# Genetic Variations on *SETD5* Underlying Autistic Conditions

**Isabella R. Fernandes**,[1] **Ana C.P. Cruz**,[2] **Adriano Ferrasa**,[2,3] **Dylan Phan**,[1] **Roberto H. Herai**,[2,4] **Alysson R. Muotri**[1]

[1] Department of Pediatrics/Rady Children's Hospital San Diego, Department of Cellular & Molecular Medicine, Stem Cell Program, University of California San Diego, School of Medicine, La Jolla, California, 92037-0695

[2] Experimental Multiuser Laboratory (LEM), Graduate Program in Health Sciences (PPGCS), School of Medicine, Pontifícia Universidade Católica do Paraná (PUCPR), Curitiba, Paraná, 80215-901, Brazil

[3] Department of Informatics (DEINFO), Universidade Estadual de Ponta Grossa (UEPG), Ponta Grossa, Paraná, 84030-900, Brazil

[4] Lico Kaesemodel Institute (ILK), Curitiba, Paraná, 80240-000, Brazil

*Received 19 September 2017; revised 21 February 2018; accepted 21 February 2018*

**ABSTRACT:** **The prevalence of autism spectrum disorders (ASD) and the number of identified ASD-related genes have increased in recent years. The *SETD5* gene encodes a SET-containing-domain 5 protein, a likely reader enzyme. Genetic evidences suggest that *SETD5* malfunction contributes to ASD phenotype, such as on intellectual disability (ID) and facial dysmorphism. In this review, we mapped the clinical phenotypes of individuals carrying mutations on the *SETD5* gene that are associated with ASD and other chromatinopathies (mutation in epigenetic modifiers that leads to the development of neurodevelopmental disorders such as ASD). After a detailed systematic literature review and analysis of public disease-related databank, we found so far 42 individuals carrying mutations on the *SETD5* gene, with 23.8% presenting autistic-like features. Furthermore, most of mutations occurred between positions 9,480,000–9,500,000 bp on chromosome 3 (3p25.3) at the *SETD5* gene *locus*. In all males, mutations in *SETD5* presented high penetrance, while in females the clinical phenotype seems more variable with two reported cases showing normal female carriers and not presenting ASD or any ID-like symptoms. At the molecular level, SETD5 interacts with proteins of PAF1C and N-CoR complexes, leading to a possible involvement with chromatin modification pathway, which plays important roles for brain development. Together, we propose that mutations on the *SETD5* gene could lead to a new syndromic condition in males, which is linked to 3p25 syndrome, and can leads to ASD-related intellectual disability and facial dysmorphism.**  Develop Neurobiol 78: 500–518, 2018

***Keywords:*** **SETD5 gene; SETD5 syndrome; autism spectrum disorder; syndromic autism; genetic variants; intellectual disability**

*Correspondence to*: R.H. Herai (roberto.herai@pucpr.br) and A.R. Muotri (muotri@ucsd.edu)

Isabella R Fernandes and Ana CP Cruz equally contributed to this work.

Contract grant sponsor: National Cooperative Reprogrammed Cell Research Groups (NCRCRG); contract grant number: U19MH107367.

Contract grant sponsor: Autism Science Foundation.

Contract grant sponsor: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Contract grant sponsor: CAPES/FA (Fundação Araucária); contract grant number: CP09/2016.

Additional Supporting Information may be found in the online version of this article.



Published online 26 February 2018 in Wiley Online Library (wileyonlinelibrary.com).
DOI 10.1002/dneu.22584

## INTRODUCTION

Autism spectrum disorders (ASD) constitute a wide variety of childhood-onset neurodevelopmental conditions, affecting nearly 1% of the human population (Gilman et al., 2011). The pathogenic mechanisms of ASD remains unknown, and no disease-modifying treatments are available. Based on ASD highly variable genetic architecture, autism can be classified as either syndromic or non-syndromic. Syndromic types of autisms are usually monogenetic and might arise in conjunction with other disease phenotypes (Sztainberg & Zoghbi, 2016), such as Fragile X Syndrome (FXS), Rett Syndrome (RTT), Prader-Willis Syndrome, Angelman Syndrome and Timothy Syndrome (Geschwind & Levitt, 2007). Non-syndromic autism, also known as "idiopathic autism" or "classical autism", is a complex genetic disorder involving many genes that are likely to contribute to the etiology of autism (Sztainberg & Zoghbi, 2016). The vast majority of described ASD cases are non-syndromic and lack a defined genetic etiology and syndromic autism correspond approximately 10%–20% (Schiff et al., 2011). Over the years, novel evidences emerged for specific genes and conditions that could be defined as novel syndromic types of autism. Here we show that the SET-containing-domain 5 (*SETD5*) gene emerged as a candidate gene that could be associated as novel syndromic type of autism, with high penetrance in males. Rare *de novo* loss of function (LoF) mutations in this gene accounts for approximately 0.7% of intellectual disability (ID) cases and is associated with ASD phenotype. The *SETD5* gene acts as an epigenetic modifier factor and is widely related to ID (Kuechler et al., 2015). To identify all the previously described genetic variants linking the *SETD5* gene with neurological disorders, we performed a detailed Systematic Literature Review (SLR) over different reference disease-related databases of biology, genetics, health, medical and life sciences to perform a deep literature content analysis and redundancy exclusion (Fig. 1). We show that individuals carrying *SETD5* mutations share clinical ID features, including similar ASD phenotypes. *SETD5* mutations have high penetrance in males, but not in females that can be differentially affected. Furthermore, analysis of *SETD5* metabolic pathways suggest that this gene is involved in chromatin modifications, and alterations on this pathway as in other chromatinopathies-related genes can lead to transcriptome instability as the cause of neurodevelopment disorders.

## SETD5 CLINICAL FEATURES

ASD are characterized by several clinical symptoms, including impairment of social interaction, communication delays, stereotyped behaviors and difficult with eye contact. Together, all these symptoms are

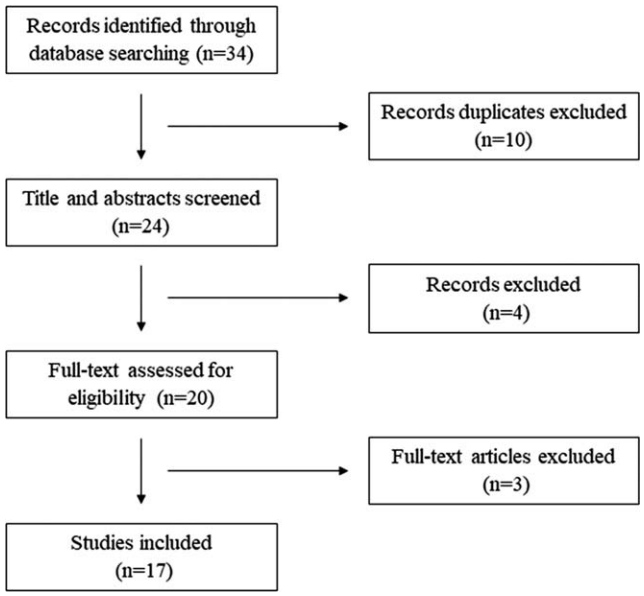


**Figure 1** Representative flowchart describing the systematic literature review for genetic variants covering *SETD5* gene. Among 32 published papers citing SETD5, only 17 studies reported mutations in *SETD5* gene.

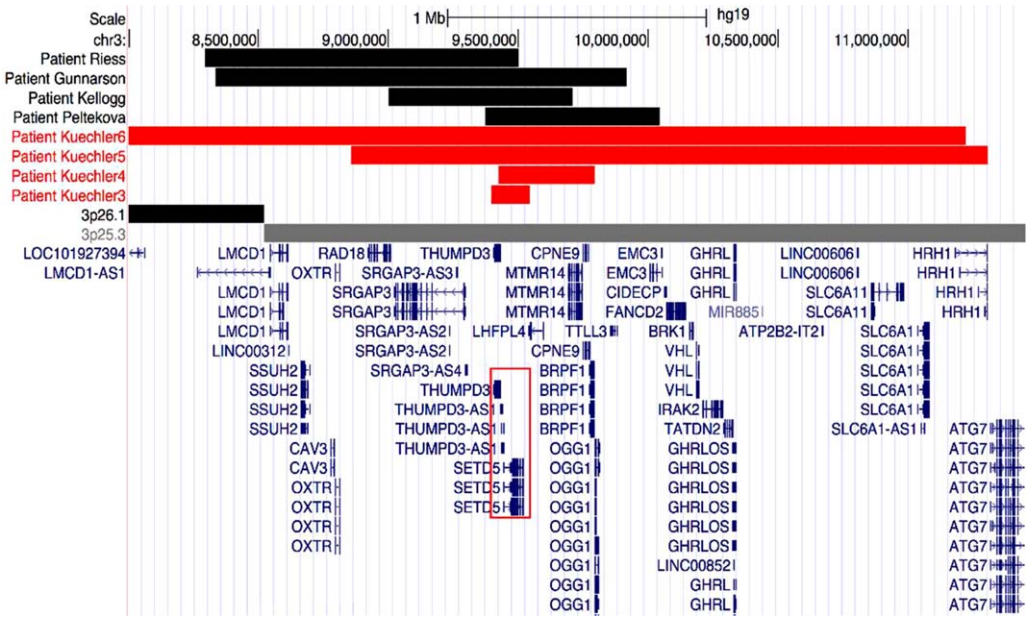


**Figure 2** Representation of 3p25 locus on autistic individuals. The locus 3p25 has, in different individuals, CNV alterations involving the genes *SETD5*, SETD5-AS1 (*THUMPD3-AS1*) and *THUMPD3* (red rectangle). *SETD5* gene was in all detected CNVs, the gene SETD5 was always involved. Adapted from Kuechler et al., 2015. [Color figure can be viewed at wileyonlinelibrary.com]

associated to the development of ID and are shared by individuals with both syndromic and non-syndromic types of autism.

Our SLR lead to the confirmation that individuals with mutations on *SETD5* share similar phenotypes with individuals with two other conditions: the 3p25 microdeletion syndrome and autosomal dominant mental retardation-23 disease (Stur et al., 2017). Recently, mutations on *SETD5* were described to overlap phenotypic features with Cornelia de Lange syndrome (CdLS; Parenti et al., 2017). The first description associating *SETD5* with ID was described within the critical chromosome region 3p25, with three genes (*SETD5*, *LOC440944,* and *THUMPD3*; Fig. 2) associated to be responsible for the condition development (Kellogg et al., 2013). Grozeva et al. (2014) was the first to show a phenotypic resemblance between individuals with 3p25 deletion and mutations in *SETD5* gene, sufficient to cause ID. Recently, 14 individuals with *SETD5* mutation were described demonstrating the variable features in phenotypes, such as facial dysmorphism (Powis et al., 2017).

In total, we found 17 articles (Fig. 1) with *SETD5* gene mutations that were described in the literature. In all the reviewed cases, patients showed ID and facial dysmorphism, implicating these clinical features as potential phenotypes associated with *SETD5* mutations (Table 1). As the possible consequence of ID, 71.4% of individuals experienced motor impairment/delay and 69.0% suffered from speech impairment/delay during childhood development. Furthermore, 23.8% of the individuals presented autistic-like features (Tables 1 and 2). Seizures, another notable clinical feature, was present in 23.8% of individuals.

Despite of only 23.8% of the individuals affected by SETD5 mutations show signs of autism, and considering mutations on this gene is rare, some reports does not describe whether the studied individual has ASD. Regardless of this fact, *SETD5* is a relatively new gene, and the first time described showing phenotypes was in 2014 (Pinto et al., 2014). ASD-related disorders present a wide percentage of individuals with ASD phenotype, such as in Rett syndrome (female individuals: 61%), tuberous sclerosis complex (36%), Angelman's syndrome (34%), Williams' syndrome (12%), Down's syndrome (16%), 22q11.2 deletion syndrome (11%; Richards et al., 2015) and Fragile X (male individuals: 46%; female individuals: 16%; Center for Disease and Control Prevention, 2017).

## GENETIC VARIANTS OF *SETD5*

SETD5 is a protein coding gene, covering a region of 82kb length at the 3p25.3 *locus*. In our SLR, we

**Table 1 Clinical features on SETD5**

| Reference | Gunnarsson and Bruun (2010) $n = 1$ | Riess et al. (2012) $n = 1$ | Peltekova et al. (2012) $n = 1$ | Kellogg et al. (2013) $n = 1$ | Grozeva et al. (2014, 2015) $n = 7$ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **SETD5 mutation** | 1.6 Mb deletion in cytoband 3p25.3-26.1 chr3:8305426-9885334 | 1.24 Mb deletion in cytoband 3p25.3-26.1 ch-r3:8250541–9491586 | 643 kb deletion ch-r3:9367274-10010209 | 684 kb interstitial 3p25.3 deletion chr3: 8,980,098-9,664,733 | c.1195A>T | c.1333C>T | c.1866C>G | c.2177_2178del | c.3001C>T | c.3771dup | c.3856del |
| **Loss of Function** | N/A | N/A | N/A | N/A | LoF | LoF | LoF | LoF | LoF | LoF | LoF |
| **Age at Evaluation** | 4 years | 3 years | 22 years | 11 years | N/A | N/A | N/A | N/A | N/A | N/A | N/A |
| **Gender** | Female | Female | Female | Female | Male | Male | Male | Male | Male | Male | Male |
| **Autistic Features** | Yes | N/A | N/A | Yes | Yes | N/A | N/A | N/A | N/A | Yes | N/A |
| **Intellectual Disability** | 1/1 | 1/1 | 1/1 | 1/1 | | | | 7/7 | | | |
| **Speech Impairment/ Delay** | 1/1 | 1/1 | 1/1 | N/A | | | | 6/7 | | | |
| **Motor Impairment/ Delay** | 1/1 | 1/1 | 1/1 | N/A | | | | 6/7 | | | |
| **Epilepsy/Seizures** | 1/1 (Transient) | 0/1 | 1/1 | 0/1 | | | | 0/7 | | | |
| **Congenital Heart Defects** | 1/1 | 0/1 | 1/1 | 0/1 | | | | 2/7 | | | |
| **Facial Dysmorphism** | 1/1 | 1/1 | 1/1 | 1/1 | | | | 7/7 | | | |
| **Hand Stereotypies/ Ritualized Behavior** | N/A | N/A | N/A | N/A | | | | 5/7 | | | |
| **Impaired Vision** | 0/1 | 1/1 | N/A | N/A | | | | N/A | | | |
| **Muscle Hypotonia** | 1/1 | 1/1 | N/A | 1/1 | | | | N/A | | | |
| **Polydactyly** | N/A | N/A | 1/1 | N/A | | | | 1/7 | | | |

**TABLE 1.** *Continued*

| Reference | | | Kuechler et al. (2015) $n = 6$ | | | | | Szczaluba et al. (2016) $n = 3$ | | Kobayashi et al. (2016) $n = 1$ | Parenti et al. (2017) $n = 2$ | | Stur et al. (2017) $n = 1$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***SETD5* Mutation** | chr3: g.9477570_ 9477650del | chr3: g.9490 270C>T | 148 kb deletion 4 genes | 371 kb deletion 10 genes | 2.45 Mb deletion 46 genes | Deletion 11.16 Mb 71 genes | c.2918 C>G | c.2918 C>G | c.2918 C>G | c.2347-7 A>G | c.2212_ 2213delAT | 54kb deletion del(3) (p25.3) chr3: 9,457, 143-9,511, 190 | c.3848_ 3849insC |
| **Loss of Function** | LoF | LoF | LoF | LoF | LoF | LoF | LoF | LoF | LoF | N/A | N/A | N/A | N/A |
| **Age at Evaluation** | 20 years | 9 years | 7 years | 1 year | 6 years | 2 years | 4 years | 1 year | 31 years | 8 years | 7 years | 6 years | 36 years |
| **Gender** | Female | Female | Male | Male | Female | Female | Male | Male | Male | Male | Male | Female | Male |
| **Autism feature** | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A |
| **Intellectual Disability** | | | 6/6 | | | | | 3/3 | | 1/1 | 2/2 | | 1/1 |
| **Speech Impairment/ Delay** | | | 5/5 (1 too young) | | | | | 1/2 (infant = N/A) | | N/A | 2/2 | | N/A |
| **Motor Impairment/ Delay** | | | 5/5 (1 too young) | | | | | 2/2 (adult father = N/A) | | 1/1 | 2/2 | | 1/1 |
| **Epilepsy/Seizures** | | | 4/6 (2 cases febrile) | | | | | 0/3 | | 1/1 | 0/2 | | N/A |
| **Congenital Heart Defects** | | | 0/6 | | | | | 2/3 | | 1/1 | 0/2 | | N/A |
| **Facial Dysmorphism** | | | 6/6 | | | | | 3/3 | | N/A | 2/2 | | N/A |
| **Hand Stereotypies/ Ritualized Behavior** | | | N/A | | | | | 0/3 | | 1/1 | N/A | | N/A |
| **Impaired Vision** | | | 5/6 | | | | | 0/3 | | N/A | 0/2 | | 1/1 |
| **Muscle Hypotonia** | | | 4/6 | | | | | 2/3 | | N/A | N/A | | N/A |
| **Polydactyly** | | | 2/6 | | | | | 2/3 | | N/A | N/A | | N/A |

**TABLE 1.** *Continued*

| Reference | Powis et al. (2017) $n = 14$ | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***SETD5* mutation** | c.1655–1656insA | c.1783-2A>T | c.3001C>T | c.582dupA | c.2347-7A>G | c.2347-7A>G | c.1967delT | c.3246delT | c.1655_1656insA | c.1655_1656insA | c.2347-7A>G | c.295 5T>A | c.295 5T>A | c.295 5T>A |
| **Loss of Function** | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A |
| **Age at Evaluation** | 7 years | 6 years | 10 years | 2 years | 3 years | 20 years | 16 years | 18 months | 8 years | N/A | 9 months | 21 years | 21 years | 49 years |
| **Gender** | Male | Male | Male | Female | Male | Male | Male | Male | Male | Female | Female | Male | Male | Female |
| **Autism feature** | Yes | No | No | No | No | Yes | Yes | No | No | No | No | Yes | Yes | No |
| **Intellectual Disability** | | | | | | | 11/14 | | | | | | | |
| **Speech Impairment/ Delay** | | | | | | | 9/14 | | | | | | | |
| **Motor Impairment/ Delay** | | | | | | | 8/14 | | | | | | | |
| **Epilepsy/Seizures** | | | | | | | 1/14 | | | | | | | |
| **Congenital Heart Defects** | | | | | | | 8/14 | | | | | | | |
| **Facial Dysmorphism** | | | | | | | 14/14 | | | | | | | |
| **Hand Stereotypies/ Ritualized Behavior** | | | | | | | N/A | | | | | | | |
| **Impaired Vision** | | | | | | | N/A | | | | | | | |
| **Muscle Hypotonia** | | | | | | | 8/14 | | | | | | | |
| **Polydactyly** | | | | | | | 4/14 | | | | | | | |

**TABLE 1.** *Continued*

| Reference | Yagasaki et al. (2017) $n = 1$ | Popp et al. (2017) $n = 1$ | Russo et al. (2017) $n = 1$ | Green et al. (2017) $n = 1$ |
|---|---|---|---|---|
| ***SETD5* mutation** | 10.1 Mb deletion (chr3:1 - 10,142,919) | c.1125dup | c.894A>G | c.1381_1388del |
| **Loss of Function** | N/A | LoF | N/A | LoF |
| **Age at Evaluation** | 4 years | N/A | 7 years | 10 years |
| **Gender** | Male | N/A | Male | Male |
| **Autism feature** | N/A | N/A | 1/1 | 0/1 |
| **Intellectual Disability** | 1/1 | 1/1 | 1/1 | 1/1 |
| **Speech Impairment/Delay** | 1/1 | N/A | 1/1 | 1/1 |
| **Motor Impairment/Delay** | 1/1 | N/A | N/A | 1/1 |
| **Epilepsy/Seizures** | 1/1 | N/A | 0/1 | 1/1 |
| **Congenital Heart Defects** | 1/1 | N/A | 0/1 | 0/1 |
| **Facial Dysmorphism** | 1/1 | N/A | N/A | 1/1 |
| **Hand Stereotypies/Ritualized Behavior** | N/A | N/A | 1/1 | 1/1 |
| **Impaired Vision** | 1/1 | N/A | N/A | 0/1 |
| **Muscle Hypotonia** | N/A | 1/1 | N/A | 0/1 |
| **Polydactyly** | N/A | N/A | N/A | 1/1 |

N/A – Phenotype not evaluated in the study. Impaired Vision = Strabismus, myopia, astigmatism. Facial dysmorphism may include: low anterior hairline, arched/thick eyebrows, synophrys, long eyelashes, depressed nasal bridge, broad nasal tip, smooth philtrum, cleft uvula, low set or abnormal ears.

**Table 2 Analysis of Clinical features related to *SETD5* mutations**

| Reference | Totals $n = 42$* | Percent Affected |
|---|---|---|
| **Gender** | Male = 66.7%, Female = 30.9%, **Not available = 2.4% | |
| **Autistic Features** | 10/42 | 23.8% |
| **Intellectual Disability** | 39/42 | 92.8% |
| **Speech Impairment/ Delay** | 29/42 | 69.0% |
| **Motor Impairment/ Delay** | 30/42 | 71.4% |
| **Epilepsy/Seizures** | 10/42 | 23.8% |
| **Congenital Heart Defects** | 16/42 | 38.0% |
| **Facial Dysmorphism** | 40/42 | 95.2% |
| **Hand Stereotypies/ Ritualized Behavior** | 9/42 | 21.4% |
| **Impaired Vision** | 8/42 | 19.0% |
| **Muscle Hypotonia** | 18/42 | 42.8% |
| **Polydactyly** | 12/42 | 28.6% |

*Number of individuals with mutation described on *SETD5* so far.
**When gender are not described.

identified three types of genetic variations on *SETD5* gene in individuals with ASD: single nucleotide variant (SNV; Fig. 3), duplication and deletion (Table 3). Most of *SETD5* SNVs were described by a single study that screened approximately one thousand individuals diagnosed with moderate to severe ID (Grozeva et al., 2015). It was found that 11% of individuals carry rare and probable pathogenic mutations in known or candidate ID-associated genes and not restricted to only to the *SETD5* gene, and ~8% of these same individuals also had pathogenic rare SNV mutation with probable loss of function (LoF). Among the LoF SNVs in *SETD5*, four correspond to nonsenses variants as follows (all genomic coordinates are based on human reference genome, build Hg19): NM_001080517:c.1195A > T (p.Lys399*), NM_001080517:c.1333C > T (p.Arg445*), NM_001080517:c.1866C > G (p.Tyr622*) and NM_001080517:c.3001C > T (p.Arg1001*). The mutation c.3001C > T, was also reported in a male with ten years old with developmental disability and dysmorphic features (Powis et al., 2017). In the same study, in patients with developmental delay, behavioral/psychiatric issues, it was found the following likely LoF *SETD5 de novo* variants in *SETD5*: NM_001080517:c.1655_1656insA, NM_001080517:c.1783-2A > T, NM_001080517:c.582dupA (p.Ala195-Serfs*), NM_001080517:c.1967delT (p.Leu656*) and NM_001080517:c.3246delT (p.Ala1083Leufs*61). Among the LoF deletions and duplication, it was also found the following three corresponding frameshift-inducing variants: deletion NM_001080517:c.2177_2178del (p.Thr726Asnfs*39), deletion NM_001080517:c.3856del (p.Ser1286Leufs*84), duplication

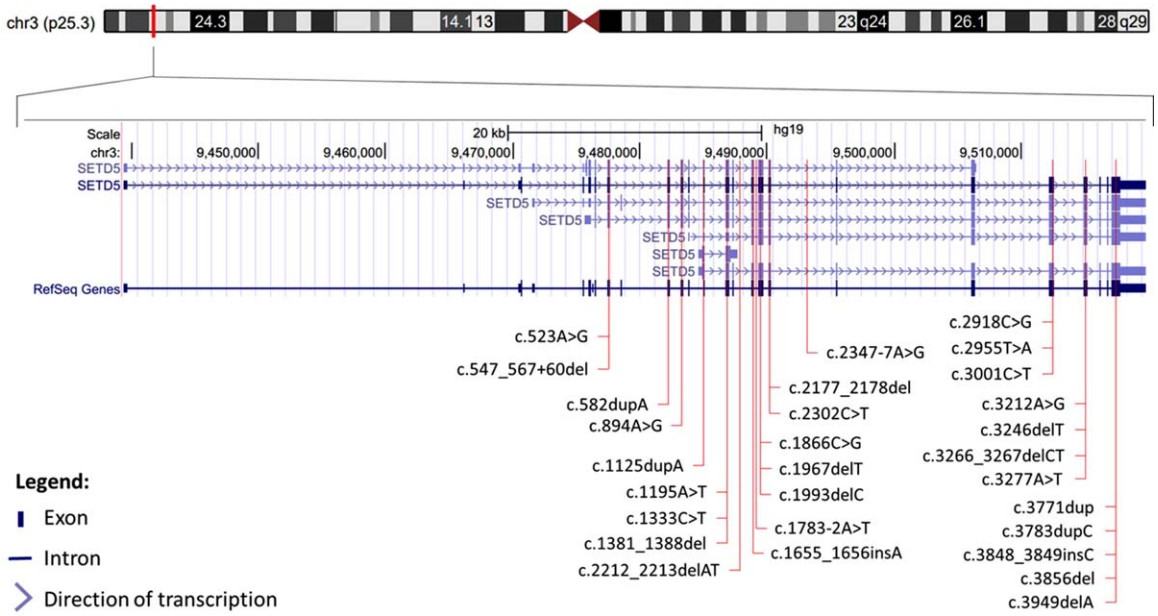


**Figure 3** *Locus* of *SETD5* gene indicating the genetic variants identified by a systematic literature review. All variants where listed within exonic and intronic regions of *SETD5* gene *locus* (chromosome 3, positions 9,439,403–9,519,838 on human reference genome, build GRCH37), with 80,436 bp length. [Color figure can be viewed at wileyonlinelibrary.com]

**Table 3 Regions and genomic locations of the mutations found in *SETD5***

| Variant | *Locus* | Type | Classification | Region | Reference |
|---|---|---|---|---|---|
| Deletion of 10.1 Mb | chr3:1 - 10,142,919 | Deletion | Not available | 3p.26.3–25.3 | Yagasaki et al., 2017 |
| Deletion of 11.16 Mb | chr3:61,891-11,220,006 | Deletion | Not available | 3p.26.3–25.3 | Kuechler et al., 2015 |
| Deletion of 1.24 Mb | chr3: 8,275,541-9,516,586 | Deletion | Not available | 3p25.3–26.1 | Riess et al., 2012 |
| Deletion of 1.6 Mb | chr3:8,330,426-9,910,334 | Deletion | Not available | 3p25.3–26.1 | Gunnarsson and Bruun, 2010 |
| Deletion of 2.45 Mb | chr3:8,856,000–11,305,600 | Deletion | Not available | 3p25.3 | Kuechler et al., 2015 |
| Deletion of 684 kb | chr3:9,005,098-9,689,733 | Deletion | Not available | 3p25.3 | Kellogg et al., 2013 |
| Deletion of 643 kb | chr3:9,392,274-10,035,209 | Deletion | Not available | 3p25.3 | Peltekova et al., 2012 |
| Deletion of 148 kb | chr3:9,394,944-9,542,885 | Deletion | Not available | 3p25.3 | Kuechler et al., 2015 |
| Deletion of 371 kb | chr3:9,422,487-9,793,524 | Deletion | Not available | 3p25.3 | Kuechler et al., 2015 |
| c.523A>G | chr3:9,477,546 | SNV | Not available | CDS | Kuechler et al., 2015 |
| c.547_567 + 60del | chr3:9,477,570 - 9,477,650 | Deletion | Pathogenic | CDS+Intron | Landrum et al., 2014 |
| Deletion of 54 kb | chr3: 9,457,143–9,511,190 | Deletion | Not available | 3p25.3 | Parenti et al., 2017 |
| c.582dupA | chr3:9,482,153-9,482,154 | Duplication | Pathogenic | CDS | Powis et al., 2017 |
| c.1125dup | chr3: 9,485,038 | Duplication | Not available | CDS | Popp et al., 2017 |
| c.1195A>T | chr3:9,486,739 | SNV | Pathogenic | CDS | Grozeva et al., 2014, 2015 |
| c.1333C>T | chr3:9,486,877 | SNV | Pathogenic | CDS | Grozeva et al., 2014, 2015 |
| c.1381_1388del | chr3:9,486,924-9,486,932 | Deletion | Not available | CDS | Green et al., 2017 |
| c.2212_2213delAT | chr3:9,487,967 | Deletion | Not available | Intron | Parenti et al., 2017 |
| c.1655_1656insA | chr3:9,488,864-9,488,865 | SNV | Pathogenic | CDS | Powis et al., 2017 |
| c.1783-2A>T | chr3:9,489,368 | SNV | Pathogenic | Intron | Landrum et al., 2014 |
| c.1866C>G | chr3:9,489,453 | SNV | Not available | CDS | Grozeva et al., 2014, 2015 |
| c.1967delT | chr3:9,489,554 | Deletion | Pathogenic | CDS | Powis et al., 2017 |

**TABLE 3.** *Continued*

| Variant | *Locus* | Type | Classification | Region | Reference |
|---|---|---|---|---|---|
| c.1993delC | chr3:9,489,580 | Deletion | Pathogenic | CDS | Landrum et al., 2014 |
| c.2177_2178del | chr3:9,490,145 - 9,490,146 | Deletion | Pathogenic | CDS | Grozeva et al., 2014 |
| c.2302C>T | chr3:9,490,270 | SNV | Pathogenic | CDS | Kuechler et al., 2015 |
| c.2347-7A>G | chr3:9,495,416 | SNV | Pathogenic | Intron | Kobayashi et al., 2016; Powis et al., 2017 |
| c.2918C>G | chr3:9,512,336 | SNV | Not available | CDS | Szczaluba et al., 2016 |
| c.2955T>A | chr3:9,512,373 | SNV | Pathogenic | CDS | Powis et al., 2017 |
| c.3001C>T | chr3:9,512,419 | SNV | Pathogenic | CDS | Grozeva et al., 2014, 2015; Powis et al., 2017 |
| c.3212A>G | chr3:9,514,936 | SNV | Not available | CDS | Szczaluba et al., 2016; Powis et al., 2017 |
| c.3246delT | chr3:9,514,970 | Deletion | Pathogenic | CDS | Landrum et al., 2014 |
| c.3266–3267delCT | chr3:9,514,990 - 9,514,991 | Deletion | Pathogenic | CDS | Landrum et al., 2014 |
| c.3277A>T | chr3:9,515,001 | SNV | Likely pathogenic | CDS | Landrum et al., 2014 |
| c.3771dup | chr3:9,517,217 | Duplication | Pathogenic | CDS | Grozeva et al., 2014, 2015 |
| c.3783dupC | chr3:9,517,229 | Duplication | Pathogenic | CDS | Landrum et al., 2014 |
| c.3848_3849insC | chr3:9,517,294 | SNV | Likely pathogenic | CDS | Stur et al., 2017 |
| c.3856del | chr3:9,517,302 | Deletion | Not available | CDS | Grozeva et al., 2014, 2015 |
| c.3949delA | chr3:9,517,395 | Deletion | Likely pathogenic | CDS | Landrum et al., 2014 |
| c.894A>G | Not available | SNV | Not available | CDS | Russo et al., 2017 |

Reported variants follow the coordinates of human reference genome, build GRCH37. Variant: correspond to the genetic variant found; *Locus*: genomic location of the genetic variant found; Type: describe the variant type, if SNV, Duplication or Deletion; Region: describe the *SETD5* region were the Variant was described: CDS or Intron for point mutations and cytoband for large deletions (<100 bp).

NM_001080517:c.3771dup (p.Ser1258Glufs*65; Grozeva et al., 2014, 2015) and deletion NM_0010 80517:c.1381_1388del (p.Asn461fs; Green et al., 2017). All reported variants were heterozygous within analyzed male individual genomes, suggesting a dominant inheritance pattern of SETD5 genetic variants (Grozeva et al., 2014, 2015). Another LoF *de novo* variant (NM_001080517:c.1125dup; Popp et al., 2017), was found in a patient with hypotonia and mild intellectual disability. This duplication change the CTG reference allele by CTTG and results in a frameshift (p.Val376Cysfs*9).

Analysis of the mutations described as LoF in *SETD5* individuals revealed that 50% had a deletion and another 50% had a nonsense or frameshift variation. Furthermore, direct correlations between LoF SNVs on *SETD5* and ASD features were not found. The prevalence of ASD in children is around 1:68 with 62% in males (Source: CDC – Center for Disease Control and Prevention, 2017). We found in this review a similar gender ratio of 66.7% of males, 30.9% of females and only 2.4% was not available when considering all individuals carrying deleterious *SETD5* mutations (Table 2). A description of individuals carrying these *SETD5* mutations and their corresponding phenotypes revealed ID as the prominent clinical feature among other conditions (Supporting Information Table S1).

An investigative study was performed in individuals with developmental delay/ID (Kuechler et al., 2015). Using a whole exome sequencing approach, two individuals with *de novo* SNVs on *SETD5* gene were found. The first individual, a 9-year-old female, contained a NM_001080517:c.2302C > T (p.Arg768*) nonsense SNV mutation; the second individual, a 20-year-old female contained a NM_001080517:c.523A > G (p.Ser 175Gly) nonsense SNV variant and, in addition to intellectual disability he also had mild attention deficit disorder (Kuechler et al., 2015).

In another case, it was reported a male individual with several clinical conditions, such as epilepsy, developmental delay, cognitive impairment, with a West Syndrome diagnosis. After genetic analysis, a *de novo* insertion of a 6-bp (TTATAG) sequence within intron 16 of the *SETD5* gene was found (Kobayashi et al., 2016). The insertion NM_00108 0517:c.2347-7A > G resulted in a premature stop codon within the transcript coding sequence (p.Arg783Leufs*2). The reported mutation expanded the phenotypic spectrum of mutations previously found in the *SETD5* gene, allowing the inclusion of the early-onset epileptic encephalopathies (EOEEs) as a novel condition associated with *SETD5* mutations (Kobayashi et al., 2016). This same insertion was found as a *de novo* event in other three patients: two males and one female (Powis et al., 2017). In the first male patient, with 3 years old, the variant was associated with developmental delay and dysmorphic features. The second male patient, with 20 years old, had ASD, ADHD (Attention-deficit/hyperactivity disorder), epilepsy and low IQ (57). The female patient did not show signs of developmental delay neither behavioral/psychiatric issues but presented dysmorphic features.

More recently, different *SETD5* mutations in a 36-year-old male whose phenotypic features included only mild motor and intellectual disability was described (Stur et al., 2017). This phenotype was associated with a novel frameshift event caused by a single nucleotide insertion (NG_034132:c.3848_ 3849insC) within *SETD5* gene coding region, that resulted in the replacement of the amino acid Serine by the Leucine (p.Ser1286Leu), and inducing a premature stop codon at position 1322 of the gene. It is hypothesized by the authors that the truncated protein leads to a partial protein activity that contributed with the mild phenotype of the individual. A clinical and molecular characterization study involving seven individuals with overlapping features with CdLS (Parenti et al., 2017), a genetically heterogeneous disease that causes growth retardation and intellectual disability (Schrier et al., 2011), revealed that none of them presented disease-causing mutations in CdLS-related genes. The only pathogenic mutation found was in a 7-year-old male child carrying a dinucleotide deletion (AT; NM_001080517:c.2212_2213delAT) on the *SETD5* gene *locus*, that induced a premature stop codon after 27 translated amino acids (p.Met738Valfs*27; Parenti et al., 2017). The child presented clinical features that included mild delayed development and intellectual disability.

In another study, whole genome sequencing of two brothers and the biological father allowed the identification of two mutations in the *SETD5* gene. The older brother presented a more severe phenotype related with delayed motor development, speech and ID. The younger brother was severely affected with ID and the father had mild intellectual impairment. All three have in the *SETD5* gene a potential nonsense LoF mutation NM_001080517:c.2918C > G (p.Ser973*). It was found a nonsense SNV mutation NM_0010 80517:c.3212A > G that causes a substitution of a tyrosine by a cysteine aminoacid at position 1,071 of the encoded protein (p.Tyr1071Cys; Szczaluba et al., 2016).

Other two cases of inherited *SETD5* alterations were analyzed (Powis et al., 2017). In a male patient, with global developmental delay, hypotonia, mild

spastic diplegia, dysarthria, and anxiety, received from his unaffected mother, the frameshift variant NM_001080517:c.1655_1656insA. The other case of inherited alteration, it was found a stop gained variant (NM_001080517:c.2955T > A (p.Tyr985*)) in dizygotic male twins with a mildly affected mother. Both twins have cognitive delay, ritualized behaviors, obsessive compulsive disorder and dysmorphic features. In both families, the passed *SETD5* mutation from their corresponding mothers presented high penetrance level in males as observed in other reported male individuals. Furthermore, a normal carrier, the mother suggests that in females, mutations on this gene might have low penetrance. And finally, the nonsynonymous mutation NM_001080517:c.894G > A (Russo et al., 2017), was found in a male with ASD, ADHD, language and cognitive delay, stereotypic behavior and sleep disturbance.

Complementary to this SLR, we also investigated the public databank of clinical variants, ClinVar (Landrum et al., 2014), to check for *SETD5* genetic variants with any clinical significance. After analysis and data selection, we found 16 genetic variants classified as pathogenic or probably pathogenic according to ClinVar classification, and 9 has also been found by our systematic literature review as follow: NM_001080517:c.1195A > T (p.Lys399*), NM_001080517:c.1333C > T (p.Arg445*), NM_001080517:c.3001C > T (p.Arg1001*), NM_001080517:c.2177_2178del (p.Thr726Asnfs*39), NM_001080517:c.3771dup (p.Ser1258Glufs*65), NM_001080517:c.3856del (p.Ser1286Leufs*84), NM_001080517:c.2302C > T, p.(Arg768*), NM_001080517:c.2347-7A > G p.(Arg783Leufs*2) and NM_001080517:c.3246delT (p.Ala1083Leufs) (Table 3).

Seven of these variants were found to be pathogenic with clinical significance. Among them, four presented frameshift mutations: NM_001080517:c.1993delC (p.Leu1775Terfs), NM_001080517:c.3266–3267delCT (p.Ser1089Cysfs) and NM_001080517:c.3783dupC (p.Phe1262Leufs). One of the *SETD5* variant was found in the acceptor splice gene region (NM_001080517:c.1783-2A > T), and another was a 141 nucleotide deletion in total, corresponding to the elimination of 81 nucleotides of the *SETD5* gene, that covers seven codons of exon 7, and 60 nucleotides of the adjacent intron 7 (NM_001080517:c.547–567 + 60del). Finally, two variants of probable pathogenic impact were found: one nonsense SNV, NM_001080517:c.3277A > T (p.Lys1093Ter), and one frameshift deletion, NM_001080517:c.3949delA (p.Thr1317Hisfs).

Within the variants found in ClinVar, we did not found gender information, and only two were associated with some clinical condition: the deletion NM_001080517:c.547–567 + 60del was associated with mental retardation and the SNV NM_001080517:c.1783-2A > T was associated with autosomal dominant 23 disease.

Large deletions (>100 bp) are also reported that removed out part or the entire *SETD5* gene. In one female case, a 54 kb deletion that removed part of the *SETD5* gene between introns 2 and 19, resulting in the loss of the first 16 exons of the corresponding coding region (Parenti et al., 2017). Another type of large deletion involved a chromosomic region not restricted to the *SETD5* gene *locus*. One female case reported a deletion of 643 kb (Peltekova et al., 2012), that depleted the following genes (position chr3: 9,392,274–10,035,209): *THUMPD3, SETD5, LHFPL4, MTMR14, CPNE9, BRPF1, OGG1, CAMK1, TADA3, ARPC4, TTLL3, CIDEC, JAGN1, IL17RE, CRELD,* and *PRRT3*. Another reported female case presented a deletion of 684 kb (chr3: 9,005,098–9,689,733), that depleted the following genes: *SRGAP3, THUMPD3, SETD5,* and *LHFPL4* (Kellogg et al., 2013). In another female case, it was reported a deletion of 1.6 Mb (chr3: 8,330,426–9,910,334; Gunnarsson & Bruun, 2010), that depleted the 16 genes as follow: *LMCD1, CAV3, OXTR, SRGAP3, SETD5, THUMPD3, LHFPL4, MTMR14, BRPF1, OGG1, CAMK1, TADA3L, ARPC4, TTLL3, RPUSD3,* and *CIDEC*.

These reported deletions of 643, 684, and 1.6 Mb, share a common region with the following four reported deletions (Kuechler et al., 2015): deletions of 148 kb (4 genes), 371 kb (10 genes), 2.45 Mb (46 genes), and 11.16 Mb (71 genes). In all these female cases, the common deleted region is on the chromosome 3, between position 9,422,487–9,542,885, composed by the three genes: *SETD5*, *THUMPD3,* and *LHFPL4*.

Finally, another female case presenting a large deletion of 1.24 Mb (Riess et al., 2012) containing seven genes (chr3: 8,275,541–9,516,586), which differed from the previously described, did not delete the *LHFPL4* gene but the genes *SETD5* and *THUMPD3* along with the following genes: *LMCD1*, *CAV3*, *OXTR, RAD18,* and *SRGAP3*.

Recently, a boy with a peculiar-appearing, delayed psychomotor development and speech and severe intellectual disability was described (Yagasaki et al., 2017) with a case of 3p25 distal deletion. The deletion of 10.1 Mb covering the *locus* 3p26.3p25.3 (chr3:1 – 10,142,919) occurs in a region that contains genes such as *CHL1 (CALL), CTCL4, LRRN1, ITPR1, SRGAP3*, and *SETD5*.

**Table 4 Variants LoF in SETD5 gene found on EXAC database described in healthy humans**

| *Locus* | Variant | Aminoacid change | Mutation type | MAF |
|---|---|---|---|---|
| chr3:9,506,250 | c.2618C>A | p.Ser873Ter | stop gained | 0.000008280 |
| chr3:9,517,551 | c.4107_4110dupGAGT | p.Ser1371GlufsTer10 | Frameshift | 0.000008282 |
| chr3:9,517,551 | c.4106dupT | p.Ser1370GlufsTer10 | Frameshift | 0.000008282 |
| chr3:9,517,687 | c.4243_4244delCC | p.Pro1415ThrfsTer38 | Frameshift | 0.000058260 |
| chr3:9,517,722 | c.4276C>T | p.Arg1426Ter | stop gained | 0.000008716 |

*Locus* – correspond to genomic location of the mutation based on human reference genome, build Hg19; Variant – indicates the nucleotide change for the transcript NM_ 001080517; Aminoacid change – indicates the position where an aminoacid was changed within the protein coded by the transcript NM_ 001080517; Mutation type – indicated the protein translation impact of the reported variant, if stop gained or frameshift; MAF – indicates the Minimum Allele Frequency for the described mutation.

To investigate whether mutations of *SETD5* gene could also be affecting the same genomic *loci* of SETD5-AS1 transcript, we analyzed different genetic databanks such as EXAC, GeneCards, EMBL and OMIM, but did not found loss of function mutations. We also performed a SLR to find an association between SETD5-AS1 transcript and ASD, but there is no research describing such correlation. In our SLR we report nine large deletions involving *SETD5* gene, that is also occurring in *SETD5-AS1* transcript (deletions of 643 kb, 684 kb, 1.6 Mb, 148 kb, 371 kb, 2.45 Mb, 11.16 Mb, 1.24 Mb, and 10.1 Mb; Table 1). The function of this transcript is still unknown in human, and functional experiments are required to correlate malfunctions in the transcript with ASD.

To verify whether the reported *SETD5* variants of this SLR are common in the typical development individuals (TDI) population, we analyzed the EXAC (Exome Aggregation Consortium) databank (Lek et al., 2016) and did not found the same mutations within the database. Furthermore, we also investigated whether LoF mutations are present on TDI and found five rare mutations associated with stop gain or frameshift mutations (Table 4). However, none of these variants have any functional studies describing their impact in cellular or molecular level that could affect cellular phenotype.

In all the reported cases of large deletions (>100 bp), *SETD5* gene *locus* was totally or partially deleted. Both, the deletion and SNV mutations in the *SETD5* gene, were found in patients sharing common clinical symptoms, including developmental delay, ID and/or behavioral deficiency problems. An interesting observation are the eight terminal deletions ranging from ~6 to 12 Mb in length (Shuib et al., 2009). In seven of these deletions the 3p.25.2-pter and 3p.25.3-pter cytobands were identified in patients with ID. Interestingly, the eighth deletion found covered the chromosomal band 3p26.1-pter that did not delete the *SETD5* gene *locus*, and the considered patient did not present clinical evidences of ID, supporting the raised evidences here that *SETD5* gene play a significant role on the development of this clinical symptom.

## MOLECULAR INTERACTIONS OF *SETD5*

As reported in a previous section, different mutations underlying *SETD5* gene are leading to similar phenotypes. This observation suggests that, at molecular level, this gene play an important role in development and maintenance of the control nervous system. Although still unknown the function of *SETD5* gene, current findings suggest that SETD5-encoded protein orchestrate gene transcription and expression that were previously associated with ASD and ID. At molecular level, this protein contains a SET domain, which is often found in nuclear proteins that interact with nuclear chromatin (Jenuwein et al., 1998; Bienz, 2006; Kuechler et al., 2015). This interaction can change the chromatin structure and organization and could lead to the modulation of gene expression, possibly causing ASD-related disorders (Balan et al., 2014). *SETD5* is highly expressed in several mouse and human tissues, including intestine, eye, and cerebral cortex in adult mouse tissues (Kuechler et al., 2015) and, in humans, is highly expressed in the brain (Nagase et al., 2000).

Current findings suggest the *SETD5*-encoded protein orchestrate gene transcription and expression that act as essential elements on development and maintenance of the nervous system, the affected tissue by ASD and ID. Analysis of encoded protein demonstrates that SETD5 contain a SET domain, which is often found in nuclear proteins that interact with chromatin (Jenuwein et al., 1998; Bienz, 2006; Kuechler et al., 2015). The *SETD5* gene size is around 82Kb and produces a protein with 1,442 amino acid residues. The protein has a predicted mass of 158 kDa. SETD5 is highly expressed in several mouse and human tissues, including intestine,

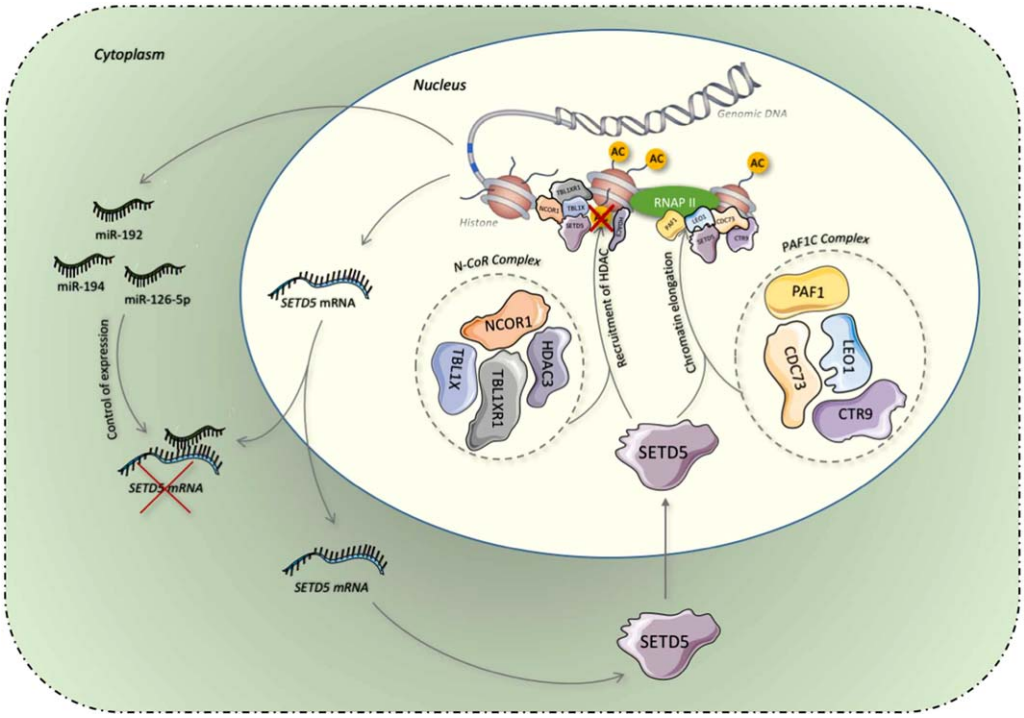


**Figure 4** Molecular interactions of SETD5 with microRNAs, N-CoR and PAF1C complex proteins. The expression of Setd5 in mouse can be regulated through the interaction of miRNAs miR-126-5p, miR-194, and miR-192 (Poissonnier et al., 2014). The interaction of *SETD5* with the HDAC-containing complex N-CoR in human, suggests that *SETD5* might be necessary for the recruitment of HDAC proteins. This interaction leads to the depletion of chromatin acetylation marks as the RNAP II proceeds to the elongation stage and moves towards the downstream region of transcribed genes. *SETD5* interactions with PAF1C and N-CoR complexes are involved in development of cell lineages, regulation of chromatin accessibility, maintenance of pluripotency, myocyte, and vascular lineage specification (Osipovich et al., 2016). [Color figure can be viewed at wileyonlinelibrary.com]

eye, and cerebral cortex in adult mouse tissues (Kuechler et al., 2015) and, in humans, is highly expressed in the brain (Nagase et al., 2000).

Current findings on SETD5 activity postulates that by interacting with the polymerase-associated factor 1 complex (PAF1C) and the nuclear receptor co-repressor (N-CoR), SETD5 may be involved in chromatin accessibility during transcription (Kim et al., 2012; Rincon-Arano et al., 2012; Osipovich et al., 2016; Yu et al., 2017).

The PAF1C is a multi-functional protein complex with a diversity of functions, such as communication with transcriptional activators, recruitment and activation of histone modification factors and the association of transcriptional elongation, cleavage and polyadenylation factors with RNA polymerase II (RNAP II; Jaehning, 2010). SETD5 also interacts with proteins of the PAF1C protein complex (Fig. 4): PAF1 (PAF1 Homolog, Paf1/RNA Polymerase II Complex Component), CTR9 (CTR9 Homolog, Paf1/RNA Polymerase II Complex Component), LEO1 (LEO1 Homolog, Paf1/RNA Polymerase II Complex Component), and CDC73 (Cell Division Cycle 73; Osipovich et al., 2016). In this HEK 293T cell model experiment, it was also found that SETD5 interacts with certain factors of the N-CoR repressor complex (Yoon et al., 2003; Fig. 4), such as NCOR1 (Nuclear Receptor Corepressor 1), HDAC3 (Histone Deacetylase 3), TBL1X (Transducin Beta Like 1 X-Linked Receptor 1), and TBL1XR1 (Transducin Beta Like 1 X-Linked Receptor 1; Osipovich et al., 2016). The N-CoR complex is known to mediate transcriptional repression by promoting chromatin condensation (Mottis et al., 2013), and its essential for the development of multiple organs such as the nervous system and the heart (Jepsen et al., 2000; Perissi et al., 2010). Moreover, near the transcription start site of genes, SETD5 is required by the N-CoR complex that contain histone deacetylases (HDAC) proteins. This interaction leads to the depletion of chromatin

**Table 5 Related physical protein interaction with *SETD5* curated by BioGRID (Chatr-Aryamontri et al., 2017)**

| Symbol | Description | Subcellular locations | Related Pathways | Dataset |
|---|---|---|---|---|
| KRT40 | Keratin 40 | Cytosol, cytoskeleton, nucleus. | Keratinization; Developmental Biology. | (Rual et al., 2005) |
| LDOC1 | Leucine Zipper Down-Regulated in Cancer 1 | Nucleus, extracellular, mitochondrion, endoplasmatic reticulum, cytosol. | - | (Venkatesan et al., 2009) |
| PRR20E | Proline Rich 20E | - | - | (Huttlin et al., 2017) |
| NXF1 | Nuclear RNA Export Factor 1 | Nucleus, cytosol. | Transport of the SLBP independent Mature mRNA; Transport of Mature Transcript to Cytoplasm. | (Castello et al., 2012) |
| NAGK | N-Acetylglucosamine Kinase | Extracellular, cytosol, cytoskeleton, peroxisome, nucleus. | Transport to the Golgi and subsequent modification; Amino sugar and nucleotide sugar metabolism. | (Huttlin et al., 2017) |
| SKAP1 | Src Kinase Associated Phosphoprotein 1 | Plasma membrane, nucleus, cytosol, cytoskeleton, mitochondrion, lysosome. | Ras signaling pathway; B cell receptor signaling pathway. | (Huttlin et al., 2017) |
| TRAF2 | TNF Receptor Associated Factor 2 | Nucleus, endoplasmic reticulum, cytosol, plasma membrane. | Regulation of activated PAK-2p34 by proteasome mediated degradation; Immune System. | (Rual et al., 2005) |
| MTUS2 | Microtubule Associated Tumor Suppressor Candidate 2 | Nucleus, cytoskeleton, cytosol, mitochondrion. | - | (Rolland et al., 2014) |
| CCDC85B | Coiled-Coil Domain Containing 85B | Cytoskeleton, nucleus, extracellular, cytosol. | - | (Rual et al., 2005) |
| PRMT2 | Protein Arginine Methyltransferase 2 | Nucleus, cytosol, plasma membrane, extracellular, cytoskeleton, mitochondrion. | RB in Cancer; mRNA Splicing. | (Huttlin et al., 2015, 2017) |
| CEP70 | Centrosomal Protein 70 | Cytoskeleton, cytosol, nucleus, plasma membrane. | Regulation of PLK1 Activity at G2/M Transition; Organelle biogenesis and maintenance. | (Rolland et al., 2014) |
| DPPA2 | Developmental Pluripotency Associated 2 | Nucleus, cytosol, cytoskeleton. | - | (Rual et al., 2005) |

**E**ach protein is described by: Symbol – correspond to the know protein symbol; Description – describe the protein abbreviation; Subcellular locations – information about subcellular compartments of protein provided by GeneCards (Stelzer et al., 2016); Related Pathways – main related pathway from Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto, 2000); Dataset – reported protein interaction dataset.

acetylation marks, with the RNAP II moving towards an elongation stage through the downstream region of transcribed genes (Osipovich et al., 2016; Fig. 4).

The N-CoR complex has already been associated with genes known to cause some ASD related conditions, such as *MeCP2* (methyl CpG binding protein 2), implicated in RTT. MeCP2 can interact with DNA and N-CoR/SMRT co-repressors and the loss of this interaction can contribute to aberrant brain function and RTT phenotypes, such as impairments in mobility, hindlimb clasp and tremor (Lyst et al., 2014).

SETD5 can also interacts with several other molecules, including proteins, RNA and microRNAs. Several of these proteins, in which interaction with SETD5 has unknown function, were identified (Table 5). Other interacting molecules demonstrating cellular/molecular phenotype were found using mouse models, but none of them were associated with the brain. In *Setd5* knockout mouse embryonic stem cells (mESCs), genes important for specifying cellular lineages and for embryonic morphogenesis, gastrulation, pattern specification and regionalization such as *Mixl1*(Mix Paired-Like Homeobox), *Mesp1* (Mesoderm Posterior BHLH Transcription Factor 1), *Flt1* (Fms Related Tyrosine Kinase 1), *Gsc* (Goosecoid Homeobox), and *T* (T Brachyury Transcription Factor) were upregulated, while pluripotency genes such as *Klf4* (Kruppel Like Factor 4), *Nanog* (Nanog Homeobox), and *Sox2* (SRY-Box 2) were downregulated (Osipovich et al., 2016). Together, these findings suggest that *Setd5* might also contribute to the maintenance of pluripotency in mice. Moreover, in these same *Setd5* knockout mouse cell lines, it was also found that genes involved in myocyte and vascular lineage specification, such as *Mixl1* and *Mesp1*, had altered expression, contributing to abnormalities in cardiovascular phenotypes (Osipovich et al., 2016). Additionally, it has also been shown that, in mice, the Setd5 interacts with different micro RNAs (miRNA) such as the following: miR-126-5p, miR-194, and miR-192 (Fig. 4). The miRNA miR-126-5p is mostly expressed in blood vessel endothelial cells and participates in the control of leucocyte trafficking by regulating the expression of the *Setd5* gene (Poissonnier et al., 2014). Although with unknown significance, *Setd5* gene is a target for the miRNAs miR-194 and miR-192, which are also found in humans (Morimoto et al., 2017).

## SETD5 SYNDROME AS A CHROMATINOPATHY DISORDER

Chromatinopathies are classified as mutation on chromatin-associated epigenetic modifiers genes that leads to the development of neurodevelopmental disorders such as ASD. These genes are very known as the cause for several ASD-related syndromes, such as *MECP2*, *KMT2A* (Lysine Methyltransferase 2A) that is associated with Wiedemann-Steiner syndrome and others. The *MECP2* gene act as a transcriptional regulator and controls genomic stability (Nan et al., 1998). Loss of *MECP2* leads to impaired sociability, motor abnormalities, cognitive defects (Marchetto et al., 2010; Veeraragavan et al., 2015). When duplication of *MECP2* occurs, it also causes several autistic like symptoms such as severe neurodevelopmental delay with limited or absent speech, motor dysfunction and autistic behavior (Nageshappa et al., 2016). *KMT2A* is a transcriptional coactivator gene that encodes multiple proteins with conserved functional domains, such as SET domain. This domain plays a role in methylatransferase activity on histone H3 lysine 4 (H3K4), being able to mono- di- or trimethylation (Wiersma et al., 2016). Mutations in *KMT2A* were identified in individuals with Wiedemann–Steiner syndrome, which leads a developmental disorder including ID, microcephaly, short stature and autism features (Dunkerton et al., 2015).

*SETD5,* as well as *MECP2* and *KMT2A*, are epigenetic readers responsible for the link between DNA methylation and histone modifications. Mutation on those genes could have a severe impact on gene activation, leading to ID and autistic features, some characteristics of chromatinopathies.

## CONCLUSION

In summary, all mutations found on *SETD5* gene by our SLR and public databank analysis are associated with a wide spectrum of ID and ASD-like symptoms, with a high penetrance in males (disease-causing mutation). In females, although most of reported cases also presented high penetrance, we found two cases of normal female carriers that segregated the mutation to their corresponding boy children that presented ID and ASD-like phenotypes, enforcing the idea of *SETD5* mutations as disease-causing in males.

Most of mutations, when described, are loss of function deletions representing around 43% of all identified genetic variations found in this gene. However, it is still unclear how SETD5 lead to ASD phenotypes in humans. More individuals with mutations on *SETD5* will provide a better understanding of the SETD5 role in the development of the nervous system and on correlating in both gender the penetrance level of *SETD5* gene mutations with ASD symptoms.

Dr. Muotri is supported by U19MH107367, part of the National Cooperative Reprogrammed Cell Research Groups (NCRCRG) to Study Mental Illness and by a NARSAD Independent Investigator Award. Dr. Fernandes is supported by a post-doctoral fellowship from the Autism Science Foundation. Ana C. P. Cruz is recipient of a master fellowship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Adriano Ferrasa is recipient of a doctoral fellowship from CAPES/FA (Fundação Araucária). RHH is supported by FA (grant #CP09/2016).

## CONFLICT OF INTEREST

The authors declare that they have no competing interests.

## LITERATURE CITED

Balan S, Iwayama Y, Maekawa M, Toyota T, Ohnishi T, Toyoshima M, Shimamoto C, et al. 2014. Exon resequencing of H3K9 methyltransferase complex genes, EHMT1, EHTM2 and WIZ, in Japanese autism subjects. Mol Autism 5:49.

Bienz M. 2006. The PHD finger, a nuclear protein-interaction domain. Trends Biochem Sci 31:35–40.

Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, et al. 2012. Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. Cell 149:1393–1406.

Center for Disease Control and Prevention. Fragile X Syndrome (FXS) – Data & Statistics. https://www.cdc.gov/ncbddd/fxs/data.html. December 5, 2017.

Chatr-Aryamontri A, Oughtred R, Boucher L, Rust J, Chang C, Kolas NK, et al. 2017. The BioGRID interaction database: 2017 update. Nucleic Acids Res 45:D369–D379.

Dunkerton S, Field M, Cho V, Bertram E, Whittle B, Groves A, Goel H. 2015. A de novo Mutation in KMT2A (MLL) in monozygotic twins with Wiedemann-Steiner Syndrome. Am J Med Genet Part A 167:2182–2187.

Geschwind DH, Levitt P. 2007. Autism spectrum disorders: Developmental disconnection syndromes. Curr Opin Neurobiol 17:103–111.

Green C, Willoughby J, DDD S, Balasubramanian M. 2017. De novo SETD5 loss-of-function variant as a cause for intellectual disability in a 10-year old boy with an aberrant blind ending bronchus. Am J Med Genet A 173: 3165–3171.

Gilman SR, Iossifov I, Levy D, Ronemus M, Wigler M, Vitkup D. 2011. Rare De Novo Variants Associated with Autism Implicate a Large Functional Network of Genes Involved in Formation and Function of Synapses. Neuron 70:898–907.

Grozeva D, Carss K, Spasic-Boskovic O, Parker MJ, Archer H, Firth HV, Park S-M, et al. 2014. De novo loss-of-function mutations in SETD5, encoding a methyltransferase in a 3p25 microdeletion syndrome critical region, cause intellectual disability. Am J Hum Genet 94:618–624.

Grozeva D, Carss K, Spasic-Boskovic O, Tejada M-I, Gecz J, Shaw M, Corbett M, et al. 2015. Targeted next-generation sequencing analysis of 1,000 individuals with intellectual disability. Hum Mutat 36:1197–1204.

Gunnarsson C, Bruun CF. 2010. Molecular characterization and clinical features of a patient with an interstitial deletion of 3p25.3-p26.1. Am. J. Med. Genet 152A:3110–3114.

Huttlin EL, Bruckner RJ, Paulo JA, Cannon JR, Ting L, Baltier K, Colby G, et al. 2017. Architecture of the human interactome defines protein communities and disease networks. Nature 545:505–509.

Huttlin EL, Ting L, Bruckner RJ, Gebreab F, Gygi MP, Szpyt J, Tam S, et al. 2015. The BioPlex network: A systematic exploration of the human interactome. Cell 162: 425–440.

Jaehning JA. 2010. Gene regulatory mechanisms. Biochim Biophys Acta 1799: 353–486.

Jenuwein T, Laible G, Dorn R, Reuter G. 1998. SET domain proteins modulate chromatin domains in eu- and heterochromatin. Cell Mol Life Sci 54:80–93.

Jepsen K, Hermanson O, Onami TM, Gleiberman AS, Lunyak V, McEvilly RJ, Kurokawa R, et al. 2000. Combinatorial Roles of the Nuclear Receptor Corepressor in Transcription and Development. Cell 102:753–763.

Kanehisa M, Goto S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27–30.

Kellogg G, Sum J, Wallerstein R. 2013. Deletion of 3p25.3 in a individuals with intellectual disability and dysmorphic features with further definition of a critical region. Am J Med Genet Part A 161:1405–1408.

Kim T, Xu Z, Clauder-Münster S, Steinmetz LM, Buratowski S. 2012. Set3 HDAC Mediates Effects of Overlapping Noncoding Transcription on Gene Induction Kinetics. Cell 150:1158–1169.

Kobayashi Y, Tohyama J, Kato M, Akasaka N, Magara S, Kawashima H, Ohashi T, et al. 2016. High prevalence of genetic alterations in early-onset epileptic encephalopathies associated with infantile movement disorders. Brain Dev 38:285–292.

Kuechler A, Zink AM, Wieland T, Lüdecke H-J, Cremer K, Salviati L, Magini P, et al. 2015. Loss-of-function variants of SETD5 cause intellectual disability and the core phenotype of microdeletion 3p25.3 syndrome. Eur J Hum Genet 23:753–760.

Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, et al. 2014. ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res 42:980–985.

Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, et al. 2016. Analysis of protein-coding genetic variation in 60,706 humans. Nature 536:285–291.

Lyst MJ, Ekiert R, Ebert DH, Merusi C, Nowak J, Guy J, Kastan NR, et al. 2014. Rett syndrome mutations abolish the interaction of MeCP2 with the NCoR/SMRT co-repressor. Nat Neurosci 16:1–14.

Marchetto MCN, Carromeu C, Acab A, Yu D, Yeo GW, Mu Y, Chen G, Gage FH, Muotri AR. 2010. A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. Cell 143:527–539.

Morimoto A, Kannari M, Tsuchida Y, Sasaki S, Saito C, Matsuta T, Maeda T, et al. 2017. An HNF4α-microRNA-194/192 signaling axis maintains hepatic cell function. J Biol Chem 292:10574–10585.

Mottis A, Mouchiroud L, Auwerx J. 2013. Emerging roles of the corepressors NCoR1 and SMRT in homeostasis. Genes Dev 27:819–835.

Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A. 1998. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 393:386–389.

Nagase T, Kikuno R, Hattori A, Kondo Y, Okumura K, Ohara O. 2000. Prediction of the coding sequences of unidentified human genes, XIX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. DNA Res 7:347–355.

Nageshappa S, Carromeu C, Trujillo CA, Mesci P, Pasciuto E, Vanderhaeghen P, Raitano S, et al. 2016. HHS Public Access. 21:178–188.

Osipovich AB, Gangula R, Vianna PG, Magnuson MA. 2016. Setd5 is essential for mammalian development and the co-transcriptional regulation of histone acetylation. Development 143:4595–4607.

Parenti I, Teresa-Rodrigo ME, Pozojevic J, Ruiz Gil S, Bader I, Braunholz D, Bramswig NC, et al. 2017. Mutations in chromatin regulators functionally link Cornelia de Lange syndrome and clinically overlapping phenotypes. Hum Genet 136:307–320.

Peltekova IT, Macdonald A, Armour CM. 2012. Microdeletion on 3p25 in a patient with features of 3p deletion syndrome. Am J Med Genet Part A 158A:2583–2586.

Perissi V, Jepsen K, Glass CK, Rosenfeld MG. 2010. Deconstructing repression: Evolving models of co-repressor action. Nat Rev Genet 11:109–123.

Pinto D, Delaby E, Merico D, Barbosa M, Merikangas A, Klei L, Thiruvahindrapuram B, et al. 2014. Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. Am J Hum Genet 94:677–694.

Poissonnier L, Villain G, Soncin F, Mattot V. 2014. MiR126-5p repression of ALCAM and SetD5 in endothelial cells regulates leucocyte adhesion and transmigration. Cardiovasc Res 102:436–447.

Popp B, Ekici AB, Thiel CT, Hoyer J, Wiesener A, Kraus C, Reis A, Zweier C. 2017. Exome Pool-Seq in neurodevelopmental disorders. Eur J Hum Genet 25: 1364–1376.

Powis Z, Farwell Hagman KD, Mroske C, McWalter K, Cohen JS, Colombo R, Serretti A, et al. 2017. Expansion and further delineation of the SETD5 phenotype leading to global developmental delay, variable dysmorphic features, and reduced penetrance. Clin Genet doi: 10.1111/cge.13132. [Epub ahead of print]

Richards C, Jones C, Groves L, Moss J, Oliver C. 2015. Prevalence of autism spectrum disorder phenomenology in genetic disorders: A systematic review and meta-analysis. Lancet Psychiatry 2:342.

Riess A, Grasshoff U, Schäferhoff K, Bonin M, Riess O, Horber V, Tzschach A. 2012. Interstitial 3p25.3-p26.1 deletion in a patient with intellectual disability. Am J Med Genet Part A 158A:2587–2590.

Rincon-Arano H, Halow J, Delrow JJ, Parkhurst SM, Groudine M. 2012. UpSET recruits HDAC complexes and restricts chromatin accessibility and acetylation at promoter regions. Cell 151:1214–1228.

Rolland T, Taşan M, Charloteaux B, Pevzner SJ, Zhong Q, Sahni N, Yi S, et al. 2014. A proteome-scale map of the human interactome network. Cell 159:1212–1226.

Rual J-F, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, Li N, Berriz GF, et al. 2005. Towards a proteome-scale map of the human protein–protein interaction network. Nature 437:1173–1178.

Russo FB, Freitas BC, Pignatari GC, Fernandes IR, Sebat J, Muotri AR, Beltrão-Braga PCB. 2017. Modeling the interplay between neurons and astrocytes in autism using human induced pluripotent stem cells. Biol Psychiatry doi: 10.1016/j.biopsych.2017.09.021. [Epub ahead of print]

Schiff M, Benoist JF, Aïssaoui S, Boespflug-Tanguy O, Mouren MC, de Baulny HO, Delorme R. 2011. Should metabolic diseases be systematically screened in nonsyndromic autism spectrum disorders? PLoS One 6:e21932.

Schrier SA, Sherer I, Deardorff MA, Clark D, Audette L, Gillis L, Kline AD, et al. 2011. Causes of death and autopsy findings in a large study cohort of individuals with Cornelia de Lange syndrome and review of the literature. Am J Med Genet A 155A:3007–3024.

Shuib S, McMullan D, Rattenberry E, Barber RM, Rahman F, Zatyka M, Chapman C, et al. 2009. Microarray based analysis of 3p25-p26 deletions (3p- syndrome). Am J Med Genet Part A 149A:2099–2105.

Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, Stein TI, et al. 2016. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. Current Protocols in Bioinformatics, Hoboken, NJ: John Wiley & Sons, Inc., p 1.30.1–1.30.33.

Stur E, Soares LA, Louro ID. 2017. SETD5 gene variant associated with mild intellectual disability – A case report. Genet Mol Res 16:2.

Szczaluba K, Brzezinska M, Kot J, Rydzanicz M, Walczak A, Stawinski P, Werner B, et al. 2016. SETD5 loss-of-function mutation as a likely cause of a familial syndromic intellectual disability with variable phenotypic expression. Am J Med Genet Part A 170:2322–2327.

Sztainberg Y, Zoghbi HY. 2016. Lessons learned from studying syndromic autism spectrum disorders. Nat Neurosci 19:1408–1418.

Veeraragavan S, Wan YW, Connolly DR, Hamilton SM, Ward CS, Soriano S, Pitcher MR, et al. 2015. Loss of MeCP2 in the rat models regression, impaired sociability and transcriptional deficits of Rett syndrome. Hum Mol Genet 25:3284–3302.

Venkatesan K, Rual J-F, Vazquez A, Stelzl U, Lemmens I, Hirozane-Kishikawa T, Hao T, et al. 2009. An empirical

framework for binary interactome mapping. Nat Methods 6:83–90.

Wiersma M, Bussiere M, Halsall JA, Turan N, Slany R, Turner BM, Nightingale KP. 2016. Protein kinase Msk1 physically and functionally interacts with the KMT2A/MLL1 methyltransferase complex and contributes to the regulation of multiple target genes. Epigenetics Chromatin 9:52.

Yagasaki H, Toda T, Koizumi K, Sugiyama T, Ohyama T, Hoshiai M, Nakane T, et al. 2017. A de novo 10.1-Mb 3p25 terminal deletion including SETD5 in a patient with ptosis and psychomotor retardation. Pediatr Neonatol doi: 10.1016/j.pedneo.2017.09.004. [Epub ahead of print]

Yoon HG, Chan DW, Huang ZQ, Li J, Fondell JD, Qin J, Wong J. 2003. Purification and functional characterization of the human N-CoR complex: The roles of HDAC3, TBL1 and TBLR1. Embo J 22:1336–1346.

Yu SE, Kim MS, Park SH, Yoo BC, Kim KH, Jang YK. 2017. SET domain-containing protein 5 is required for expression of primordial germ cell specification-associated genes in murine embryonic stem cells. Cell Biochem Funct 35:1–7.