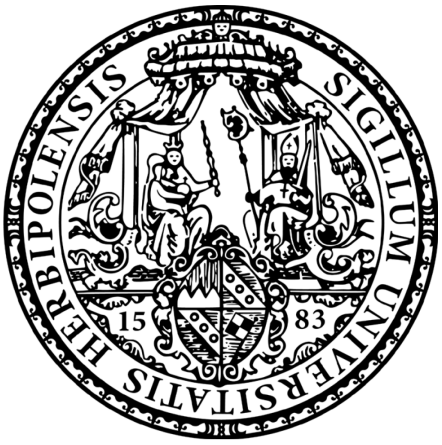


Mutations in the DROSHA/DGCR8 microprocessor complex in high-risk blastemal Wilms tumor

Mutationen des DROSHA/DGCR8 Mikroprozessor-Komplexes in blastemreichen Hochrisiko-Wilms Tumoren

Doctoral thesis for a doctoral degree
at the Graduate School of Life Sciences,
Julius-Maximilians-Universität Würzburg,
Section Biomedicine

submitted by

Romina Vardapour

from

Nürnberg

Würzburg 2020

Submitted on: ……..……..

Office stamp

Members of the *Promotionskomitee*:

Chairperson: Prof. Dr. Alexander Buchberger

Primary Supervisor: Prof. Dr. Manfred Gessler

Supervisor (Second): Prof. Dr. Eva Geissinger

Supervisor (Third): Prof. Dr. Manfred Alsheimer

Date of Public Defence: ……

Date of Receipt of Certificates: ……..

Index

1 Summary

1.1 Summary (English)

Wilms tumor (WT) or nephroblastoma is the most common kidney tumor in childhood. Several genetic alterations have been identified in WT over the past years. However, a clear-cut underlying genetic defect has remained elusive. Growing evidence suggests that miRNA processing genes play a major role in the formation of pediatric tumors, including WT.

We and others have identified the microprocessor genes *DROSHA* and *DGCR8* as key players in Wilms tumorigenesis. Exome sequence analysis of a cohort of blastemal-type WTs revealed the recurrent hotspot mutations *DROSHA* E1147K and *DGCR8* E518K mapping to regions important for catalyic activity and RNA-binding. These alterations were expected to affect processing of miRNA precursors, ultimately leading to altered miRNA expression. Indeed, mutated tumor samples were characterized by distinct miRNA patterns. Notably, these mutations have been observed almost exclusively in WT, suggesting that they play a specific role in WT formation.

The aim of the present work was to first examine the mutation frequency of *DROSHA* E1147K and *DGCR8* E518K in a larger cohort of WTs, and to further characterize these microprocessor gene mutations as potential oncogenic drivers for WT formation.

Screening of additional 700 WT samples by allele-specific PCR revealed a high frequency of *DROSHA* E1147K and *DGCR8* E518K mutations, with the highest incidence found in tumors of high-risk histology. *DROSHA* E1147K was heterozygously expressed in all cases, which strongly implies a dominant-negative effect. In contrast, *DGCR8* E518K exclusively exhibited homozygous expression, suggestive for the mutation to act recessive.

To functionally assess the mutations of the microprocessor complex *in vitro*, I generated stable HEK293T cell lines with inducible overexpression of *DROSHA* E1147K, and stable mouse embryonic stem cell (mESC) lines with inducible overexpression of *DGCR8* E518K. To mimic the homozygous expression observed in WT, DGCR8 mESC lines were generated on a DGCR8 knockout background. Inducible overexpression of wild-type or mutant *DROSHA* in HEK293T cells showed that *DROSHA* E1147K leads to a global downregulation of miRNA expression. It has previously been shown that the knockout of DGCR8 in mESCs also results in a significant downregulation of canonical miRNAs. Inducible overexpression of wild-type *DGCR8* rescued this processing defect. *DGCR8* E518K on the other hand, only led to a partial rescue. Differentially expressed miRNAs comprised members of the ESC cell cycle (ESCC) and let-7 miRNA families whose antagonism is known to play a pivotal role in the regulation of stem cell properties. Along with altered miRNA expression, DGCR8-E518K mESCs exhibited alterations in target gene expression potentially affecting various biological processes.

We could observe decreased proliferation rates, most likely due to reduced cell viability. DGCR8-E518K seemed to be able to overcome the block of G1-S transition and to rescue the cell cycle defect

in DGCR8-KO mESCs, albeit not to the full extent like DGCR8-wild-type. Moreover, DGCR8-E518K appeared to be unable to completely block epithelial-to-mesenchymal transition (EMT). Embryoid bodies (EBs) with the E518K mutation, however, were still able to silence the self-renewal program rescuing the differentiation defect in DGCR8-KO mESCs.

Taken together, I could show that *DROSHA* E1147K and *DGCR8* E518K are frequent events in WT with the highest incidence in high-risk tumor entities. Either mutation led to altered miRNA expression *in vitro* confirming our previous findings in tumor samples. While the *DROSHA* E1147K mutation resulted in a global downregulation of canonical miRNAs, *DGCR8* E518K was able to retain significant activity of the microprocessor complex, suggesting that partial reduction of activity or altered specificity may be critical in Wilms tumorigenesis.

Despite the significant differences found in the miRNA and mRNA profiles of DGCR8-E518K and DGCR8-wild-type mESCs, functional analysis showed that *DGCR8* E518K could mostly restore important cellular functions in the knockout and only slightly differed from the wild-type situation. Further studies in a rather physiological environment, such as in a WT blastemal model system, may additionally help to better assess the subtle differences between *DGCR8* E518K and *DGCR8* wild-type observed in our mESC lines. Together with our findings, these model systems may thus contribute to better understand the role of these microprocessor mutations in the formation of WT.

1.2 Zusammenfassung (Deutsch)

Der Wilms Tumor (WT), auch Nephroblastom genannt, ist der häufigste Nierentumor im Kindesalter. In den letzten Jahren wurden bereits mehrere genetische Veränderungen in Wilms Tumoren festgestellt. Bisher konnte jedoch keine eindeutige genetische Ursache gefunden werden. Immer mehr Erkenntnisse deuten darauf hin, dass miRNA Prozessierungsgene eine wichtige Rolle bei der Entstehung von pädiatrischen Tumoren, einschließlich von WT, spielen.

Uns ist es gelungen die Mikroprozessor-Gene *DROSHA* und *DGCR8* als entscheidende Faktoren in der WT-Entstehung zu identifizieren. Mit Hilfe der Exom-Sequenzierung einer Kohorte blastemreicher Wilms Tumoren konnten die wiederkehrenden Hotspot-Mutationen *DROSHA* E1147K and *DGCR8* E518K gefunden werden. Diese Mutationen betreffen Regionen, die für die katalytische Aktivität und die Bindung von RNA wichtig sind. Diese Veränderungen beeinflussen vermutlich die Prozessierung von Vorläufer-miRNAs und führen letztendlich zu einer veränderten miRNA Expression. In der Tat waren mutierte Tumorproben durch auffällige Expressionsmuster gekennzeichnet.

Bemerkenswerterweise wurden diese Mutationen fast ausschließlich in WT beobachtet, was darauf hindeutet, dass sie eine spezifische Rolle bei der Entstehung von WT spielen.

Ziel der vorliegenden Arbeit war es, zunächst die Mutationshäufigkeit von *DROSHA* E1147K und *DGCR8* E518K in einer größeren Kohorte von Wilms Tumoren zu untersuchen und anschließend diese Mikroprozessor-Genmutationen als mögliche onkogene Treiber für die WT-Entstehung näher zu charakterisieren.

Das Screening von zusätzlichen 700 WT-Proben mittels allelspezifischer PCR ergab eine hohe Häufigkeit von *DROSHA* E1147K und *DGCR8* E518K Mutationen. Dabei traten sie vermehrt bei Tumoren mit einer Hochrisikohistologie auf. Die *DROSHA* E1147K Mutation wurde in allen Fällen heterozygot exprimiert, was einen dominant-negativen Effekt impliziert. Im Gegensatz dazu zeigte die *DGCR8* E518K Mutation ausschließlich eine homozygote Expression, was darauf hindeutet, dass die Mutation rezessiv wirkt.

Um die Mutationen des Mikroprozessorkomplexes *in vitro* funktionell zu untersuchen, generierte ich stabile HEK293T-Zelllinien mit induzierbarer Überexpression von *DROSHA* E1147K, und stabile embryonale Mausstammzelllinien mit induzierbarer Überexpression von *DGCR8* E518K. Um die in WT beobachtete homozygote Expression nachzustellen, wurden für die Erzeugung der DGCR8-Zelllinien Mausstammzellen mit einem DGCR8-Knockout verwendet.

Die induzierbare Überexpression von Wildtyp oder mutiertem *DROSHA* in HEK293T-Zellen zeigte, dass *DROSHA* E1147K zu einer globalen Herunterregulierung der miRNA-Expression führt. Es wurde zuvor gezeigt, dass der Knockout von DGCR8 in Mausstammzellen ebenfalls zu einer signifikanten Herunterregulierung kanonischer miRNAs führt. Die induzierbare Überexpression von Wildtyp *DGCR8* konnte diesen Prozessierungsdefekt aufheben. *DGCR8* E518K führte dagegen nur zu einer partiellen

Behebung des Prozessierungsdefekts. Differenziell exprimierte miRNAs umfassten hierbei Mitglieder aus der stammzellspezifischen ESCC-Familie und der let-7-miRNA-Familie, deren Antagonismus bekanntermaßen eine bedeutende Rolle bei der Regulation von Stammzelleigenschaften spielt. Zusammen mit einer veränderten miRNA-Expression zeigten DGCR8-E518K-Zellen Veränderungen in der Zielgenexpression, welche möglicherweise verschiedene biologische Prozesse beeinflussen können.

Wir konnten verringerte Proliferationsraten beobachten, höchstwahrscheinlich aufgrund einer verringerten Lebensfähigkeit der Zellen. DGCR8-E518K schien in der Lage zu sein, den Block des G1-S Übergangs zu überwinden und den Zellzyklusdefekt in DGCR8-KO-Zellen zu beheben, wenn auch nicht in vollem Umfang wie Wildtyp DGCR8. Darüber hinaus schien DGCR8-E518K nicht in der Lage zu sein, die epithelial-mesenchymale Transition (EMT) vollständig blockieren zu können. Embryoid-Körper (EBs) mit der E518K Mutation konnten das Selbsterneuerungsprogramm jedoch noch unterdrücken und somit den Differenzierungsdefekt in DGCR8-KO-Zellen beheben.

Zusammenfassend konnte ich zeigen, dass *DROSHA* E1147K und *DGCR8* E518K häufige Mutationen in WT sind und dabei am häufigsten in Hochrisiko-Tumoren vorkommen. Jede dieser Mutationen führte *in vitro* zu einer veränderten miRNA-Expression, was unsere vorherigen Befunde in Tumorproben bestätigte. Während die *DROSHA* E1147K Mutation zu einer globalen Herunterregulierung kanonischer miRNAs führte, konnte die *DGCR8* E518K Mutation die Aktivität des Mikroprozessorkomplexes größtenteils wiederherstellen, was darauf hindeutet, dass eine teilweise Verringerung der Aktivität oder eine veränderte Spezifität bei der WT-Entstehung kritisch sein könnte.

Trotz der signifikanten Unterschiede in den miRNA- und mRNA-Profilen von DGCR8-E518K- und DGCR8-Wildtyp-Zellen, ergab die Funktionsanalyse, dass *DGCR8* E518K viele wichtige Zellfunktionen im Knockout wiederherstellen konnte und sich nur geringfügig von der Wildtyp-Situation unterschied. Weitere Studien unter physiologischeren Bedingungen, wie beispielsweise in einem WT-Blastem-Modellsystem, könnten zusätzlich helfen, die in unseren Mausstammzelllinien beobachteten feinen Unterschiede zwischen *DGCR8* E518K und *DGCR8* Wildtyp besser bewerten zu können. Zusammen mit unseren Erkenntnissen könnten diese Modellsysteme somit dazu beitragen, die Rolle dieser Mikroprozessormutationen bei der WT-Entstehung besser zu verstehen.

2 Introduction

2.1 Wilms tumor

Wilms tumor (WT), also referred to as nephroblastoma, is named after the German surgeon and pathologist Dr. Karl Maximilian Wilhelm Wilms (1867-1918). In 1899, Dr. Wilms published his monograph “Mischgeschwülste der Niere“ in which he gave a profound review on pediatric renal tumors. He described that renal mixed tumors are caused by developmental defects during embryogenesis. While nephroblastoma had already been described previously, he was the first to identify that all tissues present in this type of childhood renal tumor originated from cells of the middle germ layer. Comparable to the development of an embryo, all different cell types of the tumor arise from a common undifferentiated origin. Thus, for the first time, he unified morphologically diverse pediatric renal tumors and defined them as entities caused by a developmental anomaly during embryogenesis (Coppes-Zantinga and Coppes 1999).

2.1.1 Epidemiology

Wilms tumor is the most common kidney cancer in children. It affects approximately 1:10,000 children worldwide under the age of 15, with female patients being slightly more affected. The incidence of WT varies among different ethnic groups, with African Americans having the highest, and East Asian populations showing the lowest WT rates (Breslow et al. 1993). In Germany, there are 100 new cases of renal tumors diagnosed each year, of which 90% are Wilms tumors (Yiallouros 2017). Most cases are unilateral (90–95%) and sporadic (98–99%). Bilateral WTs occur less frequently (5–10%) and are diagnosed at an earlier age. The mean age of children diagnosed with WT is approximately 44 months in unilateral, and 31 months for bilateral cases. WT can also occur in association with other congenital defects, mostly genitourinary anomalies or with syndromes, such as the WAGR (Wilms tumor, aniridia, genitourinary anomalies, and mental retardation) syndrome, Denys-Drash syndrome, and somatic overgrowth disorders like the Beckwith-Wiedemann and the Simpson-Golabi-Behmel syndrome. Children with congenital anomalies and a family history of WT have a higher risk to develop bilateral tumors (Breslow et al. 1993; Huff 1998).

2.1.2 Clinical features and diagnosis

Most WT patients present with an asymptomatic abdominal mass. The tumor is usually detected by a parent or during a routine physical examination due to a protrusive abdomen. In approximately 20-30% of the cases, symptoms can include abdominal pain, malaise, and either microscopic or macroscopic hematuria. About 25% of the children have hypertension, likely due to an activated renin-angiotensin system. In some cases, children can develop an acute abdomen including symptoms, such as rapidly enlarging abdominal mass, anemia, hypertension, pain and fever (Davidoff 2012). The initial step of investigation is an abdominal ultrasound scan to determine whether the

abdominal mass is intrarenal or extrarenal. It also allows the evaluation of the contralateral kidney to exclude bilateral renal lesions and to discover possible liver metastases, which are present in <5% of the cases. Patients suspected of having a WT are further scanned via computed tomography (CT) or preferably via magnetic resonance imaging (MRI) to reduce radiation exposure. Diagnostic biopsies are not standard practice in Europe and their need is controversially discussed. However, under certain circumstances, such as older age, infection or unusual imaging finding, core needle biopsies are considered as an option (Szychot, Apps, and Pritchard-Jones 2014).

2.1.3 Histology and Staging

Wilms tumors arise from undifferentiated kidney precursor cells with significant mitotic activity. Metanephric blastema usually disappears completely by 36 weeks' gestation. Its abnormal postnatal persistence and dysregulated differentiation is called nephrogenic rest (NR) or nephroblastomatosis (NB), which occur in 40% of WT patients (vs. 0.6% in children without WT). These lesions are considered as precursors of WT. The primitive blastema contained in these NRs can potentially differentiate into different tissue types (Beckwith, Kiviat, and Bonadio 1990; Stone et al. 1990). A classic WT exhibits a triphasic histology with portions of blastemal, epithelial and stromal components (Figure 1). For histological subtyping of a WT, the proportion of the viable tumor mass has to comprise more than 1/3 of the resected tumor. If any one of the cell types comprises at least 2/3 of the viable specimen, the tumor is then accordingly subclassified as blastemal, epithelial or stromal type WT (Vujanic et al. 2002).

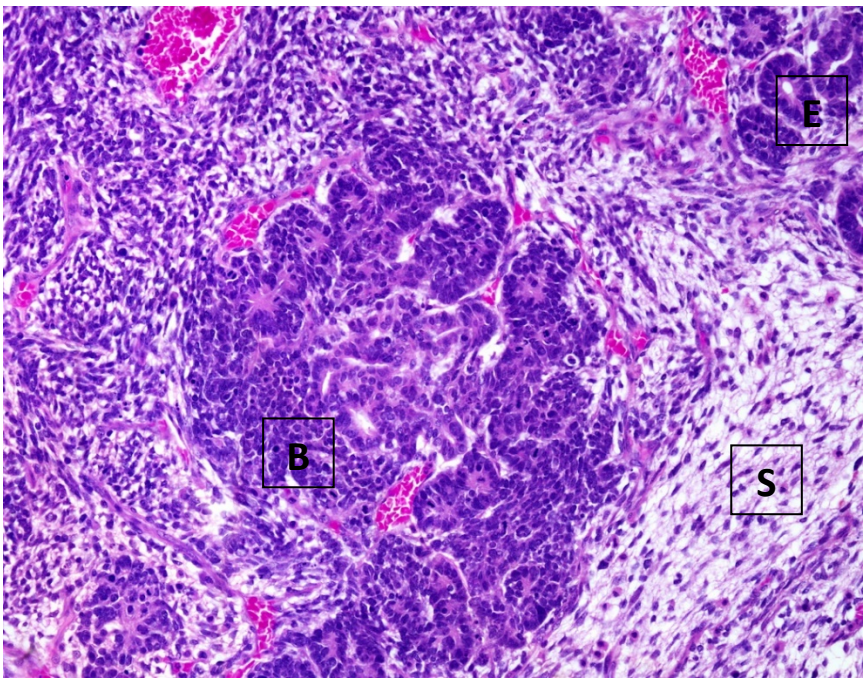


Figure 1: Histology of Wilms tumor
Example of a triphasic Wilms tumor with blastemal (B), epithelial (E) and stromal elements (S). (Image from (https://www.webpathology.com/image.asp?case=73&n=13) with permission of Dr. Dharam Ramnani, received on 03/04/2019).

Accurate histological subtyping is essential for patient prognosis and suitable treatment selection. There are three prognostic groups of WTs: Low risk, intermediate risk and high-risk. Since preoperative chemotherapy alters the histology of the original tumor, it is important to distinguish the risk assessment guidelines from those of directly resected tumors (Table 1). The main difference of the two protocols lies in the risk classification of blastemal-type WTs. Blastema per se is not of prognostic relevance in primarily resected tumors and therefore considered as intermediate risk by the COG (Children's Oncology Group). However, remaining viable blastema after preoperative chemotherapy is associated with adverse prognosis and hence defined as high-risk according to the SIOP (International Society of Pediatric Oncology) protocol.

Table 1: Risk classification based on histological subtypes of Wilms tumor (Vujanic et al. 2002)

Risk Group	Preoperative chemotherapy	Primary Surgery
Low risk	Cystic partially differentiated Completely necrotic	Cystic partially differentiated
Intermediate risk	Epithelial Stromal Mixed Regressive Focal anaplasia	Epithelial Stromal Mixed Regressive Focal anaplasia Blastemal
High-risk	Blastemal Diffuse anaplasia	Diffuse anaplasia

Another important prognostic factor is the stage of the tumor. Staging is based on tumor extension, involvement of lymph nodes, metastases and bilateralism of the tumor (Vujanic et al. 2002) (Table 2).

Table 2: Staging criteria for Wilms tumor according to SIOP (International Society of Pediatric Oncology) (Vujanic et al. 2002)

Stage	Criteria
I	Limited to kidney, renal capsule intact, complete resection, vessels of the renal sinus not involved
II	Extends beyond kidney, penetrates through renal capsule into perirenal fat, infiltrates renal sinus, invades blood and lymphatic vessels outside the renal parenchyma, infiltrates adjacent organs or vena cava, complete resection
III	Incomplete excision, abdominal lymph nodes involved, tumor rupture, penetrates through peritoneal surface
IV	Hematogenous metastases (lung, liver, bone, brain, etc.), lymph node metastases beyond abdomen
V	Bilateral renal tumors at diagnosis

2.1.4 Therapy

Treatment of WT mainly consists of a combination of chemotherapy, surgery and in some cases radiotherapy. There are two prevailing approaches to the management of WT: While in North America, WT patients are treated according to recommendations of the COG (Children's Oncology Group), children in Europe and other countries are treated according to the guidelines stated from the SIOP (Société Internationale d'Oncologie Pédiatrique) (Dome et al. 2013; Vujanic and Sandstedt 2010). The COG protocol provides for an up-front primary resection of the tumor before chemotherapy and/or radiotherapy is given, thus allowing immediate histological assessment and accurate staging of the untreated primary tumor. The SIOP protocol, on the other hand, includes preoperative chemotherapy to reduce tumor size and to minimize the risk of tumor rupture during surgery. Administration of preoperative chemotherapy also leads to tumor downstaging and hence reduces the requirement for aggressive postoperative therapy (Sonn and Shortliffe 2008; Spreafico and Bellani 2006). Postoperative treatment is individually tailored based on tumor stage and histological risk group. While low-risk stage I tumors do not require further treatment, all other stages and risk groups need to be either treated with chemotherapy alone or in combination with radiotherapy, depending on the degree of malignancy. Particularly, high-risk blastemal WTs require intensified treatment, consisting of a three or even four-drug postoperative regimen with additional radiotherapy for stage III–IV (De Kraker et al. 2001; van den Heuvel-Eibrink et al. 2017). The reason for this strict treatment regimen is due to the high aggressiveness observed in chemotherapy-resistant blastemal-type WTs. Preoperative chemotherapy alters the histology of the tumor and shifts the proportion of blastemal-type WTs from 39.4% in directly resected tumors, to only 9.3% in tumors treated with preoperative chemotherapy. Despite of the high responsiveness to chemotherapy, this remaining fraction of blastemal-type tumors is considered as a high-risk group associated with adverse prognosis (Weirich et al. 2001).

2.1.5 Prognosis

The survival rates of WT patients have drastically improved over the last decades. Despite the differences in their protocols, the COG as well as the SIOP studies contributed to a better management of WT and increased overall survival from approximately 30% in the early 20th century, to about 90% today (Metzger and Dome 2005). Still, treatment success comes at a price. More than 60% of WT survivors have to deal with long-term side effects related to chemotherapeutic agents and radiotherapy (Wright, Green, and Daw 2009). For instance, treatment with the drug doxorubicin can have adverse impact on the heart due to its known cytotoxic properties, which can lead to cardiomyopathy and congestive heart failure even 20 years after treatment (Green et al. 2001). Also, radiotherapy can be harmful to healthy cells and can affect normal tissue growth in children, which in turn can lead to decreased height, particularly if the spinal region was irradiated (Shalet et al. 1987).

More severely, radiation can lead to second malignant neoplasms (SMNs), such as soft-tissue and bone sarcoma, breast cancer, melanoma and other cancers (Termuhlen et al. 2011). These treatment-related chronic health issues again illustrate the importance of accurate risk stratification for selection of appropriate therapy modalities. In previous SIOP studies, it was shown that patients with the high-risk histological subtypes diffuse anaplasia and chemotherapy-resistant blastema had a reduced 5-year relapse-free survival compared to those with low-risk or intermediate risk tumors (38.1 and 58.4% vs. 86.7%) (Weirich et al. 2004). Therefore, diffuse anaplasia as well as resistant blastema serve as adverse prognostic indicators. A deep understanding of the genetic basis of these tumors could help to find new biomarkers that contribute to better risk-adapted therapeutic approaches.

2.1.6 Genetics

In most of the cases, Wilms tumors occur as a sporadic disease. Only 1–2% of the patients have a relative who is also affected. Familial predisposition is inherited in an autosomal-dominant manner, mostly with incomplete penetrance (Breslow et al. 1996; Matsunaga 1981). Wilms tumor is known to be genetically heterogeneous. The first WT gene being identified and cloned, was the Wilms tumor suppressor gene *WT1* at the 11p13 locus due to deletions detected in patients with WAGR syndrome (Call et al. 1990; Gessler et al. 1990). Yet, mutations in the *WT1* gene only appear in less than 20% of sporadic and non-syndromic WTs, implying other genes being involved in the development of WT (Gessler et al. 1994; Huff 1998). It was years later, when two further genes were discovered to be involved in Wilms tumorigenesis, namely *CTNNB1* (β-catenin) and *WTX*, both genes related to the Wnt-signaling pathway (Koesters et al. 1999; Rivera et al. 2007). Interestingly, tumors with mutations in *WT1* mostly carry alterations in *CTNNB1* (Maiti et al. 2000). There is no overlap, however, between tumors with *WTX* and *WT1* mutations. It was shown that *WTX* mutations occur with a similar frequency as *WT1* and *CTNNB1* mutations and that these three genes account for approximately only one third of WTs emphasizing the complexity of this tumor (Ruteshouser, Robinson, and Huff 2008). There is still a great fraction of cases where the genetic cause is yet to be revealed. Different studies have made efforts to a better understanding of the etiology of WT. Several genetic alterations have been associated with WT, such as loss of heterozygosity (LOH) of chromosomes 1p and 16q as well as gain of chromosome 1q. Both chromosomal aberrations are associated with adverse outcome (Grundy et al. 2005; Segers et al. 2013). Furthermore, epigenetic alterations at the 11p15 locus like loss of imprinting (LOI) in the *IGF2/H19* gene cluster and deletion of the maternal allele are associated with the development of WT (Scott et al. 2012). Also, mutations in the tumor suppressor gene *TP53* are implicated in WT and frequently occur in anaplastic WTs, which carry a high risk for relapse and dismal outcome (Bardeesy et al. 1994; Wegert et al. 2017). Despite of all the efforts, additional critical genes as oncogenic drivers had yet to be uncovered. We and others could identify

further recurrent mutations affecting genes of different pathways such as the *SIX1/SIX2* pathway and mutations of the miRNA processing complex *DROSHA* and *DGCR8* as promising candidate genes in high-risk WTs (Walz et al. 2015; Wegert et al. 2015).

2.2 Wilms tumor mutation screening

In order to better understand the genetic basis of high-risk blastemal WTs, we analyzed a large cohort of such tumors (n=58) by exome and transcriptome sequencing, in cooperation with the DKFZ-Heidelberg Center for Personalized Oncology (DKFZ-HIPO) (Wegert et al. 2015). We could find several genes being mutated at least once, however just a small fraction of genes showed more than two mutational events. Functionally relevant genes were then further analyzed. The two most prominent pathways were the *SIX1/SIX2* and the miRNA processing pathway. Recurrent hotspot mutations (A to G transition -> Q177R) were found in the homeodomains of *SIX1* and *SIX2,* both transcription factor genes involved in kidney development. In our study, we could show that these mutations correlate with increased proliferation of blastemal cells and hence play a critical role in blastemal-type Wilms tumors. We could further detect recurrent mutations in genes affecting the *SIX1/SIX2* pathway, including *SALL1, CITED1* and *MYCN* which are known to play a crucial role in normal kidney development. Other prevalent hotspot mutations were detected in a set of genes that are part of the miRNA processing machinery, with the most prevalent mutations found in *DROSHA* and *DGCR8* which together form the microprocessor complex (2.3). The characterization of these mutations will be covered in detail as part of this work. Further alterations were found in *DICER1* and *DIS3L2*, both important factors in the processing of miRNAs (Wegert et al. 2015). Mutations in these genes had already been implicated in WT before, such as mutations in *DICER1* affecting the RNase IIIb domain leading to defective miRNA processing, indicated by the shift of mature miRNAs towards 3p-miRNAs (Anglesio et al. 2013; Wu et al. 2013). Also, *DIS3L2* which is important for let-7 miRNA degradation is altered in patients with Perlman syndrome, a disorder that causes a predisposition to the development of WT (Astuti et al. 2012). This data underlines the significance of accurate miRNA processing and illustrates that its disruption is clearly implicated in the formation of WT. Therefore, it is important to get a deeper knowledge about these alterations in miRNA-related genes, particularly in *DROSHA* and *DGCR8*. A thorough characterization of these mutations and their consideration as oncogenic drivers may lead to a better risk stratification for WT and thus to further reduction of overall treatment burden for children with tumors that have a favorable prognosis.

2.3 The microprocessor complex DROSHA and DGCR8

The RNase III enzyme DROSHA together with its double-stranded RNA-binding (dsRBD) partner DGCR8 form a protein complex, known as the microprocessor complex (MC). The MC plays a key role

in the canonical pathway of miRNA biogenesis (Gregory et al. 2004; Landthaler, Yalcin, and Tuschl 2004).

2.3.1 MiRNA biogenesis

Mature microRNAs (miRNAs) are a group of small non-coding RNA molecules with a length of about 22 nucleotides (nt) which play integral roles in a broad range of biological processes in animals and other organisms (Bartel 2004; Ruvkun 2001). Most of mammalian miRNAs are embedded in introns of either coding or non-coding transcripts, whereas a few are encoded by exons (Rodriguez et al. 2004). Often, multiple miRNA loci tend to cluster in close proximity to each other and are co-expressed as polycistronic miRNAs. Mammalian miRNA genes are transcribed by the RNA Polymerase II into long, hairpin-structured primary miRNA transcripts (pri-miRNA) which are both capped and polyadenylated (Cai, Hagedorn, and Cullen 2004; Lee et al. 2004). Spliced and unspliced pri-miRNAs can serve as transcripts for miRNA processing (Cullen 2004; Kim and Kim 2007). In the canonical pathway of miRNA biogenesis, pri-miRNAs are processed in two sequential steps (Figure 2). In the nucleus, the pri-miRNA is recognized and processed by the microprocessor complex DROSHA/DGCR8 into about 70 nt long hairpin-structured precursor miRNAs (pre-miRNA) with a 2 nt overhang at the 3'end (Han et al. 2004; Lee et al. 2002). After being exported into the cytoplasm by Exportin-5 (XPO5) in complex with RanGTP, the pre-miRNA is then further processed into a 22 nt long miRNA-duplex by the RNase III DICER (Bohnsack, Czaplinski, and Gorlich 2004; Park et al. 2011). In mammals, DICER interacts with the transactivation-response RNA-binding protein (TRBP) and the protein kinase R-activating protein (PACT) to facilitate and fine-tune pre-miRNA cleavage (Fukunaga et al. 2012). The miRNA-duplex is subsequently loaded onto Argonaute proteins (AGO) to form an RNA-induced silencing complex (RISC). Only one miRNA strand, either the 5p- or the 3p-arm, however, is retained and incorporated into the RISC. Strand selection depends on the relative thermodynamic stability of the ends of the miRNA strands. The strand with the less stable 5'end is selected as the guide-strand, whereas the other strand, the passenger strand is discarded upon unwinding (Ha and Kim 2014; Kawamata and Tomari 2010). The mature miRNA then binds to the 3'UTR of its target mRNAs through partially complementary binding sites and negatively regulates their expression via degradation or translational repression. Complementarity is mostly restricted to the seed sequence of the mature miRNA comprising the nucleotides at the positions 2–8 at the 5'end of the miRNA (Huntzinger and Izaurralde 2011; Lewis, Burge, and Bartel 2005).

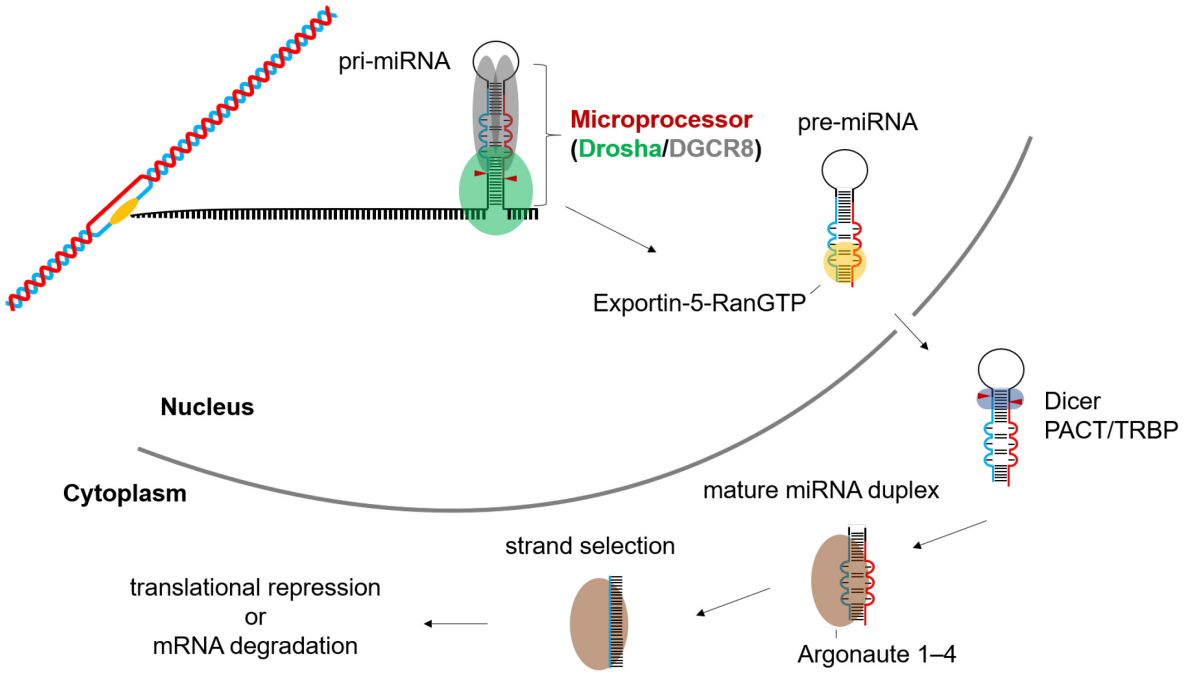


Figure 2: Biogenesis of canonical miRNAs
Pri-miRNAs are transcribed by the RNA Polymerase II and are co-transcriptionally processed by the microprocessor complex Drosha/DGCR8 in the nucleus. Cleavage leads to a 2 nt 3' overhang which is recognized by Exportin-5-RanGTP, resulting in the export of the pre-miRNA into the cytoplasm. Dicer further cleaves the pre-miRNA into a miRNA duplex. Only one miRNA strand is loaded onto Argonaute proteins and is used as a guide to target mRNAs for translational repression or degradation (Illustration adapted from (Macias, Cordiner, and Caceres 2013) with microprocessor complex being updated according to (Nguyen et al. 2015)).

2.3.2 Nuclear miRNA processing by the microprocessor complex

It was shown that the MC is a heterotrimeric complex composed of one DROSHA and two DGCR8 molecules (Figure 3) (Nguyen et al. 2015). Both DROSHA and DGCR8 are important for the specific recognition of pri-miRNAs and their subsequent processing into pre-miRNAs in the nucleus (Han et al. 2004). Pri-miRNAs are up to several kilobases long hairpin-structured RNA molecules in which the mature miRNA sequences are embedded and serve as a substrate for the MC. Typically, a pri-miRNA consists of a double-stranded stem, an apical loop and single-stranded RNA basal segments at the 5' and 3'termini. The MC interacts with several structural features of the pri-miRNA. The basal elements of the pri-miRNA are recognized by DROSHA binding to the basal UG motif and the basal junction between the single-stranded RNA segments and the double-stranded stem (ssRNA-dsRNA junction). DROSHA serves as a ruler to measure and cleave 11 bp away from the ssRNA-dsRNA junction, whereas the DGCR8 dimer interacts with the apical elements through its RNA binding heme domain (Rhed) recognizing the UGU motif (Nguyen et al. 2015). The interaction between DGCR8 and the apical UGU motif is further enhanced by hemin and ensures proper positioning of the MC on pri-miRNAs (Nguyen et al. 2018; Partin et al. 2017). In addition, DGCR8 stabilizes DROSHA through its C-terminal tail (CTT) and interacts with the stem of the pri-miRNA via double-stranded RNA-binding

domains (dsRBD) (Nguyen et al. 2015). DGCR8 therefore is indispensable for efficient and accurate processing of pri-miRNAs in the canonical miRNA pathway.

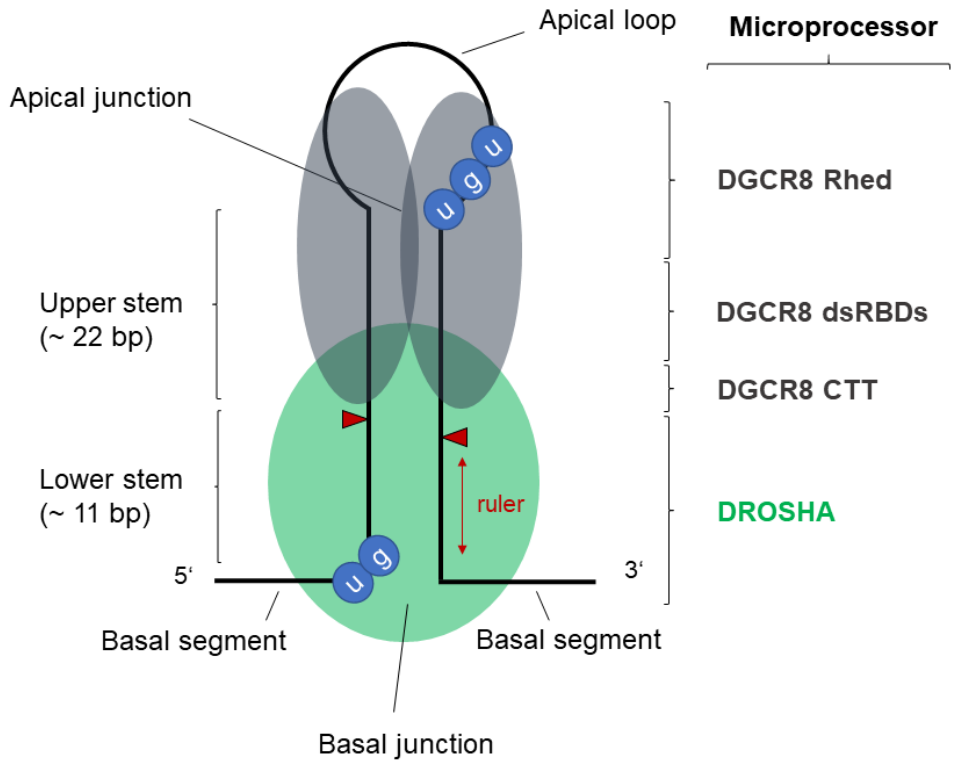


Figure 3: Functional structure of the human microprocessor complex
Heterotrimeric protein complex composed of one DROSHA and two DGCR8 molecules. DROSHA measures 11 bp away from the basal ssRNA-dsRNA junction, mediating cleavage of the pri-miRNA. DGCR8 interacts with the apical junction and the stem to ensure proper pri-miRNA processing (Illustration based on (Nguyen et al. 2015)).

2.3.3 DGCR8 (DiGeorge syndrome critical region 8)

The microprocessor subunit *DGCR8* (DiGeorge syndrome critical region 8) is located on the chromosome 22q11.2. Deletions of this region are the leading cause of DiGeorge syndrome, a defect that is associated with a wide range of conditions, including mental disorders, cardiac malformations, and immunological defects. About one third of patients with DiGeorge syndrome also suffer from congenital kidney and urinary tract anomalies (Lopez-Rivera et al. 2017). These defects are at least partially due to decreased expression levels of *DGCR8* and altered miRNAs reported in various tissues and organs of individuals with the 22q11.2 deletion syndrome, underlining the significant role of DGCR8-dependent miRNAs in the regulation of biological processes (Sellier et al. 2014; Stark et al. 2008). In addition, human *DGCR8* and its mouse ortholog are ubiquitously expressed in all tissues from fetus and adult implying its relevance for development and homeostasis in human and mouse (Shiohama et al. 2003). In line with this, mouse embryos with a *Dgcr8* knockout arrest early in their development and die at around embryonic day 6.5 (E6.5). To be still able to investigate the role of DGCR8, a conditional knockout of *Dgcr8* was generated in mouse embryonic stem cells (mESCs)

(Wang et al. 2007). For this purpose, loxP sites were inserted flanking exon 3 and subsequently treated with Cre recombinase to excise exon 3, leading to a frameshift and consequently to multiple premature stop codons downstream of the targeted region. RNA blot analysis and microarray experiments showed global loss of miRNAs confirming the essential role of DGCR8 in the processing of the majority of miRNAs (canonical miRNAs). In contrast, heterozygous deletion in mESCs did not result in significantly altered DGCR8 nor miRNA expression levels compared to the wild-type situation, suggesting a feedback control of *Dgcr8* mRNA levels by the microprocessor complex (Guo and Wang 2019; Han et al. 2009; Triboulet et al. 2009; Wang et al. 2007). The DGCR8 KO system allows to reveal phenotypical functions of specific miRNAs through rescue experiments with canonical miRNAs of interest and to subsequently investigate their impact, for instance, on self-renewal and proliferation.

In general, mESCs represent a valuable *in vitro* tool to study molecular mechanisms of early development. They are characterized with highly proliferative, yet pluripotent properties and mimic early embryogenesis (Bradley 1990; Evans and Kaufman 1981). For stem cells to be able to differentiate, fine-tuning of a variety of mRNA is required. MiRNAs are well known for their ability to control gene expression via translational repression or degradation of several targets (Valencia-Sanchez et al. 2006). Experiments with *Dgcr8* and *Dicer* knockout mESCs showed that mESCs with defective miRNA processing were unable to silence their self-renewal program and failed to undergo differentiation (Kanellopoulou et al. 2005; Wang et al. 2007). In addition to these characteristic features, stem cells share similar expression signatures with cancer cells making them a suitable tool for cancer research (Kim and Orkin 2011). Particularly, Wilms tumor, as a tumor of the developing kidney, is considered as the prototype of differentiation failure in human cancers. In this work, conditional *Dgcr8* knockout mESCs (Wang et al. 2007) were used as an *in vitro* model system to investigate the role of the hotspot mutation *DGCR8* E518K on miRNA processing and consequent effects on cellular functions such as maintenance of pluripotency and proliferation behavior.

2.3.4 MiRNA function

MiRNAs are key regulators of gene expression and play essential roles in a plethora of biological processes such as development, metabolism, apoptosis, cell proliferation and differentiation. Aberrant expression of miRNAs is implicated in various diseases, including different cancer types (Guarnieri and DiLeone 2008; Huang et al. 2011). Dysregulated miRNAs could affect one or more hallmarks of cancer and initiate tumorigenesis, for instance through sustaining proliferative signaling, resisting apoptosis and metastasis (Hanahan and Weinberg 2011; Peng and Croce 2016). Depending on their tissue-specific target gene, miRNAs can act either as oncogenes or tumor suppressors (Esquela-Kerscher and Slack 2006). The regulatory function of miRNAs is also essential for normal kidney development, homeostasis, and physiology. Their deregulation is associated with the

pathogenesis of several renal diseases, such as chronic kidney diseases, polycystic kidney disease, and renal cancers (Yu et al. 2016). Several studies have shown that alterations in miRNA expression are also major contributors to Wilms tumorigenesis (Torrezan et al. 2014; Walz et al. 2015; Wegert et al. 2015). MiRNA profiling of high-risk blastemal-type WTs vs. intermediate-risk WTs revealed groups of differentially expressed miRNAs that could serve as potential predictors for therapeutic responsiveness and could hence contribute to a better risk-stratification in the pre-treatment stage (Watson et al. 2013). A special focus should be thus placed on the microprocessor components DROSHA and DGCR8 whose proper functioning is pivotal for miRNA biogenesis. Molecular characterization of mutations affecting the miRNA processing pathway could therefore help to further define the genetic landscape of WTs.

2.4 Aim of the thesis

The recurrent hotspot mutations *DROSHA* E1147K and *DGCR8* E518K detected through our exome sequence analysis of a cohort of high-risk blastemal WTs represent potential oncogenic drivers in WT affecting important domains of the microprocessor genes. In order to confirm that these mutations are indeed frequent events in WT, a larger cohort needed to be analyzed. For mutation screening of up to 700 additional WT samples, allele-specific PCR was the method of choice. Potentially mutated samples were subsequently analyzed by traditional Sanger sequencing. In addition, cDNA of tumor samples that harbored genomic mutations of *DROSHA* or *DGCR8* was sequenced to check for either heterozygous or homozygous expression. To functionally assess the mutations of the microprocessor complex *in vitro*, I generated stable HEK293T cell lines with inducible overexpression of *DROSHA* E1147K, and stable mouse embryonic stem cell (mESC) lines with inducible overexpression of *DGCR8* E518K. To mimic the homozygous expression of the E518K mutation observed in WT, DGCR8 mESC lines were generated on a DGCR8 knockout background. The impact of these mutations on the expression of miRNAs was assessed by microarray and RNA-Seq analysis. Additionally, RNA-Seq was used to examine the effects of altered miRNA patterns on the expression of their target mRNAs. Since miRNAs are known to play essential roles in different biological processes, I investigated the proliferation behavior of the stable cell lines with different methods. Furthermore, I generated embryoid bodies (EBs) to study the impact of the *DGCR8* E518K mutation on stem cell differentiation, since stem cells without functioning miRNA processing are reported to be unable to switch off their self-renewal.

3 Material

3.1 Equipment

Table 3: Equipment used

Name	Company
Beckman Coulter FC500 Flow Cytometer	Beckman Coulter, USA
Berthold Tristar Multimode Reader	Berthold, Germany
Criterion™ Blotter	Bio-Rad, Germany
Eppendorf Centrifuge 5804 and 5424	Eppendorf, Germany
Excelitas X-Cite 120 LED Boost	Excelitas Technologies, USA
Heracell 150i CO_2 Incubator	Thermo Fisher Scientific, Germany
Image Station 4000MM PRO	Kodak, USA
Infors HT Celltron™	Infors HT, Germany
NanoDrop™ ND-1000 Spectrophotometer	NanoDrop Technologies, USA
Nikon Eclipse Ti C1 confocal microscope	Nikon, Japan
Nikon Eclipse TS100	Nikon, Japan
Nucleofector™ II/2b Device	Lonza, Switzerland
PerfectBlue™ Vertical Double Gel Systems	Peqlab, Germany

3.2 Disposables and chemicals

All disposables were either purchased from Sarstedt (Germany) or Eppendorf (Germany). Chemicals not included in Table 4 were either ordered from ROTH (Germany) or Sigma-Aldrich (Germany).

Table 4: Chemicals used

Name	Company
Aprotinin	ROTH, Germany
β-Mercaptoethanol	Sigma-Aldrich, Germany
CHIR99021	Sigma-Aldrich, Germany
DMEM (Dulbecco's Modified Eagle Medium)	Sigma-Aldrich, Germany
DNase I (1 U/µL)	Thermo Fisher Scientific, Germany
Doxycycline	Applichem, Germany
Fetal Calf Serum (FCS) Low Endotoxin A.H.	Sigma-Aldrich, Germany
GlutaMax Supplement 200 mM	Thermo Fisher Scientific, Germany
Hoechst 33342	Applichem, Germany
Leukemia Inhibitory Factor (LIF)	Self-made

Leupeptin hemisulfate	ROTH, Germany
MEM NEAA, Non-essential Amino acids 100x	Sigma-Aldrich, Germany
Mowiol	Applichem, Germany
MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide)	Sigma-Aldrich, Germany
PD0325901	Axon Medchem, Netherlands
Penicillin/Streptomycin 100x	Sigma-Aldrich, Germany
Pepstatin A	ROTH, Germany
peqGold TriFast™	Peqlab, Germany
Phenylmethylsulfonyl fluoride (PMSF)	AppliChem, Germany
Polyethylenimine (PEI)	Polysciences Europe, Germany
Propidium iodide	ROTH, Germany
Puromycin	Sigma-Aldrich, Germany
RNase A	AppliChem, Germany
Rotiphorese Gel 30	ROTH, Germany
SYBR® Green	Sigma-Aldrich, Germany
Trypsin-EDTA solution (0.25 %)	Sigma-Aldrich, Germany

3.3 Buffers

Table 5: Buffers used

Buffer	Components
Standard buffers	
PBS	150 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, 1.8 mM KH_2PO_4
TE	10 mM Tris pH 8.0, 1 mM EDTA
DNA buffers	
Base buffer 50x	1.25 M NaOH, 10 mM EDTA
DNA loading dye 10x	50% glycerol, 15% ficoll, 10 mM EDTA pH 8.0
Neutral buffer 50x	2 M Tris-HCl, pH 5.0
PCR buffer 10x	200 mM Tris pH 8.8, 100 mM KCl, 100 mM $(NH_4)_2SO_4$, 20 mM $MgSO_4$, 1% TritonX-100; Sterile-filtered; 1% acetylated BSA
SB 20x	200 mM NaOH pH 8.0 with boric acid
TAE 50x	50 mM EDTA, 2 M Tris acetate pH 8.0
Buffers for PI staining	
RNase A buffer	10 mM Tris pH 7.4, 150 mM NaCl
Sodium citrate buffer	38 mM $C_6H_5Na_3O_7 \cdot 2\ H_2O$; Sterile-filtered

Buffers for Western blot	
Blotting buffer	25 mM Tris pH 8.3, 150 mM glycine; *Add fresh*: 10% methanol
ECL substrate	100 mM Tris pH 8.0; *Add fresh*: 250 mM luminol, 90 mM coumaric acid, 0.01% H_2O_2
Protein loading buffer 4x	200 mM Tris pH 6.8, 8% SDS, 0.50% bromphenol blue, 50% glycerine, 400 mM DTT
RIPA buffer	50 mM Tris pH 8.0, 1% Nonidet P40, 0.5% Sodium deoxycholate, 0.1% SDS, 150 mM NaCl, 1 mM EDTA; *Protease inhibitors need to be added fresh to the lysis buffer:* 50 µg/mL PMSF, 1 µg/mL Aprotinin, Leupeptin hemisulfate and Pepstatin mix
SDS running buffer	25 mM Tris pH 8.3, 192 mM glycine, 1% SDS

3.4 Kits

Table 6: Kits used

Name	**Company**
Allprep DNA/RNA Mini Kit	Qiagen, Germany
Cell Proliferation ELISA, BrdU (colorimetric)	Roche, Germany
Gel Extraction Kit	Omega Bio-Tek, USA
Mouse ES Cell Nucleofector™ Kit	Lonza, Switzerland
Plasmid Midi Kit	Omega Bio-Tek, USA
Plasmid Mini Kit	Omega Bio-Tek, USA
RevertAid RT Reverse Transcription Kit	Thermo Fisher Scientific, Germany

3.5 Antibodies

Table 7: Antibodies used

Name	**Species**	**ID**	**Company**	**Dilution**
Primary Antibodies				
α-Drosha	rabbit	ab12286	Abcam, Germany	1:1000
α-Drosha [EPR12794]	rabbit	ab183732	Abcam, Germany	1:4000
α-Flag-M2	mouse	F3165	Sigma-Aldrich, Germany	1:2000
α-GFP (AA246)	goat	ABIN100085	Antibodies-online, USA	1:3000
α-Tubulin	mouse	T6074	Sigma-Aldrich, Germany	1:4000
Secondary Antibodies				
α-goat-POD	rabbit	A5420	Sigma-Aldrich, Germany	1:5000
α-mouse-POD	goat	AP124P	Sigma-Aldrich, Germany	1:5000
α-rabbit-POD	goat	7074P2	Cell Signaling Technology, Germany	1:2000

3.6 Oligonucleotides

All oligonucleotides listed in Tables 8–13 were purchased from Sigma-Aldrich (Germany). Primer sequences are annotated from 5' to 3'.

Table 8: Primer sets used for allele-specific PCR of genomic DNA of wild-type (wt) or mutant (mut) alleles

Gene	Primer sets	Primer sequence	$T_{annealing}$
DGCR8	hPasha-ASP-G	AATCCGAGGTCTGCATCCTGCATG (wt)	69°C
	hPasha-ASP-A	GAAATCCGAGGTCTGCATCCTGCATA (mut)	
	hPasha-ASP-rev	CTCCCCAGCCCTGACCAAAGTTACA	
DROSHA	hDrosha-1147-ASP-G	TTTTCCAGAGGCCACAATCAGAGAATTG (wt)	65°C
	hDrosha-1147-ASP-A	TCTTTTTCCAGAGGCCACAATCAGAGAATTA (mut)	
	hDrosha-1147-ASP-rev	ATGTGCTTTGTATACAATTTGCACAATGAAATGA	

Table 9: Primer sets for sequence verification of mutations in genomic DNA (sequencing primers are indicated in bold)

Gene	Primer sets	Primer sequence	$T_{annealing}$
DGCR8	hPasha-ex7-for	GGGAAATCCGAGGTCTGCAT	60°C
	hPasha-in7-rev	**CTTCCCTTTCCTCCCGTTCC**	
DROSHA	hDrosha-1147-Eco-test	TCGAGGGGCCTTAGGGAATTGTAT	65°C
	hDrosha-1147-ASP-rev	**ATGTGCTTTGTATACAATTTGCACAATGAAATGA**	

Table 10: Primer sets for sequence verification of mutations in cDNA (sequencing primers are indicated in bold)

Gene	Primer sets	Primer sequence	$T_{annealing}$
DGCR8	hPasha-ex7-for	GGGAAATCCGAGGTCTGCAT	60°C
	hPasha-ex8-rev	**GCTGCTTGCAGTTCCAGATC**	
DROSHA	hDrosha-for2	ACTGGCATCCGTTCTGATGT	60°C
	hDrosha-rev1	GAAGCTGGGATTTGGGGTCA	
	hDrosha-rev2	**TTTCAATCGTGGAAAGAAGCAGAC**	

Table 11: Primer sets for qPCR

Mouse specific primers, unless otherwise indicated.

Gene	Primer sets	Primer sequence
AFP	mAFP_ex11	CTCAGCGAGGAGAAATGGTC
	mAFP_ex12	GGTGATGCATAGCCTCCTGT
BRACHYURY	mBrachyuryT-real5'	CAGCCCACCTACTGGCTCT
	mBrachyuryT-real3'	GCGTCAGTGGTGTGTAATGTG

DGCR8 (Flag-tagged)	Flag_for	TGGACTACAAAGACGATGACG
	hDGCR8_rev	GGAGAGGGGCTCTCATCTG
DROSHA (human)	hDrosha-seq1	TGAGTTTGAAGAAGCAATTGGAGT
	hDrosha-1147-rev	TGGAGTCACCTAGGAATTTCATTCTCTGA
DROSHA (mouse/human)	hmDrosha-f1	TAGGAGCTGTTTACTTGGAGGG
	hmDrosha-r1	TGGAGTGGGTGGAGAGGATAAT
DROSHA (Myc-tagged)	myc-tag_for	GCTTATGGAACAAAAACTTATTTCTGA
	hDrosha_real-rev	CTCCATGTCCTCCTCGTCCT
FGF5	mFGF5-real-f	CGAGGAGTTTTCAGCAACAA
	mFGF5-real-r	CGCGGACGCATAGGTATTAT
FLT1	FLT1neu_5'real	CTTCACCTGGACTGACAGCA
	FLT1neu_3'real	ACAGCCCCGACTCCTTACTT
GATA4	mGATA4-real5'	TCAAACCAGAAAACGGAAGC
	mGATA4-real3'	CTGCTGTGCCCATAGTGAGA
GATA6	mGATA6-real5'	GCCAACTGTCACACCACAAC
	mGATA6-real3'	TGTTACCGGAGCAAGCTTTT
GFP	GFP-real-for	ACGTAAACGGCCACAAGTTC
	GFP-real-rev	AAGTCGTGCTGCTTCATGTG
HPRT	mHPRT-real-ex8	TGTTGTTGGATATGCCCTTG
	mHPRT-real-ex9	ACTGGCAACATCAACAGGACT
HPRT (human)	hu-HPRT1_neu	TGTTGTAGGATATGCCCTTGACT
	hu-HPRT2_neu	CTAAGCAGATGGCCACAGAAC
NANOG	mNanog-real5'	TTGCCTAGTTCTGAGGAAGCA
	mNanog-real3'	GAGGAAGGGCGAGGAGAG
NESTIN	mNestin _ex3	ACTCTCGCTTGCAGACACCT
	mNestin_ex4	ATTAGGCAAGGGGGAAGAGA
OCT4	mOct4-real5'	CCGTGAAGTTGGAGAAGGTG
	mOct4-real3'	GAAGCGACAGATGGTGGTCT
SOX1	mSox1-real-for	AGTGGAAGGTCATGTCCGAG
	mSox1-real-rev	TGTAATCCGGGTGTTCCTTC
SOX2	mSox2_fwd	AAGGGTTCTTGCTGGGTTTT
	mSox2_rev	AGACCACGAAAACGGTCTTG
TBX20	mTbx20-real-for	CAGCAGTCACAGCCTACCAG
	mTbx20-real-rev	GAATCGGTGTCGCTATGGAT

Table 12: Primers used for qPCR of selected mature miRNAs
a) Primers used for hemi-nested qPCR

miRNA	RT oligonucleotide	Primer sets	Primer sequence
miR-130b	CACCGTTCCCCGCCGTCGGTGATGCCC	mmu-miR-130b-3p_for	TCCGCCCAGTGCAATGAT
		mmu-miR-130b-3p_rev	CCGTCGGTGATGCCCTTT
miR-200b	CACCGTTCCCCGCCGTCGGTGTCATCA	mmu-miR-200b-3p_for	CCCGCCTAATACTGCCTG
		mmu-miR-200b-3p_rev	TCGTCGGTGTCATCATTACC
miR-24	CACCGTTCCCCGCCGTCGGTGCTGTTC	miR-24_for	CCCGCCTGGCTCAGTTC
		miR-24_rev	CCGTCGGTGCTGTTCCTG
miR-302d	CACCGTTCCCCGCCGTCGGTGACACTC	mmu-miR-302d-3p_for	CCGCCCTAAGTGCTTCCA
		mmu-miR-302d-3p_rev	CCGTCGGTGACACTCAAAC
miR-320	CACCGTTCCCCGCCGTCGGTGTCGCCC	miR-320_for	CGCCGAAAAGCTGGGTTGAG
		miR-320_rev	CCGTCGGTGTCGCCCTCT
miR-425	GACCGTTCCCCGCCGTCGGTCTCAACG	miR-425_for	GGGCGAATGACACGATCAC
		miR-425_rev	CGTCGGTCTCAACGGGAG
miR-92a	CACCGTTCCCCGCCGTCGGTGCAGGCC	miR-92a_for	CCCGCCTATTGCACTTGTC
		miR-92a_rev	GTCGGTGCAGGCCGGG

b) Primers used for two-tailed qPCR

miRNA	RT oligonucleotide	Primer sets	Primer sequence
let-7a	AACCTACTCTATGCTCTCCAGGTACAG	mmu-let-7a-for	CGAACCTACTCTATGCTCTCCAG
	TTGGTACCTGTCTCCACTTAACTA	mmu-let-7a-rev	CGGGTGAGGTAGTAGGTTGTA
miR-196a	ACTACCTACAACGACCAGAGCTAGAG	mmu-miR-196a-for	CGACTACCTACAACGACCAGAG
	AACCTAGCTCACCCACTACCCCAA	mmu-miR-196a-rev	GGGCTAGGTAGTTTCATGTTGTT

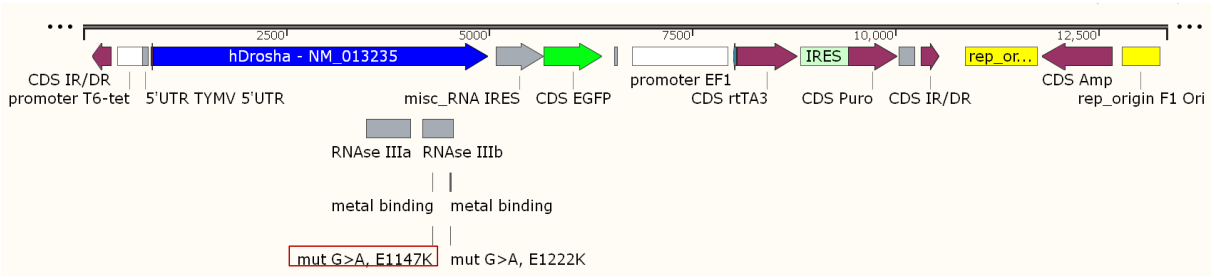
Table 13: Primer sets used for qPCR of pri-miRNAs

Pri-miRNA	Primer sets	Primer sequence
pri-miRNA-24	pri-miR-24-real-for	TCTCATTTCACACACTGGCTCA
	pri-miR-24-real-rev	GCCCTTCATCTTCTTCTCCGT
pri-miRNA-92a	pri-miR-92a-real-for	ACAGGTTGGGATTTGTCGCA
	pri-miR-92a-real-rev	GTGGAAATCGGCATCTTCAGC

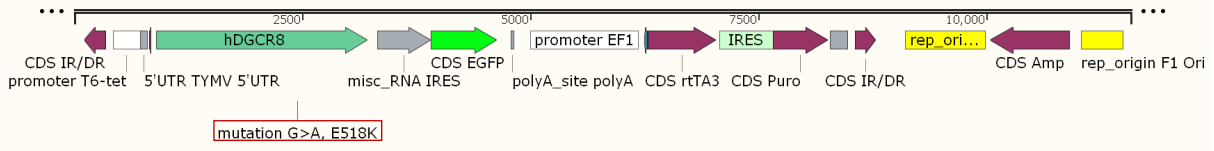
3.7 Plasmids

To generate the stable cell lines HEK293T DROSHA-wild-type/-E1147K, and DGCR8-wild-type/-E518K mESCs, the *Sleeping Beauty* transposon system was used, composed of a *Sleeping Beauty (SB)* transposase and a transposon (Ivics et al. 1997). Myc-tagged *DROSHA* and Flag-tagged *DGCR8* were expressed under the control of a tetracycline-dependent T6-tet promoter, making use of the Tet-on system (Gossen et al. 1995; Loew et al. 2010). In the presence of doxycycline, the rtTA (reverse tetracycline-controlled transactivator) activates the T6-tet promoter, which in turn drives the expression of downstream *DROSHA/DGCR8*. In both cases, induction of the respective transgene also led to EGFP expression which enabled visualization of positive cell clones by UV excitation. In addition, constitutive expression of the puromycin-resistance gene allowed for positive selection of successfully transfected cell clones. The doxycycline-inducible *(SB)*- constructs and the transposase expressing vector used in this work are shown in Figure 4 and were generated in our working group by Anja Winkler.

A) pSB-ET-iE-myc DROSHA (13,345 bp)



B) pSB-ET-iE-flag DGCR8 (11,586 bp)



C) pCMV(CAT)T7-SB100X (4756 bp)

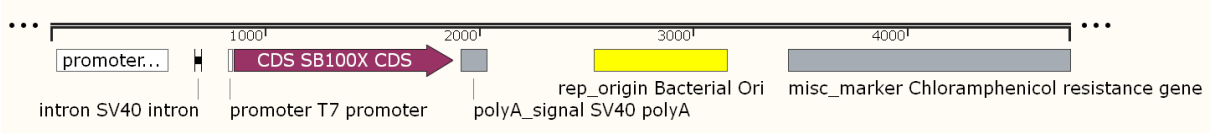
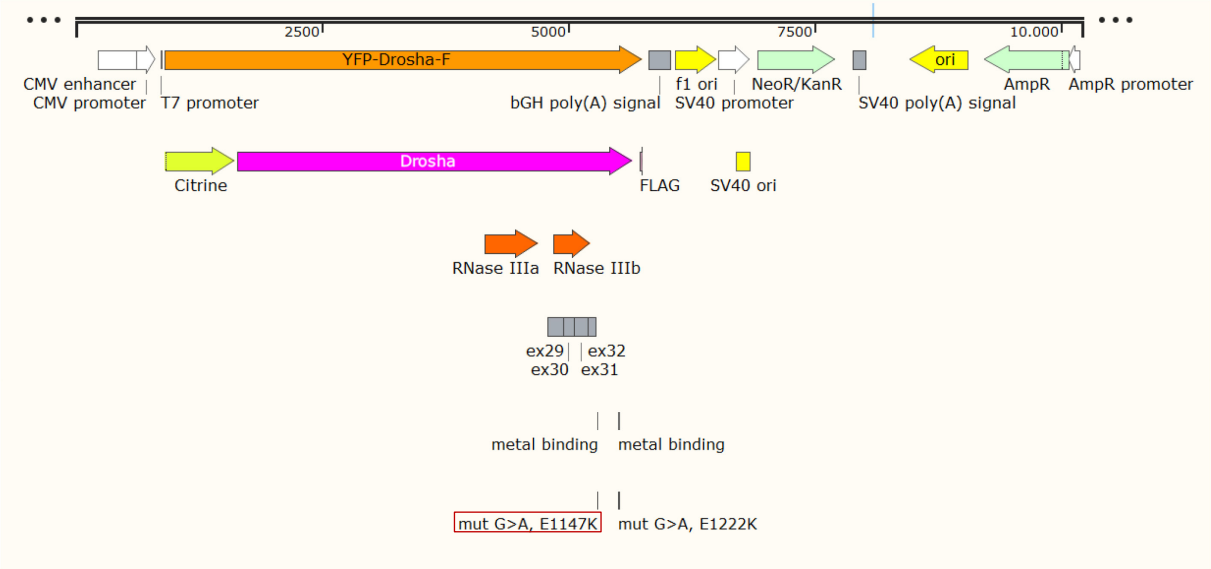


Figure 4: Doxycycline inducible *Sleeping Beauty (SB)* constructs for DROSHA (A), DGCR8 (B), and the transposase vector (C)
The *DROSHA* E1147K and *DGCR8* E518K mutation are marked by a red box in the respective vector construct.

For transient transfection of DROSHA and DGCR8 into HEK293T cells, pcDNA3.1-YFP-Drosha-Flag and pCS2P-HA2-mCherry-hDGCR8 constructs were used, respectively. Each plasmid harbored either the wild-type or the mutant version of DROSHA or DGCR8 which were constitutively expressed under the control of a CMV promoter (Figure 5).

A) pcDNA3.1-YFP-Drosha-Flag (10,220 bp)



B) pCS2P-HA2-mCherry-hDGCR8 (7245 bp)

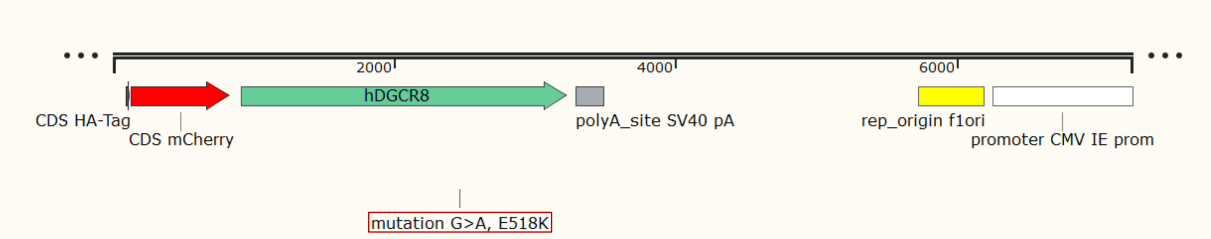


Figure 5: Plasmids used for transient transfection of DROSHA and DGCR8 in HEK293T cells
The *DROSHA* E1147K and *DGCR8* E518K mutation are marked by a red box in the respective vector construct.

3.8 Cell lines

DGCR8 knockout mouse embryonic stem cells; NBA1-19349 (Novus Biochemicals)

HEK293T cells (Human embryonic kidney 293 cells with constitutive expression of SV40 large T antigen and neomycin resistance gene)

Puromycin resistant SNL (PSNL) feeder cells (Dr. Cornelia Wiese)

3.9 Wilms tumor samples

Tumor and control (blood or normal kidney) samples were acquired from the German SIOP93-01/GPOH and SIOP2001/GPOH studies. Clinical data were available for all patients through the Wilms tumor study data center. All subjects (or their parents) provided formal consent for tumor banking and research use in compliance with national guidelines (Ethikkommission der Ärztekammer des Saarlandes; Germany 136/01). DNA and RNA were isolated from snap frozen tumor samples using standard procedures (e.g. Allprep DNA/RNA mini kit, Qiagen).

3.10 Software

Table 14: Software used

Name	Version	Source
Adobe Illustrator CC	23.0	Adobe, USA
CXP analysis software	2.2	Beckman Coulter, USA
Fiji ImageJ	1.52n	Open Source, NIH, USA
GraphPad Prism	7.0	GraphPAD Software, USA
Mastercycler™ ep Realplex	2.2	Eppendorf, Germany
Nikon NIS-Elements AR	3.2	Nikon, Japan
SnapGene software	1.1.3	GSL Biotech, USA

4 Methods

4.1 Cell culture

HEK293T cell lines were cultured at 37°C and 5% CO_2 in DMEM medium (high glucose) containing 10% FCS and penicillin/streptomycin (100 U/mL/100 µg/mL).

Mouse embryonic stem cell (mESC) lines were cultured at 37°C and 5% CO_2 in DMEM medium (high glucose) supplemented with 10% heat inactivated FCS, penicillin/streptomycin (100 U/mL/100 µg/mL), 1% MEM non-essential amino acids, 2 mM glutamine, 100 µM β-mercaptoethanol and leukemia inhibitory factor LIF (1000 U/mL).

In order to maintain pluripotency, mESCs were grown either on mouse embryonic fibroblast (MEF) feeder cells, here on PSNL feeder cells, or under feeder-free conditions in 2i-medium on porcine skin gelatin (0.1% in PBS) coated cell culture dishes. For the preparation of 2i-medium, 3 µM of the glycogen synthase kinase 3 (GSK3) inhibitor CHIR99021, and 1 µM of the mitogen-activated protein kinase (MEK) inhibitor PD0325901 were added to the medium (Ying et al. 2008). Since all experiments conducted in this work required a highly pure mESC population, it was necessary to minimize feeder cell contamination. To wean off mESCs from feeder cells, dissociated cells were plated for 30–45 min at 37°C and 5% CO_2. Supernatant was then reseeded to a new plate. This process was repeated over 3–4 passages until mESCs were depleted from feeder cells. For all experiments in this work, completely feeder-free mESCs that had been cultured in 2i-medium were used, with the exception of mESCs harvested for microarray analysis. Here, the supernatant of dissociated mESCs and feeder cells that had been reseeded on a new plate for 45 min was used for RNA extraction.

PSNL feeder cells were cultured at 37°C and 5% CO_2 in DMEM medium (high glucose) supplemented with 10% heat inactivated FCS, penicillin/streptomycin (100 U/mL/100 µg/mL), 1% non-essential amino acids on porcine skin gelatin (0.1% in PBS) coated cell culture dishes.

Before being used as feeder cells for mESCs, PSNL cells were mitotically inactivated with 10 µg/mL mitomycin C for at least 2h and washed three times with PBS.

Medium was changed daily and cultures were split every other day to maintain the cells in a subconfluent state. DGCR8-wild-type mESCs were split in a 1:4 ratio, whereas DGCR8-knockout and -E518K mESCs were split in a 1:2 or 1:3 ratio. For splitting, cells were washed with PBS and gently treated with 0.25% trypsin-EDTA for approximately 1 min. To inhibit trypsin activity, cells were washed with medium and collected in a 15 mL centrifugation tube. After centrifugation at 900 rpm for 3 min, the cell pellet was resuspended in fresh medium and seeded in the desired ratio on either feeder cell coated or gelatinized cell culture dishes.

4.2 Transfection

To generate stable HEK293T DROSHA cell lines, $1.0x10^5$ HEK293T cells per well were seeded on a 12-well-dish and transfected with 2 µg plasmid DNA. For transfection, DNA and polyethylenimine (PEI) (1 mg/mL) were used in a 1:2 ratio. In this case, 1 µg of pSB-ET-iE-mycDROSHA wild-type or E1147K mutant was co-transfected with 1 µg of the transposase vector pCMV(CAT)T7-SB100X. DNA and PEI were diluted separately in serum-free DMEM in a total volume of 50 µL and mixed by vortexing. After 5 min incubation at room temperature, PEI solution was added to the diluted DNA and incubated for another 15 min. The DNA/PEI mixture was then added dropwise to each well. The medium was changed after 5 to 8 h.

For transient transfection with pcDNA3.1-YFP-Drosha-Flag and pCS2P-HA2-mCherry-hDGCR8 and subsequent microscopic analysis, HEK293T cells were grown on coverslips in 24-well plates. 1 µg of plasmid DNA and PEI (1 mg/mL) were used in a 1:2 ratio.

For transfection of DGCR8 knockout mESCs, the Nucleofector™ Kit for Mouse Embryonic Stem Cells and the Nucleofector™ II/2b device were used. Before being used for transfection, plasmid DNA of pSB-ET-iE-FlagDGCR8 was co-precipitated with the transposase vector pCMV(CAT)-SB100X in a ratio of 1:1 using standard ethanol precipitation (1/10 volume of 3M NaAc (pH 5.2) and 2x volumes of 98% EtOH). $2.0x10^6$ mESCs were counted and centrifuged at 900 rpm for 3 min. The pellet was washed once with PBS and resuspended in 90 µL of Nucleofector solution. Cell suspension was mixed with 10 µL of vector DNA (20 µg total DNA) by pipetting and then transferred into a Lonza certified cuvette. Cells were electroporated in the Nucleofector™ II/2b device (program: A-013). After electroporation, cells were resuspended in 0.5 mL of pre-warmed medium and seeded on PSNL feeder coated 10 cm dishes using Lonza certified plastic pipettes. Medium was changed on the next day. Transfected mESCs were cultured on PSNL feeder over 2–3 passages to prevent differentiation.

Selection with puromycin (1 µg/mL) was started 48h post-transfection. Resistant clones were picked approximately after one week of selection. Non-transfected cells were used as negative controls.

Prior to each experiment, HEK293T DROSHA-wild-type/-E1147K cells, and DGCR8-wild-type/-E518K mESCs were repsectively treated with 1000 ng/mL or 300 ng/mL doxycycline for 48h to induce expression of the according transgene. HEK293T cells and DGCR8 knockout mESCs treated with doxycycline served as negative controls.

4.3 Fixation of HEK293T cells and nuclear staining

Transiently transfected HEK293T cells were fixed in 4% paraformaldehyde (PFA) for 20 min at room temperature. After three washing steps with PBS for 5 min, cell nuclei were stained with Hoechst (1:10,000 in PBS) for 10 min, protected from light. Coverslips were washed twice with PBS and

subsequently embedded in Mowiol on microscope slides. Images were analyzed using confocal laser scanning microscopy.

4.4 Embryoid body differentiation

For differentiation into embryoid bodies (EBs), $6.5x10^5$ mESCs were seeded on 6 cm bacterial petri dishes. Number of seeded cells was increased for DGCR8 knockout ($9.0x10^5$) and E518K ($7.5x10^5$) mESCs to compensate for their proliferation deficiency. EBs were cultured in the absence of LIF and 2i at 37°C and 5% CO_2 under gentle agitation (35 rpm) on a Infors HT Celltron™ shaker. To maintain transgenic expression of *DGCR8* wild-type and E518K, 300 ng/mL doxycycline was added to the medium. Medium was changed every other day. Therefore, EBs were collected and sedimented in a 50 mL centrifugation tube. After carefully aspirating the old culture medium, EBs were gently resuspended in fresh medium and returned onto the original petri dish. On day 3, 6 and 9, EBs were harvested for RNA extraction. In addition, on day 6 of EB cultivation, 100 µL of EB suspension were plated onto gelatinized 12-well cell culture dishes in ESC-medium without LIF nor 2i. After 6 days, adherent EBs were harvested for RNA extraction.

4.5 DNA extraction from cells

The Base/Neutral technique based on alkaline lysis was used for DNA extraction. To extract DNA, the cell pellet was washed with PBS and lysed in 50 µL 1xBase buffer for 30 min at 95°C. The sample was then cooled down on ice and 50 µL of Neutral buffer were added. The volume of added buffer may vary depending on the size of the cell pellet, whereas the ratio of Base to Neutral buffer always remained 1:1.

4.6 RNA extraction from cells

For RNA extraction, cell pellets were resuspended depending on their size, either in 0.5 mL or 1 mL peqGold TriFast™ and incubated for 10 min at room temperature. After lysis, RNA was isolated according to the manufacturer's recommendations and diluted in 20 µL DEPC-treated ddH_2O. The concentration of the isolated RNA was determined by measuring the absorbance using a NanoDrop™ spectrophotometer. RNA samples with an A_{260}:A_{280} ratio of approximately 2.0, and an A_{260}:A_{230} ratio in the range of 2.0–2.2 were considered as pure RNA. Residual DNA was eliminated by DNase I (1 U/µL) (Thermo Fisher Scientific) digestion for 30 min at 37°C. All steps were carried out according to the manufacturer's protocol.

4.7 Quantitative real-time PCR

1 µg of total RNA was reverse transcribed using the Thermo Scientific™ RevertAid Reverse Transcriptase (RT) kit, with oligo$(dT)_{18}$ primers or random hexamer primers according to the manufacturer's protocol. Samples were diluted to a final volume of 500 µL with ddH_2O. For

quantitative real-time PCR, 5 µL of diluted cDNA were added to the reaction mixture containing SYBR® Green (Table 15). Annealing/elongation temperature was set on 60°C (Table 16). Melting curve analysis and gel-electrophoresis were used to assess specificity of the PCR products. All reactions were performed in technical duplicates of biological replicates and expression levels were normalized to the housekeeping gene HPRT. PCR reactions were performed and analyzed with the Mastercycler™ ep Realplex 2.2. Primer sequences are listed in Tables 11–13 (3.6).

Table 15: Single qPCR reaction (25 µL)

2.5 µL	10x PCR buffer
0.25 µL	dNTPs (100 mM)
1.50 µL	Ethylenglycol
0.75 µL	SYBR Green 1:2000 in ddH_2O; 0.45 % DMSO
0.75 µL	5’ Primer (10 pmol)
0.75 µL	3’ Primer (10 pmol)
0.25 µL	His-Taq polymerase (15 U/µL)
13.25 µL	ddH_2O
5.0 µL	cDNA template

Table 16: Cycling program

95°C	2 min	Initial denaturation	
95°C	10 sec	Denaturation	40 cycles
60°C	30 sec	Annealing/Elongation	
95°C	15 sec	Denaturation	
60-95°C	10 min	Melting curve (+ 3.5°C/min)	
16°C	∞	End	

4.8 Standard PCR and allele-specific PCR (ASP)

To analyze DNA of transfected cells for correct integration of the transgene into the genome, standard PCR was performed (Table 17 and 18). Allele-specific PCR was used to screen large cohorts of Wilms tumor samples for the point mutations *DROSHA* E1147K and *DGCR8* E518K. In general, two separate PCR reactions were performed per each DNA sample, one to detect the wild-type allele (wt) and the other to identify the mutant allele (mut). Primers harbored one mismatch at the 3’ end. It was therefore pivotal to find stringent PCR conditions for each primer set to ensure only DNA

sequences to be amplified that perfectly match to the primers. Primer sequences and PCR conditions are listed in Tables 8–10 (3.6). Following amplification, PCR products were analyzed via gel electrophoresis. Samples with known alterations were used as positive controls. Potentially mutated samples were analyzed by Sanger sequencing.

Table 17: Standard PCR reaction (20 µL)

2.0 µL	10x PCR buffer
0.2 µL	dNTPs (100 mM)
0.5 µL	5’ Primer (10 pmol)
0.5 µL	3’ Primer (10 pmol)
0.2 µL	His-Taq polymerase (15 U/µL)
15.6 µL	ddH_2O
1.0 µL	DNA template (50 ng/µL)

Table 18: Standard PCR program

95°C	3 min	Initial denaturation	
95°C	30 sec	Denaturation	35 cycles
60°C	30 sec	Annealing *	
72°C	30 sec	Elongation *	
72°C	5 min	Final Elongation	
16°C	∞	End	

* Annealing temperature can vary depending on the length and the base composition of the primers. Gradient PCRs with different annealing temperatures can help to determine the optimal temperature to increase binding specificity and efficiency of the primers; Elongation time depends on the size of the PCR product and can be accordingly adjusted.

4.9 Hemi-nested real-time PCR

To quantify and validate expression of selected miRNAs, hemi-nested real-time PCR was performed as described before (Wan, Lim, and Too 2010). For multiplex reverse transcription, 100 ng of DNase I-treated total RNA was reverse transcribed using the Thermo Scientific™ RevertAid Reverse Transcriptase (RT) kit and 100 nM of each RT-oligonucleotide in a volume of 20 µL. RT was carried out at 42°C for 30 min. Following cDNA synthesis, samples were diluted 1:10 in a volume of 200 µL. 5 µL of the diluted cDNA were used in the subsequent qPCR reaction under the same experimental conditions listed in Tables 15–16 (4.7). All miRNA expression levels were normalized to miRNA-320, since it was previously shown that its synthesis is independent of DGCR8 (Castellano and Stebbing

2013). Primer sequences for both RT and qPCR are listed in Table 12a (3.6) and were designed according to the authors' guidelines (Wan, Lim, and Too 2010). Primer3web 4.1.0 was used for primer design (http://primer3.ut.ee/). Secondary structures of the RT-oligonucleotides were predicted using the mfold web server (http://unafold.rna.albany.edu/?q=mfold) (Zuker 2003).

4.10 Two-tailed real-time PCR

Two-tailed real-time PCR (Androvic et al. 2017) was alternatively used to quantify miRNA expression. Primers were designed according to the authors' guidelines and are listed in Table 12b (3.6).

10 ng of DNase I-treated total RNA was reverse transcribed using the Thermo Scientific™ RevertAid Reverse Transcriptase (RT) kit and 1 µM of two-tailed RT primer mix in a volume of 20 µL (Table 19). The reaction was incubated for 45 min at 25°C and inactivated for 5 min at 70°C. Subsequently, cDNA was 10x diluted with 180 µL ddH_2O. Quantitative real-time PCR was performed with 5 µL of diluted cDNA template and 20 µL reaction mix with an annealing/elongation temperature of 60°C (cycling program shown in Table 20).

Table 19: cDNA synthesis reaction (20 µL)

x µL	10 ng total RNA
1.0 µL	Two-tailed RT primer mix (1 µM)
ad 12 µL with DEPC-ddH_2O	
4.0 µL	5x RT buffer
0.5 µL	Ribolock RNase Inhibitor (40 U/µL)
2.0 µL	dNTPs (10 mM)
1.0 µL	Reverse Transcriptase (200 U/µL)
0.5 µL	DEPC-ddH_2O

Table 20: Cycling program for two-tailed qPCR

95°C	30 sec	Initial denaturation	
95°C	5 sec	Denaturation	⎫
60°C	15 sec	Annealing/Elongation	⎭ 45 cycles
95°C	5 sec	Denaturation	
60-95°C	10 min	Melting Curve (+3.5°C/min)	
16°C	∞	End	

4.11 Western blot

Cells of confluent 10 cm dishes were washed three times with ice-cold PBS containing PMSF (50 µg/µL). 1 mL of ice-cold RIPA buffer was added and cells were harvested with a plastic cell scraper. The suspension was then transferred into a pre-cooled 1.5 mL tube and lysed for 30–60 min on ice. After centrifugation at R_{max} for 15 min at 4°C, the supernatant was mixed with protein loading buffer and heated to 95°C for 5 min. Samples were subsequently separated on a 10% SDS-polyacrylamide gel for 1–2h at 45 mA per gel in SDS-running buffer. Following SDS-PAGE, proteins were blotted onto a nitrocellulose membrane using the Criterion™ Blotter at 1000 mA for 1h at 4°C. The membrane was blocked in 5% milk powder in PBS for 1h at room temperature and incubated overnight at 4°C in blocking solution containing the primary antibody at the recommended dilution. The membrane was washed three times in PBS for 15 min and incubated with the peroxidase coupled secondary antibody (diluted in 5% milk powder in PBS) for 1–2 h at room temperature. After another washing cycle with PBS, the membrane was incubated with an ECL (enhanced chemiluminiscence) substrate for approximately 1 min and protein bands were visualized using the Image Station 4000 MM PRO (Kodak). Antibodies used for Western blotting are listed in Table 7 (3.5).

4.12 MTT Assay

Cell metabolic activity and hence proliferation behavior of the stable cell lines was assessed by MTT assay. $0.5x10^3$ mESCs per well were seeded at equal amounts (five technical replicates per condition) on gelatinized 96-well plates. Medium without cells was used as a background control. One day after seeding, cells were treated with doxycycline (300 ng/mL). Medium was changed every day with a final volume of 100 µL. On each timepoint (day 0, 1, 2, 3, 4), 20 µL of MTT-solution were added to the wells and incubated for 2h at 37°C and 5% CO_2. MTT containing medium was carefully removed and plates were kept protected from light at –20°C until measurement. 150 µL DMSO were added to the wells and incubated for 1h at room temperature while agitating on a shaker. Plates were covered with tinfoil during incubation. Absorbance of the purple formazan solution was measured at 540 nm with a reference filter at 690 nm using a Berthold Tristar multimode reader.

4.13 Crystal Violet Assay

Crystal violet assay was additionally performed to assess cell proliferation and cell viability of the stable cell lines. Therefore, cells were seeded in triplicates at equal amounts ($1.0x10^4$ per well) on gelatinized 24-well plates. Medium without cells served as a background control. Doxycycline (300 ng/mL) induction was initiated one day after seeding and medium was daily changed. For each timepoint (day 0, 1, 2, 3, 4), cells were carefully washed with PBS and fixed with 1 mL of ice-cold methanol (≥ 99.8%) for 10 min. Plates were stored at 4°C until use. To stain cells, 500 µL of 0.1% crystal dye dissolved in 20% EtOH were added to the plates and incubated for 30 min at room

temperature. After washing the cells at least three times with ddH_2O until there was no remaining staining solution, plates were air-dried at room temperature for at least 2h or preferably overnight. To solubilize the stain, 500 µL of 10% acetic acid were added and incubated for 1h upon agitating on a shaker until color was uniform. Finally, 100 µL of the solution were transferred to 96-well plates and absorbance was measured at 540 nm (reference filter at 690 nm) with a Berthold Tristar multimode reader.

4.14 Cell cycle analysis

After 48h induction with doxycycline, cells (at least $1.0x10^5$) were harvested by trypsinization and washed with 10 mL ice-cold PBS. 1 mL of ice-cold 80% ethanol was dropwise added to the cells while slowly vortexing to prevent clumping. Fixed cells were stored at –20°C until usage. Following centrifugation, all steps were carried out at 4°C. Cells were centrifuged at 1000 rpm for 10 min and the pellet was washed with ice-cold PBS. After thoroughly resuspending the pellet with a pipette, cells were spun down at 900 rpm for 5 min. Cells were then resuspended in 500 µL sodium citrate buffer (38 mM) and treated with 25 µL RNase A (10 mg/mL). After a 30 min incubation at 37°C, cells were stained with 15 µL propidium iodide (PI) (1 mg/mL). Cytomics FC 500 flow cytometer (Beckman Coulter) was used for measurement. $1.0x10^4$ cells were acquired and subsequently analyzed with the CXP 2.2 software to determine the percentages of cells in G1, S, and G2/M phase. Doublets of cells were excluded by selective gating.

4.15 BrdU proliferation assay

To monitor S-phase progression in the DGCR8 mESC lines, incorporation of bromodeoxyuridine (BrdU) during DNA synthesis was colorimetrically measured using the Cell Proliferation ELISA BrdU kit (Roche). Therefore, $0.5x10^3$ mESCs per well were seeded at equal amounts (three technical replicates per condition) on gelatinized 96-well plates. Medium without cells was used as a background control. One day after seeding, cells were treated with doxycycline (300 ng/mL). After 48h induction, 10 µl of 100 µM BrdU solution was added to the cells in a final volume of 100 µL. Cells were incubated for 6h at 37°C and 5% CO_2. Labeling medium was removed, and plates were stored at 4°C until usage. All subsequent steps were carried out according to the manufacturer's protocol. Absorbance was measured at 450 nm (reference filter at 690 nm) with a Berthold Tristar multimode reader

4.16 Microarray analysis of miRNA expression

For miRNA expression analysis, cells were harvested 48h post-induction, snap frozen and stored at -80°C until usage. Total RNA including miRNAs was extracted using the Qiagen miRNeasy Kit according to the manufacturer's protocol. Quality of the RNA samples was checked using the PicoRNA kit (Agilent) on a Bioanalyzer 2100. MiRNA expression profiles were generated using the Human miRNA Microarray Kit (release 16.0) and Mouse miRNA Microarray Kit (release 19.0),

respectively. Raw expression values were extracted with the Feature Extraction Software (Agilent) and data were subsequently analyzed using different R/bioconductor packages (http://www.bioconductor.org). Signal intensities were normalized by variance stabilization (Huber et al. 2002). Differences in miRNA expression were identified using the limma package (Ritchie et al. 2015). All the experimental procedures of the microarray analysis were performed at the Institute of Human Genetics in Homburg, Germany and data was bioinformatically analyzed by Dr. Nicole Ludwig and Dr. Susanne Kneitz. Expression data of HEK293T Drosha wild-type/E1147K can be found in GEO (https://www.ncbi.nlm.nih.gov/geo/) under the accession number GSE60081 (Wegert et al. 2015).

4.17 Transcriptome Analysis of miRNA and mRNA

For transcriptome analysis, cells were harvested 48h post-induction. Total RNA was isolated using peqGold TriFast™ as described above (4.6).

MiRNA-sequencing was performed at the Core Unit Systems Medicine in Wuerzburg. Before library preparation, quality of the RNA samples was assessed using the Bioanalyzer 2100 (Agilent). The sequencing library was generated using the NEBNext® Multiplex Small RNA Library Prep Set for Illumina®. All steps were carried out according to the manufacturer's protocol. RNA sequencing was performed using the Illumina® NextSeq® 500 sequencing platform.

MRNA-sequencing was done by the BGI NGS Lab (Hong Kong) on a BGISEQ-500 platform (hexamer-primed oligo-dT selected RNA, 100 bp paired end, approximately 70 million reads per sample).

Raw and trimmed reads were aligned to the mouse reference genome assembly (GRCm38.p6) using STAR aligner (version 2.6.0c) (Dobin et al. 2013). SAMtools (version 1.3) was used to convert SAM to BAM format (Li et al. 2009). After quality control, raw counts for each mature miRNA (miRBase version 22) and for each gene were extracted using BEDTools (version 2.17.0) (Quinlan and Hall 2010). For downstream analyses, custom python and R scripts have been created by Tim Kehl (Saarland University). Enrichment analyses were conducted using the GeneTrail2 web service (Stockel et al. 2016).

Bioinformatical processing of miRNA- and mRNA-sequencing data was conducted by Tim Kehl and Dr. Susanne Kneitz.

5 Results

5.1 Recurrent hotspot mutations in the microprocessor genes *DROSHA* and *DGCR8*

Our exome sequence analysis of high-risk blastemal tumors could reveal several hotspot mutations affecting the microprocessor genes *DROSHA* and *DGCR8*. Both genes showed somatic mutations in regions that are important for catalysis and RNA binding, respectively. The most frequent and functionally relevant hotspot mutations were the *DROSHA* E1147K and the *DGCR8* E518K mutation, which will be described in the following sections.

5.1.1 *DROSHA* E1147K

Four of five *DROSHA* mutations were located in one of the three key metal binding sites of the RNase IIIb domain which are essential for its catalytic activity (Gan et al. 2006; Wegert et al. 2015). E1147K was the most frequent *DROSHA* mutation (n=3) observed in our exome sequence analysis and showed heterozygous expression implying a dominant effect. The point mutation G->A led to an amino acid exchange from glutamic acid to lysine. The same amino acid exchange was observed in the E1222K mutation (n=1) which also affected the active site of the RNase IIIb domain. The fifth mutation D973H affected the catalytic center of the RNase IIIa domain (Wegert et al. 2015) (Figure 6).

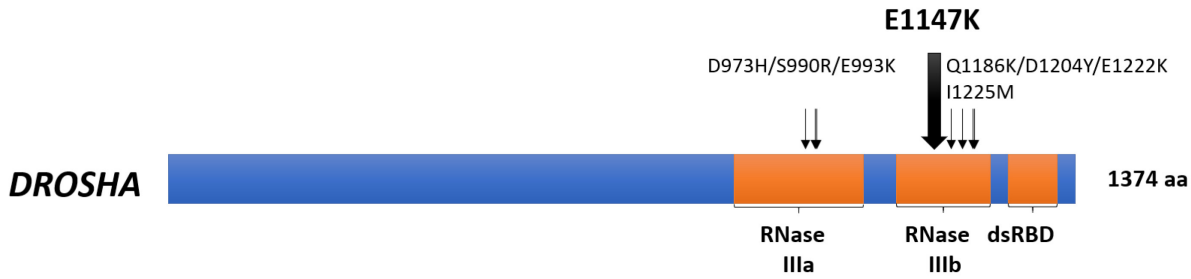


Figure 6: Position of the hotspot mutation E1147K in the RNase IIIb domain of *DROSHA* (Illustration adapted from (Wegert et al. 2015)).

Given the fact that these mutations affected key metal binding sites in the RNase IIIb as well as in the RNase IIIa domain, we directly sequenced *DROSHA* cDNA of either domain in an independent set of 100 WTs. We could further detect seven somatic mutations (S990R, 2xE993K, E1147K, Q1186K, D1204Y, I1225M). Again, alterations were located at residues in the catalytic center of *DROSHA* (E1147K, Q1186K, I1225M). Three cases were found to be outside (S990R, E993K, D1204Y).

5.1.2 *DGCR8* E518K

Four tumors of our cohort harbored the hotspot mutation E518K affecting the double stranded RNA binding domain (dsRBD) of *DGCR8*. The point mutation from G->A resulted in an amino acid alteration with a charge reversal in the RNA binding domain. Glutamic acid at position 518 was

replaced by the positively charged amino acid lysine. The E518K mutation was homozygously expressed in the tumor samples indicative for the mutation to be recessive. Additional *DGCR8* mutations which only occurred once, included a homozygous single-nucleotide insertion leading to a frameshift and early protein truncation after the amino acid 81, and a heterozygous E213X nonsense mutation (Wegert et al. 2015) (Figure 7).

Figure 7: Location of the hotspot mutation E518K in the double-stranded RNA binding domain (dsRBD) of *DGCR8* (Illustration adapted from (Wegert et al. 2015)).

5.2 Mutation frequency of *DROSHA* E1147K and *DGCR8* E518K in Wilms tumors

To test whether these microprocessor gene mutations were indeed frequent events in WT, a larger cohort of unselected WT samples were screened via allele-specific PCR. Additional 13 of 367 analyzed cases harbored the *DROSHA* E1147K mutation leading to a cumulative incidence of 3.2%. The E1147K mutation always showed heterozygous expression in genomic tumor DNA and cDNA, clearly indicative of its dominant effect. Allele-specific PCR of the E518K mutation on 719 cases followed by sequencing could identify 20 additional *DGCR8* mutated tumors adding up to 3.1%. In contrast to the *DROSHA* mutation, the *DGCR8* E518K alteration showed homozygous expression in all cases, underlining that the mutation must act recessive. Analysis of the mutation frequencies according to the histological subtypes in WTs, revealed that *DROSHA* E1147K and *DGCR8* E518K occurred more frequently in tumors with a high-risk histology (5.8%, 6/104) compared to low and intermediate risk tumors (2.7%, 11/403 and 2.8%, 18/650) (Figure 8A and Table S1). Both mutations were not correlated with increased death or relapse/metastasis rates in WT patients (p-value<0.05, χ^2 test) (Table S2).

Strikingly however, there was a significant gender bias of *DGCR8* mutated tumors, since 21 of 24 of the cases affected female patients (88%, p-value<0.05, χ^2 test), while distribution was similar between female and male patients in our entire cohort (343 vs. 430) (Figure 8B).

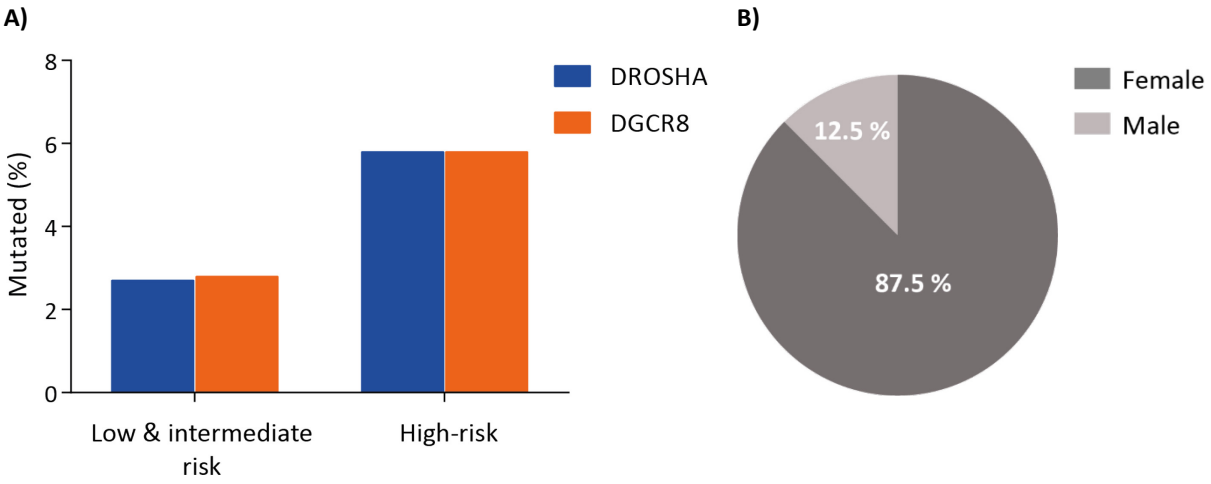


Figure 8: Mutation frequency of *DROSHA* E1147K and *DGCR8* E518K in Wilms tumors
Overview of the mutation screening data of unselected Wilms tumor samples.
(A) Diagram showing mutation rates (%) in low and intermediate risk compared to high-risk WTs.
(B) Gender bias in *DGCR8* mutated WTs; 88 % of the cases arose in female patients (p-value<0.05, χ^2 test).

5.3 Generation of stable cell lines

To investigate the functional consequences of the microprocessor mutations *DROSHA* E1147K and *DGCR8* E518K *in vitro*, stable cell lines were generated with inducible overexpression of the respective wild-type or mutant form linked with IRES-dependent EGFP expression. To mirror heterozygous expression of the *DROSHA* E1147K mutation observed in WT, HEK293T cells that already endogenously express wild-type *DROSHA* were used. Homozygously expressed *DGCR8* E518K, on the other hand, required a model system that lacks endogenous DGCR8. Therefore, DGCR8 knockout mESCs (Wang et al. 2007) were used to functionally assess the *DGCR8* E518K mutation. Within each cell system, respective wild-type and mutant transgene showed similar expression levels on mRNA as well as on protein level after induction with doxycycline for 48h (Figure 9 and 10).

Expression of DROSHA wild-type and E1147K in HEK293T cells was induced with 1000 ng/mL doxycycline (Figure 9). All replicates showed similar expression levels of DROSHA on both protein and mRNA level, comparable to endogenous DROSHA expression in non-transfected HEK293T cells (Figure 9A and C). Overexpression of the transgene therefore seemed to remain within physiological levels. Expression of transgenic Myc-tagged DROSHA was likewise comparable among DROSHA-wild-type and -E1147K replicates on protein and mRNA level (Figure 9B and C).

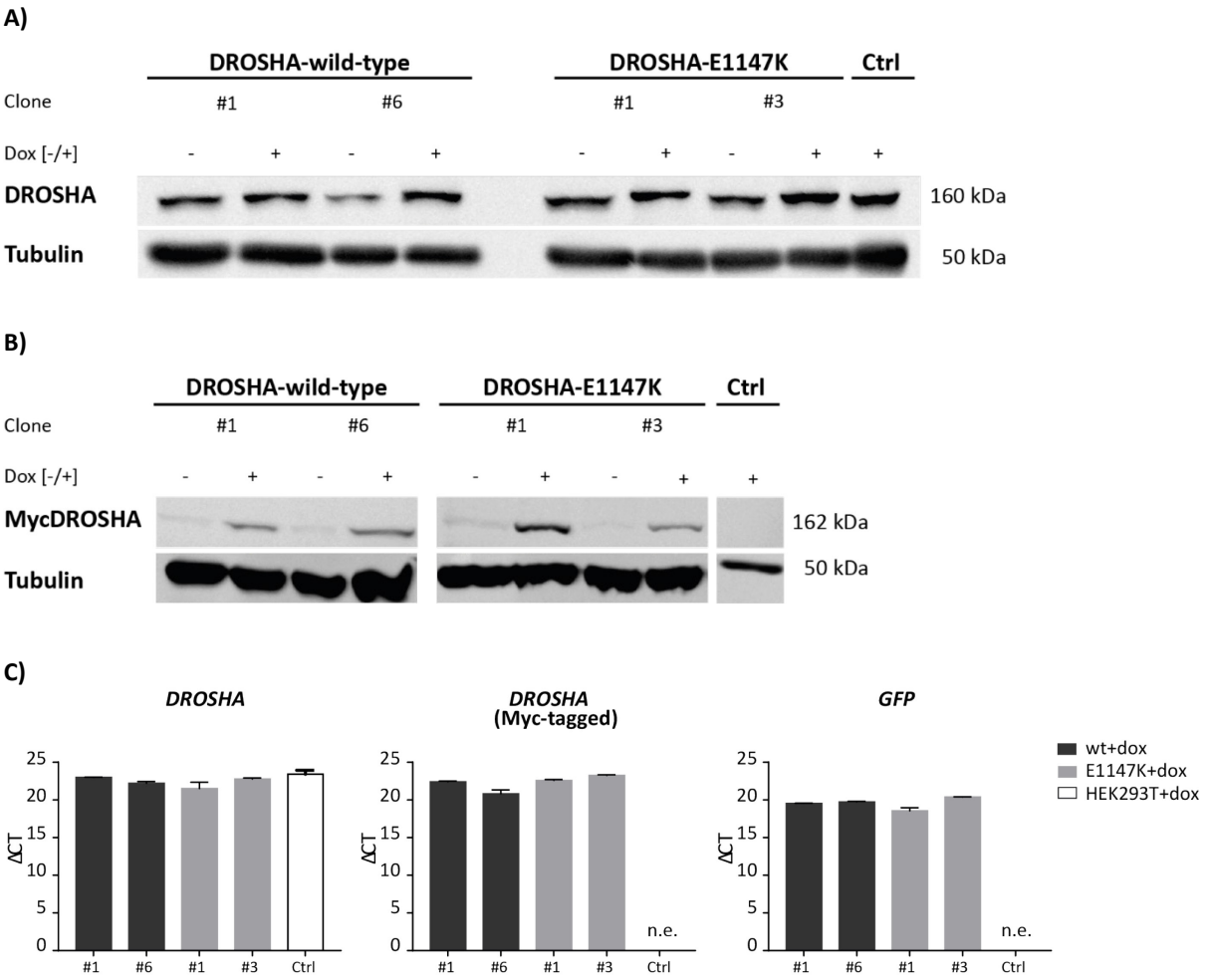


Figure 9: Stable HEK293T cells expressing DROSHA wild-type or E1147K

(A) α-Drosha Western blot analysis of HEK293T DROSHA-wild-type and DROSHA-E1147K cells treated with 1000 ng/mL doxycycline for 48h. Two biological replicates were used per condition. Non-transfected HEK293T cells served as a control.

(B) α-Myc Western blot analysis of HEK293T DROSHA-wild-type and DROSHA-E1147K cells demonstrating induction of Myc-tagged DROSHA after 48h of doxycycline treatment. Samples without doxycycline treatment showed leaky expression of the transgene. Non-transfected HEK293T cells served as a negative control.

(C) qPCR data showing mRNA expression levels of *DROSHA*, Myc-tagged *DROSHA*, and *GFP* in biological duplicates (except HEK293T+dox–only one). ΔCT values represent CT-values normalized to the housekeeping gene *HPRT*. Error bars indicate standard deviation of two technical replicates. ΔCT values ≥30 were considered as not expressed (n.e.).

DGCR8-wild-type and DGCR8-E518K mESCs showed similar GFP expression among replicates (Figure 10A). Analysis of Flag DGCR8 via Western blotting showed increasing expression levels with higher doses of doxycycline (Figure 10B). Induction with 300 ng/mL seemed to lead to the most comparable expression levels and was therefore used for all subsequent experiments conducted with the stable DGCR8 cell lines. Also, on mRNA level *Flag DGCR8* wild-type and E518K showed similar expression levels (Figure 10C).

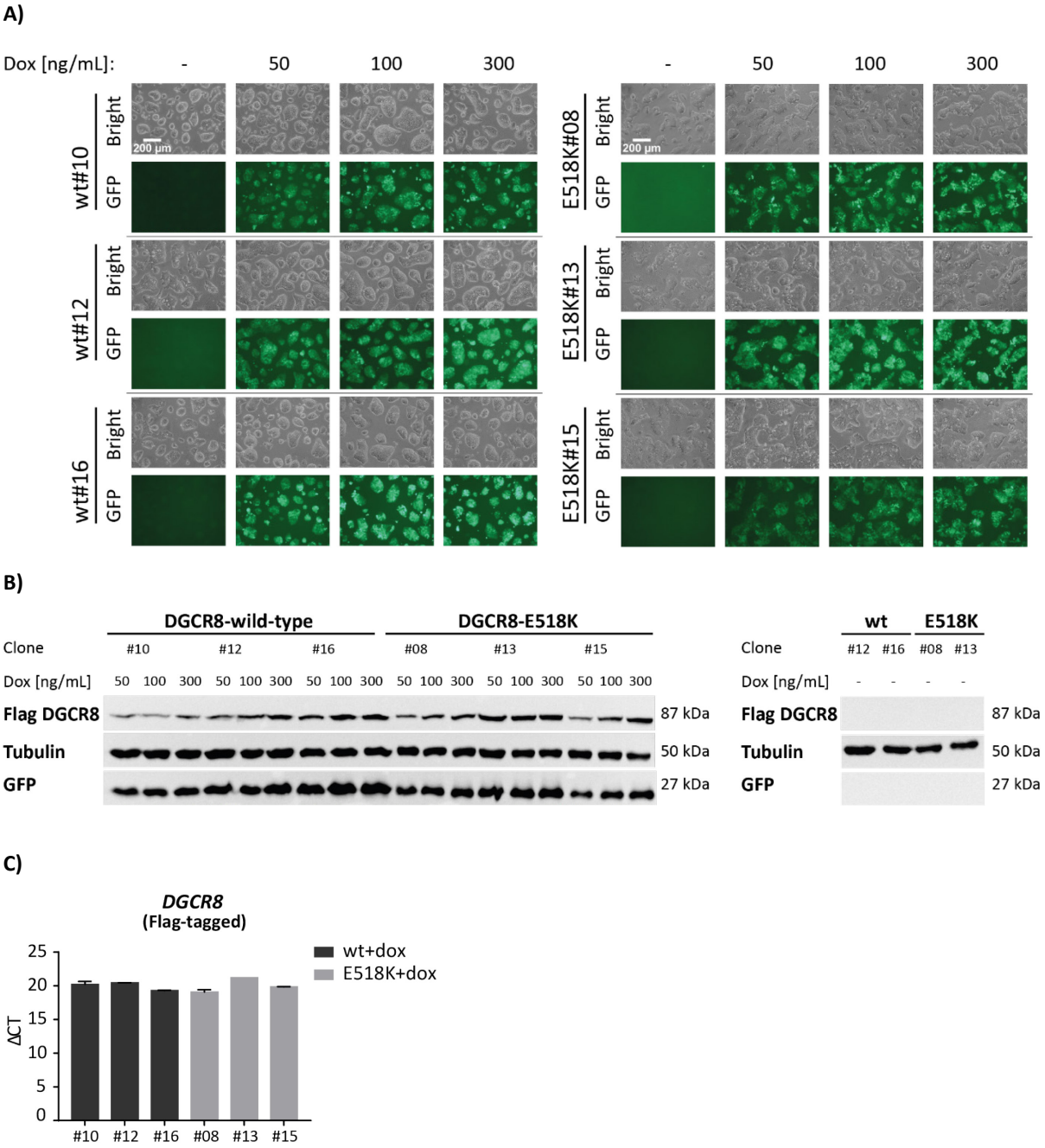


Figure 10: Stable mouse embryonic cell lines expressing DGCR8 wild-type or E518K

(A) DGCR8-wild-type (DGCR8-wt) and DGCR8-E518K mESCs upon 48h induction with 50, 100 and 300 ng/mL doxycycline. Images were taken under bright-field or dark-field showing similar GFP expression among wild-type and mutant clones of independent biological triplicates.

(B) α-Flag and α-GFP Western blot analysis demonstrating induction of Flag DGCR8 on protein level with different concentrations of doxycycline (as described in A). Right figure showing no leaky expression of the transgene without doxycycline treatment, represented by two biological replicates per condition.

(C) qPCR results showing similar mRNA expression levels of *Flag DGCR8* wild-type and E518K after induction with 300 ng/mL doxycycline for 48h. ΔCT values represent CT-values normalized to the housekeeping gene *Hprt*. Error bars indicate standard deviation of two technical replicates.

Both vector constructs seemed to have leaky expression on mRNA level, since in non-induced samples specific PCR products with CT-values under 30 (range of 25–29) could still be detected. However, leaky expression on protein level could only be appreciated in HEK293T DROSHA cells.

5.4 Effects of the microprocessor mutations on miRNA expression

To functionally assess the consequences of the microprocessor mutations, miRNA expression of the stable cell lines HEK293T DROSHA cells and DGCR8 mESCs with inducible overexpression of respective wild-type or mutant transgene was investigated by microarray analysis.

5.4.1 Microarray analysis of small RNAs in HEK293T DROSHA-wild-type and -E1147K cells

Microarray analysis revealed changes in the miRNA profiles in HEK293T cells upon inducible expression of wild-type or mutant DROSHA. While induction of wild-type DROSHA in HEK293T cells did not alter the miRNA pattern, expression of DROSHA E1147K led to a downregulation of many of the expressed small RNAs (Figure 11). Upon induction of DROSHA E1147K, all 145 significantly regulated miRNAs (adj. p-value<0.05) showed reduced expression with an average fold-change of -1.7 (Wegert et al. 2015). Interestingly, these downregulated miRNAs are mostly defined as high-confidence miRNAs according to miRBase (Kozomara and Griffiths-Jones 2014), whereas the group of small RNAs that mainly stayed unaltered upon induction of DROSHA E1147K, are not classified as canonical miRNAs.

Heterozygous expression of DROSHA E1147K therefore seemed to be sufficient to globally reduce miRNA expression implying a dominant-negative effect on miRNA processing.

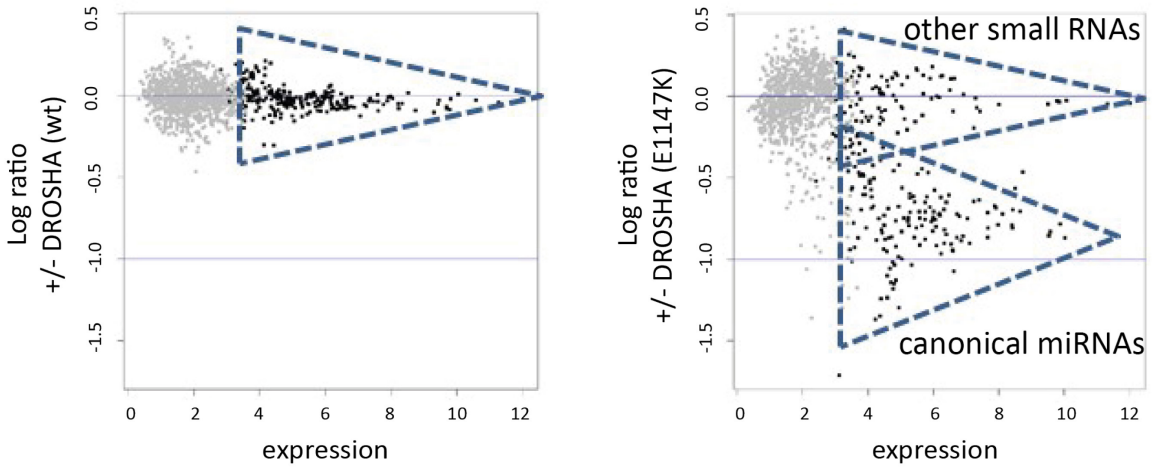


Figure 11: Effect of DROSHA E1147K on miRNA expression

Alterations of miRNA profiles in HEK293T DROSHA-wild-type and -E1147K cells upon induction. Biological duplicates were hybridized onto Agilent miRNA arrays. Black dots represent RNAs with a present call >50% by the feature extraction software; others are presented in gray. Blue triangles show cluster of small RNAs that are either defined as canonical miRNAs or other small RNAs according to miRBase (Wegert et al. 2015).

5.4.2 Microarray analysis of small RNAs in DGCR8-wild-type and -E518K mESCs

It was previously shown that deletion of DGCR8 in mouse embryonic stem cells leads to a global loss of mature miRNAs, underlining the essential role of DGCR8 for miRNA biogenesis (Wang et al. 2007). We could confirm the miRNA-deficient phenotype of DGCR8-knockout mESCs which showed reduced expression levels of high-confidence miRNAs according to miRBase (Kozomara and Griffiths-Jones 2014), compared to wild-type and mutant situation (Figure 12A). Induction of DGCR8-wild-type led to much higher miRNA expression rescuing the miRNA processing defect in DGCR8-KO cells, whereas DGCR8-E518K mutant cells provided only a partial rescue. In the absence of doxycycline induction, miRNA expression could be observed in both wild-type and mutant DGCR8, albeit at lower levels. Although leaky expression could only be observed on mRNA level, it is likely that leaky expression of the transgenes slightly raised protein levels which were insufficient to be visualized on Western blot, yet enough to restore miRNA expression to some extent.

In contrast to canonical miRNAs, DGCR8-independent miRNAs, such as mirtrons and other small RNAs that bypass DGCR8 activity in their maturation process (Castellano and Stebbing 2013), were expressed at higher levels in DGCR8-knockout and -E518K mESCs compared to DGCR8-wild-type upon induction (Figure 12B). The same trend could be observed for low-confidence miRNAs and those that are not annotated as miRNAs and are likely to belong to other groups of small RNAs. Despite known DGCR8-independent miRNAs (n=61) (Castellano and Stebbing 2013) were not included in the analysis of high-confidence miRNAs shown in Figure 12A, there were still miRNAs that exhibited a ≥ 1.5x higher expression in DGCR8-KO relative to DGCR8-wt (n=27/128). These miRNAs showed similarly high expression in DGCR8-E518K. Most likely, these miRNAs originate from unavoidable contamination with mouse embryonic fibroblast (MEF) feeder cells, since RNA used for microarray analysis was extracted from mESCs that were being weaned off from feeder cells. Indeed, most of these miRNAs were likewise expressed in the PSNL feeder control group (n=21/27) (Table S3). Furthermore, higher expression of these miRNAs in DGCR8-KO and -E518K could not be confirmed by miRNA-Seq analysis of the DGCR8-mESC lines that had been cultured under feeder-free condition over several passages.

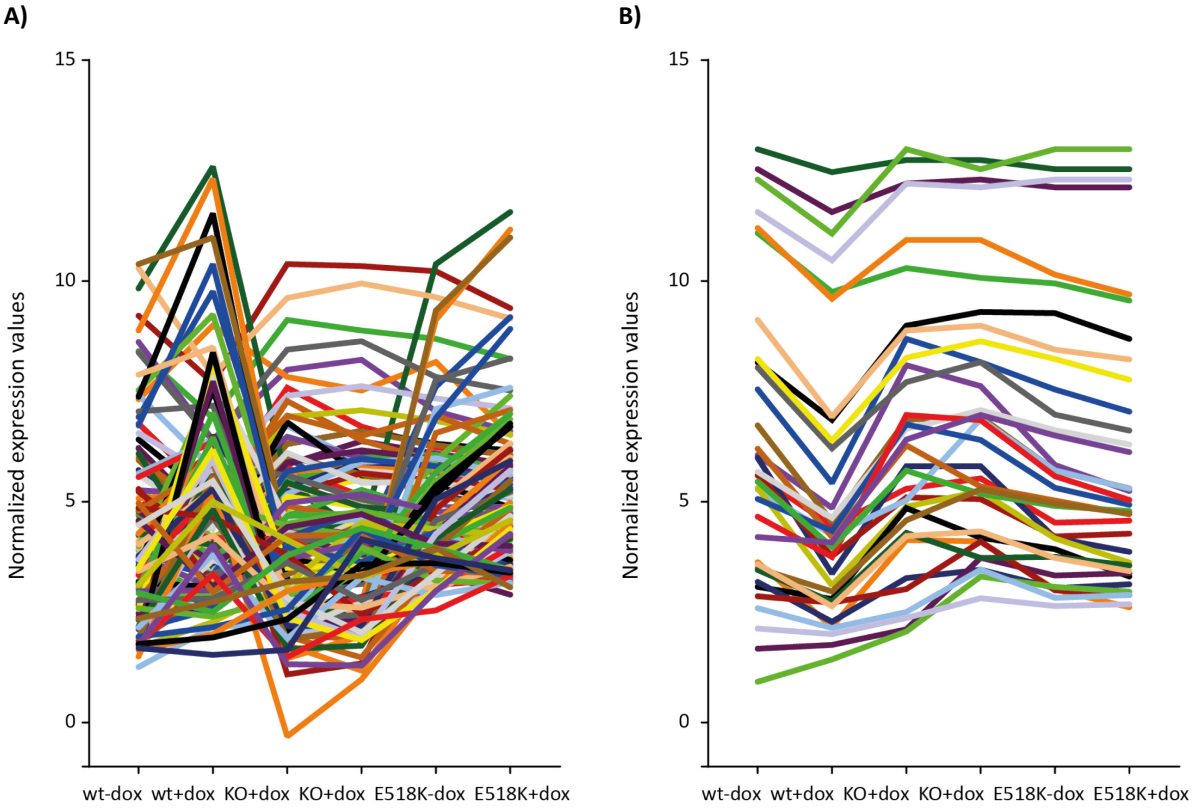


Figure 12: DGCR8 E518K partially rescues the miRNA processing defect in DGCR8-KO mESCs
MiRNA expression of DGCR8-wild-type (wt) and DGCR8-E518K (E518K) mESCs upon induction with doxycycline. DGCR8-knockout (KO) mESCs treated with doxycycline served as negative controls. Technical triplicates of DGCR8-wt/-E518K samples and duplicates of DGCR8-KO were hybridized onto Agilent miRNA arrays. Mean expression values of technical replicates were normalized using quantile-normalization.
(A) Plot showing high-confidence miRNAs according to miRBase with a present call of ≥50% (128/653) by the feature extraction software. MiRNA processing defect in DGCR8-knockout was rescued by DGCR8 wild-type and partially by E518K. DGCR8-independent small RNAs shown in Figure B are not included.
(B) Plot showing DGCR8-independent small RNAs with a present call of ≥50% (35/61). Expression of non-canonical miRNAs was highest in DGCR8-knockout and -E518K mutant cells.

5.4.3 miRNA-Sequence analysis of DGCR8-knockout, -wild-type and -E518K mESCs

To gain further insight into the altered miRNA expression profile of DGCR8-E518K identified by microarray-analysis, RNA-Seq was performed. We compared expression of miRNAs in DGCR8-knockout, -wild-type and -E518K mESCs to detect clusters of differentially expressed miRNAs and thereby reveal possible preferences in the processing of certain miRNA families. Furthermore, we looked at the length distributions of differentially expressed miRNAs to analyze whether aberrant miRNA processing in E518K would result in altered miRNA length distributions. Also, base distributions of the seed-sequence which was previously reported to be important for target gene regulation (Wang 2013; Wang 2014) was analyzed to check for altered base composition, which in turn could impact miRNA-target binding.

5.4.3.1 Normalization and clustering

Normalization solely based on miRNA counts was not possible, since miRNA ratios strongly varied between the different groups (Table S4). Therefore, proportions of other RNA classes needed to be considered. Since biogenesis of piRNAs is independent of DGCR8, it was assumed that ratios of piRNAs stayed constant among the groups and were hence used for normalization (Figure 13).

In order to find a suitable normalization, the proportions were modelled as a random experiment with n independent draws without replacement. For each sample, a population of R RNA molecules was assumed to consist of M miRNA molecules, P piRNA molecules and O other RNA molecules. The final normalization formula for each miRNA can then be derived from a hypergeometric distribution that described the probability of $\mathrm{m_i}$ successes in our n draws (Figure S1).

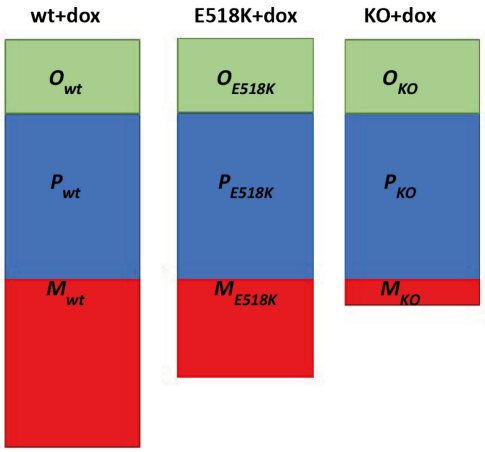


Figure 13: Schematic representation of the RNA distribution in DGCR8-wild-type, -E518K and knockout mESCs
MiRNA (M) concentrations vary among DGCR8-wild-type, -E518K and knockout mESCs. Proportions of piRNAs (P) and other groups of RNA (O) were assumed to remain stable.

Following this approach, we got a normalization formula that is closely related to the reads per million (RPM) normalization that is often used for miRNA-Seq data:

$$pRPM = p\frac{m_i}{m} * R = p\frac{m_i}{m} * 10^6$$

Here p is the estimated proportion of miRNAs in the RNA population and m is the total number of miRNA molecules in the considered sample. This proportion can either be individually estimated for each sample or collectively for all replicates. For the following analyses we used the latter approach.

Average-linkage clustering using the Euclidean distance was then performed to check if the miRNA expression values reflect the different groups. Biological replicates of each condition nicely clustered together, outliers could be excluded. DGCR8-wild-type without doxycycline treatment clustered closer to the E518K group upon induction, whereas E518K without induction clustered more towards the knockout situation (Figure 14). Proximity between non-induced wild-type and induced E518K was

most likely due to leaky expression of DGCR8-wild-type, whose expression was sufficient to partially restore miRNA expression forming a similar pattern as the E518K mutant upon induction. Leaky expression of E518K without induction, unable to rescue the processing defect, exhibited a miRNA expression profile rather comparable to that of miRNA-deficient DGCR8-knockout.

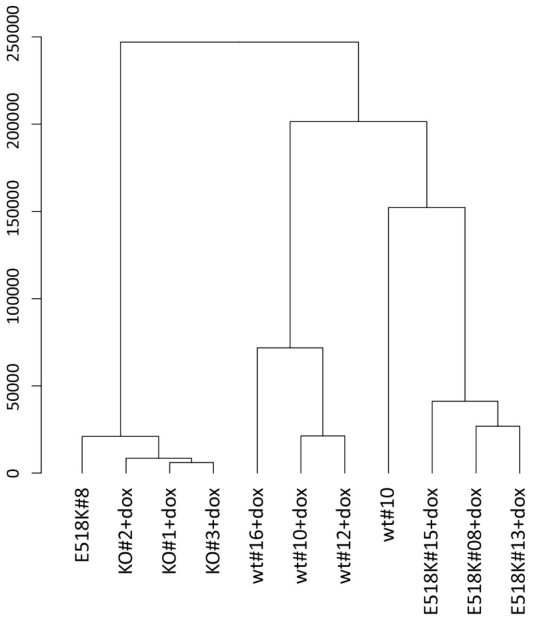


Figure 14: Cluster analysis of DGCR8-knockout, - wild-type and -E518K mESCs
Euclidean distance-based hierarchical clustering of all RNA-Seq samples consisting of three biological replicates per condition (except DGCR8-wild-type and -E518K without doxycycline treatment–only one sample each). Y-axis represents the distance between the clusters.

5.4.3.2 Group comparison

After normalization, the three groups exhibited the same pattern as already seen by raw data analysis. MiRNA expression strongly differed between the groups, with DGCR8-knockout mESCs containing the lowest total amounts of miRNAs, mainly expressing non-canonical miRNAs. Introduction of DGCR8-wild-type nicely restored miRNA expression, whereas the E518K mutation only led to a partial rescue with moderate miRNA levels (Figure 15A and S2).

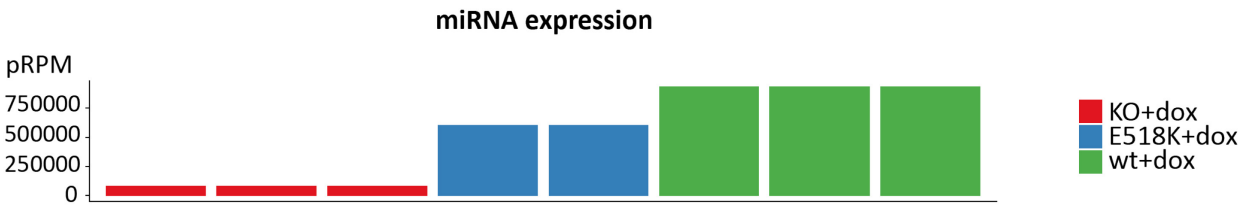
For further analysis, only miRNAs that had a pRPM (proportional reads per million) of at least 20 in one of the three groups (n=382/1966) were considered. Due to deviant piRNA counts in one of the E518K samples (Table S4), only two replicates of E518K were taken for subsequent calculations. Normalized read counts were scaled to values ranging from 0–1 among all the samples and were visualized in a heatmap. MiRNAs with a scaled expression value ≥0.50 were considered as highly expressed, whereas those with values ≤0.25 were defined as weakly expressed miRNAs. Based on these classifications, five different miRNA clusters emerged (Figure 15B and S2). The biggest cluster was formed by miRNAs that showed the highest expression levels in the wild-type situation (1,

N=153), followed by miRNAs that were highly expressed in both wild-type and E518K (4, N=66). MiRNAs that were highest expressed in E518K formed an additional cluster (2, N=24). Another group of miRNAs was formed by non-canonical miRNAs that showed the strongest expression in the knockout situation (3, N=26). Finally, five non-canonical miRNAs were expressed higher in both knockout and E518K compared to the wild-type situation (5). Separate analysis of 5p- or 3p-arm miRNAs (n=199; n=183) showed that differentially expressed miRNAs formed the same clusters with similar miRNA distribution as when both arms were analyzed together (Figure S2).

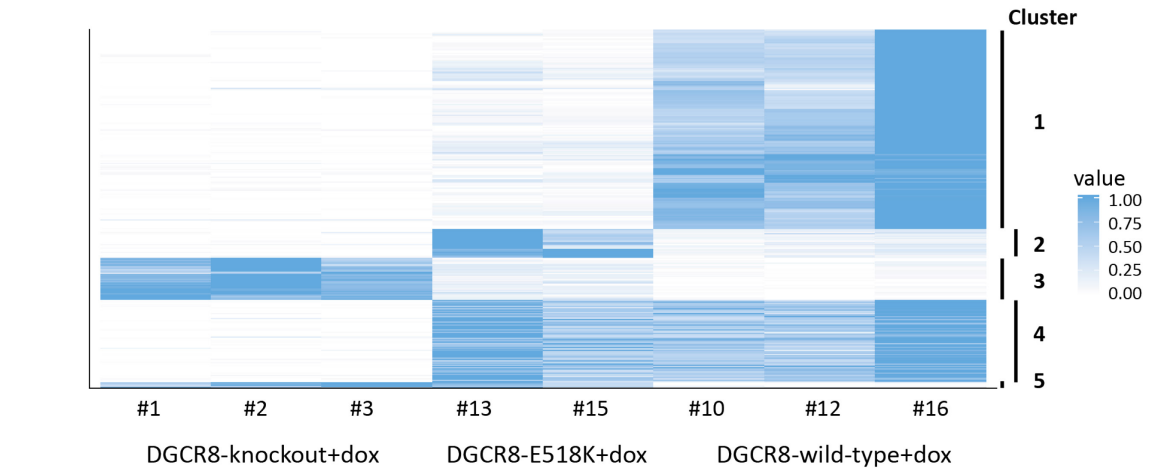
Analysis of non-canonical miRNAs (Castellano and Stebbing 2013) (N=20) showed that expression was highest in DGCR8-KO, followed by moderate levels in E518K, whereas DGCR8-wild-type showed the lowest expression of non-canonical miRNAs (Figure 15C).

To validate differential miRNA expression between DGCR8-wildtype and -E518K, miRNAs from clusters 1 and 2 were selected and analyzed by qPCR (Figure 15D and E). Expression of canonical miRNAs miR-302d-3p, miR-130b-3p, miR-200b-3p, miR-425-5p, and miR-196a-5p was significantly reduced in DGCR8-KO. Rescue of these canonical miRNAs by DGCR8-wild-type shown by RNA-Seq could be confirmed by qPCR. DGCR8-E518K, on the other hand, exhibited only a partial rescue with decreased expression levels compared to DGCR8-wt. Expression level of let-7a was higher in E518K with generally smaller differences, however.

A)



B)



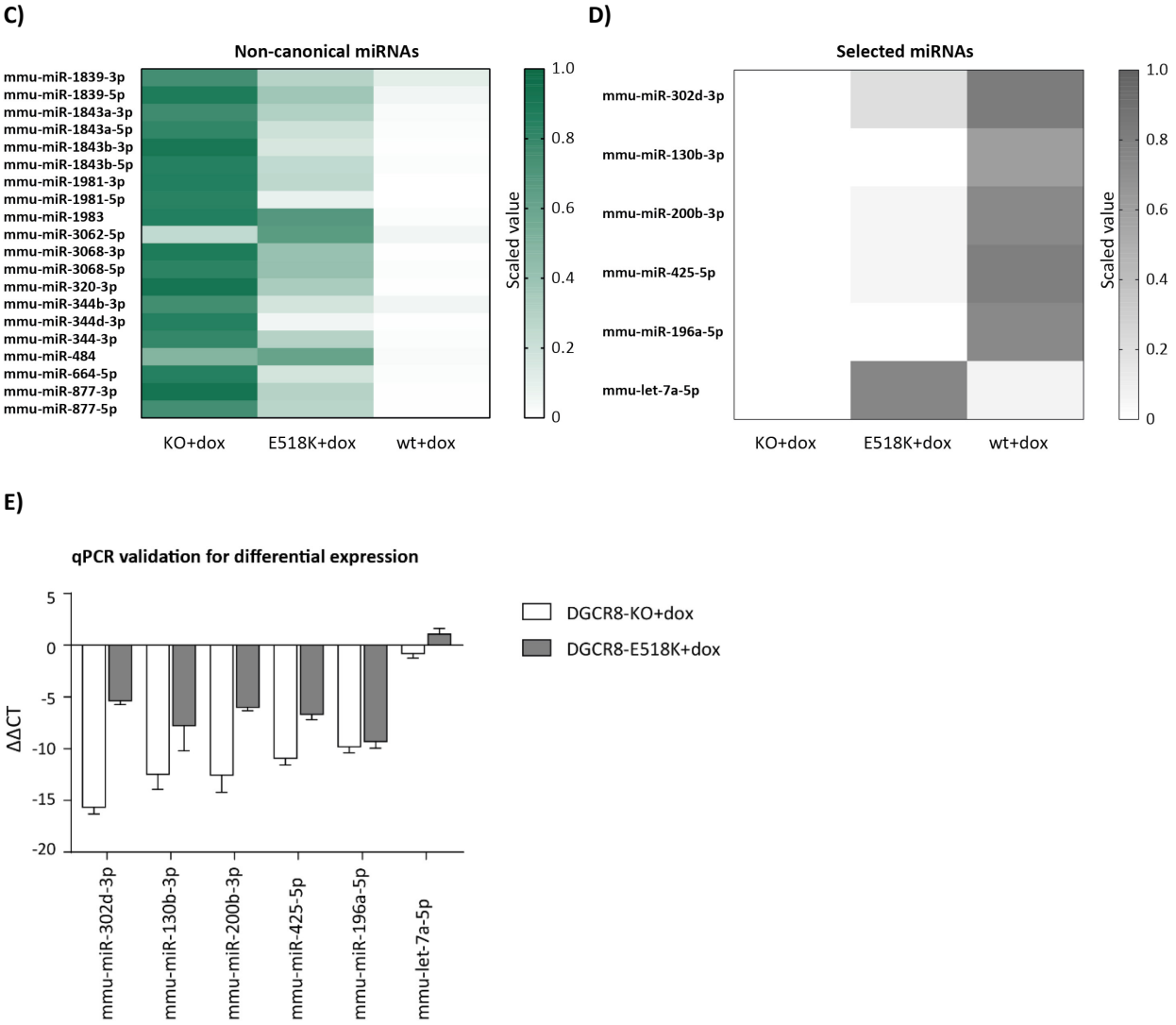


Figure 15: Comparison of miRNA expression between DGCR8-knockout, -E518K and -wild-type mESCs

(A) Bar chart showing differences in miRNA expression levels between DGCR8-KO, -E518K and -wt upon induction. Expression values are presented as proportional reads per million (pRPM) normalized to piRNAs. Three biological replicates were used per group (except E518K–only two samples).

(B) Heatmap showing clusters of differentially expressed miRNAs in DGCR8-KO, -E518K and -wt with a pRPM≥20. Values represent scaled miRNA expression ranging from 0–1. Cluster 1: KO≤0.25, E518K≤0.25, wt≥0.50 (N=153); Cluster 2: KO≤0.25, E518K≥0.50, wt≤0.25 (N=24); Cluster 3: KO≥0.50, E518K≤0.25, wt≤0.25 (N=26), Cluster 4: KO≤0.25, E518K≥0.50, wt≥0.50 (N=66); Cluster 5: KO≥0.50, E518K≥0.50, wt≤0.25 (N=5).

C) Heatmap showing non-canonical miRNAs (Castellano and Stebbing 2013) in DGCR8-KO,-E518K and -wt mESCs upon induction. Values represent scaled miRNA expression levels from RNA-Seq data, ranging from 0–1. Three biological replicates were used per condition (except E518K–only two samples).

(D) Heatmap of selected canonical miRNAs miR-302d-3p, miR-130b-3p, miR-200b-3p, miR-425-5p, miR-196a-5p let-7a. Values represent scaled miRNA expression levels from RNA-Seq data, ranging from 0–1.

(E) qPCR results of selected miRNAs shown in (D). MiRNA expression was normalized to non-canonical miR-320-3p. Bar chart represents delta delta CT values of DGCR8-KO and DGCR8-E518K relative to DGCR8-wt. Error bars indicate standard deviation of three biological replicates.

Furthermore, miRNAs that were higher expressed in DGCR8-wild-type concomitantly showed lower pri-miRNA expression levels, indicating that processing of pri-miRNAs to mature miRNAs was more efficient in wild-type. In contrast, DGCR8-KO and -E518K exhibited reduced expression of these mature miRNAs while showing an accumulation of their pri-miRNAs (Figure 16A). This was later confirmed by mRNA-Seq analysis (Figure 16B).

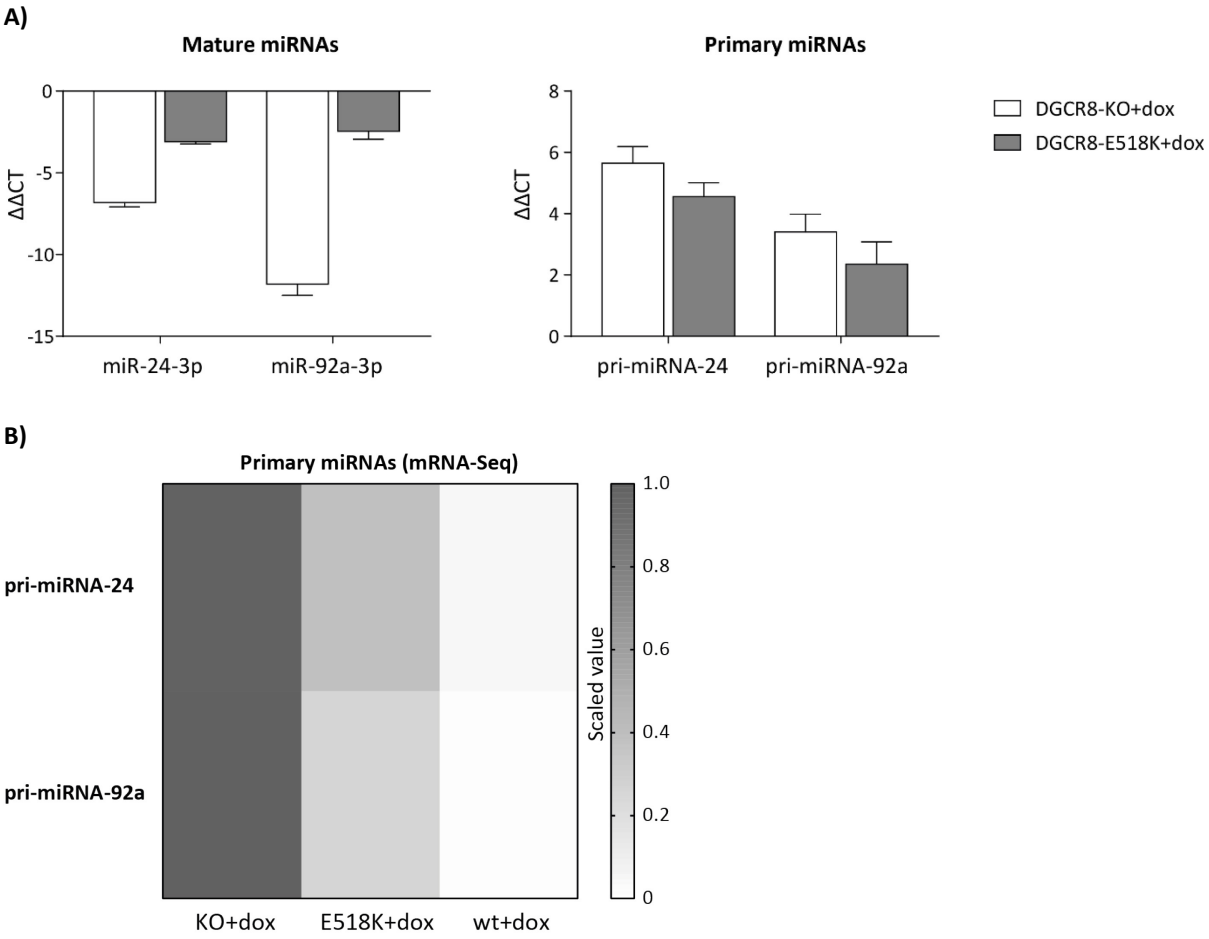


Figure 16: Mature miRNA vs. pri-miRNA expression in DGCR8-wild-type, -E518K and -knockout mESCs
(A) qPCR results showing reduced expression of mature miR-24-3p and miR-92a-3p in DGCR8-KO and -E518K relative to DGCR8-wt. By contrast, pri-miRNA expression was reduced in DGCR8-wt and increased in DGCR8-KO and -E518K. Bar chart represents delta delta CT values of DGCR8-KO and DGCR8-E518K relative to DGCR8-wt. Error bars indicate standard deviation of three biological replicates.
Mature miRNA and pri-miRNA expression were normalized to miR-320-3p or to the housekeeping gene *Hprt*, respectively.
(B) Heatmap confirming higher expression of pri-miRNA-24 and pri-miRNA-92a in DGCR8-KO and -E518K compared to DGCR8-wild-type. Values represent scaled RNA expression levels from mRNA-Seq data, ranging from 0–1. Two biological replicates were used per condition. Pri-miRNAs were defined as pre-miRNAs plus 1kb of upstream and downstream flanking regions.

5.4.3.3 Embryonic stem cell specific ESCC miRNAs

MiRNAs play essential roles in a plethora of biological processes, among others they are important regulators of proliferation and self-renewal of ESCs. Specifically, the antagonism between ESC cell cycle (ESCC) and let-7 miRNAs was shown to control the switch between self-renewal and differentiation in mESCs (Melton, Judson, and Blelloch 2010).

Expression levels of ESCC- and let-7 miRNAs were compared between DGCR8-knockout, -wild-type and -E518K mESCs. Both miRNA families were present in DGCR8-wild-type as well as in the mutant situation. However, there was a difference in the distribution of ESCC- and let-7 miRNAs between DGCR8-wild-type and -E518K cells. ESCC miRNAs showed higher expression levels in wild-type, whereas let-7 miRNAs were stronger expressed in E518K. In DGCR8-knockout cells, miRNAs of both groups were absent (Figure 17). The difference in these opposing miRNA families might have an impact on proliferation and differentiation behavior in DGCR8-wild-type and -E518K mutant cells.

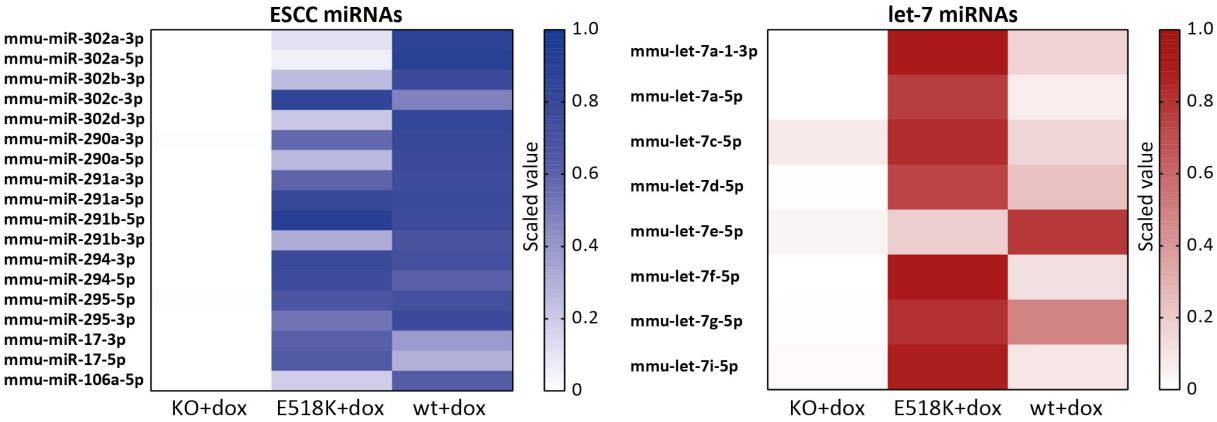


Figure 17: ESC cell cycle (ESCC) miRNAs vs. let-7 miRNAs

Heatmaps showing expression levels of ESCC- and let-7 miRNAs in DGCR8-knockout, -E518K and -wild-type mESCs. DGCR8-wt mESCs showed stronger expression of ESCC miRNAs, whereas let-7 miRNAs were higher expressed in E518K mESCs. Values represent mean scaled expression levels of three biological replicates (except E518K–only two samples).

5.4.3.4 Length distribution

To test whether the DGCR8 E518K mutation would lead to aberrant processing of miRNAs resulting in different length distribution compared to the wild-type situation, soft-clipped reads were analyzed using miRPro (Shi et al. 2015). MiRNA reads with a maximum of 10 bases exceeding their canonical reference were counted and were compared between DGCR8-wt, -E518K and -KO. Therefore, percentages of 5'/3'soft-clips in 5p- and 3p- arm miRNAs among total read mappings were calculated for each condition (Figure 18). DGCR8-wild-type and -E518K showed similar distribution of 3'soft-clips with approximately 16% on 5p-arm, and 7% on 3p-arm miRNAs. Both wild-type and E518K overall exhibited low percentages of 5'soft-clips. By contrast, miRNAs in DGCR8-knockout particularly harbored soft-clips on 3p-arm miRNAs with about 14% on the 5' or 3'end, and 22% on both ends.

Since DGCR8-wild-type and -E518K did not significantly differ in their distribution of soft-clips, there are no indications of altered miRNA length due to aberrant miRNA processing by DGCR8 E518K.

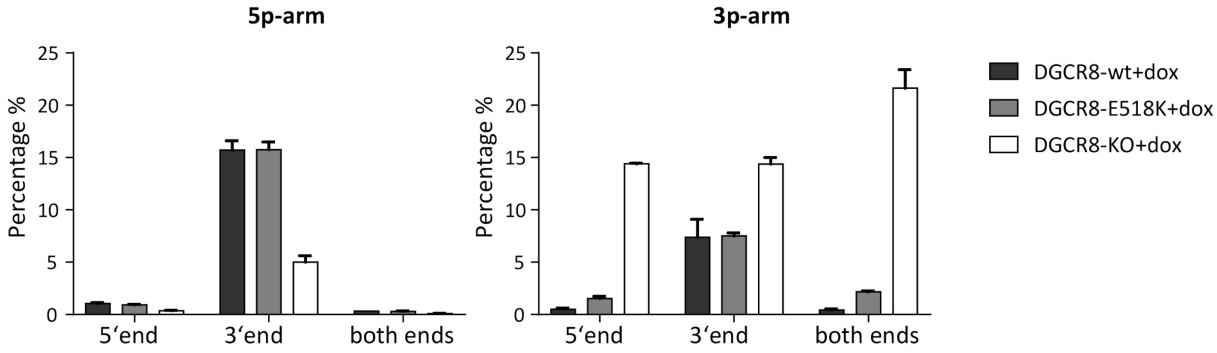


Figure 18: Soft-clipping of 5p-and 3p-arm miRNAs in DGCR8-wild-type, -E518K, and -knockout mESCs
Bar charts showing percentages of soft-clips in read mapping at the 5’- and 3’end, or both ends of either 5p- or 3p-arm miRNAs. Error bars indicate SD of three biological replicates.

5.4.3.5 Base composition of the seed sequence

The seed sequence of miRNAs (nucleotides 2–8 at the 5’end) plays an essential role for miRNAs binding to the 3’UTR of their target mRNAs to regulate gene expression (Lewis, Burge, and Bartel 2005). Base composition of the seed region can alter specificity and stability of miRNA-target binding (Wang 2013). Alterations in the seed sequence could thus have an impact on the functionality of miRNA-mediated gene expression regulation. Also, bases beyond the seed sequence, particularly at positions 13–16, were reported to impact miRNA-target interaction, when base pairing in the seed sequence is suboptimal (Broughton et al. 2016; Grimson et al. 2007). To check whether the seed sequence and beyond of mature miRNAs higher expressed in DGCR8-wild-type or -E518K mutant exhibited differences in base composition, nucleotides at the positions 2–8 and 13–16 were analyzed (Figure 19 and 20).

Overall, differentially expressed miRNAs exhibited a similar base distribution in DGCR8-wild-type and -E518K. In contrast to the seed sequence of miRNAs higher expressed in the mutant (cluster 2), however, miRNAs with higher expression in DGCR8-wild-type (cluster 1) showed a more evenly distributed base composition. At the positions 4 and 6, ratio of A residue was elevated which could imply high target specificity. Nucleotide ratios in the seed sequence of miRNAs higher expressed in the mutant, on the other hand, peaked out at various positions. At position 4, ratio of G was highly elevated, whereas at position 6, E518K also showed a high ratio of A as observed in the wild-type (Figure 19). Analysis of the seed without taking coverage into account exhibited a similar profile as when sequences were analyzed based on coverage, but with a rather evenly distributed base composition (Figure S3). This implies that those miRNAs with higher expression account for the respective seed sequence profiles in wild-type and mutant DGCR8. Both wild-type and E518K,

showed similar overall GC distribution, suggesting that thermostability of seed-target binding could be ensured in both conditions.

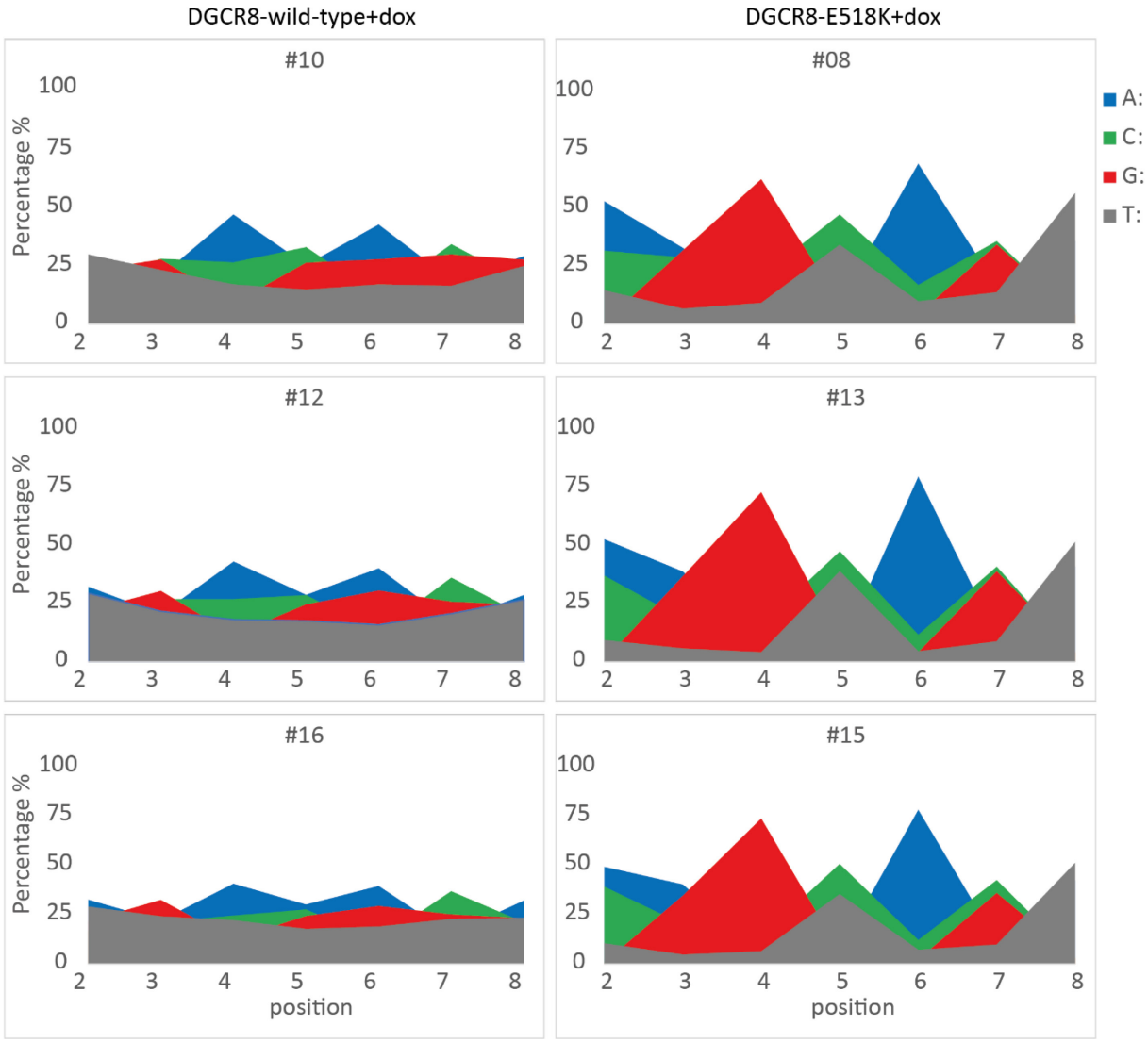


Figure 19: Base composition of the seed sequence in DGCR8-wild-type, and -E518K mESCs
Area charts showing distribution of A, C, G, and T in miRNAs upregulated in DGCR8-wild-type or -E518K mESCs encompassing the seed sequence (position 2–8 at the 5'end). Percentages were calculated based on read counts per nucleotide position. Three biological replicates are shown for each condition.

Base composition beyond the seed sequence was evenly distributed in DGCR8-wild-type with increased ratios of A at the positions 13 and 14. At position 14, however, G content was similarly high as the A ratio. MiRNAs higher expressed in E518K likewise showed increased ratio of G at position 14. Opposing to the wild-type situation, however, E518K exhibited highly elevated A ratios and low G content towards the 3'end (Figure 20).

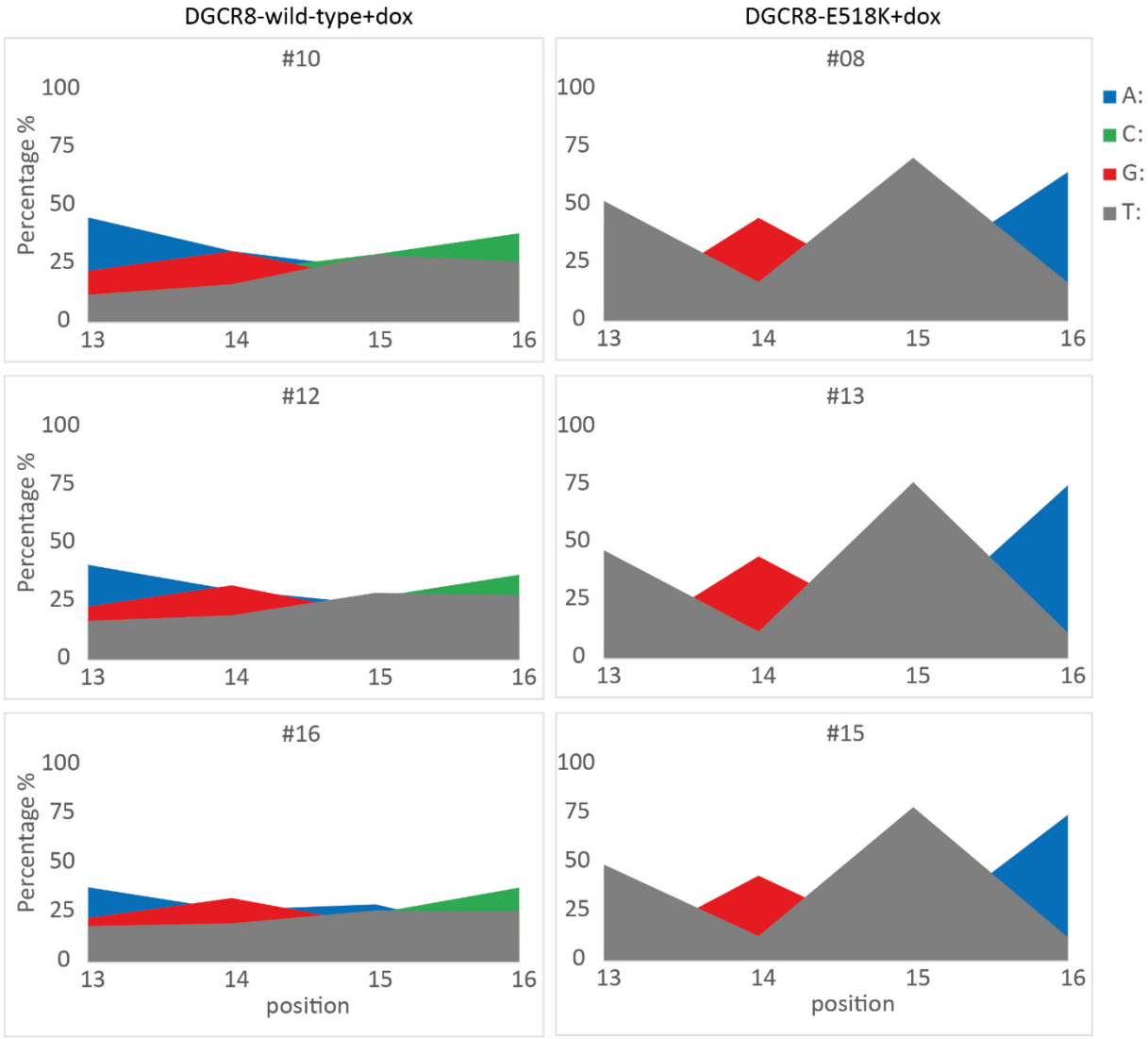


Figure 20: Base composition beyond the seed sequence in DGCR8-wild-type, and -E518K mESCs
Area charts showing distribution of A, C, G, and T in miRNAs upregulated in DGCR8-wild-type or -E518K mESCs encompassing the positions 13–16 at the 3'end Percentages were calculated based on read counts per nucleotide position. Three biological replicates are shown for each condition.

These base composition profiles of DGCR8-wild-type and -E518K miRNAs could be observed in all three respective biological replicates. The differences appreciated in the base distribution of the seed sequence and beyond could possibly lead to altered miRNA-target binding in DGCR8-E518K, which will be further analyzed in the following sections.

5.4.4 mRNA-Sequence analysis of DGCR8-knockout, -wild-type and -E518K mESCs

To test whether altered miRNA processing observed in E518K mutated mESCs would lead to changes in mRNA expression, RNA-Seq analysis was performed. For each condition, two biological replicates were used. Average-linkage clustering using the Euclidean distance was performed to check if the mRNA expression values represent the different groups. Biological replicates of each condition nicely clustered together, outliers could be excluded (Figure 21).

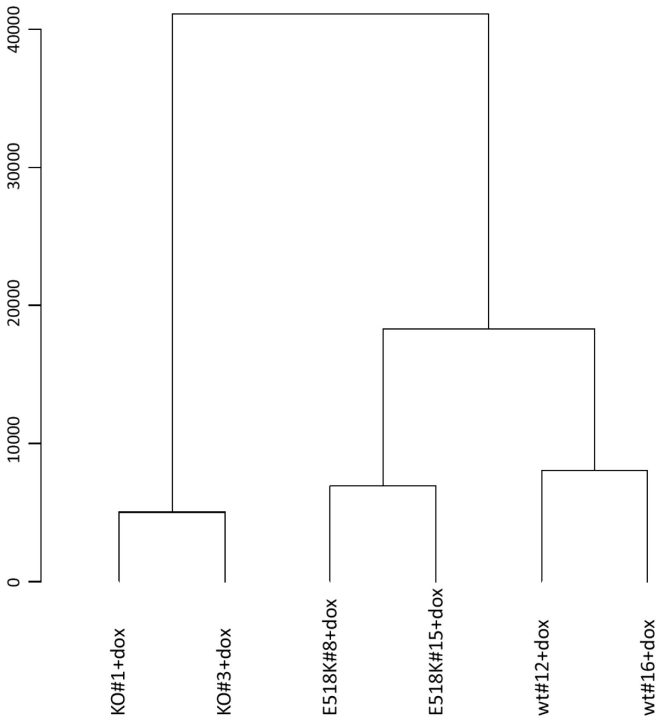


Figure 21: Cluster analysis of DGCR8-knockout, -E518K and -wild-type mESCs
Euclidean distance-based hierarchical clustering of all RNA-Seq samples consisting of two biological replicates per condition. Y-axis represents the distance between the clusters.

For each gene, expression values were calculated as transcripts per million (TPM). Only transcripts that had a TPM≥1 in one of the three groups, were considered for further analysis (N=12079/21964). Normalized read counts were scaled to values ranging from 0–1 among all samples and visualized in a heatmap (Figure 22). Genes with scaled expression values ≥0.50 were defined as highly expressed, whereas genes with values ≤0.25 were considered as weakly expressed. Upon scaling, five different clusters (A–F) emerged comprising 7066 differentially expressed genes in total.

Comparing DGCR8-wild-type and DGCR8-E518K with the knockout situation, wild-type and E518K together showed 3561 differentially expressed genes (1488 down- and 2073 upregulated genes) represented in clusters C and D. Additionally, 1323 genes were exclusively higher expressed in wild-type, whereas 184 genes showed higher expression in E518K (clusters A and B). Furthermore, DGCR8-wild-type and -E518K respectively shared about 1000 genes with DGCR8-knockout that showed higher expression (clusters E and F).

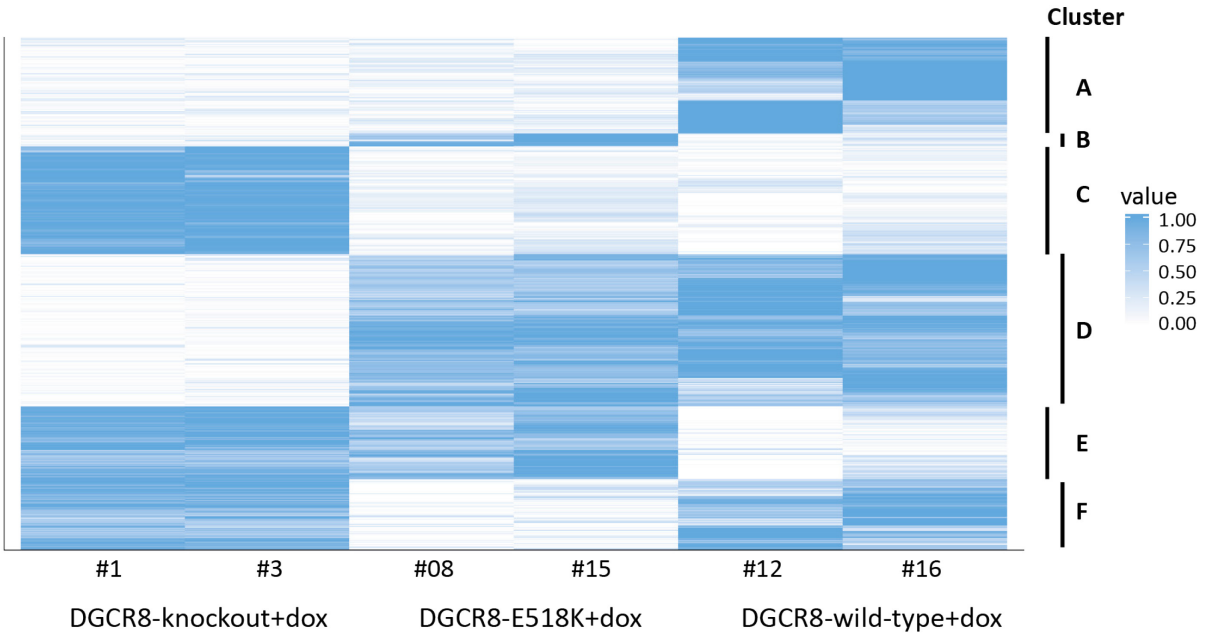


Figure 22: Differences in mRNA expression between DGCR8-knockout, -wild-type and -E518K mESCs
Heatmap showing clusters of differentially expressed mRNAs in DGCR8-knockout, -wild-type and -E518K with a TPM≥1. Values represent scaled mRNA expression ranging from 0–1. Cluster A: KO≤0.25, E518K≤0.25, wt≥0.50 (N=1323); Cluster B: KO≤0.25, E518K≥0.50, wt≤0.25 (N=184); Cluster C: KO≥0.50, E518K≤0.25, wt≤0.25 (N=1488), Cluster D: KO≤0.25, E518K≥0.50, wt≥0.50 (N=2073); Cluster E: KO≥0.50, E518K≥0.50, wt≤0.25 (N=997); Cluster F: KO≥0.50, E518K≤0.25, wt≥0.50 (N=1001).

5.4.5 Integrative analysis of miRNA and mRNA expression

In order to better understand the relationship between differentially expressed miRNAs and altered mRNA expression, miRNA target interactions (MTIs) were analyzed using the prediction tools TargetScan and miRanda (John et al. 2004; Lewis et al. 2003). Target genes of the different clusters A–F (Figure 22) that showed the opposite expression pattern to their miRNAs were selected for further analysis. This resulted in 132,994 MTIs in total (5413 genes, 291 miRNAs). The number of MTIs for each cluster are shown in Table 21. Clusters A and B were expected to lack target-binding sites serving as negative controls. Downregulation of genes in both DGCR8-knockout and either DGCR8-wild-type or -E518K was likely to be related to common biological variations in cell culture, and not due to miRNA-specific target regulation. Indeed, cluster B did not show any MTIs confirming that DGCR8-knockout and DGCR8-wild-type did not share miRNAs, which could already be observed in the miRNA-Seq analysis. On the other hand, knockout and E518K did share common MTIs (919 genes, 5 miRNAs) in cluster A. However, the small number of shared miRNAs and the fact that out of the 919 miRNA-mRNA-pairs only two were annotated as experimentally validated MTIs in miRTarBase (Hsu et al. 2011), implied that this group of mRNAs was unlikely to be targeted by specific miRNAs. In contrast, E518K shared 42,199 MTIs (1,235 genes, 67 miRNAs) with DGCR8-wild-type confirming the mutation's ability to partially rescue the knockout phenotype. As already seen in the miRNA-Seq analysis, DGCR8-wild-type led to a rescue of the miRNA processing defect, expressing

a large number of miRNAs that were preferably processed by DGCR8-wild-type. This was reflected in the nearly 57,000 MTIs that were found in the wild-type situation (cluster E). MTIs detected in cluster D and F, on the other hand, seemed to be more specific for knockout and E518K, respectively. Almost one third of the miRNAs (7/23) that were responsible for target regulation in E518K, belonged to the let-7-family. MRNAs found to be downregulated in the knockout situation were particularly targeted by non-canonical miRNAs whose expression was highest in DGCR8-KO as a result of its deficiency in processing canonical miRNAs.

Table 21: Overview of miRNA target interactions (MTIs) in the different clusters

Cluster	# of MTIs	# of mRNAs in MTIs	# of miRNAs in MTIs
A	2,694	919	5
B	none	n.a.	n.a.
C	42,199	1,235	67
D	22,384	1,650	34
E	56,791	823	162
F	8,926	786	23

In order to investigate pathways that were potentially affected by miRNA-regulated genes, overrepresentation analysis (ORA) was performed using the GeneTrail2 webserver (Stockel et al. 2016). Biological processes that were significantly enriched in the different clusters are illustrated in Figure 23.

MiRNAs that were upregulated in both DGCR8-wild-type and -E518K potentially target mRNAs of cluster C, which play essential roles in the regulation of the cytoskeleton, cell division, apoptotic signaling, and catabolic processes. At least for this group of MTIs, the E518K mutant was likewise able to regulate putative target genes important for normal stem cell function. MiRNAs that showed high expression in DGCR8-wild-type additionally regulated genes related to RNA-processing, including mRNA- and non-coding RNA processing, RNA-splicing, and ribonucleoprotein complex biogenesis (cluster E). Interestingly, cluster D also contained genes related to the same biological processes, however consisting of different gene sets than in cluster E. RNA-processing-related genes in cluster D that were absent or downregulated namely included DGCR8 and DROSHA, which was to be expected for DGCR8-KO mESCs. Both genes of the microprocessor complex are known to have additional non-canonical functions, such as the interaction with the spliceosome (Kataoka, Fujita, and Ohno 2009). Indeed, components of the spliceosome were partly downregulated in the knockout situation, implying secondary effects due to the depletion of DGCR8. In addition, genes related to cell cycle regulation including cyclin-dependent kinases (CDKs) and cell division cycle (CDC) genes, were significantly downregulated in the knockout. MTIs in E518K of cluster F play a role in protein ubiquitination, kinase activity, and metabolic processes. Ubiquitination was shown to affect self-

renewal in embryonic stem cells (Wang, Bu, and Zhang 2019). Also, kinases are known to impact a variety of cellular processes such as proliferation and differentiation, making it worthwhile to functionally test whether these regulated genes indeed lead to respective phenotypical changes in the mutant. Genes downregulated in cluster A were predicted to be involved in the regulation of apoptotic signaling pathway, cell morphogenesis, and metabolic processes. Initially considered as a negative control not containing target-binding sites, cluster A represents common mRNAs between knockout and E518K which are potentially downregulated by shared non-canonical miRNAs.

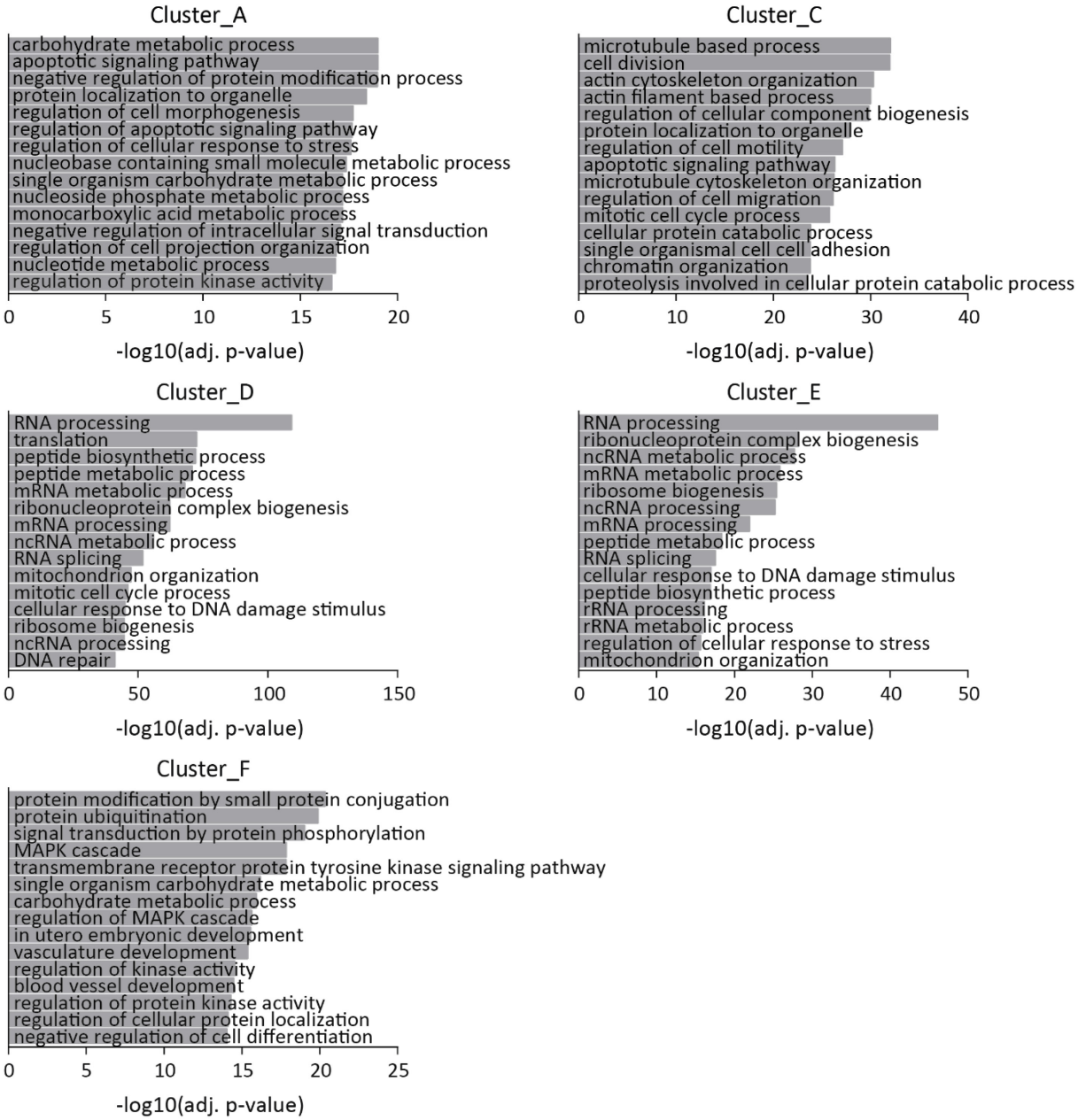


Figure 23: Overrepresentation analysis (ORA) of potential target genes for differentially expressed miRNAs
Bar chart showing significantly enriched biological processes in the clusters A–F. Target genes were analyzed using Genetrail2 webserver (Stockel et al. 2016). P-values were adjusted using the Benjamini-Hochberg adjustment method (Hochberg and Benjamini 1990). Significance is shown as -log10(adj.-p-value). Values ≥ 1.3 were considered as significant.

To get a deeper insight into cellular processes, epithelial-to-mesenchymal transition (EMT) markers were analyzed. It was shown that DGCR8 knockout mESCs tend to develop mesenchymal-morphology during passaging under feeder-free condition, suggesting mESCs undergo EMT upon loss of miRNAs. Particularly, ESCC miRNAs and let-7 miRNAs play essential roles in the regulation of EMT and can therefore have an impact on self-renewal in mESCs (Guo et al. 2015). Since DGCR8-wild-type and -E518K mESCs exhibited different distributions of ESCC- and let-7 miRNAs, it was worthwhile to investigate whether these differences would lead to regulatory changes in EMT. Expression of representative epithelial and mesenchymal genes (Guo et al. 2015) was therefore analyzed (Figure 24). DGCR8-KO mESCs exhibited an upregulation of mesenchymal markers, whereas expression levels were reduced in DGCR8-wt and DGCR8-E518K. Epithelial markers were the highest expressed in DGCR8-KO and DGCR8-wt mESCs. DGCR8-E518K exhibited a distinct pattern with reduced levels of epithelial gene expression. These results suggest that EMT can be blocked in the mutant, but not to the full extent as in the wild-type situation.

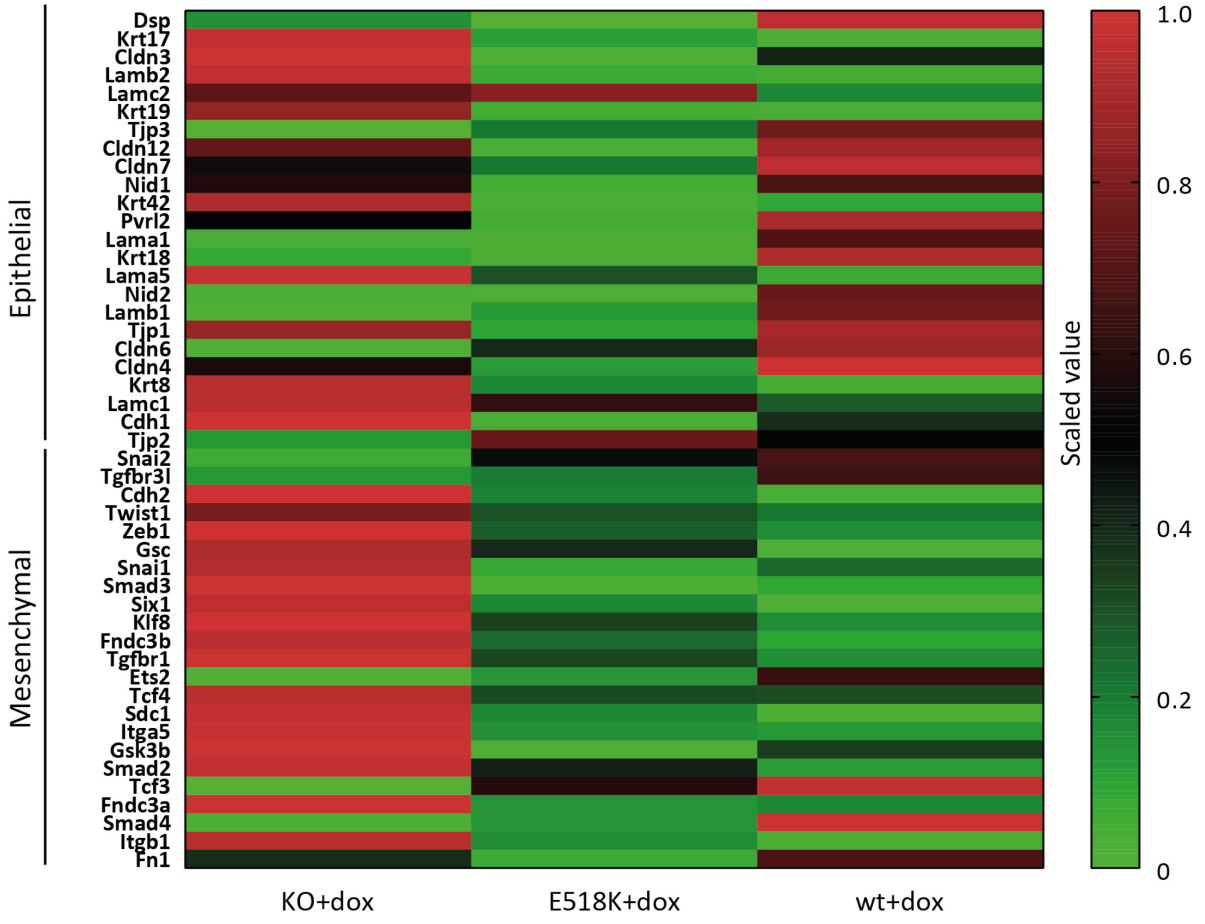


Figure 24: Suppression of epithelial mesenchymal transition in DGCR8-wild-type and -E518K mESCs
Heatmap showing changes in mRNA expression of epithelial and mesenchymal markers (Guo et al. 2015) between DGCR8-KO, -E518K, and -wt mESCs. Values represent scaled mRNA expression ranging from 0–1 with a TPM≥1. Two biological replicates were used per condition.

Also, cell cycle progression is antagonistically regulated by ESCC- and let-7 miRNAs. It was previously shown that ESCC miRNAs were able to rescue the proliferation defect in DGCR8 knockout mESCs by suppressing several key regulators of G1-S transition, such as *Cdkn1a, Rb1*, *Rbl1*, *Rbl2*, and *Lats2* (Wang et al. 2008). Four out of these five cell cycle inhibitors were downregulated in DGCR8-wild-type and -E518K mESCs, whereas in the knockout situation these genes were upregulated (Figure 25). Concomitantly, cell cycle promoting cyclin-dependent kinases (CDKs) and cell division cycle (CDC) genes, which are important for G1-S transition, were upregulated in DGCR8-wild-type and -E518K mESCs (Figure 26). This implies that cell cycle progression may be restored by both DGCR8 wild-type and DGCR8 E518K.

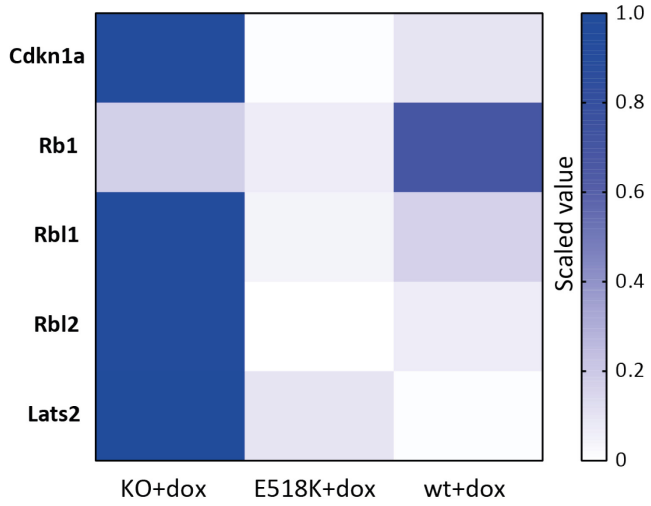


Figure 25: Suppression of cell cycle inhibitors in DGCR8-wild-type and -E518K mESCs
Heatmap showing downregulation of key regulators of the cyclin E-CDK2 regulatory pathway (Wang et al. 2008) in DGCR8-wild-type and DGCR8-E518K. Values represent scaled mRNA expression ranging from 0–1 with a TPM≥1. Two biological replicates were used per condition.

Figure 26: Upregulation of cyclin-dependent kinases (CDKs) and cell division cycle (CDC) genes in DGCR8-wild-type and DGCR8-E518K mESCs
Heatmap showing rescue of cell cycle-specific gene expression in DGCR8-wild-type and DGCR8-E518K. Values represent scaled mRNA expression ranging from 0–1 with a TPM≥1. Two biological replicates were used per condition.

In addition to increased proliferation and cell cycle progression, enhanced glycolysis is another important feature of pluripotent stem cells. Enhanced glycolysis is thought to be essential for the induction and maintenance of pluripotency (Kondoh et al. 2007; Varum et al. 2011). ESCC miRNAs were shown to stimulate glycolysis in mESCs through upregulating the expression of metabolic enzymes, such as *Ldha* and *Pkm2* (Cao et al. 2015). Our RNA-Seq analysis could confirm the downregulation of glycolytic genes in DGCR8-KO and revealed reduced expression levels in DGCR8-E518K compared to DGCR8-wt mESCs, suggestive of decreased glycolytic activity (Figure 27).

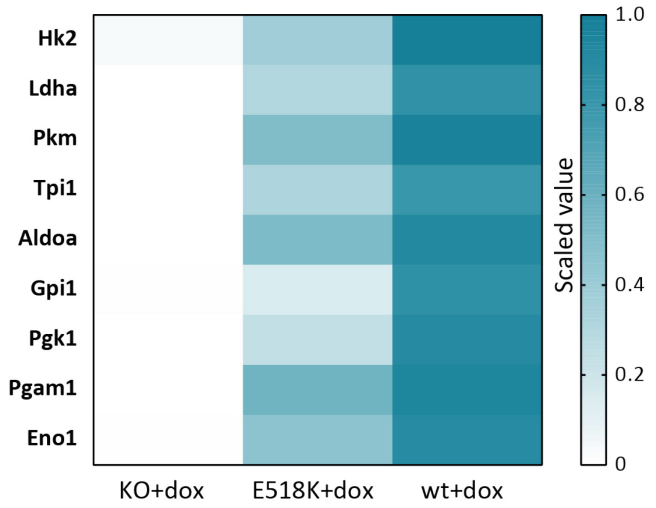


Figure 27: Downregulation of glycolytic enzymes in DGCR8-KO and DGCR8-E518K mESCs
Heatmap showing downregulation of typical glycolytic enzymes (Cao et al. 2015) in DGCR8-KO and DGCR8-E518K mESCs. Values represent scaled mRNA expression ranging from 0–1 with a TPM≥1. Two biological replicates were used per condition.

5.5 Effect of DGCR8 E518K mutation on proliferation in mESCs

Since the DGCR8 E518K mutation could only partially rescue the microprocessor defect in DGCR8-knockout mESCs, it was worthwhile to investigate whether the partial rescue had an impact on the proliferation behavior in mESCs. DGCR8-KO cells were used as a negative control, since it was previously reported that lack of miRNAs leads to severe proliferation defects in mESCs (Wang et al. 2007). For this purpose, MTT and crystal violet assays were performed as indirect methods to determine proliferation differences between DGCR8-wild-type and -E518K mESCs. In addition, cell cycle profiles of DGCR8-wild-type and -E518K mESCs were compared using FACS analysis. Furthermore, S-phase progression was monitored by quantification of BrdU incorporation.

5.5.1 MTT Assay

MTT assay showed a significantly decreased overall growth rate in DGCR8-knockout mESCs compared to DGCR8-wild-type and -E518K cells (adj. p-value: ≤ 0.0001). Overexpression of DGCR8-wild-type could rescue the proliferation defect and showed significantly increased growth rates on day 3 and 4 relative to DGCR8-KO (adj. p-value: ≤ 0.0001) and E518K mutant cells (adj.p-value: ≤ 0.001, and

≤ 0.0001). Despite the partial rescue of the knockout phenotype, E518K cells could not reach wild-type levels (Figure 28). Of note, DGCR8-wild-type without doxycycline treatment showed similar growth pattern as upon induction, indicating that leaky expression might be sufficient to restore the proliferation defect. E518K mutant cells without doxycycline induction, on the other hand, showed a similar proliferation defect as DGCR8-KO cells. 72h after treatment with doxycycline a slight increase in the absorbance rate could be appreciated in E518K. At 96h posttreatment, mutant cells exhibited a significantly higher rate compared to the uninduced situation (adj. p-value: ≤ 0.001) (Figure S4).

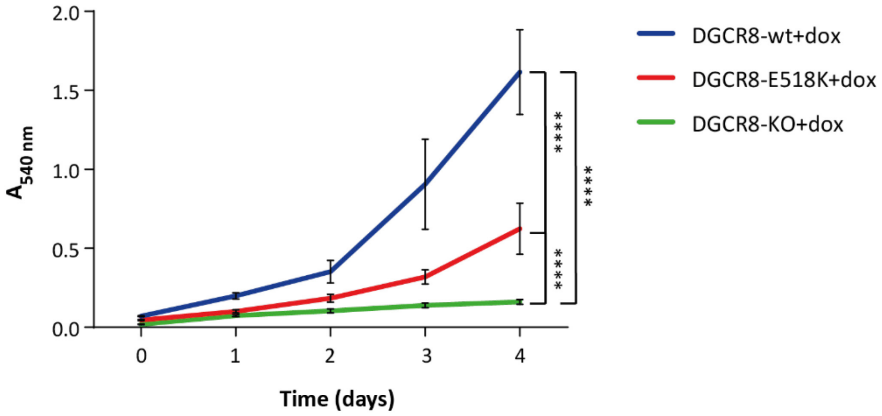


Figure 28: Effect of the DGCR8 E518K mutation on growth in mESCs
Mean growth rate of three independent biological replicates measured by MTT assay. Error bars represent standard deviation (SD) within each set of biological replicates. 2way ANOVA was performed to identify significant differences between DGCR8-wild-type, -knockout and E518K upon induction with doxycycline (adj. p-value: ≤ 0.0001 ****). Significance is shown here for day 4. For each sample, five technical replicates were used. Treatment with doxycycline (300 ng/mL) was started one day after plating. MTT assay was performed for each timepoint (day 0, 1, 2, 3, 4). Absorbance of purple formazan solution was measured at 540 nm.

5.5.2 Crystal Violet Assay

The proliferation differences between DGCR8-wild-type and -E518K observed in the MTT assay were confirmed by crystal violet assay (Figure 29). On day 3 and 4, E518K mutant cells showed significantly decreased growth rates compared to DGCR8-wild-type (adj.p-value: ≤ 0.0001). As already appreciated by MTT assay, doxycycline treatment of DGCR8-wild-type cells led to similar increase in the absorbance rate as the uninduced situation. E518K mutant cells, on the other hand, showed an increase in their growth rate 72h and 96h after doxycycline treatment, whereas without induction, growth rate of E518K cells was comparable to that of DGCR8-KO cells (Figure S5).

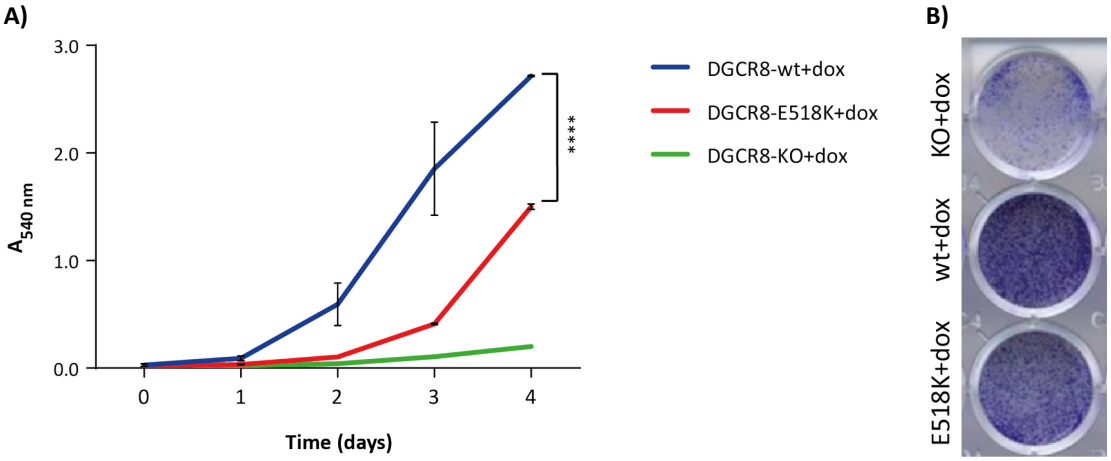


Figure 29: Effect of the DGCR8 E518K mutation on growth in mESCs

(A) Mean growth rate of two independent biological replicates measured by crystal violet assay (except DGCR8-KO–only one sample). Error bars represent SD within each set of biological replicates. 2way ANOVA was performed to identify significant differences between DGCR8-wild-type and -E518K upon induction with doxycycline (adj. p-value: ≤ 0.0001 ****). Significance is shown here for day 4. For each sample three technical replicates were used. Treatment with doxycycline (300 ng/mL) was started one day after plating. Crystal violet assay was performed for each timepoint (day 0, 1, 2, 3, 4). Absorbance was measured at 540 nm.

(B) Image of crystal violet stained cells on day 4. DGCR8-knockout, -wild-type and -E518K treated with doxycycline.

5.5.3 Cell cycle analysis

To test whether the observed differences in proliferation between DGCR8-wild-type and DGCR8-E518K cells are also displayed in different cell cycle distributions, cell cycle analysis of propidium iodide (PI) stained mESCs followed by flow cytometry was performed. In addition, BrdU assay was performed to measure S-phase progression.

DGCR8-knockout mESCs had significantly increased numbers of cells accumulating in the G1 phase from approximately 24% in DGCR8-wild-type and -E518K mutant cells to about 35% (adj. p-value: ≤ 0.05 and ≤ 0.01) (Figure 30A). Concomitantly, they showed a significant decrease (adj. p-value: ≤ 0.001, and ≤ 0.01) of cells in S-phase compared to wild-type and mutant. Based on the decreased proliferation rate of the E518K compared to DGCR8-wild-type shown via MTT and crystal violet assay (Figure 28 and 29), it was assumed that this would be reflected in blocked G1-S transition. Mutant cells did show decreased cell numbers in S-phase, however, this difference was statistically not significant. Consistent with the results obtained from the MTT and crystal violet assays, induction of DGCR8-wild-type led to similar cell cycle profiles as under uninduced condition, implying leaky expression to be sufficient to rescue the G1 phenotype. E518K mutant cells, on the other hand, showed significant decrease of cells in G1 phase (adj. p-value: ≤ 0.0001) and increase in S-phase (adj. p-value: ≤ 0.001) upon induction, whereas doxycycline treatment did not affect cell cycle distribution in DGCR8-knockout cells (Figure 30B and C).

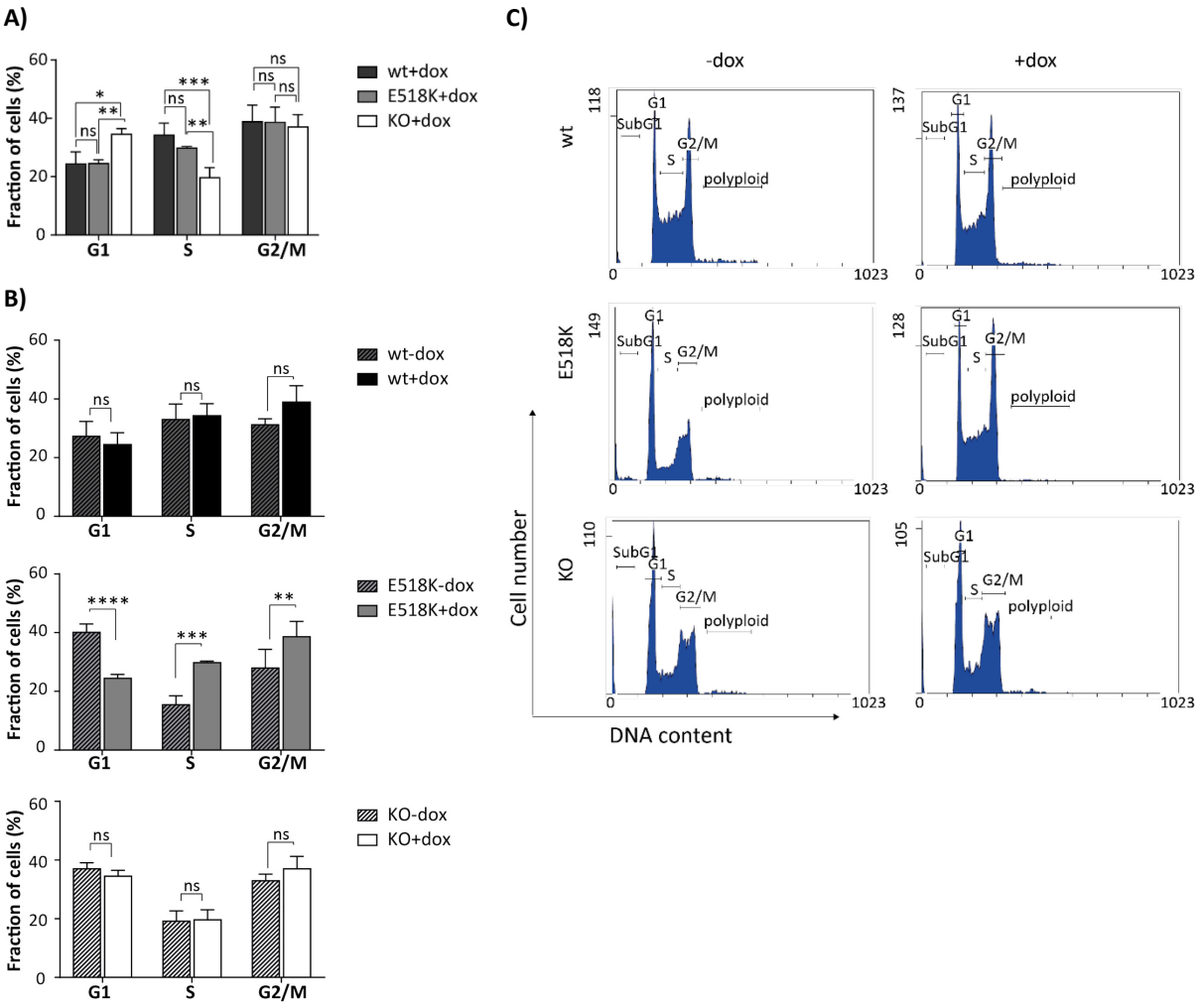


Figure 30: Cell cycle analysis of DGCR8-wild-type, -E518K and -knockout mESCs

Cells were harvested 48h after doxycycline induction (300 ng/mL). DNA was stained with propidium iodide (PI) following FACS analysis. Doublets were excluded by selective gating.

(A) Bar chart presenting fraction of cells (%) in G1, S and G2/M phase. Results are shown as the mean (SD) of three biological replicates. 2way ANOVA was performed to identify significant differences among DGCR8-wt, -E518K and -KO upon induction with doxycycline (adj. p-value ≤ 0.05 *; ≤ 0.01 **; ≤ 0.001 ***; > 0.05 ns).

(B) Bar chart presenting fraction of cells (%) in G1, S and G2/M phase. Results are shown as the mean (SD) of three biological replicates. 2way ANOVA was performed to identify significant differences between induced and uninduced conditions (adj. p-value: ≤ 0.01 **; ≤ 0.001 ***; ≤ 0.0001 ****; > 0.05 ns).

(C) Representative cell cycle profiles of DGCR8-wild-type, -E518K and -knockout ± doxycycline.

S-phase progression tested by BrdU assay led to similar results as in the cell cycle analysis of PI-stained mESCs. DGCR8-knockout cells displayed decreased BrdU incorporation compared to DGCR8-wild-type and -E518K mutant cells, confirming the proliferation defect of miRNA-deficient cells. S-phase progression was not significantly decreased in the mutant situation compared to DGCR8-wild-type as already observed by cell cycle analysis (Figure 31).

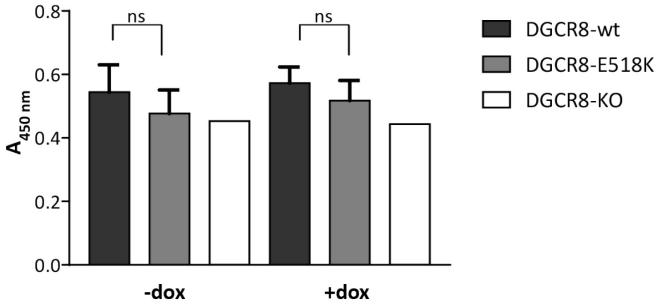


Figure 31: BrdU incorporation to monitor S-phase progression

Cells were incubated with BrdU for 6h after 48h induction with doxycycline (300 ng/mL). BrdU incorporation was colorimetrically determined using the Cell Proliferation ELISA BrdU kit (Roche).

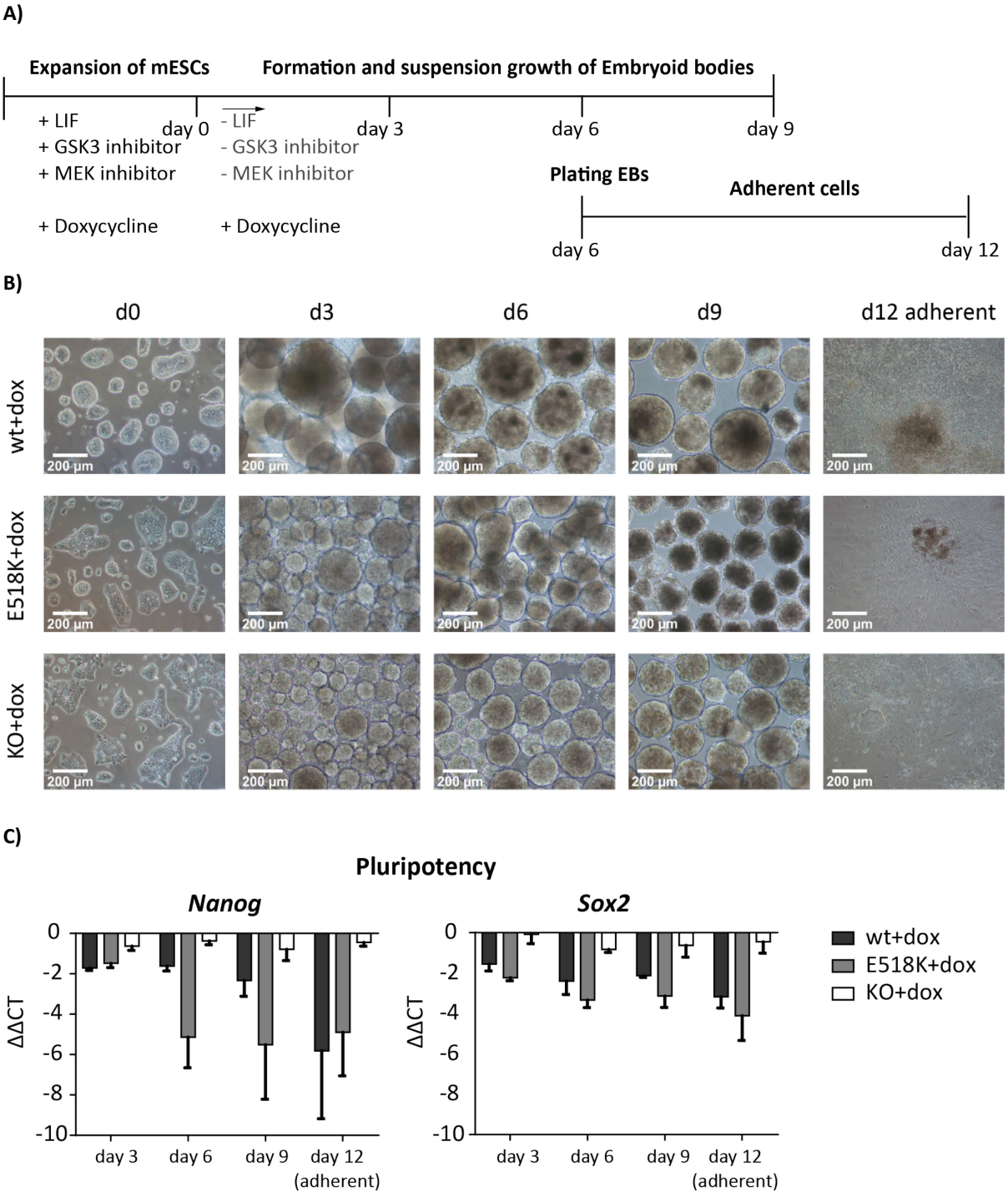
Bar chart showing mean absorbance rate of DGCR8-wild-type, -E518K and -knockout cells upon induction with doxycycline. Error bars indicate SD of three biological replicates (except DGCR8-KO–only one sample). 1way ANOVA was performed to identify significant differences between DGCR8-wt and -E518K (adj. p-value: > 0.05 ns). Absorbance was measured at 450 nm.

5.6 Embryoid body differentiation

To assess the differentiation capacity of DGCR8-E518K mESCs, they were cultured as embryoid bodies (EBs) in the absence of leukemia inhibitory factor (LIF) and GSK3 and MEK inhibitors for 9 days. In addition, on day 6, EBs were plated on cell culture dishes and grown for further 6 days (Figure 32A). Already on day 3, morphological differences between DGCR8-wild-type, -E518K and -knockout EBs could be appreciated. DGCR8-wild-type EBs were bigger in size, whereas knockout cells formed small and light-colored EBs. E518K mutant EBs did not reach the size of DGCR8-wild-type EBs either, but unlike DGCR8-knockout, mutant cells were likewise able to form cystic EBs as wild-type EBs throughout the differentiation process. After 6 days of plating, adherent DGCR8-wild-type and -E518K EBs showed a heterogeneous mix of differentiated cells, whereas DGCR8-knockout cells still formed stem cell-like colonies in addition to differentiated cells (Figure 32B).

To test whether these morphological differences could be observed in altered expression of pluripotency- and differentiation-related genes, qPCR analysis was performed (Figure 32C). DGCR8-knockout mESCs were unable to downregulate pluripotency markers during EB differentiation. However, they did express some differentiation markers, albeit not to wild-type nor E518K level. Both DGCR8-wild-type and -E518K cells, on the other hand, were able to differentiate into all three germ layers. DGCR8-wt and -E518K EBs exhibited increased expression levels of differentiation markers, while simultaneously downregulating the pluripotency markers *Nanog*, *Sox2* and *Oct-4*. E518K EBs showed a stronger downregulation of pluripotency markers than DGCR8-wild-type EBs, similar or higher expression levels of mesodermal (*Flt1*, *Tbx20* and *Brachyury*) and endodermal markers (*Afp*, *Gata4* and *Gata6*). At a later timepoint, however, DGCR8-wild-type EBs could reach comparable levels as in the E518K situation. Expression of ectodermal markers *Nestin* and *Fgf5* was strong in both wild-type and mutant EBs from day 3 on (Figure 32C and S6). The differentiation

defect in DGCR8 knockout cells could be hence rescued by introduction of DGCR8 wild-type as well as DGCR8 E518K.



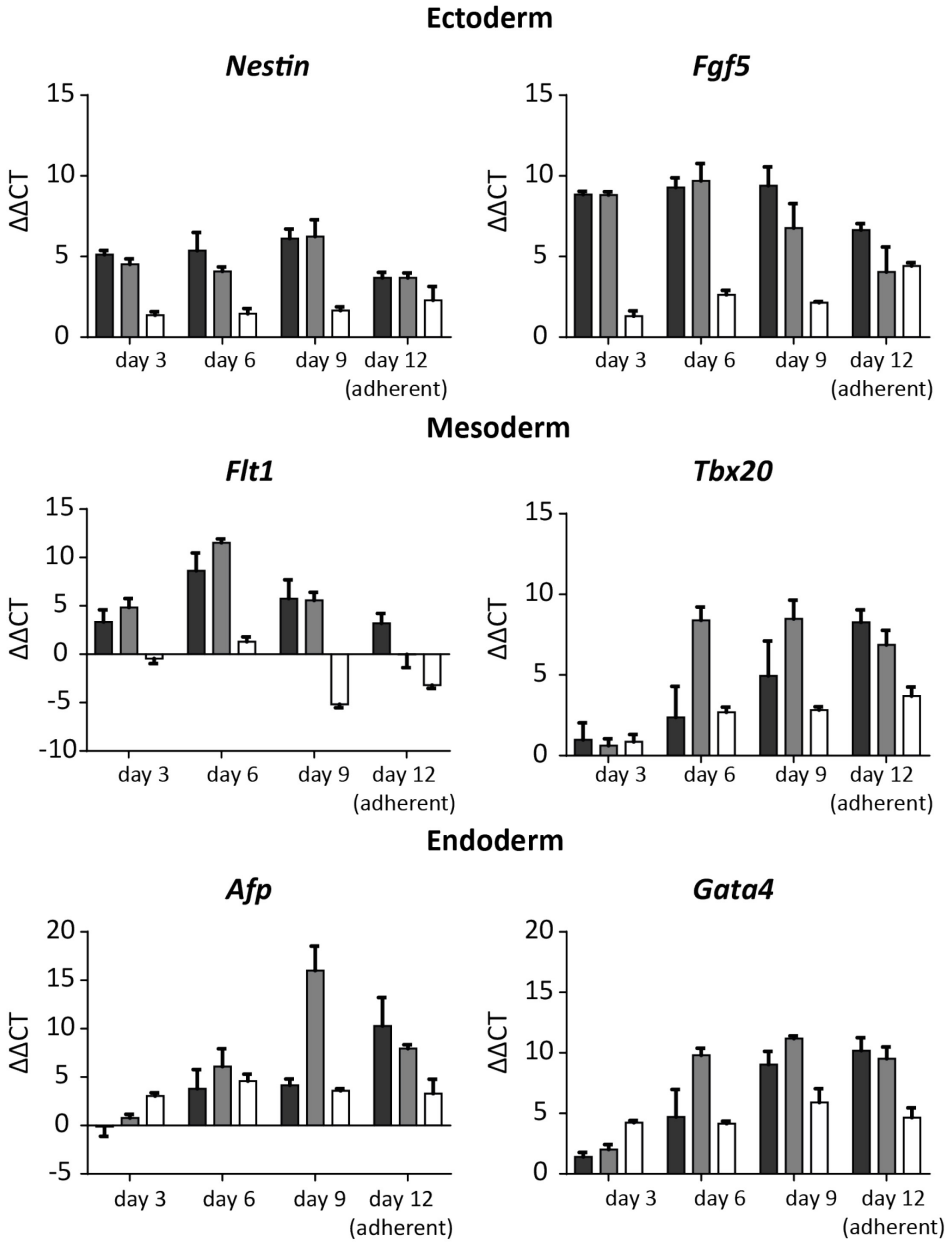


Figure 32: Embryoid body differentiation of DGCR8-wild-type, -E518K, and -knockout mESCs

(A) Schematic overview of the EB differentiation protocol

(B) Representative images of DGCR8-wild-type, -E518K and -knockout EBs.

(C) qPCR results of pluripotency markers *Nanog* and *Sox2* and differentiation markers (ectoderm: *Nestin* and *Fgf5*; mesoderm: *Flt1* and *Tbx20*; endoderm: *Afp* and *Gata4*) on day 3, 6, and 9 of EB differentiation. On day 6, EBs were collected and grown as monolayers for additional 6 days (day 12 adherent). CT-values were normalized to the housekeeping gene *Hprt.* Bars represent mean ΔΔCT-values relative to DGCR8-wild-type mESCs on day 0. Error bars indicate SD of three biological replicates.

5.7 Nuclear localization of DGCR8 and DROSHA

Nuclear localization of DROSHA and its binding partner DGCR8 is pivotal for canonical miRNA processing. It has been previously shown that nuclear localization and/or nuclear retention signal of DGCR8 is restricted to the N-terminal region of DGCR8 (Yeom et al. 2006).

Although this region is located upstream of its dsRBDs, it was interesting to investigate whether the E518K mutation would lead to an abnormal localization pattern in the cell through secondary effects. Therefore, HEK293T cells were transiently co-transfected with either pCS2P-HA2-mcherry-hDGCR8 wild-type or E518K and a pcDNA3.1-YFP-DROSHA-Flag construct. Subsequently, protein signals were visualized using confocal laser scanning microscopy (Figure 33). Aberrant subcellular localization by the E518K mutation could be excluded, as both wild-type and mutant DGCR8 were nicely colocalized with DROSHA wild-type in the nucleus. Co-transfection of DGCR8 wild-type or E518K and DROSHA E1147K was also found to be restricted to the nucleus only (Figure S7). Therefore, effects of both mutations on the subcellular localization of the microprocessor could be excluded.

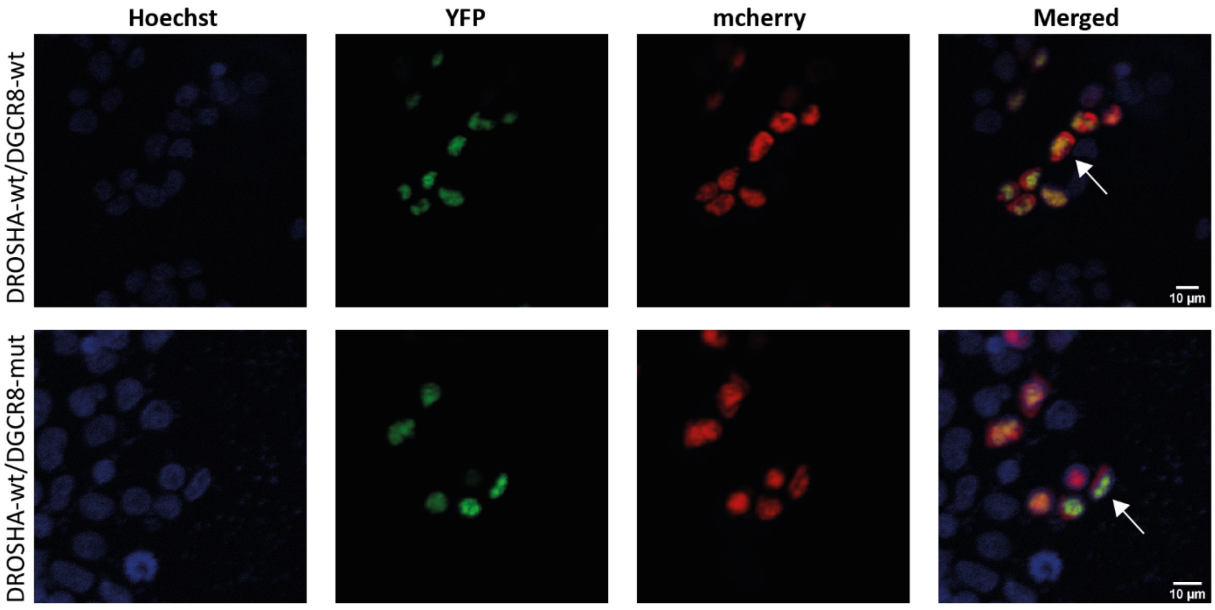


Figure 33: Nuclear localization of the microprocessor complex DROSHA/DGCR8
Transiently co-transfected HEK293T cells with pCS2P-HA2-mcherry-hDGCR8 wild-type (upper row) or E518K (bottom row) and pcDNA3.1-YFP-DROSHA-Flag. Localization of the nucleus was shown using Hoechst staining. DGCR8 wild-type and E518K were found to be both restricted to the nucleus. Arrows indicate nuclear co-localization of DGCR8 and DROSHA.

5.8 Posttranscriptional stabilization of DROSHA by DGCR8

DGCR8 was reported to stabilize DROSHA posttranscriptionally via protein-protein interaction (Han et al. 2009). To test whether the E518K mutant had an impact on the stabilization of DROSHA, Western blot analysis of DGCR8-wild-type and -E518K mESCs was performed (Figure 34A and S8). Downregulation of DROSHA could be confirmed in DGCR8-KO mESCs. Also, DGCR8-wild-type and DGCR8-E518K mutant cells showed similarly weak expression of DROSHA protein under uninduced

condition. Upon induction, however, both DGCR8-wild-type and -E518K exhibited increased DROSHA protein levels implying the E518K mutation had a similar stabilizing effect on DROSHA as wild-type DGCR8. Changes in protein level were not accompanied by altered mRNA expression levels (Figure 34B and C).

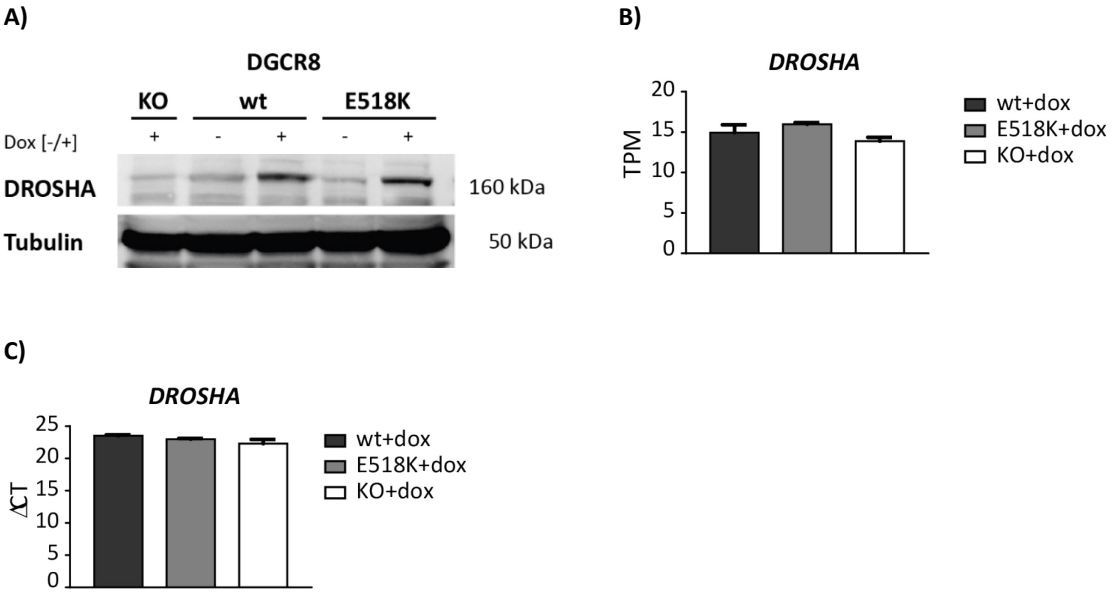


Figure 34: Posttranscriptional stabilization of DROSHA protein by DGCR8 wild-type and E518K

(A) α-DROSHA Western blot analysis of DGCR8-knockout, -wild-type and -E518K mutant mESCs upon induction with doxycycline (300 ng/mL) for 48h.

(B) RNA-seq data showing similar mRNA expression levels of *DROSHA* in DGCR8-wt, -E518K and -KO mESCs. Mean expression TPM values and SD of two biological replicates are shown.

(C) qPCR results validating similar *DROSHA* mRNA expression levels among DGCR8-wt, -E518K and -KO mESCs. ΔCT values represent mean CT-values normalized to the housekeeping gene *Hprt*. Error bars indicate standard deviation of three biological replicates.

6 Discussion

Several studies in the past few years have greatly contributed to a better characterization of the genetic landscape of Wilms tumor and have helped to expand the so far small repertoire of known candidate genes. A clear-cut underlying genetic defect, however, remained elusive. Accumulating evidence suggests a crucial role of the miRNA biogenesis pathway in tumor formation. We and others could identify recurrent hotspot mutations in Wilms tumors affecting the microprocessor genes *DROSHA* and *DGCR8* (Walz et al. 2015; Wegert et al. 2015). The *DROSHA* E1147K and the *DGCR8* E518K mutations were the most frequent and functionally relevant alterations mapping to regions that are important for catalytic activity and RNA-binding, respectively. These mutations are almost exclusively found in WT, implying a specific role in Wilms tumorigenesis.

Despite these substantial findings, the functionality of the detected mutations still needs to be characterized. The present work therefore deals with the *in vitro* functional analysis of the miRNA microprocessor mutations *DROSHA* E1147K and *DGCR8* E518K in HEK293T cells and mouse embryonic stem cells, respectively. Given the complexity and multitude of cellular pathways that miRNAs could affect, the effects of the detected mutations are likely to be complex and heterogeneous, as well. In this work, first the frequency of the *DROSHA* E1147K and *DGCR8* E518K mutation was analyzed in a larger cohort of Wilms tumors to check whether these are indeed frequent events in WT. Subsequently, the impact of these mutations on miRNA expression was investigated. Furthermore, the functional consequences of altered miRNA expression in DGCR8-E518K mESCs on target gene regulation and various biological processes were assessed.

6.1 DROSHA

Evidence has been emerging that normal function of the microprocessor machinery plays an essential role in kidney development, homeostasis, and physiology. Loss or deregulation of its key components have been reported to be implicated in developmental defects of the kidney and renal cancers, including Wilms tumor (Bartram et al. 2015; Chu et al. 2014; Walz et al. 2015; Wegert et al. 2015). For instance, podocyte-specific deletion of Drosha was shown to result in proteinuria and renal failure, ultimately leading to premature death in mice (Zhdanova et al. 2011). The strong phenotype of targeted deletion of Drosha in kidney precursor cells as well as embryonic lethality of Drosha-deficient mice (Chong et al. 2010; Kruber et al. 2019) underscores the importance of functional Drosha in development and other physiological processes.

6.1.1 Recurrent somatic *DROSHA* E1147K mutation in Wilms tumors

Our exome sequence analysis of a cohort of high-risk blastemal Wilms tumor revealed somatic *DROSHA* mutations in domains that are important for normal DROSHA function (Wegert et al. 2015). Based on superposition models, the E993K and the D1204Y mutations involve a putative contact to

the dsRNA substrate, which in turn may lead to altered binding and processing of pri-miRNA molecules (Wegert et al. 2015). Additionally, recently published 3D-superimposition of the microprocessor complex with pri-miRNA revealed that point mutations of key metal-coordinating residues (E1045Q/E1222Q) weaken the interaction of DROSHA with the pri-miRNA stem. It was shown that without properly coordinated divalent ions, the core domains of DROSHA, the RNase III domains and the central domain (CED) move away from the RNA, reducing direct contact with the RNA substrate. Improper docking of the pri-miRNA in turn leads to conformational changes in DROSHA where the helical Belt, a subregion of the CED, can no longer interact with the basal junction and the rest of DROSHA, which ultimately result in decreased efficiency and accuracy of pri-miRNA processing (Partin et al. 2020).

These findings therefore imply that the somatic mutations found through our screening most likely result in altered activity of the microprocessor complex and consequently in altered miRNA expression.

Recurrence of mutations in the catalytic center of the RNase IIIa and b domains found in Wilms tumor, suggest that disrupted *DROSHA* activity might be implicated in Wilms tumorigenesis. For further investigation the most frequent *DROSHA* mutation was selected.

Consistent with other studies where the *DROSHA* E1147K mutation was reported to make up to 80% of all *DROSHA* mutations in WT (Rakheja et al. 2014; Torrezan et al. 2014; Walz et al. 2015), the E1147K missense mutation was likewise the most prevalent alteration in our cohort of high-risk blastemal cases. Our large-scaled mutation screening of unselected WT samples could confirm that *DROSHA* E1147K was indeed a frequent event in Wilms tumor. Importantly, analysis of tumor DNA and RNA exclusively showed a heterozygous expression of the *DROSHA* mutations which strongly implies a dominant-negative effect.

Comparing the occurrence of this alteration among the different histological subtypes showed that in the low/intermediate risk group, tumors with regressive histology, which mainly represents dead blastemal cells, and with focal anaplasia had the highest rate of E1147K mutation. Compared to low/intermediate WTs, the E1147K mutation overall occurred more frequently in tumors with high-risk histology, namely chemotherapy resistant blastemal tumors and those with diffuse anaplasia. The blastemal subtype showed the highest frequency in this group. A preference for blastemal histology could likewise be shown in the study by Walz et. al (2015).

6.1.2 *DROSHA* E1147K leads to a global miRNA processing defect

Multiple human cancers have been reported to exhibit altered miRNA expression profiles, mainly with a global downregulation of mature miRNA expression compared to normal tissues (Lu et al. 2005). In Wilms tumor, the *DROSHA* E1147K mutation was shown to induce changes in miRNA expression profiles, which could also be confirmed by *in vitro* studies in human cell lines where

miRNA expression was significantly reduced compared to the wild-type situation (Rakheja et al. 2014; Torrezan et al. 2014). Comparing miRNA expression of heterozygous *DROSHA* E1147K with heterozygous *DROSHA* null cell lines additionally showed that the effects of heterozygously expressed E1147K was more deleterious on miRNA expression than heterozygous null mutations, confirming the dominant-negative effect of *DROSHA* E1147K (Rakheja et al. 2014).

Microarray analysis of our cohort of unselected Wilms tumors likewise exhibited a significant global downregulation of miRNAs forming a separate cluster from non-mutated tumors. Tumors either harboring the *DROSHA* E1147K or *DGCR8* E518K mutation were intermingled in one cluster, implying that both mutations might have similar effects on miRNA expression (Wegert et al. 2015). Through functional analysis in stable HEK293T DROSHA cell lines, I could show that upon induction of the *DROSHA* E1147K mutation miRNA expression was globally reduced, whereas overexpression of the *DROSHA* wild-type did not alter the miRNA profile. The E1147K mutation thus led to altered miRNA expression when co-expressed with the endogenous wild-type *DROSHA* allele in HEK293T cells, likely compromising its function. In contrast to *DICER1* mutations of the RNase IIIb domain, the *DROSHA* E1147K mutation did not result in 5p/3p skewing of miRNA expression. The alteration in *DROSHA* rather leads to global reduction of miRNAs irrespective from their derivation of 5'- or 3'-strand.

Affecting the key metal binding site of the RNase IIIb domain and therefore the cleavage function of DROSHA, the E1147K mutation may lead to incompletely processed RNA-substrates lacking the structural features of precursor miRNAs. Given that the 2 nt 3'overhang in pre-miRNAs is needed for DICER1 to recognize and cleave the RNA-substrate (Han et al. 2004; Park et al. 2011), dsRNA molecules derived from defective cleavage by mutant DROSHA will most likely, if exported at all, not be recognized by DICER1. This in turn, may affect cleavage efficiency and ultimately lead to decreased amounts of mature miRNAs.

Abrogation of global miRNA processing has been shown to accelerate cellular transformation and promote tumorigenesis (Kumar et al. 2007). In addition, reduced miRNA levels were shown to lead to impaired kidney differentiation and development, consequently enhancing the risk for the formation of renal malignancies (Bartram et al. 2016; Bartram et al. 2015; Chu et al. 2014; Nagalakshmi et al. 2011).

We assumed that reduced miRNA expression caused by the E1147K mutation might lead to altered regulation of genes that are important for the differentiation process of renal precursor cells into the developing kidney. The *DROSHA* E1147K mutation might therefore facilitate the persistence of undifferentiated metanephric blastema in postnatal kidney, which is thought to ultimately give rise to Wilms tumor. Our group could confirm the predicted dominant-negative effect of *DROSHA* E1147K *in vivo*, leading to impaired kidney development, kidney failure and lethality within the first two months of life. These mice exhibited impaired differentiation of nephron progenitor cells which

would be consistent with the formation of embryonal neoplasms that derive from rapidly growing undifferentiated precursor cells. However, the embryonic kidney cortex of these mice showed reduced proliferation and an increase in apoptosis, implying that additional genetic alterations in mitogenic and antiapoptotic pathways may be needed for malignant transformation. Despite the strong phenotype, the E1147K mutation alone did therefore not seem to be sufficient to induce WT formation (Kruber et al. 2019).

6.2 DGCR8

In line with the significance of functional miRNA processing for normal development and homeostasis in most tissues, studies have shown that deletion of the microprocessor subunit DGCR8 is associated with a variety of diseases including congenital kidney and urinary tract anomalies, as well as developmental defects due to aberrant miRNA-processing (Bartram et al. 2013; Sellier et al. 2014). Knockout of *Dgcr8* in the developing renal tubular system was shown to lead to severe hydronephrosis, kidney cysts and progressive renal failure in mice. These mice died within the first two months after birth (Bartram et al. 2015). This severe phenotype was also found in a *Pax8Cre* mediated *Dgcr8* knockout mouse model due to loss of miRNAs in renal epithelial cells. Postnatal kidneys exhibited increased proliferation and apoptosis leading to end-stage renal disease and premature death (Bartram et al. 2016). Interestingly, the renal phenotype was likewise observed in *Pax8Cre* mediated knockout of *Dicer1* (Iervolino et al. 2015) as well as in other *Dicer1* knockout models using different cre-alleles that led to premature depletion of nephron progenitors (Chu et al. 2014; Nagalakshmi et al. 2011). This is in line with the *Six2Cre* mediated metanephric deletion of *Drosha* and activation of *DROSHA* E1147K mutation which led to kidney agenesis and premature death in mice due to apoptotic loss of progenitor cells and impaired differentiation (Kruber et al. 2019).

Conclusively, alterations in miRNA processing genes share many phenotypical similarities underlining the essential role of miRNAs in kidney development and renal disease. *DGCR8* as one of the key microprocessor genes and binding partner of *DROSHA* seems to be indispensable for normal kidney development, and alterations in this gene might play a role in Wilms tumorigenesis with an impact on the renal architecture.

6.2.1 Recurrent somatic *DGCR8* E518K mutations in Wilms tumors

Mutation screening of our exome sequencing cohort revealed four tumors with hotspot mutations leading to a single amino acid exchange with charge reversal (E518K) in the RNA binding domain of *DGCR8*. Based on the known crystal structure of DGCR8 the assumption has been that the E518K mutation may disrupt the strong hydrogen bond interaction of three adjacent amino acids (Wegert et al. 2015). The solution structure of the ADAR2 double-stranded RNA binding motif (dsRBM)-RNA

complex (Stefl et al. 2010) further proposes that the mutated residue in DGCR8 may be implicated in the readout of the dsRNA via the minor groove (Wegert et al. 2015). This could severely impact target selectivity and functioning of the microprocessor complex. Furthermore, based on the recently solved cryoelectron microscopy (cryo-EM) structure of the microprocessor complex with pri-miRNA (Jin et al. 2020; Partin et al. 2020), we and the working group of Dr. Clemens Grimm, were able to visualize the 3D-structure of DGCR8 bound to pri-miRNA with a close-up of the position E518. We could appreciate that glutamic acid at position 518 is strikingly directed away from the RNA substrate. An amino acid exchange by lysine (E518K) which leads to a charge reversal is likely to interact with the neighboring phosphate backbone. This in turn, would most likely reduce target specificity and result in altered miRNA processing.

The frequency of this missense mutation which made up to 70% of all *DGCR8* mutated WTs was similarly high as the *DROSHA* E1147K alteration in the large set of unselected tumors screened and was likewise observed by Walz et al. (2015). Tumors of high-risk histology exhibited an overall higher frequency of *DGCR8* E518K mutations compared to low and intermediate risk WTs just as in *DROSHA* mutated tumors. A correlation with increased death or relapse/metastasis rates, however, could not be observed. Given that both mutations affect important functional domains of the microprocessor complex and occur with comparable frequencies in WTs imply that they may play similar roles in Wilms tumorigenesis.

In contrast to the *DROSHA* mutated tumors, *DGCR8* E518K exclusively showed homozygous expression. The few cases that still did harbor the wild-type allele only expressed the mutated allele, suggestive for the mutation to act recessive. In addition, we could observe a striking sex bias in *DGCR8* mutated cases with a significant female predominance (Walz et al. 2015; Wegert et al. 2015). However, there was no evidence for sex-specific expression or imprinting alterations in the affected genomic region nor proof for monoallelic expression that could explain this bias (Wegert et al. 2015).

6.2.2 *DGCR8* E518K partially rescues miRNA processing defect in DGCR8 KO mESCs

Analysis of the *DGCR8* E518K hotspot mutation requires a model system that lacks DGCR8 to ensure homozygous expression of the mutation, mimicking the situation in the tumor and to exclude secondary effects through endogenous miRNAs. Knockout of DGCR8 was shown to lead to a significant downregulation of all canonical miRNAs in mouse embryonic stem cells (Wang et al. 2007) representing a suitable tool to investigate the impact of the E518K mutation on miRNA processing and other biological processes in a clean background. To assess miRNA expression, microarray and RNA-Seq analysis were performed in this work. It could be shown that inducible overexpression of wild-type *DGCR8* was able to rescue the processing defect of the knockout situation expressing all canonical miRNAs. Inducible overexpression of mutant *DGCR8*, on the other hand, exhibited only a partial rescue of canonical miRNAs with overall moderate miRNA expression levels. While we could

observe a significant downregulation of many canonical miRNAs in E518K, there was still a set of miRNAs that were highly expressed in both wild-type and mutant situation, and only a smaller cluster of miRNAs that showed higher expression in the mutant. These results indicate that the E518K mutation is able to retain significant activity of the microprocessor complex, suggesting that partial reduction of activity or altered specificity might be critical. In our study, a dislocation of DGCR8 E518K from the nucleus or destabilization of DROSHA, which might have contributed to aberrant miRNA processing in the mutant, could be excluded.

6.2.3 *DGCR8* E518K impairs miRNA biogenesis at the pri-miRNA processing step

Mutations in the microprocessor genes leading to decreased levels of mature miRNAs are expected to exhibit higher expression levels of pri-miRNAs. It has previously been reported that the *DROSHA* E1147K mutation, which results in a global downregulation of miRNAs, showed increased levels of unprocessed pri-miRNA substrates (Rakheja et al. 2014; Walz et al. 2015). Also, miRNA-deficient DGCR8 knockout mESCs were shown to accumulate pri-miRNA transcripts as a result of their inability to efficiently process pri-miRNAs to pre-miRNAs (Wang et al. 2007). Comparing the ratio of mature miRNAs and pri-miRNAs in our DGCR8-E518K mESCs confirmed that also the E518K mutation leads to an accumulation of unprocessed pri-miRNAs, while showing reduced mature miRNA expression levels relative to DGCR8-wild-type. However, pri-miRNA expression levels were not as high as in the knockout, supportive of the E518K mutant's partial processing defect. The DGCR8 E518K mutation therefore is likely to impair miRNA biogenesis at the pri-miRNA processing step, possibly due to inefficient binding of its mutated RNA binding domain.

6.2.4 Reduced ESCC miRNA expression levels in DGCR8-E518K mESCs

Several studies have demonstrated the pivotal role of ESC cell cycle (ESCC) miRNAs in the regulation of stem cell properties, such as rapid proliferation and maintenance of self-renewal (Gruber et al. 2014; Marson et al. 2008; Wang et al. 2008). Consistent with their important function, ESCC miRNAs are highly expressed in mESCs. The most abundant miRNAs in undifferentiated mESCs are encoded by the miR-290-295 cluster. They share an identical or a similar seed sequence (5'-AAGUGCU-3') with other ESCC or ESCC-like miRNAs, such as those of the miR-302- and the miR-17-92-cluster (Calabrese et al. 2007; Greve, Judson, and Blelloch 2013; Houbaviy, Murray, and Sharp 2003).

Given that the E518K mutation only led to a partial rescue of the processing defect, where a great number of miRNAs were downregulated, it was worthwhile to check whether the expression of this specific group of miRNAs was altered in DGCR8-E518K mESCs.

MiRNA-Seq analysis showed that ESCC miRNAs were expressed in DGCR8-E518K, however, their expression was mostly reduced compared to DGCR8-wild-type. Particularly, expression of ESCC miRNAs belonging to the miR-302 family was lower in the mutant. Also, members of the miR-290-295

cluster partly showed reduced expression levels. However, the E518K mutant was still able to process a subset of the miR-290-295 family at similar levels as the wild-type. The ESCC-like miR-17 even showed slightly higher expression than in DGCR8-wild-type.

Opposing to ESCC miRNAs, members of the let-7 family almost exclusively exhibited higher expression in DGCR8-E518K. Let-7 miRNAs are known to be highly expressed in somatic cells and differentiating stem cells (Chen et al. 2007). Studies have shown that ESCC- and let-7 miRNAs act in an antagonistic manner to control the switch between self-renewal and differentiation in mESCs. However, let-7 miRNAs are only able to silence the self-renewal program in DGCR8 KO and not in wild-type mESCs. It was shown that ESCC miRNAs were able to prevent let-7-induced suppression of self-renewal by indirectly activating many of the same downstream targets (Guo et al. 2015; Melton, Judson, and Blelloch 2010).

Considering the ability of DGCR8-E518K to still process ESCC miRNAs, albeit at lower levels, capacity of let-7 miRNAs to suppress self-renewal is most likely blocked. However, altered expression of these two opposing miRNA families might still have an impact on proliferation behavior, cell fate decisions, and other cellular processes in DGCR8-E518K mESCs and was further addressed in this work.

6.2.5 *DGCR8* E518K shows normal length distribution among miRNAs

As already mentioned above, the known crystal structure of DGCR8 and superposition models with bound dsRNA, suggest that the E518K mutation might lead to altered sterical interaction with the pri-miRNA substrate. We first assumed that this in turn could lead to a spatial shift of the whole microprocessor complex and would eventually result in altered length distribution by shifted DROSHA cleavage. Alterations in length would then be expected on the 5'end of the 5p-arm and 3'end of the 3p-arm which correspond to the cleavage sites of DROSHA. However, analysis of the miRNA length distribution could not confirm this assumption. E518K processed miRNAs exhibited a similar length distribution as miRNAs that are processed by wild-type DGCR8, indicating that cleavage sites of the microprocessor are not shifted through altered binding of the mutation. This supports the idea of a rather auxiliary role of DGCR8 in cleavage site selection and implies that DGCR8 primarily affects cleavage efficiency (Jin et al. 2020; Partin et al. 2020).

6.2.6 Altered base distribution in the seed sequence of E518K processed miRNAs

We next asked, whether miRNAs processed by DGCR8-E518K exhibited a different seed sequence preference compared to DGCR8-wild-type miRNAs. Alterations in the seed region (nucleotides 2–8 at the 5'end) can alter specificity and stability of miRNA-target binding and thus have a great impact on gene expression. While G and U residues are able to form wobble pairs (C/U and A/G) with the mRNA target, A residues can only pair with U residues in the target sequence and therefore are able to increase target specificity (Wang 2013). GC content of the miRNA seed, on the other hand, has a

major impact on canonical target recognition and downregulation through thermostability of seed-target binding (Wang 2014).

Overall, miRNAs with higher expression in DGCR8-wild-type exhibited a more evenly distributed base composition in their seed sequences compared to those higher expressed in the mutant. The less even base distribution observed in the seed sequences of miRNAs preferentially processed by the mutant, may result from a less diverse miRNA pool.

While the seed sequences of wild-type miRNAs showed elevated ratios of A residues, which could be indicative of high target selectivity, DGCR8-E518K miRNAs exhibited a more extreme base composition profile. Single base enrichments could be appreciated in the seed sequences of miRNAs higher expressed in E518K, with likewise high ratios of A and G nucleotides at various positions. A definite conclusion with regard to altered target specificity can therefore not be drawn. Overall GC distribution in the seed sequences of both DGCR8-wild-type and -E518K miRNAs, however, was similar which implies that thermostability of seed-target binding most likely was not affected.

Bases beyond the seed sequence, particularly the 3'supplementary region at positions 13–16, were reported to also play an essential role in the stabilization of miRNA-mRNA interaction, when base pairing in the seed sequence is suboptimal. Therefore, 3'pairing may enhance miRNA functionality (Broughton et al. 2016; Grimson et al. 2007). DGCR8-wild-type processed miRNAs again exhibited an overall evenly distributed base composition compared to miRNAs higher expressed in E518K. MiRNAs processed by the mutant, on the other hand, showed highly elevated ratios of A nucleotides towards the 3'end. This might allow specific target binding beyond the seed in miRNAs processed by the mutant.

Altogether, analysis of the seed sequences and beyond showed differences in the base composition of miRNAs preferentially processed by either DGCR8-wt or DGCR8-E518K. These differences, however, most likely rather represent a shift in the spectrum of expressed miRNAs in the mutant than alterations in target binding specificity itself. Still, a shift in the range of processed miRNAs can lead to different miRNA-mRNA interactions and ultimately result in altered target gene expression.

6.2.7 Altered target gene regulation in DGCR8-E518K mESCs

In order to better understand the relationship between differentially expressed miRNAs and altered mRNA expression, miRNA target interactions (MTIs) and putatively affected pathways were analyzed in this work. Despite the significant downregulation of many canonical miRNAs in E518K, there was still a set of miRNAs that were not affected by the mutation, showing similarly high expression levels as in the wild-type. Targets of these miRNAs play essential roles in normal cell function, such as regulation of cytoskeleton, cell division, apoptotic signaling, and catabolic processes. The E518K mutation therefore seems to be able to regulate important biological processes in the cell and would suggest that residual miRNA activity in Wilms tumor might be needed for the tumor's cell viability

and growth. Consistent with this, it was shown that *DROSHA* and *DICER1* mutations in Wilms tumor did not lead to complete loss of miRNAs, which supports the idea that Wilms tumorigenesis likely depends on residual miRNA expression (Rakheja et al. 2014).

DGCR8-wild-type additionally showed increased expression levels of miRNAs with target genes related to RNA-processing, including RNA-splicing, and ribonucleoprotein complex biogenesis. A large proportion of these genes are involved in the pre-mRNA splicing process, partly as components of the spliceosome. Processing of intronic miRNAs by the microprocessor complex and splicing of their host introns by the spliceosome were previously shown to be two competing processing reactions (Kataoka, Fujita, and Ohno 2009). Downregulation of genes associated with RNA splicing in DGCR8-wild-type may therefore represent a secondary effect related to increased pri-miRNA processing. Furthermore, translational regulatory genes, such as members of the eukaryotic translation initiation factors (eIFs) and RNA helicases of the DEAD box family, were targeted in DGCR8-wild-type mESCs. Studies have shown that global translation needs to be tightly regulated in embryonic stem cells to maintain an undifferentiated state, but also to be able to rapidly respond to differentiation signals during cell fate decisions (Ingolia, Lareau, and Weissman 2011; Sampath et al. 2008). MiRNAs higher expressed in DGCR8-wild-type therefore seem to target translation related genes to maintain stem cell properties and homeostasis.

MiRNAs that are less affected by the E518K mutation are involved in the regulation of genes that are associated with protein ubiquitination and kinase activity. Further analysis showed that these genes are involved in the regulation of various cellular processes, such as apoptosis, DNA damage response, and cytoskeletal organization. Apoptotic-related genes that were targeted by miRNAs in E518K, normally lead to ubiquitination and subsequent degradation of p53, preventing p53-mediated cell cycle arrest and apoptosis. Downregulation of these antiapoptotic factors in turn would result in increased cell death. Also, genes that are important for DNA damage response, such as histone deacetylase 6 (HDAC6) were downregulated in DGCR8-E518K. Depletion or inhibition of HDAC6 was previously shown to induce DNA damage and apoptosis (Namdar et al. 2010; Zhang et al. 2014). Decreased expression levels might therefore lead to reduced cell viability in the mutant. Finally, genes important for cytoskeletal organization were downregulated in DGCR8-E518K which might be an indication of disrupted cytoskeletal architecture and function. Altered cell shape and cytoskeletal regulation in turn may have an impact on self-renewal and cell fate decisions (Murray et al. 2013).

In summary, based on the predicted MTIs and overrepresentation analysis (ORA), DGCR8-E518K appears to be still capable to regulate most of the same target genes as DGCR8-wild-type. However, given that the miRNA processing defect was only partially rescued by the mutation, DGCR8-E518K mESCs exhibited an altered miRNA expression profile, which in turn led to altered target gene regulation. Ultimately, this could result in aberrant biological processes and impact important cellular

functions. Yet, it is important to note that these pathway associations are solely based on mRNA level requiring further testing on protein level, e.g. through proteome analysis in order to fully understand the role of the deregulated miRNA landscape in DGCR8-E518K mESCs. Furthermore, studies have revealed that the function of the microprocessor complex is not exclusively restricted to miRNA biogenesis, but also involves miRNA-independent activity of DROSHA/DGCR8 (Cirera-Salinas et al. 2017; Han et al. 2009; Macias et al. 2012). Through cross-linking immunoprecipitation (CLIP) experiments it was shown that DGCR8 can directly bind mRNA substrates to control their abundance and is able to modulate the relative abundance of alternatively spliced isoforms. In addition, DGCR8 is most likely involved in cellular complexes with endonucleases other than DROSHA (Macias et al. 2015; Macias et al. 2012). Changes in mRNA expression observed in our DGCR8-E518K mESCs may therefore not solely result from altered miRNA processing, but also be partly due to the impact of the mutation on the non-canonical functions of DGCR8.

6.2.8 Altered cell viability in E518K mESCs

To functionally assess the impact of the DGCR8 E518K mutation on the proliferation behavior in mESCs, MTT and crystal violet assay were performed. Additionally, cell cycle distributions and S-phase progression were analyzed.

MTT and crystal violet assay confirmed the proliferation deficit reported in DGCR8 KO mESCs (Wang et al. 2007). Cell cycle analysis additionally confirmed an accumulation of cells in G1 phase and concomitant low fraction of cells in S-Phase in the knockout, underlining the essential role that miRNAs play in cell cycle progression. Furthermore, MTT and crystal violet assay revealed reduced absorption rates in DGCR8-E518K mESCs compared to DGCR8-wild-type, while rates were still higher than in the knockout. Surprisingly however, DGCR8-E518K mESCs did not show an accumulation of cells in G1 that could explain the observed decrease in the proliferation rates of E518K mESCs. Fraction of cells in S-phase were slightly reduced in the mutant, which was also observed by BrdU assay. The differences between DGCR8-E518K and DGCR8-wt, however, were statistically not significant. The discrepancy observed between MTT/crystal violet assay and cell cycle analysis may thus be related to the fact that both MTT and crystal violet assay are rather indirect methods to measure cell proliferation. They are actually measures of metabolic activity and cell viability (Berridge, Herst, and Tan 2005; Feoktistova, Geserick, and Leverkus 2016). Therefore, the observed differences between DGCR8-E518K and -wild-type most likely display changes in cell viability and seem to be cell cycle independent.

Analysis of cell cycle specific genes indeed showed a significant downregulation of upstream inhibitors of the cycline E/Cdk2 pathway, including *Cdkn1a*, *Rb1*, *Rbl1*, *Rbl2*, and *Lats2* in both DGCR8-E518K and DGCR8-wild-type mESCs compared to the knockout. It was previously shown that these genes are targeted and regulated by ESCC miRNAs in mESCs (Wang et al. 2008). Rapid G1-S

transition by ESCC miRNAs, however, was later shown to be only *Rb*-dependent under cytostatic culture conditions. Under standard culture conditions therefore, ESCC miRNAs most likely promote cell cycle progression through additional pathways (Wang et al. 2013). Opposing to these inhibitors of G1-S transition, cell-cycle promoting CDKs (cyclin-dependent kinases) and CDCs (cell division cycle genes) were upregulated in both DGCR8-E518K and -wild-type mESCs, and showed decreased expression in the knockout.

Based on these findings, miRNA expression in DGCR8-E518K appears to be sufficient to rescue the cell cycle defect in DGCR8-KO mESCs, albeit not to the full extent as in DGCR8-wt. Yet, the results from MTT and crystal violet assay indicate that cell viability may be reduced in DGCR8-E518K. PI staining complemented with annexin V staining could help to elucidate possibly increased apoptosis rates via flow cytometry.

6.2.9 EMT can most likely not be fully blocked in DGCR8-E518K mESCs

Morphologically, E518K mESCs slightly differed from DGCR8-wild-type mESCs and exhibited a tendency to establish mesenchymal-like features, albeit to a much milder extent than DGCR8-KO mESCs. Still, these morphological differences suggested that cells possibly underwent epithelial-to-mesenchymal-transition (EMT).

Expression of EMT regulatory genes was shown to be antagonistically regulated by ESCC- and let-7 miRNAs in mESCs. DGCR8 KO mESCs were reported to be unable to block EMT and to exhibit mesenchymal-like morphology under feeder-free conditions which was reversed by the introduction of ESCC miRNAs. Let-7 miRNAs, on the other hand, promoted EMT in mESCs by upregulating EMT regulators (Guo et al. 2015). Given the different distribution of these miRNA families in DGCR8-E518K versus DGCR8-wild-type, we checked for differential expression of representative epithelial- and mesenchymal genes to understand how EMT is regulated in E518K mESCs.

We could confirm the inability of DGCR8-KO mESCs to block EMT exhibiting a mesenchymal-like morphology and could show that mesenchymal genes, including key EMT regulators, such as *Snai1* and *Zeb1* were highest expressed in the knockout. Despite exhibiting the highest expression of mesenchymal markers, DGCR8-KO mESCs still showed high expression of epithelial genes. Simultaneous expression of mesenchymal and epithelial genes indicate an intermediate EMT process (Sha et al. 2019) in DGCR8 KO mESCs, which was likewise shown by Guo et al. (2015). A rescue with DGCR8-wt with a more prominent ESCC miRNA signature led to a suppression of EMT showing strongly reduced expression of mesenchymal and high levels of epithelial markers. DGCR8-E518K cells likewise exhibited decreased levels of mesenchymal gene expression, however, with concomitantly reduced levels of epithelial markers. The distinct miRNA pattern with more prominent let-7 miRNAs and decreased ESCC levels in DGCR8-E518K therefore most likely results in a less

complete block of EMT, suggesting that stem cell characteristics cannot be fully restored in the mutant.

6.2.10 *DGCR8* E518K rescues differentiation defect in mouse embryonic bodies (EBs)

To assess the differentiation capacity of DGCR8-E518K mESCs, they were cultured as embryoid bodies in the absence of differentiation inhibiting factors and adhesive substrates. Despite their smaller size, they were likewise able to form cystic EBs as DGCR8-wild-type. In contrast, DGCR8-knockout mESCs failed to form a cyst, suggesting incomplete differentiation.

MiRNA deficient DICER1 and DGCR8 knockout mESCs were previously shown to exhibit severe differentiation defects unable to silence their self-renewal program, not even under stringent differentiation conditions (Kanellopoulou et al. 2005; Wang et al. 2007). Multiple differentiation promoting miRNAs were detected that were able to silence self-renewal in DGCR8 KO mESCs. These miRNAs were shown to be enriched in different cell lineages and tissues implying that different miRNAs regulate differentiation of mESCs towards different lineages. However, the majority could only silence self-renewal in a DGCR8 KO background, since their ability to promote differentiation is blocked by the highly expressed ESCC miRNAs in wild-type mESCs (Ma et al. 2015; Melton, Judson, and Blelloch 2010; Wang et al. 2013).

DGCR8-E518K mESCs still expressed ESCC miRNAs, albeit at lower levels than DGCR8-wild-type. This change could already have an impact on the differentiation capacity, even if a much milder phenotype was expected than in the knockout. However, RNA analysis of EBs at different timepoints barely showed a difference between mutant and wild-type, indicating that miRNA expression in E518K mESCs is likely to be sufficient to restore differentiation capacity and to silence self-renewal. Due to the ESCC miRNAs present in E518K mESCs, cell fate decisions can thus still be regulated leading to a normal differentiation process comprising all three germ layers.

Our findings, therefore, suggest that the *DGCR8* E518K mutation alone may not be sufficient to impair differentiation of renal progenitors leading to a blastemal phenotype. Nevertheless, the mutation might still be relevant in the onset of other histological subtypes of Wilms tumors with more differentiated structures.

6.3 Outlook

In this work, it could be shown that the hotspot mutations *DROSHA* E1147K and *DGCR8* E518K are frequent events in Wilms tumor that lead to altered miRNA expression *in vitro*, confirming our previous findings in WT samples. While the findings in this study underline the implication of microprocessor gene mutations in WT, further studies are needed to better understand the function of mutated pathways in the onset of WT. Here, mouse embryonic stem cells (mESCs) were used as a model system to investigate the differentiation capacity of DGCR8-E518K. Functional analysis showed that despite significant differences in the miRNA and mRNA expression profiles of DGCR8-E518K and DGCR8-wt, DGCR8-E518K was mostly able to restore essential biological processes. However, it is important to consider that aberrant miRNA processing might have a different impact in mESCs than in Wilms tumor. Reduced miRNA expression, specifically of ESCC miRNAs, leads to impairment of stem cell identity in mESCs. Functioning miRNA processing is therefore essential in mESCs to maintain their characteristic features as pluripotent stem cells, such as highly proliferative activity and unlimited self-renewal capacity. The embryonic progenitor cell clusters in WT share these properties with normal embryonic stem cells, however, WTs would rather develop these features due to aberrant expression of miRNAs. Deregulated miRNAs in the developing kidney are likely to contribute to continuous proliferation and self-renewal of renal progenitor cells, preventing the cells from differentiating. Therefore, altered miRNA expression may lead to different phenotypical outcomes in the different cell systems. It would be worthwhile to functionally assess the *DGCR8* E518K mutation in a Wilms tumor blastemal model system, such as WT blastemal cancer stem cells (Shukrun et al. 2014) or the 3D WT spheroid culture system, which was established by our working group (Wegert et al. 2020). This way, mutations can be analyzed in a rather physiological environment mirroring the *in vivo* state of the tumor and can help to better comprehend the functional consequences of the microprocessor mutations in Wilms tumorigenesis.

7 References

Androvic, P., L. Valihrach, J. Elling, R. Sjoback, and M. Kubista. 2017. 'Two-tailed RT-qPCR: a novel method for highly accurate miRNA quantification', *Nucleic Acids Res*, 45: e144.

Anglesio, M. S., Y. Wang, W. Yang, J. Senz, A. Wan, A. Heravi-Moussavi, C. Salamanca, S. Maines-Bandiera, D. G. Huntsman, and G. B. Morin. 2013. 'Cancer-associated somatic DICER1 hotspot mutations cause defective miRNA processing and reverse-strand expression bias to predominantly mature 3p strands through loss of 5p strand cleavage', *J Pathol*, 229: 400-9.

Astuti, D., M. R. Morris, W. N. Cooper, R. H. Staals, N. C. Wake, G. A. Fews, H. Gill, D. Gentle, S. Shuib, C. J. Ricketts, T. Cole, A. J. van Essen, R. A. van Lingen, G. Neri, J. M. Opitz, P. Rump, I. Stolte-Dijkstra, F. Muller, G. J. Pruijn, F. Latif, and E. R. Maher. 2012. 'Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility', *Nat Genet*, 44: 277-84.

Bardeesy, N., D. Falkoff, M. J. Petruzzi, N. Nowak, B. Zabel, M. Adam, M. C. Aguiar, P. Grundy, T. Shows, and J. Pelletier. 1994. 'Anaplastic Wilms' tumour, a subtype displaying poor prognosis, harbours p53 gene mutations', *Nat Genet*, 7: 91-7.

Bartel, D. P. 2004. 'MicroRNAs: genomics, biogenesis, mechanism, and function', *Cell*, 116: 281-97.

Bartram, M. P., E. Amendola, T. Benzing, B. Schermer, G. de Vita, and R. U. Muller. 2016. 'Mice lacking microRNAs in Pax8-expressing cells develop hypothyroidism and end-stage renal failure', *BMC Mol Biol*, 17: 11.

Bartram, M. P., C. Dafinger, S. Habbig, T. Benzing, B. Schermer, and R. U. Muller. 2015. 'Loss of Dgcr8-mediated microRNA expression in the kidney results in hydronephrosis and renal malformation', *BMC Nephrol*, 16: 55.

Bartram, M. P., M. Hohne, C. Dafinger, L. A. Volker, M. Albersmeyer, J. Heiss, H. Gobel, H. Bronneke, V. Burst, M. C. Liebau, T. Benzing, B. Schermer, and R. U. Muller. 2013. 'Conditional loss of kidney microRNAs results in congenital anomalies of the kidney and urinary tract (CAKUT)', *J Mol Med (Berl)*, 91: 739-48.

Beckwith, J. B., N. B. Kiviat, and J. F. Bonadio. 1990. 'Nephrogenic rests, nephroblastomatosis, and the pathogenesis of Wilms' tumor', *Pediatr Pathol*, 10: 1-36.

Berridge, M. V., P. M. Herst, and A. S. Tan. 2005. 'Tetrazolium dyes as tools in cell biology: new insights into their cellular reduction', *Biotechnol Annu Rev*, 11: 127-52.

Bohnsack, M. T., K. Czaplinski, and D. Gorlich. 2004. 'Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs', *RNA*, 10: 185-91.

Bradley, A. 1990. 'Embryonic stem cells: proliferation and differentiation', *Curr Opin Cell Biol*, 2: 1013-7.

Breslow, N. E., J. Olson, J. Moksness, J. B. Beckwith, and P. Grundy. 1996. 'Familial Wilms' tumor: a descriptive study', *Med Pediatr Oncol*, 27: 398-403.

Breslow, N., A. Olshan, J. B. Beckwith, and D. M. Green. 1993. 'Epidemiology of Wilms tumor', *Med Pediatr Oncol*, 21: 172-81.

Broughton, J. P., M. T. Lovci, J. L. Huang, G. W. Yeo, and A. E. Pasquinelli. 2016. 'Pairing beyond the Seed Supports MicroRNA Targeting Specificity', *Mol Cell*, 64: 320-33.

Cai, X., C. H. Hagedorn, and B. R. Cullen. 2004. 'Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs', *RNA*, 10: 1957-66.

Calabrese, J. M., A. C. Seila, G. W. Yeo, and P. A. Sharp. 2007. 'RNA sequence analysis defines Dicer's role in mouse embryonic stem cells', *Proc Natl Acad Sci U S A*, 104: 18097-102.

Call, K. M., T. Glaser, C. Y. Ito, A. J. Buckler, J. Pelletier, D. A. Haber, E. A. Rose, A. Kral, H. Yeger, W. H. Lewis, and et al. 1990. 'Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumor locus', *Cell*, 60: 509-20.

Cao, Y., W. T. Guo, S. Tian, X. He, X. W. Wang, X. Liu, K. L. Gu, X. Ma, D. Huang, L. Hu, Y. Cai, H. Zhang, Y. Wang, and P. Gao. 2015. 'miR-290/371-Mbd2-Myc circuit regulates glycolytic metabolism to promote pluripotency', *EMBO J*, 34: 609-23.

Castellano, L., and J. Stebbing. 2013. 'Deep sequencing of small RNAs identifies canonical and non-canonical miRNA and endogenous siRNAs in mammalian somatic tissues', *Nucleic Acids Res*, 41: 3339-51.

Chen, C., D. Ridzon, C. T. Lee, J. Blake, Y. Sun, and W. M. Strauss. 2007. 'Defining embryonic stem cell identity using differentiation-related microRNAs and their potential targets', *Mamm Genome*, 18: 316-27.

Chong, M. M., G. Zhang, S. Cheloufi, T. A. Neubert, G. J. Hannon, and D. R. Littman. 2010. 'Canonical and alternate functions of the microRNA biogenesis machinery', *Genes Dev*, 24: 1951-60.

Chu, J. Y., S. Sims-Lucas, D. S. Bushnell, A. J. Bodnar, J. A. Kreidberg, and J. Ho. 2014. 'Dicer function is required in the metanephric mesenchyme for early kidney development', *Am J Physiol Renal Physiol*, 306: F764-72.

Cirera-Salinas, D., J. Yu, M. Bodak, R. P. Ngondo, K. M. Herbert, and C. Ciaudo. 2017. 'Noncanonical function of DGCR8 controls mESC exit from pluripotency', *J Cell Biol*, 216: 355-66.

Coppes-Zantinga, A. R., and M. J. Coppes. 1999. 'The eponym "Wilms": a reminder of a surgeon's lifelong contributions to medicine', *Med Pediatr Oncol*, 32: 438-9.

Cullen, B. R. 2004. 'Transcription and processing of human microRNA precursors', *Mol Cell*, 16: 861-5.

Davidoff, A. M. 2012. 'Wilms tumor', *Adv Pediatr*, 59: 247-67.

De Kraker, J., N. Graf, K. Pritchard-Jones, and F. Pein. 2001. "Nephroblastoma clinical trial and study SIOP 2001, Protocol." In *SIOP RTSG 2001*.

Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, and T. R. Gingeras. 2013. 'STAR: ultrafast universal RNA-seq aligner', *Bioinformatics*, 29: 15-21.

Dome, J. S., C. V. Fernandez, E. A. Mullen, J. A. Kalapurakal, J. I. Geller, V. Huff, E. J. Gratias, D. B. Dix, P. F. Ehrlich, G. Khanna, M. H. Malogolowkin, J. R. Anderson, A. Naranjo, E. J. Perlman, and C. O. G. Renal Tumors Committee. 2013. 'Children's Oncology Group's 2013 blueprint for research: renal tumors', *Pediatr Blood Cancer*, 60: 994-1000.

Esquela-Kerscher, A., and F. J. Slack. 2006. 'Oncomirs - microRNAs with a role in cancer', *Nat Rev Cancer*, 6: 259-69.

Evans, M. J., and M. H. Kaufman. 1981. 'Establishment in culture of pluripotential cells from mouse embryos', *Nature*, 292: 154-6.

Feoktistova, M., P. Geserick, and M. Leverkus. 2016. 'Crystal Violet Assay for Determining Viability of Cultured Cells', *Cold Spring Harb Protoc*, 2016: pdb prot087379.

Fukunaga, R., B. W. Han, J. H. Hung, J. Xu, Z. Weng, and P. D. Zamore. 2012. 'Dicer Partner Proteins Tune the Length of Mature miRNAs in Flies and Mammals', *Cell*, 151: 912.

Gan, J., J. E. Tropea, B. P. Austin, D. L. Court, D. S. Waugh, and X. Ji. 2006. 'Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III', *Cell*, 124: 355-66.

Gessler, M., A. Konig, K. Arden, P. Grundy, S. Orkin, S. Sallan, C. Peters, S. Ruyle, J. Mandell, F. Li, and et al. 1994. 'Infrequent mutation of the WT1 gene in 77 Wilms' Tumors', *Hum Mutat*, 3: 212-22.

Gessler, M., A. Poustka, W. Cavenee, R. L. Neve, S. H. Orkin, and G. A. Bruns. 1990. 'Homozygous deletion in Wilms tumours of a zinc-finger gene identified by chromosome jumping', *Nature*, 343: 774-8.

Gossen, M., S. Freundlieb, G. Bender, G. Muller, W. Hillen, and H. Bujard. 1995. 'Transcriptional activation by tetracyclines in mammalian cells', *Science*, 268: 1766-9.

Green, D. M., Y. A. Grigoriev, B. Nan, J. R. Takashima, P. A. Norkool, G. J. D'Angio, and N. E. Breslow. 2001. 'Congestive heart failure after treatment for Wilms' tumor: a report from the National Wilms' Tumor Study group', *J Clin Oncol*, 19: 1926-34.

Gregory, R. I., K. P. Yan, G. Amuthan, T. Chendrimada, B. Doratotaj, N. Cooch, and R. Shiekhattar. 2004. 'The Microprocessor complex mediates the genesis of microRNAs', *Nature*, 432: 235-40.

Greve, T. S., R. L. Judson, and R. Blelloch. 2013. 'microRNA control of mouse and human pluripotent stem cell behavior', *Annu Rev Cell Dev Biol*, 29: 213-39.

Grimson, A., K. K. Farh, W. K. Johnston, P. Garrett-Engele, L. P. Lim, and D. P. Bartel. 2007. 'MicroRNA targeting specificity in mammals: determinants beyond seed pairing', *Mol Cell*, 27: 91-105.

Gruber, A. J., W. A. Grandy, P. J. Balwierz, Y. A. Dimitrova, M. Pachkov, C. Ciaudo, Ev Nimwegen, and M. Zavolan. 2014. 'Embryonic stem cell-specific microRNAs contribute to pluripotency by inhibiting regulators of multiple differentiation pathways', *Nucleic Acids Res*, 42: 9313-26.

Grundy, P. E., N. E. Breslow, S. Li, E. Perlman, J. B. Beckwith, M. L. Ritchey, R. C. Shamberger, G. M. Haase, G. J. D'Angio, M. Donaldson, M. J. Coppes, M. Malogolowkin, P. Shearer, P. R. Thomas, R. Macklis, G. Tomlinson, V. Huff, D. M. Green, and Group National Wilms Tumor Study. 2005. 'Loss of heterozygosity for chromosomes 1p and 16q is an adverse prognostic factor in favorable-histology Wilms tumor: a report from the National Wilms Tumor Study Group', *J Clin Oncol*, 23: 7312-21.

Guarnieri, D. J., and R. J. DiLeone. 2008. 'MicroRNAs: a new class of gene regulators', *Ann Med*, 40: 197-208.

Guo, W. T., X. W. Wang, Y. L. Yan, Y. P. Li, X. Yin, Q. Zhang, C. Melton, A. Shenoy, N. A. Reyes, S. A. Oakes, R. Blelloch, and Y. Wang. 2015. 'Suppression of epithelial-mesenchymal transition and apoptotic pathways by miR-294/302 family synergistically blocks let-7-induced silencing of self-renewal in embryonic stem cells', *Cell Death Differ*, 22: 1158-69.

Guo, W. T., and Y. Wang. 2019. 'Dgcr8 knockout approaches to understand microRNA functions in vitro and in vivo', *Cell Mol Life Sci*.

Ha, M., and V. N. Kim. 2014. 'Regulation of microRNA biogenesis', *Nat Rev Mol Cell Biol*, 15: 509-24.

Han, J., Y. Lee, K. H. Yeom, Y. K. Kim, H. Jin, and V. N. Kim. 2004. 'The Drosha-DGCR8 complex in primary microRNA processing', *Genes Dev*, 18: 3016-27.

Han, J., J. S. Pedersen, S. C. Kwon, C. D. Belair, Y. K. Kim, K. H. Yeom, W. Y. Yang, D. Haussler, R. Blelloch, and V. N. Kim. 2009. 'Posttranscriptional crossregulation between Drosha and DGCR8', *Cell*, 136: 75-84.

Hanahan, D., and R. A. Weinberg. 2011. 'Hallmarks of cancer: the next generation', *Cell*, 144: 646-74.

Hochberg, Y., and Y. Benjamini. 1990. 'More powerful procedures for multiple significance testing', *Stat Med*, 9: 811-8.

Houbaviy, H. B., M. F. Murray, and P. A. Sharp. 2003. 'Embryonic stem cell-specific MicroRNAs', *Dev Cell*, 5: 351-8.

Hsu, S. D., F. M. Lin, W. Y. Wu, C. Liang, W. C. Huang, W. L. Chan, W. T. Tsai, G. Z. Chen, C. J. Lee, C. M. Chiu, C. H. Chien, M. C. Wu, C. Y. Huang, A. P. Tsou, and H. D. Huang. 2011. 'miRTarBase: a database curates experimentally validated microRNA-target interactions', *Nucleic Acids Res*, 39: D163-9.

Huang, Y., X. J. Shen, Q. Zou, S. P. Wang, S. M. Tang, and G. Z. Zhang. 2011. 'Biological functions of microRNAs: a review', *J Physiol Biochem*, 67: 129-39.

Huber, W., A. von Heydebreck, H. Sultmann, A. Poustka, and M. Vingron. 2002. 'Variance stabilization applied to microarray data calibration and to the quantification of differential expression', *Bioinformatics*, 18 Suppl 1: S96-104.

Huff, V. 1998. 'Wilms tumor genetics', *Am J Med Genet*, 79: 260-7.

Huntzinger, E., and E. Izaurralde. 2011. 'Gene silencing by microRNAs: contributions of translational repression and mRNA decay', *Nat Rev Genet*, 12: 99-110.

Iervolino, A., F. Trepiccione, F. Petrillo, M. Spagnuolo, M. Scarfo, D. Frezzetti, G. De Vita, M. De Felice, and G. Capasso. 2015. 'Selective dicer suppression in the kidney alters GSK3beta/beta-catenin pathways promoting a glomerulocystic disease', *PLoS One*, 10: e0119142.

Ingolia, N. T., L. F. Lareau, and J. S. Weissman. 2011. 'Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes', *Cell*, 147: 789-802.

Ivics, Z., P. B. Hackett, R. H. Plasterk, and Z. Izsvak. 1997. 'Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells', *Cell*, 91: 501-10.

Jin, W., J. Wang, C. P. Liu, H. W. Wang, and R. M. Xu. 2020. 'Structural Basis for pri-miRNA Recognition by Drosha', *Mol Cell*.

John, B., A. J. Enright, A. Aravin, T. Tuschl, C. Sander, and D. S. Marks. 2004. 'Human MicroRNA targets', *PLoS Biol*, 2: e363.

Kanellopoulou, C., S. A. Muljo, A. L. Kung, S. Ganesan, R. Drapkin, T. Jenuwein, D. M. Livingston, and K. Rajewsky. 2005. 'Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing', *Genes Dev*, 19: 489-501.

Kataoka, N., M. Fujita, and M. Ohno. 2009. 'Functional association of the Microprocessor complex with the spliceosome', *Mol Cell Biol*, 29: 3243-54.

Kawamata, T., and Y. Tomari. 2010. 'Making RISC', *Trends Biochem Sci*, 35: 368-76.

Kim, J., and S. H. Orkin. 2011. 'Embryonic stem cell-specific signatures in cancer: insights into genomic regulatory networks and implications for medicine', *Genome Med*, 3: 75.

Kim, Y. K., and V. N. Kim. 2007. 'Processing of intronic microRNAs', *EMBO J*, 26: 775-83.

Koesters, R., R. Ridder, A. Kopp-Schneider, D. Betts, V. Adams, F. Niggli, J. Briner, and M. von Knebel Doeberitz. 1999. 'Mutational activation of the beta-catenin proto-oncogene is a common event in the development of Wilms' tumors', *Cancer Res*, 59: 3880-2.

Kondoh, H., M. E. Lleonart, Y. Nakashima, M. Yokode, M. Tanaka, D. Bernard, J. Gil, and D. Beach. 2007. 'A high glycolytic flux supports the proliferative potential of murine embryonic stem cells', *Antioxid Redox Signal*, 9: 293-9.

Kozomara, A., and S. Griffiths-Jones. 2014. 'miRBase: annotating high confidence microRNAs using deep sequencing data', *Nucleic Acids Res*, 42: D68-73.

Kruber, P., O. Angay, A. Winkler, M. R. Bosl, S. Kneitz, K. G. Heinze, and M. Gessler. 2019. 'Loss or oncogenic mutation of DROSHA impairs kidney development and function, but is not sufficient for Wilms tumor formation', *Int J Cancer*, 144: 1391-400.

Kumar, M. S., J. Lu, K. L. Mercer, T. R. Golub, and T. Jacks. 2007. 'Impaired microRNA processing enhances cellular transformation and tumorigenesis', *Nat Genet*, 39: 673-7.

Landthaler, M., A. Yalcin, and T. Tuschl. 2004. 'The human DiGeorge syndrome critical region gene 8 and Its D. melanogaster homolog are required for miRNA biogenesis', *Curr Biol*, 14: 2162-7.

Lee, Y., K. Jeon, J. T. Lee, S. Kim, and V. N. Kim. 2002. 'MicroRNA maturation: stepwise processing and subcellular localization', *EMBO J*, 21: 4663-70.

Lee, Y., M. Kim, J. Han, K. H. Yeom, S. Lee, S. H. Baek, and V. N. Kim. 2004. 'MicroRNA genes are transcribed by RNA polymerase II', *EMBO J*, 23: 4051-60.

Lewis, B. P., C. B. Burge, and D. P. Bartel. 2005. 'Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets', *Cell*, 120: 15-20.

Lewis, B. P., I. H. Shih, M. W. Jones-Rhoades, D. P. Bartel, and C. B. Burge. 2003. 'Prediction of mammalian microRNA targets', *Cell*, 115: 787-98.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and Subgroup Genome Project Data Processing. 2009. 'The Sequence Alignment/Map format and SAMtools', *Bioinformatics*, 25: 2078-9.

Loew, R., N. Heinz, M. Hampf, H. Bujard, and M. Gossen. 2010. 'Improved Tet-responsive promoters with minimized background expression', *BMC Biotechnol*, 10: 81.

Lopez-Rivera, E., Y. P. Liu, M. Verbitsky, B. R. Anderson, V. P. Capone, E. A. Otto, Z. Yan, A. Mitrotti, J. Martino, N. J. Steers, D. A. Fasel, K. Vukojevic, R. Deng, S. E. Racedo, Q. Liu, M. Werth, R. Westland, A. Vivante, G. S. Makar, M. Bodria, M. G. Sampson, C. E. Gillies, V. Vega-Warner, M. Maiorana, D. S. Petrey, B. Honig, V. J. Lozanovski, R. Salomon, L. Heidet, W. Carpentier, D. Gaillard, A. Carrea, L. Gesualdo, D. Cusi, C. Izzi, F. Scolari, J. A. van Wijk, A. Arapovic, M. Saraga-Babic, M. Saraga, N. Kunac, A. Samii, D. M. McDonald-McGinn, T. B. Crowley, E. H. Zackai, D. Drozdz, M. Miklaszewska, M. Tkaczyk, P. Sikora, M. Szczepanska, M. Mizerska-Wasiak, G. Krzemien, A. Szmigielska, M. Zaniew, J. M. Darlow, P. Puri, D. Barton, E. Casolari, S. L. Furth, B. A. Warady, Z. Gucev, H. Hakonarson, H. Flogelova, V. Tasic, A. Latos-Bielenska, A. Materna-Kiryluk, L. Allegri, C. S. Wong, I. A. Drummond, V. D'Agati, A. Imamoto, J. M. Barasch, F. Hildebrandt, K. Kiryluk, R. P. Lifton, B. E. Morrow, C. Jeanpierre, V. E. Papaioannou, G. M. Ghiggeri, A. G. Gharavi, N. Katsanis, and S. Sanna-Cherchi. 2017. 'Genetic Drivers of Kidney Defects in the DiGeorge Syndrome', *N Engl J Med*, 376: 742-54.

Lu, J., G. Getz, E. A. Miska, E. Alvarez-Saavedra, J. Lamb, D. Peck, A. Sweet-Cordero, B. L. Ebert, R. H. Mak, A. A. Ferrando, J. R. Downing, T. Jacks, H. R. Horvitz, and T. R. Golub. 2005. 'MicroRNA expression profiles classify human cancers', *Nature*, 435: 834-8.

Ma, Y., N. Yao, G. Liu, L. Dong, Y. Liu, M. Zhang, F. Wang, B. Wang, X. Wei, H. Dong, L. Wang, S. Ji, J. Zhang, Y. Wang, Y. Huang, and J. Yu. 2015. 'Functional screen reveals essential roles of miR-27a/24 in differentiation of embryonic stem cells', *EMBO J*, 34: 361-78.

Macias, S., R. A. Cordiner, and J. F. Caceres. 2013. 'Cellular functions of the microprocessor', *Biochem Soc Trans*, 41: 838-43.

Macias, S., R. A. Cordiner, P. Gautier, M. Plass, and J. F. Caceres. 2015. 'DGCR8 Acts as an Adaptor for the Exosome Complex to Degrade Double-Stranded Structured RNAs', *Mol Cell*, 60: 873-85.

Macias, S., M. Plass, A. Stajuda, G. Michlewski, E. Eyras, and J. F. Caceres. 2012. 'DGCR8 HITS-CLIP reveals novel functions for the Microprocessor', *Nat Struct Mol Biol*, 19: 760-6.

Maiti, S., R. Alam, C. I. Amos, and V. Huff. 2000. 'Frequent association of beta-catenin and WT1 mutations in Wilms tumors', *Cancer Res*, 60: 6288-92.

Marson, A., S. S. Levine, M. F. Cole, G. M. Frampton, T. Brambrink, S. Johnstone, M. G. Guenther, W. K. Johnston, M. Wernig, J. Newman, J. M. Calabrese, L. M. Dennis, T. L. Volkert, S. Gupta, J. Love, N. Hannett, P. A. Sharp, D. P. Bartel, R. Jaenisch, and R. A. Young. 2008. 'Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells', *Cell*, 134: 521-33.

Matsunaga, E. 1981. 'Genetics of Wilms' tumor', *Hum Genet*, 57: 231-46.

Melton, C., R. L. Judson, and R. Blelloch. 2010. 'Opposing microRNA families regulate self-renewal in mouse embryonic stem cells', *Nature*, 463: 621-6.

Metzger, M. L., and J. S. Dome. 2005. 'Current therapy for Wilms' tumor', *Oncologist*, 10: 815-26.

Murray, P., M. Prewitz, I. Hopp, N. Wells, H. Zhang, A. Cooper, K. L. Parry, R. Short, D. J. Antoine, and D. Edgar. 2013. 'The self-renewal of mouse embryonic stem cells is regulated by cell-substratum adhesion and cell spreading', *Int J Biochem Cell Biol*, 45: 2698-705.

Nagalakshmi, V. K., Q. Ren, M. M. Pugh, M. T. Valerius, A. P. McMahon, and J. Yu. 2011. 'Dicer regulates the development of nephrogenic and ureteric compartments in the mammalian kidney', *Kidney Int*, 79: 317-30.

Namdar, M., G. Perez, L. Ngo, and P. A. Marks. 2010. 'Selective inhibition of histone deacetylase 6 (HDAC6) induces DNA damage and sensitizes transformed cells to anticancer agents', *Proc Natl Acad Sci U S A*, 107: 20003-8.

Nguyen, T. A., M. H. Jo, Y. G. Choi, J. Park, S. C. Kwon, S. Hohng, V. N. Kim, and J. S. Woo. 2015. 'Functional Anatomy of the Human Microprocessor', *Cell*, 161: 1374-87.

Nguyen, T. A., J. Park, T. L. Dang, Y. G. Choi, and V. N. Kim. 2018. 'Microprocessor depends on hemin to recognize the apical loop of primary microRNA', *Nucleic Acids Res*, 46: 5726-36.

Park, J. E., I. Heo, Y. Tian, D. K. Simanshu, H. Chang, D. Jee, D. J. Patel, and V. N. Kim. 2011. 'Dicer recognizes the 5' end of RNA for efficient and accurate processing', *Nature*, 475: 201-5.

Partin, A. C., T. D. Ngo, E. Herrell, B. C. Jeong, G. Hon, and Y. Nam. 2017. 'Heme enables proper positioning of Drosha and DGCR8 on primary microRNAs', *Nat Commun*, 8: 1737.

Partin, A. C., K. Zhang, B. C. Jeong, E. Herrell, S. Li, W. Chiu, and Y. Nam. 2020. 'Cryo-EM Structures of Human Drosha and DGCR8 in Complex with Primary MicroRNA', *Mol Cell*.

Peng, Y., and C. M. Croce. 2016. 'The role of MicroRNAs in human cancer', *Signal Transduct Target Ther*, 1: 15004.

Quinlan, A. R., and I. M. Hall. 2010. 'BEDTools: a flexible suite of utilities for comparing genomic features', *Bioinformatics*, 26: 841-2.

Rakheja, D., K. S. Chen, Y. Liu, A. A. Shukla, V. Schmid, T. C. Chang, S. Khokhar, J. E. Wickiser, N. J. Karandikar, J. S. Malter, J. T. Mendell, and J. F. Amatruda. 2014. 'Somatic mutations in DROSHA and DICER1 impair microRNA biogenesis through distinct mechanisms in Wilms tumours', *Nat Commun*, 2: 4802.

Ritchie, M. E., B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, and G. K. Smyth. 2015. 'limma powers differential expression analyses for RNA-sequencing and microarray studies', *Nucleic Acids Res*, 43: e47.

Rivera, M. N., W. J. Kim, J. Wells, D. R. Driscoll, B. W. Brannigan, M. Han, J. C. Kim, A. P. Feinberg, W. L. Gerald, S. O. Vargas, L. Chin, A. J. Iafrate, D. W. Bell, and D. A. Haber. 2007. 'An X chromosome gene, WTX, is commonly inactivated in Wilms tumor', *Science*, 315: 642-5.

Rodriguez, A., S. Griffiths-Jones, J. L. Ashurst, and A. Bradley. 2004. 'Identification of mammalian microRNA host genes and transcription units', *Genome Res*, 14: 1902-10.

Ruteshouser, E. C., S. M. Robinson, and V. Huff. 2008. 'Wilms tumor genetics: mutations in WT1, WTX, and CTNNB1 account for only about one-third of tumors', *Genes Chromosomes Cancer*, 47: 461-70.

Ruvkun, G. 2001. 'Molecular biology. Glimpses of a tiny RNA world', *Science*, 294: 797-9.

Sampath, P., D. K. Pritchard, L. Pabon, H. Reinecke, S. M. Schwartz, D. R. Morris, and C. E. Murry. 2008. 'A hierarchical network controls protein translation during murine embryonic stem cell self-renewal and differentiation', *Cell Stem Cell*, 2: 448-60.

Scott, R. H., A. Murray, L. Baskcomb, C. Turnbull, C. Loveday, R. Al-Saadi, R. Williams, F. Breatnach, M. Gerrard, J. Hale, J. Kohler, P. Lapunzina, G. A. Levitt, S. Picton, B. Pizer, M. D. Ronghe, H. Traunecker, D. Williams, A. Kelsey, G. M. Vujanic, N. J. Sebire, P. Grundy, C. A. Stiller, K. Pritchard-Jones, J. Douglas, and N. Rahman. 2012. 'Stratification of Wilms tumor by genetic and epigenetic analysis', *Oncotarget*, 3: 327-35.

Segers, H., M. M. van den Heuvel-Eibrink, R. D. Williams, H. van Tinteren, G. Vujanic, R. Pieters, K. Pritchard-Jones, N. Bown, Cancer Children's, Group Leukaemia, and U. K. Cancer Cytogenetics Group the. 2013. 'Gain of 1q is a marker of poor prognosis in Wilms' tumors', *Genes Chromosomes Cancer*, 52: 1065-74.

Sellier, C., V. J. Hwang, R. Dandekar, B. Durbin-Johnson, N. Charlet-Berguerand, B. P. Ander, F. R. Sharp, K. Angkustsiri, T. J. Simon, and F. Tassone. 2014. 'Decreased DGCR8 expression and miRNA dysregulation in individuals with 22q11.2 deletion syndrome', *PLoS One*, 9: e103884.

Sha, Y., D. Haensel, G. Gutierrez, H. Du, X. Dai, and Q. Nie. 2019. 'Intermediate cell states in epithelial-to-mesenchymal transition', *Phys Biol*, 16: 021001.

Shalet, S. M., B. Gibson, R. Swindell, and D. Pearson. 1987. 'Effect of spinal irradiation on growth', *Arch Dis Child*, 62: 461-4.

Shi, J., M. Dong, L. Li, L. Liu, A. Luz-Madrigal, P. A. Tsonis, K. Del Rio-Tsonis, and C. Liang. 2015. 'mirPRo-a novel standalone program for differential expression and variation analysis of miRNAs', *Sci Rep*, 5: 14617.

Shiohama, A., T. Sasaki, S. Noda, S. Minoshima, and N. Shimizu. 2003. 'Molecular cloning and expression analysis of a novel gene DGCR8 located in the DiGeorge syndrome chromosomal region', *Biochem Biophys Res Commun*, 304: 184-90.

Shukrun, R., N. Pode-Shakked, O. Pleniceanu, D. Omer, E. Vax, E. Peer, S. Pri-Chen, J. Jacob, Q. Hu, O. Harari-Steinberg, V. Huff, and B. Dekel. 2014. 'Wilms' tumor blastemal stem cells dedifferentiate to propagate the tumor bulk', *Stem Cell Reports*, 3: 24-33.

Sonn, G., and L. M. Shortliffe. 2008. 'Management of Wilms tumor: current standard of care', *Nat Clin Pract Urol*, 5: 551-60.

Spreafico, F., and F. F. Bellani. 2006. 'Wilms' tumor: past, present and (possibly) future', *Expert Rev Anticancer Ther*, 6: 249-58.

Stark, K. L., B. Xu, A. Bagchi, W. S. Lai, H. Liu, R. Hsu, X. Wan, P. Pavlidis, A. A. Mills, M. Karayiorgou, and J. A. Gogos. 2008. 'Altered brain microRNA biogenesis contributes to phenotypic deficits in a 22q11-deletion mouse model', *Nat Genet*, 40: 751-60.

Stefl, R., F. C. Oberstrass, J. L. Hood, M. Jourdan, M. Zimmermann, L. Skrisovska, C. Maris, L. Peng, C. Hofr, R. B. Emeson, and F. H. Allain. 2010. 'The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove', *Cell*, 143: 225-37.

Stockel, D., T. Kehl, P. Trampert, L. Schneider, C. Backes, N. Ludwig, A. Gerasch, M. Kaufmann, M. Gessler, N. Graf, E. Meese, A. Keller, and H. P. Lenhof. 2016. 'Multi-omics enrichment analysis using the GeneTrail2 web service', *Bioinformatics*, 32: 1502-8.

Stone, M. M., B. L. Beaver, C. C. Sun, and J. L. Hill. 1990. 'The nephroblastomatosis complex and its relationship to Wilms' tumor', *J Pediatr Surg*, 25: 933-7; discussion 37-8.

Szychot, E., J. Apps, and K. Pritchard-Jones. 2014. 'Wilms' tumor: biology, diagnosis and treatment', *Transl Pediatr*, 3: 12-24.

Termuhlen, A. M., J. M. Tersak, Q. Liu, Y. Yasui, M. Stovall, R. Weathers, M. Deutsch, C. A. Sklar, K. C. Oeffinger, G. Armstrong, L. L. Robison, and D. M. Green. 2011. 'Twenty-five year follow-up of

childhood Wilms tumor: a report from the Childhood Cancer Survivor Study', *Pediatr Blood Cancer*, 57: 1210-6.

Torrezan, G. T., E. N. Ferreira, A. M. Nakahata, B. D. Barros, M. T. Castro, B. R. Correa, A. C. Krepischi, E. H. Olivieri, I. W. Cunha, U. Tabori, P. E. Grundy, C. M. Costa, B. de Camargo, P. A. Galante, and D. M. Carraro. 2014. 'Recurrent somatic mutation in DROSHA induces microRNA profile changes in Wilms tumour', *Nat Commun*, 5: 4039.

Triboulet, R., H. M. Chang, R. J. Lapierre, and R. I. Gregory. 2009. 'Post-transcriptional control of DGCR8 expression by the Microprocessor', *RNA*, 15: 1005-11.

Valencia-Sanchez, M. A., J. Liu, G. J. Hannon, and R. Parker. 2006. 'Control of translation and mRNA degradation by miRNAs and siRNAs', *Genes Dev*, 20: 515-24.

van den Heuvel-Eibrink, M. M., J. A. Hol, K. Pritchard-Jones, H. van Tinteren, R. Furtwangler, A. C. Verschuur, G. M. Vujanic, I. Leuschner, J. Brok, C. Rube, A. M. Smets, G. O. Janssens, J. Godzinski, G. L. Ramirez-Villar, B. de Camargo, H. Segers, P. Collini, M. Gessler, C. Bergeron, F. Spreafico, N. Graf, and Group International Society of Paediatric Oncology - Renal Tumour Study. 2017. 'Position paper: Rationale for the treatment of Wilms tumour in the UMBRELLA SIOP-RTSG 2016 protocol', *Nat Rev Urol*, 14: 743-52.

Varum, S., A. S. Rodrigues, M. B. Moura, O. Momcilovic, C. A. th Easley, J. Ramalho-Santos, B. Van Houten, and G. Schatten. 2011. 'Energy metabolism in human pluripotent stem cells and their differentiated counterparts', *PLoS One*, 6: e20914.

Vujanic, G. M., and B. Sandstedt. 2010. 'The pathology of Wilms' tumour (nephroblastoma): the International Society of Paediatric Oncology approach', *J Clin Pathol*, 63: 102-9.

Vujanic, G. M., B. Sandstedt, D. Harms, A. Kelsey, I. Leuschner, J. de Kraker, and Siop Nephroblastoma Scientific Committee. 2002. 'Revised International Society of Paediatric Oncology (SIOP) working classification of renal tumors of childhood', *Med Pediatr Oncol*, 38: 79-82.

Walz, A. L., A. Ooms, S. Gadd, D. S. Gerhard, M. A. Smith, J. M. Guidry Auvil, D. Meerzaman, Q. R. Chen, C. H. Hsu, C. Yan, C. Nguyen, Y. Hu, R. Bowlby, D. Brooks, Y. Ma, A. J. Mungall, R. A. Moore, J. Schein, M. A. Marra, V. Huff, J. S. Dome, Y. Y. Chi, C. G. Mullighan, J. Ma, D. A. Wheeler, O. A. Hampton, N. Jafari, N. Ross, J. M. Gastier-Foster, and E. J. Perlman. 2015. 'Recurrent DGCR8, DROSHA, and SIX homeodomain mutations in favorable histology Wilms tumors', *Cancer Cell*, 27: 286-97.

Wan, G., Q. E. Lim, and H. P. Too. 2010. 'High-performance quantification of mature microRNAs by real-time RT-PCR using deoxyuridine-incorporated oligonucleotides and hemi-nested primers', *RNA*, 16: 1436-45.

Wang, B. 2013. 'Base Composition Characteristics of Mammalian miRNAs', *J Nucleic Acids*, 2013: 951570.

Wang, D., F. Bu, and W. Zhang. 2019. 'The Role of Ubiquitination in Regulating Embryonic Stem Cell Maintenance and Cancer Development', *Int J Mol Sci*, 20.

Wang, X. 2014. 'Composition of seed sequence is a major determinant of microRNA targeting patterns', *Bioinformatics*, 30: 1377-83.

Wang, Y., S. Baskerville, A. Shenoy, J. E. Babiarz, L. Baehner, and R. Blelloch. 2008. 'Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation', *Nat Genet*, 40: 1478-83.

Wang, Y., R. Medvid, C. Melton, R. Jaenisch, and R. Blelloch. 2007. 'DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal', *Nat Genet*, 39: 380-5.

Wang, Y., C. Melton, Y. P. Li, A. Shenoy, X. X. Zhang, D. Subramanyam, and R. Blelloch. 2013. 'miR-294/miR-302 promotes proliferation, suppresses G1-S restriction point, and inhibits ESC differentiation through separable mechanisms', *Cell Rep*, 4: 99-109.

Watson, J. A., K. Bryan, R. Williams, S. Popov, G. Vujanic, A. Coulomb, L. Boccon-Gibod, N. Graf, K. Pritchard-Jones, and M. O'Sullivan. 2013. 'miRNA profiles as a predictor of chemoresponsiveness in Wilms' tumor blastema', *PLoS One*, 8: e53417.

Wegert, J., N. Ishaque, R. Vardapour, C. Georg, Z. Gu, M. Bieg, B. Ziegler, S. Bausenwein, N. Nourkami, N. Ludwig, A. Keller, C. Grimm, S. Kneitz, R. D. Williams, T. Chagtai, K. Pritchard-Jones, P. van Sluis, R. Volckmann, J. Koster, R. Versteeg, T. Acha, M. J. O'Sullivan, P. K. Bode,

F. Niggli, G. A. Tytgat, H. van Tinteren, M. M. van den Heuvel-Eibrink, E. Meese, C. Vokuhl, I. Leuschner, N. Graf, R. Eils, S. M. Pfister, M. Kool, and M. Gessler. 2015. 'Mutations in the SIX1/2 pathway and the DROSHA/DGCR8 miRNA microprocessor complex underlie high-risk blastemal type Wilms tumors', *Cancer Cell*, 27: 298-311.

Wegert, J., C. Vokuhl, B. Ziegler, K. Ernestus, I. Leuschner, R. Furtwangler, N. Graf, and M. Gessler. 2017. 'TP53 alterations in Wilms tumour represent progression events with strong intratumour heterogeneity that are closely linked but not limited to anaplasia', *J Pathol Clin Res*, 3: 234-48.

Wegert, J., L. Zauter, S. Appenzeller, C. Otto, S. Bausenwein, C. Vokuhl, K. Ernestus, R. Furtwangler, N. Graf, and M. Gessler. 2020. 'High-risk blastemal Wilms tumor can be modeled by 3D spheroid cultures in vitro', *Oncogene*, 39: 849-61.

Weirich, A., I. Leuschner, D. Harms, G. M. Vujanic, J. Troger, U. Abel, N. Graf, D. Schmidt, R. Ludwig, and P. A. Voute. 2001. 'Clinical impact of histologic subtypes in localized non-anaplastic nephroblastoma treated according to the trial and study SIOP-9/GPOH', *Ann Oncol*, 12: 311-9.

Weirich, A., R. Ludwig, N. Graf, U. Abel, I. Leuschner, G. M. Vujanic, O. Mehls, J. Boos, J. Beck, B. Royer-Pokora, and P. A. Voute. 2004. 'Survival in nephroblastoma treated according to the trial and study SIOP-9/GPOH with respect to relapse and morbidity', *Ann Oncol*, 15: 808-20.

Wright, K. D., D. M. Green, and N. C. Daw. 2009. 'Late effects of treatment for wilms tumor', *Pediatr Hematol Oncol*, 26: 407-13.

Wu, M. K., N. Sabbaghian, B. Xu, S. Addidou-Kalucki, C. Bernard, D. Zou, A. E. Reeve, M. R. Eccles, C. Cole, C. S. Choong, A. Charles, T. Y. Tan, D. M. Iglesias, P. R. Goodyer, and W. D. Foulkes. 2013. 'Biallelic DICER1 mutations occur in Wilms tumours', *J Pathol*, 230: 154-64.

Yeom, K. H., Y. Lee, J. Han, M. R. Suh, and V. N. Kim. 2006. 'Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing', *Nucleic Acids Res*, 34: 4622-9.

Yiallouros, Maria. 2017. 'Wilms Tumour (Nephroblastoma) – Brief information', Accessed November 6. https://www.kinderkrebsinfo.de/diseases/solid_tumours/wilms_tumour_nephroblastoma/pohwilms_patinfokurz120120611/index_eng.html#e103059.

Ying, Q. L., J. Wray, J. Nichols, L. Batlle-Morera, B. Doble, J. Woodgett, P. Cohen, and A. Smith. 2008. 'The ground state of embryonic stem cell self-renewal', *Nature*, 453: 519-23.

Yu, X., Z. Li, M. T. Chan, and W. K. Wu. 2016. 'The roles of microRNAs in Wilms' tumors', *Tumour Biol*, 37: 1445-50.

Zhang, M., S. Xiang, H. Y. Joo, L. Wang, K. A. Williams, W. Liu, C. Hu, D. Tong, J. Haakenson, C. Wang, S. Zhang, R. E. Pavlovicz, A. Jones, K. H. Schmidt, J. Tang, H. Dong, B. Shan, B. Fang, R. Radhakrishnan, P. M. Glazer, P. Matthias, J. Koomen, E. Seto, G. Bepler, S. V. Nicosia, J. Chen, C. Li, L. Gu, G. M. Li, W. Bai, H. Wang, and X. Zhang. 2014. 'HDAC6 deacetylates and ubiquitinates MSH2 to maintain proper levels of MutSalpha', *Mol Cell*, 55: 31-46.

Zhdanova, O., S. Srivastava, L. Di, Z. Li, L. Tchelebi, S. Dworkin, D. B. Johnstone, J. Zavadil, M. M. Chong, D. R. Littman, L. B. Holzman, L. Barisoni, and E. Y. Skolnik. 2011. 'The inducible deletion of Drosha and microRNAs in mature podocytes results in a collapsing glomerulopathy', *Kidney Int*, 80: 719-30.

Zuker, M. 2003. 'Mfold web server for nucleic acid folding and hybridization prediction', *Nucleic Acids Res*, 31: 3406-15.

8 Abbreviations

A

aa	Amino acid
A	Adenine
ASP	Allele-specific PCR

B

bp	Base pair
BrdU	Bromodeoxyuridine

C

C	Cytosine
cDNA	Complementary DNA
ct-value	Cycle threshold value

D

DEPC	Diethyl pyrocarbonate
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
dox	Doxycycline
dsRBD	Double stranded RNA binding domain

E

EBs	Embryoid bodies
eGFP	Enhanced fluorescent protein
EMT	Epithelial-to-mesenchymal-transition
ESCC miRNAs	ESC cell cycle miRNAs
ESCs	Embryonic stem cells

F

FCS	Fetal calf serum

G

G	Guanine
GFP	Green fluorescent protein

I

IRES	Internal ribosome entry site

K

KO	Knockout

L

LIF	Leukemia Inhibitory Factor

M

MC	Microprocessor complex
MEFs	Mouse embryonic fibroblasts
mESCs	Mouse embryonic stem cells
miRNA	MicroRNA
mRNA	Messenger RNA
MTIs	MiRNA target interactions
MTT	(3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
mut	Mutant

N

NB	Nephroblastomatosis
n.e.	Not expressed
NEAA	Non-essential amino acids
nt	Nucleotide

P

PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEI	Polyethylenimine
PI	Propidium iodide
POD	Peroxidase
pre-miRNA	Precursor miRNA
pri-miRNA	Primary miRNA

Q

qPCR	Quantitative real-time PCR

R

RNA	Ribonucleic acid
RNA-Seq	RNA-Sequencing
pRPM	Proportional reads per million

S

SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis

T

T	Thymine
TPM	Transcripts per million
Tris	Tris(hydroxymethyl)aminomethan

U

UTR	Untranslated Region

W

WT	Wilms tumor
wt	Wild-type

Y

YFP	Yellow fluorescent protein

9 Supplement

9.1 Supplementary data

Table S1: Mutation frequency of *DROSHA* E1147K and *DGCR8* E518K in WTs according to histological subtype

Histotype		Completely Necrotic	Epithelial Type	Stromal Type	Mixed Type	Regressive Type	Focal Anaplasia	Blastemal. pre-chemotherapy	Blastemal Type	Diffuse Anaplasia	Nephroblastomatosis	Total
DROSHA	Mutant	1/25	0/64	1/76	1/101	5/102	2/9	1/26	4/77	2/27	0/18	17/525
	Mutant (%)	4.0	0.0	1.3	1.0	4.9	22.2	3.8	5.2	7.4	0.0	3.2
DGCR8	Mutant	0/26	0/62	0/74	8/213	10/235	0/14	0/26	5/77	1/27	0/23	24/777
	Mutant (%)	0.0	0.0	0.0	3.8	4.3	0.0	0.0	6.5	3.7	0.0	3.1

Table S2: Death and relapse frequencies among DROSHA/DGCR8-mutated and -non-mutated WT cases

	death		relapse/metastasis	
DROSHA E1147K	yes	no	yes	no
yes	3	14	3	12
no	38	451	78	400
DGCR8 E518K				
yes	0	23	3	21
no	55	678	107	604

Table S3: MiRNAs higher expressed in DGCR8-KO relative to DGCR8-wt most likely derived from feeder cells

ID	wt-dox	wt+dox	KO+dox	E518K-dox	E518K+dox	PSNL-dox	PSNL+dox
mmu-miR-135a-1-3p	4.56	4.39	7.00	6.83	6.53	3.62	3.71
mmu-miR-149-3p	3.37	3.61	5.61	5.81	5.49	4.13	4.26
mmu-miR-125a-3p	4.88	3.73	5.75	5.72	5.48	4.64	4.74
mmu-miR-139-3p	3.67	3.77	5.73	6.06	5.70	2.81	2.87
mmu-miR-483-5p	3.91	4.00	6.03	5.96	5.55	3.51	3.65
mmu-miR-351-3p	2.59	3.43	5.54	5.98	5.61	2.98	3.10
mmu-miR-328-5p	3.73	3.35	5.83	5.81	5.39	4.48	4.57
mmu-let-7i-5p	3.01	2.78	4.75	4.65	3.71	9.35	9.39
mmu-let-7b-5p	3.63	2.87	4.71	4.60	3.83	8.64	8.66
mmu-miR-3058-3p	3.81	3.03	4.53	4.41	4.36	3.05	3.11
mmu-miR-574-3p	3.63	2.53	4.49	3.94	3.57	3.98	3.60
mmu-miR-30c-1-3p	3.20	2.79	4.21	3.88	3.30	2.10	1.90
mmu-miR-452-5p	1.68	1.53	2.90	3.72	3.45	5.42	4.29
mmu-miR-874-3p	1.70	2.28	3.82	4.38	4.24	2.82	2.83
mmu-miR-669c-3p	1.94	2.16	3.47	3.56	3.32	2.59	2.57
mmu-miR-770-3p	1.26	2.16	3.44	4.20	3.95	2.44	2.48
mmu-miR-1306-3p	3.76	3.06	5.07	4.78	4.59	2.82	2.88
mmu-miR-504-3p	4.24	3.89	6.20	5.42	4.87	1.38	1.41
mmu-miR-669e-5p	4.68	3.80	5.82	5.18	4.70	2.20	2.23
mmu-miR-669o-5p	4.11	3.20	5.00	4.50	3.88	1.85	1.96
mmu-miR-3572-5p	1.78	1.92	2.96	3.61	3.39	2.56	2.55

Microarray results of DGCR8-wt, -KO, -E518K mESCs and PSNL feeder cells upon induction with doxycycline. High-confidence miRNAs that have high expression ratios (≥1.5x) in DGCR8-KO relative to DGCR8-wt were also expressed in PSNL feeder cells (n=21/27). DGCR8-E518K similarly showed high expression of these miRNAs. Technical triplicates of DGCR8-wt/-E518K and duplicates of DGCR8-KO and PSNL feeder cells were hybridized onto Agilent miRNA arrays. Mean expression values of technical replicates were normalized using quantile-normalization. MiRNAs are shown with a present call of ≥50%.

Table S4: RNA distribution in DGCR8-knockout, -wild-type, and -E518K mESCs

Statistic of read counts											
	1	2	3	4	5	6	7	8	9	10	11
	KO#1+dox	KO#2+dox	KO#3+dox	wt#10	E518K#8	wt#10+dox	wt#12+dox	wt#16+dox	E518K#8+dox	E518K#13+dox	E518K#15+dox
Total reads	20102548	18407940	17975494	10411340	20674923	9231668	14721249	19591452	11219863	7868890	8890643
Mapped reads	10925369	9404835	9494599	6041158	10811621	6264075	10381883	12515167	6047454	4648299	5202695
Alignment rates in %	54.35	51.09	52.82	58.02	52.29	67.85	70.52	63.88	53.9	44.37	58.52
piRNA counts	2118285	1817426	1928829	650909	1645961	286414	323512	403044	376755	827594	771455
piRNA in %	19.39	19.32	20.32	10.77	15.22	4.57	3.12	3.22	6.23	17.80	14.83
miRNA counts	208846	166676	169316	1421247	264507	3394216	6128734	5933843	1797240	1050564	1418587
miRNA in %	1.91	1.77	1.78	23.53	2.45	54.19	59.03	47.41	29.72	22.60	27.27
miRNA transcript counts	211096	168420	171066	1424318	266919	3403628	6145995	5964800	1804012	1054145	1422153
miRNA transcripts in %	1.93	1.79	1.80	23.58	2.47	54.34	59.20	47.66	29.83	22.68	27.33
Unknown counts	8595988	7418989	7394704	3965931	8898741	2574033	3912376	6147323	3866687	2766560	3009087
Unknown in %	78.68	78.88	77.88	65.65	82.31	41.09	37.68	49.12	63.94	59.52	57.84

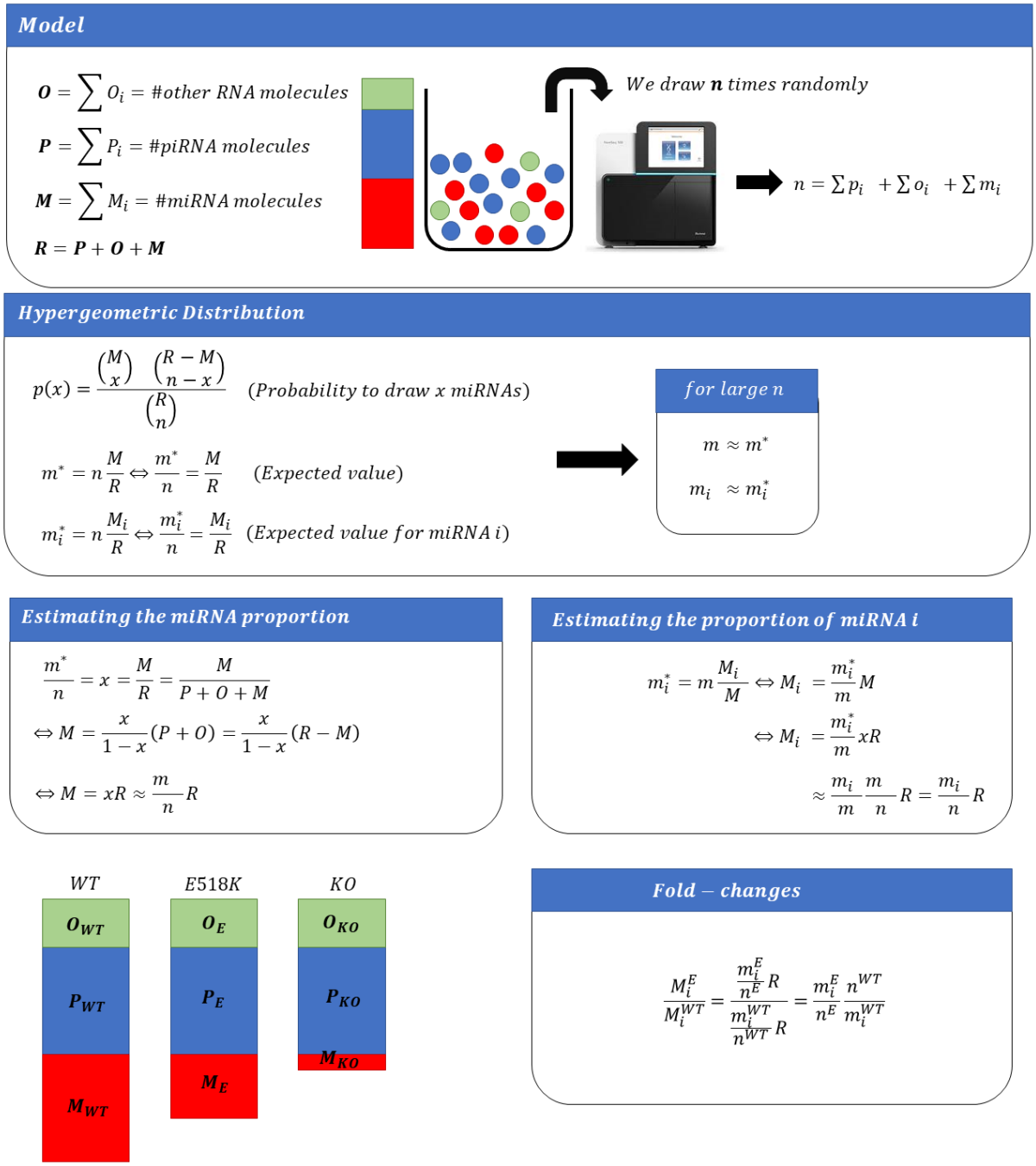


Figure S1: Hypergeometric distribution analysis to model miRNA-Seq data

Proportion of miRNAs (M) in a mixture of piRNAs (P) and other RNA molecules (O) ($P + O + M$) was modelled using the hypergeometric distribution model. Probability to randomly draw x miRNAs out of $P + O + M$ was calculated and used for subsequent calculations between DGCR8-wild-type, -E515K, and -knockout.

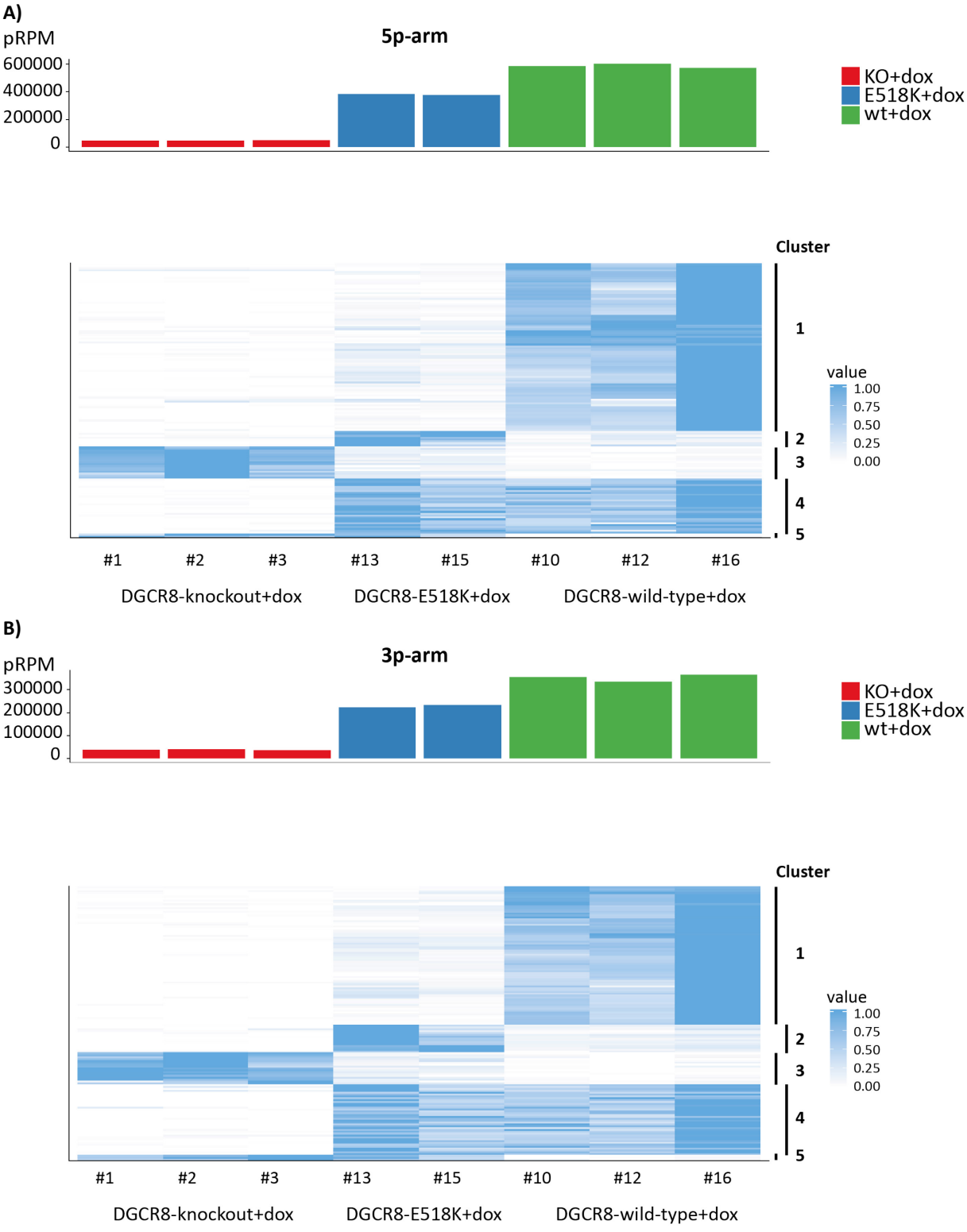


Figure S2: Comparison of miRNA expression between DGCR8-knockout, -wild-type and -E518K mESCs
(A) 5p-arm, and (B) 3p-arm miRNAs. Bar charts showing differences in miRNA expression levels between DGCR8-KO, -wt and -E518K upon induction. Expression values are presented as proportional reads per million (pRPM) normalized to piRNAs. Three biological replicates were used per group (except E518K–only two samples). Heatmaps showing clusters of differentially expressed miRNAs in DGCR8-KO, -wt and -E518K with a pRPM≥20. Values represent scaled miRNA expression ranging from 0–1. Cluster 1: KO≤0.25, E518K≤0.25, wt≥0.50 (n=82/n=71); Cluster 2: KO≤0.25, E518K≥0.50, wt≤0.25 (n=9/n=15); Cluster 3: KO≥0.50, E518K≤0.25, wt≤0.25 (n=16/n=10); Cluster 4: KO≤0.25, E518K≥0.50, wt≥0.50 (n=31/n=35); Cluster 5: KO≥0.50, E518K≥0.50, wt≤0.25 (n=2/n=3).

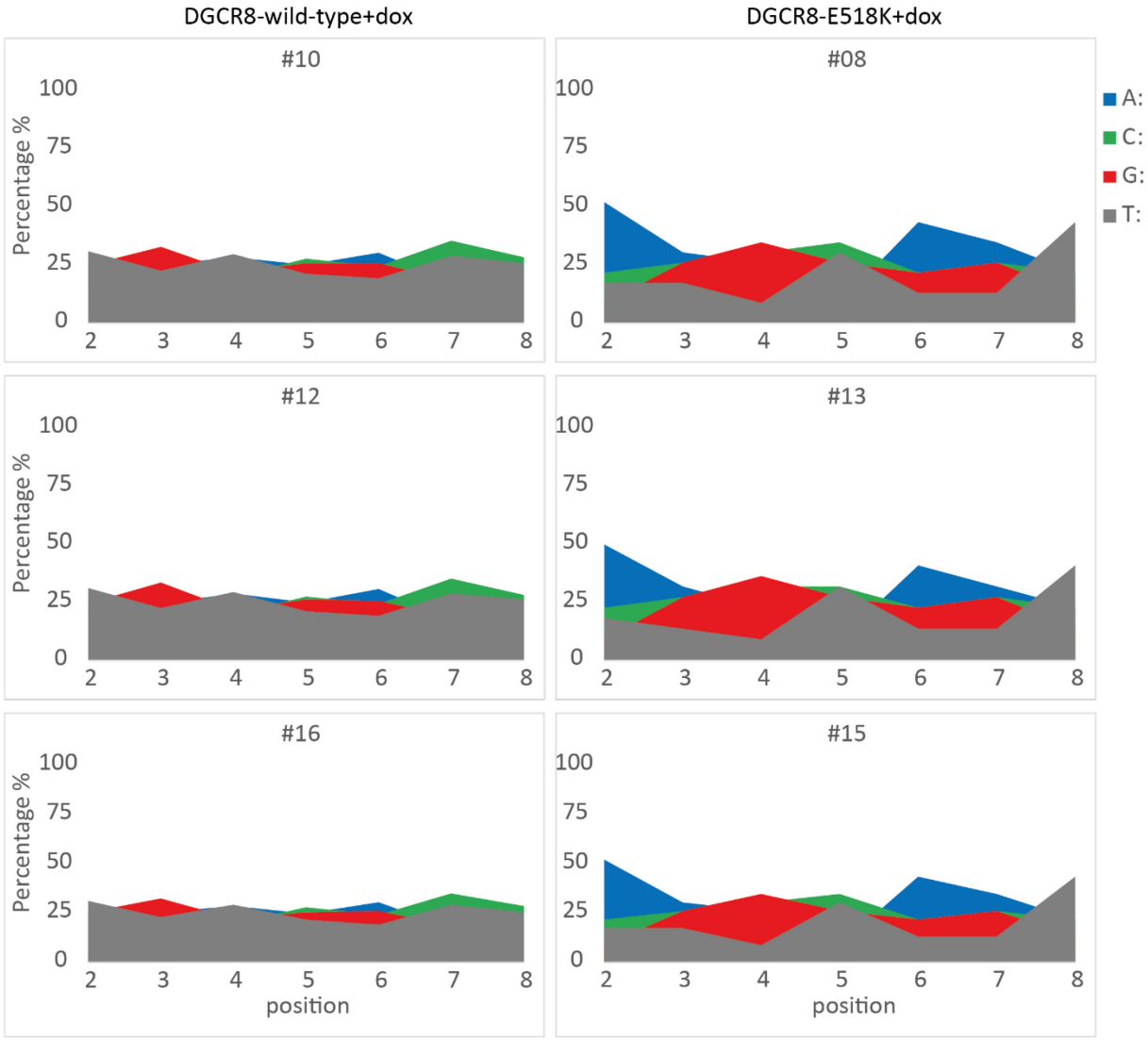


Figure S3: Base composition of the seed sequence in DGCR8-wt, and -E518K mESCs without read counts
Area charts showing distribution of A, C, G, and T in miRNAs upregulated in DGCR8-wild-type or -E518K mESCs encompassing the seed sequence (position 2–8 at the 5'end). Percentages were calculated based on mapped reads per nucleotide position. Three biological replicates are shown for each condition.

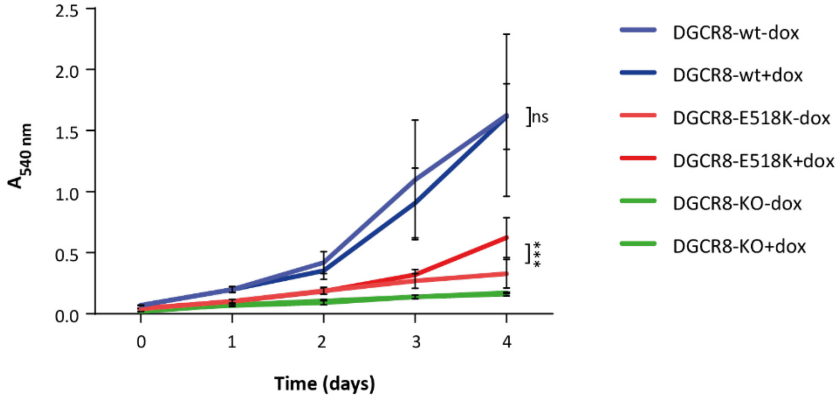


Figure S4: Effect of the DGCR8 E518K mutation on growth in mESCs

Mean growth rate of three independent biological replicates measured by MTT assay. Error bars represent SD within each set of biological replicates. 2way ANOVA was performed to identify significant differences between induced and uninduced DGCR8-wild-type or -E518K (adj. p-value ≤ 0.001 ***; > 0.05 ns). Significance is shown here for day 4. For each sample five technical replicates were used. Treatment with doxycycline (300 ng/mL) was started one day after plating. MTT assay was performed for each timepoint (day 0, 1, 2, 3, 4). Absorbance of purple formazan solution was measured at 540 nm.

A)

B)

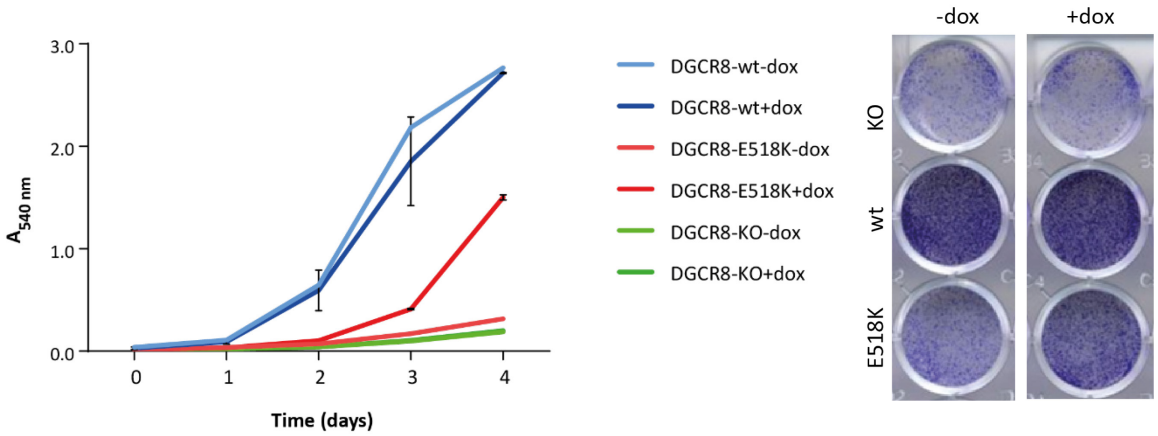


Figure S5: Effect of the DGCR8 E518K mutation on growth in mESCs

(A) Mean growth rate of two independent biological replicates measured by crystal violet assay (except DGCR8-KO, uninduced DGCR8-wild-type and -E518K–only one sample each). Error bars represent SD within each set of biological replicates. For each sample three technical replicates were used. Treatment with doxycycline (300 ng/mL) was started one day after plating. Crystal violet assay was performed for each timepoint (day 0, 1, 2, 3, 4). Absorbance was measured at 540 nm.

(B) Image of crystal violet stained cells on day 4. DGCR8-knockout, -wild-type and -E518K upon dox treatment.

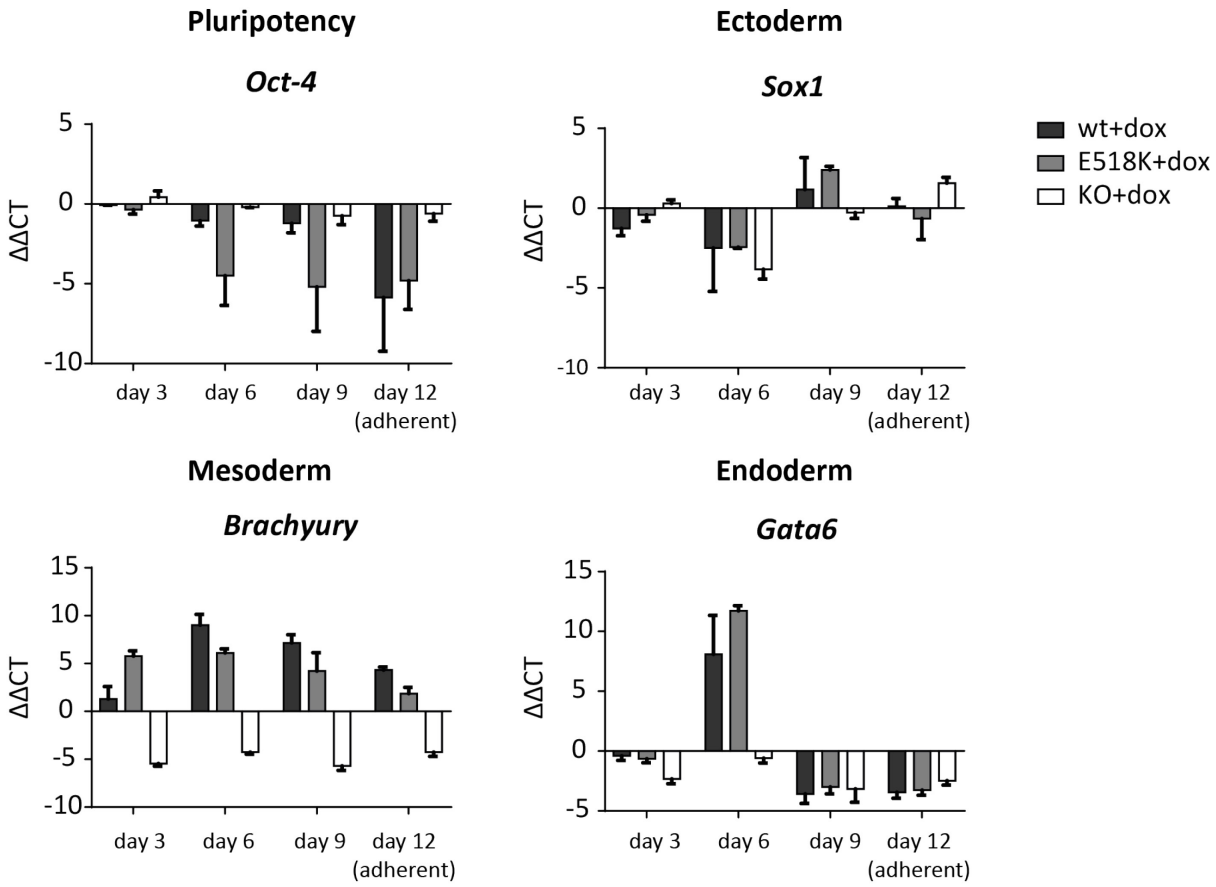


Figure S6: Embryoid body differentiation of DGCR8-wild-type, -E518K, and -knockout mESCs

qPCR results of pluripotency marker *Oct-4* and differentiation markers (ectoderm: *Sox1*; mesoderm: *Brachyury*; endoderm: *Gata6*) on day 3, 6, and 9 of EB differentiation. On day 6, EBs were collected and grown as monolayers for additional 6 days (day 12 adherent). CT-values were normalized to the housekeeping gene *Hprt.* Bars represent mean ΔΔCT-values relative to DGCR8-wild-type mESCs on day 0. Error bars indicate SD of three biological replicates.

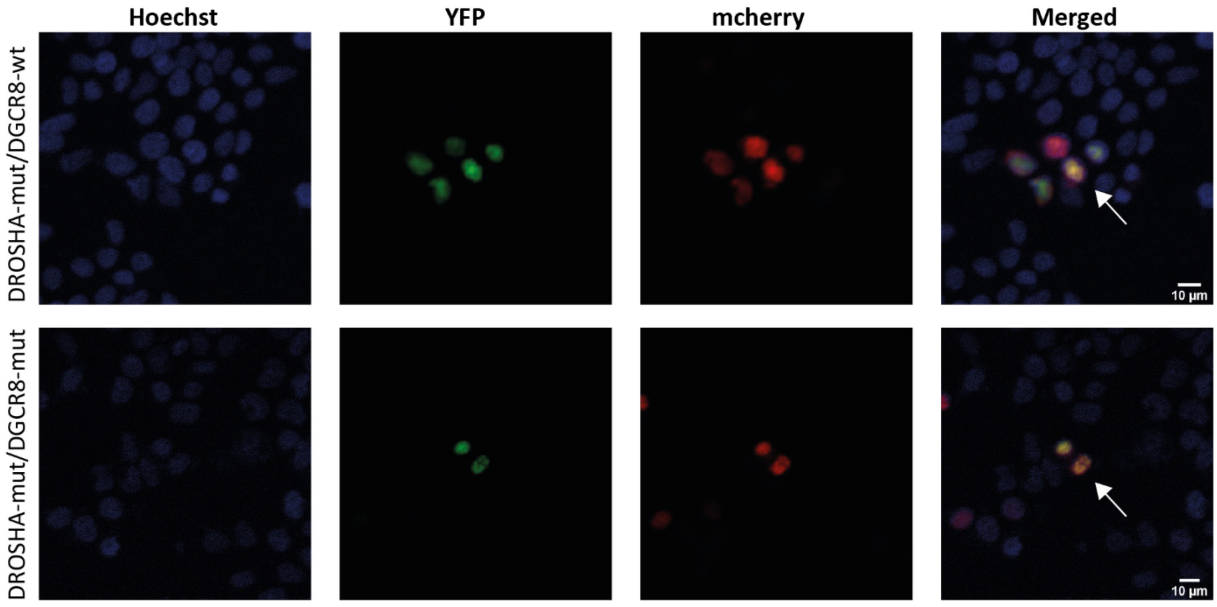


Figure S7: Nuclear localization of the microprocessor complex DROSHA/DGCR8

Transiently co-transfected HEK293T cells with pCS2P-HA2-mcherry-hDGCR8 wild-type (upper row) or E518K (bottom row) and pcDNA3.1-YFP-DROSHA-E1147K-flag. Localization of the nucleus was shown using Hoechst staining. DGCR8 E518K and DROSHA E1147K were found to be both restricted to the nucleus. Arrows indicate nuclear co-localization of DGCR8 and DROSHA.

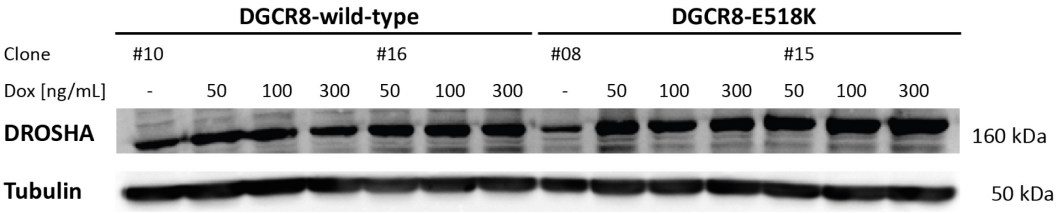


Figure S8: Posttranscriptional stabilization of DROSHA protein by DGCR8 wild-type and E518K

α-DROSHA Western blot analysis of DGCR8-wild-type and -E518K mutant mESCs upon induction with different concentrations of doxycycline (0, 50, 100 and 300 ng/mL) for 48h. For wild-type and E518K two biological replicates were used (except uninduced condition–only one sample).

9.2 Figure Index

9.3 Table Index

9.4 Publications

Wegert, J., N. Ishaque, **R. Vardapour**, C. Georg, Z. Gu, M. Bieg, B. Ziegler, S. Bausenwein, N. Nourkami, N. Ludwig, A. Keller, C. Grimm, S. Kneitz, R. D. Williams, T. Chagtai, K. Pritchard-Jones, P. van Sluis, R. Volckmann, J. Koster, R. Versteeg, T. Acha, M. J. O'Sullivan, P. K. Bode, F. Niggli, G. A. Tytgat, H. van Tinteren, M. M. van den Heuvel-Eibrink, E. Meese, C. Vokuhl, I. Leuschner, N. Graf, R. Eils, S. M. Pfister, M. Kool, and M. Gessler. 2015. 'Mutations in the SIX1/2 pathway and the DROSHA/DGCR8 miRNA microprocessor complex underlie high-risk blastemal type Wilms tumors', *Cancer Cell*, 27: 298-311.

Vardapour, R., T. Kehl, H.-P. Lenhof, S. Kneitz, M. Gessler. 2020. 'The DGCR8 E518K mutation found in Wilms tumors leads to a partial miRNA processing defect that alters gene expression and biological processes in embryonic stem cells', *in preparation.*

9.5 Oral presentations and posters

10/2014 EUREKA 9th International GSLS Students' Symposium, Wuerzburg, Germany
Poster: "Mutations in the DROSHA/DGCR8 miRNA microprocessor complex in high-risk blastemal Wilms tumor"

03/2015 AEK 18th International Cancer Congress, Heidelberg, Germany
Poster: "Functional analysis of *DROSHA* and *DGCR8* mutations in Wilms tumor"

06/2016 6th International Tuebingen-Symposium on Pediatric Solid Tumors, Tuebingen, Germany
Talk: "Functional analysis of *DROSHA* and *DGCR8* mutations in Wilms tumor"

03/2017 GSLS Students' Winter Retreat, Kochel am See, Germany
Poster: "Functional analysis of DROSHA and DGCR8 mutations in Wilms tumor"

06/2017 GSLS Students' Summer Retreat, Rothenfels, Germany
Talk: "Mutations in the DROSHA/DGCR8 miRNA microprocessor complex in high-risk blastemal Wilms tumor"

03/2018 GSLS Students' Winter Retreat, Fuessen, Germany
Talk: "Mutations in the DROSHA/DGCR8 miRNA microprocessor complex in high-risk blastemal Wilms tumor"

9.6 Curriculum vitae

9.7 Acknowledgement

An erster Stelle gilt mein Dank meinem Doktorvater Prof. Dr. Manfred Gessler, der es mir ermöglichte an diesem Projekt zu arbeiten, und der mich während meiner Promotion über die Jahre hinweg begleitet hat. Er hatte immer ein offenes Ohr für wissenschaftliche Fragestellungen und Diskussionen, die mit dem Projekt einhergingen und gab immer konstruktive Ratschläge für anfallende Herausforderungen.

Auch möchte ich mich bei Prof. Dr. Eva Geissinger und Prof. Dr. Manfred Alsheimer bedanken, dass sie als Teil des Prüfungskomitees diese Arbeit über die Jahre begleitet und unterstützt haben.

Darüber hinaus danke ich Prof. Dr. Alexander Buchberger für die Übernahme des Prüfungsvorsitzes.

Zudem möchte ich mich bei Dr. Susanne Kneitz und Tim Kehl für die bioinformatischen Auswertungen der Sequenzierdaten bedanken. Vor allem die ausgiebigen wissenschaftlichen Diskussionen und Gespräche haben in besonderem Maße zum Gelingen dieser Arbeit beigetragen.

Mein herzlicher Dank geht zudem an Dr. Gabriele Blum-Oehler und auch an die Mitarbeiter der *Graduate School of Life Sciences* (GSLS). Durch die zahlreichen Fortbildungsmöglichkeiten und Veranstaltungen konnte ich mich während meiner Promotion, neben meinen wissenschaftlichen Fähigkeiten, auch persönlich weiterentwickeln. Vor allem die Möglichkeit erhalten zu haben als Chair das EUREKA Symposium mit zu organisieren war eine bereichernde Erfahrung. Auch für die Bereitstellung ihrer *Think Tanks* möchte ich mich bei der GSLS bedanken.

Allen Arbeitskollegen aus der Entwicklungsbiochemie und Physiologischen Chemie danke ich für die angenehme Arbeitsatmosphäre und die langjährige gute Zusammenarbeit. Mein besonderer Dank gilt Barbara, die mir immer als Kollegin und Freundin zur Seite stand. Auch bei meinen Mitdoktoranden, vor allem Ovidio und Philip, möchte ich mich für die interessanten und netten Gespräche im Denkraum und in den Pausen bedanken. Insbesondere bei Philip für die gemeinsame Organisation des EUREKA Symposiums und darüber hinaus für die vielen Unternehmungen abseits der Laborarbeit. Mein Dank geht zudem an Anja für die Klonierung der Vektoren und ihre praktischen Ratschläge im Labor. Ein besonderes Dankeschön geht an Jenny, die mir stets bei meinen Fragestellungen mit gutem Rat zur Seite stand.

Mein Dank gilt auch all meinen Freunden, die immer für mich da waren und für die vielen schönen gemeinsamen Momente. Danke vor allem an Chriszandro Hofmeister für die motivierenden Gespräche.

Nun möchte ich mich ganz herzlich bei meinen Eltern und meinem Bruder bedanken, dass sie immer an mich geglaubt und mich bedingungslos unterstützt haben. Ohne sie wäre das alles nicht möglich gewesen. Auch möchte ich mich bei Harald herzlichst bedanken, dass er in den herausforderndsten Phasen meines Lebens immer an meiner Seite war, mich liebevoll unterstützt und motiviert hat, nach vorne zu blicken. Danke schön!

9.8 Affidavit

I hereby confirm that my thesis entitled “Mutations in the DROSHA/DGCR8 microprocessor complex in high-risk blastemal Wilms tumor” is the result of my own work. I did not receive any help or support from commercial consultants. All sources and/or materials applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Place, Date Signature

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, die Dissertation „Mutationen des DROSHA/DGCR8 Mikroprozessor-Komplexes in blastemreichen Hochrisiko-Wilms Tumoren“ eigenständig, d.h. insbesondere selbständig und ohne Hilfe eines kommerziellen Promotionsberaters angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Ort, Datum Unterschrift