

DIA1R Is an X-Linked Gene Related to *Deleted In Autism-1*

Azhari Aziz, Sean P. Harrop[‡], Naomi E. Bishop*

Department of Microbiology, La Trobe University, Bundoora, Victoria, Australia

Abstract

Background: Autism spectrum disorders (ASDs) are frequently occurring disorders diagnosed by deficits in three core functional areas: social skills, communication, and behaviours and/or interests. Mental retardation frequently accompanies the most severe forms of ASDs, while overall ASDs are more commonly diagnosed in males. Most ASDs have a genetic origin and one gene recently implicated in the etiology of autism is the *Deleted-In-Autism-1* (*DIA1*) gene.

Methodology/Principal Findings: Using a bioinformatics-based approach, we have identified a human gene closely related to *DIA1*, we term *DIA1R* (*DIA1-Related*). While *DIA1* is autosomal (chromosome 3, position 3q24), *DIA1R* localizes to the X chromosome at position Xp11.3 and is known to escape X-inactivation. The gene products are of similar size, with *DIA1* encoding 430, and *DIA1R* 433, residues. At the amino acid level, *DIA1* and *DIA1R* are 62% similar overall (28% identical), and both encode signal peptides for targeting to the secretory pathway. Both genes are ubiquitously expressed, including in fetal and adult brain tissue.

Conclusions/Significance: Examination of published literature revealed point mutations in *DIA1R* are associated with X-linked mental retardation (XLMR) and *DIA1R* deletion is associated with syndromes with ASD-like traits and/or XLMR. Together, these results support a model where the *DIA1* and *DIA1R* gene products regulate molecular traffic through the cellular secretory pathway or affect the function of secreted factors, and functional deficits cause disorders with ASD-like symptoms and/or mental retardation.

Citation: Aziz A, Harrop SP, Bishop NE (2011) *DIA1R* Is an X-Linked Gene Related to *Deleted In Autism-1*. PLoS ONE 6(1): e14534. doi:10.1371/journal.pone.0014534

Editor: Jörg Hoheisel, Deutsches Krebsforschungszentrum, Germany

Received: July 21, 2010; **Accepted:** December 21, 2010; **Published:** January 17, 2011

Copyright: © 2011 Aziz et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: NEB and SPH were supported by La Trobe University and AA by La Trobe University, the Malaysian Ministry of Higher Education, and the Islamic Science University of Malaysia. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: N.Bishop@latrobe.edu.au

[‡] Current address: Infectious Disease Screening Laboratory, Australian Red Cross Blood Service, Southbank, Victoria, Australia

Introduction

Autism spectrum disorders (ASDs) encompass a variety of syndromes, diagnosed on the basis of three core symptoms: deficits in social skills, impaired use of language, and restricted and stereotypic behaviours and/or interests [1]. Classical ‘autistic disorder’ (AD) falls under the ASD umbrella and, in addition to the three core symptoms, is frequently accompanied by mental retardation, as indicated by an intelligence quotient (IQ) of <70 [2–11]. By contrast, presence of the core ASD symptoms accompanied by an average (or above) IQ is typically classified as Asperger syndrome or high-functioning autism [6,12,13]. The term pervasive development disorder (PDD) is often used as a broader diagnostic descriptor, to encompass not only all ASDs, but also PDD-not-otherwise-specified (PDD-NOS or atypical autism), Rett disorder and childhood disintegrative disorder [14]. ASDs cause morbidity in as many as 1:60 people, with the incidence being as high as 1:40 in males [15–19].

ASDs have an almost exclusive genetic origin [20–25]. The concordance rate for monozygotic twins, from various studies, ranges from 36% to as high as 100% in one study on female twins, while concordance rates of dizygotic twins are reported to vary from 0% to 31%, with values as high as 40% for male-male twin pairs, and heritability estimates of ~90% have been calculated

[21,25–35]. In twin studies, when a broader definition of ASD is used, higher concordance rates are found; while lower rates are determined when a stricter definition of autistic disorder is used. These findings indicate that the severity of ASD symptoms (e.g. AD versus AS/HFA or PDD-NOS) and the age of diagnosis can differ between monozygotic twins diagnosed with ASD [26,28–30,34]. As a result, epigenetic influences have been suggested to contribute to the overall ASD phenotype [36–39]. Recent data also indicate a further ~10% of ASD cases are caused by spontaneous genetic mutations [40–43]. Currently, more than 25 different loci encode ASD-susceptibility genes, with many more being investigated [44]. Some, but not all, of the genes implicated in ASD overlap with those implicated in MR [45–48]. Nonetheless, no genetic susceptibility locus accounts for more than a small percentage of ASD cases [49]. While these findings demonstrate an etiological heterogeneity for ASDs, many of the genes appear to interact at the level of molecular pathways, making it likely that mutations of one ASD-implicated gene may affect the function of others [50]. These findings suggest a common pathway(s) for the etiology of ASDs and raise hope that common treatments may be developed.

It is rare for those with ASDs to lead independent lives in adulthood, and those with severe intellectual disability (IQ<50)

have the poorest outcomes [51–55]. By contrast, the outcome is improved for those with an ASD who lack accompanying mental retardation (MR), and 10–15% of those with Asperger syndrome/high-functioning ASD achieve independence in adulthood [55–57]. While these studies highlight the debilitating life-long effects of ASDs, they emphasize the additional problems conferred by the presence of intellectual disability. As with ASDs, X-linked mental retardation (XLMR) is more common in males and also has a heterogeneous genetic etiology [58]. MR affects around 2% of the population and accounts for 5–10% of the health care budget in developed countries [58], with XLMR representing between 5–15% of all MR [59]. Mutations in approximately 10% of the 2,000 genes on the human X chromosome can cause XLMR, although the cellular role of most of these genes remains uncharacterized [60].

We report the identification of a human gene related to the recently identified *Deleted In Autism 1 (DIA1)* gene. The *DIA1* gene was identified in a genetic study for ASD genes [61], where it was found that hemizygous *DIA1* deletions were asymptomatic, while a homozygous *DIA1* deletion coincided with a classical autism diagnosis. Using a bioinformatics-based approach, we have identified a *DIA1*-related gene, we term *DIA1R*, which localizes to the human X chromosome at position Xp11.3. Deletion and/or mutation of human *DIA1R* is associated with ASD-like syndromes and/or XLMR.

Results

Identification of a human *DIA1*-related gene

The human *c3orf58* gene, found at chromosome position 3q24, has recently been renamed *DIA1* on the basis of its deletion in ASD [61]. Human *DIA1* has a known orthologue in both the mouse (*Mus musculus*) and rat (*Rattus norvegicus*) genomes, where it is

alternatively called *1190002N15Rik*, *Ab2-095* or *GoPro49* [62]. A search of the HomoloGene database [63], revealed further *DIA1* orthologues in the chimpanzee (*Pan troglodytes*), dog (*Canis familiaris*), cow (*Bos taurus*), mouse (*Mus musculus*), chicken (*Gallus gallus*), and zebrafish (*Danio rerio*) genomes (data not shown; [64]). Unexpectedly, a Basic Local Alignment Search Tool (BLAST) search [65] of the non-redundant (NCBI) database [63,66], using the sequence of the human *DIA1* gene product, revealed a significantly similar human gene product, LOC79742. This gene product was encoded by the *cXorf36* gene (for accession numbers and identifiers, see Table 1) and localized to the human X chromosome at position Xp11.3 (Figure 1). We have renamed this gene *DIA1R (DIA1-Related)*. A search of the HomoloGene database [63] revealed orthologues of *DIA1R* in the cow, mouse, chicken, and zebrafish genomes (data not shown).

Comparison of the human *DIA1* and *DIA1R* gene products

The human *DIA1* and *DIA1R* genes encode gene products of 430 and 433 residues, respectively, with predicted molecular weights of 49.5 and 48.6 kDa. To compare the human *DIA1* and *DIA1R* proteins with each other, we used two methods: BLAST analyses and amino acid alignments. Pair-wise protein BLAST analyses generate ‘expect values’ (E-values), where values less than one are considered significantly similar, and the smaller the E-value, the greater the similarity [65]. E-values of 2e-42 and 1e-42 were obtained for reciprocal BLAST searches, comparing the human *DIA1* gene product with that of *DIA1R*, or vice versa. Next, amino acid alignments were employed to compare the human *DIA1* and *DIA1R* proteins at the amino acid level (Figure 2). Amino acid alignments revealed human *DIA1* and *DIA1R* are 62% similar overall (28% identical, and 34% similar, residues), with similarity greater in the central portion of the two proteins,

Table 1. Accession numbers and identifiers for human *DIA1* and *DIA1R*.

Database	<i>DIA1</i> identifier or accession number	<i>DIA1R</i> identifier or accession number
HUGO gene symbol*	c3orf58	cXorf58
RefSeq** protein	NP_775823	NP_789789
RefSeq mRNA	NM_173552	NM_176819
RefSeq chromosome	NC_000003	NC_000023
GenInfo (GI)	27734895	193804854
ENSEMBL protein	ENSP00000320081	ENSP00000381086
ENSEMBL transcript	ENST00000315691	ENST00000398000
ENSEMBL gene	ENSG00000181744	ENSG00000147113
UniProt	Q8NDZ4	Q9H7Y0
UniGene	Hs.288954	Hs.98321
Entrez gene GeneID	205428	79742
HGNC***	28490	25866
Aliases	LOC205428 MGC33365 GoPro49 Ab2-095 DIA1	LOC79742 DKFZp313K0825 EPQL1862 FLJ55198 FLJ14103 PRO3743 bA435K1.1 FLJ55198

*Human genome organization (HUGO) official gene name.

**NCBI non-redundant reference sequence [63].

***HUGO gene nomenclature committee (HGNC) identifier.

doi:10.1371/journal.pone.0014534.t001

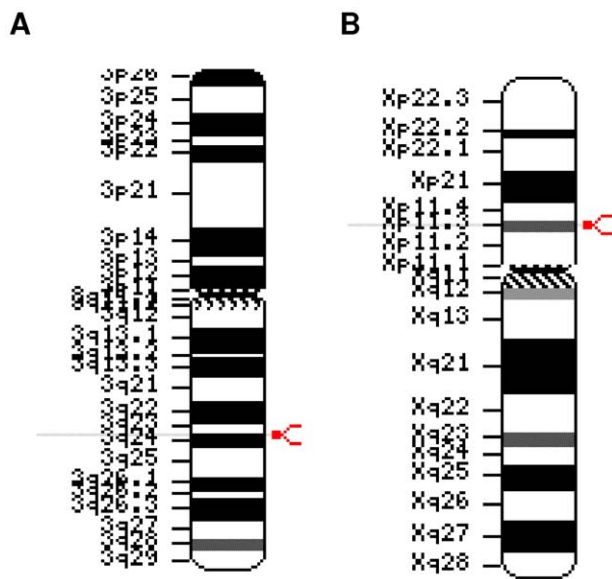


Figure 1. Chromosomal ideogram representation of the location of human *DIA1* (*c3orf58*) and *DIA1R* (*cXorf36*). (A) *DIA1* localization on human chromosome 3 at position 3q24. (B) *DIA1R* localization on the human X chromosome at position Xp11.3. Gene localization is indicated by a modified red arrowhead on the right-hand side, and a light-grey line extending through the text on the left-hand side.

doi:10.1371/journal.pone.0014534.g001

compared to the amino- or carboxy-terminal regions. This similarity suggests that the central region of *DIA1* and *DIA1R* encodes a biological function common to the two gene products.

Signal peptides in human *DIA1* and *DIA1R*

Most proteins transported to the endoplasmic reticulum (ER) have a stretch of hydrophobic amino acids at their amino-terminus, known as the signal peptide (SP), which are recognized by the signal recognition particle (SRP) to facilitate ER import [67]. The length of signal peptides usually falls within a range of 5–60 residues. Subsequent to SRP recognition, SPs are typically cleaved from the imported pre-protein by signal peptide peptidase activity [68]. Most imported proteins are then transported from the ER to the Golgi apparatus [69]. As determined above, the most disparate region found when *DIA1* and *DIA1R* are aligned (Figure 2) is the extreme amino-terminus. In a previous assessment of the rat *DIA1* gene product, the presence of a predicted signal peptide was detected, and functionally supported by localization of *DIA1* to the lumen of the Golgi apparatus [62]. We therefore analyzed the human *DIA1* and *DIA1R* gene products for the presence of amino-terminal signal peptides.

There are a variety of algorithms designed to predict the presence of signal peptides, and we used three of the best-performing algorithms [70,71] to both analyze the *DIA1* and *DIA1R* gene products for signal peptides, and to predict the signal peptide cleavage sites. The methodologies used were: (i) the neural network (NN) algorithm of SignalP v3.0 [72], (ii) the hidden Markov model (HMM) of SignalP v3.0 [72], and (iii) the SigCleave algorithm [73] at EMBOSS [74]. No trans-membrane domains were predicted, nor ER-retrieval or retention motifs, in either *DIA1* or *DIA1R* (data not shown).

Our analyses revealed SPs in human *DIA1* and *DIA1R*, using all three prediction methods (Figures 3 and S1; data not shown).

However, while all methods predicted a SP in both *DIA1* and *DIA1R*, some methods used had varying success in pin-pointing the exact cleavage site used to generate the mature protein. The HMM algorithm of SignalP confidently (score of 0.544, where 0.5 is considered significant) predicted the signal peptide cleavage site of human *DIA1* (Figure 3A) as between amino acids 37 and 38 (SLLA-SWQR using single-letter amino-acid code, where the hyphen indicates the site of cleavage). The same site was also the highest-scoring cleavage site predicted by SigCleave (score 7.79, where a score over 3.50 is considered significant), providing confidence that this is the site of human *DIA1* SP cleavage. By contrast, the NN algorithm did not predict a *DIA1* cleavage site with a significant score (Figure S1A). For human *DIA1R*, two cleavage sites were suggested by the HMM algorithm (Figure 3B), but neither score passed into the significant range (0.26 for cleavage between amino acids 31 and 32, 0.35 for between amino acids 38 and 39, where 0.5 is considered significant). The NN algorithm (Figure S1B) only predicted a single *DIA1R* cleavage site (between amino acids 31 and 32) with a Y-score close to significance (a Y-score of 0.44, where a Y-max value of 0.5 is considered significant). SigCleave analysis of *DIA1R* (data not shown) provided a highly significant score for cleavage between amino acids 31 and 32 (CSFS-LPAS, using single-letter amino-acid code, where the hyphen indicates the site of cleavage) of 10.23 (where values of greater than 3.5 are considered significant), suggesting this may represent the true site of SP cleavage for human *DIA1R*. Therefore, while all methods indicate amino-terminal SPs in both *DIA1* and *DIA1R*, the exact SP cleavage sites will need to be determined experimentally. Together, our data indicate that both the *DIA1R* and *DIA1* gene products will be translocated into the ER, and transported to the Golgi apparatus.

Ubiquitous tissue expression of human *DIA1* and *DIA1R*

To determine the cell types and tissues in which human *DIA1* and *DIA1R* are expressed, we explored the relevant expressed sequence tag (EST) profiles via UniGene [63] and searched two microarray databases: the GeneNote normal human expression database at the Weizmann Institute [75] and human Gene Atlas [76] at BioGPS [77].

EST profile data detected *DIA1* expression at levels of 61 transcripts per million (TPM) isolated from brain tissue (Table 2) and *DIA1R* at 25 TPM from brain tissue (Table 3). Of note, a wide analysis of human gene expression levels, in 14 different tissues, selected cut-off values for gene expression, with 'low' being ≤ 37 TPM and 'high' being ≥ 134 TPM [78]. Another human expression study indicated the majority (64%) of genes are expressed at between 3–50 TPM, 6% at levels between 100–5,000 TPM, and only 0.03% expressed at levels above 10,000 TPM [79]. Using these values as a guide, *DIA1R* can be described as being expressed at low-to-intermediate levels. By contrast, while *DIA1* is frequently expressed at low-to-intermediate levels, expression is high in mammary gland (135 TPM), 'mouth' (223 TPM; where libraries included in this tissue category are often derived from head or tongue), and vascular tissue (135 TPM). Ear expression of *DIA1* was the highest of those tissues falling in the intermediate range, and is close (123 TPM) to the 'high expression' category. These four tissue types, apparently expressing high levels of human *DIA1*, are all tissues with a high secretory output. One explanation is that Golgi apparatus-localized *DIA1* has a role in an aspect of secretion common to these tissues. Abnormalities in secreted levels of vascular growth factors and soluble vascular receptors have been reported in patients with ASD [80], as have abnormal levels of the endothelial cell protein PECAM-1 (platelet endothelial cell adhesion molecule-1) [81]. Defects in secretion or

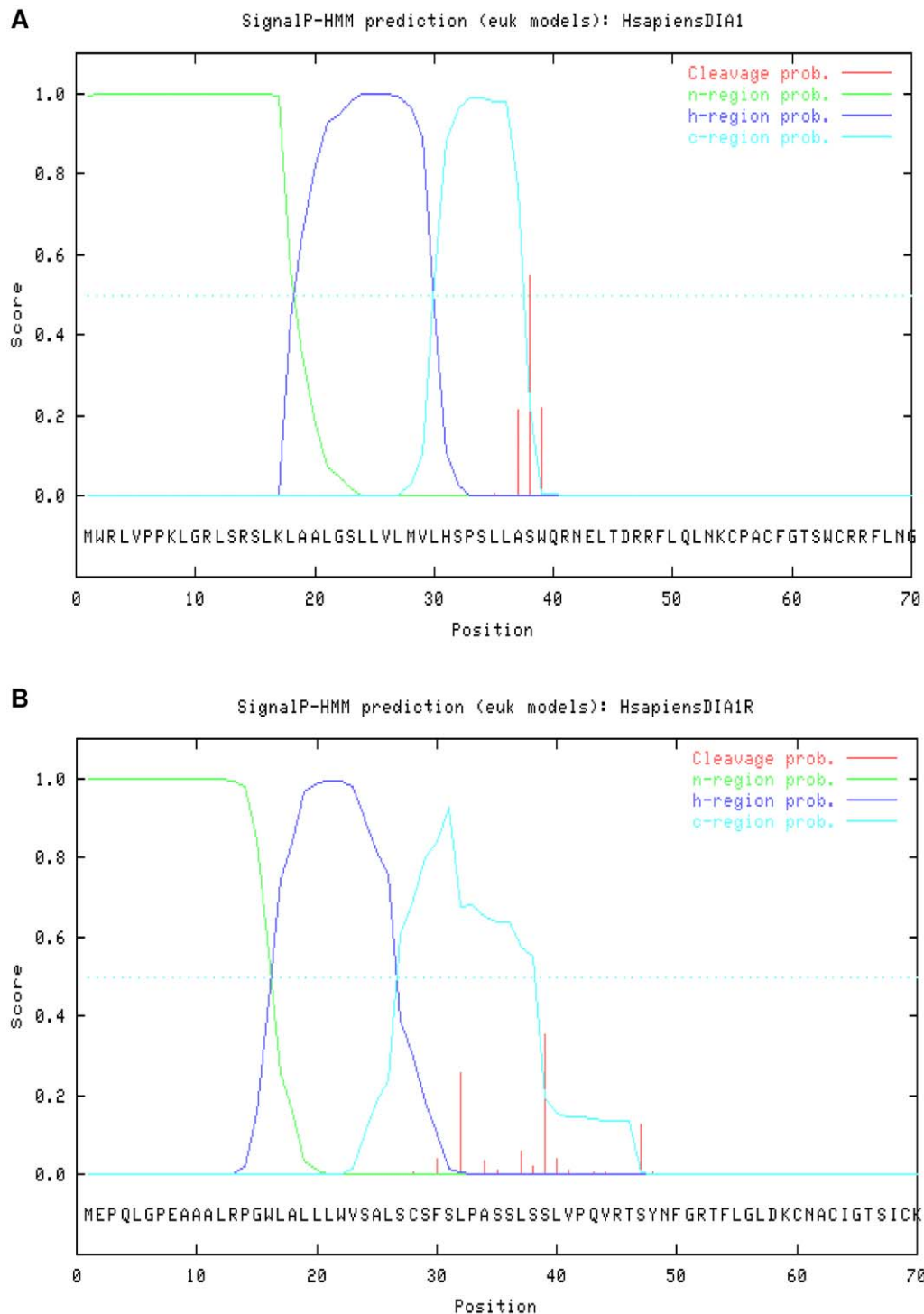


Figure 3. Prediction of signal peptides in human *DIA1* and *DIA1R*. The amino acid sequence of (A) human *DIA1* or (B) human *DIA1R*, was evaluated for amino-terminal signal peptides using the hidden Markov model of SignalP v3.0 [72]. C-region (cleavage-recognition region) scores are in aqua, H-region (hydrophobic region) scores are in dark blue, N-region (N-terminal signal peptide sequence) scores are in green, and the cleavage probability (that a given amino acid is the first in the mature protein) indicated in red. The significance cut-off value for all probability scores is 0.5, and is indicated by a blue dotted line. Standard single-letter amino acid abbreviations and numbering is provided below the graph.
doi:10.1371/journal.pone.0014534.g003

found by microarray analysis. By microarray analysis, the highest level of *DIA1* was found in BDCA-4 (blood dendritic cell antigen 4 or neuropilin-1 or CD304)-positive plasmacytoid dendritic cells (Figure S2), which secrete high levels of type I interferons [86]. By

contrast to the *DIA1* microarray data, microarray expression levels of *DIA1R* were less uniform (Figures 4 and S3). However, the highest *DIA1R* expression level found in the Gene Atlas microarray, which was found in liver (Figure S3), was not

Table 2. Expressed sequenced tag (EST) profiles showing *DIA1* gene expression in human tissues.

cDNA library source*	TPM**	EST/total***
adipose tissue	0	0/13105
adrenal gland	0	0/33195
ascites	0	0/40013
bladder	0	0/29757
blood	0	0/123476
bone	13	1/71655
bone marrow	81	4/48798
brain	59	66/1100969
cervix	41	2/48171
connective tissue	6	1/149254
ear	123	2/16212
embryonic tissue	23	5/215722
esophagus	0	0/20208
eye	37	8/211052
heart	33	3/89625
intestine	17	4/234477
kidney	56	12/211769
larynx	41	1/24144
liver	19	4/207739
lung	17	6/336969
lymph	0	0/44269
lymph node	54	5/91607
mammary gland	135	20/153267
mouth	223	15/67053
muscle	9	1/107711
nerve	0	0/15768
ovary	68	7/102050
pancreas	4	1/214811
parathyroid	48	1/20540
pharynx	0	0/41328
pituitary gland	0	0/16584
placenta	42	12/280828
prostate	21	4/189352
salivary gland	0	0/20155
skin	4	1/210574
spleen	0	0/53953
stomach	10	1/96622
testis	39	13/330449
thymus	36	3/81130
thyroid	0	0/81130
tonsil	0	0/16999
trachea	38	2/52412
umbilical cord	0	0/13680
uterus	12	3/232876
vascular	135	7/51779

*The EST profiles show approximate gene expression patterns, as inferred from EST counts from cDNA library sources, reported by sequence submitters to Unigene at the NCBI [63]. Libraries known to be normalized, subtracted, or otherwise biased are not included.

**TPM = transcripts per million.

***EST/total = number of *DIA1* ESTs in the total EST pool.

doi:10.1371/journal.pone.0014534.t002

Table 3. EST profiles of *DIA1R* gene expression in human tissues.

cDNA library source*	TPM**	EST/total***
adipose tissue	76	1/13105
adrenal gland	0	0/33195
Ascites	0	0/40013
Bladder	0	0/29757
Blood	0	0/123476
Bone	0	0/71655
bone marrow	0	0/48798
Brain	25	28/1100969
Cervix	0	0/48171
connective tissue	40	6/149254
Ear	0	0/16212
embryonic tissue	0	0/215722
esophagus	0	0/20208
Eye	0	0/211052
Heart	78	7/89625
Intestine	8	2/234477
Kidney	9	2/211769
Larynx	0	0/24144
Liver	4	1/207739
Lung	5	2/336969
Lymph	0	0/44269
lymph node	0	0/91607
mammary gland	32	5/153267
Mouth	29	2/67053
Muscle	27	3/107711
Nerve	0	0/15768
Ovary	0	0/102050
pancreas	4	4/214811
parathyroid	0	0/20540
Pharynx	0	0/41328
pituitary gland	0	0/16584
Placenta	60	17/280828
Prostate	36	7/189352
salivary gland	0	0/20155
Skin	0	0/210574
Spleen	0	0/53953
Stomach	10	1/96622
Testis	0	0/330449
Thymus	0	0/81130
Thyroid	21	1/81130
Tonsil	0	0/16999
Trachea	19	1/52412
umbilical cord	73	1/13680
Uterus	17	4/232876
Vascular	57	3/51779

*The EST profiles show approximate gene expression patterns, as inferred from EST counts from cDNA library sources, reported by sequence submitters to Unigene at the NCBI [63]. Libraries known to be normalized, subtracted, or otherwise biased are not included.

**TPM = transcripts per million.

***EST/total = number of *DIA1* ESTs in the total EST pool.

doi:10.1371/journal.pone.0014534.t003

supported by the GeneNote microarray data (Figure 4). Within brain tissue, the highest expression level of *DIA1R* was in the cerebellar peduncles (Figure S3), a region of the brain abnormal in those with ASD [87–89]. However, microarray studies are plagued by a lack of reproducibility and accuracy [90–93], and we can not place too much emphasis on differences between gene expression levels, but must focus on overall expression profiles. While quantitative reverse-transcriptase polymerase chain reactions (RT-PCR) can provide more accurate and reproducible results [94], and can be used in conjunction with commercially available human tissue panels, analysis of data from this approach relies on the choice of reference gene for normalization [95]. Changes in the comparable housekeeping gene can lead to changes in the significance and expression of the target gene using RT-PCR approaches to quantifying gene expression [96,97]. We can therefore only reliably conclude that human *DIA1* and *DIA1R* are both ubiquitously expressed genes.

Expression of human *DIA1* and *DIA1R* during development

To examine expression of *DIA1* and *DIA1R* during human development, microarray data from the Gene Atlas [76] at the BioGPS [77] database (Figure S2 and S3) and EST profile data from UniGene [63], were examined (Tables 4 and 5). As EST data was restricted by the depth at which each developmental stage could be sampled, absences of expression cannot be assumed to indicate a lack of expression. Despite the limitations of EST profiling, the data revealed that both *DIA1* and *DIA1R* are expressed in fetal tissue, with *DIA1* expression also detected in the blastocyst (Tables 4 and 5). Microarray data confirmed expression

Table 4. EST profiles of *DIA1* gene expression during human development.

cDNA library source*	TPM**	EST/total***
embryoid body	42	3/70760
blastocyst	32	2/62318
fetus	44	25/564004
neonate	0	0/31097
infant	0	0/23620
juvenile	0	0/55555
adult	31	62/1939097

*The EST profiles show approximate gene expression patterns, as inferred from EST counts from cDNA library sources, reported by sequence submitters to Unigene at the NCBI [63]. Libraries known to be normalized, subtracted, or otherwise biased are not included.

**TPM = transcripts per million.

***EST/total = number of *DIA1* ESTs in the total EST pool.

doi:10.1371/journal.pone.0014534.t004

of *DIA1* and *DIA1R* in fetal brain (Figure S2 and S3). Therefore, human *DIA1* and *DIA1R* are expressed in both adult and fetal tissues, including adult and fetal brain.

Discussion

The *DIA1* gene, recently identified in a genetic study detecting autism genes by tracing recent shared ancestry, localizes to chromosome 3 at position 3q24 [61]. Using a bioinformatics

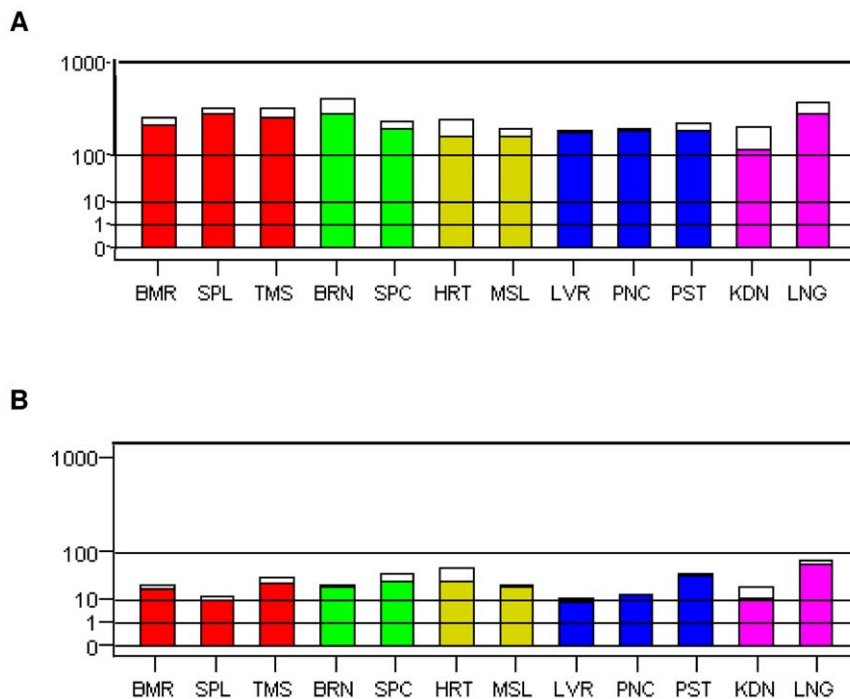


Figure 4. Microarray gene expression data for *DIA1* and *DIA1R*. Graphical presentation of microarray expression data for (A) *DIA1* or (B) *DIA1R* in normal, healthy human tissues. The intensity values (shown on the root-scale y-axis) are average values, where variation in the range of measurements is represented by a white box above the coloured minimal-measurement bar. Tissues are grouped according to their origin and the groups coloured: red = immune system; green = nervous system; yellow = muscle; blue = secretory glands; pink = 'other'. Tissue abbreviation (on the x-axis) legend: BMR = bone marrow; SPL = spleen; TMS = thymus; BRN = brain; SPC = spinal cord; HRT = heart; MSL = skeletal muscle; LVR = liver; PNC = pancreas; PST = prostate; KDN = kidney; LNG = lung. Data were obtained from the GeneNote database of human genes at the Weizmann Institute [75], and are MAS5.0 normalized data obtained using the Affymetrix HG-U95 set A–E.
doi:10.1371/journal.pone.0014534.g004

Table 5. EST profiles of *DIA1R* gene expression during human development.

cDNA library source*	TPM**	EST/total***
embryoid body	0	0/70760
blastocyst	0	0/62318
Fetus	33	19/564004
Neonate	0	0/31097
Infant	0	0/23620
Juvenile	0	0/55555
Adult	23	45/1939097

*The EST profiles show approximate gene expression patterns, as inferred from EST counts from cDNA library sources, reported by sequence submitters to Unigene at the NCBI [42]. Libraries known to be normalized, subtracted, or otherwise biased are not included.

**TPM = transcripts per million.

***EST/total = number of *DIA1* ESTs in the total EST pool.

doi:10.1371/journal.pone.0014534.t005

approach, we have identified a *DIA1*-related gene, *DIA1R*, which localizes to the human X chromosome at position Xp11.3. As has been reported for mouse *DIA1* [64], we find both human *DIA1* and *DIA1R* are ubiquitously expressed, suggesting each gene has a generic cellular role. Furthermore, human *DIA1R*, like *DIA1* [62], encodes a signal peptide for targeting to the secretory pathway. Our comparison of human *DIA1* and *DIA1R* at the amino acid level found conservation greatest between the central portions of the proteins, indicating this region may contribute to a biological role shared by each protein. The presence of *DIA1R* in the human genome may therefore explain why a homozygous deletion of a ubiquitously-expressed gene such as *DIA1* is not lethal [61]. However, human *DIA1R*, only has 62% similarity overall to human *DIA1*, and presumably cannot fully compensate for loss of *DIA1*, which would explain why *DIA1* deletion is associated with the neurological deficits of autism [61].

Role of the *DIA1R* locus in autism and X-linked mental retardation

Around 10% of genes on the X chromosome cause XLMR if inactivated [60]. Human *DIA1R* is an X-linked gene at position Xp11.3, and is therefore part of the wider Xp11.3-Xp11.23 region. This wider region is implicated in a growing number of genetic diseases, including several eye disorders, XLMR, X-linked neuromuscular diseases, and susceptibility loci for schizophrenia and type 1 diabetes, with the Xp11.3 locus implicated specifically in XLMR [98,99]. However, in addition to 19 characterized genes, there are at least four uncharacterized genes in this region of the X chromosome, including *DIA1R*, whose contribution to the disease phenotypes mapped to this region have not been assessed.

We next highlight a number of publications that support the hypothesis that deletion of the Xp11.3 region encompassing *DIA1R* causes ASD-like symptoms and/or XLMR. Essentially, we have carried out a retrospective genotype-phenotype analysis based on previously published studies which characterized the phenotypic traits of patients with deletions encompassing *DIA1R*. Genotype-phenotype mapping is a concept first introduced in 1991 [100], and genotype-phenotype correlation is widely used to examine the contribution of a certain mutation (genotype) to the resulting clinical traits (phenotype) [101–106]. Deletions encompassing a certain chromosomal region have variable phenotypes, depending on the size and position of the deleted region. However,

features common to patients with deletions overlapping a locus of interest provide evidence for the specific effect of the loss of that gene or genes. The findings outlined below suggest *DIA1R* is joining an increasing number of genes able to cause both ASD and/or MR [45–48].

In the first study, examination of a patient with developmental delay, no expressive speech, mental retardation, and microcephaly, revealed a 4.0-Mb heterozygous deletion at Xp11.3–p11.4. This region included 5 candidate genes, including *DIA1R* and *CASK*, the latter of which was suggested to be causative due to its known role in synaptic function [107]. We suggest *DIA1R* deletion contributes to the phenotype of this patient.

Secondly, a linkage study of a four-generation family, implicated chromosomal region Xp11.3–11.21 in a novel neurological syndrome [108]. The syndrome is characterized by mild mental retardation, poor verbal skills, aggressive and/or agitated behaviour, and involuntary movements referred to as choreoathetosis. These symptoms overlap with those of ASD, and choreoathetosis is associated with Rett syndrome [109]. Reyniers and colleagues [108] discussed the possible role of 11 known genes with roles in neural signalling and neurotransmission, and how they might contribute to patient symptoms. The region does not include the *CASK* locus, but does include *DIA1R*, which can now also be considered a candidate gene for this novel syndrome.

A third study, by Zhang and colleagues [110], linked the region Xp11.3–Xq22.2 of male members of a large, cross-generational family group to poor verbal skills, mental retardation, and microcephaly. Zhang and coworkers discussed *OPHN1* as a candidate gene in this region, as it had previously been implicated in XLMR [110]. Nonetheless, the authors acknowledged that uncharacterized genes in this region may contribute to patient symptoms, and this may include *DIA1R*.

A fourth study further strengthens the argument that deletion of the Xp11.3 region, including *DIA1R*, is implicated in ASD-like symptoms and XLMR. This study identified a 5 Mb deletion at Xp11.3–11.4 in Turner syndrome (TS) patients [111], a region that does not involve *OPHN1*. TS is a relatively common genetic disorder, caused by X chromosome monosomy or other X chromosome defects. TS does not typically affect intellectual function or expressive verbal abilities, although it can be associated with reduced visual-spatial and executive skills, and demonstrable impairments in social skills [112–113]. Overall, TS is associated with a substantially increased risk of ASD [114]. Good and colleagues [111] found that deletion of Xp11.3–11.4 is associated with the development of ASD-like symptoms in TS patients, including poor eye contact, poor recognition of facial expression, and impaired social skills. Deletion of this region was also associated with increased amygdala volumes in TS patients, to levels typically associated with normal males [111]. Of note, many studies have also reported increased amygdala volumes in patients with ASD [115–118], but the phenomenon is not universal and the finding may be age-dependent [119].

Good and co-workers [111] hypothesized that there is a critical gene in the Xp11.3–11.4 region that is expressed in two copies in normal (46,XX) females, and that full dosage compensation does not occur in normal (46,XY) males or in Turner syndrome females (45,X), resulting in the increased amygdala volume found in normal males [120] and TS females [111]. They further suggested that this unknown gene has implications for autism and may contribute to the increased male susceptibility to ASD [111]. This gene cannot be *CASK*, as *CASK* undergoes X-inactivation [111]. As human *DIA1R* has recently been shown to escape X inactivation [121], *DIA1R* is a good candidate within this locus. Together, these

findings demonstrate that *DIA1R* is a candidate gene for ASD-like symptoms and/or XLMR.

A more direct link between *DIA1R* and XLMR comes from two independent studies, where point mutations in *DIA1R* were implicated in XLMR [60,122]. It should be noted that any ASD-like symptoms were not commented on in either of these studies, and may not have been evaluated or specifically ruled-out. In the first study, a mutation affecting the *DIA1R* signal peptide (a change of serine at position 24 to proline or 'S24P', using single amino acid abbreviations) was reported in a single XLMR patient, while the same mutation was absent in controls [122]. Our analyses reveal that the mutant *DIA1R* gene product (*DIA1R*-S24P) is still expected to have a functional signal peptide, but the mutation is predicted to change the signal peptide cleavage site (data not shown). Therefore, the *DIA1R*-S24P mutant is still expected to translocate to the secretory pathway, but the mutation may affect the size of the mature protein, with resulting impact on structure and function. In the absence of knowledge about ASD-like symptoms in this patient, the data currently suggest that mutation of *DIA1R* can cause XLMR, but complete deletion of *DIA1R* causes ASD-like symptoms with or without MR.

In the second study, a systematic large-scale screen for mutations implicated in XLMR found recurrent mutations in *DIA1R* [60]. In 67 XLMR patients a 383G>A mutation in *DIA1R* was found [60], where the mutation is part of an arginine-coding residue (codon: AGA), resulting in an arginine to lysine (mutated codon: AAA) change in the *DIA1R* gene product. The amino acid mutation (R128K) falls within the amino-terminal one-third of the *DIA1R* protein. While we are currently investigating the conservation of this amino acid in the wider *DIA1* and *DIA1R* family, the findings of Tarpey and colleagues [60] suggest an R128K mutation alters *DIA1R* function in the human brain. These two studies demonstrate that independent *DIA1R* mutations are specifically associated with human XLMR, while the deletion studies implicate *DIA1R* deletion in autism-like syndromes and/or XLMR.

Overall, multiple lines of evidence now support a role for *DIA1R* in ASD and/or mental retardation: (i) the genotype-phenotype correlation data provided above; (ii) the high degree of amino acid conservation between *DIA1* and *DIA1R*, where *DIA1* is implicated in ASD [61]; (iii) targeting of both gene products to the Golgi apparatus, due to the conservation of signal peptide sequence in *DIA1* and *DIA1R*; (iv) the ubiquitous tissue expression of both genes, including in brain tissue; (v) and data from our accompanying paper which indicates an origin of *DIA1R* from a *DIA1* gene duplication early in the vertebrate lineage [123]. Each of these findings indicate conserved gene function. Therefore, it is not surprising that mutation of *DIA1R*, appears to cause a similar phenotype to deletion *DIA1*. (How these genes might function within the Golgi apparatus to cause this phenotype will be discussed below.)

Disorders associated with the human 3q24 locus

In addition to the publication originally describing the homozygous deletion of *DIA1* in autism [61], wider deletions encompassing 3q24 (the region of human chromosome 3 encoding the *DIA1* gene), have also been reported. While each patient with a deletion encompassing 3q24 has unique features, presumably related to differences in deletion breakpoints between patients, several common features are found: developmental delay, mental retardation, and growth delay including microcephaly [124–126]. These three common symptoms overlap those described in a subset of PDD patients: females with Rett syndrome (where *MECP2* is mutated), patients with Angelman syndrome, and male Rett syndrome-like patients in which the *MECP2* gene is not implicated [127,128].

By contrast, a patient with a deletion mapping cytogenetically to 3q25, showed developmental delay, but lacked the delayed growth and microcephaly of larger deletions in this chromosomal area [129]. Another patient with a deletion around 3q25 was specifically described as having autistic traits, severe learning difficulties, hypogonadism, and dysmorphic features [130]. These latter two deletions [129,130] have only been characterized cytogenetically, which is a relatively imprecise method [131], and it would be beneficial to map the exact deletion boundaries to confirm whether the deleted regions encompass *DIA1*.

Therefore, while a single publication links homozygous *DIA1* deletion to autism [61], numerous studies report that deletion of portions of the 3q24 region encompassing *DIA1* are associated with developmental delay, ASD-like traits, and/or mental retardation [124–126]. Together, these publications provide further support for a role for *DIA1* in autism etiology, as first proposed by Morrow and colleagues in 2008 [61].

Why does *DIA1* or *DIA1R* mutation cause autism-like syndromes?

We have demonstrated that human *DIA1* and *DIA1R* both encode signal peptides for targeting to the secretory pathway. Indeed, the *DIA1* gene product has been localized, using immunofluorescence microscopy, to the lumen of the Golgi apparatus [62]. Surprisingly, for genes with roles in neurological diseases, we find both *DIA1* and *DIA1R* are ubiquitously expressed, which indicates they fulfil basic, non-specialist cellular roles. However, despite ASDs being diagnosed on the behavioural manifestations of neurological deficits [132–135], many co-morbid non-neuronal conditions occur in ASD patients, suggesting widespread physiological and biochemical abnormalities in ASD patients [136–139]. Secretory pathway malfunction is a plausible explanation for the wide-ranging deficits in ASD and, as both *DIA1* and *DIA1R* are predicted to play ubiquitous roles in some aspect of the cellular secretion pathway, they may be involved in a common cellular pathway deficient in those with ASDs.

The secretory pathway plays a key role in all cells, with essential roles in neuron development and function. Secretion in neurons, while having many factors in common to that in non-neuronal cells, has specific challenges. Recently, dendrites have been found to have satellite Golgi-like cisternal stacks, known as Golgi outposts, which have no membranous connection with the somatic Golgi [140]. What triggers Golgi outpost formation is not known, and only proteins common to the somatic Golgi have been identified in Golgi outposts [141]. Therefore, despite using the same secretory compartments as other cells, the dependence of neurons on the secretory pathway for synaptic transmission, outgrowth and remodelling of dendrites and axons, combined the vast distances involved, means the impact of mutation in many secretory pathway genes has a greater impact on neuronal, rather than non-neuronal, cell function [142]. Indeed, aberrations in cell secretion are being detected in increasing numbers of patients with ASD and/or mental retardation [143–149]. While large numbers of genes have been implicated as causative of ASD, we next discuss data suggesting that a common theme in ASD etiology may be alterations in cellular secretory pathway(s).

Recent work has found that four independent gene rearrangements, all previously implicated in ASD, induce the same morphological and functional aberrations in the large dense core granules of platelets [149]. It is also well-established that aberrant serotonin secretion from platelets is common in those with ASD [150–151], along with aberrant levels of other secreted factors, including neuropeptides and neurotrophins [137,143,144]. Furthermore, two research groups have found evidence for defective enzymatic activity in the secretory pathway of those with autism. The more recent of these two studies, found a functional deficit in

a secretory pathway-localized kinase by assaying substrates in the saliva of patients with ASD [152]. The activity of this ubiquitously expressed kinase can affect the traffic and/or function of its substrates [153–157]. The second research group found a deficit in the activity of tyrosylprotein sulfotransferase (TPST) in platelets of patients with ASD [158,159]. TPST is a ubiquitous sulfotransferase, which acts on substrates in the lumen of the Golgi apparatus [160]. Protein tyrosine sulfation is again known to affect the function of the target protein [161]. Therefore, an emerging theme in autism is defects in the delivery or function of secreted, or cell-surface, molecules. The location of *DIA1* and *DIA1R* suggests they may also affect such processes.

Further evidence supporting a role of secretory pathway deficits in ASD etiology, are studies identifying genes related to glycobiology in autism [48,162]. As the Golgi apparatus is the main site of glycosylation [163], this finding further implicates deficits in Golgi apparatus function in ASD etiology. Therefore, a unifying hypothesis is that genes involved in the pathogenesis of autism are involved in regulation of secretion, the expression of secreted proteins, or the optimal function of proteins trafficking via the secretory pathway. We propose a model where *DIA1* and *DIA1R* play a role in regulation of cargo secretion or modification of cargo molecules, which is required for optimal cargo functionality. We conclude that defects in *DIA1* or *DIA1R* function within the lumen of the Golgi apparatus of human cells causes autism and/or mental retardation. However, in the absence of known protein motifs and domains (data not shown), ongoing studies are required to determine the exact cellular roles of human *DIA1* and *DIA1R*, particularly the effects of mutation on brain function.

Materials and Methods

Detection and chromosome mapping of human *DIA1R*

The human *DIA1* amino acid sequence [61,62] was used in a BLAST search, using the BLASTP [65] program, and this identified the related human sequence, *DIA1R*, on the non-redundant protein database at the NCBI [63]. Chromosomal location and chromosome ideograms were obtained using MapViewer at NCBI [63].

Alignments, protein analysis, and homologue identification

Protein sequence alignments were generated with CLUSTALX version 1.8 [164]. The ExPASy Compute pI/MW tool [165] was used to calculate theoretical molecular weights. Three transmembrane prediction methods were used to analyze protein sequences: TMpred [166], TMAP [167], and HMMTOP version 2.0 [168]. Signal peptides were evaluated using SignalP version 3.0 [72] or the SigCleave algorithm [73] which is part of the EMBOSS software suite [74]. Amino acid motifs and domains were investigated using the following resources: MOTIF at GenomeNet [169]; PSORT-II [170]; the Conserved Domain Database at the NCBI, which also contains data from Pfam, SMART and COG [171]; and the Eukaryotic Linear Motif resource [172]. Homologs of human *DIA1* and *DIA1R* were identified using the NCBI HomoloGene resource [63].

Human tissue expression

EST profiles were obtained via the UniGene database at NCBI [63] and were accessed on 12th November 2010. Microarray data

was accessed from the GeneNote normal human expression database at the Weizman Institute [75] and the Gene Atlas database [76] via the BioGPS gene portal server [7].

Supporting Information

Figure S1 Neural network prediction of signal peptides in human *DIA1* and *DIA1R*. The amino acid sequence of (A) human *DIA1* or (B) human *DIA1R*, was evaluated for amino-terminal signal peptides using the neural network (NN) algorithm of SignalP v3.0 [72]. The C-score (cleavage score for each amino acid) is indicated in red, the S-score (signal peptide score) is indicated in green, and the Y-score (derived from the C-score and S-score, and can give a better indication of the cleavage site) is indicated in blue. The significance cut-off value for all probability scores is 0.5, and is indicated by a pink dotted line. Standard single-letter amino acid abbreviations and numbering are provided below the graph. Found at: doi:10.1371/journal.pone.0014534.s001 (0.02 MB PDF)

Figure S2 BioGPS microarray expression data for *DIA1*. Graphical presentation of expression data for *DIA1* in a variety of normal, and cancerous human tissues and cells. Expression values were obtained from an Affymetrix U133A microarray and relate to fluorescence intensity. Multiple probes were used for each transcript on the microarray and these intensity values have been normalized, background-subtracted, and summarized using the data-processing algorithm GCRMA (GeneChip Robust Multi-array Average). The identifier of the Affymetrix probe set used is indicated above the graph. Data were obtained from the Gene Atlas database of human genes [76] at the BioGPS gene portal server [77]. Found at: doi:10.1371/journal.pone.0014534.s002 (0.02 MB PDF)

Figure S3 BioGPS microarray expression data for *DIA1R*. Graphical presentation of expression data for *DIA1R* in a variety of normal, and cancerous human tissues and cells. Expression values were obtained from an Affymetrix U133A microarray and relate to fluorescence intensity. Multiple probes were used for each transcript on the microarray and these intensity values have been normalized, background-subtracted, and summarized using the GCRMA (GeneChip Robust Multi-array Average) data-processing algorithm. The identifier of the Affymetrix probe set used is indicated above the graph. Data were obtained from the Gene Atlas database of human genes [76] at the BioGPS gene portal server [77]. Found at: doi:10.1371/journal.pone.0014534.s003 (0.03 MB PDF)

Acknowledgments

We thank Cheryl Dissanayake and other researchers at the Olga Tennison Autism Research Centre, David Austin and Enzo Palombo at the Swinburne Autism Bio-Research Initiative, and Dennis Crowley of the Autism Victoria Board, for enthusiastic discussions about autism spectrum disorders. We are also indebted to all research scientists who are able to provide publicly-available high quality sequence databases, gene expression databases, and software for gene and gene-product analysis. Finally, we thank the reviewers for their thoughtful and constructive criticisms.

Author Contributions

Conceived and designed the experiments: NEB. Performed the experiments: AA SPH NEB. Analyzed the data: AA SPH NEB. Wrote the paper: NEB. Had input to the written paper content and discussion section: AA SPH.

References

- O'Hare A (2009) Autism spectrum disorder: diagnosis and management. Arch Dis Child Educ Pract Ed 94: 161–168.
- Lotter V (1966) Epidemiology of autistic conditions in young children: I. Prevalence. Soc Psychiatry Psychiatr Epidemiol 1: 124–135.

3. Gillberg C, Steffenburg S, Schaumann H (1991) Is autism more common now than ten years ago? *Br J Psychiatry* 158: 403–409.
4. Fombonne E, du Mazaubrun C (1992) Prevalence of infantile autism in four French regions. *Soc Psychiatry Psychiatr Epidemiol* 27: 203–210.
5. Nordin V, Gillberg C (1996) Autism spectrum disorders in children with physical or mental disability or both. II: Screening aspects. *Dev Med Child Neurol* 38: 314–324.
6. Chakrabarti S, Fombonne E (2001) Pervasive developmental disorders in preschool children. *JAMA* 285: 3093–3099.
7. Mayes SD, Calhoun SL (2004) Influence of IQ and age in childhood autism: Lack of support for DSM-IV Asperger's Disorder. *J Dev Phys Disabil* 16: 257–272.
8. Bryson SE, Bradley EA, Thompson A, Wainwright A (2008) Prevalence of autism among adolescents with intellectual disabilities. *Can J Psychiatry* 53: 449–459.
9. Nishiyama T, Taniai H, Miyachi T, Ozaki K, Tomita M, et al. (2009) Genetic correlation between autistic traits and IQ in a population-based sample of twins with autism spectrum disorders (ASDs). *J Hum Genet* 54: 56–61.
10. Hoekstra RA, Happé F, Baron-Cohen S, Ronald A (2009) Association between extreme autistic traits and intellectual disability: insights from a general population twin study. *Br J Psychiatry* 195: 531–536.
11. Hoekstra RA, Happé F, Baron-Cohen S, Ronald A (2010) Limited genetic covariance between autistic traits and intelligence: findings from a longitudinal twin study. *Am J Med Genet B Neuropsychiatr Genet* 153B: 994–1007.
12. World Health Organization (2007) Chapter V. Mental and behavioural disorders. In: *International Statistical Classification of Diseases and Related Health Problems. Tenth Revision*. Available: <http://apps.who.int/classifications/apps/icd/icd10online/>. Accessed April 2010.
13. American Psychiatric Association (1994) *Diagnostic and statistical manual of mental disorders*. Fourth Edition. Washington, DC: Author.
14. Santangelo SL, Tsatsanis K (2005) What is known about autism: Genes, brain, and behavior. *Am J Pharmacogenomic* 5: 71–92.
15. Baird G, Simonoff E, Pickles A, Chandler S, Loucas T, et al. (2006) Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet* 368: 210–215.
16. Honda H, Shimizu Y, Imai M, Nitto Y (2005) Cumulative incidence of childhood autism: a total population study of better accuracy and precision. *Dev Med Child Neurol* 47: 10–18.
17. Fombonne E (2005) Epidemiology of autistic disorder and other pervasive developmental disorders. *J Clin Psychiatry* 66: S3–8.
18. National Center for Health Statistics (2007) *National Survey of Children's Health (USA)*. Atlanta: Centers for Disease Control and Prevention. Available: <http://www.cdc.gov/nchs/slats/nsch.htm#2007nsch>. Accessed April 2010.
19. ADDM Network Surveillance Year 2006 Principal Investigators (2009) *Prevalence of Autism Spectrum Disorders-Autism and Developmental Disabilities Monitoring Network, United States, 2006*. *Morb Mortal Wkly Rep* 58: 1–20.
20. Rutter M (2000) Genetic studies of autism: from the 1970s into the millennium. *J Abnorm Child Psychol* 28: 3–14.
21. Folstein SE, Rosen-Sheidley B (2001) Genetics of autism: complex aetiology for a heterogeneous disorder. *Nat Rev Genet* 2: 943–955.
22. Veenstra-VanderWeele J, Cook EH, Jr. (2004) Molecular genetics of autism spectrum disorder. *Mol Psychiatry* 9: 819–832.
23. Persico AM, Bourgeron T (2006) Searching for ways out of the autism maze: genetic, epigenetic and environmental clues. *Trends Neurosci* 29: 349–358.
24. Losh M, Sullivan PF, Trembath D, Piven J (2008) Current developments in the genetics of autism: from phenome to genome. *J Neuropathol Exp Neurol* 67: 829–837.
25. Freitag CM (2007) The genetics of autistic disorders and its clinical relevance: a review of the literature. *Mol Psychiatry* 12: 2–22.
26. Folstein S, Rutter M (1977) Infantile autism: a genetic study of 21 twin pairs. *J Child Psychol Psychiatry* 18: 297–321.
27. Ritvo ER, Freeman BJ, Mason-Brothers A, Mo A, Ritvo AM (1985) Concordance for the syndrome of autism in 40 pairs of afflicted twins. *Am J Psychiatry* 142: 74–77.
28. Steffenburg S, Gillberg C, Hellgren L, Andersson L, Gillberg IC, et al. (1989) A twin study of autism in Denmark, Finland, Iceland, Norway and Sweden. *J Child Psychol Psychiatry* 30: 405–416.
29. Bailey A, Le Couteur A, Gottesman I, Bolton P, Simonoff E, et al. (1995) Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med* 25: 63–77.
30. Le Couteur A, Bailey A, Goode S, Pickles A, Robertson S, et al. (1996) A broader phenotype of autism: the clinical spectrum in twins. *J Child Psychol Psychiatry* 37: 785–801.
31. Kates WR, Burnette CP, Eliez S, Strunge LA, Kaplan D, et al. (2004) Neuroanatomic variation in monozygotic twin pairs discordant for the narrow phenotype for autism. *Am J Psychiatry* 161: 539–546.
32. Ronald A, Happé F, Bolton P, Butcher LM, Price TS, et al. (2006) Genetic heterogeneity between the three components of the autism spectrum: a twin study. *J Am Acad Child Adolesc Psychiatry* 45: 691–699.
33. Taniai H, Nishiyama T, Miyachi T, Imaeda M, Sumi S (2008) Genetic influences on the broad spectrum of autism: study of proband-ascertained twins. *Am J Med Genet B Neuropsychiatr Genet* 147B: 844–849.
34. Rosenberg RE, Law JK, Yenokyan G, McGready J, Kaufmann WE, et al. (2009) Characteristics and concordance of autism spectrum disorders among 277 twin pairs. *Arch Pediatr Adolesc Med* 163: 907–914.
35. Lichtenstein P, Carlström E, Råstam M, Gillberg C, Anckarsäter H (2010) The genetics of autism spectrum disorders and related neuropsychiatric disorders in childhood. *Am J Psychiatry* 167: 1357–1363.
36. Schanen NC (2006) Epigenetics of autism spectrum disorders. *Hum Mol Genet* 15: R138–150.
37. Ptak C, Petronis A (2010) Epigenetic approaches to psychiatric disorders. *Dialogues Clin Neurosci* 12: 25–35.
38. Grafodatskaya D, Chung B, Szatmari P, Weksberg R (2010) Autism spectrum disorders and epigenetics. *J Am Acad Child Adolesc Psychiatry* 49: 794–809.
39. Nguyen A, Rauch TA, Pfeifer GP, Hu VW (2010) Global methylation profiling of lymphoblastoid cell lines reveals epigenetic contributions to autism spectrum disorders and a novel autism candidate gene, RORA, whose protein product is reduced in autistic brain. *FASEB J* 24: 3036–3051.
40. Zhao X, Leotta A, Kustanovich V, Lajonchere C, Geschwind DH, et al. (2007) A unified genetic theory for sporadic and inherited autism. *Proc Natl Acad Sci U S A* 104: 12831–12836.
41. Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, et al. (2007) Strong association of de novo copy number mutations with autism. *Science* 316: 445–449.
42. Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, et al. (2008) Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* 82: 477–488.
43. Vorstman JA, van Daalen E, Jalali GR, Schmidt ER, Pasterkamp RJ, et al. (2010) A double hit implicates *DIAPH3* as an autism risk gene. *Mol Psychiatry*. In press.
44. Abrahams BS, Geschwind DH (2008) Advances in autism genetics: on the threshold of a new neurobiology. *Nat Rev Genet* 9: 341–355.
45. Kalscheuer VM, FitzPatrick D, Tommerup N, Bugge M, Niebuhr E, et al. (2007) Mutations in autism susceptibility candidate 2 (*AUTS2*) in patients with mental retardation. *Hum Genet* 121: 501–509.
46. Berkel S, Marshall CR, Weiss B, Howe J, Roeth R, et al. (2010) Mutations in the *SHANK2* synaptic scaffolding gene in autism spectrum disorder and mental retardation. *Nat Genet* 42: 489–491.
47. Kooy RF (2010) Distinct disorders affecting the brain share common genetic origins. *F1000 Biol Rep* 2: 11.
48. Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, et al. (2010) Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* 466: 368–372.
49. Geschwind DH (2009) Advances in autism. *Annu Rev Med* 60: 367–380.
50. Bill BR, Geschwind DH (2009) Genetic advances in autism: heterogeneity and convergence on shared pathways. *Curr Opin Genet Dev* 19: 271–278.
51. Nordin V, Gillberg C (1998) The long-term course of autistic disorders: update on follow-up studies. *Acta Psychiatr Scand* 97: 99–108.
52. Howlin P (2000) Autism and intellectual disability: diagnostic and treatment issues. *J R Soc Med* 93: 351–355.
53. Howlin P, Goode S, Hutton J, Rutter M (2004) Adult outcome for children with autism. *J Child Psychol Psychiatry* 45: 212–229.
54. Seltzer MM, Shattuck P, Abbeduto L, Greenberg JS (2004) Trajectory of development in adolescents and adults with autism. *Ment Retard Dev Disabil Res Rev* 10: 234–247.
55. Cederlund M, Hagberg B, Billstedt E, Gillberg IC, Gillberg C (2008) Asperger syndrome and autism: a comparative longitudinal follow-up study more than 5 years after original diagnosis. *Autism Dev Disord* 38: 72–85.
56. Engström I, Ekström L, Emilsson B (2003) Psychosocial functioning in a group of Swedish adults with Asperger syndrome or high-functioning autism. *Autism* 7: 99–110.
57. Renty JO, Roeyers H (2006) Quality of life in high-functioning adults with autism spectrum disorder: The predictive value of disability and support characteristics. *Autism* 10: 511–524.
58. Ropers HH (2006) X-linked mental retardation: Many genes for a complex disorder. *Curr Opin Genet Dev* 16: 260–269.
59. Stevenson RE (2000) Splitting and lumping in the nosology of *XLMR*. *Am J Med Genet* 97: 174–82.
60. Tarpey PS, Smith R, Pleasance E, Whibley A, Edkins S, et al. (2009) A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. *Nat Genet* 41: 535–543.
61. Morrow EM, Yoo SY, Flavell SW, Kim TK, Lin Y, et al. (2008) Identifying autism loci and genes by tracing recent shared ancestry. *Science* 321: 218–223.
62. Takatalo MS, Kouvonen P, Corthals G, Nyman TA, Rönnholm RH (2006) Identification of new Golgi complex specific proteins by direct organelle proteomic analysis. *Proteomics* 6: 3502–3508.
63. Sayers EW, Barrett T, Benson DA, Bolton E, Bryant SH, et al. (2010) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 38: D5–16.
64. Takatalo M, Järvinen E, Laitinen S, Thesleff I, Rönnholm R (2008) Expression of the novel Golgi protein *GoPro49* is developmentally regulated during mesenchymal differentiation. *Dev Dyn* 237: 2243–2255.
65. Altshul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402.

66. Pruitt KD, Tatusova, T, Maglott DR (2007) NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res* 35: D61–65.
67. Meacock SL, Greenfield JJ, High S (2000) Protein targeting and translocation at the endoplasmic reticulum membrane—through the eye of a needle? *Essays Biochem* 36: 1–13.
68. Halic M, Beckmann R (2005) The signal recognition particle and its interactions during protein targeting. *Curr Opin Struct Biol* 15: 116–125.
69. Teasdale RD, Jackson MR (1996) Signal-mediated sorting of membrane proteins between the endoplasmic reticulum and the Golgi apparatus. *Annu Rev Cell Dev Biol* 12: 27–54.
70. Menne KM, Hermjakob H, Apweiler R (2000) A comparison of signal sequence prediction methods using a test set of signal peptides. *Bioinformatics* 16: 741–742.
71. Klec EW, Ellis LB (2005) Evaluating eukaryotic secreted protein prediction. *BMC Bioinformatics* 6: 256.
72. Emanuelson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* 2: 953–971.
73. von Heijne G (1986) A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res* 14: 4683–4690.
74. Rice P, Longden I, Bleasby A (2000) EMBOSS: the European molecular biology open software suite. *Trends Genet* 16: 276–277.
75. Shmueli O, Horn-Saban S, Chalifa-Caspi V, Shmoish M, Ophir R, et al. (2003) GeneNote: whole genome expression profiles in normal human tissues. *C R Biol* 326: 1067–1072.
76. Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, et al. (2004) A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc Natl Acad Sci U S A* 101: 6062–6067.
77. Wu C, Orozco C, Boyer J, Leglise M, Goodale J, et al. (2009) BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol* 10: R130.
78. Brandenberger R, Khrebukova I, Thies RS, Miura T, Jingli C, et al. (2004) MPSS profiling of human embryonic stem cells. *BMC Dev Biol* 4: 10.
79. Lercher MJ, Urrutia AO, Hurst LD (2002) Clustering of housekeeping genes provides a unified model of gene order in the human genome. *Nat Genet* 31: 180–183.
80. Emanuele E, Orsi P, Barale F, di Nemi SU, Bertona M, et al. (2010) Serum levels of vascular endothelial growth factor and its receptors in patients with severe autism. *Clin Biochem* 43: 317–319.
81. Tsuchiya KJ, Hashimoto K, Iwata Y, Tsujii M, Sekine Y, et al. (2007) Decreased serum levels of platelet-endothelial adhesion molecule (PECAM-1) in subjects with high-functioning autism: a negative correlation with head circumference at birth. *Biol Psychiatry* 62: 1056–1058.
82. Lim DJ, Chun YM, Lee HY, Moon SK, Chang KH, et al. (2000) Cell biology of tubotympanum in relation to pathogenesis of otitis media - a review. *Vaccine* 19: S17–25.
83. Konstantareas MM, Homatidis S (1987) Ear infections in autistic and normal children. *J Autism Dev Disord* 17: 585–594.
84. Smith DE, Miller SD, Stewart M, Walter TL, McConnell JV (1988) Conductive hearing loss in autistic, learning-disabled, and normal children. *J Autism Dev Disord* 18: 53–65.
85. Rosenhall U, Nordin V, Sandström M, Ahlsén G, Gillberg C (1999) Autism and hearing loss. *J Autism Dev Disord* 29: 349–357.
86. Colonna M, Trinchieri G, Liu XJ (2004) Plasmacytoid dendritic cells in immunity. *Nat Immunol* 5: 1219–1226.
87. Catani M, Jones DK, Daly E, Embiricos N, Deeley Q, et al. (2008) Altered cerebellar feedback projections in Asperger syndrome. *Neuroimage* 41: 1184–1191.
88. Brito AR, Vasconcelos MM, Domingues RC, Hygino da Cruz LC, Jr., Rodrigues Lde S, et al. (2009) Diffusion tensor imaging findings in school-aged autistic children. *J Neuroimaging* 19: 337–343.
89. Sivaswamy L, Kumar A, Rajan D, Behen M, Muzik O, et al. (2010) A diffusion tensor imaging study of the cerebellar pathways in children with autism spectrum disorder. *J Child Neurol* 25: 1223–31.
90. Shi L, Tong W, Goodsaid F, Fruch FW, Fang H, et al. (2004) QA/QC: challenges and pitfalls facing the microarray community and regulatory agencies. *Expert Rev Mol Diagn* 4: 761–777.
91. Chen JJ, Hsueh HM, Delongchamp RR, Lin CJ, Tsai CA (2007) Reproducibility of microarray data: a further analysis of microarray quality control (MAQC) data. *BMC Bioinformatics* 8: 412.
92. Wilkes T, Laux H, Foy CA (2007) Microarray data quality - review of current developments. *OMICS* 11: 1–13.
93. Chiorino G, Mello Grand M, Scatolini M, Ostano P (2008) From single gene to integrative molecular concept MAPS: pitfalls and potentials of microarray technology. *J Biol Regul Homeost Agents* 22: 7–16.
94. Skrzypski M (2008) Quantitative reverse transcriptase real-time polymerase chain reaction in translational oncology: lung cancer perspective. *Lung Cancer* 59: 147–154.
95. Ohl F, Jung M, Xu C, Stephan C, Rabien A, et al. (2005) Gene expression studies in prostate cancer tissue: which reference gene should be selected for normalization? *J Mol Med* 83: 1014–1024.
96. Ståhlberg A, Zoric N, Aman P, Kubista M (2005) Quantitative real-time PCR for cancer detection: the lymphoma case. *Expert Rev Mol Diagn* 5: 221–230.
97. Passmore M, Nataatmadja M, Fraser JF (2009) Selection of reference genes for normalisation of real-time RT-PCR in brain-stem death injury in *Ovis aries*. *BMC Mol Biol* 10: 72.
98. Lubs HA, Chiurazzi P, Arena JF, Schwartz C, Tranebjærg L, et al. (1996) XLMR genes: update 1996. *Am J Med Genet* 64: 147–157.
99. Thiselton DL, McDowall J, Brandau O, Ramser J, d'Esposito F, et al. (2002) An integrated, functionally annotated gene map of the DXS8026-ELK1 interval on human Xp11.3-Xp11.23: potential hotspot for neurogenetic disorders. *Genomics* 79: 560–572.
100. Alberch P (1991) From genes to phenotype: dynamical systems and evolvability. *Genetica* 84: 5–11.
101. Shapira SK (1998) An update on chromosome deletion and microdeletion syndromes. *Curr Opin Pediatr* 10: 622–627.
102. Shaffer LG, Bejjani BA, Torchia B, Kirkpatrick S, Coppinger J, et al. (2007) The identification of microdeletion syndromes and other chromosome abnormalities: cytogenetic methods of the past, new technologies for the future. *Am J Med Genet C Semin Med Genet* 145C: 335–345.
103. Shaffer LG, Theisen A, Bejjani BA, Ballif BC, Aylsworth AS, et al. (2007) The discovery of microdeletion syndromes in the post-genomic era: review of the methodology and characterization of a new 1q41q42 microdeletion syndrome. *Genet Med* 9: 607–616.
104. Feenstra I, Vissers LE, Orsel M, van Kessel AG, Brunner HG, et al. (2007) Genotype-phenotype mapping of chromosome 18q deletions by high-resolution array CGH: an update of the phenotypic map. *Am J Med Genet A* 143A: 1858–1867.
105. Slavotinek AM (2008) Novel microdeletion syndromes detected by chromosome microarrays. *Hum Genet* 124: 1–17.
106. Bonaglia MC, Marelli S, Novara F, Commodaro S, Borgatti R, et al. (2010) Genotype-phenotype relationship in three cases with overlapping 19p13.12 microdeletions. *Eur J Hum Genet* 18: 1302–1309.
107. Hayashi S, Mizuno S, Migita O, Okuyama T, Makita Y, et al. (2008) The *CASK* gene harbored in a deletion detected by array-CGH as a potential candidate for a gene causative of X-linked dominant mental retardation. *Am J Med Genet* 146A: 2145–2151.
108. Reyniers E, Van Bogaert P, Peeters N, Vits L, Pauly F, et al. (1999) A new neurological syndrome with mental retardation, choreoathetosis, and abnormal behavior maps to chromosome Xp11. *Am J Hum Genet* 65: 1406–1412.
109. Al-Mateen M, Philippart M, Shields WD (1986) Rett syndrome. A commonly overlooked progressive encephalopathy in girls. *Am J Dis Child* 140: 761–765.
110. Zhang X, Liu Q, Chen B, Guo C, Li J, et al. (2004) A locus for nonspecific X-linked mental retardation mapped to a 22.3 cM region of Xp11.3-q22.3. *Am J Med Genet* 129A: 286–289.
111. Good CD, Lawrence K, Thomas NS, Price CJ, Ashburner J, et al. (2003) Dosage-sensitive X-linked locus influences the development of amygdala and orbitofrontal cortex, and fear recognition in humans. *Brain* 126: 2431–2446.
112. Nijhuis-van der Sanden MW, Eling PA, Otten BJ (2003) A review of neuropsychological and motor studies in Turner Syndrome. *Neurosci Biobehav Rev* 27: 329–338.
113. Hong D, Scaletta Kent J, Kesler S (2009) Cognitive profile of Turner syndrome. *Dev Disabil Res Rev* 15: 270–278.
114. Cresswell CS, Skuse DH (1999) Autism in association with Turner syndrome: Genetic implications for male vulnerability to pervasive developmental disorders. *Neurocase* 5: 511–518.
115. Mosconi MW, Cody-Hazlett H, Poe MD, Gerig G, Gimpel-Smith R, et al. (2009) Longitudinal study of amygdala volume and joint attention in 2- to 4-year-old children with autism. *Arch Gen Psychiatry* 66: 509–516.
116. Schumann CM, Barnes CC, Lord C, Courchesne E (2009) Amygdala enlargement in toddlers with autism related to severity of social and communication impairments. *Biol Psychiatry* 66: 942–949.
117. Groen W, Teluij M, Buitelaar J, Tendolkar I (2010) Amygdala and hippocampus enlargement during adolescence in autism. *J Am Acad Child Adolesc Psychiatry* 49: 552–560.
118. Ortiz-Mantilla S, Choe MS, Flax J, Grant PE, Benasich AA (2010) Associations between the size of the amygdala in infancy and language abilities during the preschool years in normally developing children. *Neuroimage* 49: 2791–2799.
119. Verhoeven JS, De Cock P, Lagae L, Sunaert S (2010) Neuroimaging of autism. *Neuroradiology* 52: 3–14.
120. Goldstein JM, Seidman LJ, Horton NJ, Makris N, Kennedy DN, et al. (2001) Normal sexual dimorphism of the adult human brain assessed by in vivo magnetic resonance imaging. *Cereb Cortex* 11: 490–497.
121. Yang F, Babak T, Shendure J, Distcheu CM (2010) Global survey of escape from X inactivation by RNA-sequencing in mouse. *Genome Res* 20: 614–622.
122. Jensen LR, Lenzner S, Moser B, Freude K, Tzschach A, et al. (2007) X-linked mental retardation: a comprehensive molecular screen of 47 candidate genes from a 7.4 Mb interval in Xp11. *Eur J Hum Genet* 15: 68–75.
123. Aziz A, Harrop SP, Bishop NE (2010) Characterization of the deleted in autism 1 protein family: implications for studying cognitive disorders. *PLoS One*. In press.
124. Alvarado M, Bocian M, Walker AP (1987) Interstitial deletion of the long arm of chromosome 3: case report, review, and definition of a phenotype. *Am J Med Genet* 27: 781–786.
125. Ko WT, Lam WF, Lo FM, Chan WK, Lam TS (2003) Wisconsin syndrome in a patient with interstitial deletion of the long arm of chromosome 3: further delineation of the phenotype. *Am J Med Genet* 120A: 413–417.

126. Zweier C, Guth S, Schulte-Mattler U, Rauch A, Trautmann U (2005) 9 Mb deletion including chromosome band 3q24 associated with unsuspected facial gestalt, persistent ductus omphaloentericus, mild mental retardation and tic. *Eur J Med Genet* 48: 360–362.
127. Jedele KB (2007) The overlapping spectrum of Rett and Angelman syndromes: a clinical review. *Semin Pediatr Neurol* 14: 108–117.
128. Santos M, Temudo T, Kay T, Carrilho I, Medeira A, et al. (2009) Mutations in the *MECP2* gene are not a major cause of Rett syndrome-like or related neurodevelopmental phenotype in male patients. *J Child Neurol* 24: 49–55.
129. Robin NH, Magnusson M, McDonald-McGinn D, Zackai EH, Spinner NB (1993) De novo interstitial deletion of the long arm of chromosome 3: 46,XX, del(3)(q25.1q26.1). *Clin Genet* 44: 335–337.
130. Slavotinek AM, Huson SM, Fitchett M (1997) Interstitial deletion of band 3q25. *J Med Genet* 34: 430–432.
131. Yen KH, Ho CL, Lee C (2009) The analysis of inconsistencies between cytogenetic annotations and sequence mapping by defining the imprecision zones of cytogenetic banding. *Bioinformatics* 25: 845–852.
132. Pickett J, London E (2005) The neuropathology of autism: a review. *J Neuropathol Exp Neurol* 64: 925–935.
133. Ramocki MB, Zoghbi HY (2008) Failure of neuronal homeostasis results in common neuropsychiatric phenotypes. *Nature* 455: 912–918.
134. Bourgeron T (2009) A synaptic trek to autism. *Curr Opin Neurobiol* 19: 231–234.
135. Sbacchi S, Acquadro F, Calò I, Cali F, Romano V (2010) Functional annotation of genes overlapping copy number variants in autistic patients: focus on axon pathfinding. *Curr Genomics* 11: 136–145.
136. Muhle R, Trentacoste SV, Rapin I (2004) The genetics of autism. *Pediatrics* 113: e472–486.
137. McDougle CJ, Erickson CA, Stigler KA, Posey DJ (2005) Neurochemistry in the pathophysiology of autism. *J Clin Psychiatry* 66: S9–18.
138. Castellani ML, Conti CM, Kempuraj DJ, Salini V, Vecchiet J, et al. (2009) Autism and immunity: revisited study. *Int J Immunopathol Pharmacol* 22: 15–19.
139. Coury D, Jones NE, Klatka K, Winklosky B, Perrin JM (2009) Healthcare for children with autism: the Autism Treatment Network. *Curr Opin Pediatr* 21: 828–832.
140. Tang BL (2008) Emerging aspects of membrane traffic in neuronal dendrite growth. *Biochim Biophys Acta* 1783: 169–176.
141. Hanus C, Ehlers MD (2008) Secretory outposts for the local processing of membrane cargo in neuronal dendrites. *Traffic* 9: 1437–1445.
142. Horton AC, Ehlers MD (2004) Secretory trafficking in neuronal dendrites. *Nat Cell Biol* 6: 585–591.
143. Nelson KB, Grether JK, Croen LA, Dambrosia JM, Dickens BF, et al. (2001) Neuropeptides and neurotrophins in neonatal blood of children with autism or mental retardation. *Ann Neurol* 49: 597–606.
144. Miyazaki K, Narita N, Sakuta R, Miyahara T, Naruse H, et al. (2004) Serum neurotrophin concentrations in autism and mental retardation: a pilot study. *Brain Dev* 26: 292–295.
145. Sadakata T, Kakegawa W, Mizoguchi A, Washida M, Katoh-Semba R, et al. (2007) Impaired cerebellar development and function in mice lacking *CAPS2*, a protein involved in neurotrophin release. *J Neurosci* 27: 2472–2482.
146. Chapleau CA, Larimore JL, Theibert A, Pozzo-Miller L (2009) Modulation of dendritic spine development and plasticity by BDNF and vesicular trafficking: fundamental roles in neurodevelopmental disorders associated with mental retardation and autism. *J Neurodev Disord* 1: 185–196.
147. Sadakata T, Furuichi T (2009) Developmentally regulated Ca^{2+} -dependent activator protein for secretion 2 (*CAPS2*) is involved in BDNF secretion and is associated with autism susceptibility. *Cerebellum* 8: 312–322.
148. Sadakata T, Furuichi T (2010) Ca^{2+} -dependent activator protein for secretion 2 and autistic-like phenotypes. *Neurosci Res* 67: 197–202.
149. Castermans D, Volders K, Crepel A, Backx L, De Vos R, et al. (2010) *SCAMP5*, *NBEA* and *AMISYN*: three candidate genes for autism involved in secretion of large dense-core vesicles. *Hum Mol Genet* 19: 1368–1378.
150. Piven J, Tsai GC, Nehme E, Coyle JT, Chase GA, et al. (1991) Platelet serotonin, a possible marker for familial autism. *J Autism Dev Disord* 21: 51–59.
151. Cook EH, Leventhal BL (1996) The serotonin system in autism. *Curr Opin Pediatr* 8: 348–354.
152. Castagnola M, Messana I, Inzitari R, Fanali C, Cabras T, et al. (2008) Hypophosphorylation of salivary peptidome as a clue to the molecular pathogenesis of autism spectrum disorders. *J Proteome Res* 7: 5327–5332.
153. Hendershot LM, Ting J, Lee AS (1988) Identity of the immunoglobulin heavy-chain-binding protein with the 78,000-dalton glucose-regulated protein and the role of posttranslational modifications in its binding function. *Mol Cell Biol* 8: 4250–4256.
154. Bishop L, Dimaline R, Blackmore C, Deavall D, Dockray GJ, et al. (1998) Modulation of the cleavage of the gastrin precursor by prohormone phosphorylation. *Gastroenterology* 115: 1154–1162.
155. Scott DB, Blanpied TA, Ehlers MD (2003) Coordinated PKA and PKC phosphorylation suppresses RXR-mediated ER retention and regulates the surface delivery of NMDA receptors. *Neuropharmacology* 45: 755–767.
156. Lee SN, Hwang JR, Lindberg I (2006) Neuroendocrine protein 7B2 can be inactivated by phosphorylation within the secretory pathway. *J Biol Chem* 281: 3312–3320.
157. McMahon HE, Sharma S, Shimasaki S (2008) Phosphorylation of bone morphogenetic protein-15 and growth and differentiation factor-9 plays a critical role in determining agonistic or antagonistic functions. *Endocrinology* 149: 812–817.
158. Waring R, Phoenix J (2001) TPST-assay for diagnosis of autism and related disorders. SHS International Ltd: International patent WO/2001/077681.
159. Waring R (2001) Sulphation and autism – what are the links? Available: www.autismfile.com/papers/Rosemary_Waring_Sulphation.asp. Accessed April 2010.
160. Stone MJ, Chuang S, Hou X, Shoham M, Zhu JZ (2009) Tyrosine sulfation: an increasingly recognised post-translational modification of secreted proteins. *N Biotechnol* 25: 299–317.
161. Kehoe JW, Bertozzi CR (2000) Tyrosine sulfation: a modulator of extracellular protein-protein interactions. *Chem Biol* 7: R57–61.
162. van der Zwaag B, Franke L, Poot M, Hochstenbach R, Spierenburg HA, et al. (2009) Gene-network analysis identifies susceptibility genes related to glycobiology in autism. *PLoS One* 4: e5324.
163. Nilsson T, Au CE, Bergeron JJ (2009) Sorting out glycosylation enzymes in the Golgi apparatus. *FEBS Lett* 583: 3764–3769.
164. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTALX windows interface: flexible strategies for multiple sequence alignments aided by quality analysis tools. *Nucleic Acids Res* 25: 4876–4882.
165. Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, et al. (1999) Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol* 112: 531–552.
166. Hofmann K, Stoffel W (1993) TMbase - A database of membrane spanning proteins segments. *Biol Chem Hoppe-Seyler* 374: 166–168.
167. Milpetz F, Argos P, Persson B (1995) TMAP: a new email and WWW service for membrane-protein structural predictions. *Trends Biochem Sci* 20: 204–205.
168. Tusnady GE, Simon I (2001) The HMMTOP transmembrane topology prediction server. *Bioinformatics* 17: 849–850.
169. Kanehisa M (1997) Linking databases and organisms: GenomeNet resources in Japan. *Trends Biochem Sci* 22: 442–444.
170. Horton P, Nakai K (1997) Better prediction of protein cellular localization sites with the k nearest neighbors classifier. *Proc Int Conf Intell Syst Mol Biol* 5: 147–152.
171. Marchler-Bauer A, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, et al. (2009) CDD: specific functional annotation with the Conserved Domain Database. *Nucleic Acids Res* 37: D205–210.
172. Gould CM, Diella F, Via A, Puntervoll P, Gemünd C, et al. (2010) ELM: the status of the 2010 eukaryotic linear motif resource. *Nucleic Acids Res* 38: D167–180.