

Retinoic Acid Induced 1, *RAI1*: A Dosage Sensitive Gene Related to Neurobehavioral Alterations Including Autistic Behavior

Paulina Carmona-Mora and Katherina Walz*

John P. Hussman Institute for Human Genomics, Dr. John T. Macdonald Foundation, Department of Human Genetics, Miller School of Medicine, University of Miami, Miami, Florida, USA

Abstract: Genomic structural changes, such as gene Copy Number Variations (CNVs) are extremely abundant in the human genome. An enormous effort is currently ongoing to recognize and catalogue human CNVs and their associations with abnormal phenotypic outcomes. Recently, several reports related neuropsychiatric diseases (i.e. autism spectrum disorders, schizophrenia, mental retardation, behavioral problems, epilepsy) with specific CNV. Moreover, for some conditions, both the deletion and duplication of the same genomic segment are related to the phenotype. Syndromes associated with CNVs (microdeletion and microduplication) have long been known to display specific neurobehavioral traits. It is important to note that not every gene is susceptible to gene dosage changes and there are only a few dosage sensitive genes. Smith-Magenis (SMS) and Potocki-Lupski (PTLS) syndromes are associated with a reciprocal microdeletion and microduplication within chromosome 17p11.2. in humans. The dosage sensitive gene responsible for most phenotypes in SMS has been identified: the Retinoic Acid Induced 1 (*RAI1*). Studies on mouse models and humans suggest that *RAI1* is likely the dosage sensitive gene responsible for clinical features in PTLS. In addition, the human *RAI1* gene has been implicated in several neurobehavioral traits as spinocerebellar ataxia (SCA2), schizophrenia and non syndromic autism. In this review we discuss the evidence of *RAI1* as a dosage sensitive gene, its relationship with different neurobehavioral traits, gene structure and mutations, and what is known about its molecular and cellular function, as a first step in the elucidation of the mechanisms that relate dosage sensitive genes with abnormal neurobehavioral outcomes.

Received on: September 19, 2010 - Revised on: October 08, 2010 - Accepted on: October 21, 2010

Keywords: Copy Number Variation, dosage sensitive gene, neurobehavioral traits, Potocki-Lupski Syndrome, *RAI1*, Smith-Magenis Syndrome, transcription factor activity.

GENOMIC STRUCTURAL CHANGES AND NEUROPSYCHIATRIC TRAITS

Genomic structural changes, such as gene Copy Number Variations (CNVs) are extremely abundant in the human genome. Approximately 14,400 CNV loci larger than 1 kb were identified to date (<http://projects.tcag.ca/variation/>) [1]. In a recent article the impressive amount of more than 1,000 CNV/individual human genome ranging from 443 bp to 1.28 megabases (Mb), with a median size of 2.9 kb was reported [2]. On the other hand, an enormous effort is currently ongoing to recognize and catalogue human CNVs associated with abnormal phenotypic outcomes [3-7]. Some conditions affecting the central nervous system, such as schizophrenia, mental retardation (MR) and autism spectrum disorder (ASD) were shown recently to be associated with both the deletion and the duplication of the same genomic segment [8-13]. There is a great level of complexity between the presence of CNV and the resulting phenotypes that is not only a direct consequence of specific altered gene dosage. Analyses of the genome-wide functional impact of these structural variants showed that copy number changes not

only cause alterations in expression levels of genes within CNVs but also influence the expression of genes in their vicinity [14, 15]. Moreover, a recent study clearly demonstrates that the presence of structural changes associated with CNV is enough to cause a phenotype independent of gene dosage [16].

Genomic disorders are the clinical manifestation of pathological CNV, they are frequent conditions (~ 1 per 1,000 births) and often sporadic resulting from *de novo* rearrangements [17]. In a subset of such conditions the rearrangements comprise multiple unrelated contiguous genes that are physically linked and thus have been referred to as Contiguous Gene Syndromes (CGS). An increasing number of CGS are being described each of them presenting a complex and specific phenotype. The beauty of these CGS is that there are specific phenotypes associated with certain genomic regions, thus making the detection of the genes involved in certain pathways easy. Although the segmental aneuploidy usually involves several genes, only a small subset of them conveys phenotypes as a function of copy number alteration. These particular genes are referred as “dosage sensitive genes”. However, the question of what makes a gene or its product a “dosage sensitive gene” remains to be elucidated.

Smith-Magenis (SMS, OMIM# 182290) and Potocki-Lupski (PTLS, OMIM# 610883) syndromes are two examples of CGS that are associated with a microdeletion and a

*Address correspondence to this author at the Miller School of Medicine, University of Miami, 1501 NW 10th Avenue, Room 418, Miami, FL 33136, USA; Tel: 305-243-1116; Fax: 305-243-2396; E-mail: kwalz@med.miami.edu

microduplication respectively within chromosome 17 band p11.2. SMS was first described in 1986, and has a birth prevalence estimated at 1/15,000-1/25,000 [18-20]. The clinical phenotype includes craniofacial abnormalities, brachydactyly, self-injurious behavior, sleep abnormalities and mental retardation. Less commonly reported is cleft palate, congenital heart defects, seizures, hearing impairment and urinary tract anomalies. Molecular studies revealed a common deleted region of ~4 Mb in the majority of SMS patients (>70-80%) [21]. Unusual sized deletions (smaller or larger) were observed in 20–25% of patients. By examining the breakpoints in unusual sized deletions, the SMS critical region was redefined to a ~950 kb interval, in which fifteen genes and eight predicted genes were present [22] (Table 1). PTLs was first described in 2000 [23] and is associated with the reciprocal duplication of the genomic interval deleted in SMS. The clinical presentation of PTLs includes mental retardation, autistic features, hyperactivity, developmental delay and dysmorphic features. In addition, PTLs patients present hypotonia, poor feeding and difficulty to thrive in infancy, oral-pharyngeal dysphasia, obstructive and central sleep apnea, structural cardiovascular abnormalities, electroencephalogram (EEG) abnormalities, behavioral abnormalities, and hypermetropia [24-26] (Table 1).

RAII IS THE DOSAGE SENSITIVE GENE MOSTLY RESPONSIBLE FOR THE SMS AND PTLs CLINICAL PRESENTATION

As mentioned above, only a few genes do appear to be clearly haploinsufficient, such that hemizygotes exhibit a recognizable phenotype. Point mutations (nonsense and

frameshift as well as missense alleles) in the Retinoic Acid Induced 1 gene (*RAII*) were identified in patients with clinical presentation of SMS but no molecular deletion found by FISH [30-33] (Tables 1 and 2), suggesting that *RAII* is the dosage sensitive gene causative of SMS. In specific, there are two cases of patients with SMS that carry missense mutations that alter amino acids within a highly conserved region of the gene (A4685G and G5423A). Both patients have a wide spectrum of neurobehavioral phenotypes, including sleep disturbance, self-injurious behaviors, mental retardation and developmental delay. This supports the fact that *RAII* is a protein directly involved in these specific neurobehavioral phenotypes, since the point mutations mentioned above generate a full length protein with only the change of one single amino acid, with stability similar to the wild type protein [34].

In addition, a recent report showed a patient with PTLs phenotype having a duplication of a very small region that only contains the *RAII* gene [29]. This is consistent with the notion of *RAII* as the dosage sensitive gene within this genomic region (Table 1).

Additional support to the idea that *RAII* is the dosage sensitive gene within this interval comes from studies in mouse models. Human chromosome 17p11.2 is syntenic to the 32-34 cM region of murine chromosome 11. The number and order of the genes are highly conserved [36-38]. By chromosome engineering, a strategy developed to introduce defined chromosomal rearrangements into the mouse genome to delineate gene haploinsufficiency effects (haploid genetics) [39-43], a deletion and reciprocal duplication in the mouse chromosome region syntenic to the SMS critical in-

Table 1. Key Neurobehavioral Features Found in SMS and PTLs Patients

Neurobehavioral features	SMS		PTLS	
	% in common deletion	% in <i>RAII</i> mutations	% in common duplication	% in small duplications
Variable mental retardation or cognitive impairment	100	100	100	100
Motor and speech delay	>90	70	100	100
Sleep disturbance	90	100	66.6	75
Infantile hypotonia	>90	61	88.8	75
Overt seizures	11–30	16.6	NR	NR
EEG abnormalities	49	NR	66.6-70	5.8
Attention seeking	80–100	100	NR	NR
Onychotillomania	25–85	100	NR	NR
Polyembolokoilomania	25–85	80	NR	NR
Hand biting/self-biting	80	60	NR	NR
Self-injurious behaviors	70–90	100	NR	NR
Self-hugging/hand wringing	50–80	100	NR	NR
Head banging/face slapping	70	90	NR	NR
Self-picking	54	100	NR	NR
Hyperactive	80	100	>70	66.6
Autistic features	NR	NR	87.5	66.6

A comparison between the phenotypes present in patients with the genomic rearrangements and point mutations of *RAII* gene is shown. For SMS, the clinical features are presented in the percentage of patients carrying either the common 17p11.2 deletion (~3.7 Mb) or a mutation in *RAII* gene. The percentages of features in PTLs are represented for patients that harbor the common 17p11.2 duplication (~3.7 Mb) or small duplications (<1 Mb). (NR: not reported) [25, 27-29].

Table 2. *RAII* Gene Variations Related with Neurobehavioral Phenotypes

Mutation	Exon	Resultant protein	Mental retardation	Sleep disturbance	Speech delay	Motor delay	Self injurious behavior	Developmental delay	Hypotonia
³² 253del19	3	Missincorporation of 60 aa	+	+	+	+	+	NR	+
³⁵ 1119delC	3	Missincorporation of 65 aa	+	+	NR	NR	+	NR	NR
³⁰ 1449delC	3	Missincorporation of 34 aa	+	+	NR	+	+	NR	+
³⁰ 2773del29	3	Missincorporation of 8 aa	+	+	+	NR	+	NR	NR
³¹ C2878T	3	R960X	+	+	NR	NR	+	NR	NR
³¹ 3103insC	3	Missincorporation of 30 aa	+	+	NR	NR	+	NR	NR
³³ 3103delC	3	Missincorporation of 28 aa	+	+	+	NR	+	NR	NR
³² 3801delC	3	Missincorporation of 46 aa	+	+	NR	+	+	NR	NR
³⁵ 4649delC	3	Missincorporation of 36 aa	+	+	NR	NR	+	NR	NR
³² A4685G	3	Q1562R	+	+	+	+	+	NR	+
³⁵ 4933delGCCG	3	Missincorporation of 35 aa	+	+	NR	NR	+	+	NR
³⁰ 5265delC	3	Missincorporation of 74 aa	+	+	NR	NR	+	NR	NR
³² G5423A	3	S1808N	+	+	+	NR	+	NR	+
³³ A3634G	3	S1212G							
³³ (PolyQ) ₁₈	3	(PolyQ) ₁₈	+	+	+		+	+	+

The mutations found in SMS patients are listed with their position in the cDNA, the protein that they generate and the neurobehavioral phenotypes associated with each mutation (+ means the presence of the characteristic, NR: not reported) [30-33, 35].

terval were generated [44]. The mouse model for SMS, *Df(11)17/+* or *Del*, carries a ~1.7 Mb deletion and presents craniofacial abnormalities, seizures, obesity, alterations in circadian rhythm and hypoactivity, decreased anxiety in the plus maze test, social interaction differences in the tube test and the social novelty test recapitulating most of the phenotypes observed in human patients [15, 44-47]. The mouse model for PTLs syndrome, *Dp(11)17/+* or *Dup*, carries a duplication of a region of ~1.7 Mb and are significantly leaner than wild type mice, hyperactive and exhibit impaired contextual fear conditioning, they display increased anxiety behavior in the plus maze test and increased dominance in the tube test. Abnormal sociability and preference for social novelty was also observed [15, 44, 45].

Interestingly enough, most of the phenotypes present in *Dup* mice were rescued in the *Dup/Rail^{-/-}* compound heterozygote (obtained by mating *Dup* mice with *Rail^{+/-}* mice, a model with *Rail* haploinsufficiency) where the gene dosage of *Rail* was 2n despite the presence of 3 copies of the other 23 genes in the region; being the first indication of *Rail* as the dosage sensitive gene responsible the PTLs syndrome [50]. Consistent with this, there have been reports of mouse models with underexpression of *Rail* [48] and overexpression of *Rail* [51] that presented several of the expected phe-

notypes. For the last one, analyses in both hemizygous and homozygous transgenic mice revealed a dosage-dependent exacerbation of the phenotype, which included growth retardation, increased anxiety and hyperactivity [51]. A summary can be seen in Table 3.

Utilizing the *Del* and *Dup* mouse models, the levels of expression of *Rail* and the other genes within this specific genomic rearrangement were studied. Genome-wide expression levels in four organs (cerebellum, heart, kidney and testis) from adult male mice (at least three animals of *Dup*, *Del*, *Dup/Del* and wild type mice) showed a positive correlation between gene dosage and expression levels [15, 16]. The transcripts within the rearranged interval are expressed on average 0.66 (SD=0.15) fold less in the *Del* (one copy) and 1.38 (SD=0.29) fold more in *Dup* animals (three copies). They were, however, unchanged in *Del/Dup* mice (1.02 (SD=0.16) fold more, two copies in *cis*) compared to normal controls (two copies in *trans*) [16]. This clearly showed that *Rail* gene copy number is correlated with levels of expression of the gene, and suggests that the protein levels are modified concomitant with gene dosage. Interestingly enough, most of the genes within this genomic interval behave in the same way, suggesting that is not the levels of

Table 3. *Rai1* CNV is Sufficient to Cause Most of the Phenotypes of SMS and PTLs also in Mice

Mouse model Phenotype	<i>Del</i>	<i>Rai1</i> ^{+/-}	<i>Dup</i>	<i>Dup/Rai1</i>	<i>Tg Rai1</i>
Overt seizures	Present (20%)	Subtle	No present	No present	No present
EEG	Abnormal	Abnormal	Normal	ND	ND
Locomotor activity	Hypoactive	Abnormal	Hyperactive	Hyperactive	Hyperactive
Anxiety	Decreased	Abnormal	Increased	Normal	Increased
Learning and memory	Normal	Normal	Impaired	Normal	ND
Social novelty recognition	Abnormal	ND	Impaired	ND	Abnormal
Dominance like behavior	Decreased	ND	Increased	ND	Increased
Reference	[16, 44, 45]	[48-50]	[15, 16, 44, 45, 50]	[50]	[51, 52]

Summary of the neurobehavioral phenotypes observed in existing mouse models for SMS and PTLs. *Del* and *Rai1*^{+/-} mice carry one gene copy of *Rai1*, *Dup* model carries three copies, *Tg Rai1* has several copies and *Dup/Rai1* has two copies or *Rai1* gene. (N/D: not determined).

expression that make the difference between *Rai1* and the other 23 genes in the region.

MOLECULAR ASPECTS OF RAI1

RAI1 lies in the middle of the Smith-Magenis critical region, spanning over 120 kb. It consists of six exons, being the third exon the one that contains more than 90% of the coding region.

As is shown in Fig. (1), there are several regions recognizable in RAI1 protein structure. A polymorphic CAG repeat encoding for a polyglutamine tract is present in the N-terminus of the RAI1 protein. Two polyserine tracts, each located at the N and C-terminal regions. The polyserine signal is similar to the one present in the *DRPLA* gene and the *Drosophila* hairless gene, both proteins involved in neuronal development [53, 54]. RAI1 has two putative bipartite nuclear localization signals (NLSs) *in silico* predicted. It also possesses a zinc finger like plant homeo domain (PHD) in the C-terminus, which is present in the trithorax family of chromatin remodeling transcriptional regulators [30, 31]. The isoform 1 of *RAI1* corresponds to the canonical sequence, and it is found in high levels mainly in brain and heart [55]. There are other three isoforms produced by alternative splicing which are described in the online database Universal Protein Resource (www.UniProt.org). These isoforms differ from the canonical one by the change or lack of several residues; the schematic representations of these structures are represented in Fig. (1). Isoform 2 is the most similar to the canonical, with a length of 1862 amino acids and domain conservation [53]. In the isoform 3 (1640 amino acids) the second polyserine tract and PHD domain are missing [56]. Isoform 4 is the shortest, containing only the first half of RAI1 protein (966 amino acids) [57] not encompassing NLSs and being comparable to the truncated protein found in some SMS patients. This short isoform has been found in muscle, while the other proteins have been described from brain tissue. In addition to the isoforms 2, 3 and 4 of RAI1,

the online database of eukaryotic genomes, Ensembl project (www.ensembl.org) reports other two transcripts (ENST00000395776 and ENST00000428973), also generated by alternative splicing and both produce the same RAI1 protein of 1993 amino acids (ENSP00000379122). This protein shares a 100% homology with the isoform 1 until the residue 1855. It is important to note that in the same database, several transcripts of RAI1 were found, and these are consistent with the records of the other variants reported in the databases of UniProt and UCSC Genome Browser (<http://genome.ucsc.edu/>).

Besides the domains described above for the protein, there are several post-translational modifications (PTM) sites found in RAI1, both *in silico* and experimentally. Some experimental evidence of PTM present in the RAI1 protein comes from *in vitro* studies where the isoform 1 and five RAI1 mutated forms of the protein (*RAI1-HA* 2687delC, *RAI1-HA* 3103delC, RAI1 R960X, RAI1-HA Q1562R, and RAI1-HA S1808N, four of them associated with SMS clinical phenotype) were studied. In this report all proteins presented a higher molecular weight than expected when run in an SDS-PAGE gel analysis. Moreover, the PTM were mapped to the N-terminal of RAI1 [34]. Through the analysis of RAI1 sequence with NetGlycate 1.0, NetOGlyc 3.1 and YinOYang 1.2 servers [58-60], we found several predicted glycation sites homogeneously located within the protein sequence and more interestingly, there are several O-glycosylation sites with high score, most of them mapping in the first amino-terminal half of RAI1. There are also O-linked β -N-acetylglucosamine (O-GlcNAc) predicted sites with high score, which are present in many transcription factors and neuronal proteins; this modification participates in transcription, translation, proteasomal degradation and apoptosis among others and is directly involved in the regulation of neuronal development and synaptic transmission [61].

Phosphorylation is a well established PTM, implicated in protein regulation in a variety of cellular processes. We ran

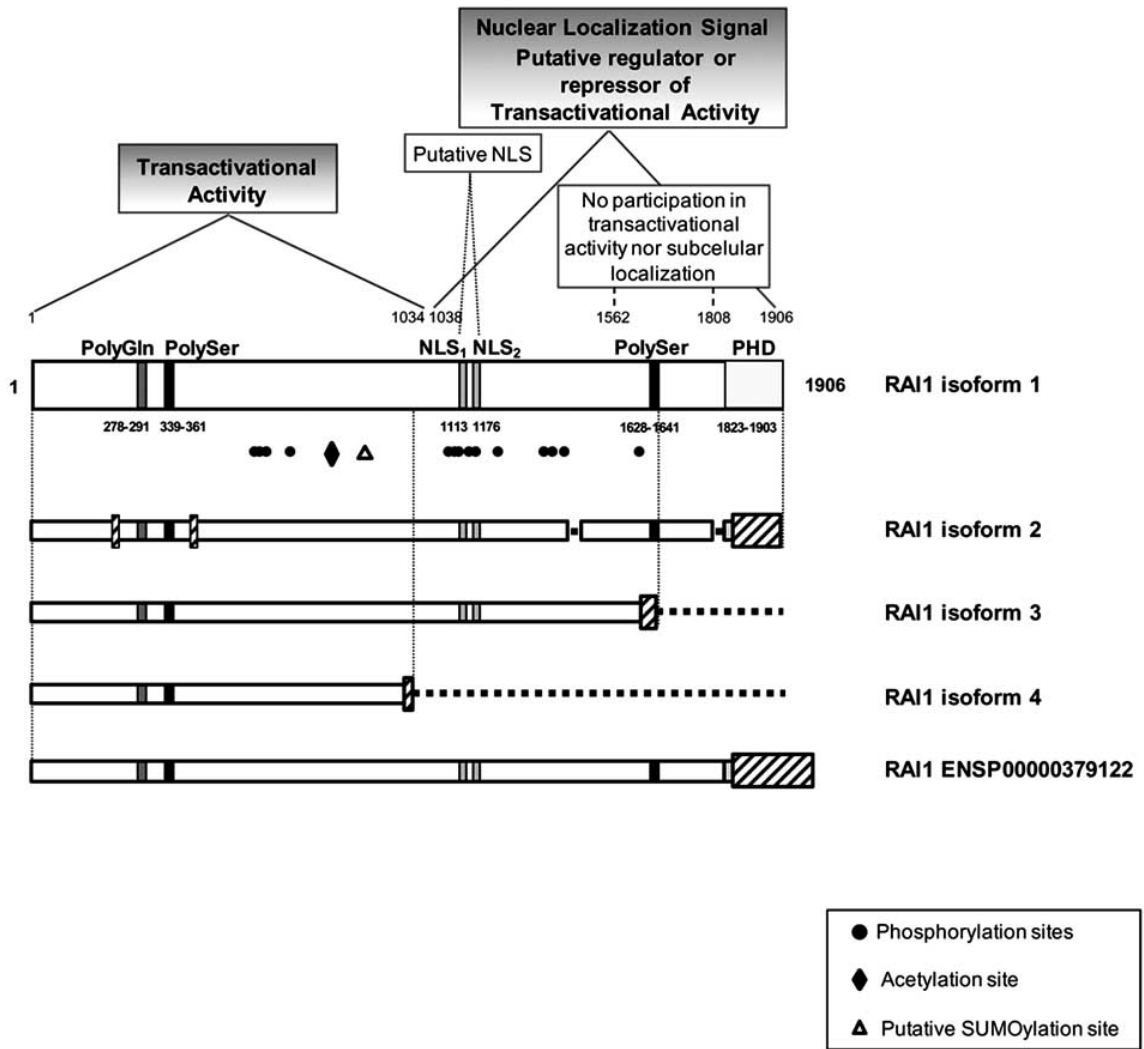


Fig. (1). Structure of RAI1. By *in silico* analyses, several domains have been found for RAI1: a polyglutamine tract at the N-terminal of the protein, two polyserine domains, a PHD domain at the C-terminal of RAI1 and two putative nuclear localization signals (NLS). RAI1 protein structure is schematically represented, and the domains described are showed with their localization in the amino acidic sequence. By the evaluation of the wild type and mutant proteins associated to SMS, two main domains were found: the N-terminal half of RAI1 is responsible for the transactivational activity and C-terminal half beginning in residue 1038 is responsible for nuclear localization [34]. The RAI1 PTMs found are depicted with black figures while the putative sites are represented by white figures. The phosphorylation sites (1Thr and 13Ser) found in HeLa cells are represented with black circles. The reported acetylation site is depicted with a black diamond. The putative SUMOylation site with the highest score is represented with a white triangle.

Below, the other four isoforms described for human RAI1 are represented. These isoforms differ from the canonical one by the change of several residues (depicted in slanted rectangles) and also by some fragments missed (depicted by discontinued lines). Isoform 2 is the most similar to the canonical, while in isoform 3 are missing the second polyserine tract and PHD domain. Isoform 4 is the shortest containing only the first half of RAI1 protein, not encompassing NLSs and being comparable to the truncated proteins found in some SMS patients. The longest variant of RAI1 (ENSP00000379122) is a protein reported by the online database of eukaryotic genomes, Ensembl, which has a length of 1993 amino acids and has complete homology with the isoform 1 until amino acid 1855.

an *in silico* analysis of putative phosphorylation sites for RAI1 (NetPhos 2.0 Server [62]), and found the number of putative phosphorylatable amino acid/total as follows: Ser128/217, Thr36/99, and Tyr13/39, indicating the existence of several possible phosphorylation sites. Consistent with these predictions, in recent studies to assess the global *in vivo* phosphoproteome one threonine (Thr1186) and thirteen serines (Ser568, 571, 574, 687, 1107, 1108, 1110, 1121, 1122, 1352, 1358, 1374 and 1616) were identified in a phosphorylated state in the RAI1 protein, and three of them

(Ser568, Ser571 and Ser574) were regulated upon EGF stimulation in HeLa cells [63-66]. Interestingly, most of these sites are consensus for CK1, GSK3, NEK6 and ERK kinases.

Lysine acetylation is a reversible PTM, which neutralizes the positive charge of this amino acid, changing protein function. This PTM is very important in the regulation of gene expression through the modification of histone proteins. Site-specific acetylation of a growing number of non-histone proteins has been shown to regulate their activity, localiza-

tion, specific interactions as well as stability/degradation, therefore controlling a variety of cellular processes, such as transcription, proliferation, apoptosis and differentiation [67]. Interestingly, in another large scale analysis [68] an acetylation modification was reported for RAI1 at the peptide KGLEQGGKASDGIISK in the Lysine 774 (underlined), which also presented a high accessibility rate in the 3D structure.

SUMOylation (Small Ubiquitin-like Modifier proteins) is a post-translational modification involved in various cellular processes, such as nuclear-cytosolic transport, transcriptional regulation, apoptosis, protein stability, response to stress, and progression through the cell cycle [69, 70]. *In silico* analysis (SUMOsp 2.0 [71]) of the RAI1 protein sequence gave seven putative SUMOylation sites of the TypeI: Ψ-K-X-E, in Lys549, 811, 819, 988, 1328, 1425 and 1500 (being Lys819 the most probable one). Therefore, the prediction of several sites for PTM within RAI1 sequence and the difference between the obtained vs. expected molecular weight of the protein, are consistent with a PTM in the N-terminal domain of RAI1 (Fig. (1)).

RAI1 MOLECULAR AND CELLULAR FUNCTION

Little is known about the cellular and developmental role of RAI1. The first study was made in a mouse embryonic carcinoma cell line (P19) where the expression of GT1, a splice variant of *Rai1* was markedly up-regulated after the treatment with retinoic acid for inducing neuronal differentiation [72].

Initially it was suggested, based in bioinformatic analyses that RAI1 could function as a transcriptional regulator [30, 31]. RAI1 and TCF20, a transcriptional cofactor, share a homology of 50% at the amino acid level, and both genes have a similar structure. Furthermore, the presence of the polyglutamine stretch at the N-terminal of RAI1 is coincident with many transcription factors containing polyglutamine activation domains and it has been shown that these stretches can modulate the transcription activation [73]. Consistent with this speculation, it was reported that mouse *Rai1* has transactivational activity in HeLa cells but in moderate levels [48]. Considering it had milder activity than classical transcription factors, this can indicate that *Rai1* might not act alone but may be part of a complex that regulates transcription. In 2010, an independent study showed that mouse *Rai1* had much stronger transactivation activity in Neuro-2a cells (a neuronal derived cell line) than in HeLa cells [34]. These results suggest specificity in the transcription machinery that RAI1 could be part of and further indicate the importance of its role in neuronal derived tissues, since *RAI1* is a dosage sensitive gene involved in neurobehavioral phenotypes. Moreover, it was also demonstrated that human RAI1 has transcription factor activity [34] and is able to localize to the nucleus. Intriguingly, the two missense mutations (RAI1 Q1562R and RAI1 S1808N) were shown to localize to the nucleus and activate transcription of a reporter gene in the same magnitude than the wild type protein, strongly suggesting that these mutations are involved in other functions yet unknown for RAI1. The analysis of the mutations *RAI1-HA* 2687delC, *RAI1-HA* 3103delC and RAI1 R960X showed that they generated a truncated polypeptide with cytoplas-

matic localization and higher transcription factor activity than the wild type protein. In addition, two different C-terminal halves of RAI1 protein (1038 aa-end and 1229 aa-end) were able to localize into the nucleus but had no trans-activation activity. Taking all this into consideration plus the mouse protein studies previously reported [48], two functional domains were assigned to each half of RAI1: the N-terminal that contains transactivational activity and the C-terminal half of RAI1 responsible for the correct subcellular localization of the protein. In a normal situation the two functional domains of the protein are together. This suggests that the pathogenic outcome of some of the mutated forms may be explained by the dissociation of these two domains. However, in this context it is intriguing, and needs further study, the contribution of the different RAI1 isoforms to the final phenotype.

Therefore, it is clear that RAI1 is regulating gene expression and there are new data from cells and mice regarding to this aspect of the protein. A genome-wide gene expression study made in HEK293T (human embryonic kidney cells) with a 50% knockdown of *RAI1* generated by siRNA was performed using microarrays [74]. The top dysregulated genes found in these *RAI1* haploinsufficient cells are involved in neuronal differentiation, circadian activity and behavior; besides growth signaling, insulin sensitivity, gene expression, cell-cycle regulation, lipid biosynthesis, fat mobilization and skeletal, cardiovascular and renal development. The top candidate genes found upregulated due to *RAI1* haploinsufficiency are *ZIC1*, *PSEN2*, *RXRβ*, *CLN8* and *SCN12A* being involved in neurogenesis, mental retardation, sensory transmission, nociceptive behavior, epilepsy and neurological function. Amongst the top candidate genes that are downregulated there are *NF1*, *MLL* and *NR1D2*, which are associated to schizophrenia, neurofibromatosis type I and circadian activity [74].

Recently in the mouse model for *Rai1* haploinsufficiency, *Rai1*^{+/-}, several genes were reported to be dysregulated in the hypothalamus [75]. These genes include *Slc6a12* and *Slc38a5*, which are related to traits like epilepsy, anxiety disorder, mental retardation, whose expression is downregulated as a consequence of the haploinsufficiency of *Rai1*. Furthermore, it was found that human RAI1 isoform 1 directly regulates the expression of BDNF (brain-derived neurotrophic factor) [75] which is a neurotrophin that regulates neuronal survival, differentiation and growth during brain development, among others. The expression of this gene is known to be reduced in patients with Alzheimer, Parkinson and Huntington diseases [76]. Additionally, BDNF has been reported to be associated with obesity and hyperphagia [77, 78].

Moreover, altered gene expression was found for many genes in a hippocampus expression profile of the *Dup* mouse model for PTLs [15]. In this study the transcript of Ataxin 2, a protein mutated in spinocerebellar ataxia type 2 (SCA2), had a fold change of 1.24 by microarray analysis; certainly suggesting a link between its expression and the overexpression of *Rai1*.

Altogether, the studies showing that CNV of *RAI1* causes dysregulation of many genes involved in many different cellular processes, suggest that RAI1 functions through specific

pathways involved in the regulation of various biological processes, and when disrupted result in the phenotypic effects observed in SMS, PTLs and other neurobehavioral traits.

RAI1 HAS BEEN IMPLICATED IN SEVERAL NEUROBEHAVIORAL TRAITS

As mentioned before, the haploinsufficiency of *RAI1* causes SMS meanwhile the duplication of a small region containing this gene is able to cause PTLs. Besides the strong association of *RAI1* with these two genomic disorders, this protein has been reported to be related with other neurobehavioral traits.

At the N-terminus of the *RAI1* protein there is a polymorphic CAG repeat, which length is associated with the age of onset of spinocerebellar ataxia type 2 [79] and the response to neuroleptic medication in schizophrenia [80]. In addition, *RAI1* has been shown to be associated with non syndromic autism [81].

Spinocerebellar ataxia type 2 (SCA2) is an autosomal dominant disease caused by the increment of a CAG repeat

tract in ataxin-2 gene, generating an expanded polyglutamine tract in this protein. A study of the association between CAG repeat length and age at onset variability in SCA2 [79] in more than 100 patients found that *RAI1* locus contributes to 4.1% of the variance in SCA2 age at onset, after the effect caused by the expanded ataxin-2 protein. Therefore, this is the only association reporting *RAI1* as a contributor in a neurodegenerative disorder, and due to a larger polyglutamine tract in *RAI1* protein, which ranges from 10 to 16 glutamines in the normal population [33].

It was reported that a CAG repeat polymorphism in *RAI1* is associated with both the severity of the phenotype and the response to medication in schizophrenic patients [80]. Compared to non-responders, schizophrenic subjects responding to neuroleptic medication are characterized by a later age of onset and a better long-term outcome. The group responding to medication carried a significantly shorter CAG repeat size compared to control groups.

It is remarkable to mention that CAG repeat polymorphism of *RAI1* have not been found to be causative of the SMS phenotype; and the length of this CAG repeat has not

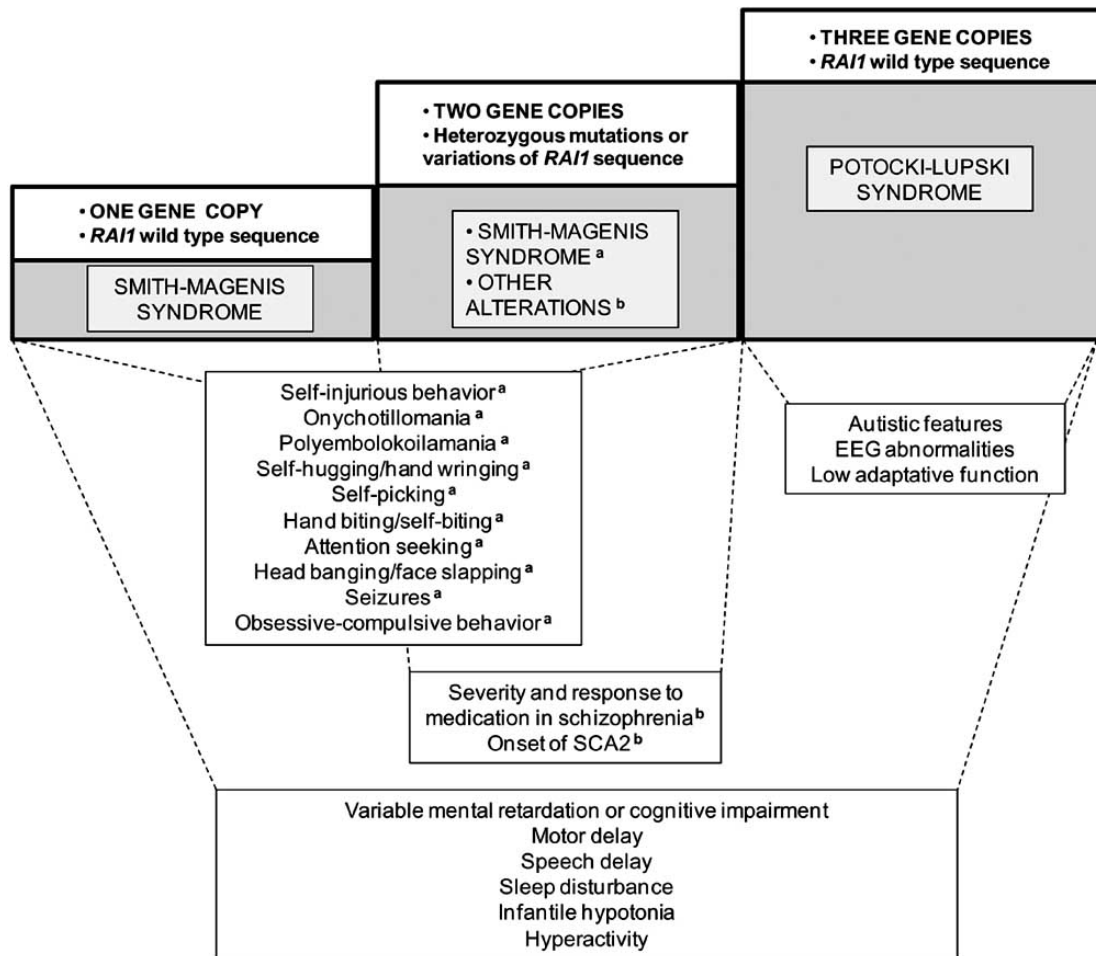


Fig. (2). *RAI1* changes associated with neurobehavioral phenotypes. The traits that have been found related to *RAI1* CNVs, mutations or sequence changes are summarized. There are phenotypes that are associated with only one copy of *RAI1* gene (SMS). Many of these phenotypes are also related to carrying two copies of the gene, but one allele with mutations (SMS caused by frameshifts, nonsense or missense mutations)^(a). Harboring two copies of *RAI1* gene but with polymorphic CAG repeats influence other neurobehavioral traits^(b). The alteration of *RAI1* dosage to three copies causes specific phenotypes of PTLs as well as features that overlap with some clinical characteristics of SMS.

been analyzed in PTLs patients. The most common polyglutamine tracts in normal and SMS patients are 13-14 Gln. The comparison of phenotypes in SMS deletion patients with different number of CAG repeats did not show any association between CAG repeat length and the severity of SMS phenotype [33]. Recently, in a study with a group of patients with non-complex-autism, *RAI1* was found to be one of the novel autism candidate genes within the autism-associated CNVs regions [81]. In this study, *RAI1* was the top most significant gene identified in the analysis of 173 non-overlapping CNV regions.

All of these data indicate that *RAI1* is participating in different aspects of behavior regulation. In Fig. (2) are summarized the neurobehavioral traits related with *RAI1* variations. In summary, there is plenty of evidence indicating that *RAI1* is the dosage sensitive gene responsible for most of the phenotypes of SMS and PTLs syndromes, including several neurobehavioral traits.

It is also possible to suggest that *RAI1* might have a role modulating some aspects of neuron function or differentiation since in addition to its contribution in both SMS and PTLs phenotypic outcome, it is highly expressed in neurons, its expression is influenced by retinoic acid and has polyglutamine polymorphisms that have been found to be associated with schizophrenia and spinocerebellar ataxia type 2. Taking into consideration the information known to date from studies *in vitro*, it can be hypothesized that at the cellular level the cause of the resulting phenotypes related to alterations in *RAI1* can be due to (either or together) an abnormal translocation of the protein, reduced transactivational activity or inability of interaction with other proteins and/or DNA due to changes in a specific motif or to modifications in the 3D structure.

However, further studies are necessary with *RAI1* and generally with all the other dosage sensitive genes known up to date in order to answer the main question that remains unanswered: what makes a gene, or its product, a dosage sensitive one?

ACKNOWLEDGEMENTS

We thank Irene Perez for English editing of the manuscript and Ana Lorenzo for administrative support. This work was supported in part by the Jérôme Lejeune Foundation.

REFERENCES

[1] Iafrate, A. J.; Feuk, L.; Rivera, M. N.; Listewnik, M. L.; Donahoe, P. K.; Qi, Y.; Scherer, S. W.; Lee, C. Detection of large-scale variation in the human genome. *Nat. Genet.*, **2004**, *36*, 949-951.

[2] Conrad, D. F.; Pinto, D.; Redon, R.; Feuk, L.; Gokcumen, O.; Zhang, Y.; Aerts, J.; Andrews, T. D.; Barnes, C.; Campbell, P.; Fitzgerald, T.; Hu, M.; Ihm, C. H.; Kristiansson, K.; Macarthur, D. G.; MacDonald, J. R.; Onyiah, I.; Pang, A. W.; Robson, S.; Stirrups, K.; Valsesia, A.; Walter, K.; Wei, J.; The Wellcome Trust Case Control Consortium; Tyler-Smith, C.; Carter, N. P.; Lee, C.; Scherer, S. W.; Hurles, M. E. Origins and functional impact of copy number variation in the human genome. *Nature*, **2009**, *464*, 704-712.

[3] Beckmann, J. S.; Estivill, X.; Antonarakis, S. E. Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. *Nat. Rev. Genet.*, **2007**, *8*, 639-646.

[4] Redon, R.; Ishikawa, S.; Fitch, K. R.; Feuk, L.; Perry, G. H.; Andrews, T. D.; Fiegler, H.; Shaper, M. H.; Carson, A. R.; Chen, W.;

Cho, E. K.; Dallaire, S.; Freeman, J. L.; Gonzalez, J. R.; Gratacos, M.; Huang, J.; Kalaitzopoulos, D.; Komura, D.; MacDonald, J. R.; Marshall, C. R.; Mei, R.; Montgomery, L.; Nishimura, K.; Okamura, K.; Shen, F.; Somerville, M. J.; Tehinda, J.; Valsesia, A.; Woodwark, C.; Yang, F.; Zhang, J.; Zerjal, T.; Zhang, J.; Armenogol, L.; Conrad, D. F.; Estivill, X.; Tyler-Smith, C.; Carter, N. P.; Aburatani, H.; Lee, C.; Jones, K. W.; Scherer, S. W.; Hurles, M. E. Global variation in copy number in the human genome. *Nature*, **2006**, *444*, 444-454.

[5] Korn, J. M.; Kuruvilla, F. G.; McCarroll, S. A.; Wysoker, A.; Nemesh, J.; Cawley, S.; Hubbell, E.; Veitch, J.; Collins, P. J.; Darvishi, K.; Lee, C.; Nizzari, M. M.; Gabriel, S. B.; Purcell, S.; Daly, M. J.; Altshuler, D. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat. Genet.*, **2008**, *40*, 1253-1260.

[6] McCarroll, S. A.; Kuruvilla, F. G.; Korn, J. M.; Cawley, S.; Nemesh, J.; Wysoker, A.; Shaper, M. H.; de Bakker, P. I.; Maller, J. B.; Kirby, A.; Elliott, A. L.; Parkin, M.; Hubbell, E.; Webster, T.; Mei, R.; Veitch, J.; Collins, P. J.; Handsaker, R.; Lincoln, S.; Nizzari, M.; Blume, J.; Jones, K. W.; Rava, R.; Daly, M. J.; Gabriel, S. B.; Altshuler, D. Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat. Genet.*, **2008**, *40*, 1166-1174.

[7] Firth, H. V.; Richards, S. M.; Bevan, A. P.; Clayton, S.; Corpas, M.; Rajan, D.; Van Vooren, S.; Moreau, Y.; Pettett, R. M.; Carter, N. P. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am. J. Hum. Genet.*, **2009**, *84*, 524-533.

[8] International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature*, **2008**, *455*, 237-241.

[9] Stefansson, H.; Rujescu, D.; Cichon, S.; Pietilainen, O. P.; Ingason, A.; Steinberg, S.; Fossdal, R.; Sigurdsson, E.; Sigmundsson, T.; Buizer-Voskamp, J. E.; Hansen, T.; Jakobsen, K. D.; Muglia, P.; Francks, C.; Matthews, P. M.; Gylfason, A.; Halldorsson, B. V.; Gudbjartsson, D.; Thorgeirsson, T. E.; Sigurdsson, A.; Jonasdottir, A.; Jonasdottir, A.; Bjornsson, A.; Mattiasdottir, S.; Blondal, T.; Haraldsson, M.; Magnusdottir, B. B.; Giegling, I.; Moller, H. J.; Hartmann, A.; Shianna, K. V.; Ge, D.; Need, A. C.; Crombie, C.; Fraser, G.; Walker, N.; Lonqvist, J.; Suvisaari, J.; Tuulio-Henriksson, A.; Paunio, T.; Touloupoulou, T.; Bramon, E.; Di Forti, M.; Murray, R.; Ruggeri, M.; Vassos, E.; Tosato, S.; Walshe, M.; Li, T.; Vasilescu, C.; Muhleisen, T. W.; Wang, A. G.; Ullum, H.; Djurovic, S.; Melle, I.; Olesen, J.; Kiemene, L. A.; Franke, B.; Sabatti, C.; Freimer, N. B.; Gulcher, J. R.; Thorsteinsdottir, U.; Kong, A.; Andreassen, O. A.; Ophoff, R. A.; Georgi, A.; Rietschel, M.; Werge, T.; Petursson, H.; Goldstein, D. B.; Nothen, M. M.; Peltonen, L.; Collier, D. A.; St Clair, D.; Stefansson, K.; Kahn, R. S.; Linszen, D. H.; van Os, J.; Wiersma, D.; Bruggeman, R.; Cahn, W.; de Haan, L.; Krabbendam, L.; Myin-Germeys, I. Genetic Risk and Outcome in Psychosis (GROUP) Large recurrent microdeletions associated with schizophrenia. *Nature*, **2008**, *455*, 232-236.

[10] Sebat, J.; Lakshmi, B.; Malhotra, D.; Troge, J.; Lese-Martin, C.; Walsh, T.; Yamrom, B.; Yoon, S.; Krasnitz, A.; Kendall, J.; Leotta, A.; Pai, D.; Zhang, R.; Lee, Y. H.; Hicks, J.; Spence, S. J.; Lee, A. T.; Puura, K.; Lehtimaki, T.; Ledbetter, D.; Gregersen, P. K.; Bregman, J.; Sutcliffe, J. S.; Jobanputra, V.; Chung, W.; Warburton, D.; King, M. C.; Skuse, D.; Geschwind, D. H.; Gilliam, T. C.; Ye, K.; Wigler, M. Strong association of *de novo* copy number mutations with autism. *Science*, **2007**, *316*, 445-449.

[11] Sharp, A. J.; Mefford, H. C.; Li, K.; Baker, C.; Skinner, C.; Stevenson, R. E.; Schroer, R. J.; Novara, F.; De Gregori, M.; Ciccone, R.; Brooker, A.; Casuga, I.; Wang, Y.; Xiao, C.; Barbacioru, C.; Gimelli, G.; Bernardina, B. D.; Torniero, C.; Giorda, R.; Regan, R.; Murday, V.; Mansour, S.; Fichera, M.; Castiglia, L.; Failla, P.; Ventura, M.; Jiang, Z.; Cooper, G. M.; Knight, S. J.; Romano, C.; Zuffardi, O.; Chen, C.; Schwartz, C. E.; Eichler, E. E. A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures. *Nat. Genet.*, **2008**, *40*, 322-328.

[12] Shinawi, M.; Liu, P.; Kang, S. H.; Shen, J.; Belmont, J. W.; Scott, D. A.; Probst, F. J.; Craigen, W. J.; Graham, B.; Pursley, A.; Clark, G.; Lee, J.; Proud, M.; Stocco, A.; Rodriguez, D.; Kozel, B.; Sparagana, S.; Roeder, E.; McGrew, S.; Kurczynski, T.; Allison, L.; Amato, S.; Savage, S.; Patel, A.; Stankiewicz, P.; Beaudet, A.; Cheung, S. W.; Lupski, J. R. Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioral

- problems, dysmorphism, epilepsy, and abnormal head size. *J. Med. Genet.*, **2009**, *47*, 332-341.
- [13] McCarthy, S. E.; Makarov, V.; Kirov, G.; Addington, A. M.; McClellan, J.; Yoon, S.; Perkins, D. O.; Dickel, D. E.; Kusenda, M.; Krastoshevsky, O.; Krause, V.; Kumar, R. A.; Grozeva, D.; Malhotra, D.; Walsh, T.; Zackai, E. H.; Kaplan, P.; Ganesh, J.; Krantz, I. D.; Spinner, N. B.; Roccanova, P.; Bhandari, A.; Pavon, K.; Lakshmi, B.; Leotta, A.; Kendall, J.; Lee, Y. H.; Vacic, V.; Gary, S.; Iakoucheva, L. M.; Crow, T. J.; Christian, S. L.; Lieberman, J. A.; Stroup, T. S.; Lehtimäki, T.; Puura, K.; Haldeman-Englert, C.; Pearl, J.; Goodell, M.; Willour, V. L.; Derosse, P.; Steele, J.; Kassem, L.; Wolff, J.; Chitkara, N.; McMahon, F. J.; Malhotra, A. K.; Potash, J. B.; Schulze, T. G.; Nothen, M. M.; Cichon, S.; Rietschel, M.; Leibenluft, E.; Kustanovich, V.; Lajonchere, C. M.; Sutcliffe, J. S.; Skuse, D.; Gill, M.; Gallagher, L.; Mendell, N. R.; Wellcome Trust Case Control Consortium; Craddock, N.; Owen, M. J.; O'Donovan, M. C.; Shaikh, T. H.; Susser, E.; Delisi, L. E.; Sullivan, P. F.; Deutsch, C. K.; Rapoport, J.; Levy, D. L.; King, M. C.; Sebat, J. Microduplications of 16p11.2 are associated with schizophrenia. *Nat. Genet.*, **2009**, *41*, 1223-1227.
- [14] Henrichsen, C. N.; Vinckenbosch, N.; Zollner, S.; Chagnat, E.; Pradervand, S.; Schutz, F.; Ruedi, M.; Kaessmann, H.; Reymond, A. Segmental copy number variation shapes tissue transcriptomes. *Nat. Genet.*, **2009**, *41*, 424-429.
- [15] Molina, J.; Carmona-Mora, P.; Chrast, J.; Krall, P. M.; Canales, C. P.; Lupski, J. R.; Reymond, A.; Walz, K. Abnormal social behaviors and altered gene expression rates in a mouse model for Potocki-Lupski syndrome. *Hum. Mol. Genet.*, **2008**, *17*, 2486-2495.
- [16] Ricard, G.; Molina, J.; Chrast, J.; Gu, W.; Gheldof, N.; Pradervand, S.; Schütz, F.; Young, J. I.; Lupski, J. R.; Reymond, A.; Walz, K. Phenotypic consequences of copy number variation: insights from Smith-Magenis and Potocki-Lupski syndrome mouse models. *PLoS Biology*, **2010**, *In press*.
- [17] Lupski, J. R. Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet.*, **1998**, *14*, 417-422.
- [18] Greenberg, F.; Guzzetta, V.; Montes de Oca-Luna, R.; Magenis, R. E.; Smith, A. C.; Richter, S. F.; Kondo, I.; Dobyns, W. B.; Patel, P. I.; Lupski, J. R. Molecular analysis of the Smith-Magenis syndrome: a possible contiguous-gene syndrome associated with del(17)(p11.2). *Am. J. Hum. Genet.*, **1991**, *49*, 1207-1218.
- [19] Chen, K. S.; Manian, P.; Koeuth, T.; Potocki, L.; Zhao, Q.; Chinault, A. C.; Lee, C. C.; Lupski, J. R. Homologous recombination of a flanking repeat gene cluster is a mechanism for a common contiguous gene deletion syndrome. *Nat. Genet.*, **1997**, *17*, 154-163.
- [20] Smith, A. C.; Magenis, R. E.; Elsea, S. H. Overview of Smith-Magenis syndrome. *J. Assoc. Genet. Technol.*, **2005**, *31*, 163-167.
- [21] Juyal, R. C.; Figueroa, L. E.; Hauge, X.; Elsea, S. H.; Lupski, J. R.; Greenberg, F.; Baldini, A.; Patel, P. I. Molecular analyses of 17p11.2 deletions in 62 Smith-Magenis syndrome patients. *Am. J. Hum. Genet.*, **1996**, *58*, 998-1007.
- [22] Vlangos, C. N.; Yim, D. K.; Elsea, S. H. Refinement of the Smith-Magenis syndrome critical region to approximately 950 kb and assessment of 17p11.2 deletions. Are all deletions created equally? *Mol. Genet. Metab.*, **2003**, *79*, 134-141.
- [23] Potocki, L.; Chen, K. S.; Park, S. S.; Osterholm, D. E.; Withers, M. A.; Kimonis, V.; Summers, A. M.; Meschino, W. S.; Anyane-Yeboah, K.; Kashork, C. D.; Shaffer, L. G.; Lupski, J. R. Molecular mechanism for duplication 17p11.2- the homologous recombination reciprocal of the Smith-Magenis microdeletion. *Nat. Genet.*, **2000**, *24*, 84-87.
- [24] Nakamine, A.; Ouchanov, L.; Jimenez, P.; Manghi, E. R.; Esquivel, M.; Monge, S.; Fallas, M.; Burton, B. K.; Szomju, B.; Elsea, S. H.; Marshall, C. R.; Scherer, S. W.; McInnes, L. A. Duplication of 17(p11.2p11.2) in a male child with autism and severe language delay. *Am. J. Med. Genet. A.*, **2008**, *146A*, 636-643.
- [25] Potocki, L.; Bi, W.; Treadwell-Deering, D.; Carvalho, C. M.; Eifert, A.; Friedman, E. M.; Glaze, D.; Krull, K.; Lee, J. A.; Lewis, R. A.; Mendoza-Londono, R.; Robbins-Furman, P.; Shaw, C.; Shi, X.; Weissenberger, G.; Withers, M.; Yatsenko, S. A.; Zackai, E. H.; Stankiewicz, P.; Lupski, J. R. Characterization of Potocki-Lupski syndrome (dup(17)(p11.2p11.2)) and delineation of a dosage-sensitive critical interval that can convey an autism phenotype. *Am. J. Hum. Genet.*, **2007**, *80*, 633-649.
- [26] Treadwell-Deering, D. E.; Powell, M. P.; Potocki, L. Cognitive and Behavioral Characterization of the Potocki-Lupski Syndrome (Duplication 17p11.2). *J. Dev. Behav. Pediatr.*, **2010**, *31*, 137-143.
- [27] Elsea, S. H.; Girirajan, S. Smith-Magenis syndrome. *Eur. J. Hum. Genet.*, **2008**, *16*, 412-421.
- [28] Edelman, E. A.; Girirajan, S.; Finucane, B.; Patel, P. I.; Lupski, J. R.; Smith, A. C.; Elsea, S. H. Gender, genotype, and phenotype differences in Smith-Magenis syndrome: a meta-analysis of 105 cases. *Clin. Genet.*, **2007**, *71*, 540-550.
- [29] Zhang, F.; Potocki, L.; Sampson, J. B.; Liu, P.; Sanchez-Valle, A.; Robbins-Furman, P.; Navarro, A. D.; Wheeler, P. G.; Spence, J. E.; Brasington, C. K.; Withers, M. A.; Lupski, J. R. Identification of uncommon recurrent Potocki-Lupski syndrome-associated duplications and the distribution of rearrangement types and mechanisms in PTLs. *Am. J. Hum. Genet.*, **2010**.
- [30] Slager, R. E.; Newton, T. L.; Vlangos, C. N.; Finucane, B.; Elsea, S. H. Mutations in RAI1 associated with Smith-Magenis syndrome. *Nat. Genet.*, **2003**, *33*, 466-468.
- [31] Bi, W.; Saifi, G. M.; Shaw, C. J.; Walz, K.; Fonseca, P.; Wilson, M.; Potocki, L.; Lupski, J. R. Mutations of RAI1, a PHD-containing protein, in nondeletion patients with Smith-Magenis syndrome. *Hum. Genet.*, **2004**, *115*, 515-524.
- [32] Girirajan, S.; Elsas, L. J., 2nd; Devriendt, K.; Elsea, S. H. RAI1 variations in Smith-Magenis syndrome patients without 17p11.2 deletions. *J. Med. Genet.*, **2005**, *42*, 820-828.
- [33] Bi, W.; Saifi, G. M.; Girirajan, S.; Shi, X.; Szomju, B.; Firth, H.; Magenis, R. E.; Potocki, L.; Elsea, S. H.; Lupski, J. R. RAI1 point mutations, CAG repeat variation, and SNP analysis in non-deletion Smith-Magenis syndrome. *Am. J. Med. Genet. A.*, **2006**, *140*, 2454-2463.
- [34] Carmona-Mora, P.; Encina, C. A.; Canales, C. P.; Cao, L.; Molina, J.; Kairath, P.; Young, J. I.; Walz, K. Functional and cellular characterization of human Retinoic Acid Induced 1 (RAI1) mutations associated with Smith-Magenis Syndrome. *BMC Mol. Biol.*, **2010**, *11*, 63.
- [35] Girirajan, S.; Vlangos, C. N.; Szomju, B. B.; Edelman, E.; Trevors, C. D.; Dupuis, L.; Nezarati, M.; Bunyan, D. J.; Elsea, S. H. Genotype-phenotype correlation in Smith-Magenis syndrome: evidence that multiple genes in 17p11.2 contribute to the clinical spectrum. *Genet. Med.*, **2006**, *8*, 417-427.
- [36] Lucas, R. E.; Vlangos, C. N.; Das, P.; Patel, P. I.; Elsea, S. H. Genomic organization of the approximately 1.5 Mb Smith-Magenis syndrome critical interval: transcription map, genomic contig, and candidate gene analysis. *Eur. J. Hum. Genet.*, **2001**, *9*, 892-902.
- [37] Bi, W.; Yan, J.; Stankiewicz, P.; Park, S. S.; Walz, K.; Boerkoel, C. F.; Potocki, L.; Shaffer, L. G.; Devriendt, K.; Nowaczyk, M. J.; Inoue, K.; Lupski, J. R. Genes in a refined Smith-Magenis syndrome critical deletion interval on chromosome 17p11.2 and the syntenic region of the mouse. *Genome Res.*, **2002**, *12*, 713-728.
- [38] Zody, M. C.; Garber, M.; Adams, D. J.; Sharpe, T.; Harrow, J.; Lupski, J. R.; Nicholson, C.; Searle, S. M.; Wilming, L.; Young, S. K.; Abouelleil, A.; Allen, N. R.; Bi, W.; Bloom, T.; Borowsky, M. L.; Bugalter, B. E.; Butler, J.; Chang, J. L.; Chen, C. K.; Cook, A.; Corum, B.; Cuomo, C. A.; de Jong, P. J.; DeCaprio, D.; Dewar, K.; FitzGerald, M.; Gilbert, J.; Gibson, R.; Gnerre, S.; Goldstein, S.; Grafham, D. V.; Grocock, R.; Hafez, N.; Hagopian, D. S.; Hart, E.; Norman, C. H.; Humphray, S.; Jaffe, D. B.; Jones, M.; Kamal, M.; Khodiyar, V. K.; LaButti, K.; Laird, G.; Lehoczyk, J.; Liu, X.; Lokyitsang, T.; Loveland, J.; Lui, A.; Macdonald, P.; Major, J. E.; Matthews, L.; Mauceli, E.; McCarroll, S. A.; Mihalev, A. H.; Mudge, J.; Nguyen, C.; Nicol, R.; O'Leary, S. B.; Osoegawa, K.; Schwartz, D. C.; Shaw-Smith, C.; Stankiewicz, P.; Steward, C.; Swarbreck, D.; Venkataraman, V.; Whittaker, C. A.; Yang, X.; Zimmer, A. R.; Bradley, A.; Hubbard, T.; Birren, B. W.; Rogers, J.; Lander, E. S.; Nusbaum, C. DNA sequence of human chromosome 17 and analysis of rearrangement in the human lineage. *Nature*, **2006**, *440*, 1045-1049.
- [39] Ramirez-Solis, R.; Liu, P.; Bradley, A. Chromosome engineering in mice. *Nature*, **1995**, *378*, 720-724.
- [40] Zheng, B.; Mills, A. A.; Bradley, A. A system for rapid generation of coat color-tagged knockouts and defined chromosomal rearrangements in mice. *Nucleic Acids Res.*, **1999**, *27*, 2354-2360.
- [41] Mills, A. A.; Bradley, A. From mouse to man: generating megabase chromosome rearrangements. *Trends Genet.*, **2001**, *17*, 331-339.

- [42] Yu, Y.; Bradley, A. Engineering chromosomal rearrangements in mice. *Nat. Rev. Genet.*, **2001**, *2*, 780-790.
- [43] van der Weyden, L.; Bradley, A. Mouse chromosome engineering for modeling human disease. *Annu. Rev. Genomics Hum. Genet.*, **2006**, *7*, 247-276.
- [44] Walz, K.; Caratini-Rivera, S.; Bi, W.; Fonseca, P.; Mansouri, D. L.; Lynch, J.; Vogel, H.; Noebels, J. L.; Bradley, A.; Lupski, J. R. Modeling del(17)(p11.2p11.2) and dup(17)(p11.2p11.2) contiguous gene syndromes by chromosome engineering in mice: phenotypic consequences of gene dosage imbalance. *Mol. Cell. Biol.*, **2003**, *23*, 3646-3655.
- [45] Walz, K.; Spencer, C.; Kaasik, K.; Lee, C. C.; Lupski, J. R.; Paylor, R. Behavioral characterization of mouse models for Smith-Magenis syndrome and dup(17)(p11.2p11.2). *Hum. Mol. Genet.*, **2004**, *13*, 367-378.
- [46] Goldman, A. M.; Potocki, L.; Walz, K.; Lynch, J. K.; Glaze, D. G.; Lupski, J. R.; Noebels, J. L. Epilepsy and chromosomal rearrangements in Smith-Magenis Syndrome [del(17)(p11.2p11.2)]. *J. Child Neurol.*, **2006**, *21*, 93-98.
- [47] Carmona-Mora, P.; Molina, J.; Encina, C. A.; Walz, K. Mouse models of genomic syndromes as tools for understanding the basis of complex traits: an example with the Smith-Magenis and the Potocki-Lupski syndromes. *Curr. Genomics*, **2009**, *10*, 259-268.
- [48] Bi, W.; Ohyama, T.; Nakamura, H.; Yan, J.; Visvanathan, J.; Justice, M. J.; Lupski, J. R. Inactivation of *Rai1* in mice recapitulates phenotypes observed in chromosome engineered mouse models for Smith-Magenis syndrome. *Hum. Mol. Genet.*, **2005**, *14*, 983-995.
- [49] Bi, W.; Yan, J.; Shi, X.; Yuva-Paylor, L. A.; Antalffy, B. A.; Goldman, A.; Yoo, J. W.; Noebels, J. L.; Armstrong, D. L.; Paylor, R.; Lupski, J. R. *Rai1* deficiency in mice causes learning impairment and motor dysfunction, whereas *Rai1* heterozygous mice display minimal behavioral phenotypes. *Hum. Mol. Genet.*, **2007**, *16*, 1802-1813.
- [50] Walz, K.; Paylor, R.; Yan, J.; Bi, W.; Lupski, J. R. *Rai1* duplication causes physical and behavioral phenotypes in a mouse model of dup(17)(p11.2p11.2). *J. Clin. Invest.*, **2006**, *116*, 3035-3041.
- [51] Girirajan, S.; Patel, N.; Slager, R. E.; Tokarz, M. E.; Bucan, M.; Wiley, J. L.; Elsea, S. H. How much is too much? Phenotypic consequences of *Rai1* overexpression in mice. *Eur. J. Hum. Genet.*, **2008**, *16*, 941-954.
- [52] Girirajan, S.; Elsea, S. H. Abnormal maternal behavior, altered sociability, and impaired serotonin metabolism in *Rai1*-transgenic mice. *Mamm. Genome*, **2009**, *20*, 247-255.
- [53] Seranski, P.; Hoff, C.; Radelof, U.; Hennig, S.; Reinhardt, R.; Schwartz, C. E.; Heiss, N. S.; Poustka, A. *RAI1* is a novel polyglutamine encoding gene that is deleted in Smith-Magenis syndrome patients. *Gene*, **2001**, *270*, 69-76.
- [54] Maier, D.; Stumm, G.; Kuhn, K.; Preiss, A. Hairless, a Drosophila gene involved in neural development, encodes a novel, serine rich protein. *Mech. Dev.*, **1992**, *38*, 143-156.
- [55] Toulouse, A.; Rochefort, D.; Roussel, J.; Joobert, R.; Rouleau, G. A. Molecular cloning and characterization of human *RAI1*, a gene associated with schizophrenia. *Genomics*, **2003**, *82*, 162-171.
- [56] Nagase, T.; Nakayama, M.; Nakajima, D.; Kikuno, R.; Ohara, O. Prediction of the coding sequences of unidentified human genes. XX. The complete sequences of 100 new cDNA clones from brain which code for large proteins *in vitro*. *DNA Res.*, **2001**, *8*, 85-95.
- [57] Gerhard, D. S.; Wagner, L.; Feingold, E. A.; Shenmen, C. M.; Grouse, L. H.; Schuler, G.; Klein, S. L.; Old, S.; Rasooly, R.; Good, P.; Guyer, M.; Peck, A. M.; Derge, J. G.; Lipman, D.; Collins, F. S.; Jang, W.; Sherry, S.; Feolo, M.; Misquitta, L.; Lee, E.; Rotmistrovsky, K.; Greenhut, S. F.; Schaefer, C. F.; Buetow, K.; Bonner, T. I.; Haussler, D.; Kent, J.; Kiekhaus, M.; Furey, T.; Brent, M.; Prange, C.; Schreiber, K.; Shapiro, N.; Bhat, N. K.; Hopkins, R. F.; Hsie, F.; Driscoll, T.; Soares, M. B.; Casavant, T. L.; Scheetz, T. E.; Brown-stein, M. J.; Usdin, T. B.; Toshiyuki, S.; Carninci, P.; Piao, Y.; Dudekula, D. B.; Ko, M. S.; Kawakami, K.; Suzuki, Y.; Sugano, S.; Gruber, C. E.; Smith, M. R.; Simmons, B.; Moore, T.; Waterman, R.; Johnson, S. L.; Ruan, Y.; Wei, C. L.; Mathavan, S.; Gunaratne, P. H.; Wu, J.; Garcia, A. M.; Hulyk, S. W.; Fuh, E.; Yuan, Y.; Sneed, A.; Kowis, C.; Hodgson, A.; Muzny, D. M.; McPherson, J.; Gibbs, R. A.; Fahey, J.; Helton, E.; Ketterman, M.; Madan, A.; Rodrigues, S.; Sanchez, A.; Whiting, M.; Madari, A.; Young, A. C.; Wetherby, K. D.; Granite, S. J.; Kwong, P. N.; Brinkley, C. P.; Pearson, R. L.; Bouffard, G. G.; Blakesly, R. W.; Green, E. D.; Dickson, M. C.; Rodriguez, A. C.; Grimwood, J.; Schmutz, J.; Myers, R. M.; Butterfield, Y. S.; Griffith, M.; Griffith, O. L.; Krzywinski, M. I.; Liao, N.; Morin, R.; Palmquist, D.; Petrescu, A. S.; Skalska, U.; Smailus, D. E.; Stott, J. M.; Scherch, A.; Schein, J. E.; Jones, S. J.; Holt, R. A.; Baross, A.; Marra, M. A.; Clifton, S.; Makowski, K. A.; Bosak, S.; Malek, J.; MGC Project Team The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). *Genome Res.*, **2004**, *14*, 2121-2127.
- [58] Julenius, K.; Molgaard, A.; Gupta, R.; Brunak, S. Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology*, **2005**, *15*, 153-164.
- [59] Johansen, M. B.; Kiemer, L.; Brunak, S. Analysis and prediction of mammalian protein glycation. *Glycobiology*, **2006**, *16*, 844-853.
- [60] Gupta, R.; Brunak, S. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac. Symp. Biocomput.*, **2002**, 310-322.
- [61] Lazarus, B. D.; Love, D. C.; Hanover, J. A. O-GlcNAc cycling: implications for neurodegenerative disorders. *Int. J. Biochem. Cell Biol.*, **2009**, *41*, 2134-2146.
- [62] Blom, N.; Gammeltoft, S.; Brunak, S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.*, **1999**, *294*, 1351-1362.
- [63] Olsen, J. V.; Blagoev, B.; Gnäd, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M. Global, *in vivo*, and site-specific phosphorylation dynamics in signaling networks. *Cell*, **2006**, *127*, 635-648.
- [64] Dephoure, N.; Zhou, C.; Villen, J.; Beausoleil, S. A.; Bakalarski, C. E.; Elledge, S. J.; Gygi, S. P. A quantitative atlas of mitotic phosphorylation. *Proc. Natl. Acad. Sci. U. S. A.*, **2008**, *105*, 10762-10767.
- [65] Mayya, V.; Lundgren, D. H.; Hwang, S. I.; Rezaul, K.; Wu, L.; Eng, J. K.; Rodionov, V.; Han, D. K. Quantitative phosphoproteomic analysis of T cell receptor signaling reveals system-wide modulation of protein-protein interactions. *Sci. Signal.*, **2009**, *2*, ra46.
- [66] Gauci, S.; Helbig, A. O.; Slijper, M.; Krijgsveld, J.; Heck, A. J.; Mohammed, S. Lys-N and trypsin cover complementary parts of the phosphoproteome in a refined SCX-based approach. *Anal. Chem.*, **2009**, *81*, 4493-4501.
- [67] Spange, S.; Wagner, T.; Heinzl, T.; Kramer, O. H. Acetylation of non-histone proteins modulates cellular signalling at multiple levels. *Int. J. Biochem. Cell Biol.*, **2009**, *41*, 185-198.
- [68] Choudhary, C.; Kumar, C.; Gnäd, F.; Nielsen, M. L.; Rehman, M.; Walther, T. C.; Olsen, J. V.; Mann, M. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science*, **2009**, *325*, 834-840.
- [69] Seeler, J. S.; Dejean, A. Nuclear and unclear functions of SUMO. *Nat. Rev. Mol. Cell Biol.*, **2003**, *4*, 690-699.
- [70] Hay, R. T. SUMO: a history of modification. *Mol. Cell*, **2005**, *18*, 1-12.
- [71] Ren, J.; Gao, X.; Jin, C.; Zhu, M.; Wang, X.; Shaw, A.; Wen, L.; Yao, X.; Xue, Y. Systematic study of protein sumoylation: Development of a site-specific predictor of SUMOsp 2.0. *Proteomics*, **2009**, *9*, 3409-3412.
- [72] Imai, Y.; Suzuki, Y.; Matsui, T.; Tohyama, M.; Wanaka, A.; Takagi, T. Cloning of a retinoic acid-induced gene, *GT1*, in the embryonal carcinoma cell line P19: neuron-specific expression in the mouse brain. *Brain Res. Mol. Brain Res.*, **1995**, *31*, 1-9.
- [73] Gerber, H. P.; Seipel, K.; Georgiev, O.; Hofferer, M.; Hug, M.; Rusconi, S.; Schaffner, W. Transcriptional activation modulated by homopolymeric glutamine and proline stretches. *Science*, **1994**, *263*, 808-811.
- [74] Girirajan, S.; Truong, H. T.; Blanchard, C. L.; Elsea, S. H. A functional network module for Smith-Magenis syndrome. *Clin. Genet.*, **2009**, *75*, 364-374.
- [75] Burns, B.; Schmidt, K.; Williams, S. R.; Kim, S.; Girirajan, S.; Elsea, S. H. *Rai1* haploinsufficiency causes reduced *Bdnf* expression resulting in hyperphagia, obesity and altered fat distribution in mice and humans with no evidence of metabolic syndrome. *Hum. Mol. Genet.*, **2010**, *In press*.
- [76] Zuccato, C.; Cattaneo, E. Brain-derived neurotrophic factor in neurodegenerative diseases. *Nat. Rev. Neurol.*, **2009**, *5*, 311-322.
- [77] Binder, D. K.; Scharfman, H. E. Brain-derived neurotrophic factor. *Growth Factors*, **2004**, *22*, 123-131.
- [78] Gray, J.; Yeo, G. S.; Cox, J. J.; Morton, J.; Adlam, A. L.; Keogh, J. M.; Yanovski, J. A.; El Gharbawy, A.; Han, J. C.; Tung, Y. C.;

- Hodges, J. R.; Raymond, F. L.; O'rahilly, S.; Farooqi, I. S. Hyperphagia, severe obesity, impaired cognitive function, and hyperactivity associated with functional loss of one copy of the brain-derived neurotrophic factor (BDNF) gene. *Diabetes*, **2006**, *55*, 3366-3371.
- [79] Hayes, S.; Turecki, G.; Brisebois, K.; Lopes-Cendes, I.; Gaspar, C.; Riess, O.; Ranum, L. P.; Pulst, S. M.; Rouleau, G. A. CAG repeat length in RAI1 is associated with age at onset variability in spinocerebellar ataxia type 2 (SCA2). *Hum. Mol. Genet.*, **2000**, *9*, 1753-1758.
- [80] Joobar, R.; Benkelfat, C.; Toulouse, A.; Lafreniere, R. G.; Lal, S.; Ajroud, S.; Turecki, G.; Bloom, D.; Labelle, A.; Lalonde, P.; Alda, M.; Morgan, K.; Palmour, R.; Rouleau, G. A. Analysis of 14 CAG repeat-containing genes in schizophrenia. *Am. J. Med. Genet.*, **1999**, *88*, 694-699.
- [81] van der Zwaag, B.; Franke, L.; Poot, M.; Hochstenbach, R.; Spierenburg, H. A.; Vorstman, J. A.; van Daalen, E.; de Jonge, M. V.; Verbeek, N. E.; Brilstra, E. H.; van 't Slot, R.; Ophoff, R. A.; van Es, M. A.; Blauw, H. M.; Veldink, J. H.; Buijzer-Voskamp, J. E.; Beemer, F. A.; van den Berg, L. H.; Wijmenga, C.; van Amstel, H. K.; van Engeland, H.; Burbach, J. P.; Staal, W. G. Gene-network analysis identifies susceptibility genes related to glycobiology in autism. *PLoS One*, **2009**, *4*, e5324.