DGCR8 microprocessor defect characterizes familial multinodular goiter with schwannomatosis

Barbara Rivera, … , Marc R. Fabian, William D. Foulkes

*J Clin Invest.* 2019. [https://doi.org/10.1172/JCI130206](https://doi.org/10.1172/JCI130206).

**Graphical abstract**

Find the latest version:

[https://jci.me/130206/pdf](https://jci.me/130206/pdf)
DGCR8 microprocessor defect characterizes familial multinodular goiter with schwannomatosis

Barbara Rivera PhD1,2,∗, Javad Nadaf PhD2,∗, Somayyeh Fahiminiya PhD3,#, Maria Apellaniz-Ruiz PhD2,3,4#, Avi Sasaki MD4,5, Anne-Sophie Chong BSc2, Sahil Sharma BSc6, Rabea Wagener PhD7, Timothée Revil PhD4,5, Vincenzo Condello PhD9, Zineb Harra BSc2, Nancy Hamel MSc3, Nelly Sabbaghian MSc2, Karl Muchantef MD10,11, Christian Thomas MD12, Leanne de Kock PhD2,4, Marie-Noëlle Hébert-Blouin MD13, Angelia V. Bassenden BSc14, Hannah Rabenstein7, Ozgur Mete MD15,16, Ralf Paschke MD, PhD17, Marc P. Pusztaszeri MD18, Werner Paulus MD, PhD12, Albert Berghuis PhD14, Jiannis Ragoussis PhD3,8, Yuri E. Nikiforov MD, PhD9, Reiner Siebert MD, PhD7, Steffen Albrecht MD19, Robert Turcotte MD20,21, Martin Hasselblatt MD, PhD12, Marc R Fabian PhD1,2,6,14, William D. Foulkes MBBS, PhD1-5.

1Gerald Bronfman Department of Oncology, McGill University, Montreal, Quebec, Canada; 2Lady Davis Institute, Segal Cancer Centre, Jewish General Hospital, Montréal, Québec, Canada; 3Cancer Research Program, Research Institute of the McGill University Health Centre Montreal, Quebec, Canada; 4Department of Human Genetics, McGill University, Montréal, Québec, Canada; 5Division of Medical Genetics, Department of Medicine, McGill University Health Centre, Montréal, Québec, Canada; 6Department of Experimental Medicine, McGill University, Montréal, Quebec, Canada; 7Institute of Human Genetics, University of Ulm and University of Ulm Medical Center, Ulm, Germany; 8McGill University and Génome Québec Innovation Centre, Montreal, Quebec, Canada; 9Department of Pathology, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA. 10Department of Diagnostic Radiology, McGill, University, Montreal, Quebec, Canada; 11Pediatric radiology, McGill University Health Centre, Montreal, Quebec, Canada; 12Institute of Neuropathology, University Hospital Münster, Münster, Germany; 13Department of Neurosurgery, McGill University Health Centre, Montreal, Quebec, Canada; 14Department of Medical Genetics, McGill University, Montréal, Québec, Canada; 15Division of Medical Genetics, Department of Medicine, McGill University Health Centre, Montréal, Québec, Canada; 16Department of Experimental Medicine, McGill University, Montréal, Quebec, Canada; 17Institute of Human Genetics, University of Ulm and University of Ulm Medical Center, Ulm, Germany; 18McGill University and Génome Québec Innovation Centre, Montreal, Quebec, Canada; 19Department of Pathology, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA. 20Department of Diagnostic Radiology, McGill, University, Montreal, Quebec, Canada; 21Pediatric radiology, McGill University Health Centre, Montreal, Quebec, Canada; 22Institute of Neuropathology, University Hospital Münster, Münster, Germany; 23Department of Neurosurgery, McGill University Health Centre, Montreal, Quebec, Canada.
Abstract: max word count = 250
**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 non-familial DGCR8-E518K positive tumors, including MNG, schwannomas, papillary thyroid cancers (PTC) and Wilms Tumors. MiRNA profiles of four tissue types were compared, and sequencing of miRNA, pre-miRNA and mRNA was performed in a subset of 9 schwannomas, four 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 two 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 PTC, 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.

**Conclusions:**

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

Funded by CIHR, Compute Canada, Alex’s Lemonade Stand Foundation, the Mia Neri Foundation for Childhood Cancer.
Introduction:

Familial multinodular goiter can either 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 multi-clonal in origin, in that different nodules possess one 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 two genes associated with this disorder are *SMARCB1* and *LZTR1* (11). Notably, the three genes lie adjacent to each other, covering 8.72 megabases of chromosome 22q. Different sequences of events that affect all three loci can occur, but the most frequent combination of genetic events that lead to schwannomatosis occur in three steps: 1) a hypomorphic germline pathogenic variant in *SMARCB1*; 2) 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 3) 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 \textit{SMARCB1} or \textit{LZTR1} are described\cite{11}, most cases remain unexplained\cite{14}.

To further understand the genetic contribution to both MNG and schwannomatosis, we extensively characterized nine 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.

\textbf{Results}

\textbf{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 six persons over three 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 one or more peripheral nerve schwannomas (Figure 1 and Supplementary Methods) and in three of these five, 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 Supplementary Text and Supplementary Figures 1 and 2. Individual III-1 was diagnosed with autism spectrum disorder, however no other features consistent with a DiGeorge syndrome diagnosis were identified in the patient.

\textbf{Molecular genetics studies}

As the major gene for familial euthyroid MNG is \textit{DICER1}, we performed segregation analysis for 4 affected members of this family using markers within and flanking \textit{DICER1}. The results were not consistent with linkage of MNG to 14q23-32, where \textit{DICER1} is situated (data not shown). Therefore, we performed whole exome sequencing (WES) using blood or saliva DNA from nine
family members (six affected, three unaffected). Only two 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 Supplementary 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, which is highly conserved (Supplementary 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 (Supplementary 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 01/April/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 the DGCR8 gene in the family, we sequenced the entire locus of the DGCR8 gene (Chr22:20067757-20099510 hg19) in individual II-2 germline using a Haloplex High Sensitivity (Haloplex HS) 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 three 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 HS 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, 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,
Supplementary Tables 3-6 and Supplementary 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), Supplementary Text and Supplementary 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 one 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 with Wilms
Tumors(18, 19), was identified in 5 schwannomas and in 1 CPT, for a combined frequency of
6/237 cases studied (Supplementary Table 8). The c.1552G>A;p.E518K variant was also
present in two cases from the Tumor Cancer Genome Atlas(20) PTC dataset. We collected
tumor and germline samples from these two cases and confirmed LOH on chromosome 22q
(Supplementary Figures 5 and 6). We genotyped the c.1552G>A;p.E518K variant in 315 PTCs
and 106 hyperplasic 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 two double-strand RNA-binding domains within
DGCR8 and the corresponding mRNA is expressed and detected both by RT-PCR followed by
Sanger sequencing and by RNA-seq, suggesting that it is translated (Supplementary 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
(Supplementary Figure 7).
microRNA and RNA profiling

We combined our microRNA sequencing data from 9 schwannomas (four with the c.1552G>A;p.E518K germline DGCR8 variant and LOH of the alternate allele and five which were DGCR8-wild type) with publicly available data from 24 Wilms Tumors (4 with the somatic c.1552G>A;p.E518K DGCR8 variant and LOH of the wild type allele (7, 19), and 20 which were DGCR8-wild type). Unsupervised consensus clustering (1000 repetitions) of miRNAs in the combined dataset of 33 tumors identified two main clusters, one including only the DGCR8-mutated tumors (n = 8) and the other with all the DGCR8-wild type tumors (n = 25) (Figure 3A and Supplementary Figure 8). Multidimensional scaling analysis supported the clustering results (Supplementary Figure 9). In total, compared to their DGCR8-wild type counterparts, DGCR8 mutated schwannomas and Wilms Tumors share 190 differentially expressed miRNAs (Supplementary 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 five DGCR8-E518K schwannomas from 3 members of the family, and five MNG samples from the affected members of the family as well as two PTCs with the c.1552G>A;p.E518K variant. Four schwannomas, three MNGs and two PTCs, all wild type 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, two additional schwannomas, one 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 wild type profile (Figure 3B).
MiRNA and mRNA profiling of DGCR8-mutated schwannomas showed differences in their mRNA and miRNA expression patterns compared to DGCR8-wild type schwannomas (Figure 4A-B and Supplementary Tables 9 and 10). Similar results were found for the expression pattern of the pre-miRNAs (Supplementary Figure 11) and, despite reduced power to identify statistically significant alterations (false discovery rate < 1%), several pre-miRNAs were differentially expressed compared to wild type DGCR8 tumors. As expected, comparison between the two 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 Supplementary 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 i) showed the most highly significant p-value among the 190 common differentially expressed miRNAs in the two tumor sets (Supplementary Figure 10 and Supplementary Tables 9 and 11) and ii) was also confirmed to be an underrepresented pre-miRNA in the mutated schwannomas (Supplementary Figure 12). As a control we chose miR-223, a miRNA that was invariant in both datasets. 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 (Supplementary 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)(21) (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 dataset (n sample = 9; 4 DGCR8-E518K and 5 DGCR8-wild type), 113 (81%) were overrepresented in DGCR8-mutated
schwannomas. No mirtrons were underrepresented in mutated tumors (Figure 4B and Supplementary Table 12). The same analysis was then applied to the Wilms Tumor datasets with similar results, validating our hypothesis (Figure 4D). In parallel, we performed the same analysis using an in-house dataset of pituitary blastomas (22) with mutations in DICER1 (n = 3) compared to normal fetal (n = 3) and adult (n = 4) pituitaries. In this scenario, 13% of expressed mirtrons were overrepresented and 8% were underrepresented in DICER1-mutated tumors (Supplementary Figure 14 and Supplementary Table 12). Taken together, these findings confirmed that the differences in the miRNA profiles of DGCR8-mutated and wild type 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 to wild type schwannomas was KRAS (Fold Change 2.6; p = 3.03e-7; FDR = 1.27e-3), and NRAS was among the top 10 (Fold Change 2.1; p = 2.42e-5; FDR = 5.89e-3) (Figure 4A and Supplementary 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 down-regulated by KRAS activation (FDR = 0.002, NES = -1.9). The other four were (MYC target genes, interferon alpha response, interferon gamma response and genes regulated by NF-kB in response to TNF) (Supplementary Figure 15 and Supplementary Table 13).

Discussion

In this three-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 $NF2$ (Supplementary Figure 16). Loss of the entire remaining wild type chromosome 22 was the accompanying somatic genetic lesion in five 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 (Supplementary 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 (Supplementary 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 wild type $DGCR8$ 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 DiGeorge 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 $TBX1$(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 Supplementary 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, two paired DGCR8 proteins bind one DROSHA, forming a trimeric nuclear complex known as the microprocessor\(^{(30)}\) (Figure 2). DGCR8 plays a key role in early development, as embryonic stem cell murine \(Dgcr8\) knock-out results in a complete loss of microRNAs and loss of self-renewal regulation\(^{(31)}\). \(DGCR8\) knock-down 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 wild type allele\(^{(7)}\) and a review of the literature (Supplementary Figure 17), lead us to conclude that biallelic loss-of-function mutations in \(DGCR8\) are unlikely 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 \(SMARC\-\) tumorigenesis model in schwannomatosis that necessitates a hypomorphic mutation and LOH of the wild type 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 re-analysis of \(DGCR8\)-E518K-mutated Wilms Tumors (Figures 4B and D and Supplementary Tables 9 and 11) as compared to their respective wild type 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 (Figure 2)(36). Unlike DICER1-mutated tumors, in which mirtrons showed no biases in representation (Figure 4B and D, Supplementary Figure 14 and Supplementary 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 two syndromes.

We compared the expression profiles of schwannomas arising in DGCR8 heterozygotes with their wild type counterparts. RAS pathway members (KRAS, NRAS) were overexpressed (Figure 4A). Our Gene Set Enrichment analysis 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 miRNAs production, thereby impacting the ratio of canonical to intron-derived miRNAs 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 non affected 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 (FFT) (4 schwannomas from II-2, 2 from III-3 and 1 from III-2) were collected. We studied series of cases composed in total of 273 cases: 74 FFPE CPTs from 74 patients (29 Choroid Plexus Papillomas (CPPs), 21 atypical Choroid Plexus Papillomas (aCPPs), 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); 181 schwannomas from 163 patients were recruited (135 samples from recruited through the Institute of Neurology in Münster (Germany), 32 samples from the McGill University Health Centre in Montreal (Canada), 14 samples from the Montreal Neurological Institute (Canada)). Thirty-six out of 181 samples were from 19 patients (17 samples from 11 patients from Montreal and 19 samples from 8 patients from Münster) with suspected and/or confirmed schwannomatosis but no family history reported. Three out of the 181 schwannomas were relapse samples. Ninety-nine out of 181 samples were spinal schwannomas and 82 were peripheral nerve schwannomas. We also studied 18 blood samples from patients with familial MNGs or that 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 two harbored the c.1552G>A;p.E518K DGCR8 mutation and two were DGCR8 wild-type cases. All four tumors histologically correspond to the follicular variant of papillary thyroid cancer and were obtained through the University Health Network (UHN), Toronto tumor bank and reviewed by an expert endocrine pathologist (O.M.). FFPE CPTs and schwannoma samples were reviewed according to 2007 WHO criteria by three senior neuropathologists (S.A., M.H., W.P). MNGs were reviewed by an expert thyroid pathologist (M.P.). 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, four schwannomas of II-2 and tumor DNA from the choroid plexus papilloma (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 Supplementary Methods. WES data has been included in the EGA repository (accession number EGAS00001004038).

DGCR8 screening methods in schwannomas, CPTs, MNGs and thyroid nodules.
This section is summarized in Supplementary Table 7.
CPT series: During the validation phase, twenty-four 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 one 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 HaloPlexHS 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 HaloplexHS 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 datasets were analyzed as previously described(44). In brief, the HaloPlexHS 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(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, ARID1A, ARID1B, DDR1, CHEK2, LATS1 schwannoma-associated genes.
HaloPlex\textsuperscript{HS} datasets were analyzed as follows: 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\cite{45}. The Variant Call File was annotated and loaded into a Gemini database\cite{46}, aggregated and selected according to snpEFF predictions. Finally, variants were manually validated against read alignments using IGV software V.2.3\cite{47}. The percentage of homozygosity was used to calculate LOH scores in the sample set that was studied by HaloPlex\textsuperscript{HS}. LOH scores for each variant were calculated as follows: $|0.5 - \text{altFreq}|$ 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 \textit{DGCR8} coding region (NM_022720) and exon–intron boundaries was polymerase chain reaction (PCR) amplified and Sanger sequenced for tumor DNA extracted from 67 schwannoma samples. In 421 fine needle aspirate DNA from thyroid tumors the c.1552G>A;p.E518K \textit{DGCR8} mutation locus was genotyped using PCR and Sanger sequencing. Finally, \textit{DGCR8} (NM_022720) Single Nucleotide Variants identified in the CPTs, schwannomas and MNGs by Fluidigm Array, Haloplex\textsuperscript{HS} 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 Supplementary Table 14.
Multiplex Ligation-Dependent Probe Amplification (MLPA) of the NF2, the SMARCB1 and the LZTR1 Gene

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 of 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 three different nodules from each case using the Haloplex<sup>HS</sup> capture and calculated as described above. LOH at the c.1552G>A;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 Supplementary Table 14.

Genome-wide allelic imbalance analysis

For genome wide LOH analysis, ExomeAl(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 two 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 (TARGET-50-PAJMIZ)(18); project (phs000218.v21.p7). Access to patient genetic data is controlled by dbGaP. Permission to access was granted through dbgap (https://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 two unrelated PTC 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 Caucasian, 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), mean = 49 and standard deviation (SD) = 18.07.

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 (cat# 4332072). Each reaction was done on 20 ng of genomic DNA using the TaqMan® Genotyping Master Mix (cat# 4381656). The following cycling was performed in an Eppendorf pro 384 thermocycler: 95°C 10min, 50 cycles (92°C 15sec, 60°C 90sec), 4°C hold. The Endpoint 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 Supplementary Table 14

RNA sequencing, miRNA sequencing and pre-miRNA sequencing

RNA sequencing was performed at the MUGQIC using 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). rRNA were depleted from 400 ng of total RNA using Ribo-Zero rRNA Removal kit specific for HMR RNA (Illumina). Residual RNA was cleaned up using the Agencourt RNACleanTM 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 100bp flow-cell lanes.

In parallel, miRNA libraries were generated from 1000 ng of 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 20-30bp insert size (standard practice for mature miRNA sequencing). The size-selected miRNA libraries were sequenced on three 50bp single end read lanes of an Illumina HiSeq2500 sequencer. Pre-miRNA sequencing was performed on the tumor samples RNA (4 DGCR8-mutated schwannomas and 4 DGCR8-wild type schwannomas). For this purpose, 1000ng of total RNA were used to prepare libraries and then subjected to a size selection of 30-90 insert size with a SAGE cassette and subsequently sequenced using 100bp single end sequencing on an Illumina HiSeq 4000 sequencer. Analysis pipeline is described in the Supplementary Methods section. RNAseq data has been included in the EGA repository (accession number EGAS00001004038).

**NanoString**

A total of 23 samples were interrogated for the miRNA profiling. The sample set included:

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-wild type schwannomas including 3 fresh frozen and 1 FFPE).

b) 4 follicular variant of papillary thyroid cancers (referred to as PTC) fresh frozen samples (2 known to be DGCR8-E518K and 2 that were DGCR8 wild type) obtained from the UHN Tumor bank. They all 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; plus 3 MNG that were DGCR8-wild type). All samples tested negative for coding sequence **DICER1** mutations.
miRNA differential expression profiles in the DGCR8-mutated tumors vs DGCR8-wild type 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 two chips. Samples from different tissue types and mutational status were also distributed among the two chips. In brief, ~100 ng of 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 (IHC)**

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 SA, blinded to
mutational analyses.

**In silico modelling**

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 modelled using the structure of the RNA-binding protein (TRBP) (PDB ID: 5N8L), with a 39% sequence identity to DGCR8, where the residue of interest (E518) is conserved (51). We modelled 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 medium (Wisent Inc) with 10% FBS (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 4th, 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 wild type *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 manufacturer protocol. Subsequently, both wild type and mutant constructs were mutated to degenerate the last base of codons along the sequence target for the shRNA. The mutated and wild type *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 co-transfected with pLKO/puro plasmid containing the hairpin against endogenous DGCR8 together with pBabe-FLAG-DGCR8-wild type or pBabe-FLAG-DGCR8-E518K plasmids using Lipofectamine 2000 (ThermoFisher Scientific) following the manufacturer’s protocol. Cell pellets were collected after 48h.

**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 1h with constant agitation. Total cell extract was incubated with 50ul of anti-FLAG M2 affinity gel (Sigma-Aldrich) for 2h with constant agitation. The affinity gel was then washed three 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/ul of 3X FLAG peptide (Sigma-Aldrich). All steps were carried out at 4 ºC. Endogenous DROSHA 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% w/v non-fat dry milk, 1X TBS, 0.1% Tween-20 (TBST) for 1h and blotted O/N at 4 ºC 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-β-tubulin, clone AA2 (1:1000, Millipore-Sigma, 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 triplicates.

In vitro cleavage

Pri-miR-30c-2 and pri-miR-223 were prepared by in vitro transcription using the MAXIscript™ T7 Transcription Kit (ThermoFisher Scientific) in the presence of 5 ul of α-32P UTP (Perkin Elmer). Pri-miRNAs were purified and folded by incubation at 95ºC 3 min, 70ºC 3 min and cooled down to 20ºC. The reaction mix comprised: 2ul of each pri-miRNA (10⁴-10⁵ c.p.m), 10 ul FLAG-DGCR8/Drosha complex and 8 ul of reaction buffer supplemented with 7 mM MgCl₂, 2 mM DTT, 0.5 U/ul Recombinant RNase Inhibitor (Takara). The reactions were incubated at 37ºC to different time points (30 min and 60 min). Then, the reactions were run on a 10% UREA-PAGE gel at 200V for 45 min. Decade Marker System was used to generate a ladder of radiolabeled RNA molecules (ThermoFisher Scientific). The gel was exposed on a Storage Phosphor Screen BAS-IP (GE Healthcare) overnight at 4 ºC 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 limma package of Bioconductor project. Regarding gene set enrichment analysis, gene sets with FDR less than 1% were reported. All the reported gene sets had absolute Normalized Enrichment Score (NES) of greater than 1.9.

Study approval
The study was approved by the Institutional Review Board (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 co-wrote the manuscript. JN performed and designed the RNA-seq bioinformatics analysis. SF performed the WES analysis. MAR, AC, SS, RW, VC, ZH, NH and NS contributed to the experiments. TR, AB contributed to the bioinformatics analysis. AS and KM, performed clinical follow-up of the family. MP, SA, OM performed pathology revision. CT, MH, WP, MNB, RP, YN provided samples and clinical data. RT assisted with protocol development, clinical samples and follow-up. RS, AB, JR, MRF supervised experiments and analysis. WDF conceived and supervised the study and co-wrote the manuscript. Co-authorship order was established based on conceptual and intellectual contributions, managerial responsibility and time allocation to the project.

**Acknowledgements:**

We would like to thank to Dr Tina Kader for clinical advice, Dr Marc Tischkowitz for help with ascertainment of the family, Dr Louise Rochon for pathological diagnosis advice, Anne-Laure Chong and HyeRim Han (Lady Davis Institute) for technical support in sample preparation. We also thank Dr 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 PhD.
References


2015;43(7):e47.

A, Hunziker J, et al. Structural basis of siRNA recognition by TRBP double-stranded RNA binding


Figure Legends

**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. Person II-1 was diagnosed with MNG. Person II-2 had MNG, a mature cystic
teratoma, nine schwannomas and an ovarian serous cystadenofibroma. Individual III-1 was
diagnosed with autism spectrum disorder (ASD), MNG, 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 loss of heterozygosity (LOH) in individuals I-1, II-1, II-2, III-1, III-2 and III-
3. III-2 and III-3’s schwannoma samples had a remnant of the wild type allele likely due 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 Supplementary Tables 3-6 and Supplementary Figures 5 and 6.

Arrow indicates a wild type 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 wild type
sequence at the c.1552G;p.E518 locus is shown for unaffected individual III-5. No tumour DNA from the schwannomas of III-1 and from the MNG of I-1 was available.

**Figure 2. Canonical vs 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 is shown in the inset.

**Figure 3. Clustering of miRNA expression in DGCR8-c.1552G>A;p.E518K mutated tumors and their wild type counterparts.**

A) Clustering of miRNA expression in 4 DGCR8-E518K mutated schwannomas, 5 DGCR8 wild type schwannomas, 4 DGCR8-E518K mutated and 20 DGCR8 wild type 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 wild type. See also Supplementary Figures 8 and 9(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 wild type cases); 4 follicular variant papillary thyroid cancers (PTC) (2 DGCR8-E518K mutated and 2 wild type) and 11 schwannomas (5 DGCR8-E518K mutated, 1 with the VUS c.1147A>G;p.S383G case indicated by *, 1 with the germline variant c.1763A>G;p.K588R previously described in Wilms Tumors indicated by #, and 4 wild type 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 wild type tumors, suggesting the variants do not disturb the miRNA processing role of DGCR8. While DGCR8-E518K is distributed randomly among tissue types, wild type samples clustered by tissue of origin (thyroid vs schwann cells), highlighting the particularity of the c.1552G>A;p.E518K profile in miRNA processing.

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 three volcano plots, Log Fold change is plotted on the x-axis and the adjusted p-value (False Discovery Rate [FDR]; -log10 scale) on the y-axis. Dotted horizontal and vertical lines indicate threshold of significance (FDR < 0.01) and absolute fold change (>2). Over- and under-expressed mRNAs/miRNAs in mutated cases compared to wild types 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 Supplementary Figure 12), miR-30c-2, is precursor to the most significantly under-expressed mature miRNA. All significant mirtrons are overrepresented in mutated cases compared to wild type. The top 25 mRNAs and miRNAs, up and down, are listed in Supplementary Tables 9 and 10, respectively. C) In vitro cleavage results of pri-miR30c-2. The microprocessor complex formed with DGCR8 wild type 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 three independent replicate experiments. In each experiment, three conditions were tested, and freshly immunoprecipitated
proteins were used in each case. DGCR8 wild type & DROSHA: pull down of the complex between DGCR8-wild type protein and endogenous DROSHA; DGCR8-E518K & DROSHA: pull down of the complex between DGCR8-E518K mutant protein and endogenous DROSHA; M: RNA-labelled ladder marker; black arrow shows the band corresponding to the primary miRNA (~100bp); gray arrow points the band corresponding to a precursor miRNA (~65bp). 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 to wild type tumors.
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. Person II-1 was diagnosed with MNG. Person II-2 had MNG, a mature cystic teratoma, nine schwannomas and an ovarian serous cystadenofibroma. Individual III-1 was diagnosed with autism spectrum disorder (ASD), MNG, had a choroid plexus papilloma (CPP) WHO stage I and multiple schwannomas. III-3 was diagnosed with MNG and multiple schwannomas. Person II-3 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 loss of heterozygosity (LOH) in individual s I-1, II-1, II-2, II-3, II-4 and II-5. II-2 and II-3’s schwannoma samples had a remnant of the wild type allele likely due 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 Supplementary Tables 3-6 and Supplementary Figures 5 and 6. Arrow indicates a wild-type genotype. Asterisk indicates presence of the mutant base. Three nodules of the MNG were sequenced for II-1, III-1, II-3 and II-3. Germline DNA with wild type sequence at the c.1552G>p.E518K locus is shown for unaffected individual III-5. No tumour DNA from the schwannomas of III-1 and from the MNG of I-1 was available.
Figure 2. Canonical vs intron derived miRNA production pathway

Schematic diagrams of canonical miRNA processing pathway (left) as well as the intronic miRNA (mirtrons) processing pathway (right). Mirtrons are processed by the spliceosome and 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.
Figure 3. Clustering of miRNA expression in DGCR8-c.1552G>A:p.E518K mutated tumors and their wild-type counterparts

A) Clustering of miRNA expression in 4 DGCR8-E518K mutated schwannomas, 5 DGCR8 wild-type schwannomas, 4 DGCR8-E518K mutated and 20 DGCR8 wild-type 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 wildtype. See also Supplementary Figures 8 and 9 (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 wild-type cases), 4 follicular variant papillary thyroid cancers (PTC) (2 DGCR8-E518K mutated and 2 wild-type) and 11 schwannomas (5 DGCR8-E518K mutated, 1 with the VUS c.1147A>G;p.S383G case indicated by *, 1 with the germline variant c.1763A>G;p.K588R previously described in Wilms tumors indicated by #, and 4 wild-type 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 wild-type tumors, suggesting the variants do not disturb the miRNA processing role of DGCR8. While DGCR8-E518K is distributed randomly among tissue types, wild-type samples clustered by tissue of origin (thyroid vs schwann cells), highlighting the particularity of the c.1552G>A:p.E518K profile in miRNA processing.
Figure 4. Differential expressed (DE) mRNA and miRNA analysis of tumors with or without DGC8 mutation and in vitro cleavage of miR-30c-2

For all three volcano plots, Log Fold change is plotted on the x-axis and the adjusted p-value (FDR; -log10 scale) on the y-axis. Dotted horizontal and vertical lines indicate threshold of significance (FDR < 0.01) and absolute fold change (>|2|). Over- and under-expressed mRNAs/miRNAs in mutated cases compared to wild-types 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 the c.1552G>A:p.E518K mutation (n = 5). B) Results of differentially expressed miRNA expression analysis between schwannomas with (n = 4) and without the c.1552G>A:p.E518K mutation (n = 5). One of the most significantly underrepresented pre-miRNAs (as shown in Supplementary Figure 12), miR-30c-2, is the precursor for the most significant under-expressed mature miRNA. All differentially expressed mirtrons are overrepresented in mutated cases compared to wild-type. The top 25 mRNAs and miRNAs up and down are listed in Supplementary Tables 9 and 10, respectively. C) In vitro cleavage results of pri-miR-30c-2. The microprocessor complex formed with DGC8-wildtype trims the primary miRNA into a precursor miRNA, but the mutant DGC8-E518K fails to cleave the primary miRNA even after 60 minutes. Reaction incubation time is shown in minutes. DGC8-wildtype & Drosha: pull down of the complex between DGC8- wild type protein and endogenous DROSHA; DGC8-E518K & DROSHA: pull down of the complex between DGC8-E518K mutant protein and endogenous DROSHA; M: RNA-labelled ladder marker; black arrow shows the band corresponding to the primary miRNA (~100bp); gray arrow points the band corresponding to a precursor miRNA (~65bp). D) Differentially expressed miRNA expression analysis between Wilms Tumors with (n = 4) and without DGC8-E518K mutation (n = 20). All differentially expressed mirtrons are overrepresented compared to wild-type tumors.