and Cellular Functions*, *Physiological System Development and Functions* e *Diseases and Disorders*. All'interno di queste categorie sono presenti 85 categorie funzionali di alto

livello, a loro volta classificate in categorie di livello inferiore. Il tool *Canonical Pathway Analysis* è stato utilizzato per identificare i pathways metabolici e le vie di segnale associate alle liste di geni. Come significatività dell'associazione, il software restituisce due valori: il p-value e il rapporto tra il numero di geni della lista associati al pathway e il numero di geni totali del pathway.

Le categorie funzionali e i pathways maggiormente arricchiti nelle liste di geni dati come input sono identificati utilizzando un test esatto di Fischer che restituisce un p-value rappresentante la probabilità che l'associazione di ogni categoria o pathway a quel data set sia dovuta al caso.

KEGG (Kyoto Encyclopedia of Genes and Genomes)

[<http://www.genome.jp/kegg/>] è un database pubblico di informazioni genomiche e chimiche costituito da 16 database organizzati in 3 categorie principali: *Genomic information*, *Chemical information* e *Systems information*. Il tool *Pathway mapping* è utilizzato per mappare datasets di molecole, derivate da esperimenti high-throughput di genomica, trascrittomica, proteomica e metabolomica, sui pathways presenti nel database per facilitarne l'interpretazione biologica. Attualmente KEGG contiene 381 pathways curati manualmente.

RISULTATI E DISCUSSIONE

L'animale modello per la MPS II utilizzato nel presente lavoro è stato ampiamente studiato nel laboratorio dove è stata svolta questa tesi [Friso et al., 2005; Friso et al., 2010]. Dal punto di vista del coinvolgimento neurologico, però, tale caratterizzazione è limitata ad analisi biochimiche ed istologiche e anche in letteratura questo aspetto è poco approfondito. Pertanto, prima di procedere all'analisi del trascrittoma delle aree cerebrali, si è ritenuto opportuno effettuare una valutazione del coinvolgimento neurologico mediante test comportamentali.

Questa sezione si articola, dunque, in una breve esposizione dei test comportamentali effettuati e in una più ampia descrizione dell'analisi del trascrittoma.

Data la complessità dell'argomento e, soprattutto, dell'analisi dei dati ottenuti in questo lavoro sperimentale, si è scelto di presentare ogni risultato seguito dalla relativa discussione, allo scopo di facilitare la comprensione dello sviluppo del progetto stesso.

TEST COMPORTAMENTALI

L'applicazione di test comportamentali nella ricerca *in vivo* è una componente importante per la caratterizzazione dell'animale modello utilizzato e per evidenziare possibili differenze tra il topo KO e quello wild-type, che potranno essere successivamente approfondite con altre metodiche.

In questo caso sono stati applicati l'Open Field test per valutare le conseguenze comportamentali del coinvolgimento neurologico e il Footprint test per evidenziare un differente pattern nella camminata dovuto alle difficoltà motorie.

Sono stati valutati 7 animali IDS-KO di 9 mesi di età e 7 rispettivi fratelli wild-type.

Il modello murino per la MPS II a 9 mesi manifesta le seguenti caratteristiche fenotipiche: irrigidimento articolare con conseguenti difficoltà motorie, ingrossamento delle zampe, muso allargato e accorciato e pelo ispido.

Open Field test (OF)

Uno dei test più comunemente usati nello studio del comportamento murino è l'Open Field (OF) test. Questo metodo comporta l'osservazione e la misurazione del comportamento dei topi introdotti in una gabbia nuova, vuota e pulita dalla quale gli animali non sono in grado di uscire, determinando così uno stato di ansia. Con questo test si può analizzare qualitativamente e quantitativamente il comportamento esplorativo e motorio degli animali esaminati.

L'arena è stata divisa in 3 zone (centro, interno, esterno) e le registrazioni sono state analizzate con il programma ANY-maze (Stoelting, IL, USA).

In **Tabella 1** sono elencati i parametri rilevati, le relative medie dei valori ottenuti per i topi IDS-KO e WT e i valori di p-value risultanti dall'applicazione del test-t di Student al confronto tra le due medie.

Parametri valutati	Media valori topi IDS-KO	Media topi WT	p-value
Distanza (m)	36.154	25.177	0.008
Velocità media (m/s)	0.060	0.043	0.009
Velocità max (m/s)	0.429	0.323	0.023
Episodi di freezing	9	20	0.065
Time di freezing (s)	36.550	91.975	0.077
Centro: entrate	15	17	0.626
Centro: tempo (s)	19.850	31.775	0.201
Centro: distanza (m)	1.598	1.767	0.721
Centro: tempo mobile (s)	19.850	31.775	0.201
Centro: velocità max (m/s)	0.397	0.375	0.650
Interno: entrate	57	58	0.963
Interno: tempo (s)	99.888	168.175	0.126
Interno: distanza (m)	9.064	8.688	0.833
Interno: tempo mobile (s)	99.888	168.175	0.126
Interno: velocità max (m/s)	0.569	0.390	0.076
Esterno: entrate	43	42	0.832
Esterno: tempo (s)	480.275	386.550	0.055
Esterno: distanza (m)	25.492	14.722	0.001
Esterno: tempo mobile (s)	480.275	386.550	0.055
Esterno: velocità max (m/s)	0.605	0.375	0.026
Peso (g)	32.125	30.925	0.424
Punto centrale: massima distanza da (m)	0.389	0.375	0.023
Punto centrale: tempo mobile verso (s)	180.688	185.525	0.766

Tabella 1: Parametri medi rilevati dopo 10 minuti di Open Field test. In grassetto le differenze che sono risultate statisticamente significative per un valore di $p \leq 0.05$.

Come si vede dalla **Tabella 1**, il confronto tra topi IDS-KO e WT ha evidenziato differenze statisticamente significative ($p\text{-value}\leq 0.05$) per 6 parametri: la distanza totale percorsa, la velocità media, la velocità massima, la distanza percorsa nella zona esterna, la velocità massima raggiunta nella zona esterna e la massima distanza dal punto centrale.

Nella **Figura 1** sono riportati i relativi grafici.

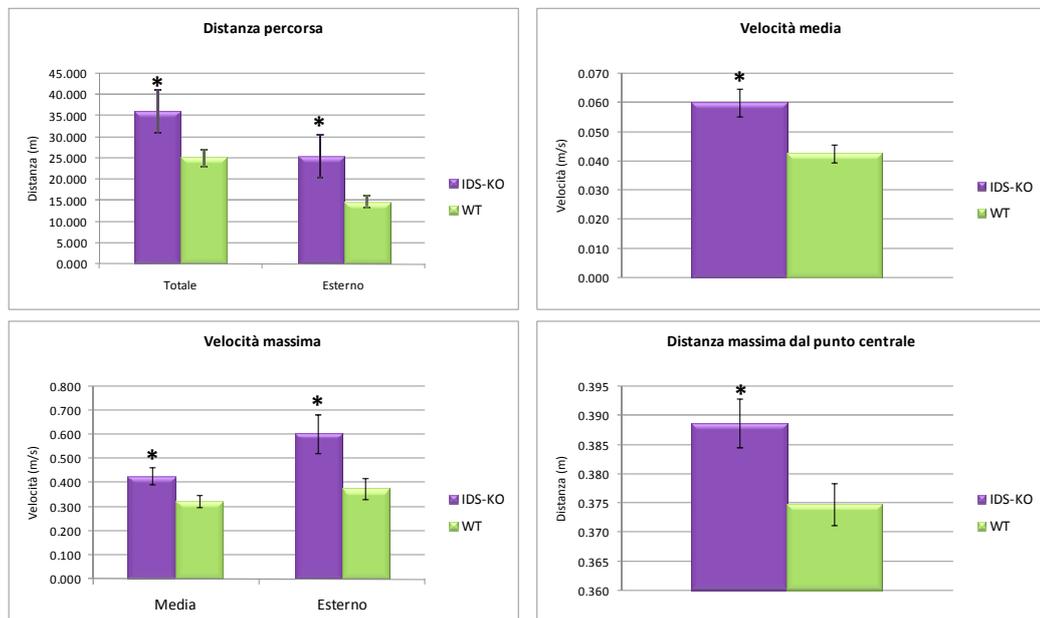


Figura 1: Rappresentazione grafica dei parametri risultati statisticamente significativi ($p \leq 0.05$) dopo 10 minuti di OF. Animali IDS-KO = 7, WT = 7. Le barre di errore rappresentano la SEM.

Per ogni animale sono stati contati, inoltre, i boli fecali quale dato ulteriore di valutazione dell'ansia. Come si vede in **Figura 2**, la defecazione nei topi IDS-KO risulta superiore ai WT in modo statisticamente significativo ($p\text{-value}=0.012$).

La distanza percorsa nella zona esterna e la massima distanza dal punto centrale evidenziano la tendenza dell'animale IDS-KO rispetto al WT a prediligere l'area perimetrale (esterno), zona in cui il topo si può sentire più protetto dalle pareti rispetto all'area aperta della zona centrale, come già ipotizzato in precedenza [Karl et al., 2003].

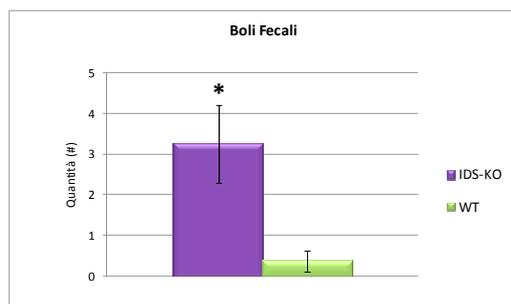


Figura 2: Media del numero di boli fecali prodotti in 10 minuti di OF ($p = 0.012$).

Animali IDS-KO = 7, WT = 7. Le barre di errore rappresentano la SEM.

L'aumento della tigmotassi associato all'aumento della defecazione è sintomo di un maggiore stato di ansia nei topi patologici rispetto ai fratelli wild-type [Denenberg, 1969].

La distanza totale percorsa, la velocità media, la velocità massima e la velocità massima raggiunta nella zona esterna sono aumentati in modo statisticamente significativo negli animali IDS-KO, indicando una maggiore attività complessiva; è ipotizzabile che questo dato rifletta l'iperattività che si riscontra nei pazienti MPS II. In **Figura 3** sono riportati i track-plot rappresentativi dei due genotipi, che mostrano evidenti differenze tra i due tipi di animali. Tale dato conferma la maggiore tendenza dell'animale IDS-KO rispetto al WT a muoversi preferenzialmente nell'area perimetrale, già precedentemente rilevato.

Infine, è da sottolineare che altri parametri, seppure non statisticamente significativi, mostrano un andamento analogo a quelli presi precedentemente in considerazione; tra questi un aumento del tempo trascorso dai topi IDS-KO nella zona esterna, che rimarca l'ansia indotta dagli spazi aperti, e la diminuzione del numero di episodi di freezing (momenti di totale immobilità), che appoggia l'ipotesi di iperattività. Tuttavia, la conferma di questi dati richiederebbe un aumento importante della numerosità campionaria.

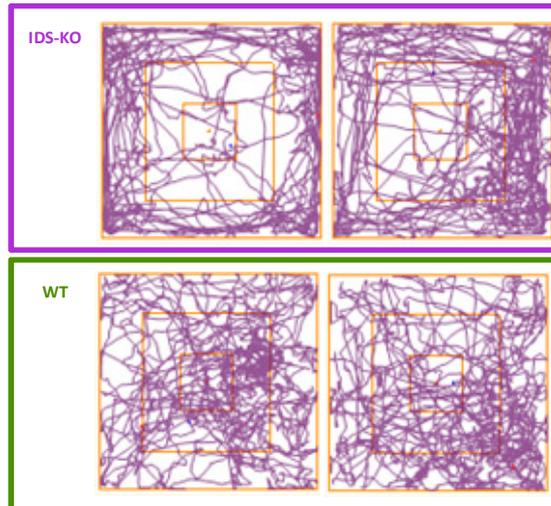


Figura 3: Track-plot: 4 rappresentazioni del tracciato ottenuto dall'analisi con ANY-maze di 10 minuti di registrazione di OF. Topi IDS-KO nel riquadro superiore, WT in quello inferiore.

Footprint test

Il footprint test valuta il pattern della camminata del topo in modo da rilevare eventuali difformità nell'andatura [Carter et al., 1999; Karl et al., 2003].

I patterns di uno dei topi IDS-KO e di uno WT è mostrato in **Figura 4**.



Figura 4: Footprint test: la linea tratteggiata permette di valutare la linearità di camminata degli animali. Le impronte delle zampe anteriori sono nere, quelle posteriori rosse. IDS-KO (superiore), WT (inferiore).

Sono stati valutati tre parametri: la lunghezza del passo, la sovrapposizione delle zampe anteriori e posteriori e la linearità della camminata.

Come mostrato in **Figura 5**, i topi IDS-KO risultano avere una lunghezza del passo ridotta in modo statisticamente significativo ($p\text{-value}=0.004$). Inoltre, i topi malati presentano un overlapping tra zampe anteriori e posteriori e un'andatura non lineare rispetto ai WT, caratteristiche probabilmente imputabili alla compromissione osteo-articolare.

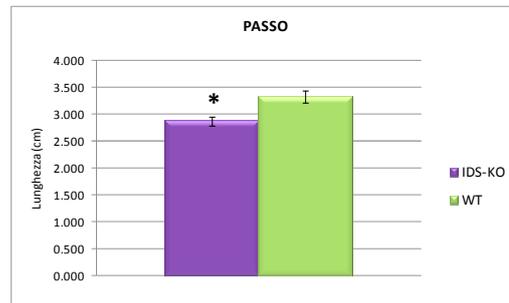


Figura 5: Media della lunghezza del passo dei topi. Misura di 3-7 passi per animale.

Animali IDS-KO = 7, WT = 7. Le barre di errore rappresentano la SEM.

ANALISI DEL TRASCRITTOMA

Rna-Seq: Preparazione e analisi dei campioni

I cervelli di 7 animali KO di 9 mesi di età e di 7 rispettivi fratelli wild-type (WT) sono stati prelevati e suddivisi in bulbi olfattivi, corteccia, mesencefalo, nuclei della base, midollo allungato e cervelletto. In questo lavoro, delle macroaree cerebrali isolate è stato deciso di valutare il trascrittoma solamente della corteccia e del mesencefalo. La corteccia cerebrale è implicata nell'apprendimento, nel pensiero, nella funzione espressiva e nel comportamento, mentre il mesencefalo comprende diverse sotto-aree (tra cui corpo striato, ippocampo, talamo, ippotalamo) coinvolte nella memoria, nell'apprendimento, nel movimento, nelle emozioni e nei ritmi circadiani; alcune di queste funzioni risultano alterate nei pazienti con sindrome di Hunter.

Il trascrittoma è stato valutato mediante RNA-Seq utilizzando una tecnologia di seconda generazione, SOLiD™ 3 Plus System (Sequencing by Oligo Ligation and Detection) (Applied Biosystems). Prima di poter sequenziare il trascrittoma l'RNA deve essere valutato in termini di integrità, arricchito in mRNA, frammentato e retroscritto per la costruzione della library.

Valutazione integrità RNA

L'integrità dell'RNA estratto dai diversi animali e dalle diverse aree è stata valutata separatamente prima di effettuare il pool dei campioni. Come si evidenzia dagli elettroferogrammi in **Figura 6** l'RNA ottenuto presenta una buona integrità (**RNA Integrity Number > 7**) con un valore medio di RIN di circa 8.7.

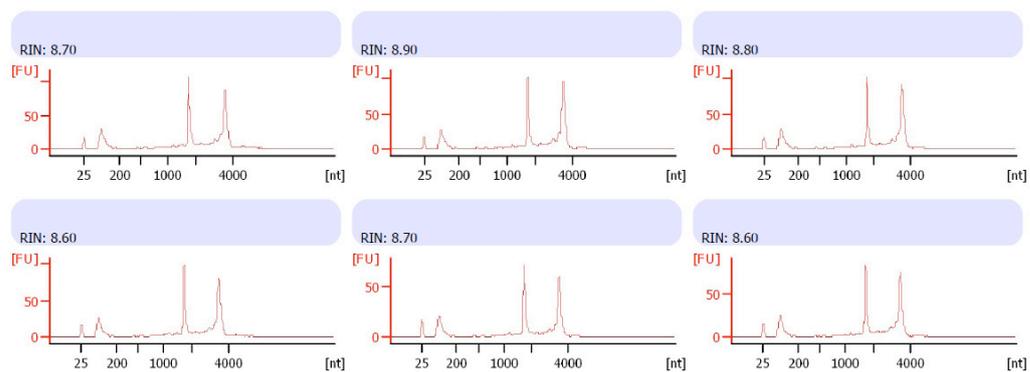


Figura 6: Elettroferogrammi ottenuti dall'analisi dell'RNA, mediante 2100 Bioanalyzer, di alcuni campioni di corteccia e di mesencefalo. In ogni grafico i picchi a destra rappresentano l'rRNA 18S e 28S rispettivamente.

Purificazione dell'mRNA

Appurata l'integrità dell'RNA sono stati costituiti i pool dei 4 diversi campioni di corteccia e di mesencefalo knock-out e wild-type (Cx_KO, Cx_WT, M_KO, M_WT). L'arricchimento in mRNA è un metodo che ha permesso di minimizzare la grande quantità di RNA ribosomale e transfer presente nei campioni, circa il 90% dell'RNA totale; ciò ha consentito di analizzare così quasi esclusivamente RNA messaggeri. L'efficienza della purificazione è stata valutata con il Bioanalyzer e come si può osservare in **Figura 7** la percentuale di rRNA residuo nei diversi campioni, dopo due cicli di arricchimento, è compreso tra 0.8% e 2.0%.

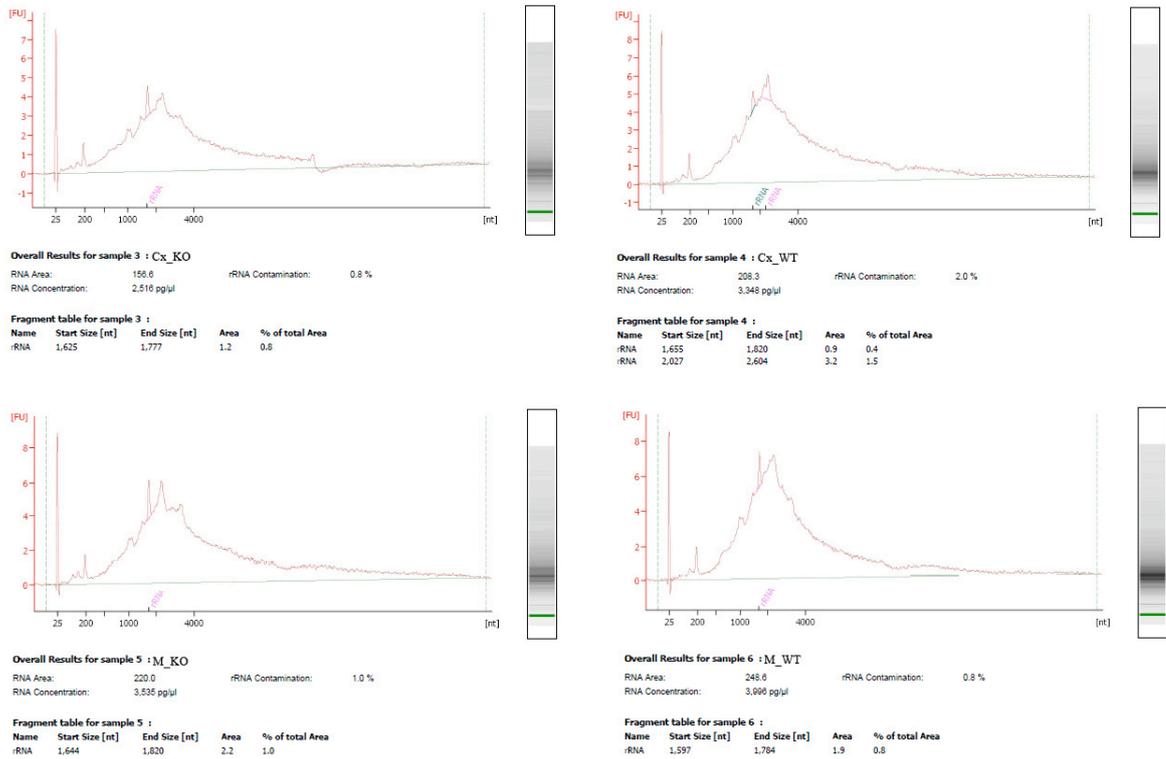


Figura 7: Elettroferogrammi ottenuti dall'analisi dell'mRNA di corteccia e di mesencefalo KO e WT, mediante 2100 Bioanalyzer (Agilent Biotechnologies).

Frammentazione dell'RNA

L'RNA purificato è stato frammentato per permettere le fasi successive di allestimento della library per il sequenziamento e il risultato della frammentazione (50-150 nucleotidi di lunghezza) è stato valutato con un RNA 6000 Pico chip - 2100 Bioanalyzer (**Figura 8**).

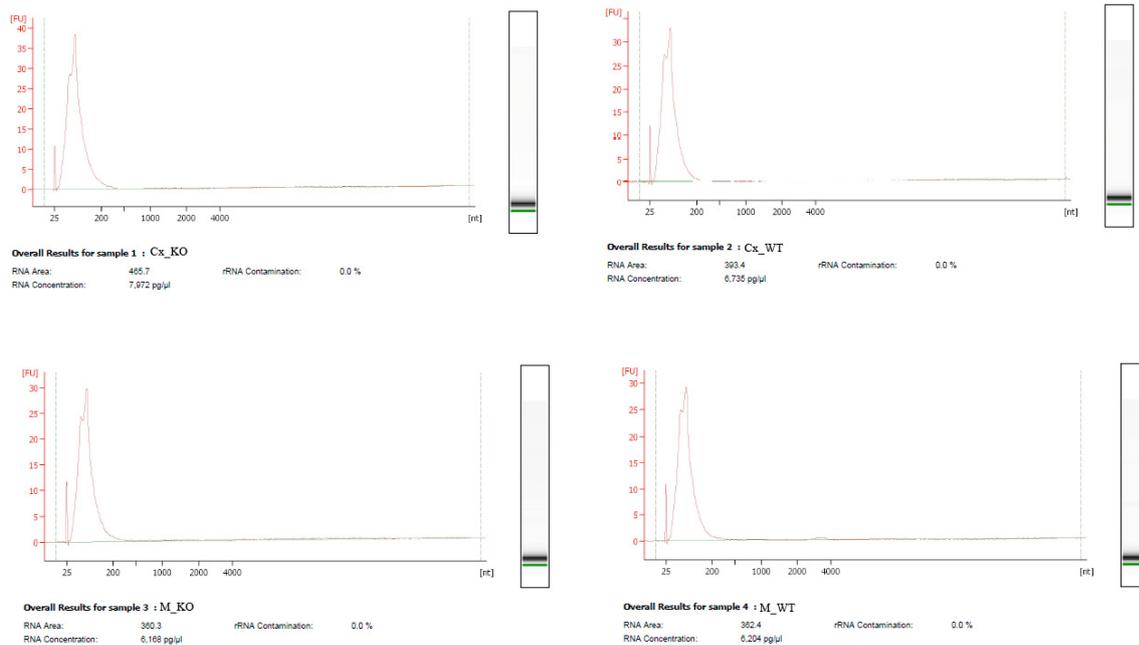


Figura 8: Elettroferogrammi ottenuti dalla frammentazione dell'mRNA di corteccia e di mesencefalo di topi IDS-KO e WT, mediante 2100 Bioanalyzer (Agilent Biotechnologies).

Preparazione della library

L'mRNA frammentato è stato retrotrascritto in cDNA, dopo 12 cicli di amplificazione è stato selezionato per lunghezza (150-200bp), come si osserva dagli elettroferogrammi in **Figura 9**, in modo da renderlo processabile per il sequenziamento con il SOLiD 3 Plus System.

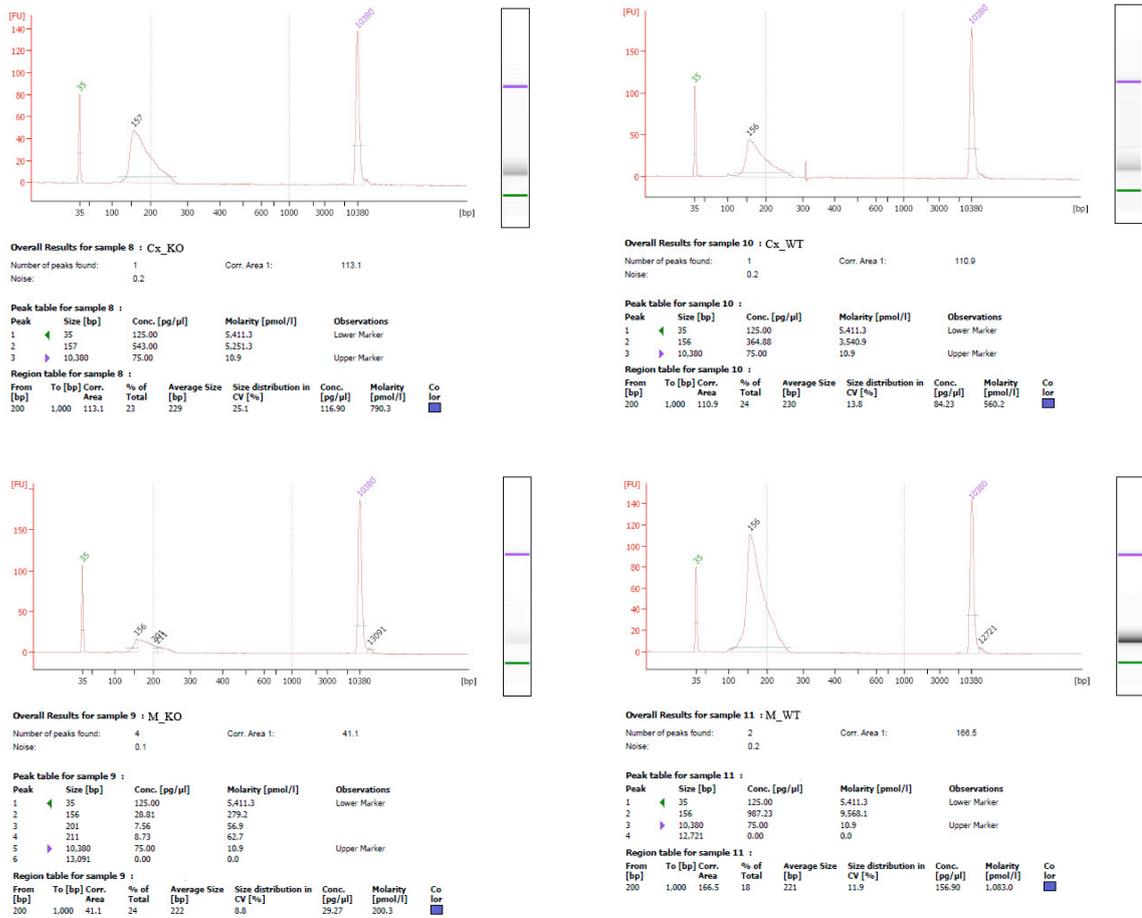


Figura 9: Elettroferogrammi ottenuti dalla library di cDNA di corteccia e di mesencefalo di topi IDS-KO e WT, mediante DNA 1000 Pico chip - 2100 Bioanalyzer (Agilent Biotechnologies).

Analisi dei dati

Il sequenziamento dell'RNA della corteccia e del mesencefalo dell'animale modello per la mucopolisaccaridosi di tipo II e del wild type ha fornito un elevato numero di dati (reads). Le reads ottenute sono state sottoposte ad allineamento, identificazione dei geni differenzialmente espressi e analisi funzionale.

Allineamento

La **Tabella 2** riporta alcune informazioni riguardanti il numero di reads ottenute dal processamento di ciascun campione e il risultato derivato dal loro allineamento sul genoma murino effettuato con il software PASS [Campagna et al., 2009;

<http://pass.cribi.unipd.it>]. In media il 56.5% delle reads ha superato il controllo di qualità, che permette di identificare le reads in grado di passare alla fase di allineamento, e il 65.6% di queste mappano su trascritti murini annotati.

ID sample	numero totale reads	numero di reads che superano il controllo di qualità (pcq)	% pcq	totale reads allineate	% allineate
Cx_KO	98,434,960	56,388,301	57.28%	38,502,647	68.28%
M_KO	98,716,308	56,155,432	56.89%	37,363,194	66.54%
Cx_WT	88,183,367	52,388,424	59.41%	33,617,996	64.17%
M_WT	87,079,859	45,727,298	52.51%	28,945,731	63.30%

Tabella 2: Processo di allineamento delle reads sul genoma murino.

pcq: passed check quality.

In **Figura 10** è riportata la distribuzione del numero di reads mappate su un gene per ognuno dei campioni. Come atteso, dato l'elevato numero di reads ottenute in ciascun campione, si può osservare dai grafici che i dati seguono una distribuzione normale.

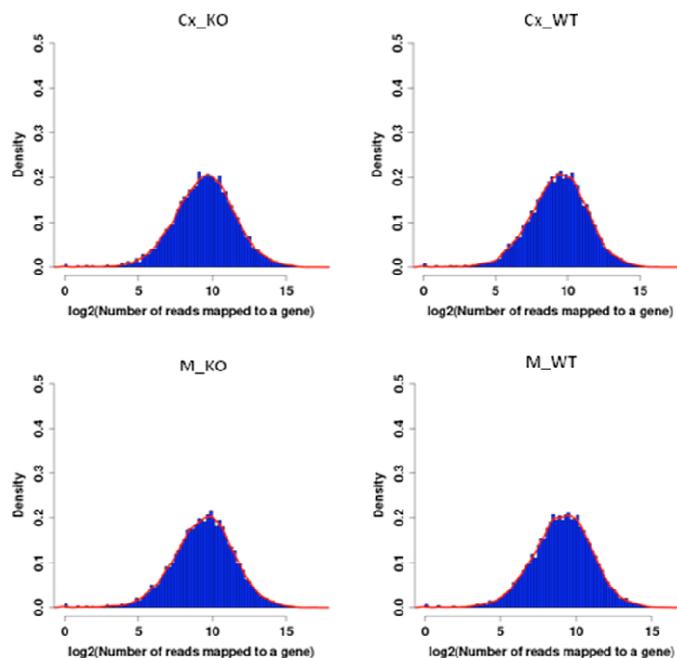


Figura 10: Distribuzione numero di reads per gene (DEGseq).

Identificazione dei geni differenzialmente espressi

Dall'allineamento è stata ottenuta una lista di 27309 geni murini con i relativi valori di espressione nei quattro campioni, dove per valore di espressione di un gene si intende il numero di reads uniche allineate sulla sequenze del gene. Le reads uniche sono reads che si allineano in un solo gene.

I 27309 geni sono stati sottoposti a filtrazione per coverage considerando solo quelli con copertura media del gene superiore al 50%. Dopo filtrazione il numero dei geni risulta 12723.

A seguito dell'analisi con il software DEGseq, per ogni confronto (Cx_KO vs Cx_WT e M_KO vs M_WT), sono stati considerati come differenzialmente espressi i geni che presentavano $p\text{-value} \leq 0.05$ e $[\log_2(\text{KO}/\text{WT})] \geq 0.7$ e ≤ -0.7 , dove KO e WT rappresentano il numero di reads uniche che si allineano sul gene per il campione KO e WT, rispettivamente.

Poiché il software DEGseq non è in grado di analizzare valori nulli e poiché alcuni geni allineati hanno valore di espressione pari a zero in un campione e diverso da zero nell'altro, si è deciso di sostituire a tutti gli zeri un valore pari a 0.5, considerato che tali geni potrebbero risultare molto interessanti ai fini dell'analisi.

In **Figura 11** sono rappresentati i MA-plot dei due confronti considerati dove per ogni gene mappato è riportata in ascissa la media dei logaritmi in base 2 dell'espressione nei due campioni e in ordinata il logaritmo in base 2 del rapporto tra i due valori di espressione. Come atteso da una distribuzione non distorta, i grafici si presentano come una nuvola ellittica di punti di simmetria rispetto alla retta $M=0$. Ciò esclude la necessità di ricorrere ad ulteriori normalizzazioni, oltre a quella effettuata dal software DEGseq, prima di procedere con l'analisi dei dati.

In rosso sono raffigurati i geni differenzialmente espressi. Come si può osservare nei due confronti si evidenzia una più importante down regolazione genica nelle aree provenienti dal topo KO rispetto a quelle provenienti dal topo WT e una minore up-regolazione.

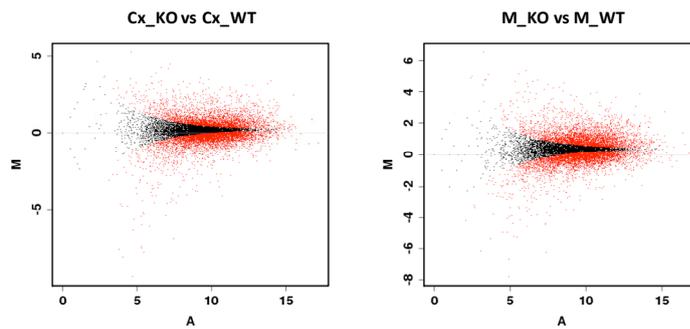


Figura 11: MA-plot dei geni differenzialmente espressi (rosso) nei due confronti (DEGseq).

L'analisi con DEGseq ha identificato 1202 e 1562 geni sovraespressi e 1463 e 1616 geni sottoespressi, rispettivamente nella corteccia e nel mesencefalo dell'animale IDS-KO confrontati con i rispettivi campioni WT (**Figura 12**).

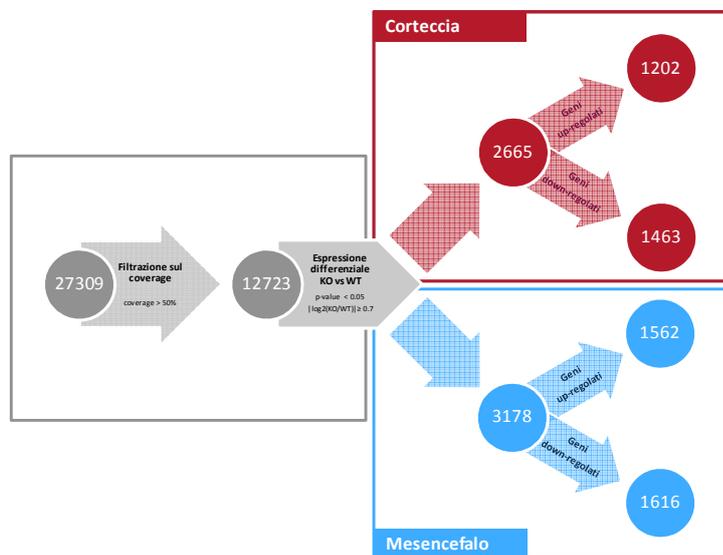


Figura 12: Schema riassuntivo del processo di filtrazione dei geni.

Il confronto tra le due aree cerebrali, inoltre, ha evidenziato 792 geni upregolati e 822 geni downregolati comuni alle due macroaree (**Figura 13**).

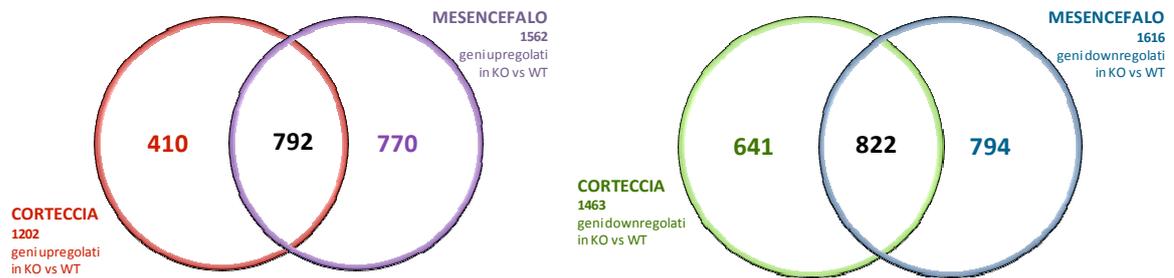


Figura 13: Diagramma di Venn dei geni upregolati (sinistra) e downregolati (destra) unici e comuni in corteccia e mesencefalo.

Analisi funzionale

L'analisi funzionale è stata effettuata avvalendosi dei software: PANTHER, ExPlain ed IPA. I software applicano metodi statistici diversi per valutare il grado di arricchimento e per semplicità e omogeneità di rappresentazione si è scelto di uniformare tutti i risultati secondo l'output di uno di questi. Poiché i tools utilizzati si avvalgono di test statistici diversi si è deciso di rappresentare il grado di arricchimento della lista di geni esaminata, riportando in grafico la significatività statistica di tale arricchimento espressa dal valore di p-value.

Tali software hanno permesso di identificare i termini della gene ontology e i pathways maggiormente interessati da espressione differenziale, nella comparazione Cx_KO vs Cx_WT e M_KO vs M_WT. Da ora in avanti, i campioni di corteccia e di mesencefalo (KO vs WT) saranno, per semplicità, denominati Cx e M rispettivamente.

Gene Ontology (GO)

La Gene Ontology permette di classificare in modo gerarchico i termini che rappresentano i geni e le proprietà dei prodotti genici. La valutazione della GO è stata eseguita utilizzando software diversi tra loro in modo da aumentare il pannello di termini arricchiti, grazie ai diversi algoritmi e database utilizzati dai vari programmi,

con l'intento di ottenere un risultato più solido. I software impiegati sono PANTHER (Protein ANalysis THrough Evolutionary Relationships) [http://www.pantherdb.org], Explain™ Analysis System (BIOBASE Biological Database) [http://www.biobase-international.com] ed IPA® (INGENUITY® Systems) [http://www.ingenuity.com].

- PANTHER

La GO ottenuta con PANTHER ha identificato, nel dominio *Biological Process*, 72 e 66 termini statisticamente significativi ($p \leq 0.001$), rispettivamente nella corteccia e nel mesencefalo KO, in **Figura 14** sono riportati i primi 20 più rappresentati.

I termini maggiormente arricchiti sono implicati nei processi, nella comunicazione nella trasmissione del segnale cellulare. A parte queste categorie generiche risultano interessanti *Neurological system process*, *Nervous system development* e *Synaptic transmission* con 17% e 19%, 27% e 22%, 24% e 27% di geni differenzialmente espressi sul totale di quelli presenti nella categoria nella corteccia cerebrale e mesencefalo rispettivamente.

Tali termini fanno pensare a una modificazione della trasmissione del segnale neuronale affiancata da una anomalia legata ai processi del comparto neurologico.

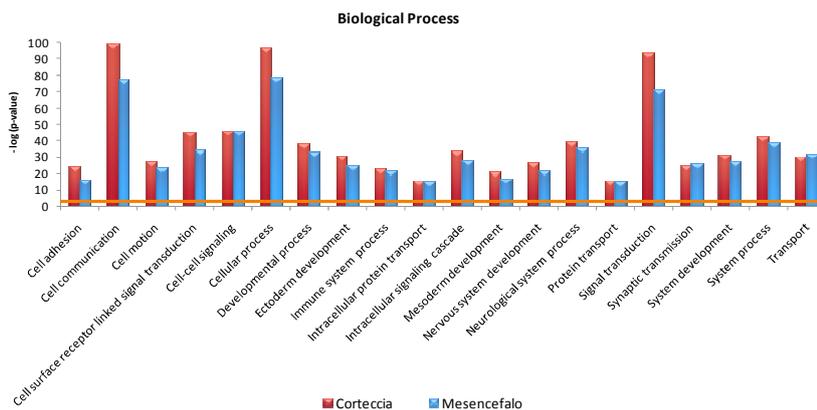


Figura 14: GO_PANTHER (*Biologic Process*). I primi 20 termini del dominio più arricchiti nella corteccia e nel mesencefalo KO rispetto al WT. L'arricchimento della categoria è espresso come $-\log(p\text{-value})$ e la linea continua arancione rappresenta $-\log(0.001)=3$.

Anche il movimento cellulare e i segnali tra le cellule sembrano subire delle variazioni dal pattern non patologico in entrambe le macroaree valutate, suggerendo un'alterazione della fisiologia cerebrale.

Il dominio *Molecular Function* (**Figura 15**) ha individuato 42 termini nella corteccia e 42 nel mesencefalo KO statisticamente significativi, dei quali in figura sono riportati i primi 20.

I termini più rappresentati riguardano legami tra le proteine, in particolare legami ai recettori.

Una funzione molecolare importante come quella dei trasportatori e canali ionici sembra essere molto compromessa data la presenza dei termini: *Cation channel activity*, *Cation transmembrane transporter activity*, *Ion channel activity*, *Transporter activity* e *Voltage-gated ion channel activity*. Queste categorie fanno pensare ad una alterazione nella regolazione delle concentrazioni ioniche cellulari con conseguente modificazione della neurotrasmissione del segnale chimico ed elettrico. A ciò si aggiungono i termini *Calcium ion binding* e *Calmodulin binding* che evidenziano l'importanza dello ione Ca^{2+} nella patofisiologia della neurodegenerazione.

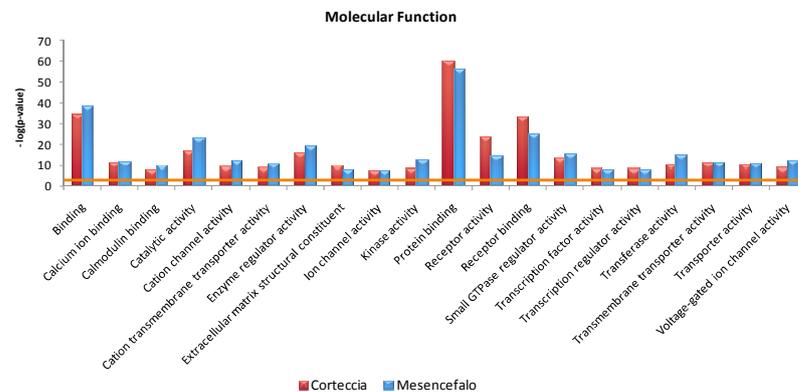


Figura 15: GO_PANTHER (*Molecular Function*). I primi 20 termini del dominio più arricchiti nella corteccia e nel mesencefalo KO rispetto al WT. L'arricchimento della categoria è espresso come $-\log(p\text{-value})$ e la linea continua arancione rappresenta $-\log(0.001)=3$.

Un altro termine che risulta arricchito è *Extracellular matrix structure constituent* con circa il 29% di geni in entrambe le aree. Tali variazioni sono probabilmente dovute alla

forte riorganizzazione cellulare che si osserva nelle cellule patologiche a causa dell'alterata via degradativa e dell'accumulo di glicosaminoglicani (GAGs) che si verifica anche nella matrice extracellulare.

Il terzo dominio, *Cellular Component*, ottenuto con l'utilizzo del software PANTHER ha fornito solo 7 termini per entrambe le aree cerebrali e di questi ne risultano in comune 5 statisticamente significativi (**Figura 16**).

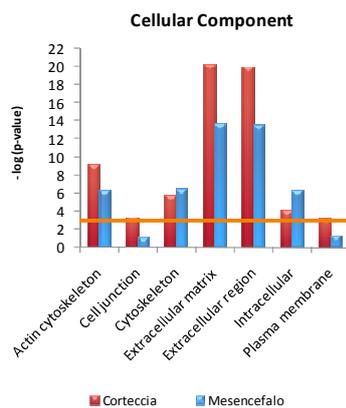


Figura 16: GO_PANTHER (*Cellular Component*). I 7 termini del dominio più arricchiti nella corteccia e nel mesencefalo KO rispetto al WT. L'arricchimento della categoria è espresso come $-\log(p\text{-value})$ e la linea continua arancione rappresenta $-\log(0.001)=3$.

Tra questi termini il citoscheletro potrebbe essere tra quelli di maggiore interesse, considerando la sua implicazione nella trasmissione dei neurotrasmettitori neuronali.

- Explain

Il software Explain utilizza le annotazioni corrispondenti ai termini della Gene Ontology, ma l'annotazione è curata manualmente nella Biobase Knowledge Library (BKL); ciò ha portato ad individuare sia termini già rilevati con il software PANTHER, sia ad identificarne di nuovi.

Sono stati considerati statisticamente significativi i termini con $p\text{-value} \leq 0.001$.

Nel dominio *Biological Process* sono stati identificati 451 e 363 termini, rispettivamente nella corteccia e nel mesencefalo KO; in **Figura 17** sono riportati i primi 22 termini più rappresentati nelle due aree.

Tale tipologia di annotazione ha evidenziato un maggior numero di termini legati alla trasmissione sinaptica, agli impulsi nervosi (in media 33% dei geni in entrambe le categorie e aree), alla memoria e apprendimento (36% Cx, 39% M) che delineano chiaramente un quadro di alterazione neurologica nella corteccia cerebrale e nel mesencefalo patologico rispetto al controllo.

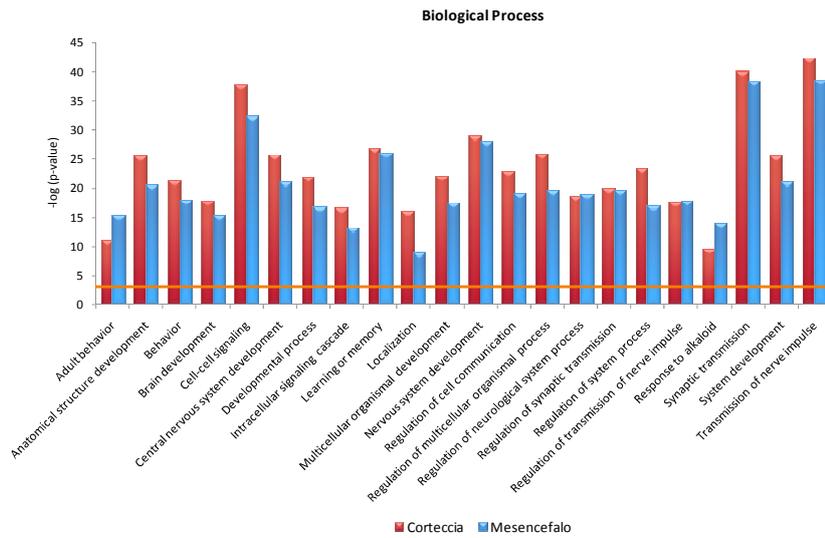


Figura 17: GO_Expain (*Biologic Process*). I primi 22 termini del dominio più arricchiti nella corteccia e nel mesencefalo KO rispetto al WT. L'arricchimento della categoria è espresso come $-\log(p\text{-value})$ e la linea continua arancione rappresenta $-\log(0.001)=3$.

Il dominio *Molecular Function* (**Figura 18**) ha individuato 57 termini nella corteccia e 45 nel mesencefalo KO statisticamente significativi; in figura sono riportati 24 termini comprendenti i primi 20 più rappresentativi di ogni macroarea. Oltre alle categorie già identificate con il software PANTHER, l'annotazione manuale di Explain porta alla luce termini importanti come *Muscarinic acetylcholine receptor activity* (Chrm1-5) il quale è accoppiato alle proteine G_q che regolano la fosfolipasi C, l'inositolo trifostato e l'accumulo del calcio intracellulare, questo potrebbe supportare l'ipotesi di una pesante alterazione dell'omeostasi del calcio.

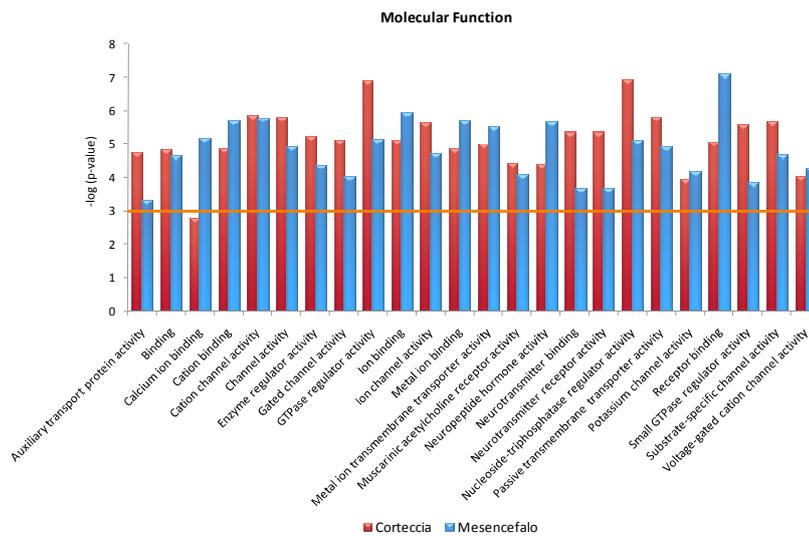


Figura 18: GO_Explain (*Molecular Function*). 24 termini del dominio comprendenti i primi 20 più rappresentativi nella corteccia e nel mesencefalo KO rispetto al WT. L'arricchimento della categoria è espresso come $-\log(p\text{-value})$ e la linea continua arancione rappresenta $-\log(0.001)=3$.

Il terzo dominio, *Cellular Component*, ottenuto con l'utilizzo del software ExPlain ha fornito termini più approfonditi rispetto al corrispondente output di PANTHER con 22 e 23 categorie rispettivamente nella corteccia e nel mesencefalo KO. In **Figura 19** sono riportati 25 termini comprendenti i primi 20 più rappresentativi di ogni macroarea.

I termini che risultano arricchiti includono geni di parti cellulari neuronali quali assoni (25% Cx, 32% M), soma (32% Cx, 37% M), dendriti (37 Cx, 35%M) e sinapsi (30% Cx, 31% M), ciò potrebbe rispecchiare un malfunzionamento di comparti cellulari differenti.

Inoltre, una classe molto interessante risulta *Neuron projection* con il 29% e il 32% di geni differenzialmente espressi sul totale di quelli presenti nella categoria, rispettivamente nella corteccia e nel mesencefalo patologico. Tale termine racchiude geni coinvolti nell'organizzazione delle proiezioni neuronali mediante regolazione del citoscheletro e delle vescicole esocitotiche calcio-dipendenti.

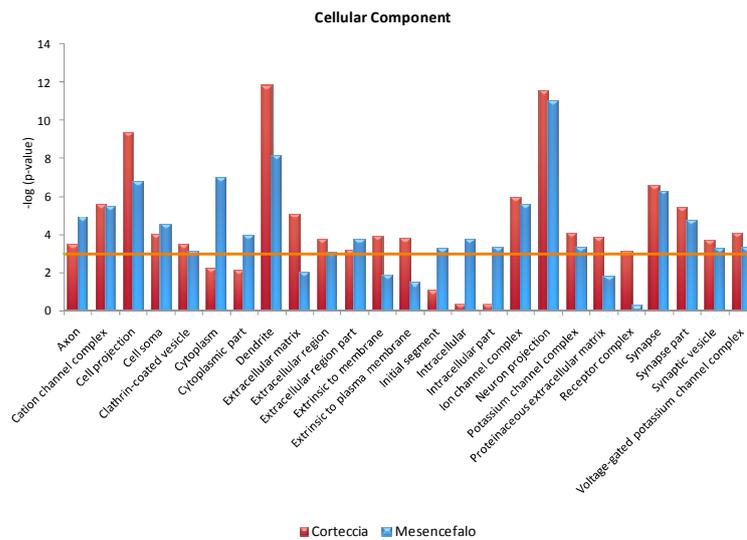


Figura 19: GO_Expain (*Cellular Component*). 25 termini del dominio comprendenti i primi 20 più rappresentativi nella corteccia e nel mesencefalo KO rispetto al WT. L'arricchimento della categoria è espresso come $-\log(p\text{-value})$ e la linea continua arancione rappresenta $-\log(0.001)=3$.

Categorie funzionali

Allo scopo di approfondire la classificazione dei geni differenzialmente espressi ottenuti dal confronto della corteccia e del mesencefalo patologico con il controllo sano, sono state eseguite ulteriori analisi avvalendosi del software IPA® (Ingenuity Pathways Analysis), oltre a quelli utilizzati precedentemente (PANTHER ed Explain).

- IPA

L'analisi funzionale di IPA è basata su tre categorie principali di funzioni: *Diseases and Disorders*, *Molecular and Cellular Functions* e *Physiological System Development and Functions*.

Sono stati considerati statisticamente significativi i domini con $p\text{-value} \leq 0.001$. In **Figura 20** sono riportate le categorie del dominio *Diseases and Disorders* ottenute dall'analisi delle due macroaree.

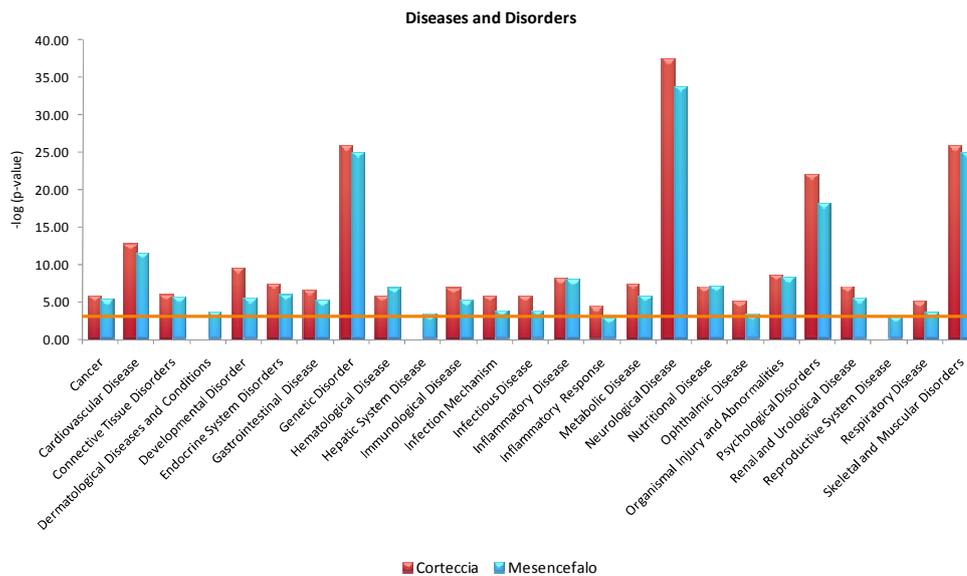


Figura 20: IPA (*Diseases and Disorders*). I termini del dominio più arricchiti nella corteccia e nel mesencefalo KO rispetto al WT. L'arricchimento della categoria è espresso come $-\log(p\text{-value})$ e la linea continua arancione rappresenta $-\log(0.001)=3$.

I termini più rappresentati e di interesse sono: *Neurological Disease*, *Genetic Disorder*, *Skeletal Muscular Disorders*, *Psychological Disorders*, *Organismal Injury and Abnormalities* e *Inflammatory Response*. Tali termini racchiudono diversi geni, alcuni in comune tra loro, che identificano principalmente le malattie neurodegenerative, come Huntington, Parkinson, Alzheimer, patologie autoimmuni, quali il Lupus Eritematoso e l'Artrite Reumatoide, e neuromuscolari come la Sclerosi Laterale Amiotrofica.

Il dominio *Molecular and Cellular Functions*, illustrato in **Figura 21**, conferma l'interessamento cellulare in termini di organizzazione, movimento e comunicazione intercellulare.

Il terzo dominio, *Physiological System Development and Functions*, presenta 3 categorie principali: *Behavior*, *Nervous system development and function* e *Organismal function* (**Figura 22**).

Tali termini sono identificativi dell'apprendimento cognitivo, della regolazione delle fasi del sonno, ma soprattutto sono correlati con lo sviluppo e le funzioni cellulari del sistema nervoso, tra cui l'attivazione degli astrociti e la plasticità sinaptica.

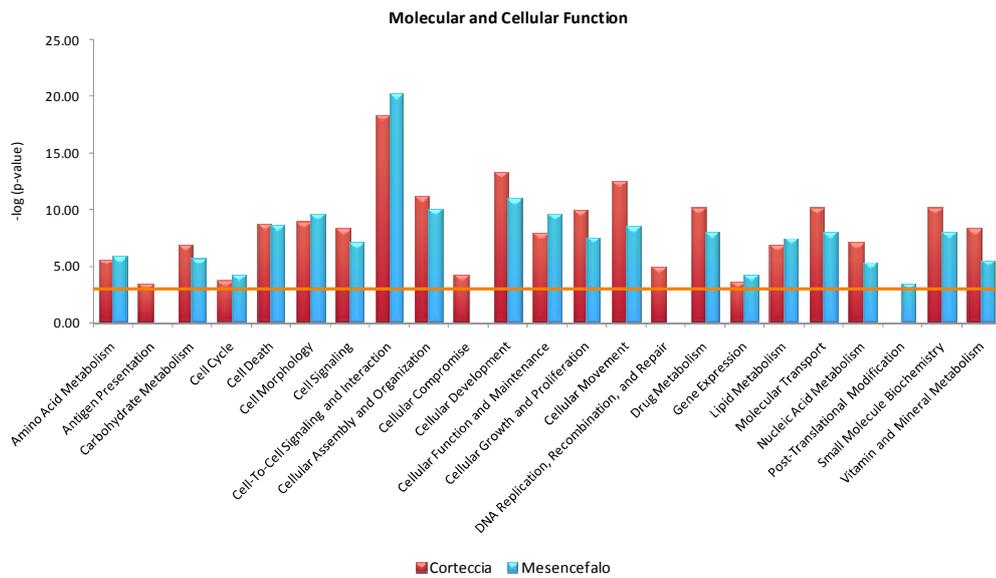


Figura 21: IPA (*Molecular and Cellular Function*). I termini del dominio più arricchiti nella corteccia e nel mesencefalo KO rispetto al WT. L'arricchimento della categoria è espresso come $-\log(p\text{-value})$ e la linea continua arancione rappresenta $-\log(0.001)=3$.

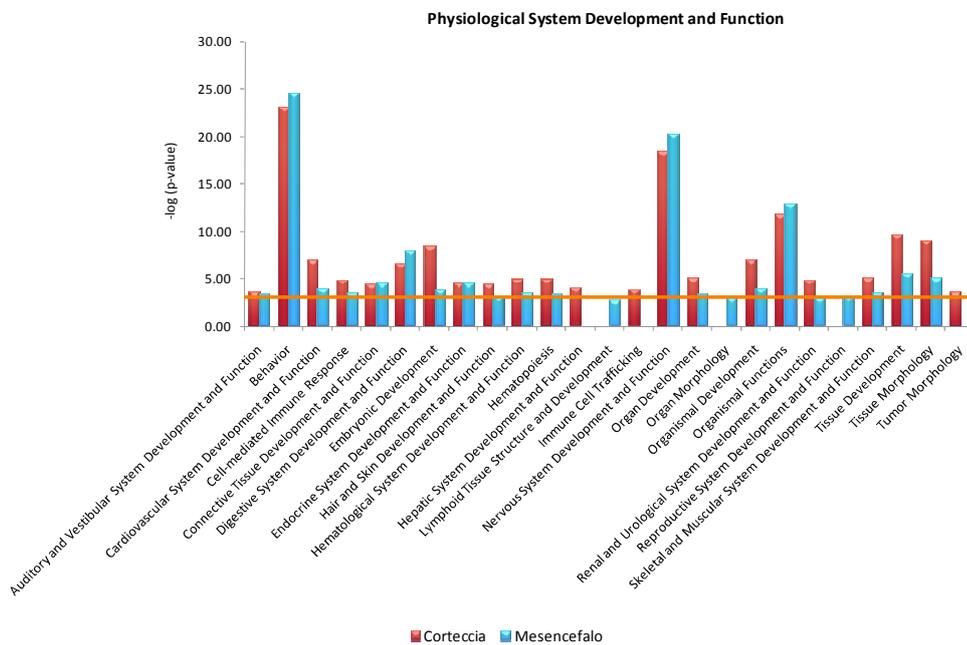


Figura 22: IPA (*Physiological System Development and Function*). I termini del dominio più arricchiti nella corteccia e nel mesencefalo KO rispetto al WT. L'arricchimento della categoria è espresso come $-\log(p\text{-value})$ e la linea continua arancione rappresenta $-\log(0.001)=3$.

- PANTHER

La classificazione proteica ottenuta con l'utilizzo del software PANTHER ha evidenziato un arricchimento nella categoria dei neuropeptidi (**Figura 23**), noti per avere effetti multipli sulla sinaptogenesi, sulla morfologia della glia e sull'espressione genica, interagendo con i recettori metabotropici e le G-protein. All'interno della classe sono presenti geni downregolati implicati nella plasticità e integrità sinaptica (Cbln1, Cerebellin-1), nello sviluppo e nella differenziazione neuronale (Pomc, pro-opiomelanocortin-alpha), nella regolazione omeostatica dell'energia metabolica e nel bilanciamento ormonale (Hcrt, hypocretin). Ciò porta ad ipotizzare una marcata debilitazione delle funzioni basali di sopravvivenza cellulare e una conseguente compromissione dell'area cerebrale coinvolta.

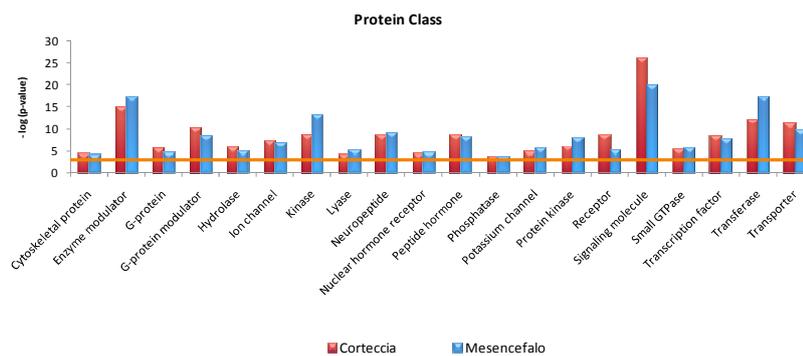


Figura 23: PANTHER (*Protein Class*). I primi 20 termini più arricchiti nella corteccia e nel mesencefalo KO rispetto al WT. L'arricchimento della categoria è espresso come $-\log(p\text{-value})$ e la linea continua arancione rappresenta $-\log(0.001)=3$.

- ExPlain

Il software ExPlain ha permesso di eseguire una valutazione dell'espressione dei geni rispetto ai diversi organi e tessuti. Sono stati considerati statisticamente significativi i termini con $p\text{-value} \leq 0.001$.

Come atteso, il dominio *Organs and Tissues* ha evidenziato il cervello come organo e tra i tessuti quelli appartenenti al sistema nervoso centrale.

In **Figura 24** sono riportati i primi 20 termini rappresentati in entrambe le aree cerebrali.

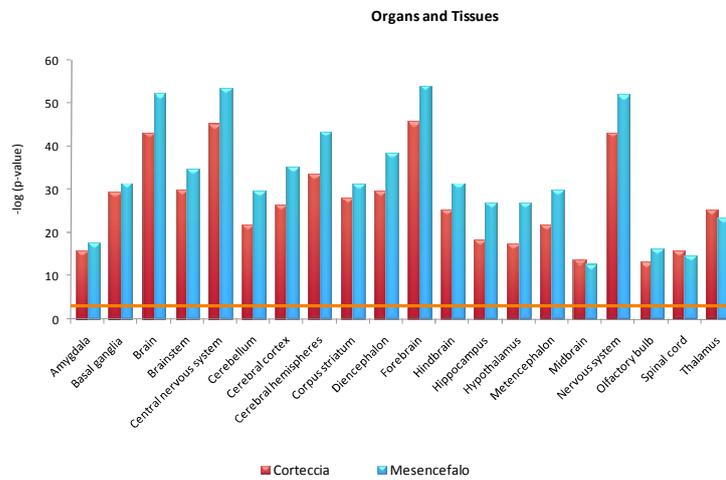


Figura 24: Explain (*Organs and Tissues*). I primi 20 termini più arricchiti nella corteccia e nel mesencefalo KO rispetto al WT. L'arricchimento della categoria è espresso come $-\log(p\text{-value})$ e la linea continua arancione rappresenta $-\log(0.001)=3$.

Le categorie più interessanti sono: talamo, ipotalamo, corpo striato ed ippocampo.

- Il talamo, collegato con l'ipotalamo e altre aree corticali, è implicato nelle vie della sensibilità somatica e della percezione sensitiva.
- L'ipotalamo controlla il sistema nervoso autonomo (riflessi, ritmo sonno-veglia, espressione degli stati emotivi) ed è interconnesso all'ipofisi che costituisce l'attività endocrina di tutto l'organismo. Tra i geni differenzialmente espressi, in particolare nel mesencefalo, si osservano downregolati vari ormoni e recettori, come il pro-ormone opioide (*Pomc*, pro-opiomelanocortina), il recettore dell'ormone della crescita (*Ghr*, growth hormone receptor) e il neuropeptide tireotropo (*Trh*, thyrotropin-releasing hormone). Tali squilibri potrebbero portare a malfunzionamento non solo del sistema nervoso centrale, ma anche di altri apparati dell'organismo come quello muscolo-scheletrico.
- Il corpo striato, noto per il suo ruolo nella pianificazione e nella modulazione dei movimenti, è anche coinvolto in una varietà di processi cognitivi relativi alla funzione esecutiva.

- L'ippocampo svolge un ruolo importante nella memoria a lungo termine e nella navigazione spaziale e nell'Alzheimer è una delle prime regioni cerebrali ad evidenziare danni.

Nella sindrome di Hunter si evidenziano ritardo cognitivo, iperattività, disturbi dell'attenzione, dell'umore e del carattere che comportano un'aggressività di tipo patologico; questi potrebbero essere imputabili all'interessamento delle strutture anatomiche evidenziate in questo studio.

Pathways

Dopo l'analisi di arricchimento per meglio approfondire le vie di segnale interessate dalla variazione di espressione è stata effettuata la pathway analysis, nelle comparazioni Cx_KO vs Cx_WT e M_KO vs M_WT, con l'utilizzo dei software PANTHER ed IPA.

- PANTHER

Per la valutazione dei pathways è stata eseguita un'analisi dei geni differenzialmente espressi nelle due aree cerebrali ottenendo per ciascuna un totale di 147 pathways. Quelli risultati statisticamente significativi ($p\text{-value} \leq 0.001$) sono riportati nelle **Figure 25** e **26**, relative, rispettivamente, alla corteccia cerebrale e al mesencefalo dell'animale IDS-KO paragonato al WT.

Nelle due aree sono risultati maggiormente rappresentati gli stessi pathways e i più interessanti sono:

- Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha mediated pathway;
- Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway;
- Muscarinic acetylcholine receptor 1 and 3 signaling pathway;
- Ionotropic glutamate receptor pathway;
- Metabotropic glutamate receptor group I pathway;
- Adrenaline and noradrenaline biosynthesis;
- Alpha adrenergic receptor signaling pathway;

- Synaptic vesicle trafficking;
- Inflammation mediated by chemokine and cytokine signaling.

Come si può notare, tutti i pathways individuati risultano implicati nella neurotrasmissione, dalla biosintesi dei neuropeptidi, passando per i loro recettori, per finire con le proteine implicate nell'attivazione della cascata di segnale del secondo messaggero che ne consegue.

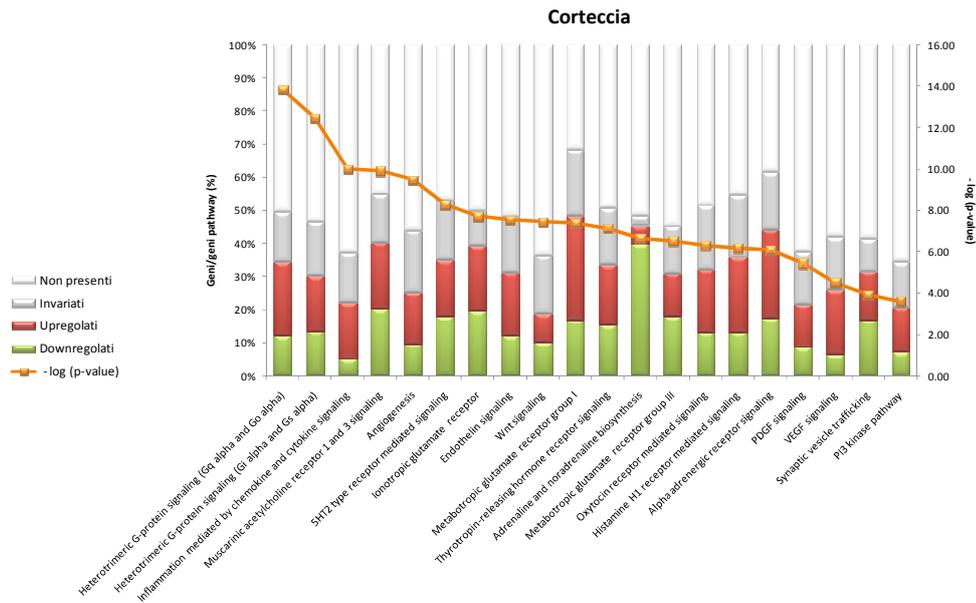


Figura 25: PANTHER (*Pathway*): I 20 pathways più rappresentativi ($p\text{-value} \leq 0.001$) nella corteccia dell'animale modello per la mucopolissaridosi di tipo II rispetto al controllo sano. In ogni pathway è mostrato il numero di geni/geni totali presenti nel pathway in percentuale. Il verde rappresenta la % di geni del pathway che risultano downregolati, il rosso gli upregolati, il grigio i geni che non soddisfano il criterio di espressione differenziale ($|\log(KO/WT)| \geq 0.7$; $p\text{-value} \leq 0.05$) e il bianco i geni non presenti/non riconosciuti nella lista totale. I punti arancione sopra ciascun istogramma rappresentano il $p\text{-value}$.

Le proteine G eterotrimeriche, come ben noto, attraverso legame con recettori transmembrana sono coinvolte in molti processi di trasmissione dei segnali chimici (ormoni, neurotrasmettitori) dall'ambiente extracellulare all'interno della cellula.

Nei due pathways riguardanti le proteine G si evidenziano soprattutto recettori a loro collegati come quelli della serotonina (Htr), recettori metabotropici del glutammato (Grm) e muscarinici per l'acetilcolina (Chrm). Inoltre, sono presenti target delle

proteine G come l'adenilato ciclastasi (Adcy), la protein chinasi (Prkc) e regolatori del segnale (Rgs, Regulator of G-protein signaling).

I pathways relativi alle G-protein sono molto ampi e comprendono già al loro interno una parte dei geni riscontrabili anche in altre vie come Muscarinic acetylcholine receptor 1 and 3 signaling, Metabotropic glutamate receptor group I e Alpha adrenergic receptor signaling.

Tutti questi recettori sono stati già associati alla neurodegenerazione riscontrabile nell'Alzheimer; nello specifico, il recettore della serotonina 7 (Htr7) e Grm1 risultano downregolati mentre Grm2 e 3 upregolati. Tali variazioni sono state confermate con la stessa direzionalità anche nel presente lavoro.

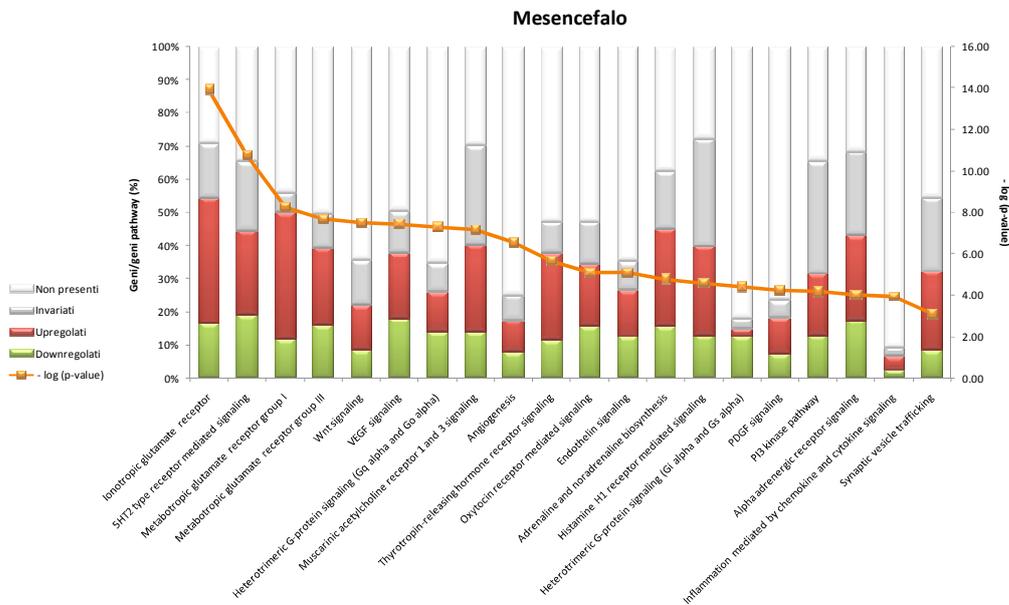


Figura 26: PANTHER (Pathway): I 20 pathways più rappresentativi ($p\text{-value} \leq 0.001$) nel mesencefalo dell'animale modello per la mucopolissacaridosi di tipo II rispetto al controllo sano. In ogni pathway è mostrato il numero di geni/geni totali presenti nel pathway in percentuale. Il verde rappresenta la % di geni del pathway che risultano downregolati, il rosso gli upregolati, il grigio i geni che non soddisfano il criterio di espressione differenziale ($|\log(KO/WT)| \geq 0.7$; $p\text{-value} \leq 0.05$) e il bianco i geni non presenti/non riconosciuti nella lista totale. I punti arancione sopra ciascun istogramma rappresentano il p-value.

Per quanto riguarda la via metabolica dell'adrenalina e noradrenalina i geni maggiormente rappresentati sono quelli coinvolti, non solo nella biosintesi delle

molecole (Th, Tyrosin hydroxylase), ma anche nella costituzione delle vescicole sinaptiche (Slc18a2, Synaptic vesicular amine transport; Vamp, Vesicle associated membrane protein) dove avviene il completamento della loro sintesi. Questi geni risultano tutti downregolati, sia nella corteccia cerebrale sia nel mesencefalo patologico.

La Vamp, come atteso, è rappresentata anche nel pathway Synaptic vesicle trafficking, associata a sinaptotagmine (Syt) e sintassine (Stx). Tali geni sono implicati nel targeting e nella fusione delle vescicole sinaptiche con le membrane presinaptiche, facendo intuire una possibile alterazione della trasmissione del segnale.

Quest'ultima ipotesi è supportata dal fatto che il mantenimento della memoria e dell'apprendimento è garantito dalla plasticità sinaptica, nella quale il recettore ionotropico del glutammato sembra avere un ruolo cruciale. Il primo meccanismo che regola la plasticità sinaptica, sia nelle sinapsi eccitatorie sia inibitorie, è dipendente dal calcio e i geni principalmente alterati nella via di segnale ionotropica sono i canali del Ca^{2+} voltaggio dipendenti (Cacng, Voltage-dependent calcium channel; Cacn, Voltage-dependent N-type calcium channel). Il secondo meccanismo è basato su una cascata di secondo messaggero che regola la trascrizione genica di proteine chiave per le sinapsi, come le protein-chinasi calcio/calmodulina dipendenti (Camk2), anch'esse presenti in questa lista.

Inflammation mediated by chemokine and cytokine signaling è il pathway che presenta maggiori differenze tra le due macroaree. Sembra, infatti, che il mesencefalo evidenzi una minore alterazione dei geni coinvolti in questa via di segnale, portando a pensare ad una più importante infiammazione della corteccia cerebrale rispetto al mesencefalo.

- IPA

Il secondo software utilizzato per l'analisi dei pathways è IPA.

Le vie di segnale identificate sono 215 nella corteccia cerebrale e 216 nel mesencefalo IDS-KO rispetto al controllo. In **Figura 27** e **28** sono riportati i primi 25 pathways annotati, statisticamente significativi ($p\text{-value} \leq 0.001$).

Anche questa analisi evidenzia il coinvolgimento delle vie di segnale delle proteine G, del recettore per la serotonina e per il glutammato, incrementando tale interessamento e aggiungendo più specificamente il segnale mediato dall'AMP ciclico e dalla fosfolipasi C. Tutti questi pathways sono legati anche alla via di segnale del calcio, che risulta ampiamente alterata soprattutto nel mesencefalo patologico.

Il Calcio (Ca^{2+}) è un comune secondo messaggero che regola molti processi cellulari (contrazione, secrezione, trasmissione sinaptica, regolazione dei pori nucleari, trascrizione) e i geni risultati differenzialmente espressi comprendono recettori per i neurotrasmettitori [oltre a quelli già trattati anche i colinergici (Chrn, cholinergic receptor nicotinic)], i canali del calcio (Itpr1, inositol 1,4,5-triphosphate receptor, type 1) e alcuni geni codificanti proteine che regolano la trascrizione, come CREB binding protein (Crebbp) e cAMP responsive element binding protein 3 (Creb3).

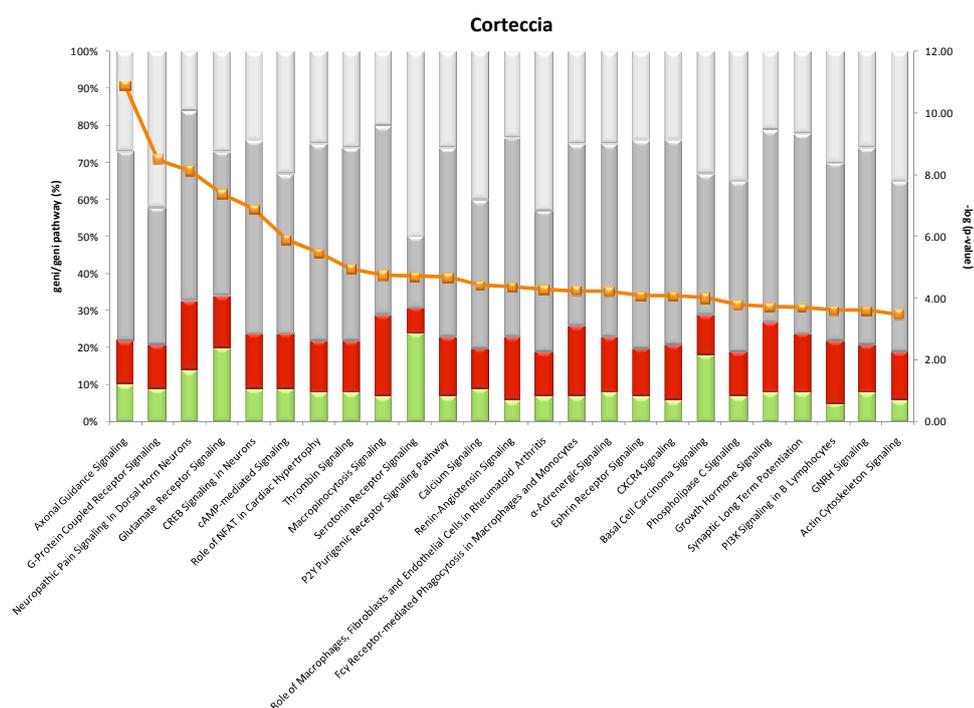


Figura 27: IPA (Pathway): I primi 25 pathways più rappresentativi ($p\text{-value} \leq 0.001$) nella corteccia dell'animale patologico rispetto al controllo sano. In ogni pathway è mostrato il numero di geni/geni totali presenti nel pathway in percentuale. Il verde rappresenta la % di geni del pathway che risultano downregolati, il rosso gli upregolati, il grigio i geni che non soddisfano il criterio di espressione

differenziale ($|\log(\text{KO}/\text{WT})| \geq 0.7$; $p\text{-value} \leq 0.05$) e il bianco i geni non presenti/non riconosciuti nella lista totale. I punti arancione sopra ciascun istogramma rappresentano il p-value.

Il pathway Neuropathic Pain Signaling in Dorsal Horn Neurons è rappresentato da geni coinvolti nella via dei neurotrasmettitori del glutammato e Bdnf (brain derived neurotrophic factor) che, rilasciati dai terminali presinaptici, inducono un aumento generale della sensibilità dei neuroni agli input afferenti e al mantenimento di un'elevata eccitabilità e responsività sinaptica.

La via Axonal Guidance concerne processi che portano i neuroni a costituire assoni direzionati verso un target definito.

Gli assoni spesso seguono precise vie nel sistema nervoso e tale processo è regolato da fini e delicati meccanismi. Le categorie di molecole e recettori che guidano la crescita assonale attrattiva e repulsiva sono molte, le principali si possono suddividere in Netrine, Efrine e Semaforine. Nella valutazione dell'espressione genica tutte queste categorie risultano alterate nelle aree cerebrali dell'animale IDS-KO rispetto al WT.

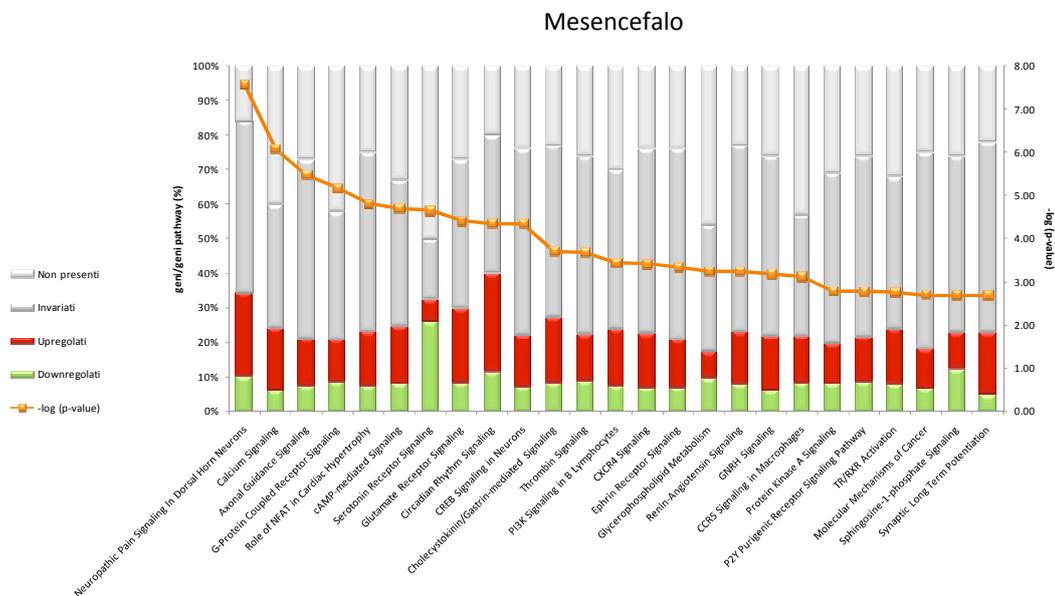


Figura 28: IPA (Pathway): I primi 25 pathways più rappresentativi ($p\text{-value} \leq 0.001$) nella corteccia dell'animale patologico rispetto al controllo sano. In ogni pathway è mostrato il numero di geni/geni totali presenti nel pathway in percentuale. Il verde rappresenta la % di geni del pathway che risultano

dowregolati, il rosso gli upregolati, il grigio i geni che non soddisfano il criterio di espressione differenziale ($|\log(\text{KO}/\text{WT})| \geq 0.7$; $p\text{-value} \leq 0.05$) e il bianco i geni non presenti/non riconosciuti nella lista totale. I punti arancione sopra ciascun istogramma rappresentano il p-value.

Analisi delle vie di segnale alterate

Le analisi effettuate in questo studio hanno permesso di confermare e ampliare la visione di insieme delle vie di segnale coinvolte nelle alterazioni dell'espressione genica della corteccia cerebrale e del mesencefalo dell'animale modello IDS-KO. Sebbene richiederà ulteriori approfondimenti, indispensabili per delucidare dubbi derivati dalla scarsa conoscenza delle interazioni tra le molecole di alcuni pathways, tale studio porrà nuove domande sulla patofisiologia della mucopolisaccaridosi di tipo II.

In **Appendice** sono riportati i geni caratteristici delle vie di segnale analizzate.

- Eparan solfato binding proteins

I glicosaminoglicani (GAGs) che si accumulano nella MPS II, eparan e dermatan solfato (HS e DS), svolgono la loro normale funzione fisiologica nella matrice extracellulare dove sono associati a core proteici trans-membrana a formare i proteoglicani. I proteoglicani, in particolare quelli contenenti HS, svolgono un ruolo importante nel signaling cellulare sia come recettori sia co-recettori di diversi ligandi. Il legame con i ligandi è specifico ed è determinato dal tipo di glicosaminoglicano, dalla sua sequenza, dalla solfatazione e dalla struttura tridimensionale. Le attività biologiche modulate mediante l'interazione HS-proteine sono molteplici: proliferazione, sopravvivenza, migrazione, adesione, differenziamento, riconoscimento cellulare, angiogenesi, infiammazione, meccanismi di difesa, endocitosi, assemblaggio della matrice extracellulare [Dreyfuss et al. 2009].

Nella mucopolisaccaridosi di tipo I, patologia simile all'MPS II e con storage degli stessi substrati, è stato osservato un accumulo nella matrice extracellulare di tali molecole sottoforma di frammenti e con ridotta solfatazione. Questo potrebbe alterare i pathways regolati dai proteoglicani con HS.

Infatti, alcune delle porzioni proteiche costituenti i proteoglicani presentano espressione alterata nelle due aree cerebrali analizzate. I sindecani, i glicopirani, GAG di

superficie cellulare, e i testicani GAG di matrice extracellulare, risultano tendenzialmente downregolati. Nel contesto di questa analisi risulta particolarmente interessante la downregolazione di Spock3 (Sparc/osteonectin, cwcv and kazal-like domains proteoglycan (testican) 3) che appartiene ad una famiglia di testicani che legano il Ca^{2+} e sembra essere coinvolto in diversi step della neurogenesi.

Tra i fattori che legano l'eparan solfato risultano differenzialmente espressi vari tipi di collagene, il che evidenzia il coinvolgimento dei principali tessuti connettivi, la laminina, coinvolta in diversi processi biologici come adesione cellulare, differenziamento, migrazione, signaling e crescita dei neuriti, e la trombospondina associata al signaling cellulare nello sviluppo e nel sistema nervoso adulto. Anche diversi fattori di crescita che legano l'eparan solfato, quali FGF (Fibroblast Growth Factor, 9-12-13), PDGF β (Platelet-derived growth factor beta polypeptide), TGF β (Transforming growth factor, beta) presentano un'espressione differenziale.

Tra i fattori di crescita risulta interessante FGF9, che è downregolato in entrambe le aree e ha un ruolo nella crescita e nel differenziamento della glia, nella gliosi, nella rigenerazione del tessuto cerebrale dopo danno e nel differenziamento e nella sopravvivenza delle cellule neuronali. FGF12, probabilmente coinvolto nello sviluppo e nella funzione del sistema nervoso centrale è downregolato nella corteccia. FGF13, upregolato nel mesencefalo, è associato a ritardo mentale X-linked.

- *Omeostasi del calcio*

Il mantenimento di una corretta omeostasi del calcio nella cellula risulta fondamentale in quanto il calcio rappresenta un messaggero ubiquitario coinvolto in differenti processi biologici dalla trascrizione genica alla contrazione muscolare, dalla proliferazione alla morte cellulare. In **Figura 29** sono rappresentate le principali vie di segnale coinvolte nell'omeostasi del calcio

Le due principali fonti di Ca^{2+} sono l'ambiente extracellulare e i depositi intracellulari.

L'entrata di Ca^{2+} dall'esterno avviene attraverso l'attivazione di canali di membrana plasmatica di varia natura tra cui i voltage-gated channels e i ligand-gated channels, questi ultimi attivati dal legame con ligandi specifici; ne sono esempi i recettori ionotropici per il glutammato (NMDAR/Grin, AMPAR/Gria e Kainate receptor/Grik), il

recettore nicotinico per l'acetilcolina (Chrn β), il recettore per la serotonina (Htr), il recettore per l'ATP (P2rx) e per le diidropiridine (Cacng). Altri trasportatori di membrana sono il canale del Ca²⁺ attivato dal rilascio di Ca²⁺ (CRAC/Orai) e i TRP (Transient Receptor Potential), canali cationici non specifici che rispondono indirettamente o direttamente a ormoni, neurotrasmettitori e a numerosi stimoli sensoriali. Altri stimoli sensoriali operano attraverso la cascata del segnale che include l'attivazione di adenilato ciclasi e guanilato ciclasi e la conseguente apertura o chiusura di canali calcio CNG (cyclic nucleotide gated) regolati dai nucleotidi ciclici cAMP o cGMP.

In contrasto alla varietà di meccanismi che inducono l'entrata di Ca²⁺ nella cellula, l'estrusione è deputata essenzialmente a due famiglie proteiche: PMCA (Plasma Membrane Ca²⁺ ATPase) e NCX (scambiatore Na⁺/Ca²⁺).

I depositi di Ca²⁺ intracellulari sono localizzati principalmente nel reticolo sarcoplasmatico. Il rilascio di Ca²⁺ da questo comparto è regolato dai recettori per la rianodina (Ryr) e dal recettore dell'inositolo trifosfato (Ip3r) che rilascia Ca²⁺ in seguito al legame con l'IP3. I primi sono canali permeabili al Ca²⁺ attivati principalmente dal Ca²⁺ stesso, ma anche da altri segnali come cADP (ADP-ribosio ciclico) e sfingosina mentre il secondo coinvolge l'attivazione di una fosfolipasi (PLC) e la produzione di IP3 e DAG da PIP2 (Phosphatidylinositol (4, 5)-Bisphosphate).

Gli ormoni e i neurotrasmettitori si legano a recettori accoppiati alle proteine G (GPCR) e attivano PLC β . I fattori di crescita possono attivare PLC γ e PLC ϵ , legandosi ai recettori tirosi-chinasici (RTK) e PLC δ legandosi ai recettori TRP. PLC γ è attivata anche dal legame di antigeni a recettori non-RTK come TCR (T-cell receptors) e BCR (B-cell receptor) [Nowycky and Thomas, 2002].

L'uptake nel reticolo endoplasmatico è regolato dalla famiglia Serca (sarco- and Endoplasmic Reticulum Ca²⁺ ATPase).

Anche gli altri organelli presentano trasportatori specifici per il Ca²⁺. I mitocondri rilasciano Ca²⁺ dallo scambiatore Na⁺-H⁺/Ca²⁺ e, in alcune circostanze, dal poro di transizione di permeabilità, mentre l'uptake è mediato dal trasportatore uniporto mitocondriale.

All'interno degli organelli gli ioni Ca^{2+} sono legati a proteine tampone come la calsequestrina (Cs), la calreticulina (Cr) e la calnessina (Cn), mentre nel citosol la stessa funzione è svolta dalla calmodulina (Calm) e dalla famiglia proteica delle S100.

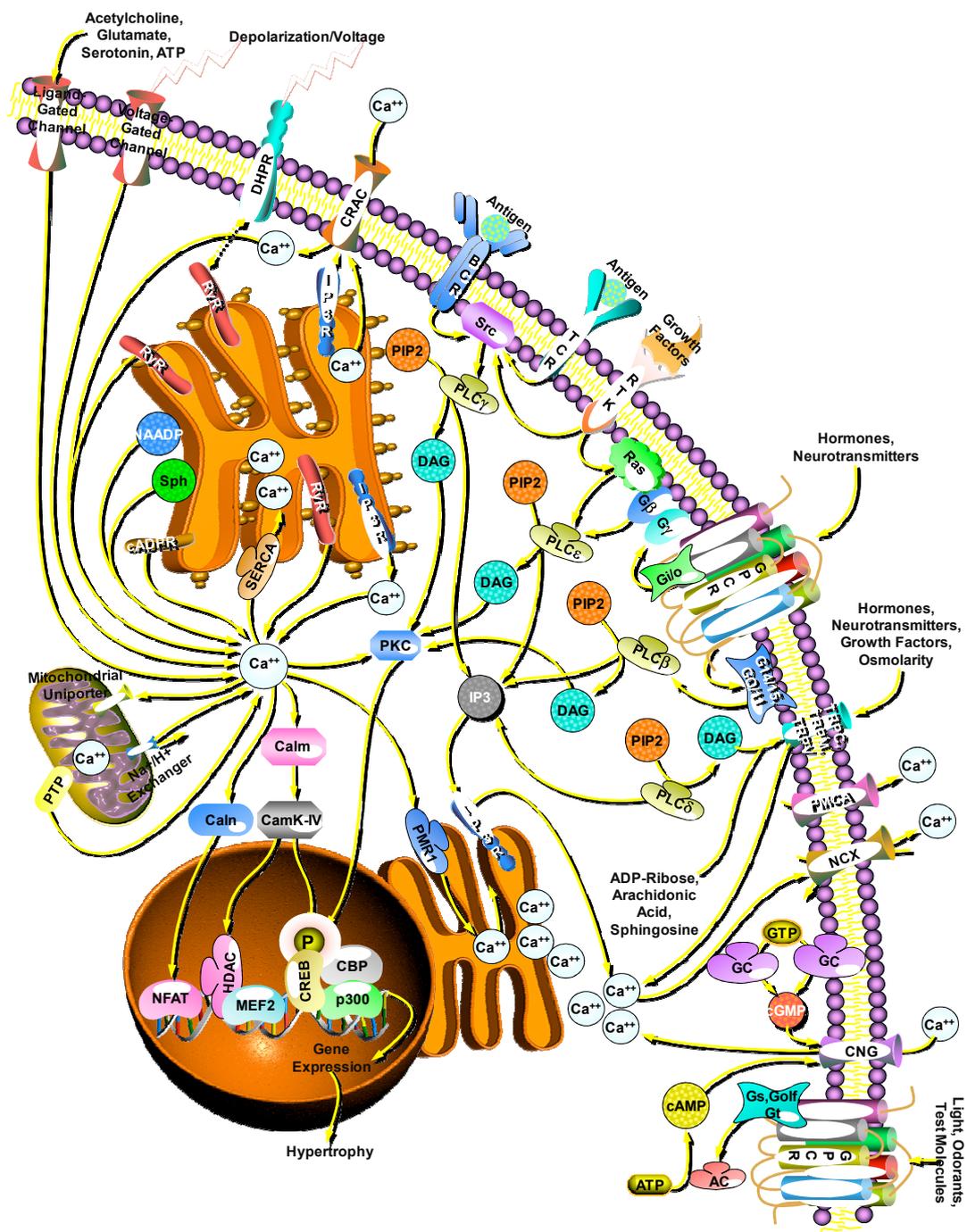


Figura 29: Via di segnale del Calcio [<http://www.sabiosciences.com>];

© 2009 Sabioscience, all rights reserved].

I segnali di Ca^{2+} citoplasmatico si propagano al nucleo dove avviene la trascrizione di geni target. L'ingresso di Ca^{2+} attraverso i recettori sinaptici Grin attiva Creb (c-AMP Response Element-Binding Protein) mentre l'ingresso attraverso i recettori Gria extrasinaptici inibisce Creb. I canali attivati da ligando usano la calmodulina per percepire le variazioni locali di Ca^{2+} e attivare il pathway PKA-dipendente Rap1-MAPK/ERK (Mitogen-Activated Protein Kinase/ Extracellular Signal-Regulated Kinase). La fosforilazione, ad opera di CalmKII e CalmKIV (calmodulina chinasi-II e IV) di CBP (CREB Binding Protein, Crebbp) e delle istone deacetilasi HDAC4, HDAC5 e HDAC7 media alcuni segnali Ca^{2+} nucleari. La fosforilazione delle HDAC permette a Mef2 (Myocyte Enhancing Factor-2) di attivare la trascrizione reclutando, nella regione CRE (cAMP Responsive Element), altri fattori sensibili al Ca^{2+} come NFAT (Nuclear Factor of Activated T-Cells) e coattivatori trascrizionali come p300. Il Ca^{2+} nucleare attiva anche la calcineurina (Caln) che defosforila NFAT e promuove l'attività trascrizionale.

Il Ca^{2+} è un importante secondo messaggero in alcuni processi come la contrazione muscolare e la secrezione di ormoni e peptidi. Alterazioni funzionali fini di diversi segnali del Ca^{2+} sono state associate alle condizioni patofisiologiche delle distrofinopatie, della malattia di Brody, di alcune forme di miopatie e dell'ipertermia maligna. La calpaina, Ca^{2+} binding protein del muscolo, è stata associata a diverse patologie quali la distrofia muscolare di tipo 2A delle giunture degli arti e il diabete di tipo 2. Inoltre, deficit dei sensori di calcio neuronali sono responsabili della degenerazione della retina [Berchtold et al., 2000].

La via di segnale del calcio risulta altamente compromessa in entrambe le aree e a diversi livelli, con alterazione dei trasportatori di membrana plasmatica, mitocondriale e del reticolo endoplasmatico. Il risultato delle variazioni di questi geni sembra comportare un aumento della concentrazione del calcio citosolico, in accordo con quanto osservato in altre LSDs.

Nella corteccia cerebrale, ma soprattutto nel mesencefalo, si riscontra upregolazione dei recettori NMDA per il glutammato (Grin) e dei canali voltaggio dipendenti Cacna, Cacnb e Cacng, che portano all'internalizzazione nel citoplasma degli ioni calcio.

Si evidenzia anche un aumento del rilascio del calcio dal reticolo endoplasmatico (ER) dovuto all'upregolazione del recettore della rianodina (RYR) e dell'inositolo-1,4,5-

trifosfato (IP3R). Analoghe variazioni sono state osservate anche in altre patologie lisosomiali come le malattie di Gaucher e di Niemann-Pick A e le gangliosidosi GM1 e GM2 [Nguyen et al., 2002; Ginzburg et al., 2005; Ballabio and Gieselmann, 2009].

Si può ipotizzare che la via che porta all'attivazione di IP3R sia quella mediata dai recettori accoppiati alle proteine G che, in seguito al legame del ligando (ormone o neurotrasmettitore), portano all'attivazione della PLC β con produzione di IP3. I geni che codificano per GPCR, PLC β e IP3R risultano, infatti, upregolati nei campioni analizzati.

Anche la via di segnale legata all'attivazione della trascrizione, indotta dal calcio, risulta alterata nella corteccia, ma soprattutto nel mesencefalo, con upregolazione dei geni Crebbp, Mef2 e Hdac4-5-7.

Lo scambiatore sodio/calcio (Slc8a2) presente sulla membrana plasmatica e mitocondriale si presenta upregolato sia nella corteccia cerebrale che nel mesencefalo dei topi IDS-KO. Sottoposto a stress, questo scambiatore può subire un'inversione di funzionalità con conseguente aumento del calcio sia all'interno dei mitocondri sia nel citoplasma cellulare, confermando l'alterazione dell'omeostasi mitocondriale del Ca²⁺ riscontrata in almeno due LSDs. In un modello neuronale per la GM1 gangliosidosi, l'accumulo di GM1 comporta un'elevata permeabilizzazione della membrana mitocondriale con conseguente apertura dei pori di membrana e attivazione del pathway apoptotico. Anche nella mucopolipidosi di tipo IV, la frammentazione dei mitocondri e l'alterazione del Ca²⁺ mitocondriale porta ad un'attivazione di segnali pro-apoptotici [Vitner et al., 2010].

In un recente lavoro Pereira et al. [2010] hanno dimostrato una variazione del Ca²⁺ e degli H⁺ nei topi IDUA-KO (mucopolisaccaridosi di tipo I) con un incremento dell'accumulo di Ca²⁺ nel reticolo endoplasmatico e nei lisosomi. Un rilascio differenziale di calcio dall'ER può rappresentare un accumulo compensatorio di ioni, come risultato di una fisiologia lisosomiale compromessa e una conseguente alterazione degli ioni citosolici. Tra le differenti funzioni del calcio nella fisiologia cellulare, il ruolo di questi ioni nel pathway endocitico può essere di grande importanza nella patofisiologia delle LSDs. La fusione tra gli endosomi precoci e gli endosomi tardivi e tra questi ultimi e i lisosomi è un importante step nel pathway

endocitico ed è stato dimostrato che entrambi gli eventi richiedono il rilascio di calcio dal lume degli organelli fusi. Gli organelli privi di Ca^{2+} non sono in grado di fondersi e le alte concentrazioni di Ca^{2+} citosoliche possono a loro volta inibire questo processo [Pryor et al., 2000].

È plausibile pertanto che nei topi IDS-KO ci sia uno squilibrio delle concentrazioni di Ca^{2+} che risultano aumentate negli organelli cellulari. Questa alterazione dell'omeostasi porta ad aumentato stress ossidativo e, di conseguenza, ad un'aumentata autofagia e morte cellulare.

- Stress ossidativo

Lo stress ossidativo consiste in uno sbilanciamento tra la produzione di specie reattive dell'ossigeno e la capacità di un sistema biologico di detossificare prontamente gli intermedi reattivi o di riparare i danni risultanti. Disturbi del normale stato redox dei tessuti possono causare effetti tossici mediante la produzione di perossidi e radicali liberi che danneggiano tutti i componenti della cellula compresi proteine, lipidi e DNA. Lo stress ossidativo (**Figura 30**) è indotto nelle cellule come conseguenza di uno di questi tre fattori:

1. un aumento della produzione di ossidanti;
2. una diminuzione della protezione antiossidante;
3. la mancata riparazione da danni ossidativi.

Il danno cellulare è indotto da specie reattive dell'ossigeno (ROS). I ROS sono radicali liberi, cioè anioni reattivi contenenti atomi di ossigeno, oppure molecole che possono produrre o essere attivate chimicamente dai radicali liberi; esempi sono il radicale idrossile, il superossido, il perossido di idrogeno e il perossinitrito. La principale fonte di ROS in vivo è la respirazione aerobica, anche se i ROS sono prodotti anche dalla beta-ossidazione degli acidi grassi, dal metabolismo degli xenobiotici mediato da citocromo P450, dalla stimolazione della fagocitosi da parte di patogeni o lipopolisaccaridi, dal metabolismo dell'arginina e da enzimi tessuto-specifici. In condizioni normali, i ROS sono eliminati dalla cellula per azione della superossido dismutasi (Sod), dalla catalasi o dalla glutatione perossidasi (GSH). I danni principali alle

cellule sono dovuti ad alterazioni indotte da ROS di macromolecole quali gli acidi grassi nelle membrane lipidiche, proteine essenziali e DNA. Inoltre, lo stress ossidativo e i ROS sono stati implicati in stati patologici come il morbo di Alzheimer, il morbo di Parkinson, il cancro e l'invecchiamento.

Nella corteccia cerebrale e nel mesencefalo dei topi IDS-KO rispetto ai WT sono risultate variare in modo statisticamente significativo la maggior parte delle molecole coinvolte in questo pathway. In particolare sono notevolmente variate la Ptgs2 (NADPH P450 reductase,) che è un enzima chiave nella biosintesi delle prostaglandine ed è coinvolta nell'infiammazione e nella mitogenesi; la LPO (lactoperoxidase), che è un agente antibatterico naturale; le Gpx 3 e 6 (GSH peroxidase), che agiscono nella detossificazione del perossido d'idrogeno e sono espresse rispettivamente nel fegato e nell'epitelio olfattivo. Le altre molecole hanno un livello di espressione inferiore, anche se statisticamente significativo, indicando una complessiva implicazione di questo pathway nella patofisiologia dell'MPS II, che dovrà però essere studiato ad un livello maggiore per meglio identificarne il coinvolgimento.

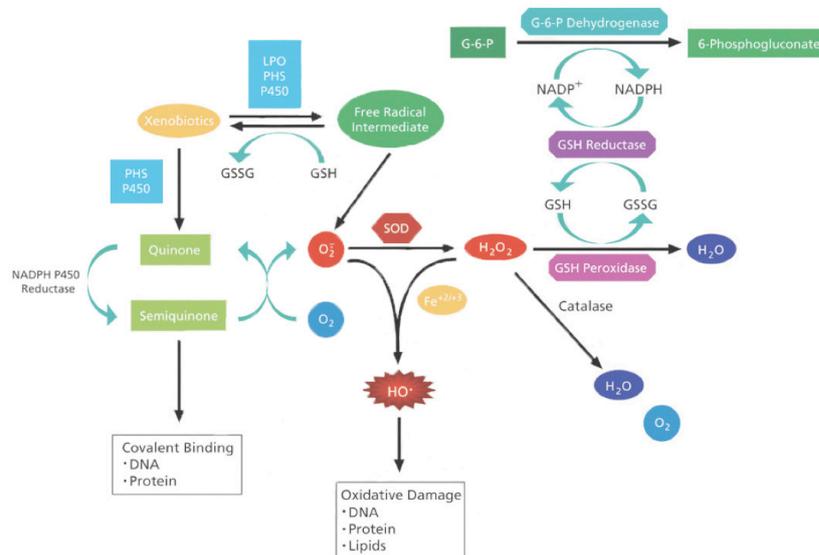


Figura 30: Pathway dello stress ossidativo [www.sigmaaldrich.com].

Evidenze riportate in diversi lavori suggeriscono che le specie reattive dell'ossigeno (ROS) giocano un ruolo fondamentale e sono mediatori comuni della morte cellulare nelle LSDs. Nella gangliosidosi GM1 e GM2, Nos (inducibile nitric oxide synthase) è

elevata nella microglia attivata e i ROS sono elevati nei modelli per la malattia di Fabry. L'analisi con microarray su fibroblasti derivanti da pazienti affetti da Niemann-Pick (NPC1) è coerente con un aumentato stress ossidativo, con ROS elevati e perossidazione lipidica, rendendo i fibroblasti più suscettibili alla morte cellulare dopo un acuto insulto ossidativo. Nella MPS III B, potenziali stress ossidativi provocano l'ossidazione delle proteine, dei lipidi e del DNA; uno squilibrio è stato rilevato anche nella MPS I [Vitner et al., 2010].

Il ruolo centrale che lo stress ossidativo gioca nell'integrare altri pathways cellulari suggerisce che, nelle LSDs, è più probabilmente attivato come via biochimica secondaria, piuttosto che come un risultato diretto di accumulo del substrato primario. Inoltre, il possibile ruolo dello stress ossidativo può essere di reale importanza nel delineare la patofisiologia delle LSDs, soprattutto dal momento che esso gioca un ruolo principale nelle malattie neurodegenerative maggiormente studiate [Vitner et al., 2010].

- Autofagia

L'autofagia è un pathway catabolico lisosoma-dipendente attraverso il quale organelli come i mitocondri e le proteine invecchiate sono sequestrati in vescicole con doppia membrana (gli autofagosomi) e infine degradati dopo la fusione autofagosoma-lisosoma. Il malfunzionamento dell'autofagia è stato collegato a un ampio range di patologie umane, compreso il cancro, la neurodegenerazione e l'infezione patogena.

La maggior parte degli aggregati proteici che causano una condizione neurodegenerativa ad insorgenza tardiva sono substrati autofagici [Rubinsztein, 2006]. Settembre et al. [2008] hanno dimostrato che, nel modello murino per la Mucopolisaccaridosi di tipo III A (MPS III A) e per la Deficienza Multipla di Solfatasi (MSD), un blocco del pathway dell'autofagia è la conseguenza di una diminuzione della capacità dei lisosomi di fondersi con gli autofagosomi. Questo comporta l'accumularsi di corpi di inclusione che contengono proteine poliubiquitinate e substrati tossici che sono i mediatori putativi della morte cellulare.

L'accumulo di proteine ubiquitinate è stato osservato in topi KO per geni dell'autofagia. Questi topi mostrano una severa neurodegenerazione, suggerendo che i

neuroni, in confronto ad altri tipi cellulari, sono più suscettibili al blocco di questo pathway. Questo potrebbe essere dovuto al fatto che cellule non mitotiche non possono diluire l'accumulo citosolico attraverso la divisione cellulare [Settembre et al., 2008].

È interessante notare che tra i dati ottenuti nel presente lavoro di RNA-Seq, A53T α -sinucleina (Snca), una proteina coinvolta nella morte dei neuroni dopaminergici e implicata in diverse malattie neurodegenerative [Dev et al., 2003], è upregolata in modo significativo sia nella corteccia sia nel mesencefalo dei topi IDS-KO rispetto ai topi WT, avvalorando l'ipotesi di similarità tra le MPS e le malattie neurodegenerative croniche, proposta in molti lavori.

Il mantenimento della morfologia e della funzionalità neuronale dipende da un efficiente trasporto intracellulare poiché i maggiori compartimenti di sintesi e degradazione sono localizzati nel soma.

La Snapina riveste un ruolo critico nel coordinare il trasporto endocitico tardivo e la maturazione dei lisosomi, due processi cellulari dinamici richiesti per il corretto funzionamento del sistema autofagico-lisosomiale nei neuroni. La deficienza di Snapina comporta una ridotta vitalità neuronale e neurodegenerazione [Cai et al., 2010].

Poiché i lisosomi sono localizzati prevalentemente nel soma e nella regione prossimale dei neuroni, il trasporto retrogrado dineina-dipendente (Dyn) è critico per il delivery del carico degli endosomi tardivi dalla periferia cellulare e dai processi distali al soma, portando gli endosomi tardivi e i lisosomi in prossimità sufficiente per fondersi. La perturbazione della funzione di dineina-dinactina (Dctn) comporta la dispersione degli organelli endocitici tardivi nella periferia e la proliferazione lisosomiale, danneggia la clearance degli aggregati proteici imputabile alla disfunzionalità autofagica-lisosomiale ed induce la degenerazione assonale. Nei neuroni in cui la Snapina è downregolata, l'associazione degli endosomi tardivi con la dineina è significativamente compromessa [Cai et al., 2010].

La mutazione della dineina o il mal funzionamento dell'autofagia porta una rilevante neurodegenerazione; il trasporto retrogrado Snapina-mediato e dineina-guidato degli endosomi tardivi è cruciale per lo sviluppo e la sopravvivenza neuronale in quanto

un'inibizione dell'autofagia. Anche i geni coinvolti nell'attivazione dell'autofagia in caso di ipossia (**Figura 32A**) sono downregolati (Hif, Hypoxia-inducibile factor; E2f5-6, Transcriptional factor E2f5-6 ; Prkcd, Protein kinase c delta; Bnip3, Bcl-2 adenovirus E1a 19kDa interacting protein 3) o upregolati (Tsc1-2, tuberous sclerosis complex 1-2) in direzione tale da indurre un'inibizione dell'autofagia.

Nel caso della risposta allo stress (**Figura 32B**), invece, non è ancora chiaro in che direzione vada la regolazione in quanto sono stati trovati sia geni che essendo downregolati comportano l'inibizione dell'autofagia (Prkcq, Protein kinase c theta ; Eif2ak2, Eukariotic translation initiation factor 2-alfa kinase 2), sia geni che upregolati ne comportano l'induzione (Camkk2, Calcium/calmodulin-dependent protein kinase 2; Dapk1, Death-associated kinase 1); soprattutto Camkk2 sembra influenzare fortemente questo processo sia nella corteccia che nel mesencefalo dei topi IDS-KO.

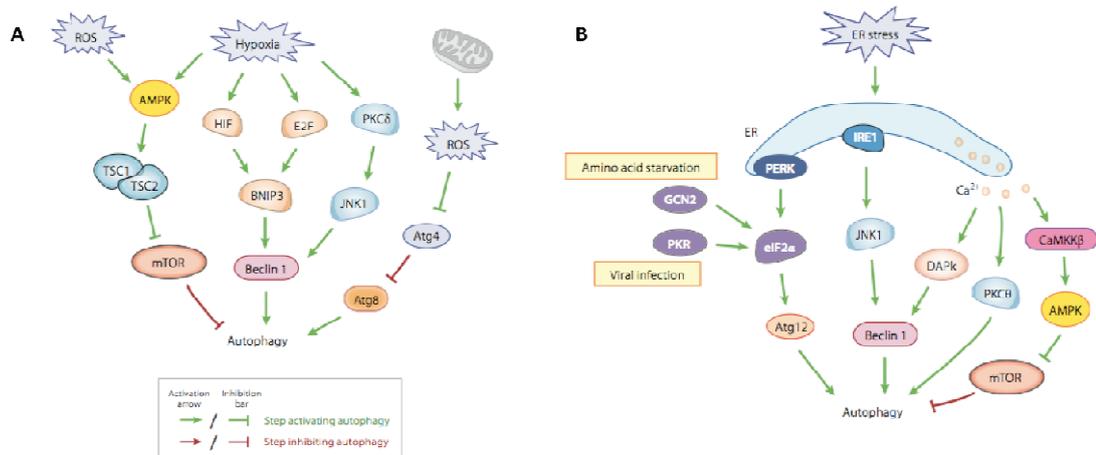


Figura 32: Rappresentazione schematica della regolazione dell'autofagia in situazione di ipossia (**A**) e di stress ossidativo (**B**) [He and Klionsky, 2009].

L'autofagia a livello di corteccia e mesencefalo nei topi IDS-KO sembra, pertanto, essere alterata a diversi livelli, non solo nell'unione autofagosoma-lisosoma, ma anche nel trafficking e nella formazione dell'autofagosoma, portando generalmente ad una inibizione di questo processo.

- Axon guidance

L'axon guidance è una fase chiave dello sviluppo neurale; è il processo mediante il quale i neuroni ridirigono i propri assoni verso la destinazione finale corretta. In corrispondenza dell'estremità di crescita gli assoni presentano una struttura altamente mobile detta cono di crescita. Il cono di crescita è ricco in recettori per una serie di molecole segnale che istruiscono l'assone sulla direzione da prendere. Queste molecole possono essere fisse o diffusibili e possono avere effetto attrattivo o repulsivo sull'assone. In ogni caso si induce, nell'assone, una risposta chemiotropica che si traduce con un rimodellamento del citoscheletro. Le molecole attrattive inducono una polimerizzazione dell'actina e un'estensione dell'assone, mentre quelle repulsive riducono la polimerizzazione dell'actina risultando nel collasso del cono di crescita e nella retrazione dell'assone. La presenza all'esterno di gradienti di molecole segnale porta ad una trasmissione asimmetrica del signalling intracellulare nel cono di crescita permettendo la regolazione della direzionalità dell'assone.

Sono state identificate quattro principali famiglie di molecole segnale responsabili dell'axon guidance:

- le netrine, molecole secrete che si legano ai recettori DCC (deleted in colonrectal carcinoma) (con effetto attrattivo) e Unc5 (unc-homolog C) (con effetto repulsivo). Il legame del ligando al recettore media anche un segnale anti-apoptotico;
- le efrine, molecole di superficie cellulare che attivano i recettori Eph (Ephrine receptor) sulla superficie di altre cellule. Questa interazione può essere sia attrattiva che repulsiva; in alcuni casi le efrine possono fungere da recettore e Eph da ligando;
- le semaforine, principalmente chemiorepellenti per l'assone. Attivano complessi di recettori di superficie chiamati Plexine e Neurofiline;
- le Slit (Srgap), proteine secrete che inducono repulsione legando il recettori Robo (roundabout homolog). [Bashaw, 2010].

Queste molecole sfruttano vie di segnalazione diverse illustrate in **Figura 33**.

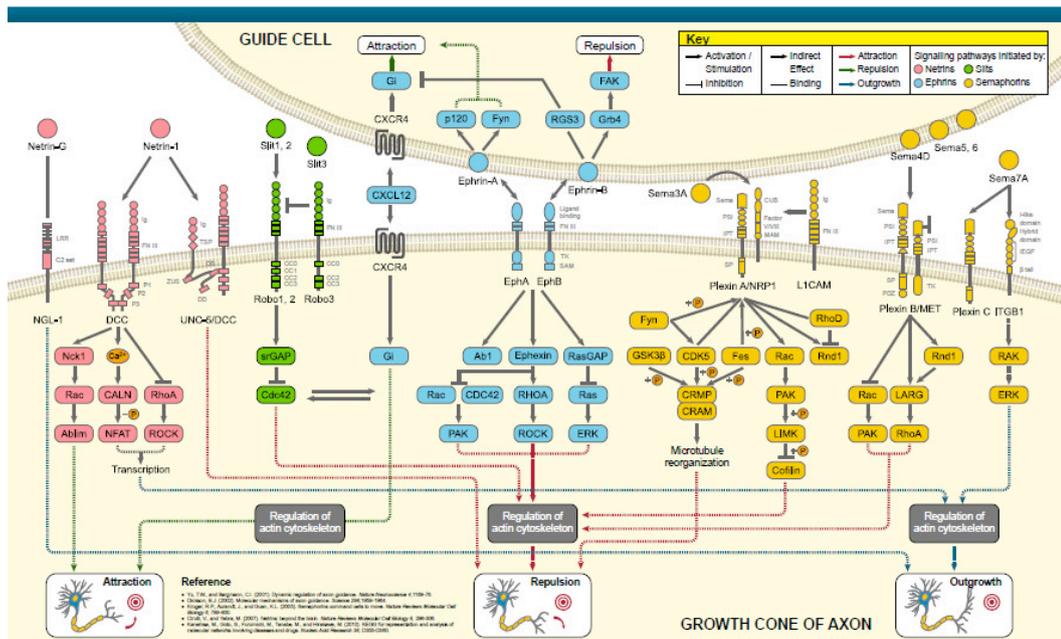


Figura 33: Axon guidance signaling [www.novusbio.com].

L'axon guidance risulta molto interessato dalle variazioni di espressione genica analizzate in questo lavoro. Le **Figure 34** e **35** rappresentano il pathway dell'axon guidance ricavato dal software KEGG, riferito alle aree corteccia e mesencefalo rispettivamente; i geni differenzialmente espressi sono mostrati in rosso, in verde sono indicati i geni upregolati nel campione IDS-KO rispetto al WT, in verde i downregolati.

Come si può notare le quattro famiglie di molecole segnale risultano tutte interessate dalla variazione. In particolare, si evidenzia una downregolazione delle netrine (Ntn1 e Ntn1g1) e delle semaforine, un'upregolazione delle efrine e delle slit e alterazioni delle vie di segnale a vari livelli.

Oltre a queste molecole sono noti anche molti altri fattori coinvolti nell'axon guidance:

- morfogeni dello sviluppo, come bone morphogenetic protein (Bmp), wingless-type MMTV integration site family (Wnt), Sonic hedgehog (Shh) e Fibroblast growth factor (Fgf)
- fattori di crescita come neural growth factor (Ngf)

- costituenti della matrice extracellulare e molecole di adesione come laminina (Lam), tenascina (Tnc), neural cell adhesion molecule 1 (Ncam) e L1 cell adhesion molecule (L1cam)

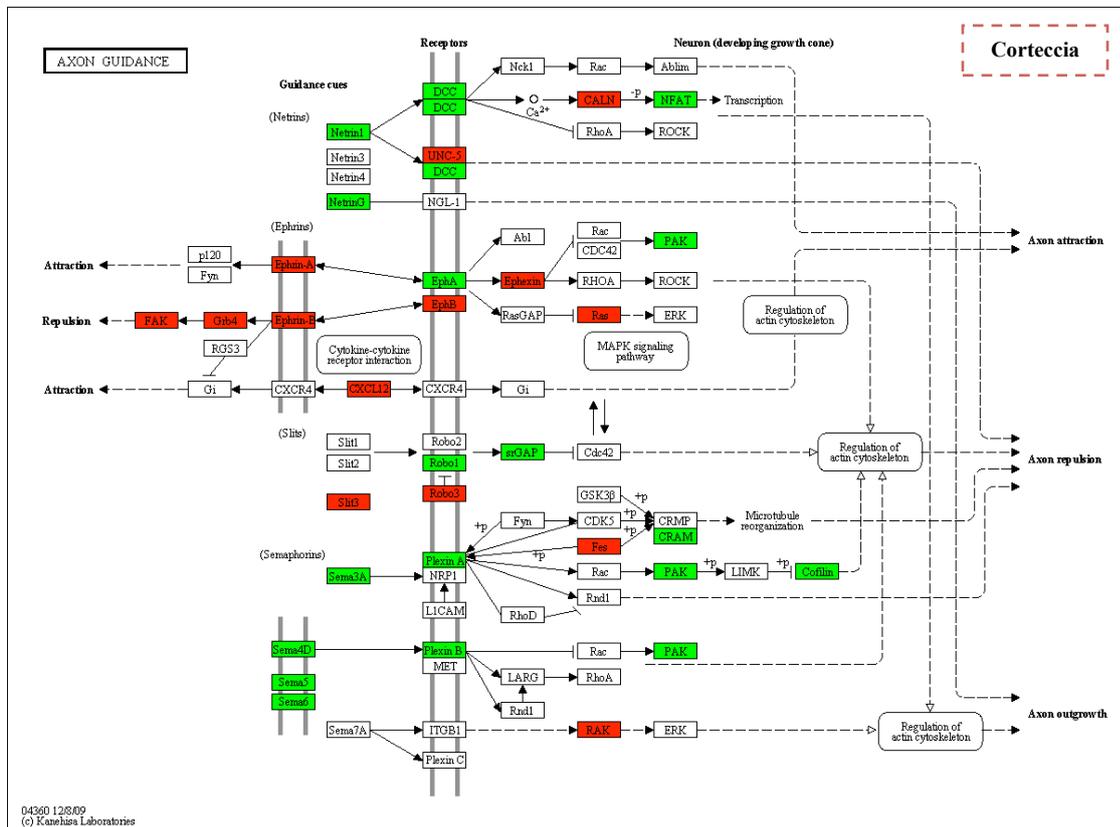


Figura 34: Rappresentazione schematica della via di segnale Axon guidance dove sono visualizzati i geni upregolati (rosso) e downregolati (verde) nella corteccia cerebrale animale IDS-KO rispetto al controllo sano [kegg. <http://www.genome.jp/kegg/>].

Anche l'espressione di alcune di queste molecole è alterata. Tra i morfogeni dello sviluppo risultano upregolati in entrambe le aree Bmp3, Wnt2, Wnt4, Wnt9a, Wnt10a, e downregolati in entrambe le aree Wnt3, Wnt9b, Shh, Fgf1 e Fgf9. Fgf22 risulta downregolato solo nella corteccia, Ngs è upregolato nel mesencefalo mentre la laminina (Lamc2) è upregolata in entrambe le macroaree.

L'alterazione dell'axon guidance nella Mucopolisaccaridosi di tipo II può indicare la presenza di compromissione neurologica anche a livello dello sviluppo del sistema

nervoso centrale e periferico. Inoltre, il rinnovo dei neuroni danneggiati potrebbe essere compromesso, a causa della possibile incapacità di generare neuroni funzionali in quanto deficitari nella proprietà di direzionare correttamente gli assoni.

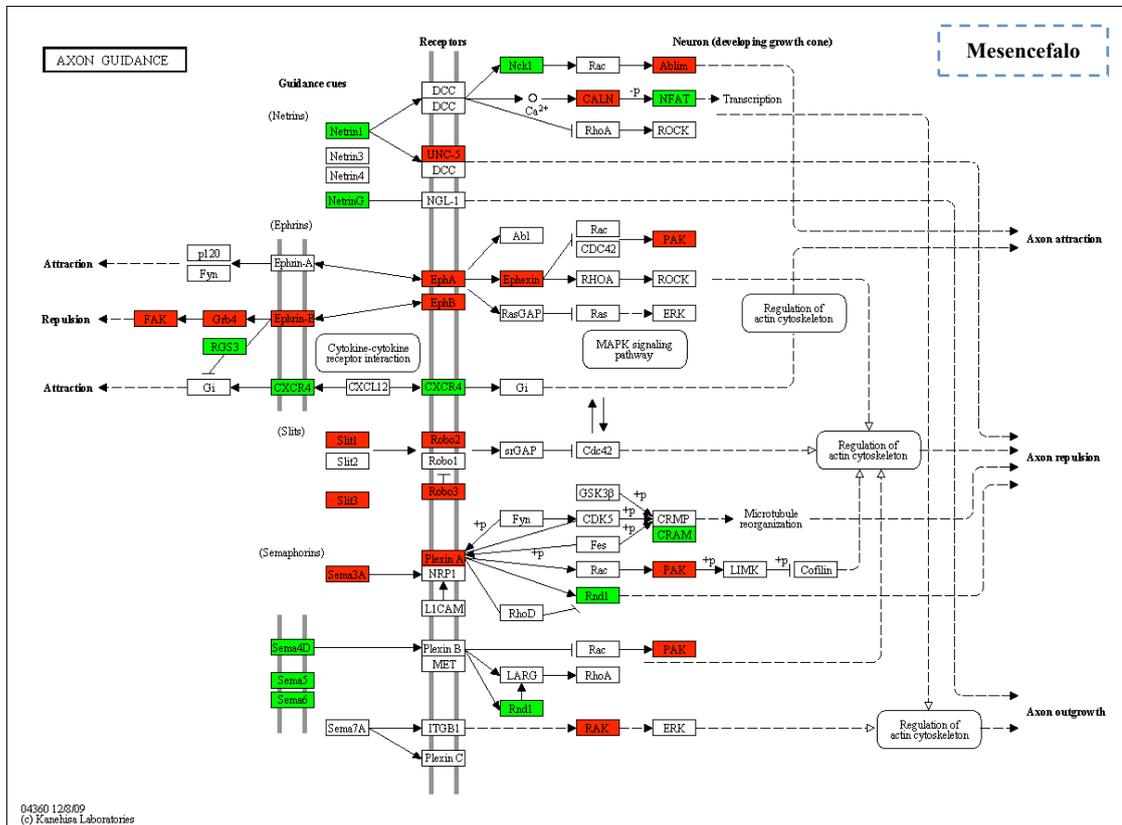


Figura 35: Rappresentazione schematica della via di segnale Axon guidance dove sono visualizzati i geni upregolati (rosso) e downregolati (verde) nel mesencefalo dell'animale IDS-KO rispetto al controllo sano [kegg. <http://www.genome.jp/kegg/>].

- Neuroinfiammazione

Una caratteristica comune alle malattie neurodegenerative è la neuroinfiammazione. Il principale mediatore cellulare dell'infiammazione cerebrale è la microglia residente che condivide molte caratteristiche con i relativi macrofagi di origine ematopoietica. Quando attivata, la microglia rilascia mediatori pro-infiammatori diffusibili che promuovono la disfunzione e la morte dei neuroni dopaminergici nel mesencefalo.

Recentemente, al recettore nucleare Nr4a2, ben conosciuto come un fattore determinante per la maturazione dei neuroni dopaminergici, sono state ascritte altre due proprietà. Primo, esso promuove la produzione e il rilascio del neuropeptide vasoattivo intestinale (Vip) che agisce sia nello stimolare la sopravvivenza dei neuroni dopaminergici sia inibendo l'infiammazione; secondo, Nr4a2 sopprime l'espressione e il rilascio di citochine pro-infiammatorie nelle cellule gliali. Vip a sua volta agisce come una molecola anti-infiammatoria inibendo il rilascio di sostanze pro-infiammatorie tra cui la chemochina RANTES (Ccl5) e attenuando l'attivazione della microglia. L'overespressione di Nr4a2 può agire sia attenuando l'espressione di molecole pro-infiammatorie dalla microglia e dagli astrociti, sia aumentando l'espressione di molecole anti-infiammatorie [Maguire-Zeiss and Federoff, 2010]. Nr4a2 è stato recentemente localizzato nella glia dove, dopo l'attivazione, questo fattore di trascrizione è traslocato nel nucleo incrementando l'espressione delle molecole anti-infiammatorie. Tuttavia, con la persistente esposizione a stress ossidativi o a misfold proteici, come quello dell' α -sinucleina, sembra esserci un apparente circolo vizioso dell'attivazione della glia che sovrasta la risposta anti-infiammatoria diretta di Nr4a2, comportando una vulnerabilità cellulare [Maguire-Zeiss and Federoff, 2010].

Il meccanismo molecolare e cellulare attraverso il quale gli accumuli lisosomiali di GAG portano alla neurodegenerazione non è ancora conosciuto. Studi nei modelli murini per la MPS I e la MPS III B hanno mostrato che i neuroni contengono vacuoli di stoccaggio e che anche altre cellule nel cervello, tra cui quelle della microglia, risultano affette. In quanto parte del meccanismo dell'immunità innata, quest'ultima può difendere il sistema nervoso centrale dai danni, ma può produrre anche sostanze tossiche. I danni ai neuroni causati dall'accumulo lisosomiale è pertanto aggravato dalla risposta immunitaria da parte delle cellule microgliali [Ohmi et al., 2003].

Così come osservato da Ohmi et al. [2003] per la MPS I e la MPS III B, nella corteccia dei topi IDS-KO rispetto ai fratelli WT c'è un aumento dei trascritti associati all'attività e alla linea macrofagica/monocitica; tra questi geni se ne possono nominare alcuni. La catepsina S (Ctss), così come il lisozima M (Lyz2) e il gene 1 specifico per i macrofagi (Mpeg1), sono preferenzialmente espressi proprio in queste cellule. Il citocromo b558 (Cyba) fa parte del complesso NADPH ossidasi, che è responsabile del burst ossidativo

dei fagociti, compresa la microglia. Tyrobp, una tirosin-chinasi binding protein, originariamente fu descritta tra le cellule natural killer ma più tardi fu trovata anche nei monociti e nella microglia. La catepsina Z (Ctsz) è stata trovata nei fagosomi macrofagici e il complemento C1q (C1qb) è drammaticamente incrementato nella microglia in caso di ischemia cerebrale; C1qb, ma anche il complemento C4 (C4a), sono prodotti dalla microglia ma si trovano anche in altre cellule nel cervello nell'Alzheimer [Ohmi et al., 2003].

Tuttavia, al contrario di quello che succede nella corteccia, è di rilevante importanza che nel mesencefalo tutti questi trascritti sono downregolati, indicando una diversa alterazione neuroinfiammatoria in quest'area del cervello.

Questi dati sono confermati anche dai valori di Nr4a2, Vip e Ccl5. Quest'ultimo, infatti, è upregolato nella corteccia e downregolato nel mesencefalo dei topi IDS-KO, mentre gli altri due sono upregolati in entrambe le aree, ma si può giustificare il valore nel mesencefalo ricordando quanto detto da Maguire-Zeiss and Federoff, cioè che con la persistente esposizione a stress ossidativi o a misfold proteici sembra esserci un apparente circolo vizioso dell'attivazione della glia che sovrasta la risposta anti-infiammatoria diretta di Nr4a2.

- Comparazione con altre malattie neurodegenerative

Molte caratteristiche delle malattie da accumulo lisosomale legate ad eventi secondari sono comuni ad altre malattie neurodegenerative non collegate agli accumuli lisosomiali, come ad esempio Alzheimer, Parkinson e Huntington. In tutte queste malattie, il principale problema è l'accumulo di una quantità eccessiva di molecole che comportano una cascata patogenetica e la progressiva perdita del normale funzionamento cellulare [Bahr and Bendiske, 2002].

La malattia di Alzheimer

La malattia di Alzheimer (AD) è la più comune forma di demenza e la sua prevalenza aumenta esponenzialmente con l'età. A livello neuropatologico, è caratterizzata dalla presenza di depositi extracellulari di proteina β -amiloide ($A\beta$) nelle placche neuritiche

e da cambiamenti neurofibrillari, tra cui depositi intracellulari di proteine tau iperfosforilate e agglomerate e neuriti distrofici [Braak et al., 1998].

In un recente lavoro Bossers et al. [2010] hanno realizzato un profilo di espressione genica genome-wide nella corteccia prefrontale di pazienti con Alzheimer riscontrando molti geni alterati che portano all'accumulo di placche di proteina A β . Confrontando i loro dati con quelli ottenuti dalla corteccia e dal mesencefalo, la maggior parte dei geni risultano differenzialmente espressi anche nel confronto IDS-KO contro WT indicando una comunanza con questa patologia. Tuttavia, la direzionalità di queste alterazioni è in parte diversa.

In un altro lavoro [Thathiah e De Strooper, 2011] viene valutato il coinvolgimento delle proteine G nella patologia dell'Alzheimer. Tali proteine sono coinvolte in molte vie di segnale e nell'AD sono implicate nella cascata del secondo messaggero derivata dai neurotrasmettitori.

Alcuni recettori glutammatergici (Grm, Metabotropic glutamate receptor), serotonergici (Htr, 5-hydroxytryptamine (serotonin) receptor), adrenergici (Adr, adrenergic receptor) e peptidergici (Oprd, opioid receptor), individuati nella corteccia cerebrale e nel mesencefalo dell'animale modello per la mucopolisaccaridosi di tipo II, risultano alterati come nei pazienti affetti da AD (Grm1 e Htr7 downregolati; Grm2-3, Adbr1, Oprd1 upregolati) confermando la possibile correlazione esistente tra le malattie neurodegenerative ad insorgenza tardiva e le LSDs ad insorgenza giovanile.

Nonostante questi primi dati positivi, risultano necessari ulteriori approfondimenti sull'argomento, poiché esiste un'incongruenza sull'andamento dell'espressione di alcuni geni tra diversi lavori presenti in letteratura.

La malattia di Parkinson

La malattia di Parkinson (PD) è la seconda più comune patologia neurodegenerativa.

La degenerazione dei neuroni dopaminergici della sostanza nera compatta porta a questa patologia, in quanto viene a mancare il rigido controllo della via nigro-striatale sulla facilitazione del movimento.

Attualmente si conoscono 10 geni associati all'insorgenza della PD. Tra questi α -sinucleina (Snca), Leucin-rich repeat kinase 2 (Lrrk2) e Nurr1 (Nr4a2) risultano upregolati in modo significativo sia nella corteccia sia nel mesencefalo dei topi IDS-KO rispetto ai fratelli WT.

In un recente lavoro Vahedi et al. [2010] hanno identificato 20 nuovi geni come possibili geni malattia. Tra questi Tspan6 (Tetraspanin 6), Arhgap24 (Rho GTPase activating protein 24), Klhl13 (Kelch-like 13) e Prph (Peripherin) risultano downregolati in modo rilevante in entrambe le aree analizzate nei topi IDS-KO, Nell2 (NEL-like 2) risulta upregolato, mentre altri 12 geni risultano significativamente variati ma in grado inferiore.

In un precedente studio di espressione genica, Gherbassi et al. [2007] hanno evidenziato in tutto 70 possibili geni malattia. In entrambe le aree dei topi IDS-KO questi geni risultano per la maggior parte downregolati.

In particolare tra questi geni è fortemente downregolato Slc18a2 (solute carrier family 18 member 2); questo gene codifica per il trasportatore vescicolare delle monoamine che agisce accumulando queste molecole citosoliche nelle vescicole sinaptiche, usando un gradiente protonico che è mantenuto attraverso la membrana sinaptica vescicolare. Il suo corretto funzionamento è essenziale per la corretta attività dei sistemi monoaminergici che sono stati implicati in numerosi disordini neuropsichiatrici umani [Peter et al., 1993] e in caso di ridotta espressione può essere correlato ad una più alta sensibilità agli stress ambientali.

Anche il gene Plp1 (Proteolipid protein 1) è risultato downregolato ed esso codifica per una proteina transmembrana proteolipidica che è la proteina mielinica predominante nel sistema nervoso centrale. Essa può giocare un ruolo importante nella compattazione, nella stabilizzazione e nel mantenimento delle guaine mieliniche, così come nello sviluppo degli oligodendrociti e nella sopravvivenza assonale. Mutazioni in questo gene causano la malattia X-linked di Pelizaeus-Merzbacher e la paraplegia spastica di tipo 2.

Il gene Icam5 (Intercellular adhesion molecule 5, telencephalin) è risultato invece upregolato; la proteina codificata è un membro della famiglia delle molecole di adesione intracellulari (ICAM). Questa proteina è espressa sulla superficie dei neuroni

telencefalici e mostra due tipi di attività di adesione, un legame omofilico tra i neuroni e uno eterofilico tra neuroni e leucociti. Essa può essere un componente critico nelle interazioni cellulari neuroni-microglia nel corso del normale sviluppo o nel caso di malattie neurodegenerative.

Il confronto effettuato con i geni normalmente correlati al morbo di Parkinson's e con i geni rilevati in questi ultimi studi consente di confermare l'associazione fatta tra le malattie lisosomiali e il morbo di Parkinson in lavori precedenti [Bellettato e Scarpa, 2010] e di evidenziare alterazioni a diversi livelli sia nella corteccia sia nel mesencefalo.

La malattia di Huntington

La Malattia di Huntington (HD) è una patologia ereditaria degenerativa caratterizzata dalla presenza di un disturbo del movimento progressivamente invalidante, disturbi del comportamento talvolta molto gravi e decadimento delle funzioni intellettive fino alla demenza. Un'espansione della poliglutammina nella regione amino-terminale della proteina huntingtina (Htt) è la causa genetica dell'HD. La proteina Htt interagisce con un'ampia varietà di proteine e mutazioni nel gene che la codifica causano alterazioni in diverse vie di neurotrasmissione, inclusi i sistemi dopaminergici, glutaminergici e cannabinoidi [Ribeiro et al., 2011].

L'espansione in poliglutammina nella proteina Htt comporta perturbazioni del signaling cellulare alterando l'omeostasi neuronale. È possibile che molte di queste alterazioni compaiano nella fase pre-sintomatica della malattia e pertanto possono essere cruciali per determinarne la progressione.

Mutazioni in Htt portano ad un decremento nella funzionalità di BDNF (brain-derived neurotrophic factor) e possono contribuire alla perdita di cellule neuronali nello striato e nella corteccia cerebrale e ad alterazioni motorie. Le modificazioni del sistema glutaminergico coinvolgono sia recettori ionotropici per il glutammato sia quelli metabotropici. I recettori NMDA (N-methyl-D-aspartate receptor) sono sensibilizzati da mutazioni in Htt e il signaling del gruppo I degli mGluRs (metabotropic glutamate receptors), che fanno parte della famiglia dei recettori accoppiati a proteine G, è significativamente alterato in modelli murini per l'HD. Il sistema glutaminergico è strettamente connesso alla regolamentazione della morte delle cellule neuronali;

tuttavia, l'attivazione dei recettori del glutammato può essere sia eccitotossica sia di protezione a seconda del contesto di attivazione. L'eccitotossicità è caratterizzata da un eccessivo rilascio sinaptico di calcio, che porta alla over-stimolazione dei recettori per il glutammato e un aumento del Ca^{2+} , compromettendo la funzionalità mitocondriale e portando alla morte cellulare [Ribeiro et al., 2011].

Mutazioni in Htt possono comportare, inoltre, variazioni delle funzioni correlate all'endocitosi clatrina-mediata, al processo di trasporto neuronale, al signaling post-sinaptico, all'apoptosi e alla sopravvivenza cellulare. Mutazioni possono, inoltre, causare un incremento nella concentrazione intracellulare di Ca^{2+} destabilizzando la regolazione del Ca^{2+} mitocondriale e sensibilizzandone il rilascio dagli accumuli intracellulare [Ribeiro et al., 2011].

Tutte queste alterazioni possono essere riassunte nel pathway dell'attivazione del signalling cellulare da parte del gruppo I degli mGluRs (**Figura 36**). Quasi tutti i geni presenti in questo pathway sono significativamente upregolati nella corteccia e nel mesencefalo dei topi IDS-KO rispetto ai fratelli WT, indicando un'alterazione comune di questo signalling con la malattia di Huntington che porta in entrambi i casi ad una modificazione della concentrazione intracellulare di Ca^{2+} con tutte le conseguenze neuronali che ne derivano.

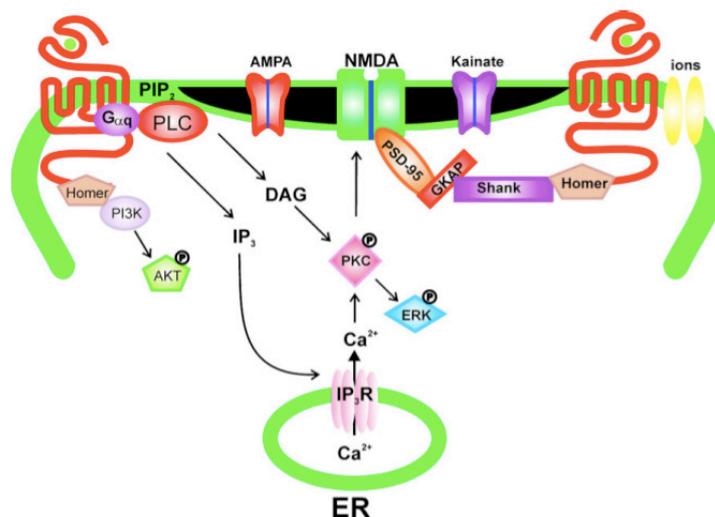


Figura 36: Attivazione del signalling cellulare da parte del gruppo I degli mGluRs [Ribeiro et al., 2011].

Ormone della crescita

La somatotropina (STH, nota anche come GH dal nome inglese *growth hormone*, cioè ormone della crescita) è un ormone peptidico composto da 191 amminoacidi, sintetizzato, accumulato e secreto dalle cellule della adenoipofisi in modo pulsatile e secondo un ritmo circadiano, che stimola lo sviluppo dell'organismo umano e di molti vertebrati. Durante l'infanzia, l'iposecrezione di GH determina nanismo ipofisario, mentre l'ipersecrezione causa gigantismo ipofisario. Come le altre tropine ipofisarie, il GH non è un ormone anabolizzante e non causa direttamente alcuna crescita, però stimola la sintesi epatica di un fattore di crescita insulino-simile (più conosciuto come IGF-1 o come somatomedina C) che è il vero ormone anabolicamente attivo e che è il responsabile degli effetti sulla crescita attribuiti al GH.

Il difetto di produzione di GH determina una bassa statura e un rallentamento della maturazione ossea, oltre a possibili alterazioni metaboliche quali l'ipoglicemia, la riduzione dello sviluppo muscolare e l'alterazione del metabolismo lipidico [Polgreen e Miller, 2010].

Un grave difetto nell'ormone della crescita determina anche un'insufficiente funzionalità cardiaca con riduzione dello spessore del muscolo cardiaco e conseguente riduzione della forza di contrazione del cuore.

Difetti nel signaling del GH possono portare a nanismo ed è stato supposto che la diminuzione dei livelli di GH con l'età giochi un ruolo importante nel ridurre la funzionalità di alcuni sistemi fisiologici. Il GH agisce nelle cellule attraverso growth hormone receptor (Ghr) che è un membro della famiglia dei recettori per le citochine. Il GH causa la dimerizzazione del recettore, attiva la protein-chinasi Jak2 (Janus kinase 2), la quale media molte delle risposte a valle attraverso la fosforilazione dei fattori di trascrizione Stat (Signal transducer and activator of transcription), delle Map-chinasi (Mapk), di altre cascate chinasiche e di molecole coinvolte nel metabolismo come Irs1 (Insulin receptor substrate 1). Fattori simili a Socs (Suppressor of cytokine signaling) e Shp1 sembrano giocare un ruolo nella downregolazione del segnale attraverso il GH e le citochine (**Figura 37**).

La maggior parte di questi geni, ed in particolare il Ghrh (growth hormone releasing hormone), risultano downregolati nella corteccia e soprattutto nel mesencefalo dei

topi IDS-KO rispetto ai controlli WT indicando una complessiva downregolazione di questo pathway. Ciò è in accordo con quanto riscontrato in varie mucopolisaccaridosi dove è stata rilevata un riduzione dei livelli di GH, IGF1 e TH [Polgreen e Miller, 2010].

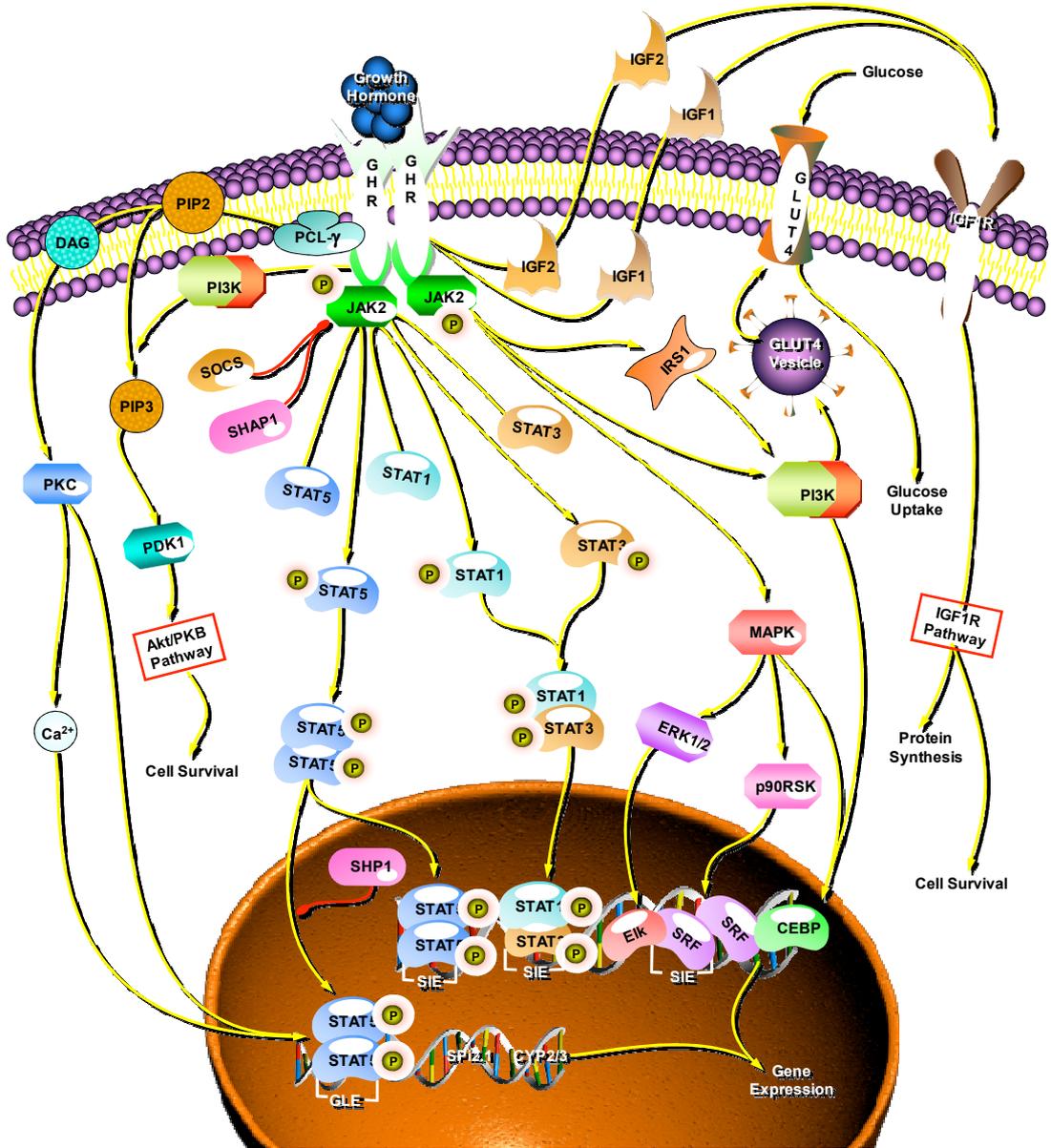


Figura 37: Growth Hormone Signaling Pathway [<http://www.sabiosciences.com>]

© 2009 Sabioscience, all rights reserved.

CONCLUSIONI

Come già sottolineato più volte nel corso della stesura di questo lavoro, la comprensione della componente neurologica delle malattie da accumulo lisosomiale, è stata e rimane ancora un obiettivo molto importante della comunità medico-scientifica del settore. Infatti, anche se per le più diffuse tra queste patologie sono oggi disponibili protocolli terapeutici per il trattamento della sintomatologia somatica, nessuna terapia è stata finora in grado di approssicare il deficit cognitivo e comportamentale che caratterizza buona parte dei pazienti. Inoltre, spesso le malattie da accumulo lisosomiale sono prese di esempio per la valutazione di nuove strategie terapeutiche di “sostituzione di prodotti genici difettivi” per le malattie ereditarie [Dietz, 2010]. È facile intuire come la comprensione della componente neurologica delle LSD potrebbe svelare nuovi target terapeutici specifici per questo gruppo di malattie, ma anche essere di ispirazione e supporto per la comprensione della patogenesi e lo sviluppo di nuove terapie per altre patologie a coinvolgimento neurologico.

Questa tesi è nata dall’opportunità di effettuare una valutazione del comparto neurologico nell’animale modello per la sindrome di Hunter, animale già ampiamente studiato nel laboratorio in cui questo progetto è stato svolto. Inoltre, questo progetto è stato reso possibile dal supporto di alcuni ricercatori del CRIBI (Centro di Ricerca Interdipartimentale per le Biotecnologie Innovative) dell’Università di Padova, che hanno fornito la strumentazione e le competenze bioinformatiche necessarie a realizzarlo.

La tecnologia di sequenziamento mediante SOLiD è l'ultima tra le prime tre tecnologie sviluppate nell'era dei sequenziatori di seconda generazione e attualmente, per questo, la meno utilizzata per il sequenziamento dell'RNA.

Il lavoro presentato in questa tesi è perciò nuovo nel settore delle LSDs e in generale nella storia del sequenziamento dell’RNA mediante tecnologia SOLiD per studi di patofisiologia su tessuto di mammifero. A tutt'oggi sono stati riportati complessivamente una ventina di lavori che impiegano il SOLiD per l’RNA-seq, di cui il 90% rilasciati nel 2010 [www.appliedbiosystem.com].

Fino a poco tempo fa la valutazione dell'RNA attraverso tecniche cosiddette "high throughput" era realizzato quasi esclusivamente con la tecnologia del microarray e solo recentemente è stato introdotto il sequenziamento del trascrittoma.

In generale, le tecniche di sequenziamento dell'RNA hanno portato ad una rivoluzione nell'analisi di questa molecola. Come già accennato nella parte introduttiva di questa tesi, il vantaggio principale, rispetto alla tecnica del microarray, consiste nella possibilità di non limitare la propria indagine a molecole note, come è proprio di una tecnologia che lavora mediante sonde, ma di amplificare e registrare tutte le specie molecolari di RNA presenti in un dato momento nel campione fornendo un dato quantitativo e affidabile.

Due sono i limiti principali della tecnica, i costi ancora piuttosto elevati e l'analisi dei dati ottenuti, che richiede ampie competenze bioinformatiche e capacità di gestione dei software.

Un altro limite che si rileva nel corso dell'analisi è il fatto che mentre la tecnologia ha fatto grandi passi avanti, i supporti bioinformatici non sono stati in grado di rispondere adeguatamente alle richieste di questo mercato. Inoltre i componenti di molti pathways biochimici e le effettive relazioni molecolari sono ancora in parte sconosciuti e tali valutazioni vengono effettuate manualmente con l'ausilio di pochi web-site funzionali.

Lo studio qui riportato è di tipo puramente osservazionale e ha proposto il confronto di 2 aree cerebrali provenienti dal modello murino per la MPS II e da un corrispondente campione di controllo proveniente da un pool di 7 cervelli derivati da fratelli sani.

Seppure l'analisi presenti il limite, attualmente non superabile, d'essere stata eseguita in un animale modello, ha permesso di fotografare una situazione priva di interferenze esterne donando la possibilità di chiarire parte della districata matassa di eventi che interessa il coinvolgimento neurologico nella MPS II.

Lo studio di trascrittomica è stato preceduto da una breve valutazione comportamentale degli animali allo scopo di ottenere qualche evidenza del coinvolgimento neurologico nel modello, che risulta scarsamente caratterizzato da questo punto di vista. Uno studio accurato sul comportamento richiederebbe altri test,

effettuati su un gruppo più ampio di animali; tuttavia, i risultati qui ottenuti hanno evidenziato un'iperattività e uno stato di ansia che differenziano, in modo significativo, l'animale malato dal sano. Ciò è stato poi confortato dall'analisi del trascrittoma che ha rilevato un numero elevato di geni e di pathways con differenze statisticamente significative di espressione tra i 2 pool di campioni.

L'analisi mediante RNA-Seq ha messo in luce nell'animale modello per la MPS II confrontato con l'animale sano, geni differenzialmente espressi in modo analogo nelle 2 aree cerebrali valutate, corteccia e mesencefalo, e geni differenzialmente espressi in una sola delle 2 aree valutate. I primi sono più probabilmente correlati allo stato di accumulo delle cellule, alterazioni che perciò potrebbero essere osservate anche in altre aree cerebrali dell'animale, i secondi hanno più probabilmente messo in luce i geni con espressione differenziale che, invece, rappresentano pathways specifici per le funzioni svolte da quell'area cerebrale.

Molto interessante risulta, a questo stadio dell'analisi, l'evidenza dell'alterazione dell'omeostasi del calcio, già riscontrata nelle LSDs, che vede coinvolte le proteine G non solo nell'aumento del calcio citosolico ma anche nella variazione della trasmissione del segnale mediato da neurotrasmettitori. A questo si collega il deficit dell'autofagia, l'elevato stress ossidativo e il difetto del citoscheletro nel trasporto vescicolare con dannose implicazioni nel rilascio dei neuropeptidi nello spazio intersinaptico e nella trasmissione dell'impulso nervoso, nonché il coinvolgimento di recettori ormonali riguardanti l'ormone della crescita e tiroideo già riscontrati diminuiti in altre mucopolisaccaridosi.

Questo studio potrebbe richiedere ulteriori approfondimenti, indispensabili per delucidare dubbi derivati dalla scarsa conoscenza delle interazioni tra le molecole di alcuni pathways. La grossa mole di dati ottenuti pone e porrà nuove domande sulla patofisiologia della mucopolisaccaridosi di tipo II, ma soprattutto metterà in evidenza il coinvolgimento di pathways metabolici trascurati fino ad ora, nella patogenesi di questa malattia, che potrà essere presa presa anche come esempio per le altre di più generale malattie da accumulo lisosomiale. E' abbastanza probabile, infatti, che una volta formatosi l'accumulo endo-lisosomiale, che avrà per ogni patologia cause

diverse, le conseguenze per il comparto endo- ed eso-cellulare possano seguire percorsi sovrapponibili, determinati dallo stato di accumulo patologico.

Inoltre, il presente studio mette in evidenza relazioni importanti tra queste patologie e altre malattie neurologiche, quali la malattia di Alzheimer o di Parkinson o di Huntington con cui potrebbero condividere meccanismi di neurogenerazione.

In particolare, per quanto riguarda la sindrome di Hunter, è ipotizzabile che l'alterazione dell'axon-giudance, qui rilevata, possa compromettere il potenziale rinnovo dei neuroni danneggiati, nelle aree in cui ciò sia possibile e, nel contempo, possa impedire la formazione di nuove interconnessioni tra i neuroni esistenti, limitando la plasticità cerebrale.

Il coinvolgimento dell'autofagia nella patogenesi delle LSD è confermato dal presente studio sulla MPS II, nel quale sono state osservate alterazioni del processo in entrambe le aree analizzate. Il malfunzionamento dell'autofagia è stato collegato a un ampio range di patologie umane, compreso il cancro, la neurodegenerazione e l'infezione patogena.

Infine, questo studio conferma, per la corteccia cerebrale, un processo neuroinfiammatorio in atto, già precedentemente descritto per i modelli murini di MPS I e MPS III B (Ohmi et al, 2003) analizzati mediante microarray, mentre non ha rilevato lo stesso tipo di alterazione per il mesencefalo.

L'enorme quantità di dati forniti da questo tipo di studi rende praticamente impossibile un'analisi completa di tutte le specie evidenziate e delle loro correlazioni nei pathways. Questo, tuttavia, rappresenta anche il suo valore aggiunto in quanto l'analisi potrebbe suggerire informazioni e spunti di studio utili non solo nell'ambito delle malattie da accumulo lisosomiale ma anche in altri settori della neuropatologia.

APPENDICE

In appendice sono riportati i geni più rappresentati nei termini della GO (altri geni) e più caratteristici nei pathways valutati. In grassetto sono identificati i geni che soddisfano il criterio di espressione differenziale ($|\log(\text{KO}/\text{WT})| \geq 0.7$; $p\text{-value} \leq 0.05$), in corsivo i geni statisticamente non significativi ($p\text{-value} > 0.05$).

GENE SYMBOL		ALTRI GENI			
Mus musculus	Homo sapiens	Cx_KO vs Cx_WT_log2Norm	Cx_KO vs Cx_WT_p-value	M_KO vs M_WT_log2Norm	M_KO vs M_WT_p-value
Bdnf	BDNF	0.776	4.20E-14	0.624	6.91E-10
Cacng2	CACNG2	<i>0.046</i>	<i>0.446319617</i>	0.597	9.66E-15
Cacng3	CACNG3	1.628	0	1.127	0
Cacng4	CACNG4	<i>0.099</i>	<i>0.336987374</i>	-0.586	9.45E-07
Cacng5	CACNG5	-1.358	0	-1.337	0
Cacng7	CACNG7	0.109	0.000898266	0.585	0
Cacng8	CACNG8	0.528	6.28E-07	2.379	0
Camk2a	CAMK2A	2.262	0	2.378	0
Camk2b	CAMK2B	0.787	0	0.876	0
Camk2d	CAMK2D	-1.037	0	-0.655	0
Chrm1	CHRM1	2.599	0	2.947	0
Chrm2	CHRM2	-1.795	0	-1.443	0
Chrm3	CHRM3	0.770	0	1.019	0
Chrm4	CHRM4	1.831	0	2.009	0
Chrm5	CHRM5	-2.977	0	-1.261	1.77E-05
Ghr	GHR	-0.781	1.44E-15	-0.832	0
Hcrt	HCRT	-7.693	0	-7.069	0
Pomc	POMC	-2.047	2.31E-11	-1.115	0.000509553
Rgs10	RGS10	-0.472	1.67E-08	-1.060	0
Rgs11	RGS11	<i>0.061</i>	<i>0.429267404</i>	1.291	0
Rgs12	RGS12	0.954	0	1.330	0
Rgs14	RGS14	2.304	0	2.375	0
Rgs16	RGS16	-2.398	0	-1.577	0
Rgs17	RGS17	-1.359	0	-1.137	0
Rgs19	RGS19	<i>0.039</i>	<i>0.679523834</i>	-0.538	5.55E-08
Rgs20	RGS20	0.844	0	0.102	<i>0.249491668</i>
Rgs3	RGS3	-0.426	5.17E-05	-0.711	9.52E-11
Rgs4	RGS4	0.752	0	0.559	0
Rgs8	RGS8	-0.659	0	-0.466	0
Rgs9	RGS9	0.759	0	1.098	0
Slc18a2	SLC18A2	-4.223	0	-3.381	0
Stx12	STX12	-0.539	0	-0.334	0
Stx17	STX17	-0.708	0	-0.549	2.78E-15
Stx1a	STX1A	2.429	0	2.742	0
Stx3	STX3	-0.569	0	0.007	<i>0.905983847</i>
Stx6	STX6	0.685	0	0.610	0
Stx7	STX7	-0.384	0	-0.585	0
Syt1	SYT1	0.189	0	0.518	0
Syt10	SYT10	-0.467	0.003840182	-0.548	0.00035889
Syt12	SYT12	0.693	0	0.612	2.22E-16
Syt16	SYT16	1.092	0	1.127	0
Syt17	SYT17	0.723	0	1.048	0
Syt12	SYT2	-1.728	0	-1.328	0
Syt4	SYT4	-0.979	0	-0.189	4.46E-13
Syt5	SYT5	0.567	0	1.206	0
Syt7	SYT7	0.355	0	0.778	0
Syt9	SYT9	-2.305	0	-1.820	0
Th	TH	-2.846	0	-2.771	0
Trh	TRH	-1.687	0	-2.671	0
Vamp1	VAMP1	-2.010	0	-1.595	0
Vamp2	VAMP2	0.458	0	0.648	0
Vamp3	VAMP3	-0.712	0	-0.550	0
Vamp4	VAMP4	-0.708	0	-0.546	0
Vamp7	VAMP7	-0.821	0	-1.537	0

EPARAN SOLFATO BINDING PROTEINS					
GENE SYMBOL		Cx_KO vs Cx_WT_log2Norm	Cx_KO vs Cx_WT_p-value	M_KO vs M_WT_log2Norm	M_KO vs M_WT_p-value
Mus musculus	Homo sapiens				
Proteoglicani con eparan solfato					
Gpc1	GPC1	-0.012	0.729883981	0.121	0.005734145
Gpc3	GPC3	-1.835	0	-1.318	7.43849E-15
Gpc4	GPC4	1.691	7.03881E-14	0.290	0.205371671
Gpc5	GPC5	-0.370	3.3266E-10	-0.437	2.32037E-13
Gpc6	GPC6	-0.122	0.265594675	0.372	0.001319101
Sdc2	SDC2	-0.970	0	-0.806	0
Sdc3	SDC3	-0.291	0	-0.174	5.41234E-13
Sdc4	SDC4	0.009	0.861535083	-0.673	0
Sdcbp	SDCBP	0.131	0.000958313	-0.957	0
Spock2	SPOCK2	-0.611	0	0.103	4.18198E-11
Spock3	SPOCK3	-1.004	0	-1.030	0
Matrici extracellulari					
Col11a1	COL11A1	-1.781	0	-0.426	0.000174204
Col11a2	COL11A2	-0.635	4.79837E-10	0.731	2.46676E-10
Col12a1	COL12A1	0.535	2.01643E-08	0.856	1.11022E-16
Col15a1	COL15A1	-1.249	0	-0.678	0.000147656
Col16a1	COL16A1	-0.890	6.99441E-15	0.290	0.021448554
Col18a1	COL18A1	-1.180	3.33067E-16	-0.066	0.71295187
Col19a1	COL19A1	1.075	0	1.523	0
Col1a2	COL1A2	-0.180	0.086618182	0.604	3.89272E-07
Col23a1	COL23A1	0.189	0.084264036	0.653	4.90819E-10
Col25a1	COL25A1	-1.495	0	-0.926	1.70752E-13
Col4a1	COL4A1	0.066	0.243716877	0.879	0
Col4a2	COL4A2	-0.148	0.001623152	0.779	0
Col4a3bp	COL4A3BP	-1.069	0	-0.806	0
Col5a1	COL5A1	1.796	0	3.195	0
Col6a1	COL6A1	1.272	0	1.780	0
Col6a2	COL6A2	0.985	0	2.006	0
Col8a1	COL8A1	0.105	0.553157071	0.269	0.113178308
Col8a2	COL8A2	0.263	0.135990432	0.515	0.010478515
Col9a2	COL9A2	-0.830	5.22773E-08	-0.868	8.34317E-08
Col9a3	COL9A3	-0.254	0.025848703	0.524	3.47428E-05
Fn1	FN1	0.969	0	0.287	2.46928E-07
Lama2	LAMA2	-0.115	0.273547939	-0.170	0.120297813
Lama5	LAMA5	0.232	0.011674083	0.601	1.19619E-07
Lamb1-1	LAMB1-1	0.351	0.000152003	0.677	1.2405E-11
Lamb2	LAMB2	0.172	0.001363963	0.627	0
Lamc1	LAMC1	0.143	0.035683954	0.325	2.77919E-05
Lamc2	LAMC2	2.106	0	2.618	0
Lamc3	LAMC3	-0.010	0.945448119	0.720	3.85573E-05
Thbs2	THBS2	-0.381	0.001317371	-0.982	2.22045E-16
Thbs3	THBS3	-0.354	0.002617941	1.425	0
Thbs4	THBS4	2.522	0	2.582	0
Tnr	TNR	0.174	8.65161E-05	0.599	0
Fattori di crescita					
Fgf1	FGF1	-1.112	0	-1.219	0
Fgf10	FGF10	-0.635	1.7643E-05	0.036	0.810713221
Fgf11	FGF11	-0.331	0.028930691	0.419	0.025947178
Fgf12	FGF12	-0.771	0	-0.453	0
Fgf13	FGF13	0.502	0	0.750	0
Fgf14	FGF14	-0.195	0.000639281	0.079	0.185938315
Fgf18	FGF18	-0.661	0.106814531	-0.641	0.106208828
Fgf2	FGF2	-0.118	0.804169022	-0.854	0.149167761
Fgf22	FGF22	-1.533	3.33047E-05	0.242	0.56812082
Fgf9	FGF9	-0.738	2.63123E-14	-0.877	0
Pdgfa	PDGFA	-0.144	0.066084689	0.483	7.19482E-08
Pdgfb	PDGFB	1.025	0	1.450	0
Pdgfc	PDGFC	-0.075	0.570718223	-0.219	0.128676417
Pdgfd	PDGFD	-0.898	1.60874E-05	-0.847	1.16962E-05
Tgfb1	TGFBI	1.141	2.60902E-14	0.078	0.694410365
Tgfb11	TGFBI1	0.837	6.83453E-13	1.088	3.33067E-16
Tgfb2	TGFBI2	0.048	0.644410452	-0.094	0.384470938
Tgfb3	TGFBI3	0.171	0.06887977	-0.175	0.112919371
Tgfb1	TGFBI	0.593	0.000957619	-0.502	0.010637707

OMEOSTATI DEL CALCIO					
GENE SYMBOL		Cx_KO vs Cx_WT_log2Norm	Cx_KO vs Cx_WT_p-value	M_KO vs M_WT_log2Norm	M_KO vs M_WT_p-value
Mus musculus	Homo sapiens				
Chrn3	CHRN3	-2.226	0	-3.751	0
Chrn4	CHRN4	-2.650	0	-2.921	0
Plce1	PLCE1	-1.085	0	-1.602	0
Plch1	PLCH1	-2.084	0	-1.406	0
Plcb4	PLCB4	-1.638	0	-1.044	0
Plcl1	PLCL1	-0.687	0	-1.028	0
Chrn2	CHRN2	-0.669	0	-0.772	0
Grin1a	GRIN1A	-0.449	0	-0.679	0
Gria4	GRIA4	-0.853	0	-0.661	0
Mapk6	MAPK6	-0.666	0	-0.514	0
Capn7	CAPN7	-0.737	0	-0.463	0
Erk2	Mapk1	0.524	0	0.459	0
Mapk1	MAPK1	0.524	0	0.459	0
Orai2	Orai2	0.689	0	0.495	0
Mapk14	MAPK14	0.372	0	0.642	0
Mapk8ip1	MAPK8IP1	0.306	0	0.657	0
Bcr	BCR	0.595	0	0.685	0
Gria3	GRIA3	0.505	0	0.712	0
Hdac5	HDAC5	0.363	0	0.718	0
Mapk4	MAMPK4	0.496	0	0.759	0
Plcb1	PLCB1	0.833	0	0.829	0
Atp2b1	ATP2B1	0.730	0	0.851	0
Calm3	CALM3	0.549	0	0.871	0
Atp2b2	ATP2B2	0.670	0	0.891	0
Grin1	GRIN1	0.367	0	0.970	0
Mapk11	MAPK11	0.633	0	0.980	0
Caln1	CALN1	0.828	0	0.992	0
Camk4	CAMK4	1.229	0	1.060	0
Atp2b4	ATP2B4	0.729	0	1.127	0
Ryr2	RYR2	0.378	0	1.130	0
Plch2	PLCH2	0.665	0	1.225	0
Mef2d	MEF2D	0.854	0	1.345	0
Itpr1	ITPR1	1.516	0	1.466	0
Cacnb3	CACNB3	1.762	0	1.523	0
Grin2b	GRIN2B	1.271	0	1.524	0
Grik5	GRIK5	1.009	0	1.676	0
Grin2a	GRIN2A	1.340	0	1.771	0
Mef2c	MEF2C	2.025	0	1.822	0
Cacnb1	CACNB1	1.204	0	2.048	0
Ryr1	RYR1	1.343	0	2.114	0
Slc8a2	SLC8A2 (NCX)	1.917	0	2.115	0
Camk2a	CAMK2A	2.262	0	2.378	0
Hdac4	HDAC4	0.638	2.33E-13	1.272	0
Grin2c	GRIN2C	0.297	5.49E-09	1.284	0
Rap1b	RAP1B	-0.352	1.24E-08	-0.554	0
Crebbp	CREBBP (CBP)	0.258	3.68E-06	1.142	0
Hdac7	HDAC7	0.109	0.117877529	1.070	0
Atp2b3	ATP2B3	-0.006	0.872859092	0.520	0
Plcd4	PLCD4	-0.509	2.84E-07	-0.839	1.11E-16
Cacnb2	CACNB2	-0.537	0	-0.522	2.22E-16
Grik1	GRIK1	-1.274	0	-0.684	1.89E-15
Htr3a	HTR3A (5HT3)	1.869	0	1.777	3.44E-15
Grin3a	GRIN3A	-1.208	0	-0.639	1.28E-13
Plcd3	PLCD3	1.091	0	0.895	1.87E-12
Grin2d	GRIN2D	-1.548	0	-0.714	5.43E-12
Capn1	CAPN1	0.149	0.067494947	0.572	4.50E-09
Plcl2	PLCL2	0.925	0	0.315	8.97E-09
P2rx5	P2RX5	-2.343	0	-1.591	1.04E-07
Plcd1	PLCD1	0.423	0.000122307	0.602	1.02E-06
P2rx6	P2RX6	0.168	0.26277697	0.693	3.98E-05
Plcg2	PLCG2	0.759	6.90E-07	-0.512	0.002092288
Mapk12	MAPK12	-0.501	0.000294118	-0.426	0.005846413
Capn6	CAPN6	-0.121	0.603801525	-0.559	0.009056346
Ryr3	RYR3	-0.612	0	0.015	0.75390437

STRESS OSSIDATIVO					
GENE SYMBOL		Cx_KO vs Cx_WT_log2Norm	Cx_KO vs Cx_WT_p-value	M_KO vs M_WT_log2Norm	M_KO vs M_WT_p-value
Mus musculus	Homo sapiens				
Cat	CAT	-0.682	0	-0.091	0.060943739
Gpx3	GPX3	-2.031	0	-3.012	0
Gpx6	GPX6	3.126	5.57E-08	1.719	0.017114577
Gss	GSS (GSH)	-0.599	2.71E-11	-0.578	4.90E-09
Lpo	LPO	-2.318	0	-1.881	1.70E-12
Nos3	NOS3	0.891	9.99E-16	0.522	0.000114171
Ptgs2	PTGS2	2.230	0	0.999	0
Sod3	SOD3	0.603	3.76E-13	0.604	3.74E-09

AUTOFAGIA					
GENE SYMBOL		Cx_KO vs Cx_WT_log2Norm	Cx_KO vs Cx_WT_p-value	M_KO vs M_WT_log2Norm	M_KO vs M_WT_p-value
Mus musculus	Homo sapiens				
Atg10	ATG10	-0.752	9.15E-06	-0.857	3.91E-08
Atg16l1	ATG16L1	0.083	0.121559101	0.501	0
Atg2a	ATG2A	0.045	0.408065498	0.706	0
Atg3	ATG3	-0.805	0	-0.675	0
Atg4d	ATG4D	0.291	0.001682966	0.514	5.93E-06
Atg9a	ATG9A	0.168	1.20E-06	0.773	0
Bnip3	BNIP3	-0.267	4.40E-13	-0.756	0
Camkk2	CAMKK2	1.762	0	2.008	0
Dapk1	DAPK1	0.380	0	0.614	0
Dctn1	DCTN1	-0.197	0	0.956	0
Dctn6	DCTN6	-0.670	0	0.094	0.093214002
Dync11l1	DYNC11L1	-0.526	1.23E-14	-0.054	0.440535993
Dync2h1	DYNC2H1	-1.042	0	-0.503	0
Dync2l1	DYNC2L1	-0.190	0.151253789	-0.640	4.23E-07
Dynlrb2	DYNLRB2	-0.459	0.136391633	-1.557	2.30E-06
Dynt3	DYNT3	-0.434	0	-0.745	0
E2f6	E2F6	-0.375	2.13E-07	-0.813	0
Eif2ak2	EIF2AK2	-0.480	0.000378758	-0.821	2.93E-10
Hif1a	HIF1A	-0.775	0	-0.912	0
Hif3a	HIF3A	-0.678	1.84E-08	0.007	0.955107176
Prkcd	PRKCD	-2.819	0	-2.431	0
Snapi	SNAPIN	-0.533	0	-0.731	0
Snca	SNCA	1.122	0	0.919	0
Tor1aip1	TOR1AIP1	-0.333	2.86E-08	-0.545	0
Tor1aip2	TOR1AIP2	-0.548	0	-0.580	0
Tor2a	TOR2A	0.648	6.47E-14	0.604	4.92E-09
Tsc2	TSC2	0.407	0	0.710	0

ORMONE DELLA CRESCITA					
GENE SYMBOL		Cx_KO vs Cx_WT_log2Norm	Cx_KO vs Cx_WT_p-value	M_KO vs M_WT_log2Norm	M_KO vs M_WT_p-value
Mus musculus	Homo sapiens				
Ghr	GHR	-0.781	1.44E-15	-0.832	0
Ghrh	GHRH	-5.000	7.73E-07	-4.538	1.50E-06
Igf1	IGF1	-0.548	8.64E-06	-0.564	2.77E-06
Irs1	IRS1	-0.375	7.16E-05	-0.614	4.69E-10
Mapk1	MAPK1	0.524	0	0.459	0
Mapk11	MAPK11	0.633	0	0.980	0
Mapk12	MAPK12	-0.501	0.000294118	-0.426	0.005846413
Mapk14	MAPK14	0.372	0	0.642	0
Mapk4	MAPK4	0.496	0	0.759	0
Mapk6	MAPK6	-0.666	0	-0.514	0
Mapk8ip1	MAPK8IP1	0.306	0	0.657	0
Ptpn6	PTPN6 (SHP1)	1.335	0	-0.569	0.000249827
Socs2	SOCS2	0.769	8.73E-07	0.606	0.000210137
Socs3	SOCS3	1.090	2.88E-09	0.342	0.098159337
Socs6	SOCS6	-0.842	2.34E-11	-0.889	9.22E-13
Socs7	SOCS7	0.816	0	0.881	0
Stat1	STAT1	-0.536	1.39E-13	-0.433	3.49E-09
Stat5b	STAT5B	-0.608	0	-0.383	2.43E-08
Stat6	STAT6	0.852	3.33E-16	0.290	0.007876343

AXON GUIDANCE					
GENE SYMBOL		Cx_KO vs Cx_WT_log2Norm	Cx_KO vs Cx_WT_p-value	M_KO vs M_WT_log2Norm	M_KO vs M_WT_p-value
Mus musculus	Homo sapiens				
Bmp1	BMP1	0.667	0	1.850	0
Bmp2	BMP2	-0.805	8.59176E-05	-0.290	0.172128199
Bmp3	BMP3	2.060	0	2.145	0
Bmp4	BMP4	0.342	0.148490003	0.854	0.001380392
Bmp6	BMP6	0.194	0.075172922	1.050	0
Bmp7	BMP7	0.801	4.40359E-12	0.303	0.029012136
Dcc	DCC	-0.736	0	-0.225	0.004447994
Efna1	EFNA1	0.771	0.003155174	0.302	0.246785389
Efna3	EFNA3	0.634	1.88E-13	0.613	6.13E-09
Efnb1	EFNB1	0.715	9.18E-11	1.070	2.89E-15
Efnb2	EFNB2	1.139	0	0.716	0
Epha10	EPHA10	-0.170	0.275888681	1.126	4.46E-08
Epha3	EPHA3	0.142	0.440211937	0.583	0.000576332
Epha4	EPHA4	0.618	0	0.647	0
Epha6	EPHA6	0.025	0.751778648	0.831	0
Epha7	EPHA7	0.481	0	0.689	0
Epha8	EPHA8	-1.279	0	-0.459	9.44E-05
Ephb3	EPHB3	1.057	0	1.251	0
Ephb4	EPHB4	0.648	3.32E-06	0.405	0.011978621
Ephb6	EPHB6	1.168	0	1.350	0
Fgf1	FGF1	-1.112	0	-1.219	0
Fgf10	FGF10	-0.635	1.7643E-05	0.036	0.810713221
Fgf12	FGF12	-0.771	0	-0.453	0
Fgf13	FGF13	0.502	0	0.750	0
Fgf22	FGF22	-1.533	3.33047E-05	0.242	0.56812082
Fgf9	FGF9	-0.738	2.63123E-14	-0.877	0
L1cam	L1CAM	-0.585	0	-0.230	8.91287E-13
Ngf	NGF	0.541	0.043860793	0.880	0.0008065
Ntn1	NTN1	-0.854	0	-0.847	0
Ntn5	NTN5	4.039	0	5.000	0
Ntng1	NTNG1	-2.290	0	-2.121	0
Nxph1	NXPH1	-1.188	0	-0.962	0
Nxph3	NXPH3	0.971	0	1.211	0
Nxph4	NXPH4	-2.595	0	-1.711	0
Plxna2	PLXNA2	0.576	0	0.844	0
Plxna3	PLXNA3	-0.849	0	-0.023	0.730416636
Plxnb3	PLXNB3	-0.997	0	-0.362	2.68008E-08
Plxnd1	PLXND1	2.051	0	2.422	0
Robo1	ROBO1	-0.726	0	-0.564	0
Robo2	ROBO2	0.534	0	0.713	0
Robo3	ROBO3	1.432	0	3.162	0
Sema3a	SEMA3A	0.886	6.88338E-15	0.935	1.11022E-16
Sema3d	SEMA3D	-1.002	0	-0.335	0.002425689
Sema3f	SEMA3F	-1.209	0	-0.554	1.26262E-06
Sema4b	SEMA4B	0.036	0.689421565	0.520	9.37708E-07
Sema4f	SEMA4F	0.249	1.99847E-05	0.513	2.16938E-13
Sema4g	SEMA4G	-1.612	0	-0.756	0
Sema5a	SEMA5A	-1.280	0	-0.964	0
Sema6a	SEMA6A	-1.013	0	-0.720	2.22045E-16
Sema6b	SEMA6B	0.424	0	1.151	0
Sema6d	SEMA6D	-0.999	0	-0.751	0
Sema7a	SEMA7A	0.696	0	0.680	0
Shh	SHH	-1.585	3.10862E-15	-1.166	1.06623E-07
Slit1	SLIT1	0.147	0.008099868	0.812	0
Slit3	SLIT3	1.017	0	1.835	0
Srgap1	SRGAP1	-0.897	3.21965E-15	-0.533	4.35808E-05
Srgap3	SRGAP3	0.317	0	0.595	0
Unc5a	UNC5A	1.333	0	1.898	0
Unc5b	UNC5B	-0.538	0	-0.186	0.007084761
Unc5c	UNC5C	-0.785	0	-0.914	0
Unc5d	UNC5D	-0.503	1.14E-14	-0.395	1.43E-09
Wnt10a	WNT10A	2.389	0	3.121	0
Wnt2	WNT2	1.707	3.33067E-16	1.023	2.33011E-05
Wnt2b	WNT2B	-1.047	2.79176E-10	0.146	0.368392849
Wnt3	WNT3	-2.165	0	-2.474	0
Wnt4	WNT4	0.794	8.39768E-05	1.170	1.63295E-06
Wnt5a	WNT5A	-0.582	2.70912E-05	-0.320	0.034924642
Wnt7b	WNT7B	0.794	0	0.331	3.22582E-05
Wnt9a	WNT9A	0.998	0	1.242	0
Wnt9b	WNT9B	-2.705	0	-2.918	0

NEUROINFIAMMAZIONE					
GENE SYMBOL		Cx_KO vs Cx_WT_log2Norm	Cx_KO vs Cx_WT_p-value	M_KO vs M_WT_log2Norm	M_KO vs M_WT_p-value
Mus musculus	Homo sapiens				
C1qb	C1QB	1.265	0	-0.684	0
C4a	C4A	0.824	0	-0.865	0
Ccl5	CCL5	1.233	0.005503966	-2.226	1.59E-05
Ctss	CTSS	0.795	0	-1.448	0
Ctssz	CTSZ	0.615	0	-0.412	3.33E-16
Cyba	CYBA	1.101	9.13E-12	-0.319	0.093273849
Lyz2	LYZ2	1.931	0	-1.927	0
Mpeg1	MPEG1	1.069	0	-1.281	0
Nr4a2	NR4A2	1.122	0	2.125	0
Snca	SNCA	1.122	0	0.919	0
Tyrobp	TYROBP	1.358	0	-1.497	0
Vip	VIP	1.673	0	2.134	0

COMPARAZIONE CON ALTRE MALATTIE NEURODEGENERATIVE					
GENE SYMBOL		Cx_KO vs Cx_WT_log2Norm	Cx_KO vs Cx_WT_p-value	M_KO vs M_WT_log2Norm	M_KO vs M_WT_p-value
Mus musculus	Homo sapiens				
Adra1d	ADRA1D	3.111	0	2.705	0
Adra2a	ADRA2A	-0.732	0	-0.589	2.02E-09
Adra2c	ADRA2C	1.326	0	1.313	0
Adrb1	ADRB1	1.177	0	1.211	0
Adrb1	ADRB1	1.177	0	1.211	0
Adrb3	ADRB3	0.042	0.860907892	0.649	0.01482019
Adrbk1	ADRBK1	0.515	0	0.445	0
Arhgap24	ARHGAP24	-1.693	0	-2.023	0
BDNF	BDNF	0.776	4.20E-14	0.624	6.91E-10
Grin1	GRIN1 (NMDA)	0.367	0	0.970	0
Grin2a	GRIN2A	1.340	0	1.771	0
Grin2b	GRIN2B	1.271	0	1.524	0
Grin2c	GRIN2C	0.297	5.49E-09	1.284	0
Grin2d	GRIN2D	-1.548	0	-0.714	5.43E-12
Grin3a	GRIN3A	-1.208	0	-0.639	1.28E-13
Grm1	GRM1 (mGluRs)	-1.099	0	-0.244	0.000115987
Grm2	GRM2	0.801	0	0.901	0
Grm3	GRM3	0.563	0	0.750	0
Grm4	GRM4	-0.726	0	-0.187	0.044479291
Grm5	GRM5	0.044	0.144170017	0.512	0
Grm8	GRM8	-0.500	2.12E-06	-0.580	1.20E-08
Htr1d	HTR1D	0.830	0.00029463	0.077	0.732245169
Htr1f	HTR1F	0.755	0.000757063	0.708	0.002106957
Htr2a	HTR2A	0.652	4.58E-11	1.432	0
Htr2c	HTR2C	-1.943	0	-1.598	0
Htr3a	HTR3A	1.869	0	1.777	3.44E-15
Htr4	HTR4	-1.313	9.22E-13	-0.180	0.290142519
Htr5b	HTR5B	-0.790	0.001750864	-0.889	0.001183803
Htr6	HTR6	1.042	4.53E-06	1.824	3.60E-09
Htr7	HTR7	-1.907	0	-1.502	0
Icam5	ICAM5	2.200	0	2.738	0
Klhl13	KLHL13	-1.039	0	-1.292	0
Lrrk2	LRRK2	1.309	0	1.324	0
Nr4a2	NR4A2	1.122	0	2.125	0
Oprd	OPRD	1.615	0	1.620	0
Oprd1	ADRB1	1.615	0	1.620	0
Plp1	PLP1	-1.241	0	-1.149	0
Prph	PRPH	-2.122	1.11E-15	-1.823	1.01E-10
Slc18a2	SLC18A2	-4.223	0	-3.381	0
Snca	SNCA	1.122	0	0.919	0
Tspan6	TSPAN6	-1.532	0	-1.632	0

BIBLIOGRAFIA

Auclair D, Finnie J, White J, Nielsen T, Fuller M, Kakkis E, Cheng A, O'Neill CA, Hopwood JJ. 2010. Repeated intrathecal injections of recombinant human 4-sulphatase remove dural storage in mature mucopolysaccharidosis VI cats primed with a short-course tolerisation regimen. *Mol Genet Metab* 99(2):132-41.

Baehner F, Schmiedeskamp C, Krummenauer F, Miebach E, Bajbouj M, Whybra C, Kohlschutter A, Kampmann C, Beck M. 2005. Cumulative incidence rates of the mucopolysaccharidoses in Germany. *J Inherit Metab Dis* 28(6):1011-7.

Bahr BA and Bendiske J. 2002. The neuropathogenic contributions of lysosomal dysfunction. *J Neurochem* 83(3):481-9.

Ballabio A and Gieselmann V. 2009. Lysosomal disorders: From storage to cellular damage. *Biochim Biophys Acta* 1793(4):684-96.

Bashaw GJ and Klein R. 2010. Signaling from axon guidance receptors. *Cold Spring Harb Perspect Biol* 2(5):a001941.

Beck M. 2010. Therapy for lysosomal storage disorders. *IUBMB Life* 62(1):33-40.

Bellettato CM and Scarpa M. 2010. Pathophysiology of neuropathic lysosomal storage disorders. *J Inherit Metab Dis* 33(4):347-62.

Berchtold MW, Brinkmeier H, Muntener M. 2000. Calcium ion in skeletal muscle: Its crucial role for muscle function, plasticity, and disease. *Physiol Rev* 80(3):1215-65.

Bossers K, Wirz KT, Meerhoff GF, Essing AH, van Dongen JW, Houba P, Kruse CG, Verhaagen J, Swaab DF. 2010. Concerted changes in transcripts in the prefrontal cortex precede neuropathology in Alzheimer's disease. *Brain* 133(Pt 12):3699-723.

Braak H, de Vos RA, Jansen EN, Bratzke H, Braak E. 1998. Neuropathological hallmarks of Alzheimer's and Parkinson's diseases. *Prog Brain Res* 117:267-85.

Burrow TA and Leslie ND. 2008. Review of the use of idursulfase in the treatment of mucopolysaccharidosis II. *Biologics* 2(2):311-20.

Cai Q, Lu L, Tian JH, Zhu YB, Qiao H, Sheng ZH. 2010. Snapin-regulated late endosomal transport is critical for efficient autophagy-lysosomal function in neurons. *Neuron* 68(1):73-86.

Campagna D, Albiero A, Bilardi A, Caniato E, Forcato C, Manavski S, Vitulo N, Valle G. 2009. PASS: A program to align short sequences. *Bioinformatics* 25(7):967-8.

Cardone M, Polito VA, Pepe S, Mann L, D'Azzo A, Auricchio A, Ballabio A, Cosma MP. 2006. Correction of Hunter syndrome in the MPSII mouse model by AAV2/8-mediated gene delivery. *Hum Mol Genet* 15(7):1225-36.

Carter RJ, Lione LA, Humby T, Mangiarini L, Mahal A, Bates GP, Dunnett SB, Morton AJ. 1999. Characterization of progressive motor deficits in mice transgenic for the human Huntington's disease mutation. *J Neurosci* 19(8):3248-57.

Clarke LA. 2008. The mucopolysaccharidoses: A success of molecular medicine. *Expert Rev Mol Med* 10:e1.

Costa V, Angelini C, De Feis I, Ciccodicola A. 2010. Uncovering the complexity of transcriptomes with RNA-seq. *J Biomed Biotechnol* 2010:853916.

D'Anna R, Cannata ML, Marini H, Atteritano M, Cancellieri F, Corrado F, Triolo O, Rizzo P, Russo S, Gaudio A, et al. 2009. Effects of the phytoestrogen genistein on hot flushes, endometrium, and vaginal epithelium in postmenopausal women: A 2-year randomized, double-blind, placebo-controlled study. *Menopause* 16(2):301-6.

de Duve C. 2005. The lysosome turns fifty. *Med Sci (Paris)* 21(11 Suppl):12-5.

Denenberg VH. 1969. Open-field behavior in the rat: What does it mean? *Ann N Y Acad Sci* 159(3):852-9.

Dev KK, Hofele K, Barbieri S, Buchman VL, van der Putten H. 2003. Part II: Alpha-synuclein and its molecular pathophysiological role in neurodegenerative disease. *Neuropharmacology* 45(1):14-44.

Dickson PI, Hanson S, McEntee MF, Vite CH, Vogler CA, Mlikotic A, Chen AH, Ponder KP, Haskins ME, Tippin BL, et al. 2010. Early versus late treatment of spinal cord compression with long-term intrathecal enzyme replacement therapy in canine mucopolysaccharidosis type I. *Mol Genet Metab* 101(2-3):115-22.

Dietz HC. 2010. New therapeutic approaches to mendelian disorders. *N Engl J Med* 363(9):852-63.

Dionisi-Vici C, Rizzo C, Burlina AB, Caruso U, Sabetta G, Uziel G, Abeni D. 2002. Inborn errors of metabolism in the Italian pediatric population: A national retrospective survey. *J Pediatr* 140(3):321-7.

DiRosario J, Divers E, Wang C, Etter J, Charrier A, Jukkola P, Auer H, Best V, Newsom DL, McCarty DM, et al. 2009. Innate and adaptive immune activation in the brain of MPS IIIB mouse model. *J Neurosci Res* 87(4):978-90.

Dolmetsch R. 2003. Excitation-transcription coupling: Signaling by ion channels to the nucleus. *Sci STKE* 2003(166):PE4.

Dreyfuss JL, Regatieri CV, Jarrouge TR, Cavaleiro RP, Sampaio LO, Nader HB. 2009. Heparan sulfate proteoglycans: Structure, protein interactions and cell signaling. *An Acad Bras Cienc* 81(3):409-29.

Eto Y, Shen JS, Meng XL, Ohashi T. 2004. Treatment of lysosomal storage disorders: Cell therapy and gene therapy. *J Inherit Metab Dis* 27(3):411-5.

Friso A, Tomanin R, Salvalaio M, Scarpa M. 2010. Genistein reduces glycosaminoglycan levels in a mouse model of mucopolysaccharidosis type II. *Br J Pharmacol* 159(5):1082-91.

Friso A, Tomanin R, Alba S, Gasparotto N, Puicher EP, Fusco M, Hortelano G, Muenzer J, Marin O, Zacchello F, et al. 2005. Reduction of GAG storage in MPS II mouse model following implantation of encapsulated recombinant myoblasts. *J Gene Med* 7(11):1482-91.

Froissart R, Da Silva IM, Maire I. 2007. Mucopolysaccharidosis type II: An update on mutation spectrum. *Acta Paediatr Suppl* 96(455):71-7.

Futerman AH and van Meer G. 2004. The cell biology of lysosomal storage disorders. *Nat Rev Mol Cell Biol* 5(7):554-65.

Garcia AR, Pan J, Lamsa JC, Muenzer J. 2007. The characterization of a murine model of mucopolysaccharidosis II (Hunter syndrome). *J Inher Metab Dis* 30(6):924-34.

Gherbassi D, Bhatt L, Thuret S, Simon HH. 2007. Merging mouse transcriptome analyses with Parkinson's disease linkage studies. *DNA Res* 14(2):79-89.

Ginzburg L and Futerman AH. 2005. Defective calcium homeostasis in the cerebellum in a mouse model of Niemann-Pick A disease. *J Neurochem* 95(6):1619-28.

Gould TD, Dao DT, Kovacsics CE. 2009. The open field test. In: *Mood and Anxiety Related Phenotypes in Mice: characterization using behavioral tests*, Gould TD editor, The Humana Press Inc., Totowa, NJ, pp 1-20.

Gu Y, Zhu CF, Iwamoto H, Chen JS. 2005. Genistein inhibits invasive potential of human hepatocellular carcinoma by altering cell cycle, apoptosis, and angiogenesis. *World J Gastroenterol* 11(41):6512-7.

He C and Klionsky DJ. 2009. Regulation mechanisms and signaling pathways of autophagy. *Annu Rev Genet* 43:67-93.

Herati RS, Ma X, Tittiger M, Ohlemiller KK, Kovacs A, Ponder KP. 2008. Improved retroviral vector design results in sustained expression after adult gene therapy in mucopolysaccharidosis I mice. *J Gene Med* 10(9):972-82.

HERS HG. 1965. Inborn lysosomal diseases. *Gastroenterology* 48:625-33.

Jakobkiewicz-Banecka J, Piotrowska E, Narajczyk M, Baranska S, Wegrzyn G. 2009. Genistein-mediated inhibition of glycosaminoglycan synthesis, which corrects storage in cells of patients suffering from mucopolysaccharidoses, acts by influencing an epidermal growth factor-dependent pathway. *J Biomed Sci* 16:26.

Jung SC, Park ES, Choi EN, Kim CH, Kim SJ, Jin DK. 2010. Characterization of a novel mucopolysaccharidosis type II mouse model and recombinant AAV2/8 vector-mediated gene therapy. *Mol Cells* 30(1):13-8.

Kakkis E, McEntee M, Vogler C, Le S, Levy B, Belichenko P, Mobley W, Dickson P, Hanson S, Passage M. 2004. Intrathecal enzyme replacement therapy reduces lysosomal storage in the brain and meninges of the canine model of MPS I. *Mol Genet Metab* 83(1-2):163-74.

Karl T, Hoffmann T, Pabst R, von Horsten S. 2003. Behavioral effects of neuropeptide Y in F344 rat substrains with a reduced dipeptidyl-peptidase IV activity. *Pharmacol Biochem Behav* 75(4):869-79.

Maguire-Zeiss KA and Federoff HJ. 2010. Future directions for immune modulation in neurodegenerative disorders: Focus on Parkinson's disease. *J Neural Transm* 117(8):1019-25.

Manara R, Rampazzo A, Cananzi M, Salviati L, Mardari R, Drigo P, Tomanin R, Gasparotto N, Priante E, Scarpa M. 2010. Hunter syndrome in an 11-year old girl on enzyme replacement therapy with idursulfase: Brain magnetic resonance imaging features and evolution. *J Inherit Metab Dis* .

Martin R, Beck M, Eng C, Giugliani R, Harmatz P, Munoz V, Muenzer J. 2008. Recognition and diagnosis of mucopolysaccharidosis II (Hunter syndrome). *Pediatrics* 121(2):e377-86.

Meikle PJ, Hopwood JJ, Clague AE, Carey WF. 1999. Prevalence of lysosomal storage disorders. *JAMA* 281(3):249-54.

Metzker ML. 2010. Sequencing technologies - the next generation. *Nat Rev Genet* 11(1):31-46.

Morozova O, Hirst M, Marra MA. 2009. Applications of new sequencing technologies for transcriptome analysis. *Annu Rev Genomics Hum Genet* 10:135-51.

Muenzer J, Fu H. 1999. Targeted disruption of the mouse iduronate sulfatase gene. *Am J Genet.* 65:A427.

Muenzer J, Wraith JE, Beck M, Giugliani R, Harmatz P, Eng CM, Vellodi A, Martin R, Ramaswami U, Guzsavas-Calikoglu M, et al. 2006. A phase II/III clinical study of enzyme replacement therapy with idursulfase in mucopolysaccharidosis II (Hunter syndrome). *Genet Med* 8(8):465-73.

Muenzer J, Beck M, Eng CM, Escolar ML, Giugliani R, Guffon NH, Harmatz P, Kamin W, Kampmann C, Koseoglu ST, et al. 2009. Multidisciplinary management of Hunter syndrome. *Pediatrics* 124(6):e1228-39.

Munoz-Rojas MV, Horovitz DD, Jardim LB, Raymundo M, Llerena JC,Jr, de Magalhaes Tde S, Vieira TA, Costa R, Kakkis E, Giugliani R. 2010. Intrathecal administration of recombinant human N-acetylgalactosamine 4-sulfatase to a MPS VI patient with pachymeningitis cervicalis. *Mol Genet Metab* 99(4):346-50.

Munoz-Rojas MV, Vieira T, Costa R, Fagundes S, John A, Jardim LB, Vedolin LM, Raymundo M, Dickson PI, Kakkis E, et al. 2008. Intrathecal enzyme replacement therapy in a patient with mucopolysaccharidosis type I and symptomatic spinal cord compression. *Am J Med Genet A* 146A(19):2538-44.

Neufeld ES, Muenzer J. 2001. The mucopolysaccharidosis. In: Scriver CR, ed. *The Metabolic and Molecular Bases of Inherited Disease*, New York, NY: McGraw-Hill, Chapter 136.

Nguyen HN, Wang C, Perry DC. 2002. Depletion of intracellular calcium stores is toxic to SH-SY5Y neuronal cells. *Brain Res* 924(2):159-66.

Nowycky MC and Thomas AP. 2002. Intracellular calcium signaling. *J Cell Sci.* 115(Pt 19):3715-6.

Ohmi K, Greenberg DS, Rajavel KS, Ryazantsev S, Li HH, Neufeld EF. 2003. Activated microglia in cortex of mouse models of mucopolysaccharidoses I and IIIB. *Proc Natl Acad Sci U S A* 100(4):1902-7.

Peckham HE, McLaughlin SF, Ni JN, Rhodes MD, Malek JA, McKernan KJ, Blanchard AP. 2007. SOLiD sequencing and 2-base encoding. *ASHG Poster #2624*.

Pereira VG, Gazarini ML, Rodrigues LC, da Silva FH, Han SW, Martins AM, Tersariol IL, D'Almeida V. 2010. Evidence of lysosomal membrane permeabilization in mucopolysaccharidosis type I: Rupture of calcium and proton homeostasis. *J Cell Physiol* 223(2):335-42.

Peter D, Finn JP, Klisak I, Liu Y, Kojis T, Heinzmann C, Roghani A, Sparkes RS, Edwards RH. 1993. Chromosomal localization of the human vesicular amine transporter genes. *Genomics* 18(3):720-3.

Pinto R, Caseiro C, Lemos M, Lopes L, Fontes A, Ribeiro H, Pinto E, Silva E, Rocha S, Marcao A, et al. 2004. Prevalence of lysosomal storage diseases in Portugal. *Eur J Hum Genet* 12(2):87-92.

Piotrowska E, Jakobkiewicz-Banecka J, Baranska S, Tytki-Szymanska A, Czartoryska B, Wegrzyn A, Wegrzyn G. 2006. Genistein-mediated inhibition of glycosaminoglycan synthesis as a basis for gene expression-targeted isoflavone therapy for mucopolysaccharidoses. *Eur J Hum Genet* 14(7):846-52.

Polgreen LE and Miller BS. 2010. Growth patterns and the use of growth hormone in the mucopolysaccharidoses. *J Pediatr Rehabil Med* 3(1):25-38.

Polito VA and Cosma MP. 2009. IDS crossing of the blood-brain barrier corrects CNS defects in MPSII mice. *Am J Hum Genet* 85(2):296-301.

Polito VA, Abbondante S, Polishchuk RS, Nusco E, Salvia R, Cosma MP. 2010. Correction of CNS defects in the MPSII mouse model via systemic enzyme replacement therapy. *Hum Mol Genet* 19(24):4871-85.

Poorthuis BJ, Wevers RA, Kleijer WJ, Groener JE, de Jong JG, van Weely S, Niezen-Koning KE, van Diggelen OP. 1999. The frequency of lysosomal storage diseases in the Netherlands. *Hum Genet* 105(1-2):151-6.

Poupetova H, Ledvinova J, Berna L, Dvorakova L, Kozich V, Elleder M. 2010. The birth prevalence of lysosomal storage disorders in the Czech Republic: Comparison with data in different populations. *J Inherit Metab Dis* 33(4):387-96.

Pryor PR, Mullock BM, Bright NA, Gray SR, Luzio JP. 2000. The role of intraorganellar Ca^{2+} in late endosome-lysosome heterotypic fusion and in the reformation of lysosomes from hybrid organelles. *J Cell Biol* 149(5):1053-62.

Ribeiro FM, Pires RG, Ferguson SS. 2011. Huntington's disease and group I metabotropic glutamate receptors. *Mol Neurobiol* 43(1):1-11.

Richard M, Arfi A, Rhinn H, Gandolphe C, Scherman D. 2008. Identification of new markers for neurodegeneration process in the mouse model of Sly disease as revealed by expression profiling of selected genes. *J Neurosci Res* 86(15):3285-94.

Rubinsztein DC. 2006. The roles of intracellular protein-degradation pathways in neurodegeneration. *Nature* 443(7113):780-6.

Sardiello M, Palmieri M, di Ronza A, Medina DL, Valenza M, Gennarino VA, Di Malta C, Donaudy F, Embrione V, Polishchuk RS, et al. 2009. A gene network regulating lysosomal biogenesis and function. *Science* 325(5939):473-7.

Settembre C, Fraldi A, Jahreiss L, Spampinato C, Venturi C, Medina D, de Pablo R, Tacchetti C, Rubinsztein DC, Ballabio A. 2008. A block of autophagy in lysosomal storage disorders. *Hum Mol Genet* 17(1):119-29.

Simonaro CM, D'Angelo M, He X, Eliyahu E, Shtraizent N, Haskins ME, Schuchman EH. 2008. Mechanism of glycosaminoglycan-mediated bone and joint disease: Implications for the mucopolysaccharidoses and other connective tissue diseases. *Am J Pathol* 172(1):112-22.

Thathiah A and De Strooper B. 2011. The role of G protein-coupled receptors in the pathology of Alzheimer's disease. *Nat Rev Neurosci* 12(2):73-87.

Vahedi S, Rajabian M, Misaghian A, Grbec D, Simon HH, Alavian KN. 2010. Parkinson's disease candidate gene prioritization based on expression profile of midbrain dopaminergic neurons. *J Biomed Sci* 17:66.

Vedolin L, Schwartz IV, Komlos M, Schuch A, Puga AC, Pinto LL, Pires AP, Giugliani R. 2007. Correlation of MR imaging and MR spectroscopy findings with cognitive impairment in mucopolysaccharidosis II. *AJNR Am J Neuroradiol* 28(6):1029-33.

Vera R, Sanchez M, Galisteo M, Villar IC, Jimenez R, Zarzuelo A, Perez-Vizcaino F, Duarte J. 2007. Chronic administration of genistein improves endothelial dysfunction in spontaneously hypertensive rats: Involvement of eNOS, caveolin and calmodulin expression and NADPH oxidase activity. *Clin Sci (Lond)* 112(3):183-91.

Villani GR, Gargiulo N, Faraonio R, Castaldo S, Gonzalez Y, Reyero E, Di Natale P. 2007. Cytokines, neurotrophins, and oxidative stress in brain disease from mucopolysaccharidosis IIIB. *J Neurosci Res* 85(3):612-22.

Vitner EB, Platt FM, Futerman AH. 2010. Common and uncommon pathogenic cascades in lysosomal storage diseases. *J Biol Chem* 285(27):20423-7.

Wang D, Zhang W, Kalfa TA, Grabowski G, Davies S, Malik P, Pan D. 2009. Reprogramming erythroid cells for lysosomal enzyme production leads to visceral and CNS cross-correction in mice with Hurler syndrome. *Proc Natl Acad Sci U S A* 106(47):19958-63.

Wang L, Feng Z, Wang X, Wang X, Zhang X. 2010. DEGseq: An R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* 26(1):136-8.

Wang Z, Gerstein M, Snyder M. 2009. RNA-seq: A revolutionary tool for transcriptomics. *Nat Rev Genet* 10(1):57-63.

Wang ZL, Sun JY, Wang DN, Xie YH, Wang SW, Zhao WM. 2006. Pharmacological studies of the large-scaled purified genistein from huaijiao (*sophora japonica-leguminosae*) on anti-osteoporosis. *Phytomedicine* 13(9-10):718-23.

Wilkerson MJ, Lewis DC, Marks SL, Prieur DJ. 1998. Clinical and morphologic features of mucopolysaccharidosis type II in a dog: Naturally occurring model of Hunter syndrome. *Vet Pathol* 35(3):230-3.

Wraith JE, Scarpa M, Beck M, Bodamer OA, De Meirleir L, Guffon N, Meldgaard Lund A, Malm G, Van der Ploeg AT, Zeman J. 2008. Mucopolysaccharidosis type II (Hunter syndrome): A clinical review and recommendations for treatment in the era of enzyme replacement therapy. *Eur J Pediatr* 167(3):267-77.

Young ID, Harper PS, Newcombe RG, Archer IM. 1982. A clinical and genetic study of Hunter's syndrome. 2. differences between the mild and severe forms. *J Med Genet* 19(6):408-11.

DATABASE E SITI WEB

<http://www.biobase-international.com/>

www.biocarta.com

www.bioconductor.org/packages/2.6/bioc/html/DEGseq.html

www.cellsignal.com

<http://david.abcc.ncifcrf.gov>

<http://www.genome.jp/kegg/>

<http://genome.ucsc.edu>

[http:// www.ingenuity.com](http://www.ingenuity.com)

www.novusbio.com

<http://www.pantherdb.org/>

<http://pass.cribi.unipd.it>

<http://www.sabiosciences.com>

www.sigmaaldrich.com

<http://solid.appliedbiosystems.com>

www.themedicalbiochemistrypage.org