

# DGCR8 microprocessor defect characterizes familial multinodular goiter with schwannomatosis

Barbara Rivera,<sup>1,2,3</sup> Javad Nadaf,<sup>2,3</sup> Somayyeh Fahiminiya,<sup>4</sup> Maria Apellaniz-Ruiz,<sup>2,3,4,5</sup> Avi Saskin,<sup>5,6</sup> Anne-Sophie Chong,<sup>2,3</sup> Sahil Sharma,<sup>7</sup> Rabea Wagener,<sup>8</sup> Timothée Revil,<sup>5,9</sup> Vincenzo Condello,<sup>10</sup> Zineb Harra,<sup>2,3</sup> Nancy Hamel,<sup>4</sup> Nelly Sabbaghian,<sup>2,3</sup> Karl Muchantef,<sup>11,12</sup> Christian Thomas,<sup>13</sup> Leanne de Kock,<sup>2,3,5</sup> Marie-Noëlle Hébert-Blouin,<sup>14</sup> Angelia V. Bassenden,<sup>15</sup> Hannah Rabenstein,<sup>8</sup> Ozgur Mete,<sup>16,17</sup> Ralf Paschke,<sup>18,19,20,21,22</sup> Marc P. Puztaszeri,<sup>23</sup> Werner Paulus,<sup>13</sup> Albert Berghuis,<sup>15</sup> Jiannis Ragoussis,<sup>4,9</sup> Yuri E. Nikiforov,<sup>10</sup> Reiner Siebert,<sup>8</sup> Steffen Albrecht,<sup>24</sup> Robert Turcotte,<sup>25,26</sup> Martin Hasselblatt,<sup>13</sup> Marc R. Fabian,<sup>1,2,3,7,15</sup> and William D. Foulkes<sup>1,2,3,4,5,6</sup>

<sup>1</sup>Gerald Bronfman Department of Oncology, McGill University, Montreal, Quebec, Canada. <sup>2</sup>Lady Davis Institute for Medical Research and <sup>3</sup>Segal Cancer Centre, Jewish General Hospital, Montreal, Quebec, Canada. <sup>4</sup>Cancer Research Program, McGill University Health Centre, Montreal, Quebec, Canada. <sup>5</sup>Department of Human Genetics, McGill University, Montreal, Quebec, Canada. <sup>6</sup>Division of Medical Genetics, Department of Medicine, McGill University Health Centre and Jewish General Hospital, Montreal, Quebec, Canada. <sup>7</sup>Department of Experimental Medicine, McGill University, Montreal, Quebec, Canada. <sup>8</sup>Institute of Human Genetics, University of Ulm and University of Ulm Medical Center, Ulm, Germany. <sup>9</sup>Génome Québec Innovation Centre, McGill University, Montreal, Quebec, Canada. <sup>10</sup>Department of Pathology, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA. <sup>11</sup>Department of Diagnostic Radiology, McGill University, Montreal, Quebec, Canada. <sup>12</sup>Pediatric Radiology, Montreal Children's Hospital, Montreal, Quebec, Canada. <sup>13</sup>Institute of Neuropathology, University Hospital Muenster, Muenster, Germany. <sup>14</sup>Department of Neurosurgery, McGill University Health Centre, Montreal, Quebec, Canada. <sup>15</sup>Department of Biochemistry, McGill University, Montreal, Quebec, Canada. <sup>16</sup>Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada. <sup>17</sup>Department of Pathology, University Health Network, Toronto, Ontario, Canada. <sup>18</sup>Department of Medicine, <sup>19</sup>Department of Oncology, <sup>20</sup>Department of Pathology, <sup>21</sup>Biochemistry and Molecular Biology Institute, and <sup>22</sup>Arnie Charbonneau Cancer Institute, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada. <sup>23</sup>Department of Pathology, Jewish General Hospital, Montreal, Quebec, Canada. <sup>24</sup>Department of Pathology, Montreal Children's Hospital, McGill University Health Centre, Montreal, Quebec, Canada. <sup>25</sup>Division of Orthopedic Surgery (Experimental Surgery), McGill University, Montreal, Quebec, Canada. <sup>26</sup>Department of Surgical Oncology, McGill University Health Centre, Montreal, Quebec, Canada.

**BACKGROUND.** *DICER1* is the only miRNA biogenesis component associated with an inherited tumor syndrome, featuring multinodular goiter (MNG) and rare pediatric-onset lesions. Other susceptibility genes for familial forms of MNG likely exist.

**METHODS.** Whole-exome sequencing of a kindred with early-onset MNG and schwannomatosis was followed by investigation of germline pathogenic variants that fully segregated with the disease. Genome-wide analyses were performed on 13 tissue samples from familial and nonfamilial *DGCR8*-E518K-positive tumors, including MNG, schwannomas, papillary thyroid cancers (PTCs), and Wilms tumors. miRNA profiles of 4 tissue types were compared, and sequencing of miRNA, pre-miRNA, and mRNA was performed in a subset of 9 schwannomas, 4 of which harbor *DGCR8*-E518K.

**RESULTS.** We identified c.1552G>A;p.E518K in *DGCR8*, a microprocessor component located in 22q, in the kindred. The variant identified is a somatic hotspot in Wilms tumors and has been identified in 2 PTCs. Copy number loss of chromosome 22q, leading to loss of heterozygosity at the *DGCR8* locus, was found in all 13 samples harboring c.1552G>A;p.E518K. miRNA profiling of PTCs, MNG, schwannomas, and Wilms tumors revealed a common profile among E518K hemizygous tumors. In vitro cleavage demonstrated improper processing of pre-miRNA by *DGCR8*-E518K. MicroRNA and RNA profiling show that this variant disrupts precursor microRNA production, impacting populations of canonical microRNAs and mirtrons.

**CONCLUSION.** We identified *DGCR8* as the cause of an unreported autosomal dominant mendelian tumor susceptibility syndrome: familial multinodular goiter with schwannomatosis.

**FUNDING.** Canadian Institutes of Health Research, Compute Canada, Alex's Lemonade Stand Foundation, the Mia Neri Foundation for Childhood Cancer, Cassa di Sovvenzioni e Risparmio fra il Personale della Banca d'Italia, and the KinderKrebsInitiative Buchholz/Holm-Seppensen.

**Authorship note:** BR and JN contributed equally. SF and MAR contributed equally.

**Conflict of interest:** The authors have declared that no conflict of interest exists.

**Copyright:** © 2020, American Society for Clinical Investigation.

**Submitted:** May 10, 2019; **Accepted:** November 26, 2019; **Published:** February 10, 2020.

**Reference information:** *J Clin Invest.* 2020;130(3):1479–1490.

<https://doi.org/10.1172/JCI130206>.

## Introduction

Familial multinodular goiter can occur alone or in combination with other disorders (1). An exemplar of the latter is *DICER1* syndrome, attributable to germline pathogenic variants in *DICER1*, encoding a key protein in the microRNA processing machinery (2). Euthyroid multinodular goiter (MNG) in *DICER1* syndrome

is usually diagnosed in childhood or adolescence (3) and can be accompanied by other typical syndromic features such as pleuropulmonary blastoma (2), ovarian Sertoli-Leydig cell tumor (2, 3), and cystic nephroma (2). *DICER1*-associated lesions in the thyroid appear to be multiclonal in origin, in that different nodules possess 1 of several distinct somatic missense hotspot mutations in *DICER1* (4, 5). These mutations result in aberrant cleavage of precursor microRNAs (2, 6). *DICER1* is the only miRNA biogenesis gene in which germline mutations have been identified to cause a syndrome; however, somatic mutations in other genes encoding miRNA biogenesis proteins (*DROSHA*, *TARBP2*, *XPO5*, and *DGCR8*) have been found in Wilms tumors, and *DROSHA* somatic homozygous deletions are reported in pineoblastomas (7–9).

Schwannomatosis is an inherited disease of myelin-producing Schwann cells of the peripheral nervous system occurring in the absence of bilateral vestibular schwannomas. It has been described as a third form of neurofibromatosis (10). Although somatic *NF2* mutations are frequently identified in schwannomas (11), germline variants in *NF2* do not cause familial schwannomatosis (10, 11); the 2 genes associated with this disorder are *SMARCB1* and *LZTR1* (11). Notably, the 3 genes lie adjacent to each other, covering 8.72 megabases of chromosome 22q. Different sequences of events that affect all 3 loci can occur, but the most frequent combination of genetic events that lead to schwannomatosis occur in 3 steps: (a) a hypomorphic germline pathogenic variant in *SMARCB1*; (b) complete loss of heterozygosity (LOH) of the alternate allele of chromosome 22q, leading to a tumor that is hemizygous for *NF2*, *SMARCB1*, and *LZTR1*; and (c) an inactivating somatic mutation on the remaining *NF2* allele harboring the germline *SMARCB1* variant (10–13). Schwannomatosis usually arises sporadically, and although familial instances attributable to germline variants in *SMARCB1* or *LZTR1* are described (11), most cases remain unexplained (14).

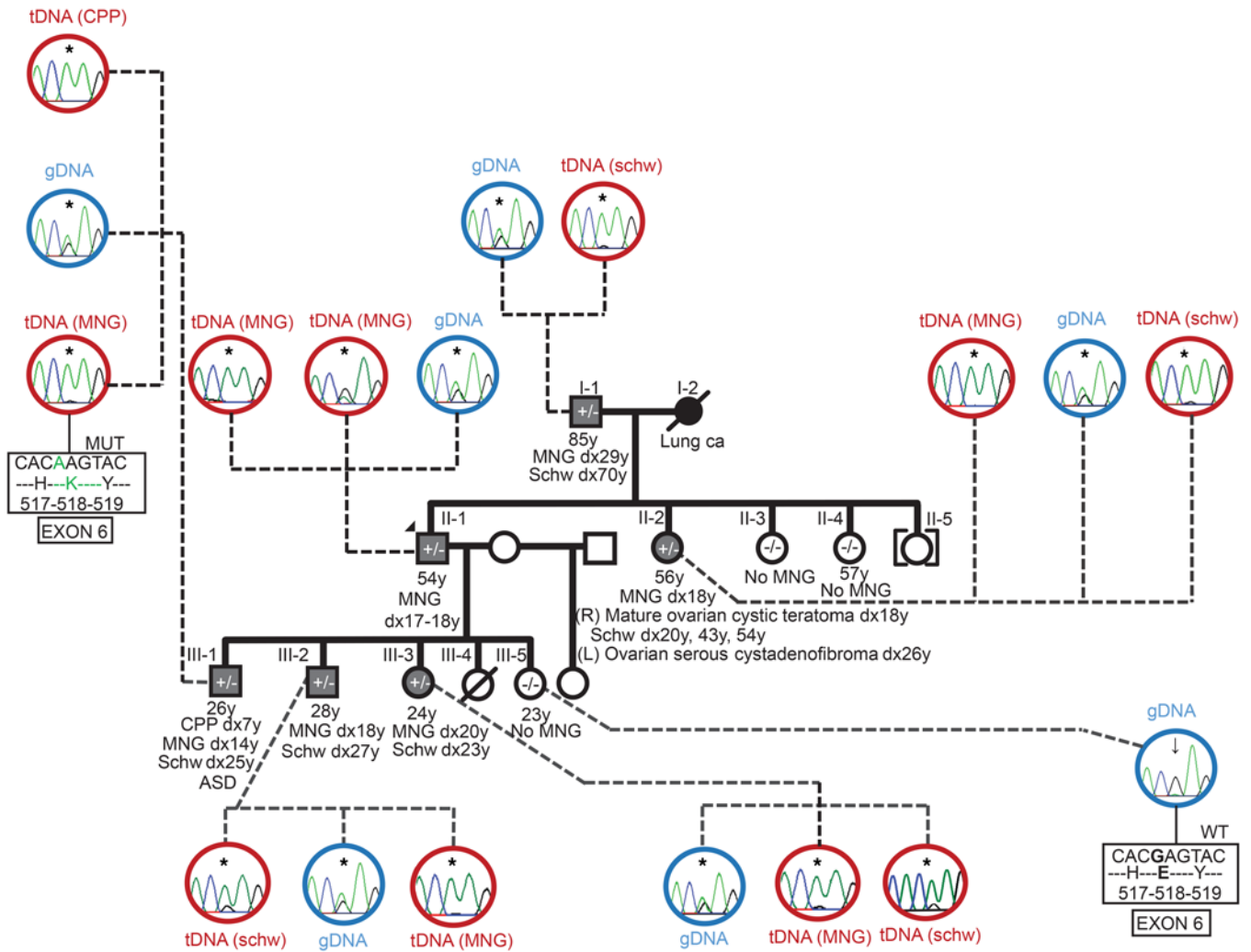
To further understand the genetic contribution to both MNG and schwannomatosis, we extensively characterized 9 members of a family with MNG, schwannomatosis, and a choroid plexus tumor (CPT), and extended the work to include detailed analysis of sporadic cases of Wilms tumors, schwannoma, and PTC, focusing on miRNA- and RNA-Seq.

## Results

**Index family.** The proband (II-1) was referred to the medical genetics service because of a personal and family history of MNG. Subsequent investigation revealed a total of 6 persons over 3 generations who had developed MNG, all resulting in total thyroidectomy. Five of these persons (I-1, II-2, III-1, III-2, and III-3) were found to have 1 or more peripheral nerve schwannomas (Figure 1 and Supplemental Methods) and in 3 of these 5, the absence of intracranial schwannomas on magnetic resonance imaging confirmed schwannomatosis. A choroid plexus papilloma was diagnosed in III-1 at age 7 years. Further clinical, imaging, and pathological details are provided in the Supplemental Text and Supplemental Figures 1 and 2; supplemental material available online with this article; <https://doi.org/10.1172/JCI130206DS1>. Individual III-1 was diagnosed with autism spectrum disorder; however, no other features consistent with a DiGeorge syndrome diagnosis were identified in the patient.

**Molecular genetics studies.** As the major gene for familial euthyroid MNG is *DICER1*, we performed segregation analysis for 4 affected members of this family using markers within and flanking *DICER1*. The results were not consistent with linkage of MNG to 14q23 to 14q32, where *DICER1* is situated (data not shown). Therefore, we performed whole-exome sequencing (WES) using blood or saliva DNA from 9 family members (6 affected, 3 unaffected). Only 2 variants passed our filters and fully segregated with MNG and schwannomas: *c.988G>A* in exon 10 of the collagen-encoding gene NM\_001849 *COL6A2*, leading to p.D330N; and *c.1552G>A* in exon 6 of NM\_022720 *Di George Critical Region 8 (DGCR8)*, resulting in p.E518K (Figure 1 and Supplemental Tables 1 and 2). Given the role of *DGCR8* in miRNA processing (Figure 2) and the importance of this pathway in the proper development of the thyroid gland, we focused on *DGCR8*. The variant *c.1552G>A*;p.E518K, located in a highly conserved amino acid (Supplemental Figure 3), is predicted to be pathogenic by multiple algorithms, is expressed at the RNA level, does not affect splicing, and is not subject to nonsense-mediated decay (Supplemental Figure 4). This variant has never been described in the germline according to public databases (15–17) and (<http://evs.gs.washington.edu/EVS/>; all accessed April 1, 2019) and was absent in 1433 (896 females and 537 males) cancer-free controls ascertained at a single Montreal hospital (mean age 44.98 years, SD 18.07). However, this variant is a recurrent somatic mutation in Wilms tumors (7). To interrogate for the presence of other intronic variants in *DGCR8* in the family, we sequenced the entire locus of *DGCR8* in chromosome 22 (nc\_000022.10:g.20067757-20099510) in individual II-2 germline using a Haloplex High Sensitivity (Haloplex<sup>HS</sup>) capture. After filtering all the promoter, intronic, and 3'UTR variants with a frequency over 1/1000 in the 1000 Genomes database, no variants remained. Finally, we ruled out the presence of large deletions in the 3 known schwannomatosis genes (*LZTR1*, *SMARCB1*, and *NF2*) by a multiplex ligation probe assay (MRC Holland) (data not shown).

All schwannomas and MNGs as well as the CPT examined in the family had biallelic alterations of *DGCR8*: *c.1552G>A*;p.E518K in the germline plus somatic loss of the whole of chromosome 22 as measured by Haloplex<sup>HS</sup> or WES. We studied 3 MNG nodules each from II-1, III-1, III-2, and III-3. All 3 nodules showed LOH in all patients except for II-1 in whom 2 MNG nodules showed LOH and 1 nodule showed absence of LOH. Individual II-1 (MNG1), and MNG samples from II-2 III-1, III-2, and III-3, as well as the CPT, showed loss of the entire chromosome 22, explaining the presence of LOH. Schwannomas from I-1 and II-2 and the CPT from III-1 showed allelic imbalance in the WES data, suggesting a loss of chromosome 22 (Figure 1, Supplemental Tables 3–6, and Supplemental Figures 5 and 6). We sequenced the *DGCR8* coding region in 181 schwannomas and 74 CPTs, as well as in the germline of 18 persons affected with MNG (history of familial MNG [*n* = 13]; personal history of *DICER1*-related lesions [*n* = 5]; Supplemental Text and Supplemental Tables 7 and 8). No cases with *DGCR8*-*c.1552G>A*;p.E518K or predicted truncating mutations were identified, although we identified a variant of uncertain significance (VUS) in 1 sporadic spinal schwannoma (*c.1147A>G*;p.S383G). Moreover, the variant *c.1763A>G*;p.K588R (rs35569747; gnomAD frequency = 0.005920), previously reported as a germline variant in children



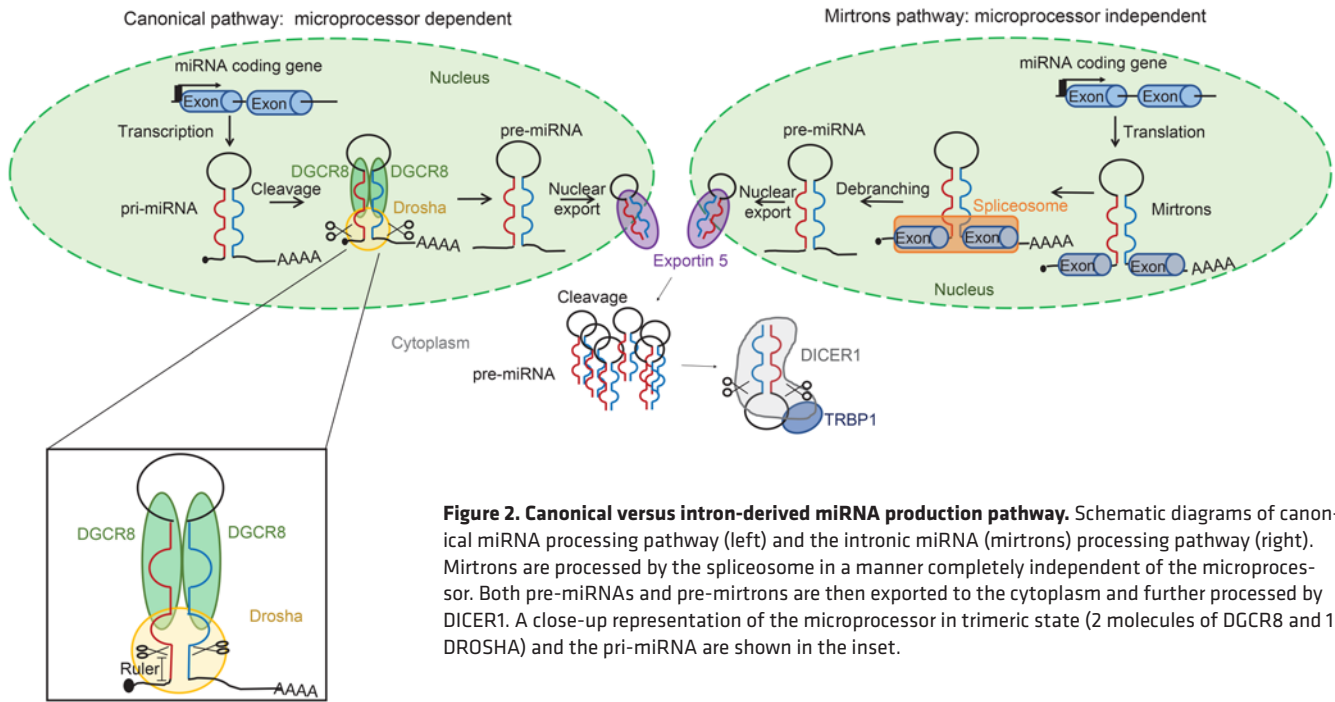
**Figure 1. Pedigree of the family: clinical data and genotypes of a family kindred with germline *DGCR8* variant c.1552G>A, p.E518K.** dx, diagnosis. Person I-1 was diagnosed with a multinodular goiter (MNG) and with a schwannoma (schw). Person II-1 was diagnosed with MNG. Person II-2 had MNG, a mature cystic teratoma, 9 schwannomas, and an ovarian serous cystadenofibroma. Individual III-1 was diagnosed with autism spectrum disorder (ASD) and MNG, and had a choroid plexus papilloma (CPP) WHO stage I and multiple schwannomas. III-3 was diagnosed with MNG and multiple schwannomas. Person III-2 was diagnosed with MNG and a single schwannoma located in the right knee. Chromatograms display the c.1552G>A;p.E518K locus in germline DNA (gDNA, blue circle) and tumor/MNG DNA (tissue DNA [tDNA], red circle) for each affected individual. Representative chromatograms show LOH in individuals I-1, II-1, II-2, III-1, III-2, and III-3. Schwannoma samples from III-2 and III-3 had a remnant of the WT allele, likely due to normal tissue contamination. The chromatograms for the MNG tDNA in III-1, III-2, and III-3 are representative of the Sanger sequencing results for all 3 nodules in each patient. All results are summarized in Supplemental Tables 3–6 and Supplemental Figures 5 and 6. Arrow indicates a WT genotype. Asterisk indicates presence of the mutant base. Three nodules of the MNG were sequenced for II-1, III-1, III-2, and III-3. Germline DNA with WT sequence at the c.1552G;p.E518K locus is shown for unaffected individual III-5. No tumor DNA from the schwannomas of III-1 and from the MNG of I-1 was available.

with Wilms tumors (18, 19), was identified in 5 schwannomas and in 1 CPT, for a combined frequency of 6/237 cases studied (Supplemental Table 8). The c.1552G>A;p.E518K variant was also present in 2 cases from the Tumor Cancer Genome Atlas (20) PTC data set. We collected tumor and germline samples from these 2 cases and confirmed somatic status of the mutation and LOH on chromosome 22q (Supplemental Figures 5 and 6). We genotyped the c.1552G>A;p.E518K variant in 315 PTCs and 106 hyperplastic thyroid nodules, but no other positive cases were found.

**Impact of *DGCR8*-c.1552G>A;p.E518K on miRNA biogenesis.** The variant is located in helix 1 of the first of 2 double-strand RNA-binding domains within *DGCR8*. The corresponding mRNA is expressed and detected both by reverse transcription PCR

(RT-PCR) followed by Sanger sequencing and by RNA-seq, suggesting that it is translated (Supplemental Figure 4). The E518 residue is responsible for forming a critical hydrogen bond with the 2' hydroxyl group of the pentose ring in the RNA molecule. In silico modeling predicts that mutating amino acid 518 from glutamate to lysine would likely reduce the affinity of RNA binding to *DGCR8* (Supplemental Figure 7).

**microRNA and RNA profiling.** We combined our microRNA sequencing data from 9 schwannomas (4 with the c.1552G>A;p.E518K germline *DGCR8* variant and LOH of the alternate allele and 5 which were *DGCR8*-WT) with publicly available data from 24 Wilms tumors (4 with the somatic c.1552G>A;p.E518K *DGCR8* variant and LOH of the WT allele (7, 19), and 20 which were



**Figure 2. Canonical versus intron-derived miRNA production pathway.** Schematic diagrams of canonical miRNA processing pathway (left) and the intronic miRNA (mirtrons) processing pathway (right). Mirtrons are processed by the spliceosome in a manner completely independent of the microprocessor. Both pre-miRNAs and pre-mirtrons are then exported to the cytoplasm and further processed by DICER1. A close-up representation of the microprocessor in trimeric state (2 molecules of DGCR8 and 1 DROSHA) and the pri-miRNA are shown in the inset.

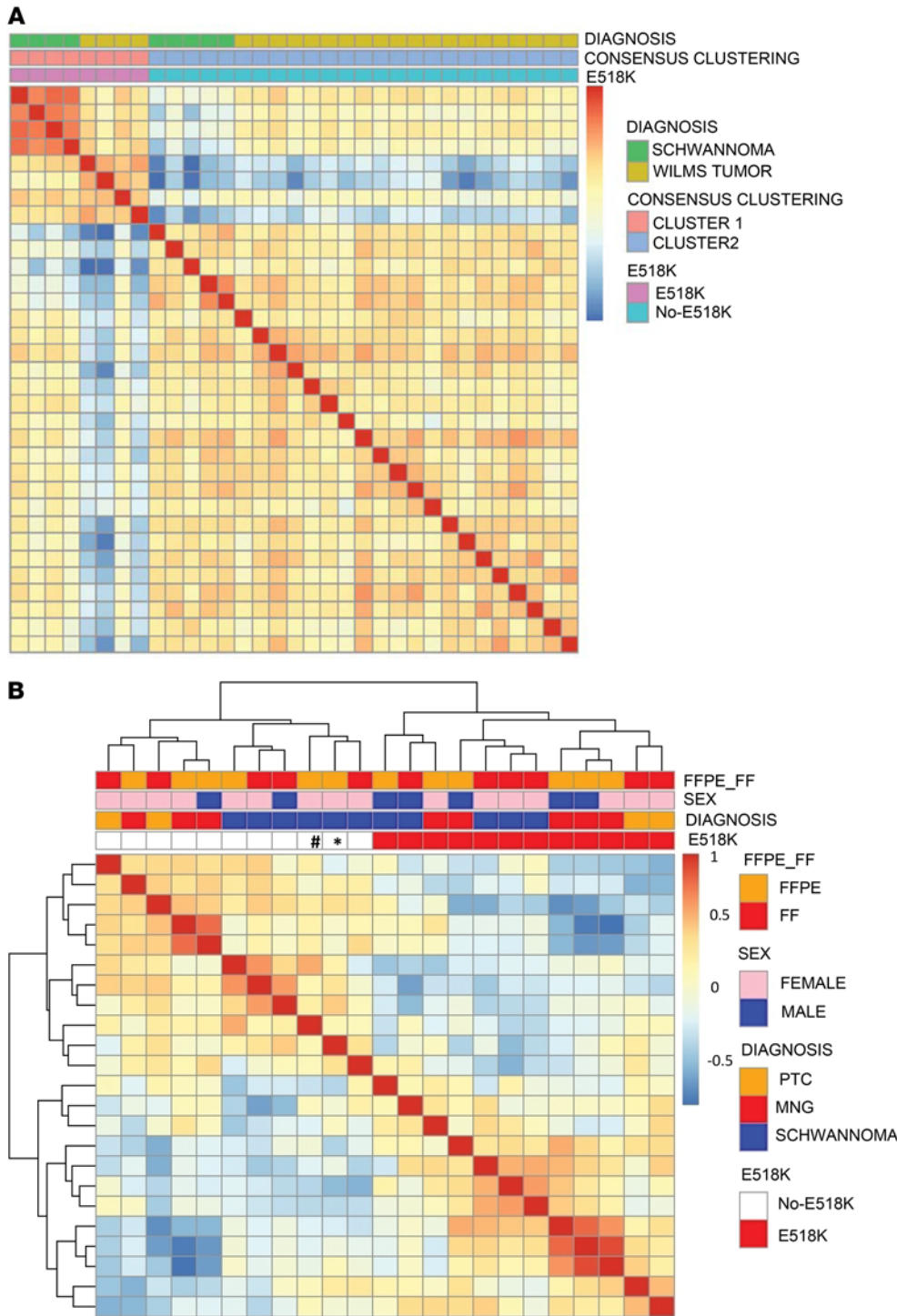
DGCR8-WT). Unsupervised consensus clustering (1000 repetitions) of miRNAs in the combined data set of 33 tumors identified 2 main clusters, 1 including only the DGCR8-mutated tumors ( $n = 8$ ) and the other with all the DGCR8-WT tumors ( $n = 25$ ) (Figure 3A and Supplemental Figure 8). Multidimensional scaling analysis supported the clustering results (Supplemental Figure 9). In total, compared with their DGCR8-WT counterparts, DGCR8 mutated schwannomas and Wilms tumors share 190 differentially expressed miRNAs (Supplemental Figure 10).

We then focused on family members. miRNA and mRNA profiling of peripheral blood mononuclear cells showed no differences between those family members heterozygous for c.1552G>A and those who did not possess this variant (data not shown). At the somatic level, we performed a NanoString analysis using 23 samples including 5 DGCR8-E518K schwannomas from 3 members of the family, and 5 MNG samples from the affected members of the family as well as 2 PTCs with the c.1552G>A;p.E518K variant. Four schwannomas, 3 MNGs, and 2 PTCs, all WT for *DGCR8*, were included as controls. In order to clarify the plausible pathogenicity of the c.1763A>G;p.K588R and the c.1147A>G;p.S383G variants, 2 additional schwannomas, 1 harboring each variant, were included in the analysis. DGCR8-E518K tissues clustered together, confirming a common miRNA profile driven by the mutation and independent of the tissue of origin. In contrast, both c.1763A>G;p.K588R and the c.1147A>G;p.S383G tumors showed a WT profile (Figure 3B).

miRNA and mRNA profiling of DGCR8-mutated schwannomas showed differences in their mRNA and miRNA expression patterns compared with DGCR8-WT schwannomas (Figure 4, A and B, and Supplemental Tables 9 and 10). Similar results were found for the expression pattern of the pre-miRNAs (Supplemental Figure 11), and, despite reduced power to identify statistically significant alterations (false discovery rate < 1%), several pre-miRNAs were differentially expressed compared with WT DGCR8

tumors. As expected, comparison between the 2 groups confirmed that differentially expressed precursor and mature miRNA levels were altered in the same direction, indicating that deficiency of the precursor miRNA likely results in underproduction of its corresponding mature miRNA (Figure 4A and Supplemental Figures 11 and 12). To confirm that the deficiency in mature miRNA is due to a defective cleavage of the corresponding pre-miRNA, we performed an in vitro cleavage experiment. For this purpose, we chose *miR-30c-2*, as it: (a) showed the most highly significant *P* value among the 190 common differentially expressed miRNAs in the 2 tumor sets (Supplemental Figure 10 and Supplemental Tables 9 and 11) and (b) was also confirmed to be an underrepresented pre-miRNA in the mutated schwannomas (Supplemental Figure 12). As a control we chose miR-223, a miRNA that was invariant in both data sets. The in vitro data validated that DGCR8-E518K is incapable of trimming primary-*miR-30c-2* into the precursor form of *miR-30c-2* (Figure 4C) while it has no effect in the processing of *pri-miR-223* (Supplemental Figure 13).

miRNAs derived from introns (known as mirtrons) are small RNAs processed by the spliceosome independently of the nuclear complex formed by DGCR8 and DROSHA (known as microprocessor) (ref. 21 and Figure 2). Therefore, we hypothesized that, unlike canonical miRNAs, mirtrons should not exhibit decreased expression from defects in microprocessor function. Out of the total of 139 mirtrons expressed in the schwannoma data set ( $n = 9$  for sample,  $n = 4$  for DGCR8-E518K, and  $n = 5$  for DGCR8-WT), 113 (81%) were overrepresented in DGCR8-mutated schwannomas. No mirtrons were underrepresented in mutated tumors (Figure 4B and Supplemental Table 12). The same analysis was then applied to the Wilms tumor data sets with similar results, validating our hypothesis (Figure 4D). In parallel, we performed the same analysis using an in-house data set of pituitary blastomas (22) with mutations in *DICER1* ( $n = 3$ ) compared with normal fetal ( $n = 3$ )

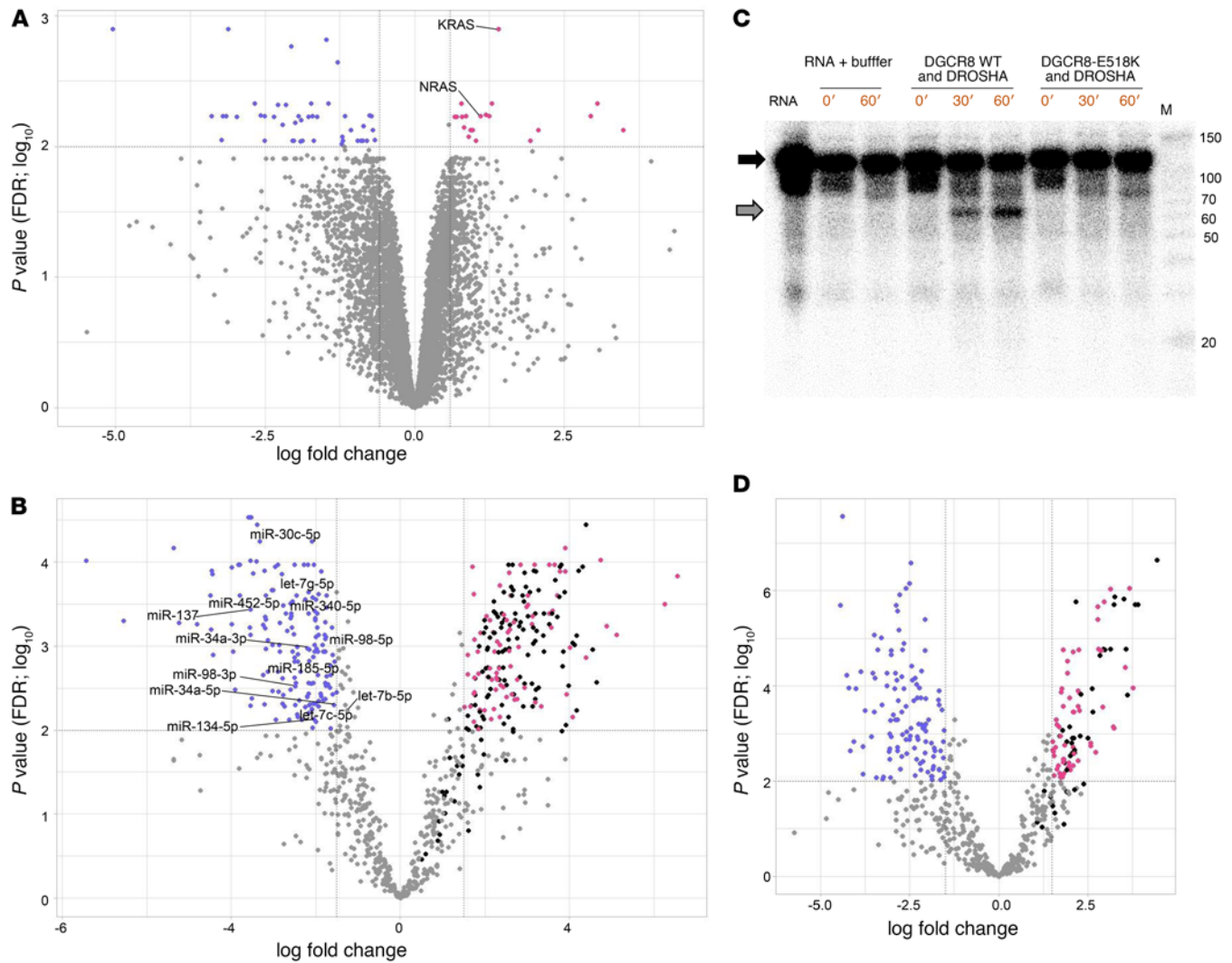


**Figure 3. Clustering of miRNA expression in *DGCR8*-c.1552G>A;p.E518K mutated tumors and their WT counterparts.** (A) Clustering of miRNA expression in 4 DGCR8-E518K mutated schwannomas, 5 DGCR8-WT schwannomas, 4 DGCR8-E518K mutated, and 20 DGCR8-WT Wilms tumors analyzed by the TARGET initiative (7). Heatmap shows sample-by-sample correlation matrix, based on Pearson correlation coefficients, using normalized values for 300 most variable miRNA expression across all samples. The DGCR8-E518K mutated tumors clustered together in the same consensus cluster. DGCR8 denotes which cases harbor the *DGCR8*-c.1552G>A;p.E518K variant (mut) and which are *DGCR8*-WT. See also Supplemental Figures 8 and 9 and ref. 7. (B) Unsupervised clustering of the top 50 most variably expressed miRNAs based on NanoString data of 8 MNG samples (5 DGCR8-E518K mutated and 3 WT cases), 4 follicular variant papillary thyroid cancers (PTC) (2 DGCR8-E518K mutated and 2 WT), and 11 schwannomas (5 DGCR8-E518K mutated; 1 with the VUS c.1147A>G;p.S383G case [\*]; 1 with germline variant c.1763A>G;p.K588R previously described in Wilms tumors [#]; and 4 WT for *DGCR8*). DGCR8-E518K mutated samples clustered together independently of the tissue of origin. Both c.1147A>G;p.S383G and c.1763A>G;p.K588R variants clustered with WT tumors, suggesting the variants do not disturb the miRNA processing role of *DGCR8*. While tissue types are distributed randomly among DGCR8-E518K, WT samples clustered by tissue of origin (thyroid vs. schwannoma cells), highlighting the particularity of the c.1552G>A;p.E518K profile in miRNA processing.

and adult ( $n = 4$ ) pituitaries. In this scenario, 13% of expressed mirtrons were overrepresented and 8% were underrepresented in DICER1-mutated tumors (Supplemental Figure 14 and Supplemental Table 12). Taken together, these findings confirmed that the differences in the miRNA profiles of DGCR8-mutated and WT tumors result from a defective microprocessor.

Finally, we interrogated the mRNA profiles of DGCR8-mutated schwannomas. The most significantly differentially expressed mRNA in DGCR8-mutated as compared with WT schwannomas

was KRAS (fold change 2.6;  $P = 3.03^{-7}$ ; FDR =  $1.27^{-3}$ ), and NRAS was among the top 10 (fold change 2.1;  $P = 2.42^{-5}$ ; FDR =  $5.89^{-3}$ ) (Figure 4A and Supplemental Table 10). In parallel, to interrogate pathways that are relevant in DGCR8-mutated schwannomas, we performed a Gene Set Enrichment Expression Analysis (GSEA) and focused in the 50 *Hallmark* gene set from the Molecular Signatures Database (MSigDB) (23, 24). The top 5 significantly enriched gene sets with FDR less than 0.01 (Normalized Enrichment Score [NES] > 1.9 or NES < -1.9) included genes downregulated by KRAS



**Figure 4. Differentially expressed mRNA and miRNA analysis of tumors with or without *DGCR8* mutation and in vitro cleavage of miR-30c-2.** For all 3 volcano plots, log fold-change is plotted on the x axis and the adjusted *P* value (FDR;  $-\log_{10}$  scale) on the y axis. Dotted horizontal and vertical lines indicate threshold of significance (FDR < 0.01) and absolute fold change (>2). Over- and underexpressed mRNAs/miRNAs in mutated cases compared with WT are shown in red and blue, respectively. Black dots represent the mirtrons. **(A)** Results of differentially expressed mRNA expression analysis between schwannomas with ( $n = 4$ ) and without ( $n = 5$ ) the c.1552G>A;p.E518K mutation. **(B)** Results of differentially expressed miRNA expression analysis between schwannomas with ( $n = 4$ ) and without ( $n = 5$ ) the c.1552G>A;p.E518K mutation. One of the most significantly underexpressed pre-miRNAs (as shown in Supplemental Figure 12), *miR-30c-2*, is a precursor to the most significantly underexpressed mature miRNA. All significant mirtrons are overrepresented in mutated cases compared with WT. The top 25 mRNAs and miRNAs, up and down, are listed in Supplemental Tables 9 and 10, respectively. **(C)** In vitro cleavage results of *pri-miR30c-2*. The microprocessor complex formed with DGCR8-WT trims the primary miRNA into a precursor miRNA, but the mutant DGCR8-E518K fails to cleave the primary miRNA even after 60 minutes. Reaction incubation time is shown in minutes. The image is representative of 3 independent replicate experiments. In each experiment, 3 conditions were tested, and freshly immunoprecipitated proteins were used in each case. DGCR8-WT and DROSHA, pull down of the complex between DGCR8-WT protein and endogenous DROSHA; DGCR8-E518K and DROSHA, pull down of the complex between DGCR8-E518K mutant protein and endogenous DROSHA; M, RNA-labeled ladder marker; black arrow shows the band corresponding to the primary miRNA (approximately 100 bp); gray arrow points the band corresponding to a precursor miRNA (approximately 65 bp). **(D)** Differentially expressed miRNA expression analysis between Wilms tumors with ( $n = 4$ ) and without ( $n = 20$ ) DGCR8-E518K mutation. All significant mirtrons are overrepresented compared with WT tumors.

activation (FDR = 0.002, NES = -1.9). The other 4 were MYC target genes, interferon alpha response, interferon gamma response, and genes regulated by NF- $\kappa$ B in response to TNF (Supplemental Figure 15 and Supplemental Table 13).

**Discussion**

In this 3-generation family with euthyroid MNG and schwannomatosis, a germline c.1552G>A;p.E518K variant in *DGCR8*

appears to have a causal role for both MNG and schwannoma susceptibility. Like *DICER1*, one of the few genes known to be implicated in familial MNG (3), *DGCR8* has a critical role in miRNA biogenesis (Figure 2).

*DGCR8* localizes to chromosome 22 next to the established schwannoma genes *LZTR1*, *SMARCB1*, and *NF2* (Supplemental Figure 16). Loss of the entire remaining WT chromosome 22 was the accompanying somatic genetic lesion in 5 *DGCR8*-mutated

schwannomas studied by WES. One schwannoma had a somatic predicted truncating mutation in *NF2* in *cis* with the c.1552G>A;p.E518K variant (Supplemental Table 4). Moreover, the only common alteration observed in PTCs, MNGs, and CPT arising in c.1552G>A;p.E518K *DGCR8* heterozygote persons at the somatic level is a loss of chromosome 22 (Supplemental Figures 5 and 6). This specific combination suggests a critical role for c.1552G>A;p.E518K in predisposing to tumor development. Somatic loss of chromosome 22, including the remaining WT *DGCR8* allele, appears to be required for tumorigenesis.

Given its location within the most common deletion implicated in the chromosome 22q11.2 deletion syndrome (25), *DGCR8* has often been discussed as a candidate gene to explain Di George syndrome (OMIM 188400) (25) (i.e., those persons with characteristics of chromosome 22q11.2 deletion syndrome but without the presence of a deletion, for example those with intragenic pathogenic variants in *TBX1*, ref. 26). Although a clinical diagnosis of autism spectrum disorder was reported in individual III-1 (Figure 1), it is not clear that the characteristics and phenotypes observed in family members are related to the variant (detailed in the Supplemental Text). A causal relation between 22q11-related syndromes and cancer development has not been established, but some instances of tumor development have been reported in 22q11 syndrome patients (27–29). Neither goiter nor schwannomas have been reported to occur in this syndrome. It is likely that biallelic loss of large stretches of chromosome 22q that would need to occur for schwannomas to arise in this situation are not compatible with cellular growth.

To exert its role in the microRNA biogenesis, 2 paired *DGCR8* proteins bind 1 DROSHA, forming a trimeric nuclear complex known as the microprocessor (ref. 30 and Figure 2). *DGCR8* plays a key role in early development, as embryonic stem cell murine *Dgcr8* knockout results in a complete loss of microRNAs and loss of self-renewal regulation (31). *DGCR8* knockdown also enhances transformation and tumor growth (32). Conditional deletion of *Dgcr8* in murine Schwann cells during development leads to an increase in proliferation and failure of Schwann cell differentiation (33). Moreover, *Dgcr8* is required for correct maintenance of myelination in adult mice (34). Similarly, an increase in proliferation as well as disruption in proper myelination after injury is observed in *Nf2* null Schwann cells (35). These findings, taken together with publicly available data showing that approximately 2.5% to 3% of Wilms tumors possess a somatic c.1552G>A;p.E518K variant accompanied by LOH of the WT allele (7) and a review of the literature (Supplemental Figure 17), lead us to conclude that biallelic loss-of-function mutations in *DGCR8* are unlikely to be selected for in most tumors. Rather, a missense alteration producing an altered protein such as p.E518K would be favored. This model resembles the *SMARCB1* tumorigenesis model in schwannomatosis that necessitates a hypomorphic mutation and LOH of the WT allele.

miRNA profiling in different tissue types demonstrated a common alteration in miRNA biogenesis driven by the c.1552G>A;p.E518K-*DGCR8* variant. Furthermore, our data from *DGCR8*-E518K-mutated schwannomas, as well as the reanalysis of *DGCR8*-E518K-mutated Wilms tumors (Figure 4, B and D, and Supplemental Tables 9 and 11) as compared with their respective

WT counterparts, showed differences in miRNA profiles whereby only canonical miRNAs were underrepresented, in keeping with the function of *DGCR8* in processing canonical microRNAs' primary transcripts. Mouse embryonic stem cells deficient for *DGCR8* show a less severe phenotype than those that are *DICER1* deficient, and some of these differences have been attributed to the existence of *DICER1*-dependent but *DGCR8*-independent small RNAs, including those known as mirtrons (ref. 36 and Figure 2). Unlike *DICER1*-mutated tumors, in which mirtrons showed no biases in representation (Figure 4, B and D, Supplemental Figure 14, and Supplemental Table 12), the lack of reduction in mirtrons in both *DGCR8*-mutated schwannomas and *DGCR8*-mutated Wilms tumors supports the pathogenicity of c.1552G>A;p.E518K and highlights mechanistic differences in the biology underlying the 2 syndromes.

We compared the expression profiles of schwannomas arising in *DGCR8* heterozygotes with their WT counterparts. RAS pathway members (*KRAS*, *NRAS*) were overexpressed (Figure 4A). Our GSEA identified 5 gene sets significantly altered in *DGCR8*-mutated schwannomas, one of which is composed of proteins downregulated upon *KRAS* activation pathway (24). Notably, germline variants of *LZTR1* have been described to cause Noonan syndrome, a classical RASopathy, and *LZTR1* has been described to drive tumorigenesis through promoting RAS/MAPK signaling either by a defective degradation of *RIT1* or a decrease in RAS ubiquitination (37–39). Moreover, a case report of a child with Costello syndrome and a germline *KRAS* pathogenic variant described the development of peripheral schwannomatosis (40) which, together with a report of somatic *BRAF* and *KRAS* mutations in peripheral schwannomas (41), highlights a potential link between schwannomas and activation of the RAS pathway that could be critical for the development of a group of peripheral schwannomas (10, 11).

In summary, we characterized what appears to be a novel, previously unreported autosomal dominant mendelian syndrome, familial multinodular goiter with schwannomatosis, which is attributable to a failure in miRNA processing. The *DGCR8* variant, c.1552G>A;p.E518K, alters canonical miRNA production, thereby impacting the ratio of canonical-to-intron-derived miRNA levels and giving rise to a miRNA profile identifiable across different tissue types.

## Methods

### Patients and samples

Blood from 6 affected members from the index family and 3 nonaffected members, 7 formalin-fixed paraffin-embedded (FFPE) blocks (1 choroid plexus tumor [WHO I] from III-1, 1 schwannoma from I-1, and 5 MNGs from I-1, II-2, III-1, III-2, and III-3) and 7 fresh frozen tumors (FFTs) (4 schwannomas from II-2, 2 from III-3, and 1 from III-2) were collected. We studied a series of cases composed of 273 cases in total: 74 FFPE CPTs from 74 patients (29 choroid plexus papillomas [CPPs], 21 atypical choroid plexus papillomas [aCPPs], and 24 choroid plexus carcinomas [CPCs]). The CPTs analyzed in this study have been registered in the choroid plexus tumor registry of the International Society of Pediatric Oncology (CPT-SIOP): a total of 181 schwannomas from 163 patients were analyzed (135 samples were from patients recruited through the Institute of Neurology in Muenster, Germany; 32 samples were from patients from the McGill University Health Centre in

Montreal, Canada; and 14 samples were from patients from the Montreal Neurological Institute, Canada). Thirty-six of 181 samples were from 19 patients (17 samples from 11 patients from Montreal and 19 samples from 8 patients from Muenster) with suspected and/or confirmed schwannomatosis but no family history reported. Three of the 181 schwannomas were relapse samples. Ninety-nine of the 181 samples were spinal schwannomas and 82 were peripheral nerve schwannomas. We also studied 18 blood samples from patients with familial MNGs or who developed MNG plus another abnormality suggesting a DICER1-like phenotype.

A set of 315 thyroid cancers and 106 benign thyroid nodules were genotyped for the c.1552G>A;p.E518K *DGCR8* locus by Sanger sequencing. For the NanoString (NanoString Technologies) experiment, 4 fresh frozen extra thyroid tumors were included, of which 2 harbored the c.1552G>A;p.E518K *DGCR8* mutation and 2 were *DGCR8*-WT cases. All 4 tumors histologically correspond to the follicular variant of papillary thyroid cancer and were obtained through the University Health Network Tumor Bank and reviewed by an expert endocrine pathologist. FFPE CPTs and schwannoma samples were reviewed according to 2007 WHO criteria by 3 senior neuropathologists. MNGs were reviewed by an expert thyroid pathologist. Tumor content was selected for DNA and miRNA extractions.

#### Whole-exome sequencing

WES was performed in the germline DNA from the family members, tumor DNA from the schwannoma of I-1, 4 schwannomas of II-2, and tumor DNA from the CPP of III-1. WES was performed at the McGill University and Génome Québec Innovation Centre (MUGQIC). Blood DNA and fresh frozen tumor-derived DNA (200 ng) or FFPE-derived DNA (50 ng) of each subject underwent exome capture using the SureSelect Human All Exon V6 kit from Agilent Technologies followed by 125-bp paired-end sequencing on an Illumina HiSeq 4000 sequencer. Analysis methods are described in detail in the Supplemental Methods. WES data have been included in the dbGAP repository (accession number EGAS00001004038).

#### *DGCR8* screening methods in schwannomas, CPTs, MNGs, and thyroid nodules

This information is summarized in Supplemental Table 7.

**CPT series.** During the validation phase, 24 CPTs (8 CPPs, 8 aCPPs, and 8 CPCs) were also studied by WES (following the protocol described above). Tumor DNA from 50 sporadic CPTs, 67 schwannomas (66 primary tumors and 1 recurrence), and germline DNA from 18 MNGs of suspected hereditary origin was sequenced using a custom Fluidigm Access Array (Fluidigm), which targets all exons and exon-intron boundaries of *DGCR8* following the same methodology as for the *DICER1* Fluidigm Access Array previously published (42).

**Schwannoma series.** During the validation phase, we also interrogated DNA from 61 schwannomas (11 cases of which were previously sequenced using the custom *DGCR8* Fluidigm Access Array) for evidence of other schwannoma genes using the Haloplex<sup>HS</sup> assay (Agilent) that incorporates molecular barcodes for high-sensitivity sequencing as a custom design (43). Tumor DNA from the 11 schwannomas that were sequenced in parallel using the Fluidigm Access Array were also sequenced using the Haloplex<sup>HS</sup> assay to cross-validate the results of both experiments. The extra 64 schwannomas were interrogated by Sanger sequencing (see below).

**MNGs and thyroid nodules.** The germline DNA of 18 patients with familial MNGs or that developed MNG plus another abnormality suggesting a *DICER1*-like phenotype was screened for variants in *DGCR8* by Sanger sequencing as described below. The E518K variant was also screened in a total of 421 thyroid tumors by Sanger sequencing.

#### Targeted sequencing analysis methods

The Fluidigm-generated data sets were analyzed as previously described (44). In brief, the Haloplex<sup>HS</sup> targeted capture method is specifically designed to identify low allele-frequency variants through the attachment of a 10-nucleotide-long molecular barcode to the captured sample DNA molecules. We followed an optimized protocol previously described in ref. 43. The design used in this study encompasses the full *SMARCB1* and *DGCR8* loci, plus the full coding region and exon-intron boundaries of *NF2*, *LZTR1*, *ARIDIA*, *ARID1B*, *DDRI*, *CHEK2*, and *LATS1* schwannoma-associated genes.

**Haloplex<sup>HS</sup> data sets.** Reads were trimmed, aligned to hg19, and deduplicated by molecular ID using Agilent SureCall v4.0.1.46. Only reads that were sequenced at least twice, i.e. for which the same molecular ID was found twice, were kept. The resulting deduplicated BAMs were used in variant calling with HaplotypeCaller and GenotypeGVCFs v3.7 per GATK best practices (45). The variant call file was annotated and loaded into a Gemini database (46), aggregated, and selected according to snpEFF predictions. Finally, variants were manually validated against read alignments using IGV software V.2.3 (47). The percentage of homozygosity was used to calculate LOH scores in the sample set that was studied by Haloplex<sup>HS</sup>. LOH scores for each variant were calculated as  $0.5 - altFreq$ , where *altFreq* represents the alternate allele frequency (alternate reads/total reads). This gives a value of 0 for perfectly heterozygous variants (50%) and 0.5 for perfectly homozygous alternate or reference variants (0% or 100%). Therefore, variants with higher values have higher LOH scores.

#### Sanger sequencing

The full *DGCR8* coding region (NM\_022720) and exon-intron boundaries were PCR-amplified and Sanger-sequenced for tumor DNA extracted from 67 schwannoma samples.

In 421 fine-needle aspirate DNA samples from thyroid tumors, the c.1552G>A;p.E518K *DGCR8* mutation locus was genotyped using PCR and Sanger sequencing. Finally, *DGCR8* (NM\_022720) single nucleotide variants identified in the CPTs, schwannomas, and MNGs by Fluidigm Array, Haloplex<sup>HS</sup> assay, or WES were amplified by PCR and followed by Sanger sequencing at the MUGQIC to verify the findings from NGS approaches. Primers are described in Supplemental Table 14.

#### Multiplex ligation-dependent probe amplification (MLPA) of the *NF2*, *SMARCB1*, and *LZTR1* genes

We used commercial MLPA kits for gene analysis (SALSA P044 *NF2*, SALSA P258-C1 *SMARCB1*, SALSA P455-A1 *LZTR1*). Information on the probe sequences and ligation sites can be found at <http://www.mlpa.com>. The MLPA protocol was performed as described by the manufacturers, using 100 ng DNA from control and patient samples. The data analysis was performed with MLPA software (JSI Medical Systems).

#### Loss-of-heterozygosity analysis

In the MNG samples from II-1, III-1, III-2, and III-3, LOH status was studied in 3 different nodules from each case using the Haloplex<sup>HS</sup> cap-



ture and calculated as described above. LOH at the c.1552G>A;p.E518K loci was validated by Sanger sequencing, including the MNG sample from II-2 and the schwannomas from III-2 and III-3. Primers are described in Supplemental Table 14.

#### Genome-wide allelic imbalance analysis

For genome-wide LOH analysis, ExomeAI (48) was applied using WES data from tumors of the index family (4 schwannomas from II-2, 1 schwannoma from I-1, 1 CPP from III-1). In addition, publicly available WES data were obtained from 2 unrelated PTCs with DGCR8-E518K mutation (TCGA-EM-A2CR, TCGA-EM-A3AP) through the TCGA data portal (20). A Wilms tumor with DGCR8-E518K mutation and its corresponding germline was also analyzed using publicly available WES data from the TARGET project (TARGET-50-PAJMIZ; phs000218.v21.p7) (18). Access to patient genetic data is controlled by dbGaP. Permission to access was granted through dbGaP (<http://dbgap.ncbi.nlm.nih.gov/aa/wga.cgi?page=login>). WES data were employed to interrogate the presence of LOH, including copy number losses and copy-neutral LOH.

#### Genome-wide copy number analysis

In parallel to the genome-wide LOH analysis, OncoScan CNV SNP arrays were performed to study somatic chromosomal aberrations in 4 MNG samples from II-2, III-1, III-2, and III-3, as well as in the 2 unrelated PTCs with the DGCR8-E518K mutation as previously described (49). Array data has been included in the GEO repository (accession number GSE135374).

#### Control population

A control population of 1433 healthy individuals (896 females and 537 males) from Montreal comprised of different ethnicities including White, Caribbean, Latino, African, and Asian ancestry was used to investigate the presence of the variant c.1552G>A;p.E518K in the germline. The population median age was 49 years (15–93 years), mean was 49 years, and SD was 18.07 years.

#### Genotyping

The c.1552G>A;p.E518K mutation was genotyped using a TaqMan custom assay. The custom TaqMan SNP genotyping assay was designed by and ordered through Life Technologies (catalog 4332072). Each reaction was done on 20 ng genomic DNA using the TaqMan Genotyping Master Mix (catalog 4381656). The following cycling was performed in an Eppendorf pro 384 thermocycler: 95°C 10 minutes, 50 cycles of 92°C 15 seconds and 60°C 90 seconds, 4°C hold. The End-point genotyping results were read and analyzed using the LightCycler 480 from Roche.

#### RNA and miRNA isolation

Total RNA was extracted from patient leukocytes and fresh frozen tissues using the MirVana Isolation Kit (Ambion). miRNA from FFPE samples was isolated using the miRNeasy FFPE kit (Qiagen).

#### mRNA analysis

RNA was reverse transcribed into cDNA using SuperScript III first-strand cDNA synthesis (Thermo Fisher Scientific). DGCR8 transcript NM\_022720 was used to design cDNA-specific PCR primers for the c.1552G>A;p.E518K mutation. Presence of a modified transcript in

blood leukocyte-derived cDNA was tested using PCR followed by Sanger sequencing. Primers are described in Supplemental Table 14.

#### RNA sequencing, miRNA sequencing, and pre-miRNA sequencing

RNA sequencing was performed at MUGQIC using the Illumina HiSeq 4000. Total RNA from blood and tumor samples was quantified using a NanoDrop Spectrophotometer ND-1000 (NanoDrop Technologies, Inc.) and its integrity was assessed on a 2100 Bioanalyzer (Agilent Technologies). rRNAs were depleted from 400 ng of total RNA using the Ribo-Zero rRNA Removal kit specific for HMR RNA (Illumina). Residual RNA was cleaned up using the Agencourt RNA-Clean™ XP Kit (Beckman Coulter) and eluted in water. cDNA synthesis was achieved with the NEBNext RNA First Strand Synthesis and NEBNext Ultra Directional RNA Second Strand Synthesis Modules (New England BioLabs). The remaining steps of library preparation were done using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England BioLabs). Adapters and PCR primers were purchased from New England BioLabs. Libraries were quantified using the Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies) and the Kapa Illumina GA with Revised Primers-SYBR Fast Universal kit (Kapa Biosystems). Average size fragment was determined using a LabChip GX (PerkinElmer) instrument. The libraries were then sequenced on an Illumina HiSeq4000 across 4 paired-end 100-bp flow-cell lanes.

In parallel, miRNA libraries were generated from 1000 ng total RNA (in blood and fresh frozen tumor RNA) using the NEBNext Multiplex Small RNA Library Prep Set for Illumina (New England Biolabs), as per the manufacturer's recommendations. cDNA construct purification was performed on a Pippin Prep instrument (SAGE Science). Final libraries were quantified using the Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies) and the Kapa Illumina GA with Revised Primers-SYBR Fast Universal kit (Kapa Biosystems). Average size fragment was determined using a LabChip GX (PerkinElmer) instrument. The resulting miRNA library was then subjected to a specific SAGE size selection of 20-bp to 30-bp insert size (standard practice for mature miRNA sequencing). The size-selected miRNA libraries were sequenced on three 50-bp single-end read lanes of an Illumina HiSeq2500 sequencer. Pre-miRNA sequencing was performed on RNA from the tumor samples (4 DGCR8-mutated schwannomas and 4 DGCR8-WT schwannomas). For this purpose, 1000 ng total RNA was used to prepare libraries and then subjected to a size selection of 30-bp to 90-bp insert size with a SAGE cassette and subsequently sequenced using 100-bp single-end sequencing on an Illumina HiSeq 4000 sequencer. Analysis pipeline is described in the Supplemental Methods section. RNAseq data have been included in the dbGAP (SRA) repository (accession number EGAS00001004038).

#### NanoString

A total of 23 samples were interrogated for the miRNA profiling. The sample set included the following: (a) 11 schwannoma samples (3 DGCR8-E518K mutated schwannomas from II-2, 1 DGCR8-E518K mutated schwannoma from I-1, and 1 DGCR8-E518K mutated schwannoma from III-2; 1 DGCR8-S383G and 1 DGCR8-K588R schwannomas from the sporadic series of cases plus 4 DGCR8-WT schwannomas including 3 fresh frozen; and 1 FFPE); (b) 4 follicular variant of papillary thyroid cancers (referred to as PTCs) fresh frozen samples (2 known to be DGCR8-E518K and 2 that were DGCR8-WT) obtained from the University Health Network Tumor Bank, all of

which tested negative for *DICER1* hotspot mutations; (c) 8 MNG samples (5 DGCR8-E518K mutated from II-1, II-2, III-1, III-2, and III-3, and 3 MNG samples that were DGCR8-WT), all of which tested negative for coding sequence *DICER1* mutations.

miRNA differential expression profiles in the DGCR8-mutated tumors versus DGCR8-WT samples of the same histology were performed using NanoString technology at the LDI Molecular Pathology Research Core. NanoString nCounter Human v3 miRNA Expression Assay (Nanostring Technologies) according to the manufacturer's instructions. To avoid bias and batch effects, FFPE and FFT samples were evenly distributed among 2 chips. Samples from different tissue types and mutational status were also distributed among the 2 chips. In brief, approximately 100 ng purified total RNA was used for miRNA sample preparation (miRNA tagging following an annealing, ligation, and purification protocol). Next, using the miRNA codeset hybridization protocol, denatured samples were hybridized with the reporter and capture probes at 65°C for 16 hours. The samples were then processed with the nCounter Preparation Station to purify the hybridized targets and affix them to the cartridge for imaging using the nCounter Digital Analyzer (CCD camera). Barcodes were counted for each target molecule at maximum resolution 555 fields of view (FOV). The code-set incorporated 800 mature miRNAs based on miRbase v21, as well as 6 positive controls, 8 negative controls, 6 ligation controls, 5 spike-in controls, and 5 mRNA housekeeping controls (*B2M*, *GAPDH*, *RPL19*, *ACTB*, and *RPLP0*). Initial data QC and extraction of raw data was performed using the nSolver Analysis Software v4.0 (NanoString Technologies). Using a linear model, the data was corrected for the effect of FFPE and Tissue Type (50), and then used for expression profiling. Sample-by-sample unsupervised hierarchical clustering was performed based on Pearson correlation coefficients, using the 50 most variable miRNAs across all samples.

### Immunohistochemistry

INI1 (SMARCB1) expression in the tumor was studied by IHC in FFPE samples. INI1 antibody (anti-BAF47) 612110 from BD Transduction Laboratories (1:50) was optimized in a BMK ultra Ventana machine following recommended protocols (16 minutes of incubation time at 36°C). INI1 expression and localization was analyzed by the pediatric neuropathologist, blinded to mutational analyses.

### In silico modeling

To assess the potential effects of the *DGCR8* mutation on RNA binding properties, we constructed models for p.E518K based on structures of DGCR8 and the known structure of the miRNA processing gene TRBP-dsRBD bound to RNA. The DGCR8 mutant p.E518K was modeled using the structure of the RNA-binding protein (TRBP) (Protein Data Bank ID: 5N8L), with a 39% sequence identity to DGCR8, where the residue of interest (E518) is conserved (51). We modeled the DGCR8-E518K mutation effects in the RNA binding using the program PyMOL (The PyMOL Molecular Graphics System, version 1.7.4.1 Schrödinger, LLC).

### Cell lines

The HEK293 cells were used to perform in vitro cleavage experiments. HEK293 cells were maintained in DMEM with 10% FBS (both from Wisent Inc.) and 1% penicillin-streptomycin (Gibco). Cells were thawed and passaged twice before transfection. All cells were last

tested for mycoplasma contamination on April 4, 2019, using the PCR Mycoplasma Detection Kit from abmGood. All cell lines were authenticated by STR DNA profiling in October 2018.

### Expression constructs and cell transfection

The cDNA encoding the full-length WT *DGCR8* were purchased from AddGene. A hairpin shRNA (clone ID: NM\_022720.4-997s1c1) against the *DGCR8* sequence was obtained from Sigma-Aldrich. Site mutagenesis was conducted to generate a p.E518K with Quickchange mutagenesis kit (Qiagen) following the manufacturer's protocol. Subsequently, both WT and mutant constructs were mutated to degenerate the last base of codons along the sequence target for the shRNA. The mutated and WT *DGCR8* cDNAs were amplified and cloned into a pBabe/FLAG/Puro retroviral vector. Clones were confirmed by bidirectional Sanger sequencing. For expression of the constructs, cells were cotransfected with pLKO/puro plasmid containing the hairpin against endogenous *DGCR8* together with pBabe-FLAG-DGCR8-WT or pBabe-FLAG-DGCR8-E518K plasmids using Lipofectamine 2000 (Thermo Fisher Scientific) following the manufacturer's protocol. Cell pellets were collected after 48 hours.

### DGCR8 immunoprecipitation

Cell pellets were lysed in NP40 buffer (20 mM Tris-HCl, pH 7.6, 150 mM NaCl, 1 mM EDTA 0.4% NP-40, and complete protease inhibitors) for 1 hour with constant agitation. Total cell extract was incubated with 50  $\mu$ L anti-FLAG M2 affinity gel (Sigma-Aldrich) for 2 hours with constant agitation. The affinity gel was then washed 3 times with NP40 buffer and twice with reaction buffer (20 mM HEPES-KOH pH 7.6, 100 mM KCl, 0.2 mM EDTA, and 5% glycerol). Elution of FLAG-DGCR8 was done using 150 ng/ $\mu$ L of  $\times$ 3 FLAG peptide (Sigma-Aldrich). All steps were carried out at 4°C. Endogenous DRO-SHA was also immunoprecipitated in complex with FLAG-DGCR8.

### Western blotting

Cells were washed with cold PBS and then lysed. After quantification, the proteins were separated on 8% SDS-PAGE and transferred onto nitrocellulose membranes. Membranes were blocked in 5% wt/vol nonfat dry milk, 1 $\times$  TBS, and 0.1% Tween-20 (TBST) for 1 hour then blotted overnight at 4°C with rabbit monoclonal anti-DGCR8 (1:1000, Abcam, ab191875), rabbit monoclonal anti-Drosha (1:1000, Cell Signaling, 3364), mouse monoclonal anti-Flag M2 (1:1000, Sigma-Aldrich, F1804), and mouse monoclonal anti- $\beta$ -tubulin, clone AA2 (1:1000, MilliporeSigma, 05-661). Membranes were washed and incubated with donkey anti-rabbit IgG secondary antibody (1:5000, GE Healthcare NA934) or sheep anti-mouse IgG secondary antibody (at 1:5000, GE Healthcare, NA931). The revealing method used was Amersham ECL detection (Amersham Biosciences). The experiments with pri-miR-30c-2 and pri-miR-223 were done in triplicate.

### In vitro cleavage

Pri-miR-30c-2 and pri-miR-223 were prepared by in vitro transcription using the MAXIscript T7 Transcription Kit (Thermo Fisher Scientific) in the presence of 5  $\mu$ L of  $\alpha$ -<sup>32</sup>P UTP (Perkin Elmer). Pri-miRNAs were purified and folded by incubation at 95°C for 3 minutes and 70°C for 3 minutes, then cooled down to 20°C. The reaction mix comprised 2  $\mu$ L of each pri-miRNA (10<sup>4</sup>-10<sup>5</sup> cpm), 10  $\mu$ L FLAG-DGCR8/Drosha complex, and 8  $\mu$ L reaction buffer supplemented with 7 mM MgCl<sub>2</sub>, 2 mM

DTT, and 0.5 U/ $\mu$ L Recombinant RNase Inhibitor (Takara). The reactions were incubated at 37°C to different time points (30 minutes and 60 minutes). Then, the reactions were run on a 10% UREA-PAGE gel at 200 V for 45 minutes. The Decade Marker System was used to generate a ladder of radiolabeled RNA molecules (Thermo Fisher Scientific). The gel was exposed on a Storage Phosphor Screen BAS-IP (GE Healthcare) overnight at 4°C and was visualized using a PhosphorImager (GE Healthcare).

### Statistics

For differential expression analysis, genes with significantly higher and or lower expression values were reported using an adjusted *P* value (false discovery rate) (52) threshold of 1%, and a minimum fold change of 2. The corresponding *P* values were obtained from empirical Bayes moderated *t* statistics as implemented in the limma package of the Bioconductor project. Regarding gene set enrichment analysis, gene sets with FDR less than 1% were reported. All the reported gene sets had an absolute Normalized Enrichment Score (NES) of greater than 1.9.

### Study approval

The study was approved by the IRB of the relevant institutions. Participants were recruited in compliance with the second edition of the Canadian Tri-Council Policy Statement of Ethical Conduct for Research Involving Humans and Eligible Persons or Designates and signed a consent form in accordance with the IRB approvals.

### Author contributions

BR conceived and performed experiments, managed the project, and cowrote the manuscript. JN performed and designed the RNA-seq bioinformatics analysis. SF performed the WES analysis. MAR, AC, SS, RW, VC, ZH, NH, LDK, HR, and NS contributed to the experiments. TR, AB contributed to the bioinformatics analysis. AS and KM performed clinical follow-up of the family. MPP, SA, and OM performed pathology revision. CT, MH, WP, MNHB, RP, and

YEN provided samples and clinical data. RT assisted with protocol development, clinical samples, and follow-up. RS, AVB, JR, and MRF supervised experiments and analysis. WDF conceived and supervised the study and cowrote the manuscript. Co-authorship order was established based on conceptual and intellectual contributions, managerial responsibility, and time allocation to the project.

### Acknowledgments

We would like to thank to Tina Kader for clinical advice, Marc Tischkowitz for help with ascertainment of the family, Louise Rochon for pathological diagnosis advice, and Anne-Laure Chong and HyeRim Han (Lady Davis Institute) for technical support in sample preparation. We also thank Blaise Clarke from University Health Network (Toronto, Canada) and Dianne Chadwick and Alisha Bhimani from the Tumor Bank at University Health Network (Toronto, Canada) for their help collecting the PTC samples. The NanoString experiment was outsourced to the Pathology Service Unit at the Lady Davis Institute and performed by Andreas Papadakis. This work was supported by Canadian Institutes of Health Research grants FDN-148390 (to WDF) and MOP-130425 (to MRF), a Compute Canada grant RAC wst-164 (to IR), Alex's Lemonade Stand Foundation, the Mia Neri Foundation for Childhood Cancer and Cassa di Sovvenzioni e Risparmio fra il Personale della Banca d'Italia (to BR), and infrastructural support by the KinderKrebsInitiative Buchholz/Holm-Seppensen (to RS).

Address correspondence to: William D. Foulkes, Room E757, Segal Cancer Centre, Jewish General Hospital, Montreal, Quebec H3T 1E2, Canada. Phone: 514.340.8222 ext. 23213 and 514.934.1934 ext. 76094; Email: william.foulkes@mcgill.ca.

JN's present address is: Department of Human Genetics, McGill University, Montreal, Quebec, Canada.

- Lever EG, Medeiros-Neto GA, DeGroot LJ. Inherited disorders of thyroid metabolism. *Endocr Rev.* 1983;4(3):213-239.
- Foulkes WD, Priest JR, Duchaine TF. DICER1 mutations, microRNAs and mechanisms. *Nat Rev Cancer.* 2014;14(10):662-672.
- Rio Frio T, et al. DICER1 mutations in familial multinodular goiter with and without ovarian Sertoli-Leydig cell tumors. *JAMA.* 2011;305(1):68-77.
- de Kock L, et al. Deep sequencing reveals spatially distributed distinct hot spot mutations in DICER1-related multinodular goiter. *J Clin Endocrinol Metab.* 2016;101(10):3637-3645.
- Khan NE, et al. Quantification of thyroid cancer and multinodular goiter risk in the DICER1 syndrome: a family-based cohort study. *J Clin Endocrinol Metab.* 2017;102(5):1614-1622.
- Zhang H, Kolb FA, Jaskiewicz L, Westhof E, Filipowicz W. Single processing center models for human Dicer and bacterial RNase III. *Cell.* 2004;118(1):57-68.
- Gadd S, et al. A Children's Oncology Group and TARGET initiative exploring the genetic landscape of Wilms tumor. *Nat Genet.* 2017;49(10):1487-1494.
- Snuderl M, et al. Recurrent homozygous deletion of DROSHA and microduplication of PDE4DIP in pineoblastoma. *Nat Commun.* 2018;9(1):2868.
- Torrezan GT, et al. Recurrent somatic mutation in DROSHA induces microRNA profile changes in Wilms tumour. *Nat Commun.* 2014;5:4039.
- Plotkin SR, et al. Update from the 2011 International Schwannomatosis Workshop: From genetics to diagnostic criteria. *Am J Med Genet A.* 2013;161A(3):405-416.
- Kehrer-Sawatzki H, Farschtschi S, Mautner VF, Cooper DN. The molecular pathogenesis of schwannomatosis, a paradigm for the co-involvement of multiple tumour suppressor genes in tumorigenesis. *Hum Genet.* 2017;136(2):129-148.
- Hadfield KD, et al. Molecular characterization of SMARCB1 and NF2 in familial and sporadic schwannomatosis. *J Med Genet.* 2008;45(6):332-339.
- Sestini R, Bacci C, Provenzano A, Genuardi M, Papi L. Evidence of a four-hit mechanism involving SMARCB1 and NF2 in schwannomatosis-associated schwannomas. *Hum Mutat.* 2008;29(2):227-231.
- Hutter S, et al. Whole exome sequencing reveals that the majority of schwannomatosis cases remain unexplained after excluding SMARCB1 and LZTR1 germline variants. *Acta Neuropathol.* 2014;128(3):449-452.
- Hunt SE, et al. Ensembl variation resources, Database, (Oxford). 2018. <https://doi.org/10.1093/database/bay119>. Accessed February 6, 2020.
- Karczewski KJ, et al. Variation across 141,456 human exomes genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes [preprint]. doi: <https://doi.org/10.1101/531210>. Posted on bioRxiv January 30, 2019.
- Firth HV, et al. DECIPHER: database of chromosomal imbalance and phenotype in humans using ensembl resources. *Am J Hum Genet.* 2009;84(4):524-533.
- Walz AL, et al. Recurrent DGCR8, DROSHA, and SIX homeodomain mutations in favorable histology Wilms tumors. *Cancer Cell.* 2015;27(2):286-297.
- Wegert J, et al. Mutations in the SIX1/2 pathway and the DROSHA/DGCR8 miRNA microprocessor complex underlie high-risk blastemal type Wilms tumors. *Cancer Cell.* 2015;27(2):298-311.
- Cancer Genome Atlas Research Network. Integrated genomic characterization of papillary

- thyroid carcinoma. *Cell*. 2014;159(3):676–690.
21. Westholm JO, Lai EC. Mirtrons: microRNA biogenesis via splicing. *Biochimie*. 2011;93(11):1897–1904.
  22. de Kock L, et al. Pituitary blastoma: a pathognomonic feature of germ-line DICER1 mutations. *Acta Neuropathol*. 2014;128(1):111–122.
  23. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545–15550.
  24. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1(6):417–425.
  25. McDonald-McGinn DM, et al. 22q11.2 deletion syndrome. *Nat Rev Dis Primers*. 2015;1:15071.
  26. Yagi H, et al. Role of TBX1 in human del22q11.2 syndrome. *Lancet*. 2003;362(9393):1366–1373.
  27. McDonald-McGinn DM, et al. Malignancy in chromosome 22q11.2 deletion syndrome (DiGeorge syndrome/velocardiofacial syndrome). *Am J Med Genet A*. 2006;140(8):906–909.
  28. Lambert MP, et al. The 22q11.2 deletion syndrome: Cancer predisposition, platelet abnormalities and cytopenias. *Am J Med Genet A*. 2018;176(10):2121–2127.
  29. Van L, et al. All-cause mortality and survival in adults with 22q11.2 deletion syndrome. *Genet Med*. 2019;21(10):2328–2335.
  30. Nguyen TA, et al. Functional anatomy of the human microprocessor. *Cell*. 2015;161(6):1374–1387.
  31. Melton C, Judson RL, Belloch R. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. *Nature*. 2010;463(7281):621–626.
  32. Kumar MS, Lu J, Mercer KL, Golub TR, Jacks T. Impaired microRNA processing enhances cellular transformation and tumorigenesis. *Nat Genet*. 2007;39(5):673–677.
  33. Lin HP, Oksuz I, Hurley E, Wrabetz L, Awatramani R. Microprocessor complex subunit DiGeorge syndrome critical region gene 8 (*Dgcr8*) is required for schwann cell myelination and myelin maintenance. *J Biol Chem*. 2015;290(40):24294–24307.
  34. Gökbuğut D, et al. The miRNA biogenesis pathway prevents inappropriate expression of injury response genes in developing and adult Schwann cells. *Glia*. 2018;66(12):2632–2644.
  35. Mindos T, et al. Merlin controls the repair capacity of Schwann cells after injury by regulating Hippo/YAP activity. *J Cell Biol*. 2017;216(2):495–510.
  36. Babiarz JE, Ruby JG, Wang Y, Bartel DP, Belloch R. Mouse ES cells express endogenous shRNAs, siRNAs, and other microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev*. 2008;22(20):2773–2785.
  37. Steklov M, et al. Mutations in LZTR1 drive human disease by dysregulating RAS ubiquitination. *Science*. 2018;362(6419):1177–1182.
  38. Motta M, et al. Dominant Noonan syndrome-causing LZTR1 mutations specifically affect the Kelch domain substrate-recognition surface and enhance RAS-MAPK signaling. *Hum Mol Genet*. 2019;28(6):1007–1022.
  39. Castel P, et al. RIT1 oncoproteins escape LZTR1-mediated proteolysis. *Science*. 2019;363(6432):1226–1230.
  40. Bertola DR, et al. Multiple, diffuse schwannomas in a RASopathy phenotype patient with germline KRAS mutation: a causal relationship? *Clin Genet*. 2012;81(6):595–597.
  41. Serrano C, et al. BRAF V600E and KRAS G12S mutations in peripheral nerve sheath tumours. *Histopathology*. 2013;62(3):499–504.
  42. de Kock L, et al. Sequencing of DICER1 in sarcoma identifies biallelic somatic DICER1 mutations in an adult-onset embryonal rhabdomyosarcoma. *Br J Cancer*. 2017;116(12):1621–1626.
  43. de Kock L, et al. High-sensitivity sequencing reveals multi-organ somatic mosaicism causing DICER1 syndrome. *J Med Genet*. 2016;53(1):43–52.
  44. McCluggage WG, et al. Ovarian microcystic stromal tumors are characterized by alterations in the beta-catenin-APC pathway and may be an extracolonic manifestation of familial adenomatous polyposis. *Am J Surg Pathol*. 2018;42(1):137–139.
  45. Van der Auwera GA, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics*. 2013;43:11.10.1–11.10.33.
  46. Paila U, Chapman BA, Kirchner R, Quinlan AR. GEMINI: integrative exploration of genetic variation and genome annotations. *PLoS Comput Biol*. 2013;9(7):e1003153.
  47. Robinson JT, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29(1):24–26.
  48. Nadaf J, Majewski J, Fahiminiya S. ExomeAI: detection of recurrent allelic imbalance in tumors using whole-exome sequencing data. *Bioinformatics*. 2015;31(3):429–431.
  49. Rivera B, et al. Functionally null *RAD51D* missense mutation associates strongly with ovarian carcinoma. *Cancer Res*. 2017;77(16):4517–4529.
  50. Ritchie ME, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
  51. Masliah G, et al. Structural basis of siRNA recognition by TRBP double-stranded RNA binding domains. *EMBO J*. 2018;37(6):e97089.
  52. Benjamini Y, Hochberg Y. The false discovery rate - a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol*. 1995;57(1):289–300.