

# Study of chromatin remodeling genes implicates SMARCA4 as a putative player in oncogenesis in neuroblastoma

Angela Bellini<sup>1,2,3\*</sup>, Nadia Bessoltane-Bentahar<sup>1,2,3\*</sup>, Jaydutt Bhalshankar<sup>1,2,3</sup>, Nathalie Clement<sup>1,2,3</sup>, Virginie Raynal<sup>4,5</sup>, Sylvain Baulande<sup>6</sup>, Virginie Bernard<sup>7</sup>, Adrien Danzon<sup>1,2,3</sup>, Mathieu Chicard<sup>1,2,3</sup>, Léo Colmet-Daage<sup>1,2,3</sup>, Gaëlle Pierron<sup>8</sup>, Laura Le Roux<sup>8</sup>, Julien M. Planchon<sup>3,7</sup>, Valérie Combaret<sup>9</sup>, Eve Lapouble<sup>3,8</sup>, Nadège Corradini<sup>10</sup>, Estelle Thebaud<sup>11</sup>, Marion Gambart<sup>12</sup>, Dominique Valteau-Couanet<sup>13</sup>, Jean Michon<sup>3</sup>, Caroline Louis-Brennetot<sup>2,3,4</sup>, Isabelle Janoueix-Lerosey<sup>1b2,3,4</sup>, Anne-Sophie Defachelles<sup>14</sup>, Franck Bourdeaut<sup>1,2,3</sup>, Olivier Delattre<sup>2,3,4,8</sup> and Gudrun Schleiermacher<sup>1b1,2,3</sup>

<sup>1</sup>Equipe SiRIC RTOP Recherche Translationnelle en Oncologie Pédiatrique, Institut Curie, Paris, France

<sup>2</sup>INSERM U830, Laboratoire de Génétique et Biologie des Cancers, Institut Curie, Paris, France

<sup>3</sup>SIREDO: Care, Innovation and Research for Children, Adolescents and Young Adults with Cancer, Institut Curie, Paris, France

<sup>4</sup>Institut Curie, PSL Research University, Inserm U830, Equipe Labellisée Ligue contre le Cancer, Paris, France

<sup>5</sup>Plateforme de Séquençage ICGex, Institut Curie, Paris, France

<sup>6</sup>Institut Curie, PSL Research University, NGS Platform, Paris, France

<sup>7</sup>Department of Biopathology, Institut Curie, PSL Research University, Paris, France

<sup>8</sup>Unité de Génétique Somatique, Institut Curie, Paris, France

<sup>9</sup>Laboratoire de Recherche Translationnelle, Centre Léon-Bérard, Lyon, France

<sup>10</sup>Institut d'Hématologie et Oncologie Pédiatrique, Lyon, France

<sup>11</sup>Service d'Oncologie Pédiatrique, Hôpital de la Mère et l'enfant, Nantes, France

<sup>12</sup>Unité d'Hémo-Oncologie, Hôpital des Enfants, Toulouse, France

<sup>13</sup>Service d'Oncologie Pédiatrique, Institut Gustave Roussy, Paris, France

<sup>14</sup>Pediatric Oncology Unit, AntiCancer Center Oscar Lambret, Lille, France

In neuroblastoma (NB), genetic alterations in chromatin remodeling (CRGs) and epigenetic modifier genes (EMGs) have been described. We sought to determine their frequency and clinical impact. Whole exome (WES)/whole genome sequencing (WGS) data and targeted sequencing (TSCA<sup>®</sup>) of exonic regions of 33 CRGs/EMGs were analyzed in tumor samples from 283 NB patients, with constitutional material available for 55 patients. The frequency of CRG/EMG variations in NB cases was then compared to the Genome Aggregation Database (gnomAD). The sequencing revealed SNVs/small InDels or focal CNAs of CRGs/EMGs in 20% (56/283) of all cases, occurring at a somatic level in 4 (7.2%), at a germline level in 12 (22%) cases, whereas for the remaining cases, only tumor material could be analyzed. The most frequently altered genes were *ATRX* (5%),

**Key words:** neuroblastoma, chromatin remodeling complex, SWI/SNF, epigenetic modifier, SMARCA4

**Abbreviations:** aCGH: array-comparative genomic hybridization; CNAs: copy number alterations; CRGs: chromatin remodeling genes; EMGs: epigenetic modifiers genes; OS: overall survival; PFS: progression-free survival; SNVs: single nucleotide variants; TSCA: TruSeq<sup>®</sup> Custom Amplicon; WES: whole exome sequencing; WGS: whole genome sequencing

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

**Conflict of interest:** The authors declare no conflict of interest.

**Grant sponsor:** Agence Nationale pour la Recherche; **Grant number:** ANR-10-INBS-518 09; **Grant sponsor:** Annenberg Foundation;

**Grant sponsor:** Equipex; **Grant number:** ANR-10-EQPX-03; **Grant sponsor:** France Génomique Consortium; **Grant number:** ANR-

10-INBS-09-08; **Grant sponsor:** Association Hubert Gouin Enfance et Cancer; **Grant sponsor:** Association Les amis de Claire;

**Grant sponsor:** Associations Enfants et Santé; **Grant sponsor:** Associations Les Bagouz à Manon; **Grant sponsor:** Canceropole Ile-de-France;

**Grant sponsor:** CEST of Institut Curie; **Grant sponsor:** Fondation Nelia et Amadeo Barletta; **Grant sponsor:** Programme Hospitalier de

recherche en Cancérologie (PHRC); **Grant number:** IC2007-09; **Grant sponsor:** SiRIC/INCa; **Grant number:** INCa-DGOS-4654;

**Grant sponsor:** SiRIC-Curie Program-SiRIC

\*A.B. and N.B.-B. contributed equally to this work

DOI: 10.1002/ijc.32361

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

**History:** Received 16 Nov 2018; Accepted 14 Mar 2019; Online 24 Apr 2019

**Correspondence to:** Gudrun Schleiermacher, Department of Pediatric Oncology, Institut Curie, 26 rue d'Ulm, 75248 Paris Cedex 05, France, Tel.: +33-1-44-32-45-50, E-mail: gudrun.schleiermacher@curie.fr

*SMARCA4* (2.5%), *MLL3* (2.5%) and *ARID1B* (2.5%). Double events (SNVs/small InDels/CNAs associated with LOH) were observed in *SMARCA4* ( $n = 3$ ), *ATRX* ( $n = 1$ ) and *PBRM1* ( $n = 1$ ). Among the 60 variations, 24 (8.4%) targeted domains of functional importance for chromatin remodeling or highly conserved domains but of unknown function. Variations in *SMARCA4* and *ATRX* occurred more frequently in the NB as compared to the gnomAD control cohort (OR = 4.49, 95%CI: 1.63–9.97,  $p = 0.038$ ; OR 3.44, 95%CI: 1.46–6.91,  $p = 0.043$ , respectively). Cases with CRG/EMG variations showed a poorer overall survival compared to cases without variations. Genetic variations of CRGs/EMGs with likely functional impact were observed in 8.4% (24/283) of NB. Our case–control approach suggests a role of *SMARCA4* as a player of NB oncogenesis.

#### What's new?

Mutations that affect chromatin remodeling can lead to cancer. In this paper, the authors investigated the impact of variations in chromatin remodeling genes and epigenetic modifier genes on neuroblastoma patients. They compared the frequency of these variations in NB cases with data from the Genome Aggregation Database (gnomAD). Neuroblastoma cases had a higher frequency of *SMARCA4* and *ATRX* gene variations than the general population. Furthermore, NB patients with CRG/EMG mutations had poorer overall survival than NB cases without such mutations. These findings highlight the importance of chromatin remodeling in neuroblastoma as an avenue for new therapeutics.

## Introduction

Alterations in chromatin remodeling genes (CRGs) have emerged as an important cause of cancer; in particular, the identification of inactivating genetic alterations in the *SMARCB1* gene in rhabdoid tumors (RTs) has highlighted the involvement of the SWI/SNF complex in tumor formation.<sup>1,2</sup> Indeed, SWI/SNF gene mutations and alterations in other epigenetic modifier genes (EMGs) are thought to occur in 20% of human cancers.<sup>3</sup>

In neuroblastoma (NB), the most common extracranial solid tumor in children, genetic alterations concern predominantly copy number alterations, with *MYCN* amplification and segmental chromosome alterations associated with a poor outcome.<sup>4</sup> Only few genes have been shown to be altered recurrently by genetic events, including *ALK*.<sup>4</sup> Genes involved in chromatin remodeling such as *ATRX*, *DAXX*, *ARID1A* and *ARID1B* genes have also been shown to be targeted by mutations or structural rearrangements,<sup>5–9</sup> but the frequency of CRG/EMG alterations has not been described in detail in NB.

The aim of our study was to determine the frequency of CRGs/EMGs variations in NB patients and to correlate findings with clinical parameters and outcome.

## Materials and Methods

### Study series

The study consisted of 283 NB patients, constituting a clinically representative cohort (Table 1, Supporting Information Table S1; Supporting Information Fig. S1).

Patients were treated in French pediatric oncology centers according to the relevant national/international protocols. Written informed consent was obtained from parents/guardians according to national law. Our study was authorized by the ethics committees “Comité de Protection des Personnes Sud-Est IV”, references L07–95/L12–171, and “Comité de Protection des Personnes Ile de France”, reference 0811728.

In 55 patients for whom paired normal and tumor tissue was available, whole-exome sequencing (WES)/whole-genome sequencing (WGS) techniques were performed, whereas for 248 diagnostic tumor samples, a TruSeq<sup>®</sup> Custom Amplicon (TSCA) approach was used. Twenty patients are common to the two subsets (Supporting Information Table S1).

Furthermore, 30 NB cell lines (Supporting Information Table S1) and six germline controls from healthy donors were studied.

To analyze the frequency of CRG/EMG variations, 28 major genes involved in chromatin remodeling processes were selected: *ACTB*, *ACTL6A*, *ACTL6B*, *ARID1A*, *ARID1B*, *ARID2*, *BCL11A*, *BCL11B*, *BCL7A*, *BCL7B*, *BCL7C*, *BRD7*, *BRD9*, *DPF1*, *DPF2*, *DPF3*, *PBRM1*, *PHF10*, *SMARCA2*, *SMARCA4*, *SMARCB1*, *SMARCC1*, *SMARCC2*, *SMARCD1*, *SMARCD2*, *SMARCD3*, *SMARCE1*, *SS18*. Five epigenetic modifier genes, *ATRX*, *DAXX*, *CHD7*, *MLL2* (*KMT2D*), *MLL3* (*KMT2C*), were also selected based to their involvement in a wide variety of human malignancies (Supporting Information Table S2).

### DNA extraction and aCGH analysis

DNA was extracted from diagnostic tumor samples harboring >20% of tumor cells using standard procedures. Genomic copy number profiles were determined by aCGH, and *MYCN* copy number status was confirmed by FISH.

### Whole exome/whole genome sequencing

Paired diagnostic tumor and normal DNA from 55 patients were whole-genome ( $n = 16$ ) and/or whole-exome ( $n = 39$ ) sequenced (Illumina<sup>®</sup> HiSeq2500; average coverage: 80×–100× per sample, respectively<sup>10</sup>). WGS for eight cases has been reported previously.<sup>10</sup>

### TruSeq<sup>®</sup> amplicon panel sequencing

Tumor DNA from 248 NB cases and six healthy donor germline DNA samples (negative controls) were analyzed using a TSCA<sup>®</sup>

panel approach covering 33 CRGs/EMGs, (average coverage >1,500×; TSCA<sup>®</sup>v1.5, Illumina Inc., San Diego, CA; Supporting Information Table S2).

### Bioinformatics analysis

Bioinformatics approaches depended on the genetic alteration (single nucleotide variations [SNVs]), small or large InDels; focal or extended copy number alterations (CNAs) and the sequencing technique.

### Whole-exome and whole-genome sequencing analyses

The WGS/WES sequencing raw reads were mapped to the reference human genome using BWA (assembly GRCh37/hg19)<sup>11,12</sup> followed by analysis of the exons of the 33 CRGs/EMGs (Supporting Information Table S2).

SNVs/InDels <30 bp were called using GATK's Haplotype-Caller v3.5 algorithm.<sup>12,13</sup> The Manta tool (1.0.3) was used to detect structural variations and large InDels.<sup>14</sup> All SNVs/small InDels with a VAF >20% were retained.

Tumor copy number profiles were calculated with FREEC (0.9) using constitutional DNA as reference,<sup>15–17</sup> with annotations of CRGs/EMGs to highlight focal variations.<sup>18</sup>

Genetic variations were termed “somatic” if no evidence of these variations was observed in the constitutional reads.

### TSCA<sup>®</sup> analysis

The TSCA<sup>®</sup> sequencing reads were mapped to the reference human genome with Bowtie.<sup>19</sup> To prevent strand or coverage bias due to noncomplete overlap of forward (FW) and reverse (RV) reads of amplicons, the bam files were split to FW and RV read bam(s).

An adapted approach was developed to enable variant calling in TSCA<sup>®</sup> data.<sup>20</sup>

For SNVs, coverage analysis was performed at each base position (GATK DepthOfCoverage). The background of noise (variability) of each amplicon was then analyzed in the six negative controls.

All possible variants (e.g., ref = A; A > T, A > C, A > G, A > -) and their alternative allele fraction (AF) were calculated for each position, and Fisher's exact two-sided tests with a Bonferroni correction were performed to compare percentages of variant allele fraction (VAF) for a given base between a case and the negative controls. Significant variations were filtered-in in case of an increase in the percentage of variant base (5% significance level). Only positions with total depth of coverage higher than 100× were considered for variant analysis. All SNVs/small InDels with a VAF >5% were retained.

Recurrent variants and variants with VAF >20% detected in more than two samples were filtered-out as they most likely indicate polymorphisms. Finally, FW and RV results were confronted and ambiguous cases were filtered-out.

Amplicon InDels Hunter tool was used to detect large InDels.<sup>21</sup> Recurrent InDels detected in more than two samples were filtered-out.

For focal CNAs, the depth of coverage of targeted regions was calculated for FW and RV bam (GATK DepthOfCoverage). For each sample, the coverage of each amplicon was normalized by the median of sample depth followed by comparison to negative controls. The FW and RV normalized data were then merged to reconstruct copy number profiles. The copy number profile obtained from TSCA data using our in-house developed pipeline was also confirmed by another technique ONCOCNV, a method that includes a multifactor normalization with respect to library size, CG-content and target length to detection of large copy number changes from amplicon deep sequencing data.<sup>22</sup>

Finally, the list of SNVs and InDels was manually curated and visually inspected in IGV (Integrated Genome Viewer).<sup>23</sup>

### Variant annotation, prioritization and classification

All SNVs/InDels were annotated using SnpEff/SnpSift.<sup>24</sup> Annotation was performed using canonical transcripts and publicly available databases dbSNPv150,<sup>25</sup> gnomADv2.0.2 (<http://gnomad.broadinstitute.org/>),<sup>26</sup> COSMICv74,<sup>27</sup> ClinVar(25-02-2018),<sup>28</sup> dbNSF Pv2.9.3<sup>18</sup> and known cancer Hotspots.<sup>29</sup> Loss of function (LoF) assessment was performed using the snpEff tool. SNVs/InDels were discarded as benign/likely benign/polymorphic based on population minor allele frequency (popmax) ≥0.1% from gnomAD database. However, SNVs/inDels reported as pathogenic/likely pathogenic in ClinVar were retained. In addition to COSMIC annotations, six prediction algorithms VEST3,<sup>30</sup> MetaSVM,<sup>31</sup> REVEL,<sup>32</sup> Mutation Taster,<sup>33</sup> M\_CAP<sup>34</sup> and CADD<sup>35</sup> were used to classify all likely protein-altering SNVs/InDels (frameshift, nonsense, splice donor/acceptor, nonsynonymous and in-frame insertions/deletions) as pathogenic, likely pathogenic. The flowchart of variant classification is indicated in Supporting Information Figure S4A and the prediction score thresholds in Supporting Information Figure S4B.

### SNVs/small InDels validation

All retained SNVs/small InDels with VAF >20% were confirmed by Sanger sequencing. SNVs with VAF <20% were validated by NGS/targeted sequencing in an independent second experiment.<sup>36</sup>

### Comparison of CRG/EMG variation frequencies between NB and gnomAD reference cohort

We compared the CRG/EMG variation frequencies from the NB cohort (Supporting Information Table S3) with those reported in a large unrelated population (gnomAD database;  $n = \sim 123,000$ ). The gnomAD VCF file was intersected with the CRG/EMG TSCA<sup>®</sup> panel bed file and the corresponding subset of variants from gnomAD were extracted. Using the same approach as in the NB cohort, SNVs/InDels from gnomAD were annotated/classified into pathogenic/likely pathogenic and their frequency was calculated by dividing the sum of the allele count by the median of total alleles. To compare the frequencies of CRG/EMG variations in the NB vs. the control cohort, for each gene an Odds Ratios was calculated: OR = (NB cases with pathogenic/likely pathogenic variant/cases without pathogenic or likely pathogenic variant)/(gnomAD

control with pathogenic or likely pathogenic variant/gnomAD control without pathogenic or likely pathogenic variant) along with 95% confidence intervals. *p*-Value and OR were obtained using two-sided Fisher's test performed in R.<sup>37</sup>

### Statistical analysis

Correlation analyses using chi-squared test was done with MedCalc (Medical Calculator 13.3.0.0). Progression-free survival (PFS) and overall survival (OS) were estimated using the Kaplan–Meier method, and comparisons were made using log-rank tests.

### Western blot

Immunoblots were done as reported previously using monoclonal rabbit SMARCA4/BRG1 (EPR3912, #GTX62750, Genetex, Irvine, CA) and HRP-conjugated GAPDH (#HRP-60004, ProteinTech, Rosemont, IL) antibodies.

## Results

### Genetic variations of CRGs/EMGs detected in 283 NB patients and in 30 NB cell lines

In a cohort of 283 patients studied by WGS/WES (*n* = 35) or by TSCA<sup>®</sup> (*n* = 228), or both (*n* = 20), pathogenic/likely pathogenic genetic variations (focal CNAs, InDels, SNVs) in at least one CRG/EMG were identified in 56/283 cases (20%). Altogether 60 pathogenic/likely pathogenic genetic variations were identified (Table 1, Fig. 1 and Supporting Information Table S3). The most frequently altered genes were *ATRX*, *SMARCA4*, *MLL3* and *ARID1B*.

*ATRX* was targeted by focal CNAs, small InDels and SNVs in 6, 2 and 6 cases, respectively (Supporting Information Table S3 and Fig. 1). *SMARCA4* was altered by 6 SNVs/InDel, 1 focal CNA and 1 large deletion. *MLL3* was targeted by SNVs in seven cases (Supporting Information Table S3, Fig. 1). *ARID1A* and *ARID1B* genes were altered in altogether 10/283 (3.5%) cases, with two SNVs and one InDel in *ARID1A* and five InDels and two SNVs in *ARID1B*.<sup>9</sup>

Variations in two different CRGs/EMGs were observed in three cases (Supporting Information Table S3, in bold; Fig. 1), targeting *MLL3* and *ATRX*, *MLL3* and *ARID1B*, and *PBRM1* and *ATRX*, respectively.

Double events encompassing both SNV/small InDels, focal CNAs and additional LOH/copy number loss of the same gene were observed in six cases (Fig. 1). The *SMARCA4* gene was targeted by double events in three cases (NBSW80, NBSW251 and NBSW77) with either a SNV and/or InDel (3 bp) in the presence of only 1 copy of the *SMARCA4* gene, corresponding to focal deletions (Figs. 2a and 2b).

Case NBSW165 showed a loss of one copy and a focal CNA covering a region from EX2 to EX9 of the second copy of *ATRX* (Fig. 1; Supporting Information Table S3). A copy number loss involving the *PBRM1* gene is observed in case NBSW23; a somatic SNV is also observed in the same gene (Fig. 1).

In another patient (NBSW26) presenting with speech delay, attention deficit and hyperactivity disorder as well as an adrenal

stage *INSS stage 1*, *INRG L1* NB diagnosed at 4 years of age, a germline deletion of 1.7 Mb at chromosome band 19p13.2 (chr19:10462524–12157782) was detected (Supporting Information Fig. S2). The germline deleted region encompassing among others the *SMARCA4* gene was confirmed by aCGH on peripheral blood lymphocytes (Supporting Information Table S2).

Targeted TSCA<sup>®</sup> sequencing of the 33 CRGs/EMGs was performed in 30 NB cell lines (Supporting Information Table S1). Nine variations were detected in seven NB cell lines (Table 1, Supporting Information Table S3, Fig. 1), most frequently targeting *SMARCA4* (3 cases), *MLL2* (2 cases) and *MLL3* (2 cases) (Fig. 1, Supporting Information Table S3).

Interestingly, a genetic event consisting of partial loss of the *MLL2* gene was detected in the NB cell line CLB-Re (Fig. 3a). SNP analysis in the same NB cell line revealed a large region of copy number loss in chr12 encompassing the *MLL2* gene at the position chr12:49449107–58196639, confirming the observation obtained by TSCA sequencing (Fig. 3b).

In the SKNFI cell line, a double event (SNV and LOH) was observed in *SMARCA4* (Figs. 4a and 4b). The absence of expression of *SMARCA4* was confirmed by expression analysis and western blot. On the other hand, expression of *SMARCA4* was maintained in the SKNSH cell line which harbors a single SNV event in this gene (Figs. 4c and 4d).

### Germline vs. somatic occurrence of variations

Altogether, among the 60 pathogenic/likely pathogenic CRG/EMG variations, the variation occurred at a somatic level in five cases, targeting *ATRX*, *PBRM1* or *MLL3*, whereas a germline variation could be confirmed in 13 other cases, targeting *ARID1B*, *MLL3*, *MLL2*, *ATRX*, *ARID1A* and *SMARCA4* (Supporting Information Table S3; Fig. 1). For the remaining 42 variations detected in tumor samples without paired germline material available, a distinction between germline and somatic variation could not be made.

We then compared the frequencies of pathogenic/likely pathogenic SNVs/inDels of the 33 studied CRGs/EMGs in the NB cohort to those observed in gnomAD as a control cohort. The analysis revealed a significantly higher frequency of variations (*p* < 0.05 after Bonferroni correction) in two CRGs, *SMARCA4* (OR 4.49, 95% CI: 1.63–9.97, *p* = 0.038) and *ATRX* (OR 3.44, 95% CI: 1.46–6.91, *p* = 0.043) in NB cases when compared to gnomAD, independent of whether the observed SNVs were of known germline, somatic or unknown origin, indicating an enrichment for variations in the genes *SMARCA4* and *ATRX* in NB cases. All other CRGs/EMGs included in our study were not altered with a higher frequency than in the general population (Supporting Information Table S4).

### Potential functional impact

The potential functional impact of the observed variations was analyzed *in silico*. For a given variation, within the studied gene, the localization of the amino acid change was studied with regards

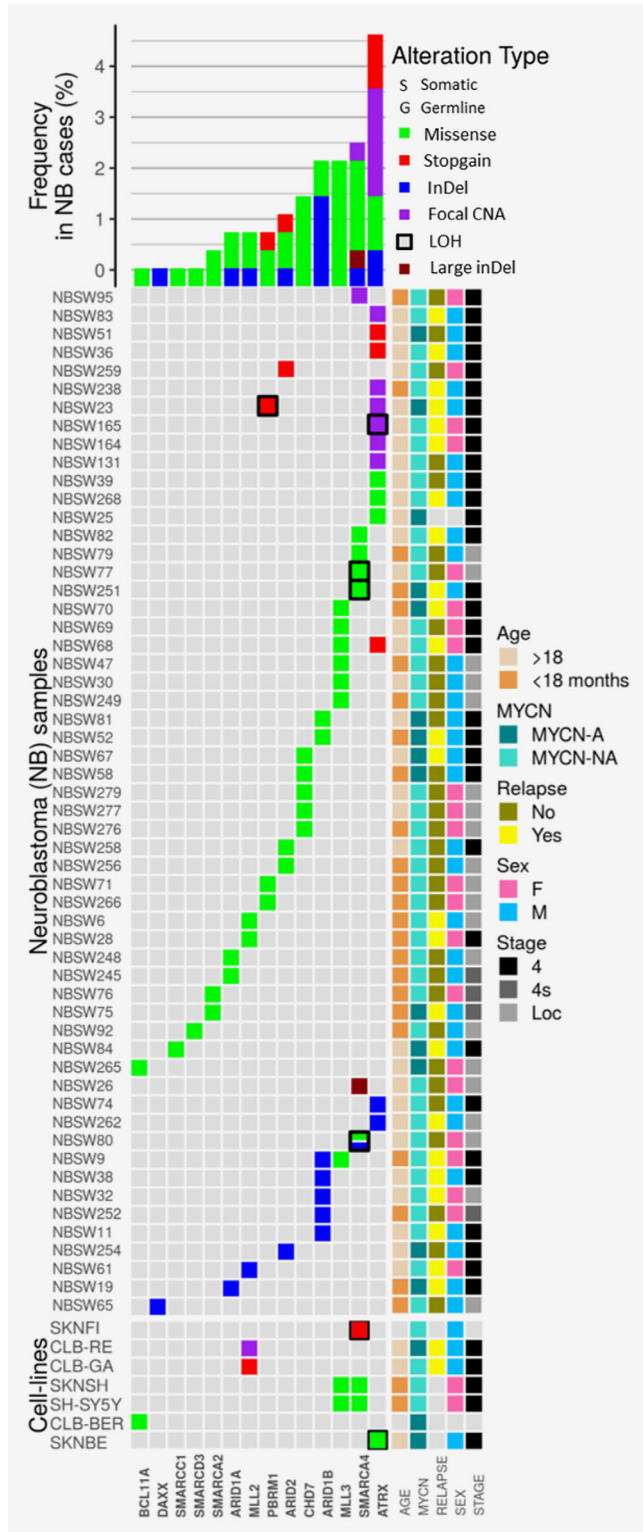
to domains of functional importance for chromatin remodeling or highly conserved domains but with unknown function.

Among the 30 variations detected in CRGs, 12 variations were localized in these domains (9 SNVs, 1 focal CNA, 1 InDel and

1 large inDel; Supporting Information Table S3). Their position was then analyzed with regards to protein-specific domains (DNA binding, bromodomain, histone binding, protein-protein interaction domains). In the *SMARCA4* gene, four identified variations mapped to these functional domains (Fig. 4d), including three missense point mutations in the ATPase domain of *SMARCA4* gene (Fig. 4d).

Among the 30 variations observed in EMGs, 12 variations were localized in domains of functional importance (10 SNVs, 1 focal CNA, 1 InDel; Supporting Information Table S3).

Overall, among all 60 variations observed in CRGs/EMGs, 24 mapped to domains of functional importance at a protein level (Supporting Information Table S3).



### Correlation of CRGs/EMGs variations with clinical parameters

There were no statistically significant correlations between the presence of a CRG/EMG variation and the main clinical prognostic parameters of NB (data not shown). Furthermore, among patients whose tumors were analyzed by WGS ( $n = 8$ ),<sup>10</sup> the presence of CRG/EMG variations, was not associated with an overall increase in the tumor mutational burden (data not shown).

A statistically significant poorer overall survival was observed for patients whose tumors harbored CRG/EMB variations in the overall cohort (Supporting Information Fig. S3A). However, there was no statistically significant difference of overall survival between the two patient groups in patient subgroups defined by age, stage or MYCN status, nor among all high-risk patients only ( $n = 124$ ; Supporting Information Fig. S3C). Furthermore, no statistically significant difference in PFS was observed (Supporting Information Fig. S3B).

### Discussion

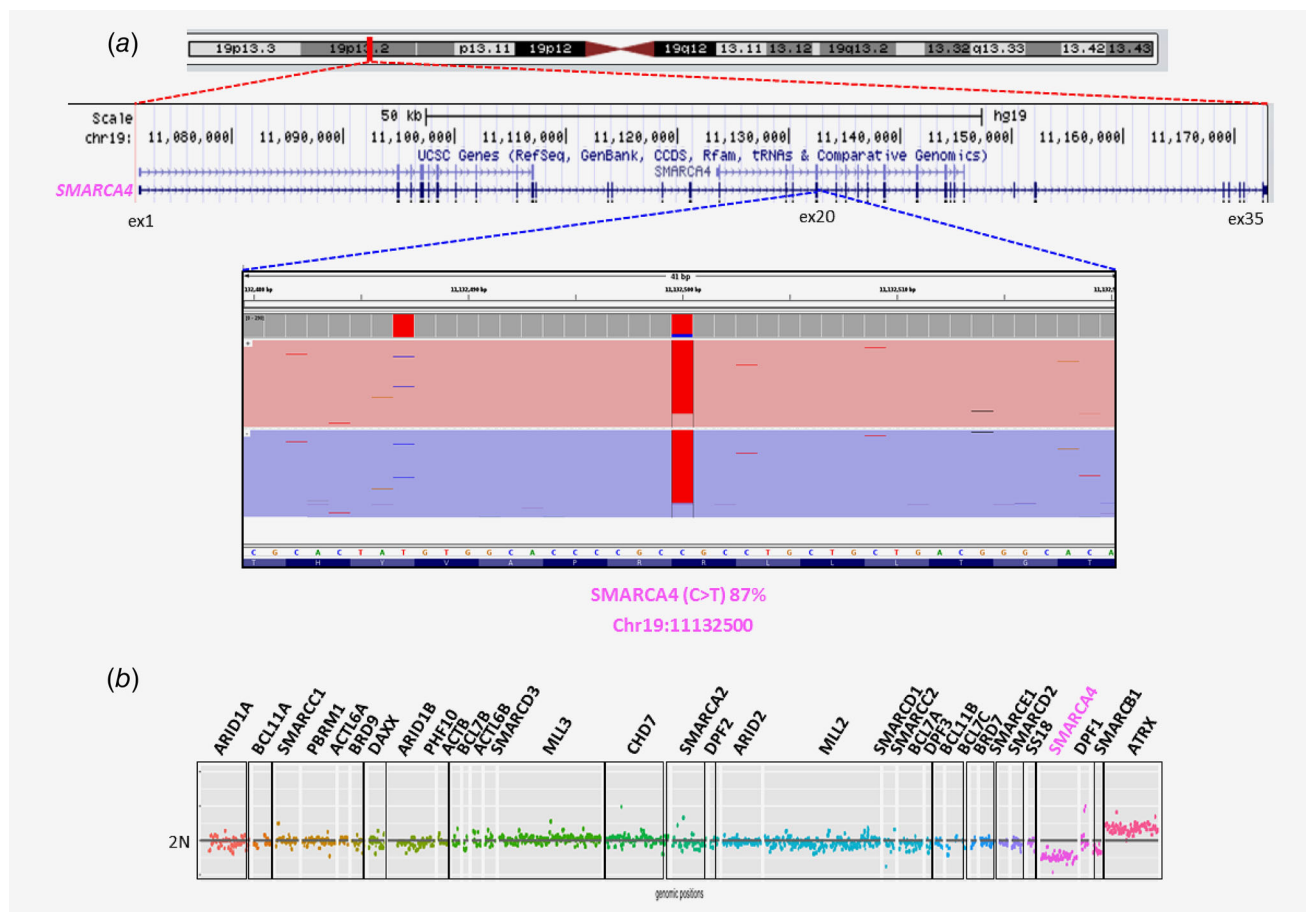
Recent reports have highlighted chromatin remodeling as an important player in oncogenesis, with the main mechanisms of action consisting in tumor suppression, but variations in CRG/EMGs might also play a role in oncogenesis via gain of function.<sup>38,39</sup>

**Figure 1.** Genetic variations in chromatin remodeling and epigenetic modifier genes across a series of 283 NB patients and 30 NB cell lines analyzed by TSCA and/or WGS/WES approaches. In the lower part of the figure, genes found to be altered ( $n = 14$ ) are arranged in rows; cases for whom a genetic variation is detected in the studied genes ( $n = 56$ ) are arranged in columns, respectively. The 60 SNVs, InDels, and focal CNAs detected in the 56 cases are represented by colored cases. Double events (SNV/inDels and LOH) are represented by colored cases surrounded by a black square. In the lower part of the graph, the 9 SNVs, InDels and CNAs detected in 7 NB cell lines are represented. The overall frequencies are indicated in the upper half of the figure. Nonsynonymous SNVs are represented in green, stop-gain SNVs are represented in red, InDels are represented in blue, focal CNA in purple and large deletions in brown. The right data grid summarizes clinical information of each neuroblastoma sample.

Table 1. Cases showing SNVs, InDels and focal CNAs in CRGs/EMGs in 283 neuroblastoma patients and 30 neuroblastoma cell lines

Cohort	MYCN status	Stage	Age at diagnosis	Analytical technique	Cases with pathogenic/likely pathogenic variations in CRGs/EMGs	Number of pathogenic/likely pathogenic variations in CRGs/EMGs	Pathogenic/likely pathogenic variations in CRGs/EMGs	Origin of variations
Patients <i>n</i> = 283	MNA <i>n</i> = 64; MN-NA <i>n</i> = 219	Stage 4 <i>n</i> = 128; Localized <i>n</i> = 130; Stage 4s <i>n</i> = 25	<18 months <i>n</i> = 159; ≥18 months <i>n</i> = 124	WES/WGS <i>n</i> = 55	16	18	ATRX <i>n</i> = 5	2/5 germline; 3/5 Somatic
							ARID1B <i>n</i> = 5	Germline
							MILL3 <i>n</i> = 3	2/3 Germline; 1/3 Somatic
							MILL2 <i>n</i> = 2	Germline
							PBRM1 <i>n</i> = 1	Somatic
							ARID1A <i>n</i> = 1	Germline
							SMARCA4 <i>n</i> = 1	Germline
				TSCA <i>n</i> = 228	40	42	ATRX <i>n</i> = 9	Unknown
							SMARCA4 <i>n</i> = 7	
							MILL2 <i>n</i> = 1	
							MILL3 <i>n</i> = 4	
							ARID1A <i>n</i> = 2	
							ARID2 <i>n</i> = 4	
							CHD7 <i>n</i> = 5	
							SMARCC1 <i>n</i> = 1	
							SMARCA2 <i>n</i> = 2	
							SMARCD3 <i>n</i> = 1	
							PBRM1 <i>n</i> = 2	
							DAXX <i>n</i> = 1	
							ARID1B <i>n</i> = 2	
							BCL11A <i>n</i> = 1	
Cell lines <i>n</i> = 30	MNA <i>n</i> = 22 MN-NA <i>n</i> = 8			TSCA <i>n</i> = 30	7	7	ATRX <i>n</i> = 1	
							MILL2 <i>n</i> = 2	
							MILL3 <i>n</i> = 1	
							SMARCA4 <i>n</i> = 2	
							BCL11A <i>n</i> = 1	

In the 55 samples analyzed by WES/WGS, 7 CRG/EMG genes show at least one variation. In the 248 samples analyzed by the TSCA approach, 14 CRG/EMG genes show a variation. Five genes harbor a variation in NB cell lines.



**Figure 2.** Example of double event detection (SNV and copy number loss in *SMARCA4* gene) by TSCA and aCGH data analysis in NBSW80. (a) The IGV profile from TSCA analysis revealed the presence of one SNV in *SMARCA4* (C>T; chr19: 11132500) gene with variant allele fraction of 87%. (b) The comparison between the copy number profiles obtained by TSCA revealed a second event with copy number loss in the *SMARCA4* gene.

CRG alterations have previously been identified in NB patients, targeting *ARIDs*, *ATRX*, *DAXX* or *SMARCA4*.<sup>7–9,40,41</sup>

Our study now focused on the genetic status of 33 CRGs/EMGs in tumor samples of 283 NB patients. Altogether, CRG/EMG variations (focal CNA, InDels, SNVs) were identified in 56 cases (20%), mapping to CRG/EMG functional domains in 8.4% (24/283) of cases and most frequently targeting *ATRX*, *SMARCA4*, *MLL3* and *ARID1B*. Other CRGs and many more EMGs, not covered by the panel used in this targeted analysis, exist throughout the genome,<sup>42</sup> and thus it is possible that variations targeting other CRG/EMGs not taken into account in our study might also be present in the analyzed samples.

In NB, *ATRX* alterations have been associated with activated alternative lengthening of telomere (ALT).<sup>8</sup> In our study, 5% (14/283) of cases showed an *ATRX* variation (SNV/InDel or focal CNA). Further studies will be required to establish the role of these variations on patient survival.

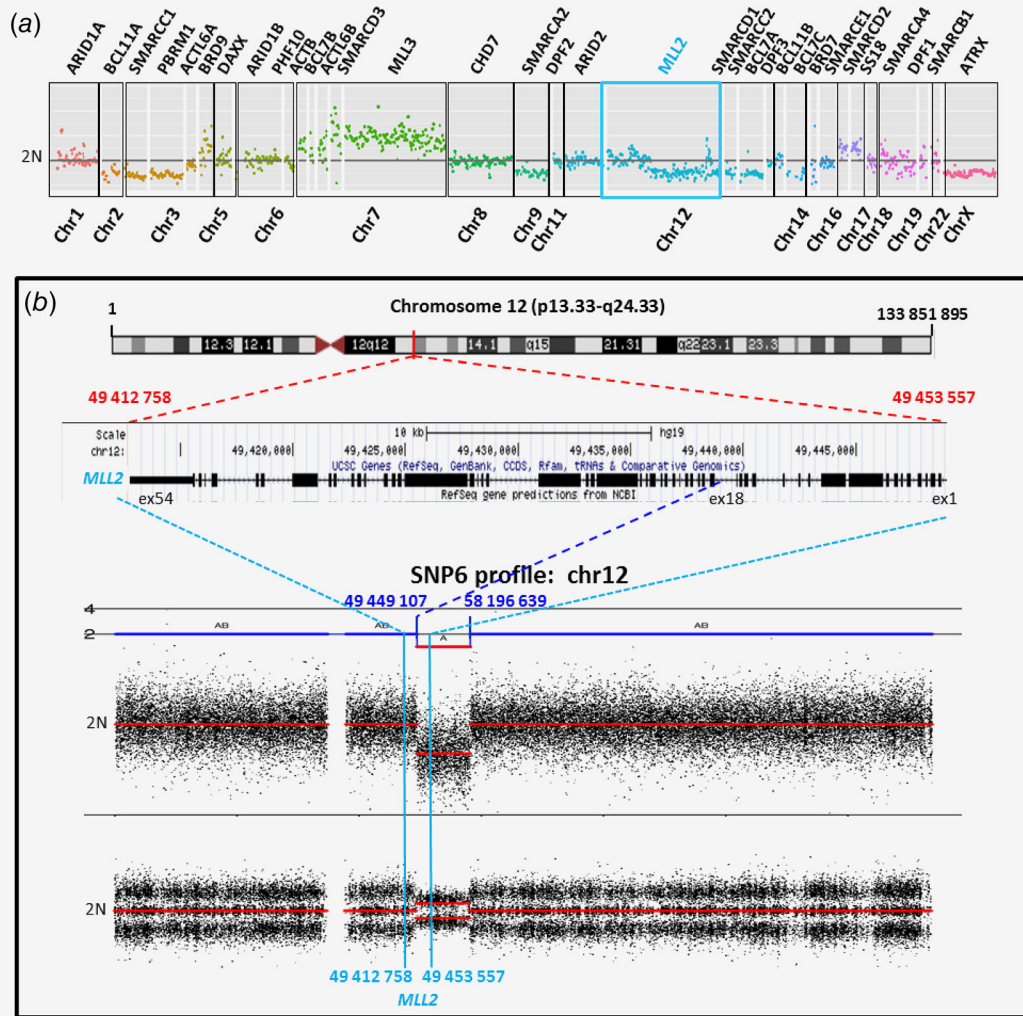
The *MLL3* gene is rarely altered in NB, with two *MLL3* variations (A293V and P309L) reported outside of annotated protein domains.<sup>43</sup> In our cohort, seven missense SNVs were detected, four of them within *MLL3* functional domains.

Alterations in *ARID1A/ARID1B* have been described in 11% of NB patients,<sup>9</sup> with the presence of variations in a single allele of *ARID1A/ARID1B* possibly corresponding to a dominant tumor suppressor.<sup>41</sup> Ten cases showed *ARID1A/ARID1B* variations, with two SNV/inDels occurring in functional domains.

Altogether, among the 56 cases with CRG/EMG variations, four patients harbored somatic and 12 showed germline variations. For the remaining 40 cases, due to the absence of paired germline DNA, it could not be established whether the observed variations were somatic or might concern private constitutional polymorphisms. Yet constitutional variations of potential functional impact might also be of oncogenic importance.

Indeed, germline events in CRGs such as *SMARCB1*, *SMARCA4* or *SMARCE1* have been described as predisposition syndromes for other cancer types, including RT.<sup>1</sup> In our series, germline events in CRGs/EMGs were observed in 23.6% (13/55) of patients with available germline material.

In RT, the recurrent germline loss of one *SMARCB1* allele followed by somatic loss of the second allele indicates a classic tumor suppressor, with loss of critical subunits of SWI/SNF.<sup>2</sup> Given the possible suppressor behavior of CRG/EMGs, we searched for



**Figure 3.** Partial loss event detected in *MLL2* gene in CLB-Re NB cell line. (a) The copy number profile obtained by TSCA shows the presence of a partial loss starting in *MLL2* gene. (b) SNP6.0 array-based copy number encompassing the Chromosome 12 confirms the presence of a large loss with a breakpoint within the *MLL2* gene.

double events possibly causing loss of function, with double events identified in 2.1% (6/283) of patients, most frequently in *SMARCA4* (3/283).

*SMARCA4* has been shown to play a role in the oncogenesis of different tumors and thus are not specific to NB. Indeed, germline alterations in *SMARCA4* conferring predisposition to SCOHT and RT.<sup>39,44,45</sup> At a somatic level, missense point mutations in *SMARCA4* mapping to the ATPase domain might cause loss of direct binding between BAF and PRC1 which could contribute to oncogenesis or epigenetic plasticity during tumor development.<sup>46</sup> Such *SMARCA4* missense point mutations have been described in medulloblastoma.<sup>47,48</sup>

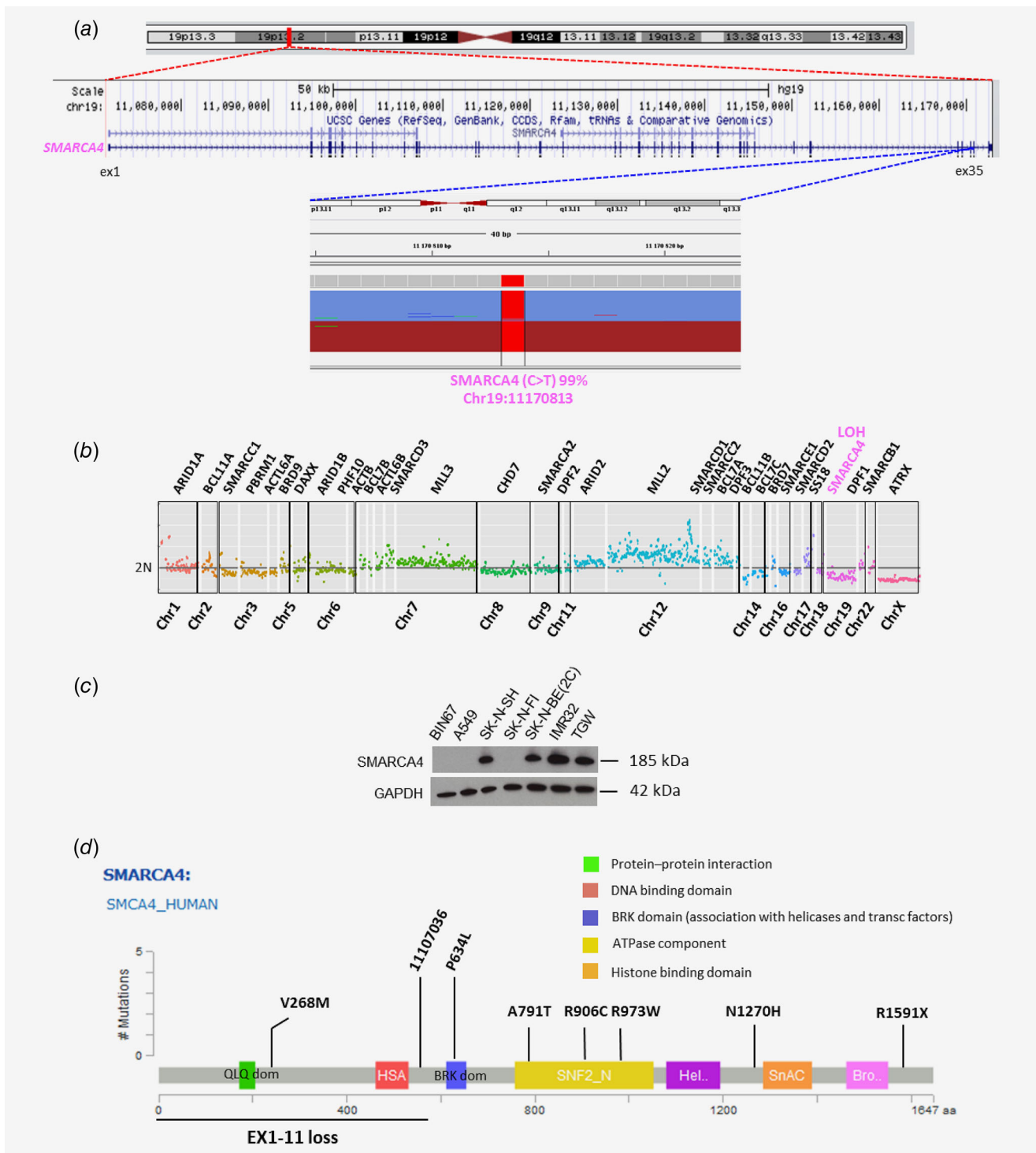
In our study, 5/8 *SMARCA4* variations concerned missense point mutations. Although it has been suggested that *SMARCA4* might function as an oncogene in NB, and the overexpression of *SMARCA4* in Stage 4 NB patients is associated with poorer outcomes,<sup>49</sup> the functional impact of genetic variations can be

difficult to determine without in-depth functional studies which are beyond the scope of this article.

Even though the somatic or germline origin of *SMARCA4* variations in our study remained undetermined, our data suggest that variations in *SMARCA4* might play a role in the oncogenesis of NB, and several arguments now underline its role. First, we describe a case of NB within the context of a germline deletion encompassing *SMARCA4*. Second, our findings highlight a significantly higher frequency of variations in *SMARCA4* in NB cases when compared to the gnomAD reference population. Third, we identified four NB cases showing double events targeting *SMARCA4*. In addition, a double event (stop-gain variation and LOH) of *SMARCA4* was observed in the SKNFI NB cell line causing the absence of *SMARCA4* expression.

Further functional studies are needed for a deeper understanding of *SMARCA4* variations in NB in order to guide the development of more effective therapies.





**Figure 4.** Example of double event (LOH and SNV in *SMARCA4* gene) in SKNFI NB cell line. (a) The IGV profile from TSCA amplicon-seq analysis revealed the presence of one SNV in *SMARCA4* gene (C>T; Chr19: 11170813) with variant allele fraction of 99%. (b) The copy number profile obtained by TSCA shows the presence of a second event with copy number loss in the same gene. (c) Immunoblot of *SMARCA4* expression in different neuroblastoma and nonneuroblastoma cell lines. BIN67: Small cell carcinoma of the ovary of hypercalcemic type; A549: lung carcinoma; SKNSH, SKNFI, SKNBE(2C), IMR32 and TGW: neuroblastoma. (d) SNVs, inDels and focal CNAs identified in *SMARCA4* gene mapping functional domains in patients enrolled in the study.

Altogether, our data demonstrate CRG/EMG variations with likely functional impact in 8.4% (24/283) of all NB patients.

Although no difference in PFS between cases with or without CRG/EMG variations was observed, a poorer overall survival in cases harboring CRG/EMG variations was observed. This suggests

that cases harboring CRG/EMG variations might present more aggressive disease or that salvage treatment used at relapse might potentially be more efficient in cases showing no CRG/EMG variation. Correlations between the presence of variations in CRG/EMG and other genetic factors associated with poor outcome in NB should be determined in future studies and larger series, such as the overall mutational burden<sup>50</sup> or factors contributing to a mechanistical classification of NB including mutations of the RAS: MAPK pathway and telomere maintenance mechanisms.<sup>51</sup>

New treatment strategies are necessary for patients with high-risk NB, and our data further highlight the potential interest of drugs modulating chromatin remodeling processes. Histone deacetylase inhibitors (HDACi) or DNA methyltransferase inhibitors were the first epigenetic compounds to reach clinical trials with potential benefits in some patients with CRG/EMG variations.<sup>52</sup> Further preclinical development of therapeutic

approaches involving CRGs/EMGs and in particular *SMARCA4* will be warranted.

### Acknowledgements

Our study was supported by the Annenberg Foundation, the Nelia and Amadeo Barletta Foundation (FNAB) and the Association Hubert Gouin Enfance et Cancer. Our study was also funded by the Associations Enfants et Santé, Les Bagouz à Manon, Les amis de Claire. Funding was obtained from SIRIC/INCa (Grant INCa-DGOS-4654), as well as from the CEST of Institute Curie, and Programme Hospitalier de recherche en Cancérologie (PHRC) IC2007-09 grant. High-throughput sequencing was performed by the ICGex NGS platform of the Institut Curie supported by the grants ANR-10-EQPX-03 (Equipex) and ANR-10-INBS-09-08 (France Génomique Consortium) from the Agence Nationale de la Recherche ("Investissements d'Avenir" program), by the Cancerpole Ile-de-France and by the SIRIC-Curie program-SIRIC Grant "INCa-DGOS-4654", as well as the grant of collaboration with CEA/IG/CNG financed by France Génomique infrastructure, as part of the 517 program "Investissements d'Avenir" from the Agence Nationale pour la Recherche (contract ANR-10-INBS-518 09).

### References

- Masliyah-Planchon J, Bièche I, Guinebretière J-M, et al. SWI/SNF chromatin remodeling and human malignancies. *Annu Rev Pathol Mech Dis* 2015;10:145–71.
- Versteeg I, Sévenet N, Lange J, et al. Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. *Nature* 1998;394:203–6.
- Kadoch C, Hargreaves DC, Hodges C, et al. Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat Genet* 2013;45:592–601.
- Matthay KK, Maris JM, Schleiermacher G, et al. Neuroblastoma. *Nat Rev Dis Primer* 2016;2:16078.
- Molenaar JJ, Koster J, Zwijnenburg DA, et al. Sequencing of neuroblastoma identifies chromothripsis and defects in neurogenesis genes. *Nature* 2012;483:589–93.
- Cheung N-KV, Dyer MA. Neuroblastoma: developmental biology, cancer genomics and immunotherapy. *Nat Rev Cancer* 2013;13:397–411.
- Pugh TJ, Morozova O, Attiyeh EF, et al. The genetic landscape of high-risk neuroblastoma. *Nat Genet* 2013;45:279–84.
- Kurihara S, Hiyama E, Onitake Y, et al. Clinical features of ATRX or DAXX mutated neuroblastoma. *J Pediatr Surg* 2014;49:1835–8.
- Sausen M, Leary RJ, Jones S, et al. Integrated genomic analyses identify ARID1A and ARID1B alterations in the childhood cancer neuroblastoma. *Nat Genet* 2013;45:12–7.
- Elefeld TF, Oldridge DA, Bernard V, et al. Relapsed neuroblastomas show frequent RAS-MAPK pathway mutations. *Nat Genet* 2015;47:864–71.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 2009;25:1754–60.
- McKenna A, Hanna M, Banks E, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20:1297–303.
- DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 2011;43:491–8.
- Chen X, Schulz-Trieglaff O, Shaw R, et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* 2016;32:1220–2.
- Boeva V, Popova T, Bleakley K, et al. Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics* 2012;28:423–5.
- Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012;22:568–76.
- Seshan V, Olshen A. *DNACopy: DNA copy number data analysis. R package version 1.46.0*. Vienna, Austria: R Foundation for Statistical Computing, 2016.
- Liu X, Wu C, Li C, et al. dbNSFP v3.0: a one-stop database of functional predictions and annotations for human nonsynonymous and splice-site SNVs. *Hum Mutat* 2016;37:235–41.
- Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. *Nat Methods* 2012;9:357–9.
- Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetic parameter estimation from sequencing data. *Bioinformatics* 2011;27:2987–93.
- Kadri S, Zhen CJ, Wurst MN, et al. Amplicon Indel hunter is a novel bioinformatics tool to detect large somatic insertion/deletion mutations in amplicon-based next-generation sequencing data. *J Mol Diagn* 2015;17:635–43.
- Boeva V, Popova T, Lienard M, et al. Multi-factor data normalization enables the detection of copy number aberrations in amplicon sequencing data. *Bioinformatics* 2014;30:3443–50.
- Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol* 2011;29:24–6.
- Cingolani P, Platts A, Wang LL, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* 2012;6:80–92.
- Sherry ST, Ward MH, Kholodov M, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001;29:308–11.
- Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016;536:285–91.
- Forbes SA, Beare D, Boutselakis H, et al. COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Res* 2017;45:D777–83.
- Landrum MJ, Lee JM, Benson M, et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res* 2016;44:D862–8.
- Chang MT, Bhattarai TS, Schram AM, et al. Accelerating discovery of functional mutant alleles in cancer. *Cancer Discov* 2018;8:174–83.
- Carter H, Douville C, Stenson PD, et al. Identifying Mendelian disease genes with the variant effect scoring tool. *BMC Genomics* 2013;14(Suppl 3):S3.
- Dong C, Wei P, Jian X, et al. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Hum Mol Genet* 2015;24:2125–37.
- Ioannidis NM, Rothstein JH, Pejaver V, et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. *Am J Hum Genet* 2016;99:877–85.
- Schwarz JM, Cooper DN, Schuelke M, et al. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods* 2014;11:361–2.
- Jagadeesh KA, Wenger AM, Berger MJ, et al. M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. *Nat Genet* 2016;48:1581–6.
- Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014;46:310–5.
- Bellini A, Bernard V, Leroy Q, et al. Deep sequencing reveals occurrence of subclonal ALK mutations in neuroblastoma at diagnosis. *Clin Cancer Res* 2015;21:4913–21.
- R Core Team (2018). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing Available from: <https://www.gbif.org/tool/81287/r-a-language-and-environment-for-statistical-computing> [date last accessed June 18, 2018].
- Biegel JA, Busse TM, Weissman BE. SWI/SNF chromatin remodeling complexes and cancer. *Am J Med Genet C Semin Med Genet* 2014;166C:350–66.

39. Holsten T, Bens S, Oyen F, et al. Germline variants in SMARCB1 and other members of the BAF chromatin-remodeling complex across human disease entities: a meta-analysis. *Eur J Hum Genet* 2018;26:1083–93.
40. Peifer M, Hertwig F, Roels F, et al. Telomerase activation by genomic rearrangements in high-risk neuroblastoma. *Nature* 2015;526:700–4.
41. Cancer Genome Atlas Research Network, Kandoth C, Schultz N, et al. Integrated genomic characterization of endometrial carcinoma. *Nature* 2013;497:67–73.
42. Kadoch C, Crabtree GR. Mammalian SWI/SNF chromatin remodeling complexes and cancer: mechanistic insights gained from human genomics. *Sci Adv* 2015;1:e1500447.
43. Chmielecki J, Bailey M, He J, et al. Genomic profiling of a large set of diverse pediatric cancers identifies known and novel mutations across tumor spectra. *Cancer Res* 2017;77:509–19.
44. Witkowski L, Goudie C, Foulkes WD, et al. Small-cell carcinoma of the ovary of hypercalcemic type (malignant Rhabdoid tumor of the ovary): a review with recent developments on pathogenesis. *Surg Pathol Clin* 2016;9:215–26.
45. Witkowski L, Carrot-Zhang J, Albrecht S, et al. Germline and somatic SMARCA4 mutations characterize small cell carcinoma of the ovary, hypercalcemic type. *Nat Genet* 2014;46:438–43.
46. Stanton BZ, Hodges C, Calarco JP, et al. Smarca4 ATPase mutations disrupt direct eviction of PRC1 from chromatin. *Nat Genet* 2017;49:282–8.
47. Robinson G, Parker M, Kranenburg TA, et al. Novel mutations target distinct subgroups of medulloblastoma. *Nature* 2012;488:43–8.
48. Pugh TJ, Weeraratne SD, Archer TC, et al. Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. *Nature* 2012;488:106–10.
49. Jubierre L, Soriano A, Planells-Ferrer L, et al. BRG1/SMARCA4 is essential for neuroblastoma cell viability through modulation of cell death and survival pathways. *Oncogene* 2016;35:5179–90.
50. Gröbner SN, Worst BC, Weischenfeldt J, et al. The landscape of genomic alterations across childhood cancers. *Nature* 2018;555:321–7.
51. Ackermann S, Cartolano M, Hero B, et al. A mechanistic classification of clinical phenotypes in neuroblastoma. *Science* 2018;362:1165–70.
52. Gryder BE, Sodji QH, Oyelere AK. Targeted cancer therapy: giving histone deacetylase inhibitors all they need to succeed. *Future Med Chem* 2012;4:505–24.