8-2015

IDENTIFICATION OF FAMILIAL WILMS TUMOR PREDISPOSITION GENES USING WHOLE GENOME SEQUENCING

Timothy B. Palculict

IDENTIFICATION OF FAMILIAL WILMS TUMOR PREDISPOSITION GENES USING WHOLE GENOME SEQUENCING

A

DISSERTATION

Presented to the Faculty of
The University of Texas
Health Science Center at Houston
and
The University of Texas
MD Anderson Cancer Center
Graduate School of Biomedical Sciences

in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

by

Timothy Blake Palculict, M.S.
Houston, Texas

August 2015

Acknowledgments

I would first like to thank my advisor, Vicki Huff, for giving me the opportunity to join her laboratory. Her guidance, advice and support over the past few years has provided the perfect training environment allowing me to grow as a scientist. I don't think I could have picked a better lab to continue my training. I would also like to thank E. Cristy Ruteshouser for her guidance while working in the Huff Lab.

I would also like to thank my family for their constant support throughout my graduate school career. My wife, Meagan, has provided endless encouragement and patience over the past eight years allowing me to follow my dream. My parents, Tim and Kim Palculict, and siblings, Kent Palculict and Erin Palculict-George, have always been my biggest supporters and I can't thank them enough.

WHOLE GENOME SEQUENCING TO IDENTIFY FAMILIAL WILMS TUMOR PREDISPOSITION GENES

Timothy Blake Palculict

Advisory Professor: Vicki Huff, Ph.D.

Wilms tumor, a childhood tumor arising from undifferentiated renal mesenchyme, is diagnosed in North America at a frequency of 1 in 10,000 live births and accounts for 5% of all pediatric cancers. The etiology of Wilms tumor is heterogeneous with multiple genes known to have an effect on Wilms tumor development; however, these genes are rarely associated with familial Wilms tumor. Gene mutations in *WT1*, *WTX*, *CTNNB1* and *TP53* are observed in a third of sporadic tumors, while the causative gene(s) responsible for familial Wilms tumor are largely unknown. Approximately 2% of Wilms tumor patients have a family history of Wilms tumor. Familial predisposition is inherited in an autosomal dominant manner and demonstrates incomplete penetrance, estimated to be 30%.

Whole genome sequencing of eleven individuals in three Wilms tumor families was performed to identify the gene(s) responsible for genetic predisposition to Wilms tumor in these families, and to increase our understanding of a genetically heterogeneous cancer. Variant calling was performed by GATK, FamSeq and CASAVA to detect single nucleotide variations (SNVs), insertions/deletions (indels) and structural variants. Shared variants within each family were identified and prioritized based on expression in fetal kidney, minor allele frequency less than 1%, and the predicted functional significance of the alteration on protein structure.

Candidate mutations were then assessed in each family for cosegregation with the affected/obligate carrier status. To examine the sequence integrity and frequency of candidate gene mutations in Wilms tumor, a comprehensive mutation analysis of each candidate gene, including Sanger sequencing of the entire coding region and quantitative PCR (qPCR) to detect exonic insertions/deletions, was performed using genomic DNA from 47 additional Wilms tumor families and tumor DNA from 175 sporadic Wilms tumors. As a result, a recurrent 570kb germline duplication including *NBAS, DDX1, MYCN* and *MYCNOS* was identified in 4% of families, as well as germline missense mutations in *DICER1* (4% of families) and *SPHK2*. Functional studies demonstrate that the missense mutation identified in *SPHK2* does not have an effect on apoptosis or proliferation, but instead alters SPHK2 subcellular localization and expression of target genes.

Our analysis resulted in the successful identification of Wilms tumor predisposition genes in each family, and provided support for the involvement of *DICER1* and *DDX1/MYCN* in Wilms tumor development. Additionally, our identification of *SPHK2* as the 19q familial Wilms tumor predisposition gene has opened up new avenues of research into the role of sphingolipid signaling in nephrogenesis and cancer development. Significantly, the communication of these results directly to patients within each family allows for the continued monitoring of new family members for the potential early detection of not only Wilms tumor, but other phenotypes associated with the specific gene mutations.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1. INTRODUCTION

1.1. Understanding Wilms tumor biology

Wilms tumor is an embryonal renal malignancy that is predominately diagnosed in children before 10 years of age.[1] It is the most common pediatric kidney cancer, occurring in 1 in 10,000 live births in North America. Unilateral cases represent 90-95% of Wilms tumor diagnoses and have a mean age of diagnosis occurring around 43 months. Bilateral cases are usually diagnosed at 33-38 months.[1] It was this kidney malignancy, in addition to neuroblastoma and retinoblastoma, which constituted the three tumor types studied in the establishment of the "two-hit" model of tumor suppression.[2] Although the survival rate for Wilms tumor is over 90% for confined disease, the development of secondary malignancies due to current therapy is common. Thus, identification of the genes and molecular pathways driving Wilms tumorigenesis is critical for 1) our understanding of the biology of Wilms tumor and other embryonal tumors, and 2) the development of targeted therapies aimed at reducing cancer mortality, while concurrently eliminating harmful side effects.

Wilms tumor is thought to originate from pluripotent embryonic renal progenitor cells that are unable to differentiate normally. Although the majority arise sporadically, roughly 2% of patients have a family history of Wilms tumor. In these instances, an earlier age of diagnosis and increased prevalence of bilateral tumors is typically observed, although this varies between families.[1-4] Until recently, the genetic etiology of most cases of non-syndromic Wilms tumor has remained a mystery, as traditional candidate gene approaches have been inefficient in revealing causal elements of this genetically heterogeneous cancer. High-throughput next-generation sequencing has represented a turning point in the identification of disease genes with the ability to

screen large numbers of genes simultaneously in an unbiased manner. Using this technology, surveying the genetic landscape of Wilms tumor has just begun to unravel the mutation events responsible for both sporadic and familial cases. However, the identification of "Wilms tumor genes" is far from complete. While the genes responsible for sporadic Wilms tumor cases are known in approximately 50% of cases, the cause of the majority of familial Wilms tumor cases is largely unknown.

1.1.1. Kidney Development

Wilms tumors originate from undifferentiated renal mesenchyme and represent an example of aberrant kidney development resulting in cancer. Kidney development commences at approximately day E36 in humans (E10.5 in mice), and is the result of complex reciprocal interactions between two tissues derived from the intermediate mesoderm: the metanephric mesenchyme, which differentiates into stromal cells and epithelial components of the nephron, and the ureteric bud, which gives rise to the collecting duct system.[5-6] Signals produced from the metanephric mesenchyme induce the ureteric bud to bifurcate as it invades the mesenchyme, resulting in the induction of mesenchyme condensation along the surface of the bud (Figure 1). A portion of the condensed mesenchyme forms pretubular aggregates adjacent and interior to the tips of the branching ureteric bud. These pretubular aggregates then undergo a mesenchymal-to-epithelial transition to form the epithelia (comma- and S-shaped bodies), which later develop into the nephron.[6] The stromal elements, derived from cells of the metanephric mesenchyme that were not induced to condense and epithelialize, develop in the periphery of the mesenchyme and between the branches

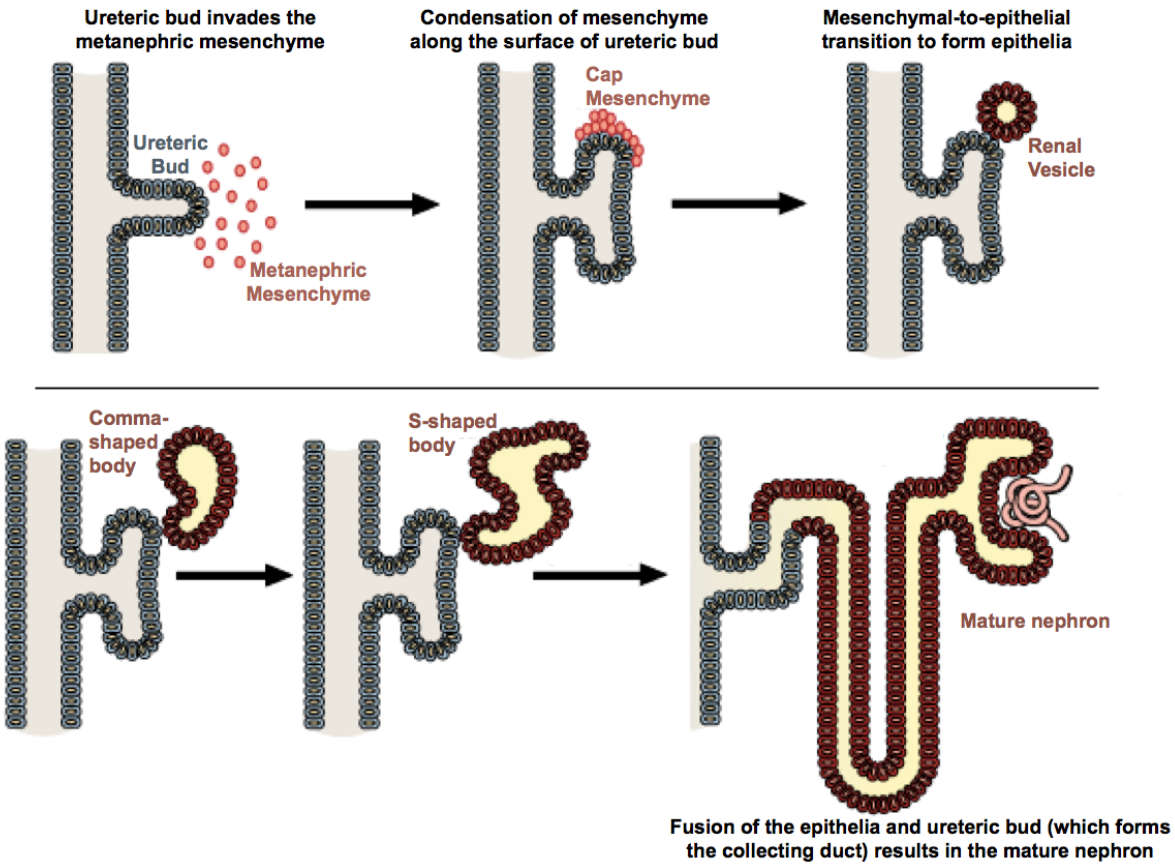


Figure 1. Kidney development. Interactions between the metanephric mesenchyme (MM) and ureteric bud (UB) result in invasion of the UB into the MM, leading to mesenchyme condensation and formation of pretubular aggregates. These pretubular aggregates go on to form the epithelial components (renal vesicle, comma-shaped bodies, S-shaped bodies) of the kidney. The distal end of the S-shaped body connects to the ureteric bud, while the proximal end joins to form the glomerulus.

of the ureteric bud. These stromal cells include pericytes and mesangial cells. The iterative branching of the ureteric bud yields the collecting duct system of the kidney. This entire process kidney development is completed by week 34 in humans, and postnatal day 7 in mice.

1.1.2. Wilms tumor resembles early kidney organogenesis

The distinctive histology of Wilms tumors is reminiscent of fetal kidney development and suggests that Wilms tumor arises from pluripotent embryonic renal progenitors that, instead of undergoing normal differentiation to form the functional kidney, show arrested differentiation and aberrant proliferation, ultimately resulting in a Wilms tumor. This is supported by a number of morphological and molecular observations. First, the characteristic triphasic appearance of Wilms tumor is histologically comparable to that of the developing kidney (Figure 2). Wilms tumors frequently contain variable proportions of three cellular components derived from the fetal mesenchyme: epithelial, blastemal, and stromal elements.[8] The epithelial elements resemble comma-shaped bodies, S-shaped bodies and glomeruli. The blastemal component is comprised of small, densely packed cells representing metanephric mesenchyme, while the stromal elements resemble fibroblasts, skeletal muscle or neural elements. Moreover, gene expression studies provide further support for a developmental-origin of Wilms tumors in that genes expressed during early kidney development (*PAX2, HOXA11, SIX1, EYA1, SALL2, MEOX1*) have a tendency to be overexpressed in Wilms tumors, while genes expressed at later stages of nephrogenesis are downregulated in Wilms tumors.[9] Additionally, genome-wide chromatin profiling of Wilms tumors, embryonic stem cells, and normal kidney tissue

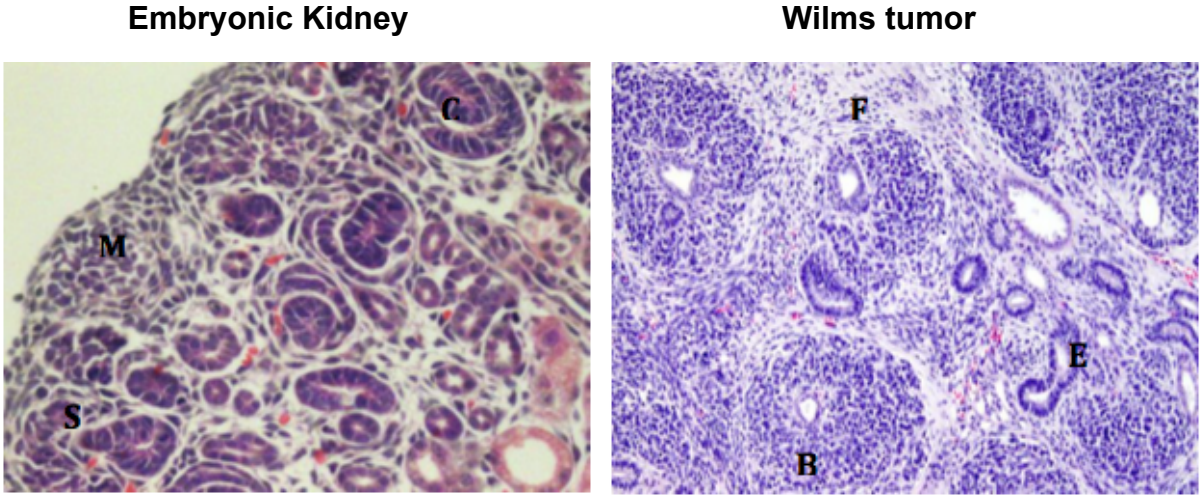


Figure 2. Wilms tumor histology. Wilms tumors (right) are histologically similar to the embryonic kidney (left). The triphasic histology of Wilms tumors shows blastemal, epithelial, and stromal components reminiscent of fetal kidney development. However, the tumors lack correct formation of glomeruli and nephrons.

indicates that Wilms tumors share an epigenetic landscape with renal stem cells, with increased domains of histone H3 lysine 4 trimethylation (K4me3) suggestive of an open chromatin environment.[10] Taken together, the histological morphology, gene expression and chromatin profiling data imply that Wilms tumors arise from a subset of cells arrested at an early metanephric mesenchymal stage during nephrogenesis. Consistent with this hypothesized ontogeny, genes mutated in Wilms tumors include those involved in directing the differentiation of early metanephric mesenchymal cells, as well as components of signaling pathways critical for kidney morphogenesis.[11]

1.1.3. Nephrogenic Rests: the connection between nephrogenesis and nephroblastoma

Wilms tumors are thought to arise from nephrogenic rests, which are benign foci of undifferentiated fetal mesenchyme that, instead of fully differentiating into renal parenchyma by 36 weeks of gestation, persists after renal embryonic development is complete. These nephrogenic rests are found in the kidneys of 1% of infants, 30-40% of sporadic Wilms tumor cases, and 100% of bilateral cases.[12] Noninvasive follow-up of these precursor lesions have shown they have the ability to progress, sclerose, or regress. Two types of nephrogenic rests have been described based on their location within the kidney. Intralobar nephrogenic rests (ILNR) are situated centrally within the renal lobe, represent approximately 10% of nephrogenic rests, and are associated with mutations or deletions in *WT1*. In contrast, perilobar nephrogenic rests (PLNR) are located at the periphery of the renal lobe and represent 90% of nephrogenic rests. Nephrogenic rests often share the same genetic defects found in Wilms tumors. These shared mutations are not present in the surrounding normal kidney tissue,

suggesting that they are responsible for the formation of Wilms tumor precursor lesions in the kidney. Additional mutations only in the tumor tissue indicates that neoplastic transformation of nephrogenic rests results in the development of Wilms tumor.[13-14]

1.2. The genetics and emergence of new genes in Wilms tumor

The etiology of Wilms tumor is heterogeneous with multiple genes known to result in Wilms tumor formation. These genes and loci, identified through molecular and cytogenetic analyses, have been found mutated primarily in sporadic Wilms tumors (Table 1) and includes somatic mutations in *WT1, CTNNB1, WTX, TP53, MYCN, SIX1/2*, miRNA pathway genes, and epigenetic abnormalities at 11p15.[15-18] Wilms tumor is a relatively euploid cancer and, like other pediatric malignancies, the number of somatic mutations found in tumors is low.[16] This suggests that these somatic alterations acquired by tumor cells are significant events in the initiation of Wilms tumor.

Gene	Effect of mutation	Somatic or Germline?
WT1	Inactivation of protein	Both
WTX	Inactivation of protein	Somatic
TP53	Inactivation of protein	Both
IGF2	Biallelic expression	Both
CTNNB1	Activating	Somatic
CTR9	Inactivation of protein	Germline
DROSHA	Alteration of miRNA biogenesis	Somatic
DGCR8	Alteration of miRNA biogenesis	Somatic
MYCN	Activating	Both
SIX1	Activating	Somatic
SIX2	Activating	Somatic

Table 1. Genes involved in Wilms tumor development.

1.2.1. *WT1* and *CTNNB1*

Wilms tumor protein 1 (*WT1*) was the first gene implicated in Wilms tumor development and initially provided most of our understanding of the relationship between kidney development and tumorigenesis. *WT1*, a zinc-finger transcription factor, is essential for renal development and plays a critical role in the regulation of expression of genes involved in nephrogenesis.[19] Moreover, *WT1* directs the mesenchymal to epithelial transition of the metanephric mesenchyme through upregulation of *WNT4*, leading to the differentiation of blastemal cells into mature nephrons.[20] Mutations in *WT1* have been identified in approximately 20% of sporadic Wilms tumors, and result in inactivation of the protein.[15]

The WNT pathway is critical for proper kidney development. Therefore, it is not surprising that components of this signaling pathway, such as *CTNBB1*, have been found to be mutated in Wilms tumors.[21] Interestingly, there is a strong association between *WT1* and *CTNNB1* mutations, as the majority of *WT1*-mutant tumors harbor activating mutations in *CTNNB1*.[15,22] These mutations result in stabilization of beta-catenin and translocation to the nucleus where it activates downstream targets.[21]

1.2.2. *WTX*

WTX, also known as family with sequence similarity 123B (FAM123B) is mutated in roughly 20% of sporadic tumors. Approximately 2/3 of these tumors harbor deletions of the entire *WTX* gene, while the others carry missense alterations, or nonsense and frameshift mutations resulting in a truncated protein.[15,23-25] *WTX* acts as a negative regulator of *CTNNB1* by interacting with the beta-catenin destruction complex in the cytoplasm to promote beta-catenin ubiquitylation and degradation.[26]

Truncating *WTX* mutations identified in Wilms tumors result in the deletion of the beta-catenin-binding region. As a result, beta-catenin is stabilized, accumulates in the nucleus, and activates downstream target genes. Germline mutations in *WTX*, underlying osteopathia striata with cranial sclerosis (OSCS), do not predispose to Wilms tumors, however most children with OSCS die at an early age.[27]

1.2.3. *TP53*

TP53 is of particular importance because, although it is only mutated in a 5% of sporadic tumors, its mutation is associated with relapse and poor prognosis due to metastasis and resistance to therapy.[28-29] The presence of anaplasia defines Wilms tumors that have unfavorable histology and poor prognosis, and somatic *TP53* mutations have been identified in ~75% of anaplastic Wilms tumors.[30-31] The mutations occur at hotspot residues or lead to truncated protein.[32]

1.2.4. *MYCN*

Approximately 10% of Wilms tumors show low-level *MYCN* amplification, with no preference for histological subtypes.[33-37] The increased copy number is typically modest when compared to other tumor types which also show *MYCN* amplification, such as neuroblastoma. *MYCN* amplification usually includes other flanking genes as well due to their close proximity, most often *DDX1*, which is telomeric to *MYCN*. Recurrent somatic point mutations in *MYCN* affecting codon 44 of the Aurora A kinase interaction domain have recently been identified in tumors, although the significance of these mutations is not known.[16-17] It has been suggested that this recurrent mutation has an activating function. *MYCN* is a transcription factor that controls

expression of genes involved in proliferation and differentiation, and is critical for brain development and essential for kidney morphogenesis.[38]

1.2.5. *SIX1/2* mutations

Two recent studies used exome sequencing of blastemal and favorable histology Wilms tumors to identify somatic mutations contributing to the development of Wilms tumors. These analyses revealed a recurrent somatic hotspot mutation in exon 1 (Q177R) of SIX1 in 10% and 7% of cases, respectively.[17,32] Chromatin immunoprecipitation sequencing (ChIP-seq) and global gene expression analysis indicated this mutation results in the alteration of both DNA binding specificity and subsequent gene expression profiles of downstream target genes. Additionally, the equivalent Q177R mutation in *SIX2* was also identified as a somatic change in Wilms tumors. Expression analysis of *SIX1/2*-mutant tumors showed an upregulation of cell-cycle and kidney developmental genes, suggestive of a renal progenitor cell state.[32]

1.2.6. Epigenetic alterations at 11p15

Epigenetic alterations at 11p15 are the most common alteration in Wilms tumor cases. These defects are usually isolated, non-heritable events and typically are only seen in sporadic cases, although not many familial tumors have been studied to date. Loss of imprinting (LOI) or loss of heterozygosity (LOH) at 11p15 is found in most tumors.[39] A cluster of imprinted genes is located at this location. This cluster includes *IGF2*, which is an embryonic growth factor, and *H19*, which expresses a noncoding RNA. The H19 imprinting control region is located at 11p15 that controls expression of IGF2 and H19. This region is differentially methylated depending on the chromosome

parent of origin. Methylation of the maternal allele prevents CTCF binding and IGF2 expression, while H19 is actively expressed. Conversely, hypomethylation of the paternal chromosome allows CTCF to bind, permitting IGF2 expression while blocking H19 expression. In Wilms tumors, loss of the maternal allele and duplication of the paternal allele by either LOH or aberrant gain of methylation at the H19-ICR results in biallelic IGF2 expression. Biallelic IGF2 expression alone is not sufficient for Wilms tumor development, as LOH of 11p15 was identified in the normal kidney and blood cells of WT patients. [40]

1.2.7. MicroRNA Pathway Genes

The most recent genes implicated in Wilms tumor development are those of the miRNA biogenesis pathway. This pathway produces mature miRNAs that negatively regulate the expression of target proteins by targeting their mRNA transcripts. This pathway was first connected to Wilms tumor formation when inactivating mutations in *DIS3L2* were identified in Perlman syndrome.[41] DIS3L2 specifically targets uridylated pre-let-7 miRNAs for degradation, and a small fraction of patients with *DIS3L2* mutations are predisposed to Wilms tumor development. Additionally, studies have demonstrated the importance of specific miRNAs in kidney development.[42]

A number of recent studies have also identified mutations in *DGCR8, DROSHA, XPO5, TARBP2* and *DICER1* in Wilms tumors.[16-17,32] Somatic heterozygous mutations in *DROSHA*, the most commonly mutated miRNA pathway gene, were identified in ~12% of Wilms tumors. These were recurrent E1147K mutations disrupting the Mg2+ binding site of the RNase IIIb domain. MicroRNA

expression profiling demonstrated loss of the miR-200 family, whose members include those involved in the mesenchymal-to-epithelial transition, consistent with the pathology of the tumors. Germline and somatic mutations in *DICER1* have been identified less frequently in Wilms tumors, but appear to contribute to Wilms tumor development through altering the balance of 5p/3p miRNAs (described in chapter 3).

1.3. Familial Predisposition to Wilms tumor

The genes responsible for sporadic Wilms tumor cases are known in approximately 50% of cases. However, these genes rarely contribute to instances of familial Wilms tumor. Approximately 2% of Wilms tumor patients have an affected relative, usually a sibling or cousin. These cases typically present at an earlier age than sporadic Wilms tumor presumably due to the inheritance of a germline mutation predisposing to cancer development. The majority of Wilms tumor families are small, often containing 2-3 affected individuals. Among these affected individuals, males and females are equally represented. Evidence suggests that the predisposition gene is not subjected to imprinting effects, as there is no gender bias in the obligate carrier parents of affected children. Familial Wilms tumor is also observed in a range of ethnicities. Although most familial Wilms tumor cases are non-syndromic, there are a minority of familial cases associated with Beckwith-Wiedeman syndrome, Denys-Drash syndrome, WAGR syndrome, Perlman syndrome, and rarely, in DICER1 Syndrome.

Familial WT predisposition displays an autosomal dominant inheritance pattern with incomplete penetrance estimated at 30%.[2,43-44] Familial predisposition genes have been localized by genetic linkage analysis to 17q12-21 (FWT1) and 19q13.3 (FWT2), although neither has been identified.[45-46] The lack of linkage of some WT

families to these regions indicates additional familial WT genes exist.[47] This genetic heterogeneity could be the reason for variation seen in families with respect to penetrance, frequency of bilateral disease, and age of diagnosis. For example, although familial Wilms tumor is typically diagnosed 12 months earlier than sporadic cases, Wilms tumors occurring in 17q-linked families are diagnosed at a median age of 6 years. On the other hand, families linked to chr19q show variable ages of diagnosis, ranging from 14 to 204 months.

1.3.1. Previous Identification of familial WT predisposition Genes

With advancements in next-generation sequencing technologies, the mutational landscape of Wilms tumor is becoming more clear. The overwhelming majority of these studies, however, focused on sporadic Wilms tumors. Nonetheless, both traditional and high-throughput sequencing efforts have contributed to a small list of genes that, when mutated, are associated with predisposition to familial Wilms tumor. These genes include *WT1* (rarely), *CTR9*, *DICER1*, and copy number alterations including *MYCN* and *DDX1*.

Germline mutations in *WT1* are known to predispose to Wilms tumor, however most germline *WT1* mutations are frequently de novo. There are rare cases in which germline *WT1* mutations inherited from an affected or obligate carrier parent results in familial Wilms tumor. These mutations have been reported in very few families, most of which are small.[48-51] In general, familial Wilms tumor cases are not due to germline mutations in *WT1*.

Recently, next-generation sequencing technology allowed for the identification of *CTR9* as a Wilms tumor susceptibility gene. In this study, inactivating mutations in

CTR9 were identified in 3/35 Wilms tumor families.[52] CTR9 is a gene involved in RNA Polymerase elongation by association with the PAF1 complex. Mutations in additional members of this complex have been identified in Wilms tumor cases, providing additional support for CTR9 as a cancer predisposition gene. Although CTR9 is a newly established Wilms tumor predisposition gene, it should be noted that these mutations were found in a fraction of familial Wilms tumor cases, which illustrates the characteristic genetic heterogeneity of Wilms tumor predisposition.

The research described in this Thesis served to expand the genetic etiology of Wilms tumor occurring in families, and successfully resulted in the identification of the predisposing alterations in three families whose whole genomes were sequenced. We identified germline mutations in *DICER1*, which have recently been found in six families with DICER1 Syndrome-associated Wilms tumors (described in chapter 3). Additionally, copy number alterations at chr2p24, encompassing the *MYCN* locus, have recently been identified in a family with a history of nephroblastomatosis and Wilms tumor. Our analysis resulted in the identification of two Wilms tumor families harboring copy number alterations of this region (chapter 4). Finally, our identification of *SPHK2* as a candidate Wilms tumor predisposition gene on chr19q (chapter 2) opens new avenues of research into the role of chromatin regulation and lipid signaling in Wilms tumor initiation.

1.4. Significance

The study of Wilms tumor has generated many concepts that are fundamental to our understanding of the genetic etiology of cancer. The first "Wilms tumor gene" identified, *WT1,* plays an important role in the development of multiple organ systems and alterations of the gene or its expression are involved in many tumor types. Interestingly, while *WT1* was originally isolated by virtue of being a "Wilms tumor gene", it is now recognized that *WT1* plays an important, although not well understood, role in a variety of cancers, including acute lymphocytic leukemia (ALL). Additionally, a *WT1* vaccine has shown therapeutic promise against AML, glioblastoma, breast and lung carcinoma in Phase I/II clinical trials. Thus it is very plausible that the identification -- and the understanding of the aberrant function of -- other "Wilms tumor genes" not only will represent a major step in our understanding of cancer predisposition, but will also greatly accelerate work in the area of cancer genetics in general and will potentially identify therapeutic targets for very common cancers.

Chapter 2: Whole genome sequencing to identify candidate Wilms tumor predisposition genes

2.1. Introduction

A number of high-throughput, massively parallel sequencing (next-generation sequencing) platforms developed since the completion of the human genome project has allowed for the generation of unprecedented amounts of data compared to sequencing by capillary electrophoresis. Not only is this technology more cost effective per base than traditional sequencing methods, but large amounts of data can now be generated in a fraction of the time compared to conventional sequencing approaches. Recent successes in disease gene identification using next-generation sequencing technology has demonstrated the power of whole exome- and genome sequencing studies to understand drivers of human disease and cancer. Fortunately, the rapid evolution of bioinformatics analysis programs combined with the falling cost of whole genome sequencing has allowed us to employ this technology to pinpoint candidate genes predisposing to familial Wilms tumor.

2.2. Illumina Whole Genome Sequencing

The Illumina sequencing platform is one of the most widely adopted next-generation sequencers used by both research and clinical laboratories due to its low error rates and cost efficiency. The sequencing process begins with library preparation, in which genomic DNA is randomly fragmented followed by the ligation of adapters to the ends of each fragment. The adapters facilitate the hybridization to an Illumina flow cell which contains surface-bound oligos complementary to the adapter

sequences. Here, solid-phase amplification of each single template creates clusters of identical copies of the fragment to generate a sufficient signal for the detection of incorporated bases. Illumina's sequencing by synthesis approach uses a reversible terminator method to distinguish single fluorescently-labeled deoxynucleoside triphosphates (dNTPs) as they are incorporated into growing DNA strands. Following the addition of a single labeled dNTP into the growing nucleic acid chain, an image is taken to identify the incorporated base. The labeled base is then cleaved for the next base incorporation event to take place. The nucleotide calls are distinguished from one another based on the signal intensity measurement, resulting in the sequence of bases from the fragment.

2.3. Whole genome sequencing in three Wilms tumor families

To identify Wilms tumor predisposition genes, we sequenced the whole genome of eleven individuals from three Wilms tumor families using the Illumina sequencing platform. The individuals from each family were chosen for whole genome sequencing based on being: 1) affected individuals or individuals known to be an obligate carriers of the predisposing mutation by virtue of having an affected child, 2) distantly related (to reduce the number of shared variants for further investigation), and 3) on our already having a DNA sample from them. Informed consent was obtained from all study participants, and approval for this research was obtained from the Institutional Review Board at MD Anderson Cancer Center.

Illumina paired-end libraries were constructed from genomic DNA by the MD Anderson Cancer Center Sequencing Core Facility. The advantage of paired-end libraries is that, in addition to high quality sequence information, they also provide

long-range positional information allowing for the highly precise alignment of sequence reads. Libraries were sequenced at approximately 30X coverage using the Illumina HiSeq 2000 system, which is sufficient coverage to allow for the detection of germline sequence variants. The resulting sequence data was then contained in fastq files for processing, alignment and variant calling.

2.4. Bioinformatic analysis

While the ability to generate larger amounts of data through next-generation sequencing has improved exponentially, the continued development of faster, more accurate bioinformatic analysis programs has not kept pace, resulting in a bottleneck at the analysis stage. Additionally, there are challenges associated with identifying rare variants to better understand the genetic basis of human diseases. Identification of variants depends on many factors. Considerations such as coverage depth and base quality scores have an impact on determining if an alteration is called or not, as well as the parameters which set the threshold for variant calling. Recognizing this, we established a collaboration with Wenyi Wang, PhD at MD Anderson, who developed FamSeq, a computational tool that performs family-based variant calling in next-generation sequence data. Of the eleven individuals sequenced in our study, three individuals comprised a trio of affected parent/child and unaffected parent. Using this family-based sequencing approach, we used information from the trio to more accurately identify germline mutations by calculating the probability of variants given the entire family's raw sequence measurements. Fortunately, the data generated from this work has not only served to advance our understanding of Wilms tumor genetics, but also to advance a computational tool aimed at improving the

identification of rare variants in family sequencing studies.

2.4.1. Read alignment and variant calling

100bp paired-end sequence reads that passed instrument QC from each individual were processed and aligned to the GRCh37/hg19 human reference genome using the Burrows-Wheeler Aligner (BWA) v0.5.9. PCR duplicates were marked by Picard v1.46. Variant calls were produced using the Genome Analysis Toolkit (GATK), FamSeq, and CASAVA, resulting in data sets which include single nucleotide polymorphism (SNP) calls, insertions, and deletions with respect to the reference sequence, as well as confidence scores for every covered position in the genome. To identify copy number variation, the total number of read counts for each position across the whole genome was extracted from the bam file. The mean read counts for a fixed window of 50 bases was calculated and used as input for a Hidden Markov Model (HMM) segmentation algorithm. This produced copy number segments assigned to gains or losses depending on the ratio of total median read counts within each segment to the mean sequencing depth across the entire genome.

2.4.2. Identification and annotation of intra-family shared variants

Using the variant calls generated by GATK, FamSeq and CASAVA, we first compared the genotypes between the affected/obligate carrier individuals in the same family to identify variants that are shared. These variant were contained in variant call format (vcf) files for each individual. This analysis generated three lists of variants specific to the three families in our study. In each family, this approach reduced the list of putative risk alleles to those which are 1) less likely to be artifacts, and b) presumably inherited from the same (carrier) relative. Using this approach, our

strategy of sequencing the most distantly related affected family members and identification of overlapping variants reduced the list of candidate variants because the more distantly they are related, the fewer genetic variants they shared. Variants that were not shared between family members were discarded from further investigation, as the criteria for the predisposition gene includes the presence of the variant within all affected/obligate carrier individuals.

Once shared variants were identified, ANNOVAR (http://www.openbioinformatics.org/annovar) was used to aid in the investigation of the functional significance of variants by annotating each variant with respect to location within the genome (intergenic, intronic, exonic, or regulatory regions) using the *annotate_variation.pl –geneanno –buildver hg19* command. Variants present in dbSNP137 were identified using the *annotate_variation.pl –dbtype snp137* command. Variant frequencies from the 1000 Genomes database were also obtained using ANNOVAR. Alterations were also examined for their presence in the NHLBI Exome Variant Server (https://esp.gs.washington.edu/drupal).

Variants in exonic regions which could potentially affect protein structure (nonsynonymous and frameshift variants) were identified and further annotated using ANNOVAR to obtain the following scores: SIFT (*annotate_variation.pl avsift*), PolyPhen (*annotate_variation.pl ljb_pp2*), MutationTaster (*annotate_variation.pl ljb_mt*), and PhyloP (*annotate_variation.pl ljb_phylop*). This produced a list of variants with predictions as to whether they are deleterious, as well as conservation scores to assess whether identified variants occur within evolutionarily conserved regions which could include promoter regions and sequences important for RNA expression, RNA stability, and RNA translation.

2.4.3. Filtering and prioritization of variants

One advantage of whole genome sequencing technology over first-generation methods, such as Sanger sequencing, is the generation of large amounts of data with variant calls across the entire genome. This typically results in >4 million variant calls for each whole genome. In this current study, with whole genome sequence data from eleven individuals, the amount of data generated required the application of a discrete set of filters to prioritize those variants which are most likely to be the cause of disease in the families. As previously described, we first identified intra-family shared variants because potential Wilms tumor predisposing mutations will be those that are present in all affected/obligate carrier individuals in a family, and absent in unaffected family members. We then filtered these shared variants using dbSNP137 and the 1000 Genomes databases to remove variants with a minor allele frequency (MAF) greater than 1% in those databases. This reduced a large list of potential candidate genes by removing all variants that are commonly occurring in the general population, with the rationale being familial Wilms tumor represents a small proportion of all Wilms tumor cases and Wilms tumor is relatively rare in the population. Consequently, we expected risk alleles to occur at a low frequency in the general population or to even be novel variants. As such, variants with a MAF greater than 1% and present in these databases were excluded from our initial analysis. Although we prioritized variants based on their presence in dbSNP, it is important to note that pathogenic variants have previously been included in dbSNP. With this in mind, our initial analysis included only those variants not present in dbSNP, although the comprehensive list of variants was archived for future examination should we be unsuccessful in identifying a promising candidate variant.

In order to identify the familial Wilms tumor predisposition mutation in each family, the variants were then screened by function and embryonic expression pattern. Similar approaches taken by other studies have proven effective in filtering out non-causal mutations. Previous studies have shown that most alleles that are known to underlie Mendelian disorders disrupt protein-coding sequences. Therefore, mutations of high interest ("first tier variants") included those that have an effect on protein via non-synonymous mutations, insertions or deletions in a coding region. "Second tier" variants included those that occur outside the coding region of genes (intronic, intergenic regions). These second tier variants were saved for future evaluation in the case that promising first tier variants were not identified. In addition, higher priority was assigned to genes expressed in fetal kidney, as Wilms tumor is a cancer involving fetal renal development.

The resulting data was initially assessed for variants that occurred in more than one of the three families or that affected the same gene/functional element in more than one family. However, this analysis failed to identify variants in the same gene from all three families. We next shifted our analysis to variants from each individual family (described in more detail in chapters 3-5). Following the identification of candidate predisposition genes, the application of a co-segregation analysis was used to further prioritize the set of key functional variants. This was because familial Wilms tumor is rare among Wilms tumor cases, and the chance that both parents of an affected familial Wilms tumor patient carry the predisposing mutation is unlikely. The results from this analysis are described in the following three chapters.

CHAPTER 3: Identification of sphingosine kinase 2 (*SPHK2*) as a putative 19q familial Wilms tumor predisposition gene.

3.1. INTRODUCTION

Following a genetic linkage analysis of a large multigenerational WT family, a 3.4Mb region of chr19q13 was identified as the location of a gene which predisposes to familial Wilms tumor.[46] Traditional approaches (i.e. Sanger sequencing) to isolate the causal 19q familial Wilms tumor are not practical due to the number of genes (~137) within the linked region. Fortunately, advancements in next-generation sequencing technologies have allowed for the cost-efficient, high throughput sequencing of individuals within the family to identify the predisposition alteration. Using whole genome sequencing, the entire genetic landscape of six individuals within the family was surveyed to pinpoint the predisposing alteration in the family. This resulted in the identification of sphingosine kinase 2 (*SPHK2*), an enzyme involved in the production of sphingosine 1-phosphate, as a putative 19q familial Wilms tumor predisposition gene.

3.1.1. S1P signaling

Sphingosine 1-phosphate (S1P), a bioactive phospholipid, has emerged as an important regulator of many fundamental cellular processes.[53] In particular, S1P has received considerable attention due to its discovery as an important extra- and intracellular signaling molecule, serving as the mechanism whereby S1P regulates cell proliferation, survival, differentiation, migration and inflammation. Invariably, with diverse roles in such cellular activities, disruptions of S1P signaling have been

implicated in the development and progression of cancer, and progress is being made to therapeutically target the biosynthesis of S1P by targeting the enzymes involved in the synthesis of S1P: the sphingosine kinases (SPHK1 and SPHK2).[54.55]

S1P is derived through a process of sphingolipid synthesis and degradation pathways (Figure 3). At the center of this network is ceramide which is deacylated to form sphingosine, the precursor to S1P. After phosphorylation of sphingosine by either SPHK1 or SPHK2, S1P is either exported from the cell or interacts with intracellular targets to elicit a number of effects. S1P can be exported out of the cell by members of the ABC transporter family where it acts extracellularly by binding to a family of five G-protein couple receptors, $S1PR_{1-5}$. This results in autocrine or paracrine signalling through the receptors, which are coupled to various G proteins. Differential expression of the various S1P receptors mediates the cellular response through the ERK, Rac and Rho, and PI3K/AKT pathways. S1P can be dephosphorylated or degraded by S1P lyase to form hexadecenal and phosphoethanolamine.

Current dogma suggests the levels of ceramide, sphingosine and S1P forms a “lipid rheostat”, whereby the regulation of synthesis and catabolism of these sphingolipids controls cell fate. Data from cell culture experiments shows that ceramide and sphingosine are pro-apoptotic and when their levels exceed that of S1P, the cell is directed towards apoptosis. Ceramide directly activates a number of targets that mediate its pro-apoptotic functions, including protein phosphatases 1, 2A and 2C, whereas sphingosine binds and initiates the inactivation of the pro-survival protein 14-3-3.[102-105] Conversely, S1P is anti-apoptotic and results in cellular proliferation, migration and survival when its production surpasses ceramide and

sphingosine levels. This is accomplished by modulation of extracellular-signal-regulated kinases 1/2 (ERK1/2), Rac and Rho, the phosphoinositides 3-kinase (PI3K)/AKT pathway and phospholipase C (PLC).[105-106] These effects also display cell- or tissue-specific variability resulting from the differential expression of S1P receptors and the various G proteins they couple to.[107] Thus a delicate balance in the levels of these lipids has the potential to direct cells toward a pro-survival or pro-apoptotic outcome.

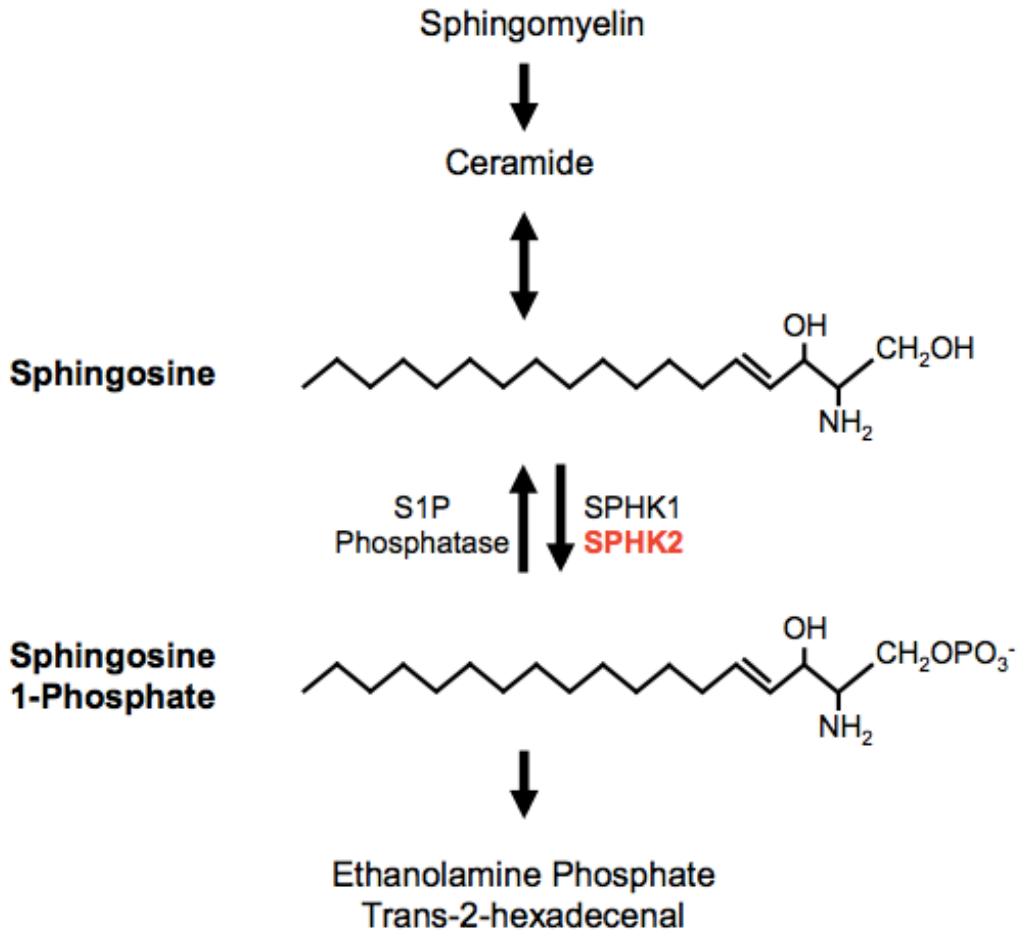


Figure 3. Sphingolipid Metabolism. The sphingosine kinases (SPHK1 and SPHK2) phosphorylate their substrate, sphingosine, producing the bioactive lipid, sphingosine 1-phosphate (S1P). The reverse reaction is catalyzed by S1P phosphatase. S1P can be irreversibly degraded to produce ethanolamine phosphate and trans-2-hexadecenal. The entire sphingolipid metabolic pathway comprises the "sphingolipid rheostat" whereby the levels of each lipid dictates cell fate. Higher levels of ceramide and sphingosine, relative to S1P, induces cellular apoptosis. Conversely, when S1P levels increase above ceramide and sphingosine, the cell is directed to proliferate.

3.1.2. Sphingosine kinase 2 (SPHK2)

There are two mammalian sphingosine kinases, *SPHK1* and *SPHK2,* which are located on chromosome 17q25 and chromosome 19q13, respectively.[56-57] These kinases share 80% similarity and 45% overall sequence identity. Both SPHKs contain five highly conserved regions which contain ATP binding and catalysis regions (Figure 4). While both catalyze the phosphorylation of sphingosine, *SPHK2* contains additional residues at its N-terminus and central region that are not present in *SPHK1*, suggesting that *SPHK2* has additional functions.[58] Opposing roles of the enzymes have also been suggested due to the presence of both a nuclear localization sequence (NLS) and a nuclear export sequence (NES) in *SPHK2* that allows shuttling between the cytoplasm and nucleus.[59-60] These sequences are not found in *SPHK1*. Additionally, differences in sphingosine kinase expression during embryonic development suggests *SPHK1* and *SPHK2* perform separate responsibilities.

Two forms of *SPHK2* have been described. The longer isoform (SPHK2-L) has an additional 36 amino acids at the N-terminus compared to the short isoform (SPHK-S).[61] Although there is much less sequence conservation in vertebrates in the region encoding these 36 amino acids compared to the rest of the protein (UCSC Genome Browser), the SPHK2-L appears to be the predominant form in most human tissues -- with the notable exception of the brain and kidney.[61] Consistent with this, we have observed a preponderance of SPHK2-S in human fetal kidney by real-time RT-PCR. Human and mouse SPHK2-S are highly homologous, with 83% identity and 90% similarity. Notably, SPHK2-S is the only isoform expressed in the mouse.

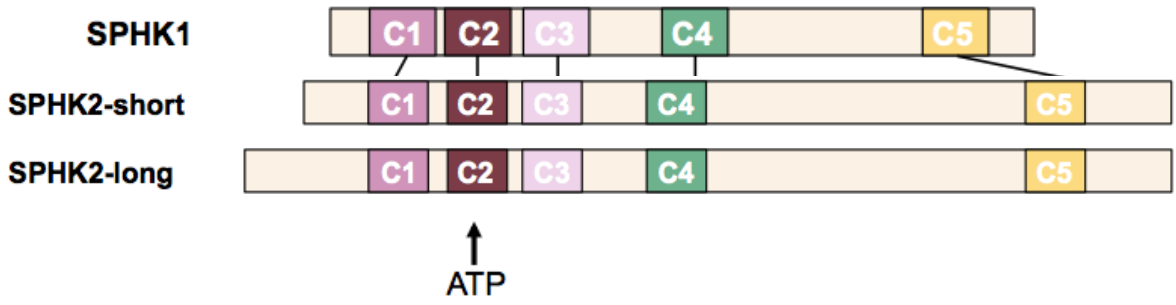


Figure 4. Comparison of SPHK1 and SPHK2. Each kinase contains five conserved domains. An ATP binding site resides within C2 of each enzyme. SPHK1 (384 amino acids) and SPHK2 (SPHK2-long: 654 amino acids, SPHK2-short: 618 amino acids) share 80% similarity and 45% overall sequence identity. Also shown are the SPHK2 long isoform (SPHK2-long) and SPHK2 short isoform (SPHK2-short). These two SPHK2 isoforms differ by 36 amino acids present at the N-terminus of SPHK2-long.

SPHK2 possesses an intrinsic catalytic activity which can be enhanced by growth factors such as EGF.[62] This is mediated by ERK1-dependent phosphorylation at Ser-351 and Thr-578, resulting in a 2-fold increase in catalytic activity.[63] Mutagenesis experiments have shown that mutating G212 interferes with ATP binding and abolishes SPHK2 catalytic activity. Genetic knockout of *SPHK1* or *SPHK2* in mice results in no obvious phenotypic abnormalities suggestive of a compensatory mechanism, while double knockout of both SPHKs is embryonic lethal around E11.5 due to defects in neurogenesis and angiogenesis.[64] While no gross phenotypic anomalies were found in *SPHK2* knockout mice, the embryonic and adult kidneys have not been extensively examined for histological abnormalities (personal communication – Dr. Richard Proia, National Institutes of Health).

While considerable attention has focused primarily on the cellular activities of *SPHK1*, less is known about the function of *SPHK2*, particularly with respect to tumorigenesis. Varied roles for *SPHK2* in tumorigenesis have been described, including its ability to induce cell cycle arrest and apoptosis or, in different cell types, to promote survival and proliferation. A mechanistic role for *SPHK2* in the nucleus was first revealed upon the demonstration that it is co-immunoprecipitated with histone H3 (but not H4, H2B, or H2A), associates with HDAC1/2 in co-repressor complexes (including those at the promoters of *p21* and *c-fos*), and results in increased acetylation at H3-K9, H3-K14 and H2B-K12.[65] This function was shown to be due to SPHK2's phosphorylation of sphingosine and resulting inhibition of HDAC1/2. Other experimental studies have demonstrated an oncogenic role of *SPHK2* in regulating the proliferation and survival of human acute lymphoblastic leukemia (ALL) cells. These effects were shown to occur via HDAC2-mediated

regulation of c-Myc.[66] Specific systemic inhibition of *SPHK2* (not affecting *SPHK1* function or expression) also resulted in a significant reduction in engraftment of human ALL cells into NOD/SCID mice. Thus, an exciting epigenetic role of *SPHK2* in regulating gene expression is emerging, as is SPHK2's epigenetic role in cancer development.

3.1.3. *SPHK2* in the developing kidney

In the mouse fetal kidney, several components of the sphingosine 1-phosphate signaling pathway are actively expressed, including the sphingosine kinases, S1P receptors (S1P1, S1P2, and S1P3), as well as S1P phosphatase and lyase, the enzymes involved in S1P metabolism. *SPHK2* expression is predominately localized to the metanephric mesenchyme at higher levels than SPHK1 at E11, and gradually decreases to P4, indicative of a role in nephrogenesis.[67] Support for a functional role also comes from pharmacological ablation of the sphingosine kinases in the embryonic kidney using small molecular inhibitors. Treatment of embryonic kidney explant cultures with sphingosine kinase inhibitors, thus blocking the production of sphingosine 1-phosphate, resulted in reduced branching morphogenesis during kidney development, underscoring the importance of S1P in development.

3.2. MATERIALS and METHODS

3.2.1 WTX524 family

Following informed consent from all study participants, blood was drawn from members of the WTX524 family and peripheral blood DNA isolated for molecular analysis. The WTX524 family is a Hispanic, multigenerational Wilms tumor family with nine total cases. The ages of diagnosis ranged from 4-37 months, with an average of

15.3 months (considerably lower than the ~33 month age of diagnosis in familial Wilms tumor). Unilateral and bilateral tumors were diagnosed in the family, with the normal kidney tissue of affected children being “unusually rich in nephrogenic rests” (personal communication with B. Beckwith). Five individuals from the family were chosen for whole genome sequencing (Figure 5). Sequencing and bioinformatic analysis was performed as described in chapter 2.

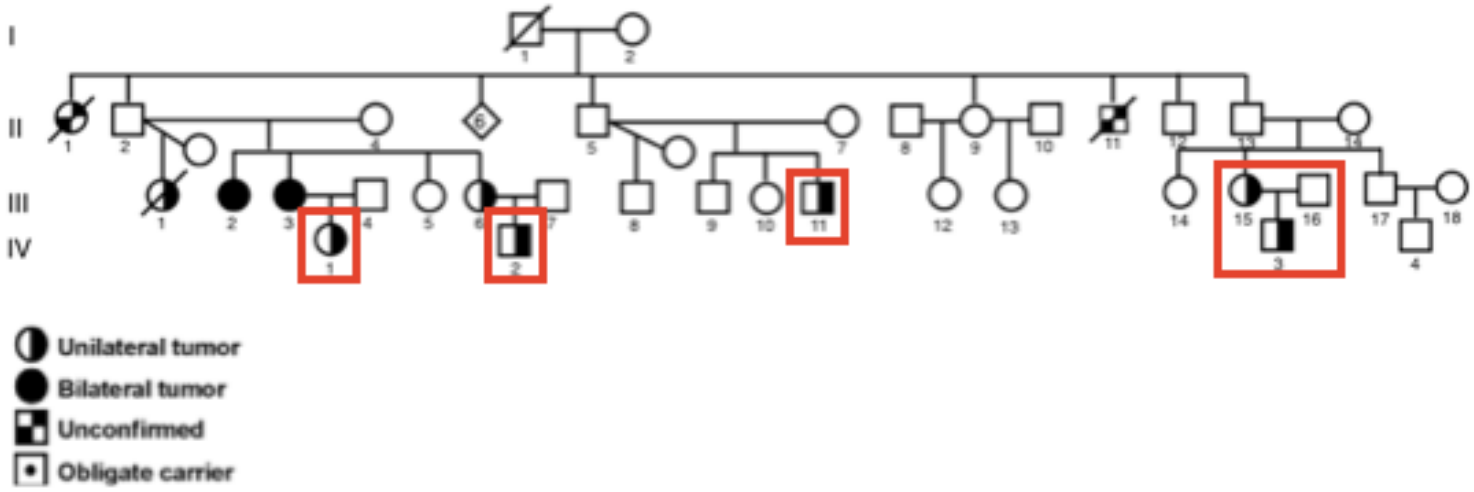


Figure 5. Pedigree of WTX524 family. Red boxes highlight the six individuals in the family whose whole genomes were sequenced. This included five affected family members and one married in control.

3.2.2. Mutation Analysis of Candidate Genes

Following the identification of candidate genes using next-generation sequencing, a complete mutation analysis of the genes was performed using Sanger sequencing of blood DNA from members of the family for segregation analysis. AmpliTaq (Life Technologies) and primers designed in Oligo (Molecular Biology Insights, Colorado Springs, CO) were used to amplify genomic DNA in a 25ul reaction volume for 35 cycles. The PCR product was then treated with Exonuclease I and Shrimp Alkaline Phosphatase. Purified PCR products were then sequenced with target-specific primers and BigDye v1.1 (Life Technologies). Sequence reactions were purified with BigDye Xterminator purification kit and run on a Genetic Analyzer 3500 (Life Technologies). Sequence analysis was next performed with MutationSurveyor (Soft Genetics). If the candidate variant co-segregated with the affected/obligate carrier status, the mutation analysis was extended to one member of 48 Wilms tumor families to assess the frequency of candidate gene mutations in familial Wilms tumor samples. Additionally, qPCR was used to detect exonic deletions and duplications of the candidate gene. qPCR primers targeting exons (including exon/intron boundaries) were designed using Primer Express. Primer sets for mutation analysis of each gene by Sanger sequencing and qPCR are listed in Tables 2-3.

SANGER PRIMERS		
	FORWARD Primer	REVERSE Primer
SPHK2 exon 2	AGCCTTCACTCCTATTAT	GATTTTGTCTTTCCACAGTAT
SPHK2 exon 3	CACCTGTCTGAGCCTGTCTG	CACATCCGGTCCTCTTGTCT
SPHK2 exons 4-6	GGGCTGGAAGTAACCTAA	GCGCTCGGGATACAG
SPHK2 exon 7a	ACCCACAGTCAGTCAAGTAA	GACGGGTGAGTGTAG
SPHK2 exon 7b	GACGGGTGAGTGTAG	CTGCCACCACAACTCTA
TRPM4 exons 1-2	TGTCCCTCTGTCCCCTTAT	CACAGACCCAAACGCTCCTA
TRPM4 exon 3	CCCTCAGAGAGAAGATAGGT	CTCTACAGGAGCATGAACAC
TRPM4 exons 4-6	CCATGTGTCCACAGTTCC	TCTCGACCAGACCAAACTAC
TRPM4 exons 7-9	ACTCCTGGGAAATGCG	AAAATATTAAGAGGGGTAAA
TRPM4 exon 10	ATGACCTTATGCCTATGATA	AGAGAGGGAGGACAGCAC
TRPM4 exons 11-12	GCCCATCTCTTGTCCTTAAC	TCGGGGAATTGTATG
TRPM4 exons 13-14	CCATCTCCAAACATTACTTCA	GAGAGTGGCTGGTCAAG
TRPM4 exons 15-16	TTGCCTTTCACCACCGTTTC	TCTCGCCATGTGATACACTG
TRPM4 exon 17	AGACTGAGTTTCGGGTCA	CTGAGTGGCGGACA
TRPM4 exons 18-19	CCGCGTCTGTCCATAC	AAACTGCTGGCGAGTGTA
TRPM4 exon 20	CGCCTGGCTAACAGT	ATCAGAGTTCAGGGACTAAA
TRPM4 exons 21-25	TACAGGCGTGACCAA	AAGGGAGGAAACGGACTC
GADD34 exon 2	CACACCGGGAGTGTTGT	AAATAGAATCTGGGCTCTCAG
GADD34 exon 3	ATAGAGAATCCCGTGACAGT	TACAGACGCTGCTCGCTACA

Table 2. Primer sets for mutation analysis of three candidate genes by Sanger sequencing.

qPCR PRIMERS

	FORWARD Primer	REVERSE Primer
SPHK2 exon 1	CTCTACCACGATCCGGACTTCT	TTGGACGTTGCGCTTGTTC
SPHK2 exon 2	CCAGGGTCCCGTTGATGTAA	ATGAGGATCTTGGCTTTTAGATGGT
SPHK2 exon 3-4	GTCTGAGCCTGTCTGTCTCTGATG	CTGTGGGAAGGGAGGAGAGAA
SPHK2 exon 5-6	CTGAGTCCTAAGGGAGCAAAGTG	CTGTCGTTCTGAAGTGGGATACC
SPHK2 exon 7	TCCGTCCCCAATCTAAAAAGC	GCTGGAAGTGCTGGAGTCTGA
TRPM4 exons 1-2	CCGGATCTGTGGAGTCAACTATG	TGCCCTCTCACCCCACTTC
TRPM4 exon 3	CGGCCCCCGTGAAGTC	CCGCAGTGGCCATGGA
TRPM4 exons 4-6	AGACCAGGGCGTCTCCTCAT	CGGAATAGAGACACCCTCATCAA
TRPM4 exons 7-9	CCAGCCACGAGGAGACATG	GGTGAATATCCTGCCTTTTCTGA
TRPM4 exon 10	GTTCCAAGCCACAGCCAAA	TTGTCCCCTCTCCCTCTAATCC
TRPM4 exons 11-12	GATGCAGGAGCGCCAAGT	CCCTGGAGCGACCTGCTT
TRPM4 exons 13-14	CCTCACCTCACTGCTGCGATA	TCACACCCTCACCCTACTCCTT
TRPM4 exons 15-16	GACAGGCCCTTCCCCATTAA	TTCTCTTCCCCTCATTTCTTGCT
TRPM4 exon 17	AGTGGGCAGGAAGGATGAGA	CTCGCCGACAGCTGGAA
TRPM4 exons 18-19	CCGTGAAAACCATGAAGTCGAT	ACACTTGACCCTTGTGGCATCT
TRPM4 exon 20	AGGAGGGTGTGCCCAGAAG	TCACCCCATCTCTGAATGTCTCT
TRPM4 exons 21-25	GCCCCGCTCTCACTTCTG	ATCGGTGCATAAGGAGAACTTTCT
GADD34 exon 1	TGGGATGGGAACCGATAAGA	TCGAAACCCCGCCTCTCT
GADD34 exon 2	CCCAAGACTCCAACTCAATTAAAAA	GCTCAAGCGCCCAGAAAC
GADD34 exon 3	CGCTGCTCGCTACAAATCAA	GGGAGGCGTGGCTGAGA

Table 3. Primer sets for the detection of exonic deletions in candidate genes by qPCR.

3.2.3. Gene Expression by qPCR

SYBR Green (Applied Biosystems) was used for analysis of SPHK2 gene expression. RNA was isolated from lymphoblastoid cell lines using Trizol, quantified and used for cDNA synthesis using MultiScribe Reverse Transcriptase and PCR Master Mix according to the protocol (Invitrogen). Forward and reverse primers for detecting expression levels of specific genes were designed using Primer Express (Applied Biosystems). Quantification of expression was performed using SYBR green and the ABI 7900 HT sequence detection system (Applied Biosystems). Gene specific Ct values were compared across samples relative to GAPDH levels and calculated as fold differences to a control reference value whose expression was designated as 1. qPCR primer sequences are listed in Table 4.

	Forward primer	Reverse primer
p21	CGATGGAACTTCGACTTTGTCA	GCACAAGGGTACAAGACAGTG
c-fos	CACTCCAAGCGGAGACAGAC	AGGTCATCAGGGATCTTGCAG
c-myc	GGCTCCTGGCAAAAGGTCA	CTGCGTAGTTGTGCTGATGT
SPHK2-L	ATGAATGGACACCTTGAAGCAG	CATGGCCTTAGCCCTGACCAG
SPHK2-Total	CTGTCTGCTCCGAGGACTGC	CAAAGGGATTGACCAATAGAAGC

Table 4. qPCR primers for expression profiling of SPHK2 isoforms and target genes.

3.2.4. Generation of SPHK2 constructs

Human SPHK2-S and SPHK2-L cDNA was amplified from a SPHK2 image clone (Origene, SC113181) using the primers listed in Table 5. A PCR product of the correct size was gel purified and enzymatically treated with XbaI and XhoI restriction enzymes for 1 hr at 37*C. This digested PCR product was then treated with Mung Bean nuclease to produce blunt ends, as well as shrimp alkaline phosphatease at 37*C for 30 min. The product was then ligated into a pcDNA3.1 V5/His expression vector (Life Technologies). The ligation reaction was performed at 16*C overnight, followed by transformation into competent E.coli cells overnight. Colonies were screened by cutting with EcoRI at 37*C for 30min, followed by gel electrophoresis. Those that appeared to have integrated the plasmid + insert of the correct size were grown overnight at 37*C and screened by PCR for the correct insertion of the plasmid. QuikChange site directed mutagenesis was performed using manufacturer's protocol to introduce the R100W and R136W mutations. Sanger sequencing confirmed the introduction of the mutation. Plasmids containing the mutation were grown overnight at 37*C and used for Midi-prep of DNA. The same procedure was repeated to generate mouse SPHK2 constructs using mSPHK2 cDNA and mouse-specific SPHK2 primers. Site directed mutagenesis was used in the generation of mutant mouse SPHK2 (R100W – the equivalent of the human R136W mutation) constructs.

3.2.5. Transient and stable-transfected cell lines

To generate transient and stable-transfected cells for functional experiments, MCF7 (a human breast cancer cell line) and M15 (a mouse renal mesenchymal cell line) cells were seeded in 6-well plates and incubated at 37*C overnight. Transfection

using Lipofectamine 3000 (Life Technologies) was performed at a ration of 3:1 (lipid complex/DNA) and incubated overnight. The following day, the transfection reagent was removed and cells were used for 1) transient expression experiments, or 2) seeded in 10cm culture plates along with 1ug/ml G418 for 2 weeks to select for stable transfectants. Following selection with G418, cells stably expressing the plasmid were seeded sparsely into 10cm plates for isolation of single colonies. Single colonies were trypsinized and seeded into 24-well culture plates, along with media supplemented to G418. Once the colonies were confluent, they were trypsinized and seeded into 12-well plates. The process was repeated using larger culture volumes until enough cells were obtained for experiments.

M15 cell lines with the R101W mutation (analogous to the human SPHK2^{R136W}) introduced into the endogenous locus of SPHK2 were generated using CRISPR/Cas9 genome engineering. These cell lines were a gift from E. Cristy Ruteshouser and used for gene expression studies described below.

3.2.6. Western Blot Analysis

SPHK2-stable transfected MCF7 cells were washed with ice-cold PBS and scrapped in SPHK2 lysis buffer containing 50mM HEPES, 150mM NaCl, 1mM EDTA, 1% Triton X-100, 2mM sodium orthovanadate, 4mM sodium pyrophosphate, 100mM NaF, and 1:500 protease inhibitor mixture. Lysates were sonicated and centrifuged at 14,000 x g for 15 min and quantified with the BioRad protein assay and spectrophotometer. Equal amounts of protein lysates (20ug) were used for subsequent analysis. Protein lysates were denatured for 10 min at 99 °C with β-mercaptoethanol before loading into BIORAD Mini-PROTEAN TGX Precast Gels with

Laemmli running buffer at 90 volts for 2 hrs. Following gel electrophoresis, proteins were transferred to nitrocellulose membranes using BIORAD Mini Trans-Blot cell assembly run at 22V at 4 °C overnight. Membranes with bound proteins were blocked for 1 hr in 5% milk in TBST (1X Tris buffered saline/0.1% Tween-20) followed by incubation with primary antibody in 5% milk in TBST for 1 hr at room temperature. Membranes were then washed 3 times with TBST, followed by a 1 hr incubation with secondary antibody (donkey anti-rabbit IgG, GE Healthcare, NA934V) diluted 1:5000 in 5% milk in TBST at room temperature. Following 3 washes in TBST, membranes were developed using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific) according to the protocol.

3.2.7. Immunofluorescence

For immunofluorescence studies of cultured cells, cells were plated in 8-well chamber slides (LabTek) and allowed to adhere overnight before fixation with 4% paraformaldehyde for 15 min. The slides were then incubated with 0.5% Bovine Serum Albumin in in PBS for 1 hr, followed by washing three times with PBS. Primary antibodies (anti-V5, 1:1000) were added to the chamber slides followed by a 1 hour incubation at room temperature. Chamber slides were then washed three times in PBS and incubated with secondary antibodies Alexa Fluor-488 (Invitrogen) for 1 hr in the dark. Subsequently, the slides were washed three times in PBS and stained using ProLong Gold anti fade reagent with DAPI (Invitrogen) and allowed to air-dry overnight at 4 degrees in the dark.

3.2.8. Proliferation Assay

To measure proliferation, MCF7 stably-transfected cells and $M15^{R101W}$ cells were plated on 8-well chamber slides (LabTek) and incubated overnight. Detection of a modified thymidine analogue (EdU) incorporated into newly synthesized DNA was performed using the Click-iT Plus EdU Alexa Fluor 488 Imaging Kit (C10637, Life Technologies) following manufacturer's protocol. Images were captured using an Olympus BX60 fluorescence microscope and quantified using MetaMorph Microscopy Automation and Image Analysis Software.

3.2.9. Apoptosis Assay

Apoptosis of M15 cells was performed using the DeadEnd TUNEL assay (Promega) according to manufacturer's instructions with the following modifications: 1) Cells were plated in 8-well chamber slides (LabTek) instead of cover slips. Images were captured using an Olympus BX60 fluorescence microscope.

3.3. RESULTS

3.3.1. Whole genome sequencing identifies candidate familial Wilms tumor predisposition gene: *SPHK2*

To identify the 19q FWT gene, we sequenced the whole genomes of five affected individuals and one married-in control within a large multigenerational WT family. The coverage was approximately 30X across all members (Figure 6). We first looked for shared variants between the 5 affected individuals, and filtered out all variants that were in common with the married-in control. This resulted in 206,860 variants

remaining across the entire genome. After excluding variants previously found in dbSNP or those present at a frequency of >1% in 1000 Genomes Project data, we were left with 20,517. A total of 18 exonic variants were identified from this analysis, of which 10 variants were nonsynonymous and 8 were synonymous (Table 6). Of these, we observed 1 in the 19q linked region (Table 7). The other variants we identified were eliminated from further consideration because they were 1) located outside of the 19q linked region and 2) not plausible candidates for having an obvious role in kidney development. In addition to *SPHK2*, two candidate variants were identified within the linked region that were novel at the time of discovery. However they have since been reported with minor allele frequencies greater than 1%. A comprehensive mutation analysis of the two genes, *TRPM4* and *PPP1R15A*, excluded these genes as candidate familial Wilms tumor predisposition genes based on 1) identification of potentially deleterious variants coming from the unaffected married-in parent and absence of the variant in the affected parent, and 2) their increased population frequency with the continued development of variant databases. Our analysis left us with one candidate, sphingosine kinase 2 (*SPHK2*).

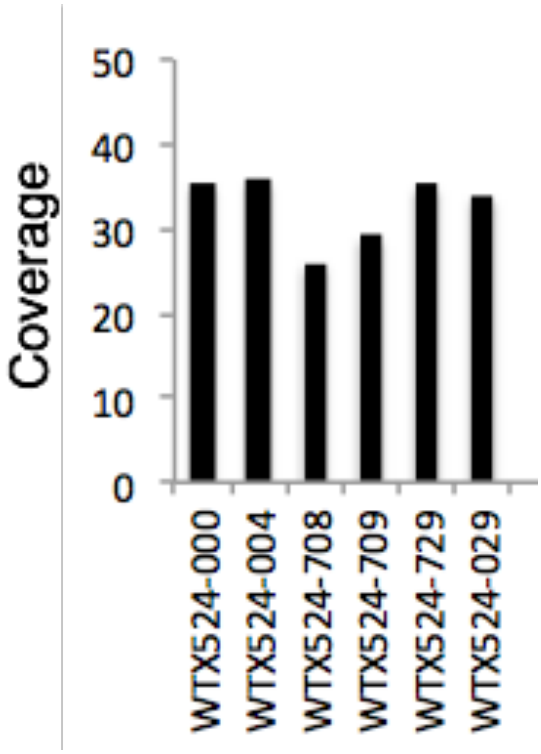


Figure 6. Coverage across the whole genome for six members of the WTX524 family. The coverage was approximately 30X for all members, which is sufficient to call germline variants at high confidence.

Table 6. Shared SNVs and Indels identified through whole genome sequencing in WTX524 family.

Type	Total Variants	Rare/Novel
Intergenic	124,493	12,438
Intronic	70,598	6,973
Downstream	1,306	156
Upstream	1,168	136
3' UTR	1,297	114
5' UTR	212	9
ncRNA_exonic	450	32
ncRNA_intronic	6,225	620
ncRNA 3' UTR	27	2
ncRNA 5' UTR	9	2
Exonic	1,012	18
Synonymous	529	8
Missense	431	10
Stop→gain	5	0
Stop→loss	1	0
Frameshift Indels	9	0
Nonframeshift Indels	16	0
TOTAL	206,860	20,517

Table 6. Shared single nucleotide variants (SNVs) and insertions/deletions (indels) across the entire genome from five WTX524 family members with Wilms tumor. Further filtering using data from the married-in control resulted in the variant numbers listed on this table. Included in the analysis are those shared across the entire genome, as well as novel and rare (minor allele frequency < 1%) variants occurring

Novel/Rare Tier 1 Variants

	Type	Gene	Location	Alteration	Conserved?	FK?
NOVEL	Missense	**ZNF717**	3p12.3	N704K	N	--
	Missense	**AGPAT5**	8p23.1	A44T	N	Y
	Missense	**PCM1**	8p22	T1507R	Y (0.999)	Y
	Missense	**GOLGA6L6**	15q11.2	I393M	N	--
	Missense	**SPHK2**	19q13.33	R136W	Y (0.955)	Y
RARE	Missense	**CLN8**	8p23.3	P229A	Y (0.973)	Y
	Missense	**PRSS55**	8p23.1	R187C	N	Y
	Missense	**SYNE1**	6q25.2	T5355M	N	N
	Missense	**SACS**	13q12.12	P3678A	Y (0.999)	N
	Missense	**SIGLEC12**	19q13.41	Y468C	N	--

Table 7. Candidate familial Wilms tumor predisposition genes following the application of all filtering criteria. Variants listed here are those that are novel/rare, shared by five affected Wilms tumor family members, and not found in the spousal control. Conservation scores are represented using PhyloP. Expression in fetal kidney (FK) was determined using gudmap.org.

The *SPHK2* variant (chr19:49129514, GRCh37/hg19, c.772C>T, p.R136W) was present in roughly 50% of the sequence reads indicative of a heterozygous germline variant and occurs at an evolutionarily conserved amino acid residue indicated by a PhyloP score of 0.999 (Figure 7-8). Substitution of an arginine for a tryptophan at this position is predicted to be deleterious (SIFT-0, PolyPhen-0.999). This variant was not identified in >60,000 individuals whose exome data were included in the Exome Variant Server (http://wvs.gs.washington.edu/EVS/). The R136W mutation does not fall within a currently defined domain of SPHK2, although it should be noted that *SPHK2* is poorly characterized at this point. Interestingly, the R136W mutation occurs within six amino acid residues from the nuclear localization sequence.

3.3.2. SPHK2^{R136W} co-segregates within the family

We confirmed the co-segregation of the germline, heterozygous mutation in all affected and obligate carrier members of the family using Sanger sequencing (Figure 9). In total, 17 mutation carriers were identified, 8 of which developed Wilms tumor (47%). The variant was not identified in 11 other family members tested, 7 of which were married-in, spousal controls. Sequencing this same exon in a tumor sample from the family showed heterozygosity for the mutation, indicating the wildtype allele was not lost somatically in the tumor. Sanger sequencing of the entire *SPHK2* coding region was then performed in the tumor sample for the assessment of possible compound heterozygosity. No somatic *SPHK2* mutations were identified in the tumor.

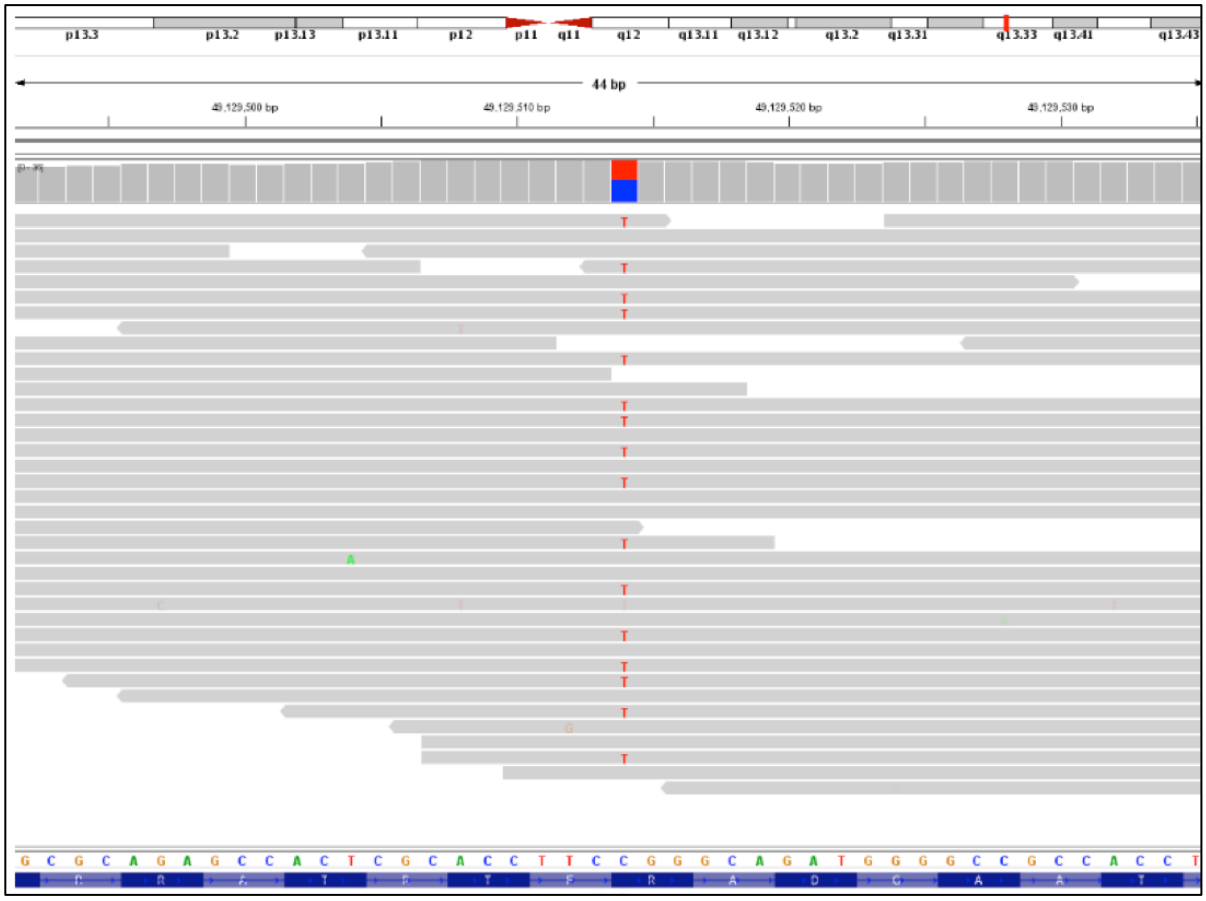


Figure 7. Visualization of the region surrounding the *SPHK2* variant (R136W) using the Integrative Genomics Viewer (IGV). The variant allele (T) is present in roughly 50% of the sequence reads, indicative of a heterozygous variant. This figure is representative of the sequence reads from five affected individuals in the WTX524 family.

Human	A T R T F **R** A D G A A	
Chimp	A T R T F **R** A D G A A	
Orangutan	A T R T F **R** A D G A A	*PhyloP: 0.999*
Rhesus	T T R T F **R** T D G A T	(conserved)
Mouse	A T R T F **R** A D G A T	
Rat	A T R T F **R** A D G A A	

Figure 8. Amino acid sequence conservation of the R136W variant and surrounding residues. Analysis by PhyloP was also used for conservation.

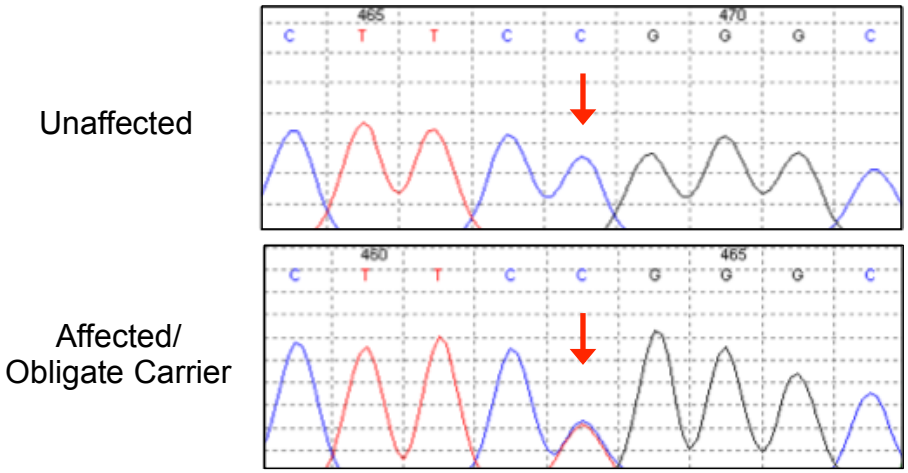


Figure 9. Cosegregation of the R136W variant by Sanger sequencing. Primers targeting exon 3 of *SPHK2* were used to assess cosegregation within the family. These sequence traces represent the unaffected and affected/obligate carriers whose DNA was assessed in the WTX524 family.

3.3.3. Comprehensive *SPHK2* Mutation Analysis in Familial Wilms tumor

A complete mutation analysis of all seven exons of *SPHK2* was performed in one member from 47 additional Wilms tumor families. This mutation analysis included Sanger sequencing, as well as qPCR for the detection of exonic deletions (Table 8). No germline deletions were detected by qPCR. However, five variants, either novel or with a minor allele frequency of less than 1%, were identified in six Wilms tumor families by Sanger sequencing. Two of these alterations, P83P (rs151257086) and L249L (rs116785119), were synonymous mutations not predicted to alter splicing (ALAMUT). A novel, heterozygous 5' UTR variant (49123707A>C) was found in one WT family, however its affect on SPHK2 expression is unknown. Additionally, a rare missense mutation (p.G601A, rs61751862) was identified in two Wilms tumor families. This rare variant has a population frequency of 0.000014 and has been reported in 1/69,600 individuals (Exome Aggregation Consortium, exac.broadinstitute.org). Like the R136W mutation, the G601A also does not occur within a currently defined domain of *SPHK2*. All germline missense mutations identified in familial Wilms tumor samples were predicted to be damaging by at least one of the prediction algorithms and segregated with the affected/obligate carrier status in each family.

Table 8. SPHK2 Mutation Analysis in 48 WT Families

Alteration	No. of Families	MAF
5' UTR (49123707 A>A/C)	1/48 (2.1%)	novel
R136W	1/48 (2.1%)	novel
P83P	1/48 (2.1%)	<1%
G601A	2/48 (4.3%)	<1%
L249L	1/48 (2.1%)	<1%
G126G	5/48 (10.6%)	4%
G21R	2/48 (4.3%)	4%
R562Q	2/48 (4.3%)	4%
L523L	8/48 (17%)	7%
L527L	1/48 (2.1%)	8%

Mutation analysis by Sanger Sequencing and qPCR.

Table 8. Comprehensive *SPHK2* mutation analysis, including Sanger sequencing and qPCR, of all seven *SPHK2* exons in one member of 47 Wilms tumor families identified novel/rare nonsynonymous and synonymous germline mutations.

3.3.4. *SPHK2* Mutation Analysis in Sporadic Wilms tumor

Many cancer genes have been identified whose germline mutation predisposes to cancer and whose somatic mutation also plays key roles in the etiology of non-familial cancers, which account for the majority of cancers in the general population. To assess the frequency of *SPHK2* variants in non-familial tumors, we extended our screening to 175 sporadic Wilms tumors. Sanger sequencing of 175 sporadic Wilms tumor samples was performed using primers targeting each exon of *SPHK2*. Eleven variants (5 missense, 6 synonymous) were identified in this analysis (Table 9). The five nonsynonymous variants were all previously reported in variant databases at a frequency of less than 1%, and all were predicted to be pathogenic by at least one prediction algorithm. Sanger sequencing of the corresponding blood sample from each individual was performed to determine if the variants occurred in the germline or were somatically acquired. The only somatic variant was a novel synonymous mutation (F135F) identified in one tumor. The potential impact of synonymous variants on splicing was assessed using ALAMUT. One variant, P380P, was predicted to activate a cryptic splice site in exon 7 of *SPHK2*. Of the five tumors in which a *SPHK2* variant was identified, three contained an additional alteration, including somatic *CTNNB1* mutation (Ser22Cys), germline *WT1* mutation, and loss of the wildtype *SPHK2* allele.

Table 9. SPHK2 Mutation Analysis in Sporadic Tumors

Alteration	No. of Tumors	MAF	Germline/Somatic	Additional Mutation
R104Q	1/176 (<1%)	novel	GERMLINE	CTNNB1 (Ser33Cys)
F135F	1/176 (<1%)	novel	SOMATIC	—
R80C	1/176 (<1%)	<1%	GERMLINE	—
P504S	1/176 (<1%)	<1%	GERMLINE	—
A506S	1/176 (<1%)	<1%	GERMLINE	Germline WT1 mut.
P654H	1/176 (<1%)	<1%	GERMLINE	Tumor LOH
P42P	1/176 (<1%)	<1%	GERMLINE	—
P83P	3/176 (1.7%)	<1%	GERMLINE	—
A95A	4/176 (2.3%)	<1%	GERMLINE	—
L249L	3/176 (1.7%)	<1%	SOMATIC	—
P380P	1/176 (<1%)	<1%	GERMLINE	—
G21R	11/176 (6.3%)	4%	GERMLINE	—
R652Q	7/176 (4%)	4%	GERMLINE	—
G126G	9/176 (5.1%)	4%	GERMLINE	—
L527L	7/176 (4%)	5%	GERMLINE	—
L523L	32/176 (18%)	10%	GERMLINE	—

Mutation analysis by Sanger Sequencing.

Table 9. Comprehensive *SPHK2* mutation analysis of all seven *SPHK2* exons in 176 sporadic Wilms tumor samples identified novel/rare nonsynonymous and synonymous *SPHK2* mutations.

3.3.5. Functional impact of *SPHK2* mutation

3.3.5.1. SPHK2-short is predominant form in fetal kidney

Before investigating the functional significance of the *SPHK2* R136W variant identified in the WTX524 family, we confirmed expression of *SPHK2* in human fetal kidney using human fetal kidney RNA purchased Clontech Laboratories. Consistent with results from mouse expression profiling, we observed expression of *SPHK2* in the embryonic kidney. It was previously reported that SPHK2-L was the predominant isoform expressed in most human adult tissues and cell lines screened to date except the kidney and brain, where the expression of both short and long isoforms appeared equal. Using isoform-specific qPCR primers, we observed 5-fold greater expression ($p = 0.001$) of SPHK-S in the human fetal kidney, suggesting the short isoform has a more significant role in kidney development (Figure 10).

3.3.5.2. SPHK2^{R136W} impairs nuclear localization and alters gene expression in MCF7 cells

The *SPHK2* variant we identified in the family has not been reported to date. As a result, there is no data describing the functional impact of this nonsynonymous mutation on protein function. Moreover, *SPHK2* has not been as extensively studied as *SPHK1*, and literature describing protein domains and their respective functions are lacking. Although the mutation we identified within the family does not fall within a currently-defined domain of *SPHK2*, it is in close proximity (within 6 amino acid residues) to a previously characterized nuclear localization signal.

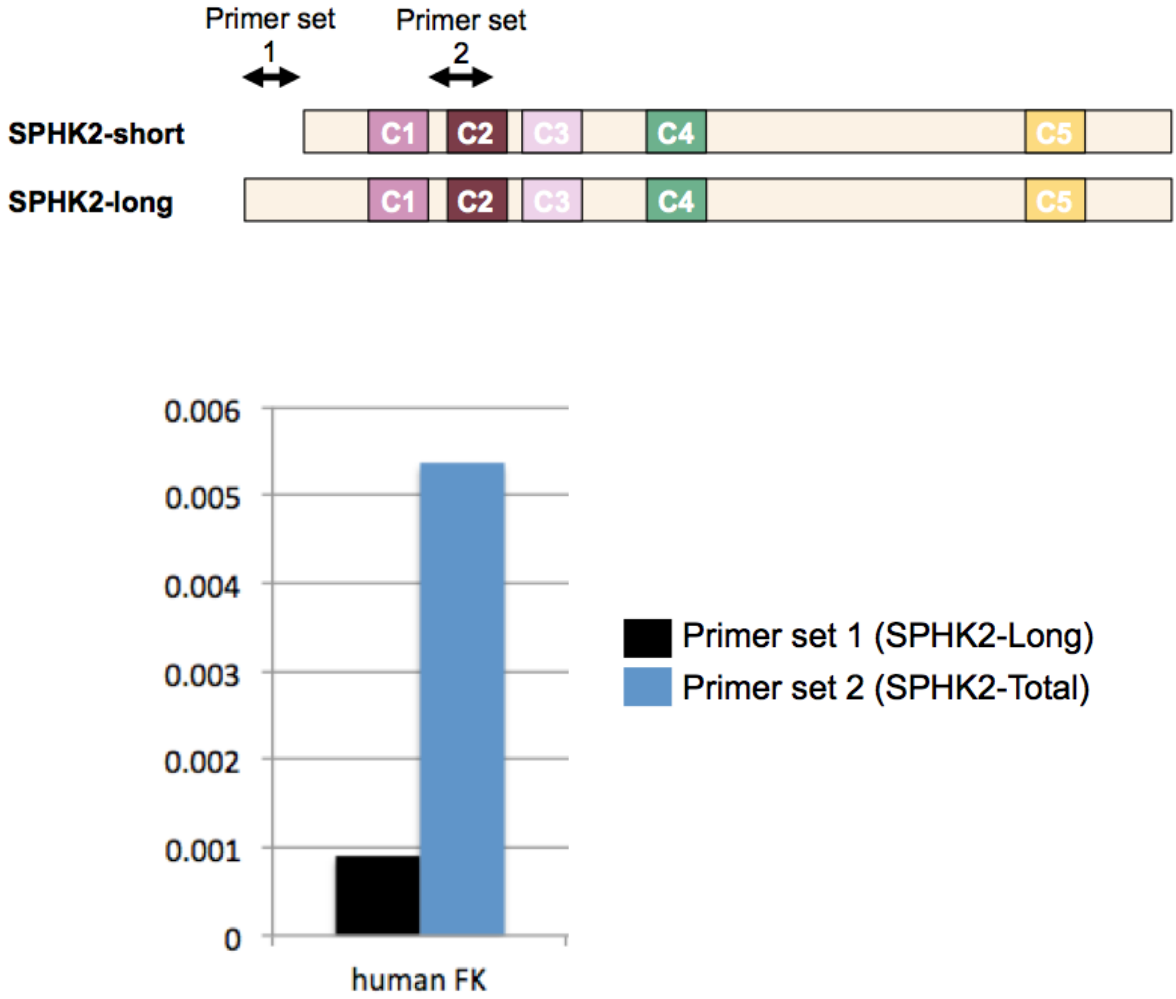


Figure 10. Expression of SPHK2-Total and SPHK2-Long in human fetal kidney. Primers were designed to target a region common to both isoforms (termed SPHK2-total) and a region specific to the long isoform (SPHK2-Long). The difference between the two represents expression of the SPHK2-short isoform.

This ability of SPHK2 to localize to the nucleus is crucial for its role in epigenetic regulation of gene expression.[65] In order to determine if this alteration is functionally significant as suggested by a number of prediction algorithms, we generated SPHK2 constructs (SPHK2-Long and SPHK2-short) containing the same mutation we identified in the family to assess the impact of the mutation on subcellular localization. Using site-directed mutagenesis, $SPHK2^{R100W}$ was introduced into the SPHK2-short isoform, and the analogous $SPHK2^{R136W}$ was introduced into the SPHK2-long isoform. These mutant constructs, as well as the wildtype controls, were transiently transfected into MCF7 cells. We chose this cell line because 1) it expresses SPHK2 at a low level (so endogenous wildtype protein is less likely to obscure the effect of the exogenous mutant protein) and 2) it was the cell line used for the initial characterization of SPHK2's role in the nucleus to interact with HDAC1/2 and histones and to regulate gene expression in an HDAC1/2- dependent manner. This allowed us to directly relate our data on $SPHK2^{R136W}$ to published findings. Expression of the constructs at the mRNA and protein levels were confirmed (Figure 11). Using immunofluorescence, we observed a significant reduction ($p = 0.029$) of mutant SPHK2 in the nucleus compared to the cytoplasm (Figure 12a). As expected, the expression of target genes (*p21* and *c-myc*) was significantly increased when wildtype *SPHK2* was expressed in MCF7 cells, an observation consistent with previous reports (Figure 12b). In contrast, expression of mutant *SPHK2* resulted in a significant reduction of p21 and c-myc, consistent with altered subcellular localization of mutant SPHK2.

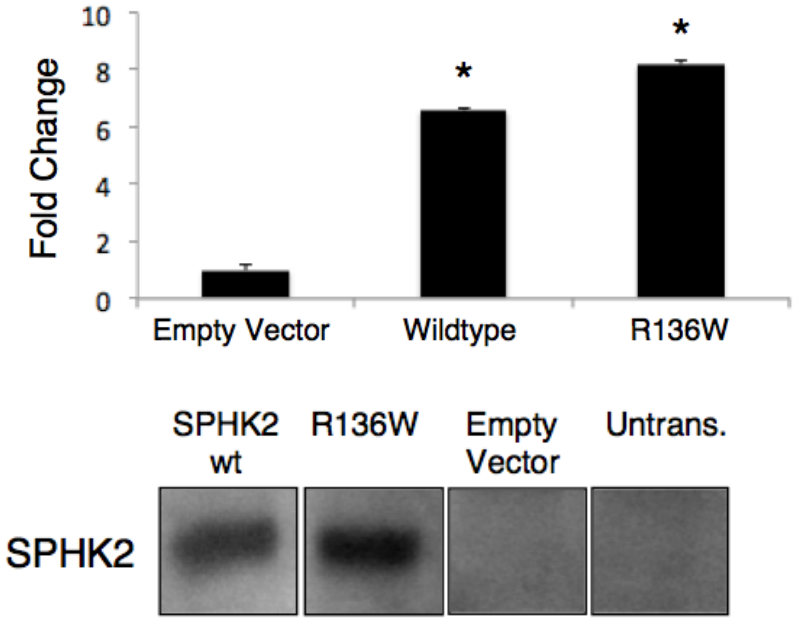


Figure 11. Expression of wildtype and mutant SPHK2 was confirmed at the mRNA and protein levels using qPCR and western blotting, respectively, in MCF7 cells transfected with human SPHK2 overexpression constructs.

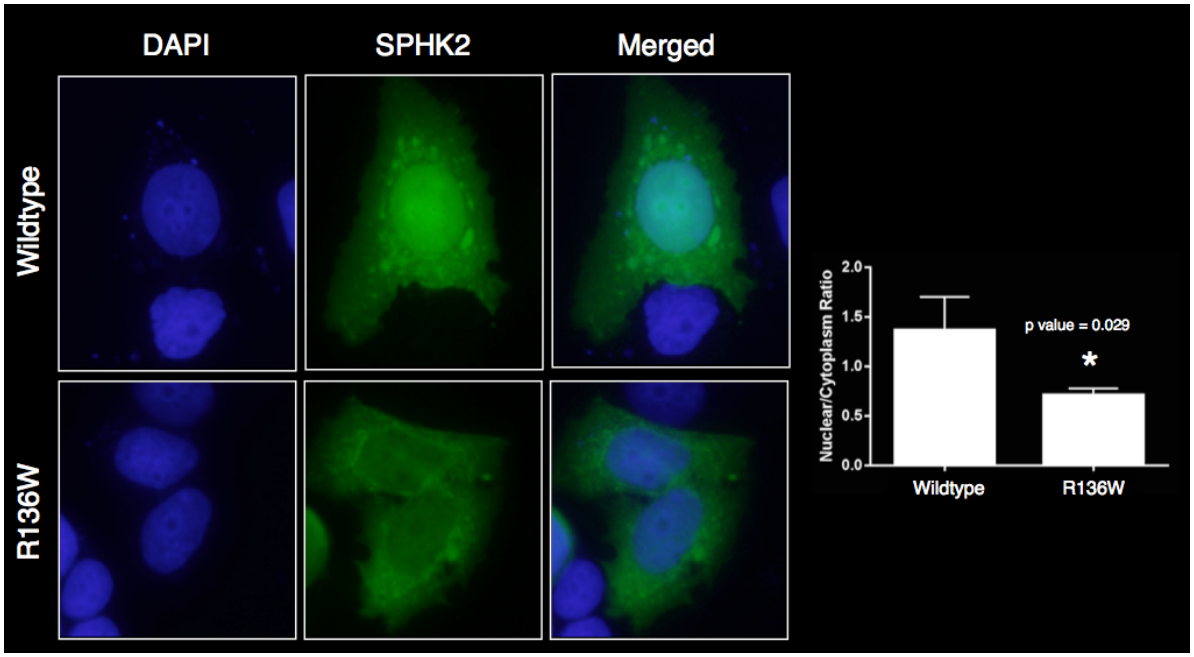


Figure 12a. Immunofluorescence using DAPI (nuclei) and anti-V5 (SPHK2) antibodies to visualize wildtype and mutant SPHK2 localization after transient transfection in MCF7 cells. Intensity of the anti-V5 signal from the nucleus and cytoplasm was quantified with ImageJ, and presented as the nuclear/cytoplasm ratio.

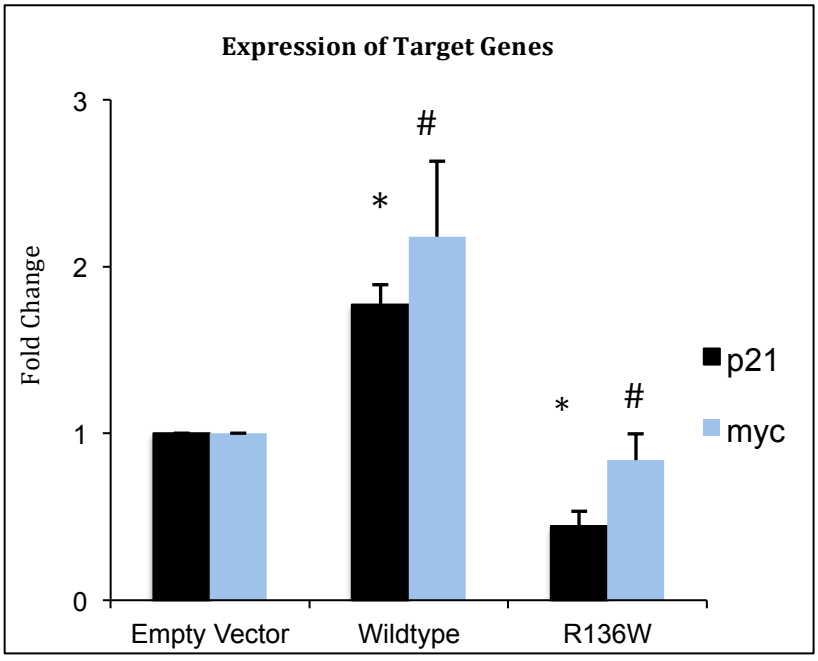


Figure 12b. Expression of SPHK2 target genes in MCF7 cells were assessed by qPCR after transfection with wildtype and mutant SPHK2 expression constructs.
(p values * = <0.02; # = 0.03)

3.3.5.3. SPHK2^{R101W} impairs nuclear localization in M15 cells

Our previous data suggests that the SPHK2 R136W mutation identified within the WTX524 family alters nuclear localization and subsequent gene expression in MCF7 cells. In order to investigate the significance of the variant in a system more relevant to that of kidney development, wildtype and mutant mouse SPHK2 expression constructs were generated and used for transfection experiments in a M15 cell line, a renal mesenchymal cell line similar to the cells in the embryonic kidney that are thought to give rise to Wilms tumors. Wildtype and SPHK2^{R100W} (analogous to the human SPHK2^{R136W}) constructs were transiently transfected into M15 cells and subcellular localization of SPHK2 protein examined using immunofluorescence. We observed a significant reduction in SPHK2^{R100W} nuclear localization compared to wildtype protein (Figure 13). These results were consistent with the analogous experiment in MCF7 cells, providing confirmation of our results in two experimental systems.

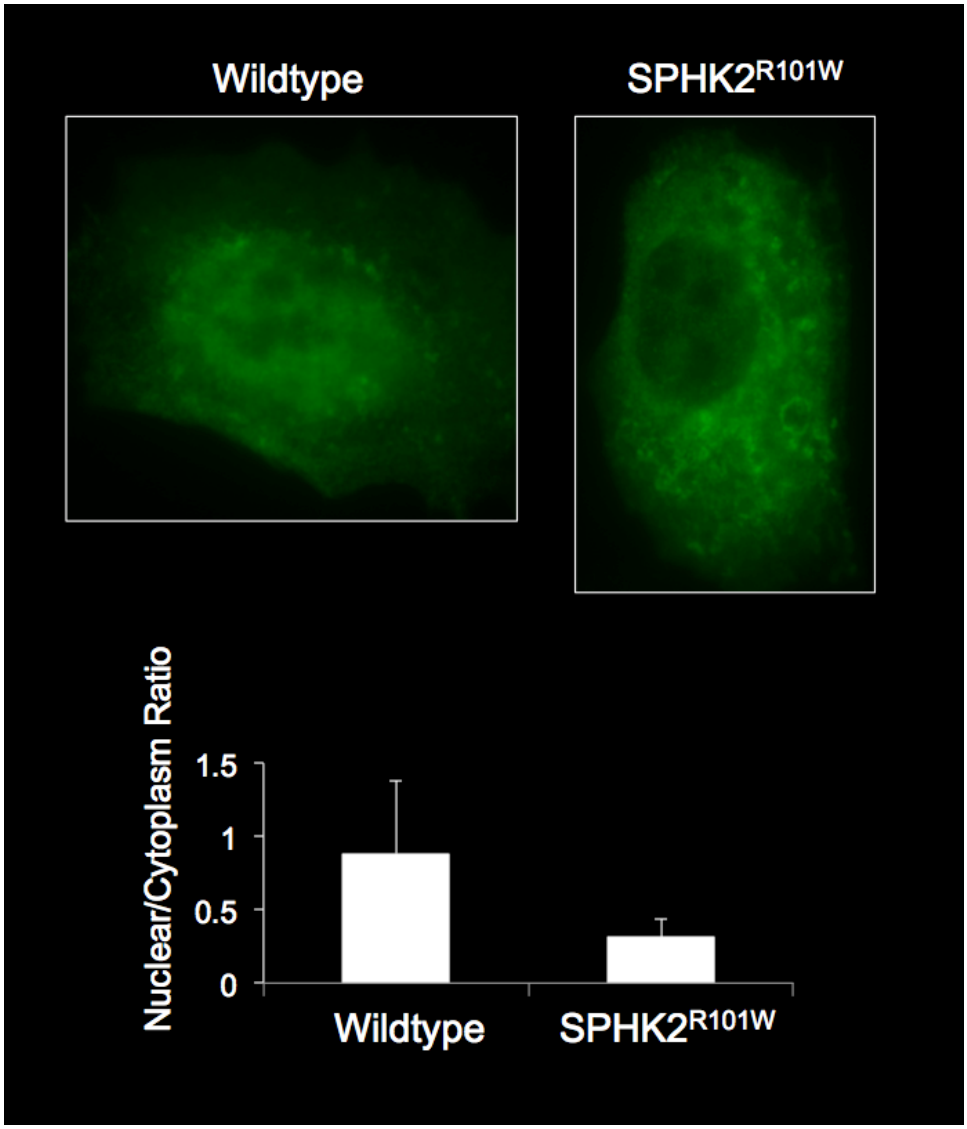


Figure 13. Immunofluorescence assay shows reduced nuclear accumulation of SPHK2^{R101W} compared to wildtype SPHK2. Intensity of the anti-V5 signal from the nucleus and cytoplasm was quantified with ImageJ and presented as the nuclear/cytoplasm ratio.

3.3.5.4. SPHK2^{R101W} alters expression of kidney developmental genes.

Data from our previous experiments suggest that SPHK2^{R136W} alters gene expression following mislocalization of mutant SPHK2 to the nucleus. These studies used overexpression of *SPHK2* in MCF7 cells, which express very low levels of *SPHK2*. In order to assess the effect of this mutation in a fetal kidney experimental system with no manipulation of gene expression level, we examined expression of genes expressed during kidney development using cells in which the SPHK2^{R101W} mutation was introduced into the endogenous locus of SPHK2 via the Clustered Regularly Interspaces Short Palindromic Repeat (CRISPR/Cas9) system. M15 cells carrying the homozygous SPHK2^{R101W} alteration were generated previously in our lab (E.C. Ruteshouser). Gene expression profiling of selected kidney developmental genes (*CDN11, CSRP1, GDNF, LHX1, MLL5, NR2F2, OSR1, SPRY1, EYA1, FOXD1, WNT4, WT1*) was assessed in M15^{wt} and M15^{R101W} cells using qPCR (Figure 14). Of these genes, EYA1 was significantly upregulated in mutant cells compared to wildtype M15 cells. One disadvantage of CRISPR/Cas9 mutagenesis is the possibility of off-target activity due to sequence similarity of the guide RNA with other regions of the genome. Although these target sites (predicted using the CRISPR Design Tool) were screened by PCR amplification and Sanger sequencing and revealed no off-target activity, our analysis included three M15^{R101W} clones generated simultaneously with CRISPR/Cas9 to account for potential off target effects not predicted by the prediction software.

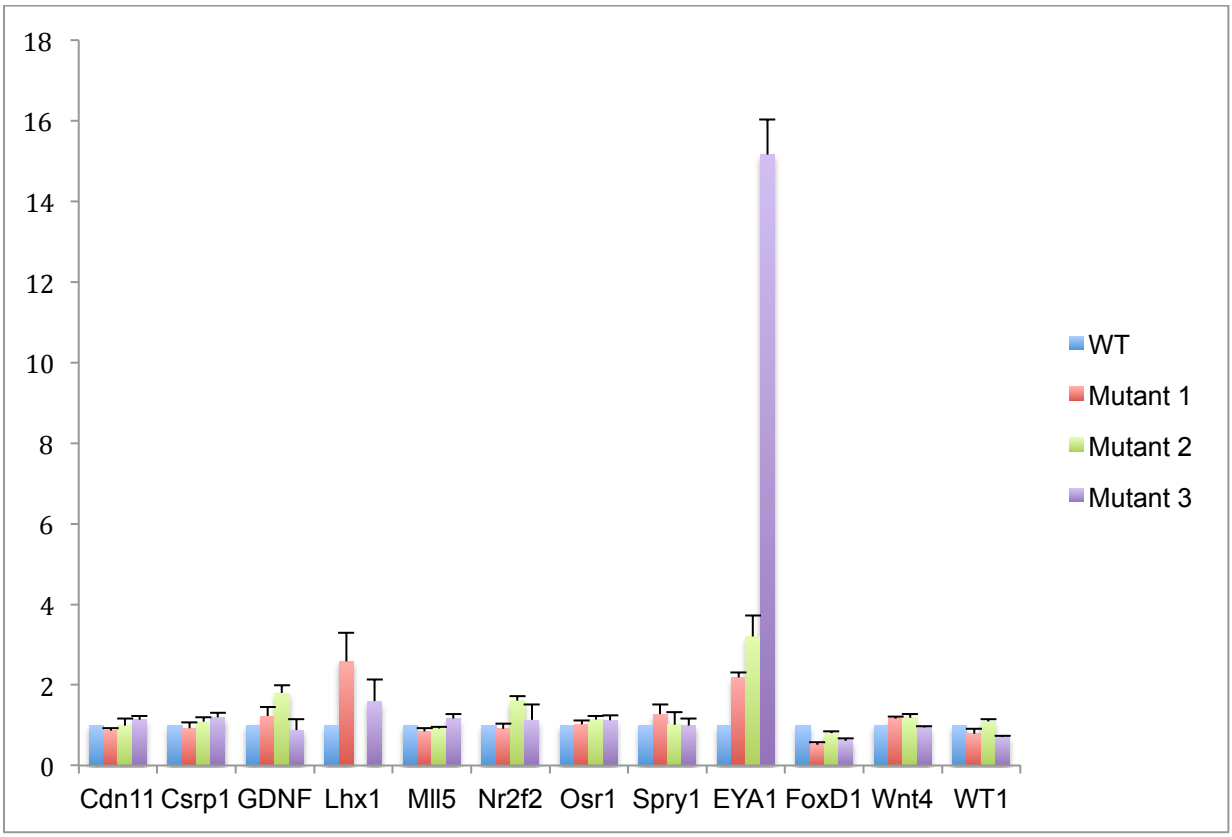


Figure 14. Gene expression profiling of 12 genes expressed during kidney development using wildtype M15 cells, as well as three M15 cell lines containing the R101W mutation generated by CRISPR/Cas9. EYA1 was consistently upregulated in each of the three mutant cell lines, whereas other genes do not appear to be altered in SPHK2 mutant cells. The expression of EYA1 in each mutant clone was significantly upregulated compared to wildtype (p value = <0.01).

3.3.5.5. SPHK2^{R100W} does not affect proliferation or apoptosis

There are several reports of phenotypic effects observed upon SPHK2 overexpression, down-regulation, or inhibition in cell lines. Wildtype SPHK2 has been shown to promote both proliferation and apoptosis, depending on the cell type used. These contradictory results appear to be due, at least in part, to the different subcellular localization of SPHK2 and its product, S1P, between cell types. We investigated the effect of mutant SPHK2 on proliferation in M15 cells containing the SPHK2 mutation knocked in to the endogenous locus and noted no significant difference between wildtype and mutant SPHK2 cells, suggesting the genes regulated by SPHK2 are not involved in cell proliferation of M15 cells (Figure 15). Additionally, apoptosis was also examined in wildtype and mutant M15 cells, and no significant differences were observed (Figure 16).

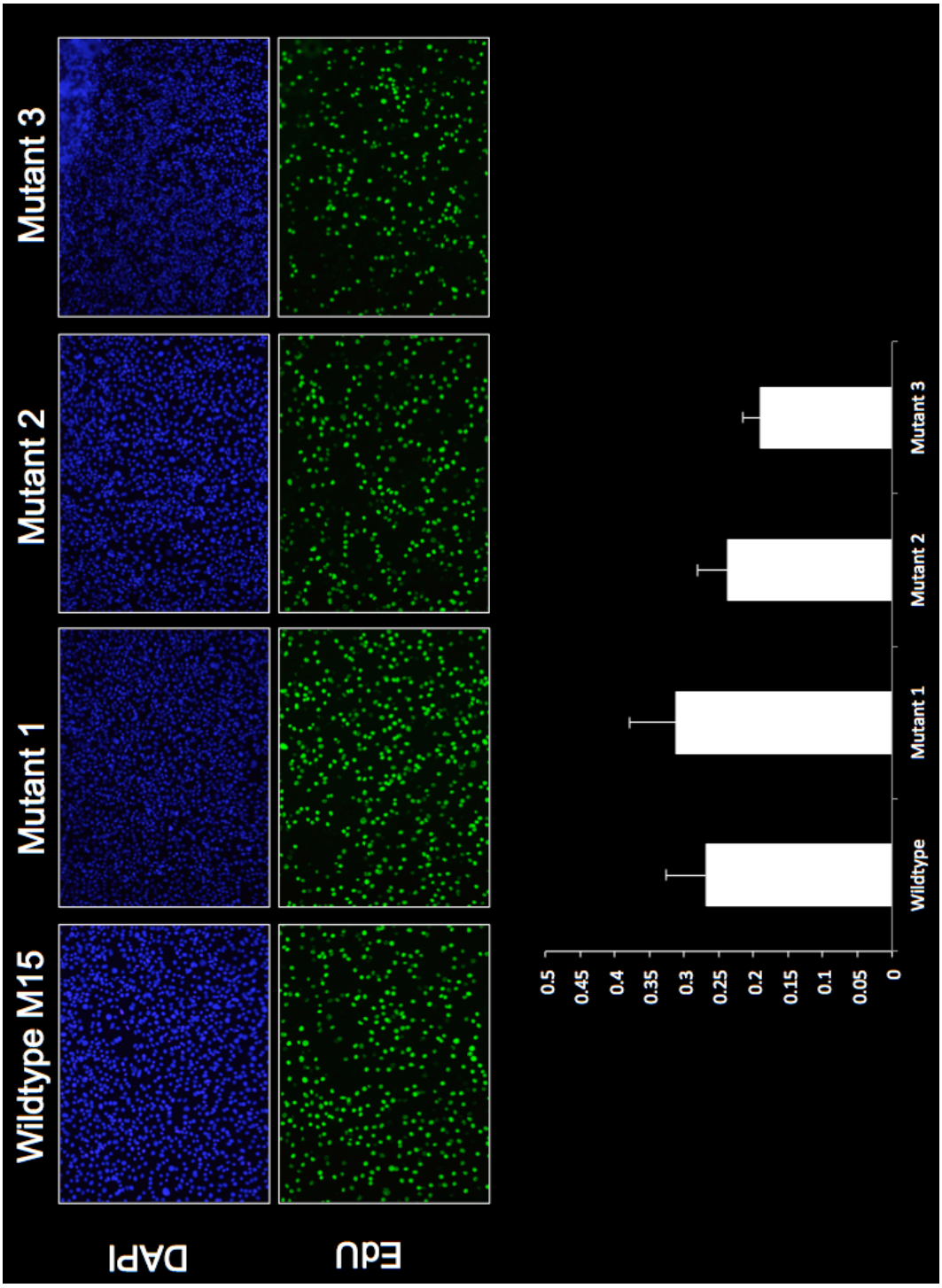


Figure 15. Cell proliferation assay using wildtype M15 cells and three SPHK2^{R101W} mutant clones. Proliferation was measured by incorporation of EdU, a thymidine analogue, into proliferating cells.

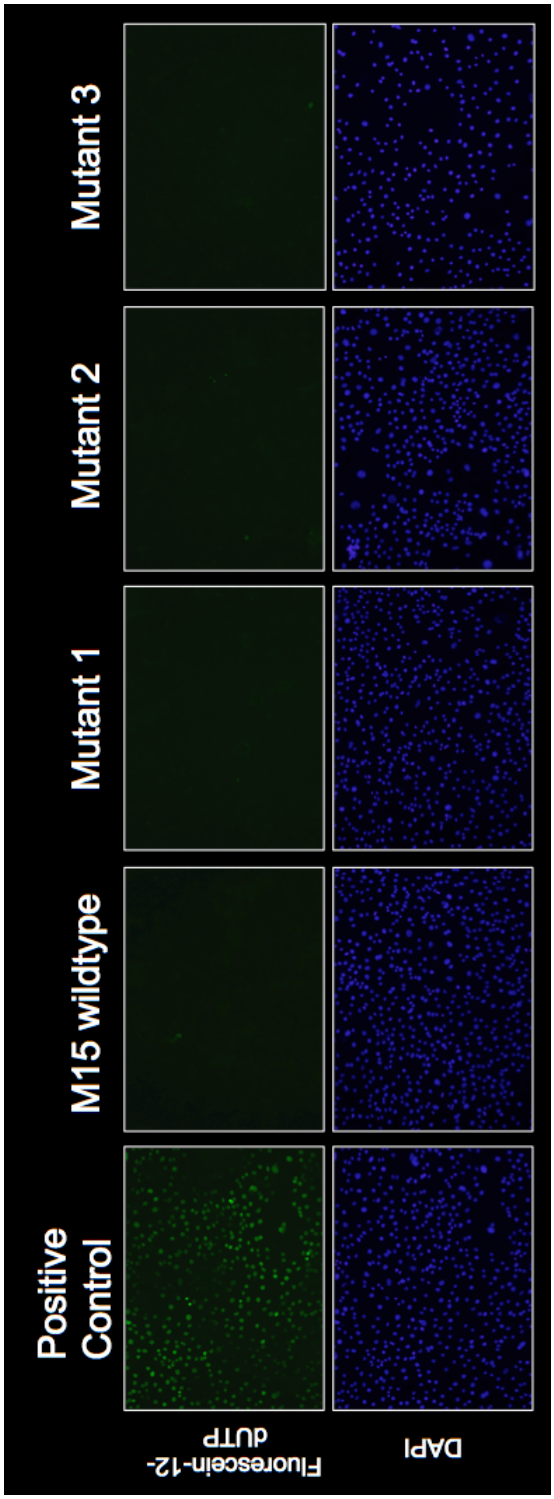


Figure 16. Apoptosis assay using wildtype M15 cells and three SPHK2^{R101W} mutant cell lines. No differences in apoptosis between wildtype and mutant cells were detected. Incorporation of fluorescein-12-dUTP at 3'-OH ends of fragmented DNA was visualized by fluorescence microscopy.

3.4. DISCUSSION

Next-generation sequencing has recently facilitated the discovery of genes involved in sporadic Wilms tumor cases, resulting in significant progress in understanding how aberrant development leads to cancer. However, the gene mutations predisposing to familial Wilms tumor are generally unknown. Familial Wilms tumor is genetically heterogeneous, and in order to understand the mechanisms of tumor formation, the mutated gene(s) involved must first be identified. Fortuitously, next-generation sequencing allows for cost-efficient, high throughput surveying of the entire genetic landscape in an unbiased manner to identify genes causing disease.

Following a whole genome sequence analysis of members of a large multigenerational 19q-linked Wilms tumor family, we identified variants shared by affected/obligate carriers in the family and analyzed their potential to be candidate Wilms tumor predisposition genes. Using this family based sequencing approach, which included 5 affected family members and 1 spousal control, we were able to significantly reduce the number of candidate variants for further investigation. Of the 10 exonic variants identified from this analysis, one was located within the 19q linked region previously characterized by our lab. As a result, we have successfully identified mutations in *SPHK2,* which encodes sphingosine kinase 2, as a candidate familial Wilms tumor predisposition gene.

The SPHK2 variant (R136W) identified in our study is novel, co-segregates with the affected/obligate carrier status in the family, occurs at an evolutionarily conserved amino acid residue, and is predicted to be deleterious by several prediction algorithms. Identification of SPHK2 mutations in other Wilms tumor families, as well as sporadic Wilms tumor cases, further implicates SPHK2 as a gene contributing to

Wilms tumor development. SPHK2 variants are noted in a variety of cancers in the COSMIC database, and these are almost exclusively missense mutations. These data argue that the *SPHK2* mutations are functionally significant due to encoding a protein with an altered function. Our data suggests that the SPHK2^{R136W} is indeed deleterious to protein function by reducing nuclear localization in MCF7 and M15 cell lines. Because mutant SPHK2 can not properly localize to the nucleus, target gene expression is altered.

SPHK2 is expressed in the embryonic kidney, with its highest levels occurring in the metanephric mesenchyme during early nephrogenesis. Therefore, SPHK2 is highly expressed in the cells thought to give rise to Wilms tumors. Previous reports show significantly increased SPHK2 mRNA in WiT49 cell and Wilms tumor samples compared to matched normal tissues, indicating that SPHK2 may have a functional role in Wilms tumor. Not only is SPHK2 actively expressed in the fetal kidney, but components of its nuclear co-repressor complex such as HDAC1/2, are differentially expressed during nephrogenesis as well. Interestingly, HDAC1/2 has only recently been recognized as critical for kidney development. Inhibition of histone deacetylase 1 and 2 in the metanephric kidney results in the specific downregulation of genes driving differentiation (*WT1, EYA1, SALL1*). SPHK2 has an important role in chromatin regulation,[65] which is becoming increasingly recognized as playing a major role in tumorigenesis. SPHK2 itself has recently been demonstrated to have a significant role in leukemogenesis through its effect on HDAC2-dependent regulation of c-Myc expression. Our identification of *SPHK2* mutations in both familial and sporadic Wilms tumors (Figure 17), and the observation that the SPHK2^{R136W} results in modification of gene expression, suggests that, like mutations in *WT1* (encoding a

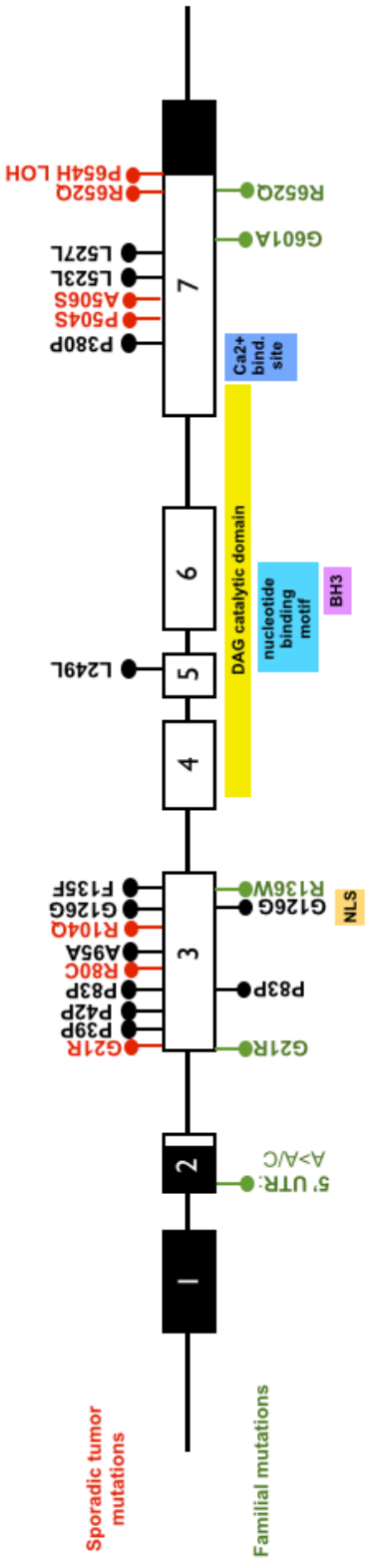


Figure 17. Summary of SPHK2 mutations identified in both familial and sporadic Wilms tumor cases from our study.

transcription factor), mutant SPHK2 alters gene expression in the developing kidney which results in the presence of cells predisposed to malignant transformation. In the case of SPHK2, the mechanism for this altered expression would be via altered chromatin regulation. Interestingly, as more pediatric sequencing studies are being completed, the number of mutations in genes involved in chromatin regulation and remodeling have outweighed the relative lack of mutations in classic oncogenes and tumor suppressors in childhood cancers, suggesting epigenetic events contribute more frequently to the incidence of pediatric cancer cases.

Analysis of tumors from families linked to chr17q, a region containing another familial Wilms tumor predisposition gene (known as FWT1), showed no loss of heterozygosity of polymorphic microsatellite markers within the 17q-linked region.[68] This was also observed with the 19q-linked tumors as well, indicating that these genes might not function as classical tumor suppressors. In our analysis, the SPHK2 R136W mutation was heterozygous in the germline, as well as in two tumor samples from affected members of the family. Furthermore, mutation analysis of all seven exons of SPHK2 showed no additional somatic mutation in either tumor. Likewise, all SPHK2 variants we identified in familial and sporadic cases were heterozygous as well, with the exception of one sporadic case (germline P654H, tumor specific LOH). Taken together, this suggests that mutation of the familial Wilms tumor predisposition genes may be necessary for the initiation of Wilms tumors, but are not sufficient to result in disease. Mutations elsewhere in the gene (regulatory or intronic regions), in additional genes, or the action of modifier genes might be necessary for tumor development from precursor lesions.

As mentioned previously, the kidneys of members of the Wilms tumor family carrying the germline SPHK2 mutation were "unusually rich" in nephrogenic rests, a precursor to Wilms tumor. In our lab's existing mouse model of Wilms tumor, in which mosaic somatic ablation of Wt1 in the context of upregulation of the fetal mitogen Igf2, we observe both tumors and foci of undifferentiated mesenchyme, like nephrogenic rests. Mice with mosaic somatic ablation of Wt1 in the absence of Igf2 upregulation also display such foci. Based on these data and the role of SPHK2 in epigenetic regulation of gene expression, we speculate that the missense mutation has an effect on normal metanephric mesenchyme differentiation such that foci of undifferentiated mesenchyme (nephrogenic rests) develop. This is consistent with our data demonstrating a significant increase in expression of *Eya1* in a M15 cell line in which the R100W mutation was knocked into the endogenous Sphk2 locus. *Eya1*, along with *Wt1* and *Pax2*, is a transcription factor expressed early in kidney development in the metanephric mesenchyme and is involved in the maintenance of nephron progenitors.[101] Deletion of *Eya1* in this population of cells results in premature differentiation and depletion of nephron progenitor cells by acting upstream of *Six2*. We can speculate that the increased expression of *EYA1* contributes to the formation of nephrogenic rests by maintaining the self-renewal of nephron progenitor cells, as deletion of *Eya1* results in the loss of *Six2* expression and premature differentiation of the progenitors.

In our model, *SPHK2* mutation carriers are at a higher risk of Wilms tumor development due to the presence of nephrogenic rests. Neoplastic transformation of these precursors then proceeds following an additional alteration acquired within the nephrogenic rest, such as upregulation of Igf2, an alteration commonly observed in

Wilms tumors, but which, alone, is not sufficient for tumor development. Additional experiments using mouse models will further clarify this hypothesis. However, it is clear that SPHK2 and sphingolipid metabolism is an emerging pathway in cancer development. Our identification of *SPHK2* as a candidate familial Wilms tumor predisposition gene and our analysis of the functional significance of its mutation in mouse mesenchymal cells suggests a role in regulation of kidney developmental genes.

CHAPTER 4: Identification of novel *DICER1* mutations in Wilms tumor families

4.1. INTRODUCTION

Since the identification of *DICER1* mutations in rare Wilms tumor cases in 2011, studies using next-generation sequencing have found that the microRNA processing pathway is sometimes mutated in Wilms tumors. As a result, these studies have provided a glimpse into the mechanism of Wilms tumor development through altered production of small RNAs regulating gene expression. The enzymes responsible for processing these microRNAs are crucial for mature microRNA expression and have thus emerged as key players of nephrogenesis and Wilms tumorigenesis.

4.1.1. MicroRNAs

MicroRNAs (miRNAs) are small, ~22 nucleotide noncoding RNAs that regulate gene expression post-transcriptionally.[69] With important roles in cellular processes such as differentiation, development and the cell cycle, dysregulation of miRNA biogenesis can serve as the key driving mechanisms of tumor formation.[70-74] Recent studies have identified miRNAs and the enzymes involved in their production as vital factors for the proper development of the kidney.[75] More than 170 miRNA families present in the embryonic kidney have been identified by deep sequencing. These miRNAs, including members of the miR-200 family, regulate the expression of genes critical for the formation of the mature kidney, as well as support differentiation and

survival of nephron progenitors, podocytes, juxtaglomerular cells and ureteric epithelium.[76-79]

The biogenesis of miRNAs proceeds in a stepwise progression that is subjected to multiple levels of regulation (Figure 18). This process begins when long, primary transcripts (pri-miRNAs) are transcribed in the nucleus by RNA Polymerase II. Interaction of pri-miRNAs with the nuclear microprocessor complex consisting of DiGeorge syndrome critical region 8 (DGCR8) and DROSHA, the nuclear counterpart of DICER1, results in the cleavage of the RNA to a 60-70 nucleotide pre-miRNA.[80-83] These hairpin-folded pre-miRNAs are exported to the cytoplasm by exportin 5 (XPO5).[84-86] In the cytoplasm, pre-miRNAs interact with DICER1, where they are cleaved into ~20 nucleotide double stranded RNAs.[87] These ~20-nucleotide mature miRNAs are loaded into the RNA-induced silencing complex (RISC), where they bind to target mRNAs. Each miRNA can regulate the expression of 50-100 mRNAs by binding through sequence complementarity of its 8-nucleotide seed sequence to target sites occurring in the 3'-UTR of genes and silencing by miRNA-mediated degradation or disruption of the ribosomal complex. Consequently, proper cleavage of the miRNAs by DICER and DROSHA is important in maintaining the correct seed sequence for recognition of target mRNAs.

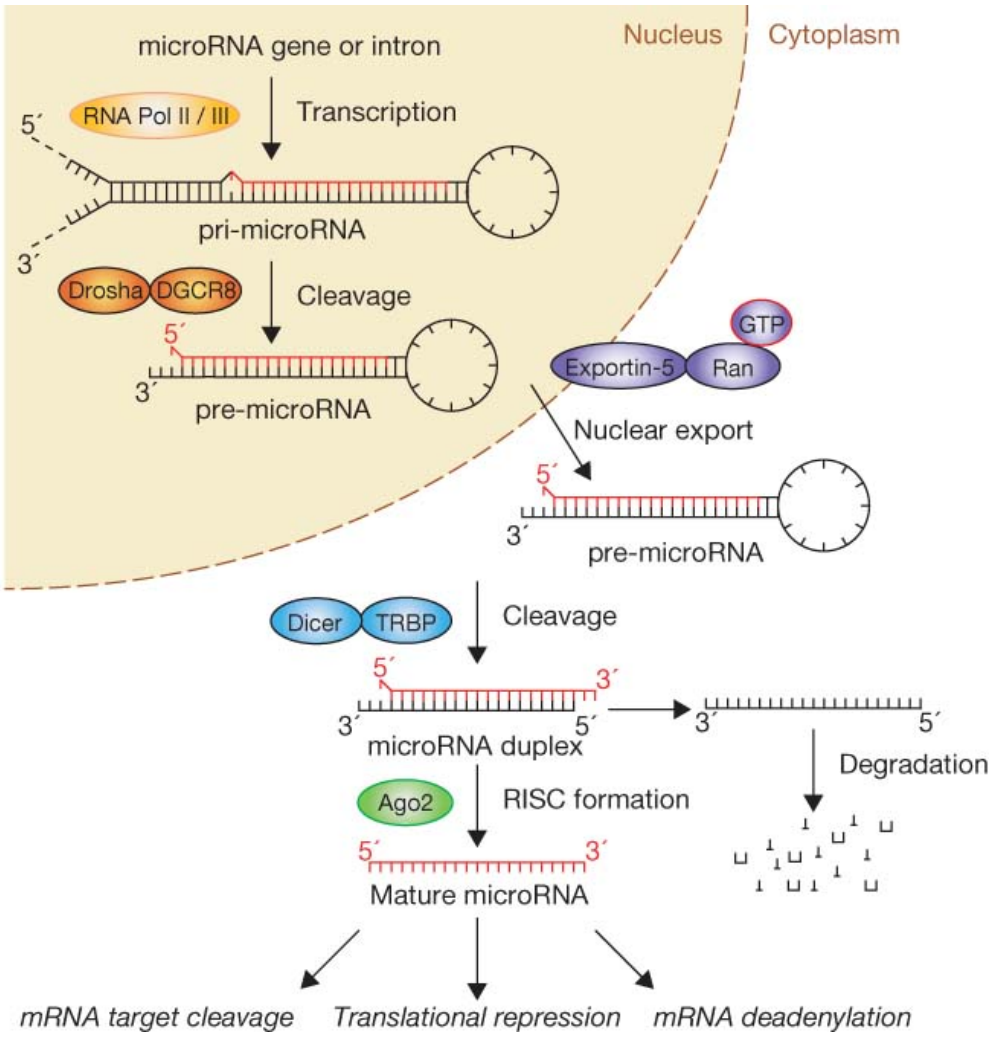


Figure 18. Micro RNA biogenesis pathway. miRNAs are transcribed in the nucleus by RNA Polymerase II as pri-miRNAs. Interaction with Drosha and DGCR8 results in the cleavage of the pri-miRNA into a shorter ~60-70 nucleotide pre-miRNA. Pre-miRNAs are exported to the cytoplasm by Exportin-5, where they interact with Dicer and TRBP2 for further cleavage into ~22 nucleotide mature miRNAs. Each mature miRNA is estimated to silence the expression of up to 100 target mRNAs.

4.1.2. The miRNA pathway in Wilms tumor

The miRNA pathway has recently emerged as a critical regulator of nephrogenesis.[17,32] Perturbation of this pathway in specific cell types and precise times during renal development results in abnormal renal morphogenesis and cancer. Indeed, mutations in components of the miRNA pathway have recently been identified in a subset of Wilms tumors. Exome sequencing of poor outcome Wilms tumors, as well as high-risk blastemal Wilms tumors revealed heterozygous mutations in *DROSHA* and *DGCR8* in 15% of these selected tumors. Heterozygous mutations in the RNase IIIb domain of *DROSHA* have been found in ~12% of favorable histology Wilms tumors.[16] In the majority of *DROSHA*-mutant tumors, a recurrent mutation (E1147K) affecting a metal ion binding residue of the RNase IIIb domain was found. These somatic mutations act through a dominant-negative mechanism to impair processing of primary miRNAs, resulting in the reduction of global miRNA biogenesis.[16-17,32] Gene and miRNA expression profiling has demonstrated that the miR-200 family is one of the subsets of miRNAs reduced in tumors, correlating with downregulation of genes involved in the mesenchymal to epithelial transition.

Given the multistep progression of miRNA processing, it is not surprising that other genes involved in the stepwise progression of miRNA biogenesis have also been found somatically mutated in Wilms tumors. Mutations in *DGCR8*, a component of the microprocessor complex and cofactor of *DROSHA*, as well as *XPO5* and *TARBP2* have also recently been identified through next-generation sequencing efforts.[16] Occasionally, mutations in several miRNA pathway genes have been found in the same tumor providing multiple hits to the miRNA pathway and underscoring the

importance of a normal functioning miRNA processing system. One component of the miRNA pathway that is mutated less frequently than *DROSHA* and *DGCR8* is *DICER1*.

4.1.3. *DICER1* and DICER1 Syndrome

One of the enzymes at the center of the micro RNA biogenesis pathway is *DICER1*. Located on chromosome 14q32.13, *DICER1* is a large multidomain enzyme containing twenty-six coding exons (Figure 19). *DICER1*, functioning downstream of *DROSHA* and *DGCR8*, encodes a cytoplasmic endoribonuclease essential for the generation of mature microRNAs from pre-miRNAs.[88] In order for DICER to perform its ribonuclease activity, the protein folds such that the RNase IIIa and IIIb domains form an intramolecular dimer, creating a catalytic core whereby each domain cleaves one strand of the double-stranded pre-miRNA substrate. The RNase IIIa and IIIb domains cleave the 3p and 5p miRNA strands, respectively. This requires proper positioning of the RNA substrate over the catalytic domains, which is accomplished by a number of DICER domains. Among these are the platform and PAZ domains, which recognize the 5' terminal phosphate and 3' 2-nucleotide overhang of the pre-miRNA, respectively. Additionally, the DExD/H box helicase domain is thought to form a claw-like structure to clamp down on the loop structure of the substrate, ultimately allowing correct positioning of the dsRNA along the connector helix, where a positively charged surface interacts with the phosphate backbone residues of the pre-miRNA. Following the favorable interaction between the pre-miRNA and DICER, cleavage of each strand takes place at a fixed distance from the terminal ends of the pre-miRNA, generating a mature miRNA of 22 nucleotides.

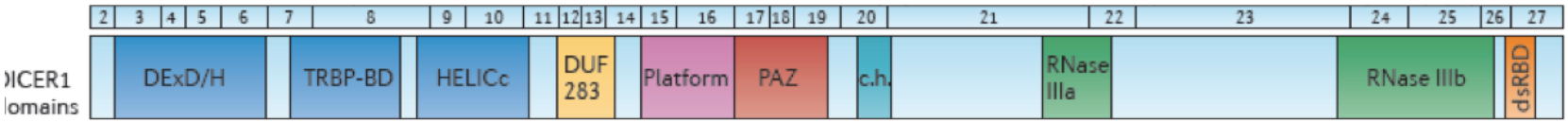


Figure 19. DICER1, located on chr14q, contains 27 exons and contains a number of well characterized domains. Figure modified from Foulkes, Nature Reviews Cancer, 2014.

Dicer1 knockout in the mouse is embryonic lethal, suggesting that loss of Dicer1 activity is incompatible with normal development. *Dicer1* is expressed in the fetal kidney, where it is critical for renal morphogenesis. Particularly, kidney-specific ablation of Dicer1 in early metanephric mesenchymal cells results in renal dysgenesis and apoptosis. These mice show nephron progenitors that are initially correctly specified with normal ureteric bud outgrowth, however subsequent ureteric branching and nephron differentiation do not occur.[70]

Most *DICER1* mutations identified in Wilms tumors to date are somatic alterations, although germline mutations are occasionally observed.[89-91] Mutations in *DICER1* were first identified to cause familial pleuropulmonary blastoma in 2009.[92] Since then, biallelic mutations have been identified in a variety of tumor types, including Wilms tumors, which suggests a “two-hit” tumor suppressor model.[91] In this model, germline mutations, frequently truncating, are inherited, whereas somatic missense mutations occurring at hotspots within the RNase IIIb domain result in a second hit contributing to the development of a well-defined tumor predisposition syndrome referred to as DICER1 Syndrome.[89] This is a pleotropic cancer syndrome which commonly includes phenotypes such as pleuropulmonary blastoma, cystic nephroma, ovarian sex cord-stromal tumors, and multinodular goiter, as well as less frequent phenotypes including differentiated thyroid carcinoma, pituitary blastoma, pineoblastoma, and Wilms tumor, among others. Penetrance of the various phenotypes of DICER1 Syndrome is incomplete, with more than half of mutation carriers exhibiting none of the phenotypes associated with the syndrome.

4.2. Materials and Methods

4.2.1. WTX593 and WTX279 Families

Following informed consent, blood was drawn and DNA isolated from affected individuals and/or obligate carriers from 48 WT families. Genomic DNA from two of these families was then used for whole genome sequencing. These two families did not carry *WT1, WTX, CTNNB1* or *TP53* mutations. The first family (WTX593) consisted of four individuals diagnosed with unilateral WT, with the age of diagnosis ranging from 38-57 months. Additionally, thyroid colloid cysts, multinodular goiter (MNG), and kidney multi colloid cysts, and/or pulmonary cysts were noted in family medical records. Fresh-frozen Wilms tumor was available from one individual.

In the second family (WTX279) two individuals were diagnosed with WT, one at 3.5 years of age and the other at 13 years of age. The latter, the proband, relapsed after two years and was found by CT scan to have two small pulmonary nodules which, however, were too small to biopsy. Additionally the father of the proband had one kidney removed at 3 years of age due to a suspected, but poorly documented, premalignant condition.

Whole genome sequencing of peripheral blood DNA was carried out on two affected individuals and one obligate carrier individual in the WTX593 family as described in chapter 2.

4.2.3. Protein modeling of DICER1 mutation

Protein modeling of the *DICER1* G803R mutation was performed using the Protein Homology/analogy Recognition Engine version 2.0 (Phyre^2;

http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) following the manufacturer's instructions.

4.2.4. Mutation analysis of DICER1 in familial Wilms tumor.

Sanger sequencing was performed (as described in chapter 2) to confirm the *DICER1* variant detected by WGS and to assess 47 additional families for mutations in the same *DICER1* exon (exon 15) or seven other *DICER1* exons (exons 8, 9, 13, 14, and 24-26) at which mutations have been reported in WT. Primer sequences are listed in Table 10.

	Forward primer	Reverse primer
DICER1 exon 7	TTCTCACTACTGCAGTATTGATACCTT	GAGCCGCATTAAGCATATTTTC
DICER1 exon 8	AAATCCCAGTTAAACCCCAC	TCACATCACAACACAGGACG
DICER1 exon 9	TAAATCACCGTCGCCAAATC	AAATCACTCTACAGCTACCTCATGG
DICER1 exon 13	TTTTACTAGGCAGGACTTTTAAAGATG	AAGTGTTCATGGTGCATGATTC
DICER1 exon 14	TTTGCAGTCCAGCTCATATTG	AAGCTGTGAATCGGAGAAAG
DICER1 exon 15	CTGAGATCCGCTGACTTAC	CCCCTCAAAATTAACAGGTA
DICER1 exon 24	TGTGGGGATAGTGTAAATGCTTC	TGCCGTCAGAACTCTGAAAC
DICER1 exons 25-26	TGGACTGCCTGTAAAAGTGG	TGAACTTTTCCCCTTTGATG

Table 10. DICER1 exon primer sequences for Sanger sequencing

4.3. RESULTS

4.3.1. Whole genome sequencing identifies DICER1 as candidate predisposition gene.

Blood DNA from three affected/obligate carrier individuals from the WTX593 family was used for whole genome, paired-end sequencing to identify the Wilms tumor predisposing alteration in the family (Figure 20a). The mean whole genome sequencing coverage across all genes for the three individuals sequenced (Figure 20b) ranged from 32-45X, with 97% coverage of genes at >10X depth. Variants shared by all three individuals were identified and filtered based on their allele frequency reported in the general population (<1%), with novel variants occurring in protein coding regions assigned a higher priority for investigation. Of the 113 variants left after filtering, a novel missense mutation in *DICER1* was of greatest interest because mutations in this gene have recently been found in *DICER1* syndrome families which include a member with WT.[16] The identified mutation (c.2407G>A, chr14: 95574690, hg19) occurred in exon 15 of *DICER1* and results in a glycine to arginine substitution at codon 803 (G803R) (Figure 21). The variant was not present in the Single Nucleotide Polymorphism database (dbSNP) (http://www.ncbi.nlm.nih.gov/SNP), NHLBI Exome Variant Server (http://evs.gs.washington.edu/EVS) nor 1000 Genomes database. This alteration was predicted to be deleterious by SIFT, PolyPhen2 and MutationTaster (0, 0.991, 0.999, respectively). The mutation occurred at an evolutionary conserved amino acid residue (PhyloP – 0.999) within the platform domain of DICER1, which is involved in

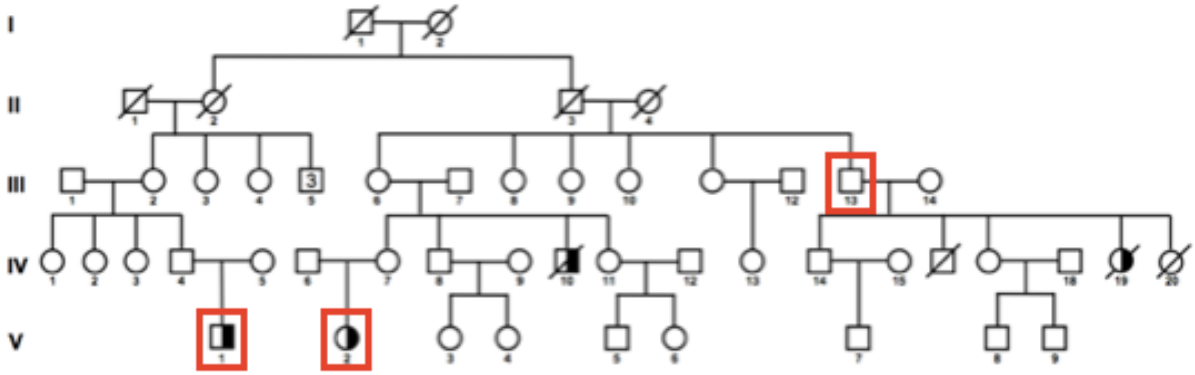


Figure 20a. WTX593 pedigree. Of the four unilateral Wilms tumor cases in the family, blood was obtained from two. Whole genome sequencing (indicated by red boxes) was performed on these family affected family members (V-1, V-2), as well as an obligate carrier (II-13)

SUMMARY OF VARIATION

Individual	V-2	V-1	III-13
Gender	Female	Male	Male
Coverage	44.5	32.3	41.5
SNPs	3,766,935	3,683,999	3,680,540
Insertions	335,055	318,537	317,365
Deletions	350,486	332,452	335,863
Shared, Novel Variants			
Synonymous SNPs	60		
Nonsynonymous SNPs	80		
Nonsense SNPs	1		
Insertion	46		
Deletion	32		

Figure 20b. The coverage and total number of variants across the whole genome for each individual sequences. This includes data for calls where Q score >20. Additionally, the number of shared, novel exonic variants are listed.

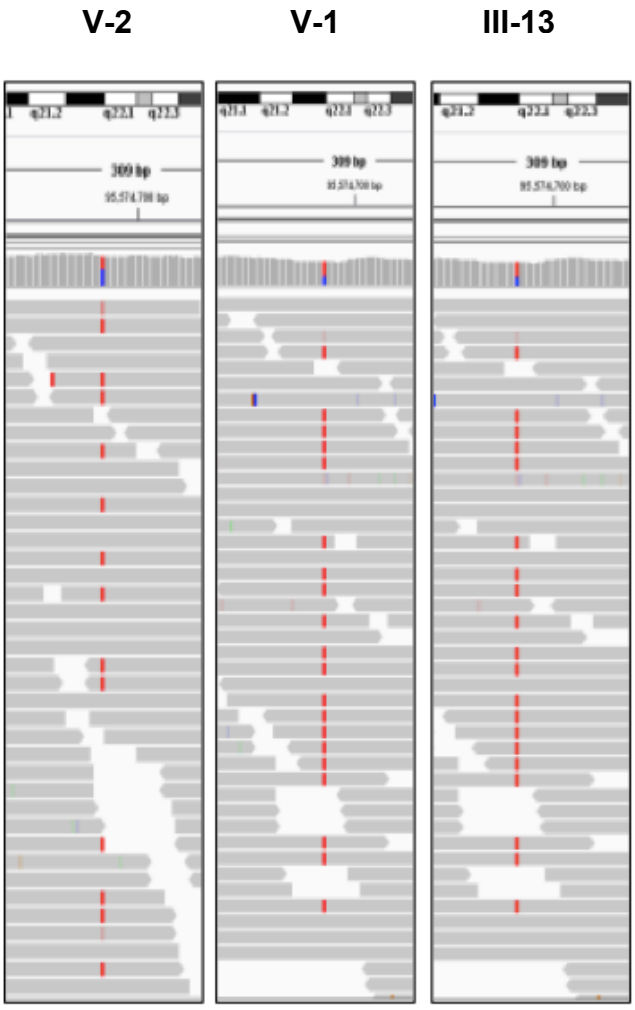


Figure 21. Next-generation sequenced reads visualized with IGV showing the *DICER1* G803R variant in each family member. The variant is present in ~50% of sequence reads.

the recognition of the 5’ phosphate group of pre-miRNAs, and between residues that interact with the 5' phosphate of miRNAs.[21]

DICER1 is actively expressed in early renal progenitor cells (Figure 22) and knockout studies show a failure of nephron differentiation, suggesting DICER1 activity is required for kidney development.

4.3.2. *DICER1* G803R cosegregates within the “DICER1 Syndrome” family

We confirmed the alteration by Sanger sequencing and extended our analysis to additional family members to estimate the penetrance of the variant in the family. It was noted that other phenotypes of DICER1 syndrome were present in some family members. In total, we detected the G803R mutation in eleven family members, some of whom displayed one or more of the phenotypes associated with familial DICER1 syndrome, including lung cysts, thyroid cysts, MNG, cystic nephroma and WT (Figure 23a).

4.3.3. Loss of wildtype allele observed in tumor

The one tumor available for study was found to be homozygous for the G803R mutation by Sanger sequencing (Figure 23b). Such homozygosity for a germline DICER1 mutation has not previously been reported for Wilms tumors, although this has been observed in cases of DICER1-associated pineoblastoma and pituitary blastoma.[99,100] Typically in Wilms tumors arising in individuals with a DICER1 truncating mutation, somatic hotspot mutations occurring within the RNase IIIb domain have been described.

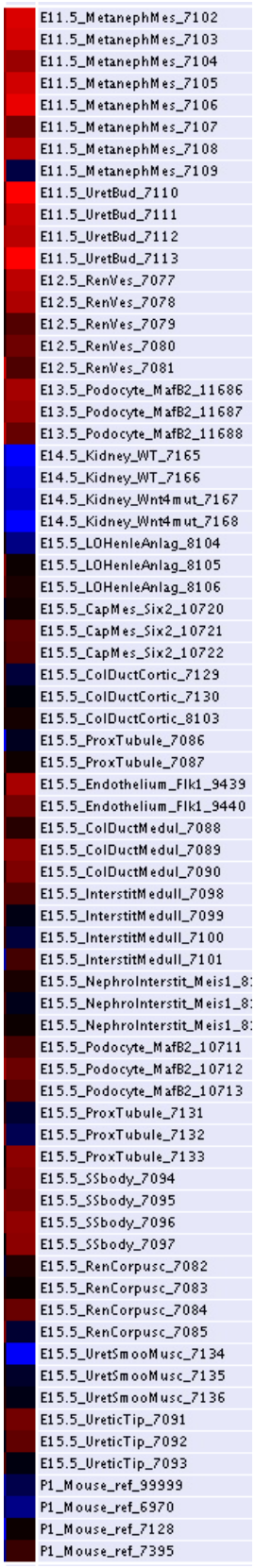


Figure 22. *DICER1* expression in the fetal kidney using gudmap.org. *DICER1* is actively expressed in the metanephric mesenchyme and ureteric bud at E11.5.

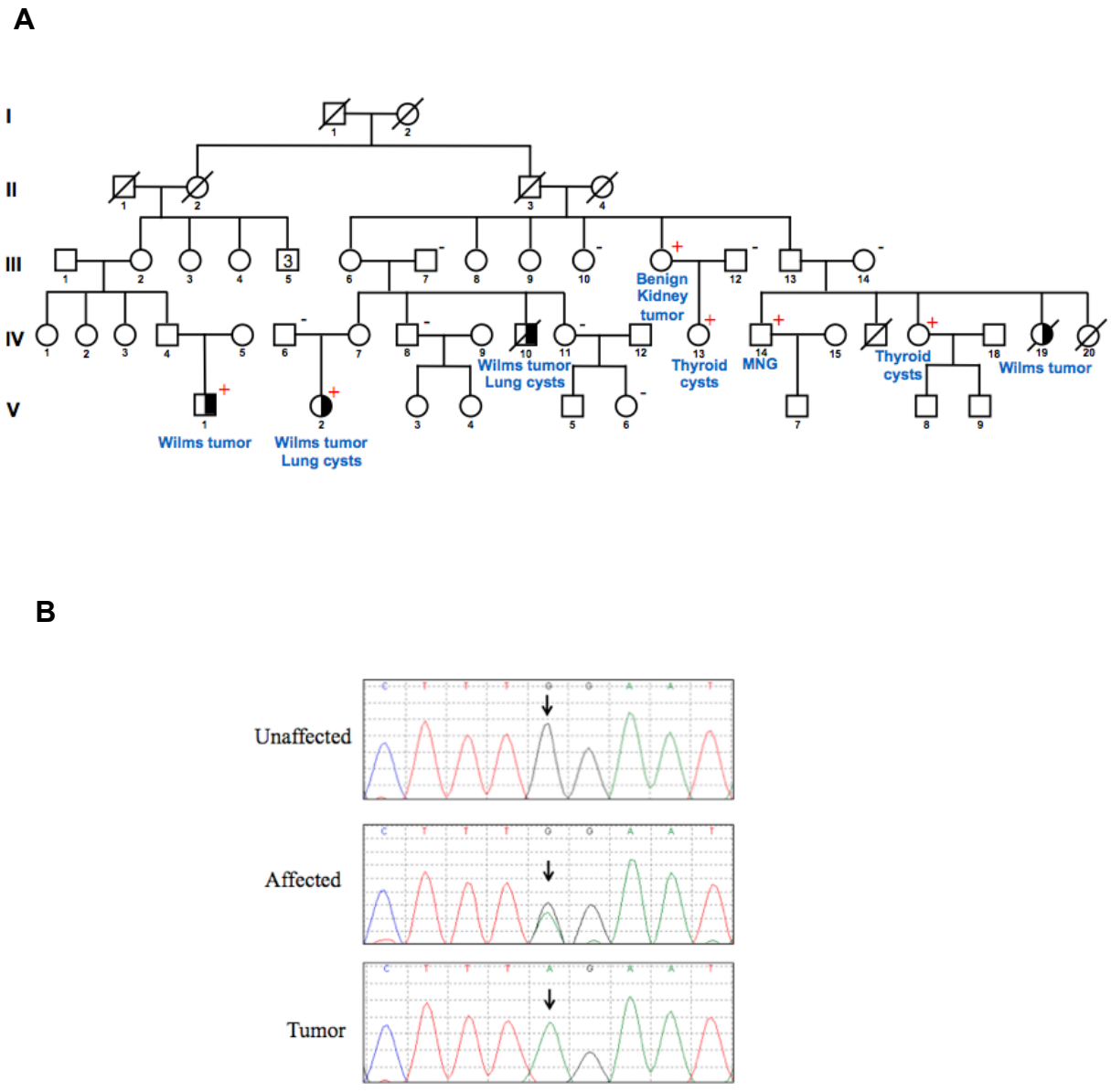


Figure 23. *DICER1* G803R co-segregates with the affected/obligate carrier status with the family. "+" indicates the individual carriers the G804R variant. "-" indicates family members who do not carry the variant. Other DICER1 Syndrome phenotypes are listed on the pedigree. WT: Wilms tumor, TC: thyroid cysts, MNG: multinodular goiter, BKT: benign kidney tumor. Sanger sequencing of the only tumor available from the family shows loss of the wildtype allele.

3.3.4. Protein modeling of DICER1 mutation suggests a functional effect

Protein modeling using the Protein Homology/analogy Recognition Engine v2.0 (Phyre2; www.sbg.bio.ic.ac.uk/phyre2) suggests that the G803R mutation disrupts an alpha helix at the end of the platform domain, potentially altering the 5' terminal phosphate recognition pocket of the domain (Figure 24). Residues within this pocket, and adjacent to the G803R variant, interact with the terminal phosphate group on the 5p miRNA strand. This is consistent with the SIFT, PolyPhen and Mutation Taster scores which predict the substitution of a glycine for an arginine at codon 803 is deleterious.

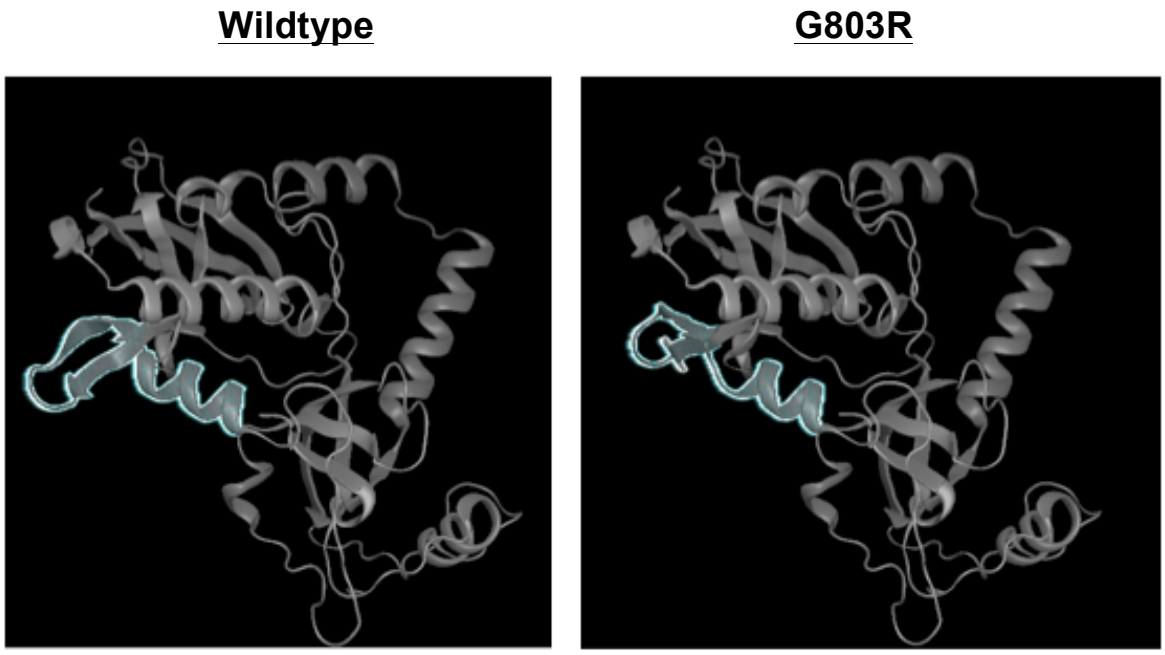


Figure 24. Protein modeling performed in Phyre2. Shown is a protein model of the platform and PAZ domains of wildtype DICER1 (left) and G803R (right). Highlighted in blue are the differences between each model, due to the introduction of an arginine residue at codon 803. This results in the alteration of an alpha-helix within the platform domain.

4.3.5. *DICER1* mutation analysis in familial Wilms tumor

Sanger sequence analysis of one member each of 47 additional WT families revealed that one carried a novel frameshift mutation (c.2399delG, p.R800fsX5) in the same exon (exon 15) as that in WTX593 (Figure 25a). This mutation also occurs in the evolutionarily conserved platform domain, was not present in the 1000 Genomes database, dbSNP or the Exome Variant Server, and segregates with the affected status in the family (Figure 25b). It is predicted to encode a truncated protein lacking an intact platform domain and the remaining half of the protein, including the PAZ domain and the RNase IIIa and IIIb domains. Tumor samples were not available for this family; LOH could not be assessed.

Sanger sequencing of seven other exons in which *DICER1* mutations have been previously identified in WT cases (exons 8, 9, 13, 14, 24, 25, 26) was also performed in 46 other Wilms tumor families. No germline variants were observed in these families.

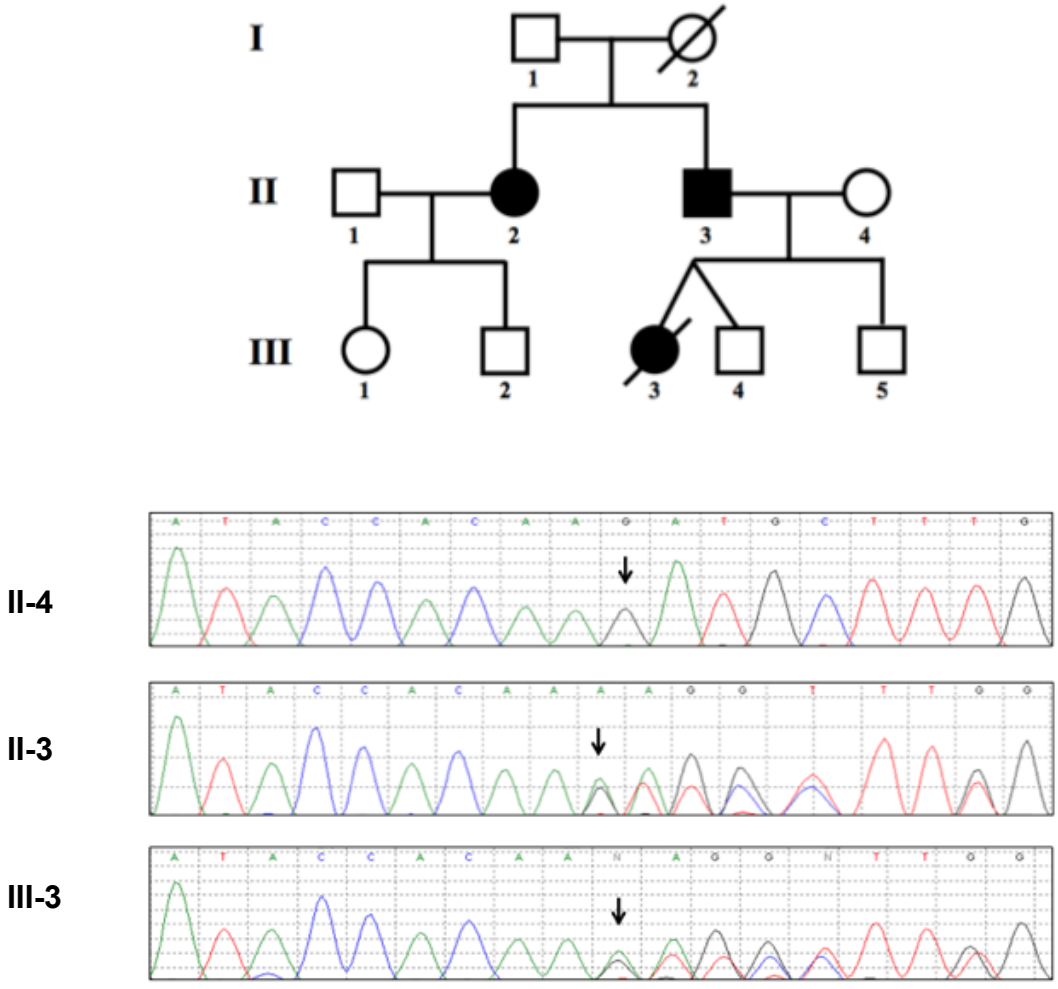


Figure 25a. Identification of an additional germline *DICER1* mutation in a Wilms tumor family. The variant is a frameshift deletion predicted to result in the premature truncation of the protein. Sanger sequencing showed the variant segregates with the affected status in the family.

4.4. DISCUSSION

Mutations in miRNA pathway genes have recently been identified in approximately 15% of favorable histology Wilms tumors.[17] The majority of these are somatic mutations clustering at RNase domains of *DROSHA* and *DICER1*, two endoribonucleases critical for proper miRNA maturation. While the vast majority of these are somatic mutations, germline mutations in *DICER1* have been identified in several tumor types and are associated with a pleiotropic tumor predisposition syndrome called DICER1 syndrome. In this syndrome, inherited truncating mutations and somatically acquired missense mutations occurring at or adjacent to metal-ion binding residues (D1705, D1709, D1713,G1809, E1813) within the RNaseIIIb domain are typically found in families. The result of these biallelic mutations is impaired 5p miRNA biogenesis, including major miRNA families such as let-7 and miR-200, while 3p miRNA levels are unaltered. This reduction of 5p miRNAs is thought to contribute to the development of DICER1-associated phenotypes through the deregulation of target genes, including members of the miR-200 family with roles in the mesenchymal to epithelial transition critical for normal kidney development.

To date, only six DICER1 families with a member diagnosed with WT have been reported in the literature and each of these has only one WT case. In these rare families, similar to non-WT DICER1 families, germline *DICER1* nonsense mutations were observed, along with an additional somatic missense mutation occurring in trans in the tumor.[16] In our study, we identified two Wilms tumor families with novel germline *DICER1* mutations. Both germline mutations occurred in the platform domain, which has a role in recognizing the terminal phosphate group of the 5p strand of pre-miRNAs for correct positioning over the RNase III domains for cleavage (Figure 26).

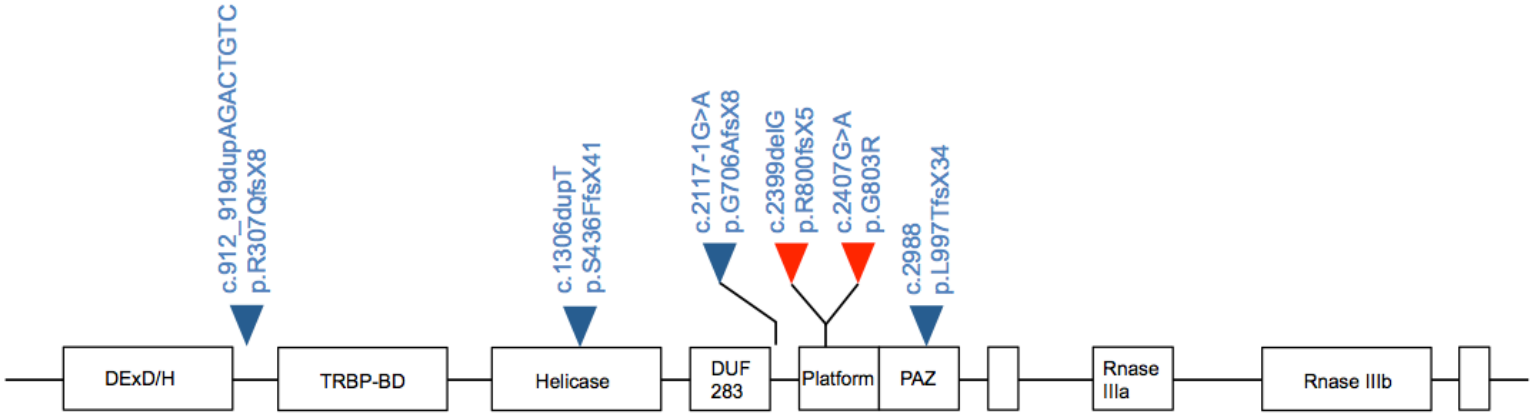


Figure 25b. Germline *DICER1* mutations found in Wilms tumors. Indicated in blue are those previously reported, while those in red were identified in our study.

Previous studies have shown mutations in this domain result in reduced processing efficiency of mature miRNAs.[21] Consistent with what has been observed in other DICER1 families, the novel mutation (R800Xfs5) we identified in WTX279 is predicted to result in the premature truncation of the protein. WTX279 is a small WT family, and, while it is not known whether individuals in this family displayed phenotypes associated with DICER1 syndrome, pulmonary nodules were detected in the proband by CT scan. These nodules, however, were too small for biopsy and it is unknown if they were WT metastases or were similar to the cystic pulmonary phenotype that is observed in DICER1 Syndrome. Uniquely, multiple family members were diagnosed with Wilms tumor in this small family; DICER1 families reported to date have only a single WT case.

The diagnosis, not only of Wilms tumor, but also of MNG, lung cysts, thyroid and kidney cysts in the WTX593 family is consistent with this being a “DICER1 syndrome family”. In total, 11 individuals in this family were found to carry the germline mutation, with 2/11 diagnosed with only WT, 2/11 diagnosed with WT and lung cysts, 2/11 with thyroid cysts, 1/11 with MNG and 1/11 with kidney cysts. In addition to the presence of multiple WT cases, WTX593 is unique on several counts. First, it carries a germline DICER1 missense mutation rather than a frameshift or nonsense mutation. Only one other family (in which only multiple nodular goiters are observed) is reported to have sustained a germline missense mutation. Interestingly, this mutation (Ser839Phe) also occurred in the platform domain. The G803R mutation we identified in WTX593 is predicted to be deleterious by a number of functional prediction algorithms and occurs at an evolutionarily conserved amino acid residue, suggesting this mutation is

functionally significant. Moreover, protein modeling suggests this alteration disrupts an alpha helix in the platform domain.

Secondly, the germline G803R mutation in WTX593 was reduced to homozygosity in a tumor from a family member. In contrast, almost all the second, somatic mutations observed in tumors from individuals with germline *DICER1* mutations are missense mutations that occur in the RNase IIIb domain of *DICER1*. To date, loss of heterozygosity with retention of the germline mutation has only been reported in two DICER1-associated phenotypes – pineoblastoma and pituitary blastoma.[99,100] LOH has not previously been reported in Wilms tumors with DICER1 mutations.

Additionally, like WTX279, WTX593 is a family where a germline *DICER1* mutation results in multiple cases of Wilms tumor. The penetrance of the G803R allele with respect to Wilms tumor is notably much higher than other germline *DICER1* mutations (which, as noted above, are almost invariably truncating mutations). The basis for this increased penetrance is not known. Obviously, ascertainment bias is a factor as the families being studied were originally selected due to the occurrence of two or more Wilms tumor cases. However, is it still curious that families ascertained for PPB and other, DICER1-associated phenotypes only rarely have a member with Wilms tumor (much less two or more), while in both families reported here, the *DICER1* mutations were relatively highly penetrant with respect to Wilms tumor.

Interestingly, while the penetrance of the two novel *DICER1* mutations we identified in these two families was unusually high, the ages at which individuals in these families were diagnosed with Wilms tumor was notably older than average Wilms tumor cases in general. The median age of diagnosis for unilateral Wilms tumor overall is ~39 months of age while that of unilateral cases in familial context is

35 months. Wilms tumor families, however, greatly vary with respect to age at diagnosis, with some families exhibiting typically younger ages, while other families consist of cases with later than average ages of onset. The observation that both families carrying *DICER1* mutation reported here have members diagnosed at late ages suggests that this may be a common feature of DICER1-related Wilms tumor.

Our data demonstrates that mutation of *DICER1* is associated with the development of familial Wilms tumor and describes interesting and novel findings with respect to the missense mutation we identified. Our identification of germline DICER1 mutations in 4% of Wilms tumor families is consistent with the recent reports describing mutations in multiple members of the miRNA biogenesis pathway in sporadic Wilms tumors.

CHAPTER 5: Identification of a recurrent 570kb duplication on chr2p encompassing *NBAS, DDX1, MYCN* and *MYCNOS* in Wilms tumor families.

5.1. INTRODUCTION

Copy number variation of specific regions of the genome have been identified in sporadic Wilms tumors using array comparative genomic hybridization (aCGH) and SNP arrays.[93] These include losses at 17p, 1p, 16q and gains at 1q, 8, 12, suggesting that these regions harbor genes that play a role in Wilms tumorigenesis. After sequencing the whole genome of two individuals from a large Wilms tumor family, we detected a 570kb duplication on chr2p24 which predisposes to the development of Wilms tumor. This region includes 4 genes; *NBAS, DDX1, MYCN* and *MYCNOS*, and was initially selected for further investigation due to the established role of *MYCN* in both development and Wilms tumorigenesis.

5.1.1. *MYCN* functional activity and it's role in kidney development

MYCN belongs to the MYC family of transcription factors with well-documented roles in the control of cell proliferation and differentiation.[94] *MYCN* is a conserved basic helix loop helix transcription factor that forms a heterodimeric complex with Max and binds to a CANNTG E-box consensus sequence found upstream of promoter target sequences. This complex acts as a transcriptional regulator by activating or repressing hundreds of target genes through the interaction with various nuclear protein complexes.

MYCN expression, primarily detected in the developing embryo, is absent from adult tissues suggestive of a principal responsibility in mediating embryonic development. Indeed, experiments in mice have shown that *MYCN* is essential for

both neurogenesis and nephrogenesis. Homozygous ablation of *MYCN* in mice results in embryonic lethality at E10 due to defects of the central nervous system, kidney, lung, gut and heart. In the developing brain, MYCN is involved in the induction of a stem-cell like state by blocking differentiation and activating the pluripotency factors, KLF2, KL4 and LIN28B. Moreover, reduction of *MYCN* expression levels in neuroblastoma cell lines leads to induction of a differentiated phenotype. *MYCN* is particularly interesting with respect to Wilms tumor biology, as it is highly expressed at early stages of development in the metanephric mesenchyme. Its expression becomes increasingly restricted as embryonic development proceeds and is undetectable in adult human kidney tissue. Reduced or loss of *MYCN* expression in cultured embryonic kidneys results in renal hypoplasia due to a decrease in cell proliferation in ureteric bud and mesenchymal tissues, while apoptosis was not affected.[38] Thus, *MYCN* expression in the embryonic kidney is critical for nephrogenesis.

5.1.2. *MYCN* alterations in Wilms tumors

5.1.2.1. Amplification

Amplification of one particular region on chromosome 2p24, including the *MYCN* locus, has been described in several pediatric cancers, including Wilms tumor.[18,33-36] In 10% of Wilms tumor cases, increased *MYCN* copy number and expression has been described, suggesting this region has a functional role in the development of Wilms tumors. *MYCN* is also somatically amplified in other childhood cancers such as neuroblastoma, medulloblastoma, retinoblastoma, alveolar rhabdomyosarcoma, and astrocytoma. Of these, *MYCN* has been most extensively

studied in the context of neuroblastoma, where it is amplified in roughly 25% of cases and is associated with poor prognosis. In Wilms tumors, somatic *MYCN* amplification has primarily been observed in the context of sporadic cases. However, a germline chr2p24 duplication encompassing *MYCN* and detected by array comparative genomic hybridization (aCGH) was reported in one family with a history of nephroblastomatosis and Wilms tumor.[95] This duplication was subsequently observed in an additional patient with bilateral Wilms tumor from an independent study, as well as in our analysis (described below).

5.1.2.2. *MYCN* is often coamplified with other flanking genes

When *MYCN* amplification has been observed in both neuroblastoma and Wilms tumor, it frequently includes additional neighboring gene(s), including orthinine decarboxylase (*ODC*), syndecan-1, ribonucleotide reductase (*RRM2*), *ALK, DDX1, NBAS*, and *MYCNOS*. Genes coamplified with *MYCN* in Wilms tumor cases include the 5' region of *NBAS*, as well as the complete coding sequences of *DDX1* and *MYCNOS*. For example, *MYCN* and *DDX1* were both somatically amplified in a Wilms tumor-derived xenotransplanted cell line.[96] Moreover, germline duplication of a 570kb region including these genes has been recently observed in Wilms tumor patients. To date, *DDX1* amplification without concomitant *MYCN* amplification has not been described, suggesting *MYCN* as the primary gene responsible for tumorigenesis. Interestingly, a genome-wide association study (GWAS) identified a significant association signal within this duplicated region when examining 1,488 Wilms tumor samples.[97] Collectively, this suggests germline copy number alterations in this region could predispose to Wilms tumor.

DDX1, a DEAD box helicase that has roles in the alteration of RNA secondary structure, is frequently coamplified with *MYCN* (*DDX* is 340kb 5' of *MYCN*, and is telomeric of *MYCN* on 2p24). Like *MYCN*, *DDX1* is also strongly expressed during early kidney development (Figure 27). A recent report described a nuclear role of *DDX1*, where it acts as a cofactor of *DROSHA* in ovarian cancer cells in the recruitment of a specific subset of pri-miRNAs for processing.[98] This provides an intriguing connection to the miRNA biogenesis pathway, which has very recently been implicated in Wilms tumor development.

5.1.2.3. Point mutations

Not only is *MYCN* amplified in sporadic Wilms tumors, but a somatic heterozygous point mutation (c.131C>T, p.P44L) within the aurora kinase A domain has recently been identified in ~4% of Wilms tumors by several groups using next-generation sequencing.[18] This variant was somatic in all cases and is believed to be a gain of function mutation due to its occurrence at a highly conserved amino acid residue. Tumors carrying the P44L mutation do not appear to have morphologically distinguishing features to suggest the alteration results in a specific histology or clinical stage. In one study, 40% of P44L-mutant tumors also had modest *MYCN* amplification. Interestingly, a bilateral Wilms tumor case described separate *MYCN* alterations in one patient: an *MYCN* point mutation (P44L) in one tumor with *MYCN* amplification in the contralateral tumor, suggesting multiple mechanisms of *MYCN* dysregulation in Wilms tumor development.[18]

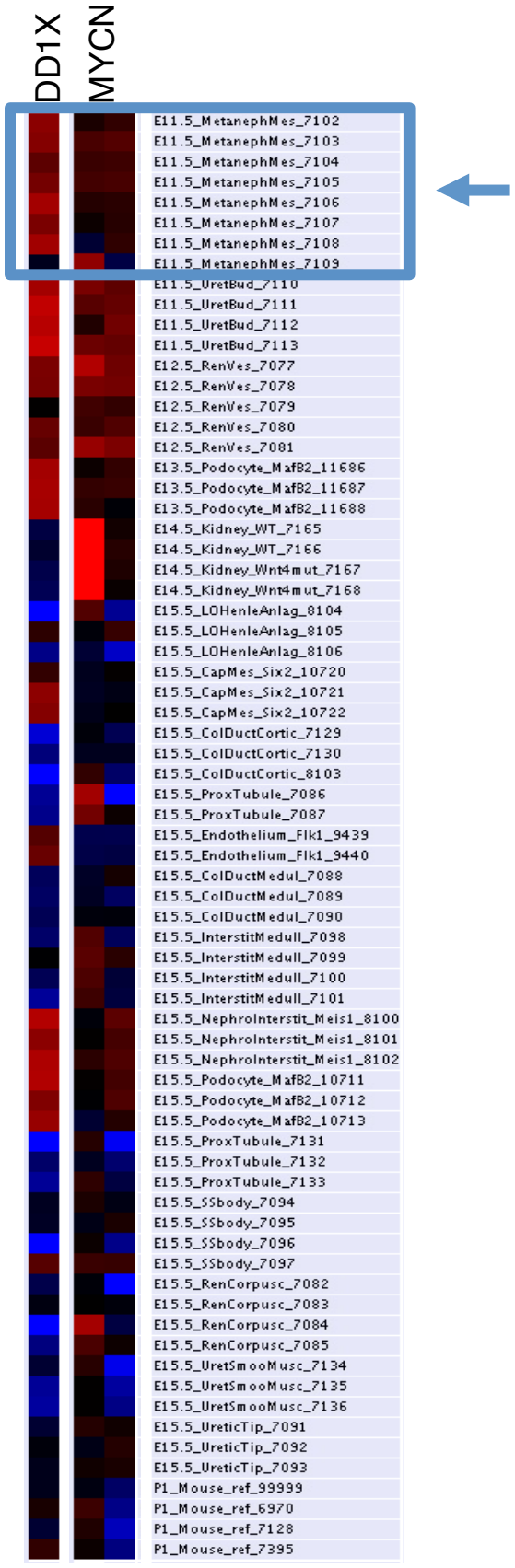


Figure 26. Expression of DDX1 and MYCN in the mouse fetal kidney from gudmap.org. Both are actively expressed throughout early kidney development, specifically in the metanephric mesenchyme (blue box).

5.2. Materials and Methods

5.2.1. WTX614 and WTX637 families

Following informed consent, blood was drawn and DNA isolated from members of the WTX614 and WTX637 families. Mutation analysis of *WT1, WTX* and *CTNNB1* revealed no germline mutations in the affected individuals, thus the predisposition gene mutation was determined to be unknown. The WTX614 family consists of eight Wilms tumor cases, with the age of diagnosis ranging from 14 to 204 months. Additional phenotypes in the family included medulloblastoma and astrocytoma. Tumor samples were not available for these patients for further analysis. The WTX637 family contained five Wilms tumor cases, with the age of diagnosis ranging from 5 to 72 months. Tumor samples were also not available from these family members.

Whole genome sequencing of peripheral blood DNA was carried out on two affected individuals in the WTX614 family (IV-3 and III-11) as described in chapter 2.

5.2.3. Copy number analysis using qPCR

Copy number variation was confirmed using a SYBR green qPCR assay with primers targeting *DDX1* and *MYCN*. Primer sequences are listed in Table 11. Genomic DNA, gene-specific primers and SYBR green were analyzed using the Applied Biosystems 3730 real time machine. Additionally, primers targeting ** on chromosome 5 (one of the regions where copy number alterations are rarely observed in Wilms tumor) served as a control.

	Forward primer	Reverse primer
MYCN	GAGAGGACACCCTGAGCGATT	CAAGACATACGAGCACTAACAAAGG
DDX1	CCCAACTGATATCCAGGCTGAA	AGTGTGTCCCCAGCTACCAATC

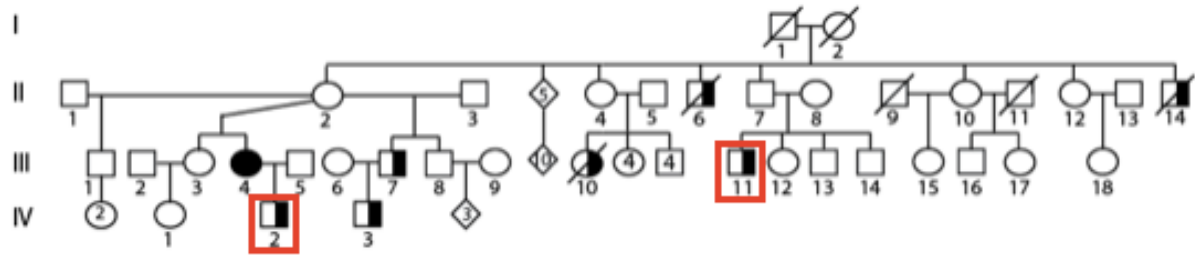
TABLE 11. qPCR primer sequences for MYCN and DDX1 copy number analysis.

5.3. Results

5.3.1. Whole genome sequencing identifies recurrent 570kb duplication on chr2p

We sequenced the whole genome of two affected individuals in a large, multi-generation WT family (Figure 27a) consisting of 8 WT cases to identify the predisposing alterations contributing to the development of FWT. The mean coverage across the genome was 30X for IV-3 and 33X for III-11, with 96.5% and 96.8% of positions covered by at least 10X (Figure 27b). Comparison of shared variants resulted in the identification of 344 novel, exonic variants (excluding synonymous variants). These genes were assessed for known developmental or differentiation-related roles in the embryonic kidney. After surveying single nucleotide variants within the family, we were not able to identify a convincing candidate gene for further evaluation. Although several genes appeared to be potential candidates based on our criteria for identifying predisposition genes, none of them had a clear role in renal morphogenesis. Next, we examined shared structural variants, including large insertions, deletions, duplications, inversions and translocation calls generated using CASAVA v.1.8. This analysis examined read depth to identify regions where copy number gains or losses were suspected due to read density deviation from the average coverage calculated across the whole genome. This resulted in the identification of a 570kb duplication on chromosome 2p24 encompassing 4 genes: *NBAS*, *DDX1*, *MYCN* and *MYCNOS* (Figure 28). This duplication was initially selected for further investigation because *MYCN* has previously been shown

A



B

SUMMARY OF VARIATION

Individual	IV-2	III-11
Gender	Male	Male
Coverage	33.6	30.5
SNPs	3,673,595	3,676,171
Insertions	315,113	317,527
Deletions	329,229	327,892
Shared Novel Variants		
Synonymous SNPs	83	
Nonsynonymous SNPs	94	
Stop Gain SNPs	0	
Stop Loss SNPs	0	
Insertion	37	
Deletion	22	

Figure 27. WTX614 pedigree. Individuals whose whole genomes were sequenced are indicated by red boxes. The summary of genetic variation is also listed including data for calls where Q score >20.

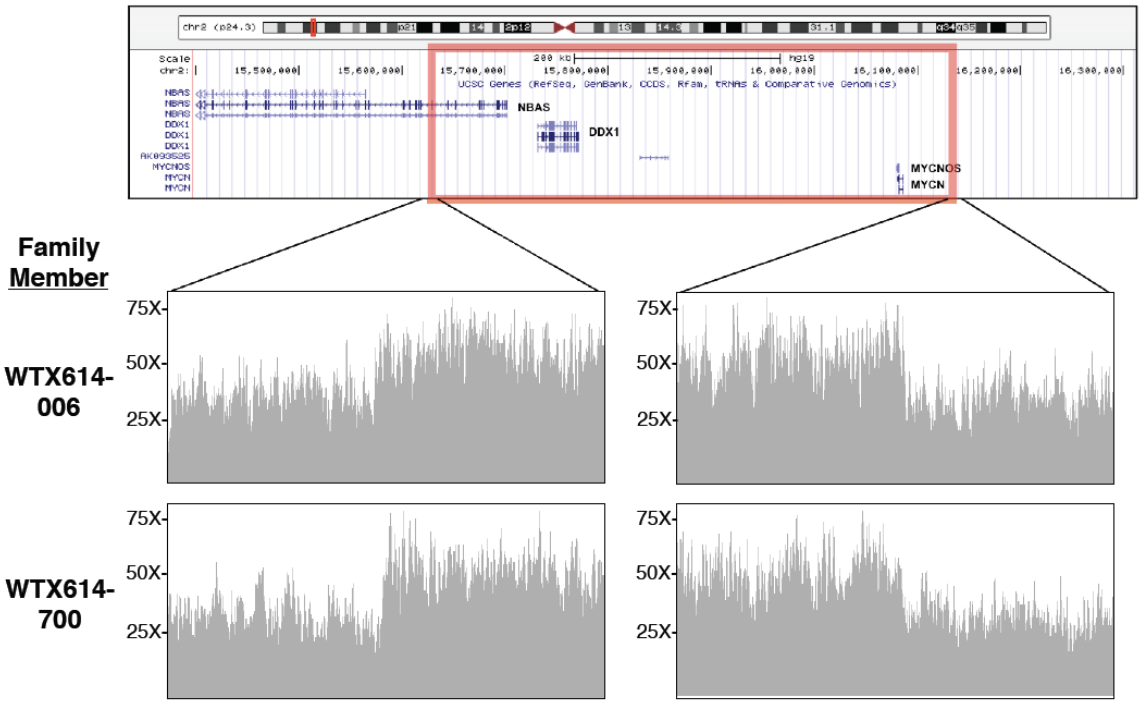


Figure 28. The chr2p24 570kb duplication includes four genes: NBAS, DDX1, MYCN and MYCNOS. Coverage plots for the boundaries of the duplication are shown from each individual sequenced. Copy number variation was identified after extracting read counts for each position across the whole genome, segmentation of the genome into 50bp segments, and calculating read depth within the 50bp segments compared to average coverage across the whole genome. Regions with increased or decreased read density compared to the whole genome were classified as gains or losses, respectively.

to be amplified in ~10% of Wilms tumors, with germline duplication of this gene potentially serving as a mechanism for increased expression.

5.3.2. Duplication co-segregates with affected/obligate carrier status

Quantitative PCR (qPCR) was used to confirm the duplication using primers targeting an exonic region of *DDX1* and *MYCN*. Extending this analysis to additional family members revealed the duplication segregates with the affected/obligate carrier status and is absent in all spousal controls tested (Fig29). In total, 9 members of the family carried the duplication and 7, including 4 spousal controls, did not carry the duplication. Each affected and obligate carrier individual in the family, for whom we had blood samples for, carried the duplication. 5 of 9 (55%) carriers developed Wilms tumor, consistent with the characteristic reduced penetrance of Wilms tumor in families.

5.3.3. Copy number analysis in familial Wilms tumor identified second family with co-segregating duplication

Additionally, we assessed this region in one member from each of 47 other WT families to determine the frequency of the duplication in familial WT (Figure 30). This analysis resulted in the identification of an additional family (WTX279) that harbors the 570kb duplication. The duplication was found to segregate with the affected/obligate carrier status in WTX279 and is absent in spousal controls (Fig30). 4/6 (66%) carriers in the family developed Wilms tumor. In total, we detected this recurrent duplication in 2/47 (4.2%) WT families.

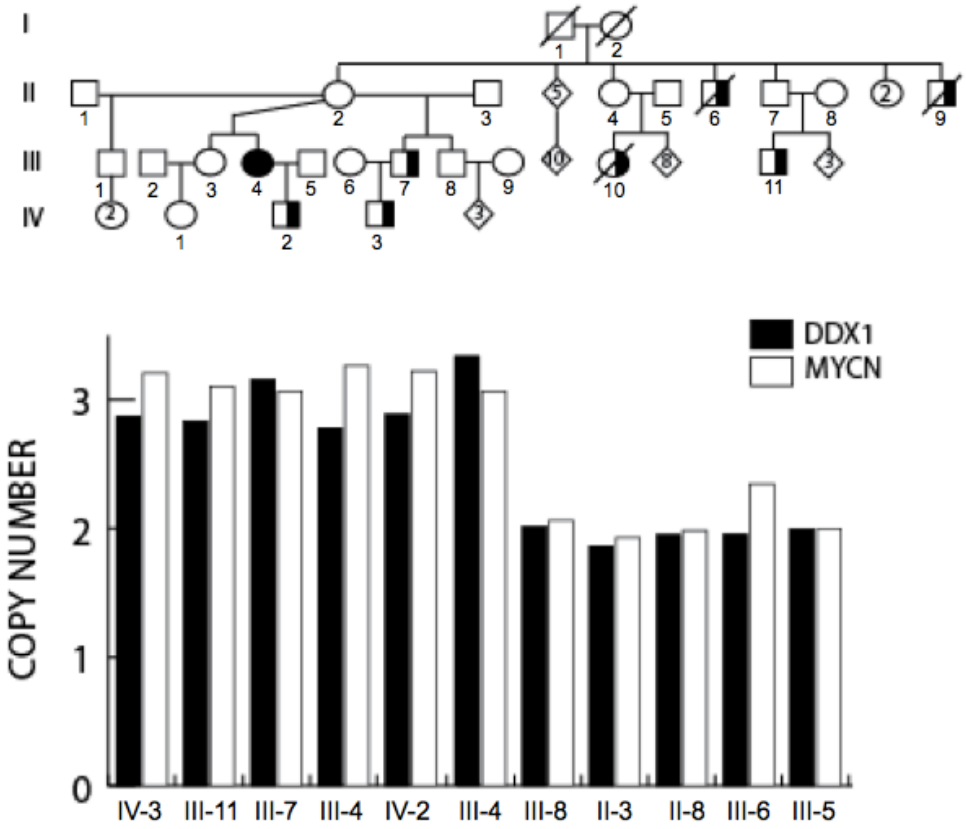


Figure 29. Copy number was confirmed using qPCR targeting DDX1 and MYCN. The 570kb duplication on chr2p24 cosegregates with the affected/obligate carrier status in the family. + indicates carriers of the duplication, - indicates those members who don't carry the duplication.

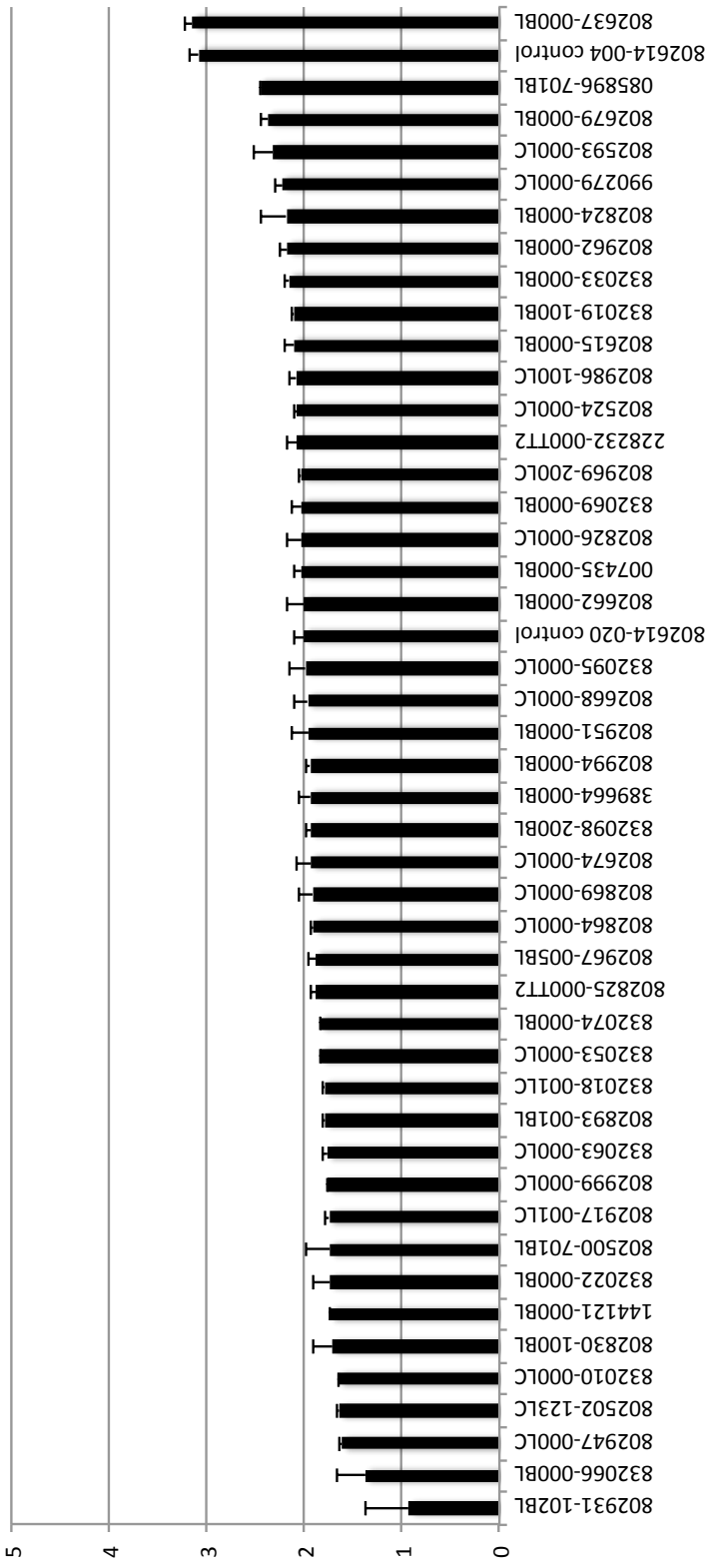


Figure 30. Copy number analysis of DDX1 in one member of 48 Wilms tumor families identified an additional family with copy number gain of DDX1.

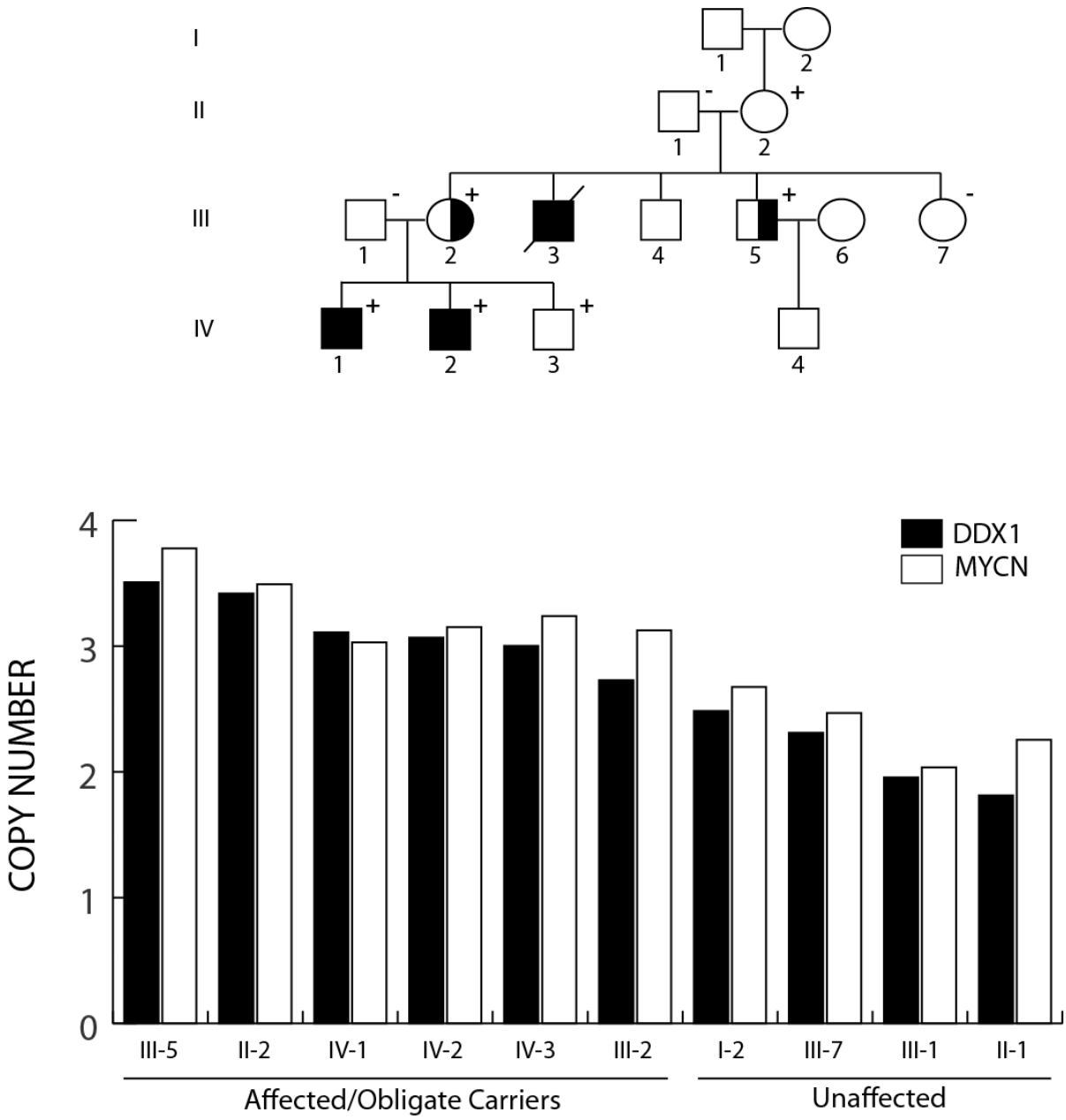


Figure 31. Copy number was confirmed using qPCR targeting *DDX1* and *MYCN*. The 570kb duplication on chr2p24 cosegregates with the affected/obligate carrier status in the WTX279 family. + indicates carriers of the duplication, - indicates those members who don't carry the duplication.

5.4. Discussion

This chapter describes two Wilms tumor families with a chr2p microduplication that cosegregates with the affected/obligate carrier status in each family. The identification of multiple families harboring this duplication from our study and others, as well as previous reports of amplification of *MYCN* in approximately 10% of Wilms tumors suggests that copy number variation of this region is associated with the development of Wilms tumor. Further support comes from recent studies that described a recurrent *MYCN* point mutation (P44L) in ~4% of Wilms tumors, which has also been identified previously in neuroblastoma. This somatic recurrent mutation, occurring at a highly conserved residue, has been suggested to be an activating mutation, although studies have not functionally validated its significance. It is likely that an activating mutation in *MYCN* could have the same functional effect as *MYCN* amplification on downstream targets. Interestingly, this recurrent point mutation occurs within the aurora kinase A interaction domain which, when bound to *MYCN*, is involved in insulating *MYCN* from *FBXW7*-mediated degradation. *FBXW7*, an element of an E3 ubiquitin ligase complex involved in the direct targeting of *MYCN* for ubiquitylation and proteasome-mediated degradation, is somatically mutated in approximately 4% of Wilms tumors. Therefore, gain of *MYCN* or the loss of genes involved in it's regulation (*FBXW7*) potentially have overlapping downstream functions allowing exploitation during kidney development to give rise to Wilms tumors.

Interestingly, alterations in *MYCN* copy number have previously been found in other childhood cancers in addition to Wilms tumor. These include retinoblastoma, neuroblastoma, astrocytoma, and medulloblastoma, where *MYCN* gain is associated with an adverse outcome. In the WTX614 family, two individuals carrying the

duplication were also diagnosed with medulloblastoma and astrocytoma. There has been no evidence reported previously suggesting a germline duplication encompassing *DDX1* and *MYCN* predisposes to these tumor types. However, two reports have described somatic amplification of this region in both medulloblastoma and astrocytoma tumors, although this appears to be a rare event. While constitutional *MYCN* duplications have been found to predispose individuals to neuroblastoma, none of the patients in this study who carry the germline duplication, to our knowledge, developed neuroblastoma.

The mechanism of the reduced penetrance in these families is not known, and tumor tissue was unavailable from both families carrying the germline duplication, eliminating the possibility of molecular analysis of tumors for additional alterations. Intriguingly, a patient carrying a germline partial 2p trisomy including *MYCN* developed a neuroblastoma with dramatic amplification of *MYCN*, suggesting the germline duplication might predispose to further somatic amplification. While we can't rule out this scenario in the Wilms tumor patients from our study, a recent report describing a bilateral Wilms tumor patient carrying this chr2p germline duplication showed no further amplification of *MYCN* in the tumor sample. Another possibility for the reduced penetrance is that, while the chr2p duplication predisposes to Wilms tumor, additional mutation events at other loci have to occur during the progression of tumorigenesis. A previous report described a bilateral Wilms case in which one tumor contained an *MYCN* P44L mutation and *TP53* V173L mutation, whereas the contralateral tumor contained MYCN amplification and a different TP53 mutation (I195T).[18] This depicts a situation where *MYCN* amplification or mutation might not be

sufficient for tumor development without an additional mutation event occurring before or after *MYCN* alteration.

Collectively, these data suggests that a constitutional chr2p duplication predisposes to Wilms tumor and patients carrying this duplication are at increased risk of developing Wilms tumor. Additionally, children with germline duplications of genes within this region potentially are at risk for the development of other tumors types, such as medulloblastoma and astrocytoma, although more studies are needed for definitive proof. The finding that *MYCN* might predispose to other tumor types in addition to Wilms tumor suggests that therapies targeting the *MYCN* pathway might be of significant value in other tumors as well.

CHAPTER 6: Summary and Future Directions

6.1. Summary of Data

The etiology of Wilms tumor is genetically heterogeneous. Like sporadic Wilms tumor cases, the genes and associated pathways contributing to familial Wilms tumor initiation and progression are diverse. The recent advancements in next-generation sequencing technologies have provided investigators with a high-throughput tool to investigate the genetic landscape of Wilms tumor, providing a glimpse into the complexity of the genetics of this childhood cancer. The research presented in this thesis is a perfect example of the characteristic genetically heterogeneous nature of Wilms tumor.

Whole genome sequencing was performed using blood DNA from members of three Wilms tumor families. Bioinformatic filtering within each family identified shared variants for further consideration as candidate predisposition genes. DNA samples from an additional 45 Wilms tumor families were then analyzed for mutations in these candidate predisposition genes. Our analysis resulted in the successful identification of Wilms tumor predisposition genes in each family.

We identified *DICER1* exon 15 germline mutations in 2/48 (4%) Wilms tumor families. These novel mutations occurred at evolutionarily conserved amino acid residues within the DICER1 platform domain and were predicted to be deleterious. Interestingly, there were multiple cases of Wilms tumor within each family, which has not been observed in DICER1-associated Wilms tumors to date. Loss of the wildtype allele in one tumor was also an additional novel finding. Our data provides support that mutations in the miRNA biogenesis machinery can give rise to familial Wilms

tumors and is consistent with recent reports describing alteration of these genes in Wilms tumor.

We also identified a germline 570kb duplication on chr2p24 in 2/48 (4%) Wilms tumor families. This duplication encompasses *NBAS, DDX1, MYCN* and *MYCNOS*. This duplication was previously observed in members of a multigenerational Wilms tumor family, as well as in a blood sample from a patient with bilateral Wilms tumor. *MYCN* and *DDX1* amplification have been associated with sporadic Wilms tumors and our data reinforces the involvement of these genes in Wilms tumor development. We also observed other phenotypes, in addition to Wilms tumor, in members of the families that carry this duplication. As a result, this duplication might increase the risk of medulloblastoma and astrocytoma as well.

Finally, we identified *SPHK2* as a 19q candidate predisposition gene. The *SPHK2* mutation we identified cosegregates in a large Wilms tumor family, has not been reported in variant databases, occurs at an evolutionarily conserved amino acid residue, and is functionally significant. This mutation alters SPHK2 subcellular localization and expression of target genes in MCF7 and M15 cell lines. Mutations identified in other Wilms tumor families and sporadic tumors provides additional evidence implicating this gene in familial Wilms tumor predisposition. Significantly, the identification of *SPHK2* mutations in Wilms tumors opens up new avenues of research into how altered sphingolipid metabolism and signaling contribute to Wilms tumor development. Our data represents a significant advancement in understanding the etiology of familial Wilms tumor and associates alteration of new pathways with Wilms tumor development.

6.2. Future directions of research

The first step in understanding the mechanism of Wilms tumor development is to identify genes responsible for its formation. The goal of the research presented in this Thesis was to expand the etiology of Wilms tumor occurring in families, thus making a significant contribution to our understanding of cancer development in families. The next step will be to build on the novel findings presented here, in order to learn more about nephrogenesis, identify how these genes and their associated pathways contribute to aberrant differentiation and tumor formation, and recognize potential therapeutic targets.

DICER1 has been previously associated with Wilms tumor development in the context of DICER1 Syndrome, and progress is currently being made to understand its role, as well as that of the other miRNA biogenesis genes, in relation to Wilms tumor development. Potential future areas of research include characterization of the G803R mutation with respect to tumor formation. Possible experiments could include transfecting a kidney cell line with *DICER1* wildtype and mutant constructs, followed by a miRNA cleavage assay and miRNA expression profiling to determine which miRNAs are deregulated. This would provide clues as to which genes are important in Wilms tumor development.

Recent studies are starting to uncover the significance of *MYCN* alterations in Wilms tumors and the mechanisms whereby they contribute to cancer. Interestingly, our data suggests that another gene within the duplicated region shared by affected family members might also contribute to Wilms tumor. *DDX1* was recently shown to interact with Drosha and DGCR8, where it recruits specific subsets of pri-miRNAs for processing. Future research could focus on the specific pri-miRNAs that DDX1 binds

to, in order to 1) determine if DDX1 has a functional effect on Wilms tumor development and is not just amplified due to the proximity of MYCN, and 2) determine potential genes regulated by the DDX1-specific subset of miRNAs.

SPHK2 has never been implicated in Wilms tumor development. Consequently, how this gene functions with respect to kidney development and initiation of tumorigenesis is unknown, although our data and others have started to suggest a role in regulation of gene expression. Future research using mouse models could focus on how *SPHK2* contributes to development of the embryonic kidney, as well as how mutating *SPHK2* results in Wilms tumors. Specifically, introduction of $SPHK2^{R101W}$ into the mouse will provide a more complete picture of the involvement of *SPHK2* in Wilms tumor development. Also, *SPHK2* has not been extensively characterized and comparison of the SPHK2 amino acid sequence to other proteins has failed to identify functional domains. Future areas of research focusing on a comprehensive characterization of *SPHK2* would greatly facilitate research into its functions. Additionally, further characterization of the known functions of *SPHK2* would clarify the apparent discrepancies reported in the literature with respect to localization, impact on proliferation and apoptosis.

Finally, our whole genome sequencing analysis revealed the predisposing alterations in a small fraction of Wilms tumor families in our cohort. Whole genome or exome sequencing in other families will no doubt uncover other novel and unexpected causal genes.

BIBLIOGRAPHY

1. Breslow N., Beckwith J.B., Ciol M., Sharples K. Age distribution of Wilms' tumor: report from the National Wilms' Tumor Study. Cancer Res. 48, 1653-1657 (1988).
2. Knudson A.G. Jr., Strong L.C. Mutation and cancer: a model for Wilms' tumor of the kidney. J Natl Cancer Inst. 48, 313-324 (1972).
3. Breslow N.E., Beckwith J.B., Perlman E.J., Reeve A.E. Age distributions, birth weights, nephrogenic rests, and heterogeneity in the pathogenesis of Wilms tumor. Pediatr Blood Cancer. 47, 260-267 (2006).
4. Huff V., Amos C.I., Douglass E.C., Fisher R., Geiser C.F., Krill C.E., Li F.P., Strong L.C., McDonald J.M. Evidence for genetic heterogeneity in familial Wilms' tumor. Cancer Res. 57, 1859-1862 (1997).
5. Saxen L., Sariola H. Early organogenesis of the kidney. Pediatr Nephrol. 1, 385-392 (1987).
6. Rivera M.N., Haber D.A. Wilms' tumour: connecting tumorigenesis and organ development in the kidney. Nat Rev Cancer. 5, 699-712 (2005).
7. Dressler G.R. The cellular basis of kidney development. Annu Rev Cell Dev Biol. 22, 509-529 (2006).
8. Md Zin R., Murch A., Charles A. Pathology, genetics and cytogenetics of Wilms' tumour. Pathology. 43, 302-312 (2011).
9. Li C.M., Guo M., Borczuk A., Powell C.A., Wei M., Thaker H.M., Friedman R., Klein U., Tycko B. Gene expression in Wilms' tumor mimics the earliest committed stage in the metanephric mesenchymal-epithelial transition. Am J Pathol. 160, 2181-2190 (2002).

10. Aiden A.P., Rivera M.N., Rheinbay E., Ku M., Coffman E.J., Truong T.T., Vargas S.O., Lander E.S., Haber D.A., Bernstein B.E. Wilms tumor chromatin profiles highlight stem cell properties and a renal developmental network. Cell Stem Cell. 6, 591-602 (2010).
11. Huff V. Wilms' tumours: about tumour suppressor genes, an oncogene and a chameleon gene. Nat Rev Can. 11, 111-121 (2011).
12. Beckwith JB, Kiviat NB, Bonadio JF. Nephrogenic rests, nephroblastomatosis, and the pathogenesis of Wilms' tumor. Pediatr Pathol. 10:1–36 (1990).
13. Park S., Bernard A., Bove K.E., Sens D.A., Hazen-Martin D.J., Garvin A.J., Haber D.A. Inactivation of WT1 in nephrogenic rests, genetic precursors to Wilms' tumour. Nat Genet. 5, 363-367 (1993).
14. Charles A.K., Brown K.W., Berry P.J. Microdissecting the genetic events in nephrogenic rests and Wilms' tumor development. Am J Pathol. 153, 991-1000 (1998).
15. Ruteshouser E.C., Robinson S.M., Huff V. Wilms tumor genetics: mutations in WT1, WTX, and CTNNB1 account for only about one-third of tumors. Genes Chromosomes Cancer. 47, 461-470 (2008).
16. Rakheja D., Chen K.S., Liu Y., Shukla A.A., Schmid V., Chang T.C., Khokhar S., Wickiser J.E., Karandikar N.J., Malter J.S., Mendell J.T., Amatruda J.F. Somatic mutations in DROSHA and DICER1 impair microRNA biogenesis through distinct mechanisms in Wilms tumours. Nat Commun. 5, 4802 (2014).
17. Walz A.L., Ooms A., Gadd S., Gerhard D.S., Smith M.A., Guidry Auvil J.M., Meerzaman D., Chen Q.R., Hsu C.H., Yan C., Nguyen C., Hu Y., Bowlby R., Brooks D., Ma Y., Mungall A.J., Moore R.A., Schein J., Marra M.A., Huff V.,

Dome J.S., Chi Y.Y., Mullighan C.G., Ma J., Wheeler D.A., Hampton O.A., Jafari N., Ross N., Gastier-Foster J.M., Perlman E.J. Recurrent DGCR8, DROSHA, and SIX homeodomain mutations in favorable histology Wilms tumors. Cancer Cell. 27, 286-297 (2015).

18. Williams R.D., Chagtai T., Alcaide-German M., Apps J., Wegert J., Popov S., Vujanic G., van Tinteren H., van den Heuvel-Eibrink M.M., Kool M., de Kraker J., Gisselsson D., Graf N., Gessler M., Pritchard-Jones K. Multiple mechanisms of MYCN dysregulation in Wilms tumour. Oncotarget. 6, 7232-7243 (2015).
19. Kreidberg J.A., Sariola H., Loring J.M., Maeda M., Pelletier J., Housman D., Jaenisch R. WT-1 is required for early kidney development. Cell. 74, 679-691 (1993).
20. Sim E., Smith A., Szilagi E., Rae F., Ioannou P, Lindsay M.H., Little M.H. Wnt-4 regulation by the Wilms' tumour suppressor gene, WT1. Oncogene. 21, 2948-2960 (2002).
21. Koesters R., Ridder R., Kopp-Schneider A., Betts D., Adams V., Niggli F., Briner J., nov Knebel Doeberitz M. Mutational activation of the beta-catenin proto-oncogene is a common even in the development of Wilms' tumors. Cancer Res. 59, 3880-3882 (1999).
22. Maiti S., Alam R., Amos C.I., Huff V. Frequent association of beta-catenin and WT1 mutations in Wilms tumors. Cancer Res. 60, 6288-6292 (2000).
23. Rivera M.N., Kim W.J., Wells J., Driscoll D.R., Brannigan B.W., Han M., Kim J.C., Feinberg A.P., Gerald W.L., Vargas S.O., Chin L., Iafrate A.J., Bell D.W., Haber D.A. An X chromosome gene, WTX, is commonly inactivated in Wilms tumor. Science. 315, 642-645 (2007).

24. Perotti D., Gamba B., Sardella M., Spreafico F., Terenziani M., Collini P., Pession A., Nantron M., Fossati-Bellani F., Radice P. Functional inactivation of the WTX gene is not a frequent event in Wilms' tumors. Oncogene. 27, 4625-4632 (2008).
25. Wegert J., Wittmann S., Leuschner I., Geissinger E., Graf N., Gessler M. WTX inactivation is a frequent, but late event in Wilms tumors without apparent clinical impact. Genes Chromosomes Cancer. 48, 1102-1111 (2009).
26. Major M.B., Camp N.D., Berndt J.D., Yi X., Goldenberg S.J., Hubbert C., Biechele T.L., Gingras A.C., Zheng N., Maccoss M.J., Angers S., Moon R.T. Wilms tumor suppressor WTX negatively regulates WNT/beta-catenin signaling. Science. 316, 1043-1046 (2007).
27. Jenkins Z.A., van Kogelenberg M., Morgan T., Jeffs A., Fukuzawa R., Pearl E., Thaller C., Hing A.V., Porteous M.E., Garcia-Miñaur S., Bohring A., Lacombe D., Stewart F., Fiskerstrand T., Bindoff L., Berland S., Adès L.C., Tchan M., David A., Wilson L.C., Hennekam R.C., Donnai D., Mansour S., Cormier-Daire V., Robertson S.P. Germline mutations in WTX cause a sclerosing skeletal dysplasia but do not predispose to tumorigenesis. Nat Genet. 41, 95-100 (2009).
28. Malkin D., Sexsmith E., Yeger H., Williams B.R., Coppes M.J. Mutations of the p53 tumor suppressor gene occur infrequently in Wilms' tumor. Cancer Res. 54, 2077–2079 (1994).
29. Sredni S.T., Camargo B., Lopes L.F., Teixeira R., Simpson A.. Immunohis-tochemical detection of p53 protein expression as a prognostic indicator in Wilms tumor. Med Pediatr Oncol. 37, 455-458 (2001).
30. Lahoti C., Thorner P., Malkin D., Yeger H. Immunohistochemical detection of p53 in Wilms' tumors correlates with unfavorable outcome. Am J Pathol. 148, 1577–

1589 (1996).

31. Bardeesy N, Falkoff D, Petruzzi MJ, Nowak N, Zabel B, Adam M, Aguiar MC, Grundy P, Shows T, Pelletier J. Anaplastic Wilms' tumor, a subtype displaying poor prognosis, harbours p53 gene mutations. Nat Genet. 7, 91–97 (1994).
32. Wegert J., Ishaque N., Vardapour R., Geörg C., Gu Z., Bieg M., Ziegler B., Bausenwein S., Nourkami N., Ludwig N., Keller A., Grimm C., Kneitz S., Williams R.D., Chagtai T., Pritchard-Jones K., van Sluis P., Volckmann R., Koster J., Versteeg R., Acha T., O'Sullivan M.J., Bode P.K., Niggli F., Tytgat G.A., van Tinteren H., van den Heuvel-Eibrink M.M., Meese E., Vokuhl C., Leuschner I., Graf N., Eils R., Pfister S.M., Kool M., Gessler M. 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 (2015).
33. Williams R.D., Al-Saadi R., Chagtai T., Popov S., Messahel B., Sebire N., Gessler M., Wegert J., Graf N., Leuschner I., Hubank M., Jones C., Vujanic G., Pritchard-Jones K.; Children's Cancer and Leukaemia Group; SIOP Wilms' Tumour Biology Group. Subtype-specific FBXW7 mutation and MYCN copy number gain in Wilms' tumor. Clin Cancer Res. 16, 2036-2045 (2010).
34. Williams R.D., Al-Saadi R., Chagtai T., Popov S., Messahel B., Sebire N., Gessler M., Wegert J., Graf N., Leuschner I., Hubank M., Jones C., Vujanic G., Pritchard-Jones K. Subtype-specific FBXW7 mutation and MYCN copy number gain in Wilms tumor. Clinical Cancer Research. 16, 2036-2045 (2010).
35. McQuaid S., O'Meara A. N-myc oncogene amplification in paediatric tumours. Ir J Med Sci. 159, 172-174 (1990).
36. Schaub R., Burger A., Bausch D., Niggli F.K., Schafer B.W., Betts D.R. Array

comparative genomic hybridization reveals unbalanced gain of the MYCN region in Wilms tumors. Cancer Genet Cytogenet. 172, 61-65 (2007).

37. Williams R.D., Al-Saadi R., Natrajan R., Mackay A., Chagtai T., Little S., Hing S.N., Fenwick K., Ashworth A., Grundy P., Anderson J.R., Dome J.S., Perlman E.J., Jones C., Pritchard-Jones K. Molecular profiling reveals frequent gain of MYCN and anaplasia-specific loss of 4q and 14q in Wilms tumor. Genes Chromosomes and Cancer. 50, 982-995 (2011).
38. Bates C.M., Kharzai S., Erwin T., Rossant K., Parada L.F. Role of N-myc in the developing mouse kidney. Dev Biol. 222, 317-325 (2000).
39. Scott RH, Murray A, Baskcomb L, Turnbull C, Loveday C, Al-Saadi R, Williams R, Breatnach F, Gerrard M, Hale J, Kohler J, Lapunzina P, Levitt GA, Picton S, Pizer B, Ronghe MD, Traunecker H, Williams D, Kelsey A, Vujanic GM, Sebire NJ, Grundy P, Stiller CA, Pritchard-Jones K, Douglas J, Rahman N. Stratification of Wilms tumor by genetic and epigenetic analysis. Oncotarget. 3, 327-335 (2012).
40. Chao L.Y., Huff V., Tomlinson G., Riccardi V.M., Strong L.C., Saunders G.F. Genetic mosaicism in normal tissues of Wilms' tumor patients. Nat Genet. 3, 127-131 (1993).
41. Astuti D., Morris M.R., Cooper W.N., Staals R.H., Wake N.C., Fews G.A., Gill H., Gentle D., Shuib S., Ricketts C.J., Cole T., van Essen A.J., van Lingen R.A., Neri G., Opitz J.M., Rump P., Stolte-Dijkstra I., Muller F., Pruijn G.J., Latif F., Maher E.R. Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility. Nat Genet. 44, 277-284 (2012).
42. Ho J., Kreidberg J.A. MicroRNAs in renal development. Pediatr Nephrol. 28, 219-225 (2013).

43. Brown WT, Puranik SR, Altman DH, Hardin HC Jr. Wilms' tumor in three successive generations. Surgery. 72, 756-761 (1972).
44. Matsunaga E. Genetics of Wilms' tumor. Hum Genet. 57, 231-246 (1981).
45. Rahman N, Arbour L, Tonin P, Renshaw J, Pelletier J, Baruchel S, Pritchard-Jones K, Stratton MR, Narod SA. Evidence for a familial Wilms' tumour gene (FWT1) on chromosome 17q21-q25. Nat Genet. 13, 461-463 (1996).
46. McDonald JM, Douglass EC, Fisher R, Geiser CF, Krill CE, Strong LC, Virshup D, Huff V. Linkage of familial Wilms' tumor predisposition to chromosome 19 and a two-locus model for the etiology of familial tumors. Cancer Res. 58, 1387-1390 (1998).
47. Rapley EA, Barfoot R, Bonaïti-Pellié C, Chompret A, Foulkes W, Perusinghe N, Reeve A, Royer-Pokora B, Schumacher V, Shelling A, Skeen J, de Tourreil S, Weirich A, Pritchard-Jones K, Stratton MR, Rahman N. Evidence for susceptibility genes to familial Wilms tumour in addition to WT1, FWT1 and FWT2. Br J Cancer. 83, 177-183 (2000).
48. Melchionda F., Spreafico F., Ciceri S., Lima M., Collini P., Pession A., Massimino M., Radice P., Perotti D. A novel WT1 mutation in familial wilms tumor. Pediatr Blood Cancer. 60, 1388-1389 (2013).
49. Zirn B., Wittmann S., Gessler M. Novel familial WT1 read-through mutation associated with Wilms tumor and slow progressive nephropathy. Am J Kidney Dis. 45, 1100-1104 (2005).
50. Pritchard-Jones K., Rahman N., Gerrard M., Variend D., King-Underwood L. Familial Wilms tumour resulting from WT1 mutation: intronic polymorphism causing artefactual consititutional homozygosity. J Med Genet. 37, 377-379

(2000).

51. Diller L., Ghahremani M., Morgan J., Grundy P., Reeves C., Breslow N., Green D., Neuberg D., Pelletier J., Li F.P. Constitutional WT1 mutations in Wilms' tumor patients. J Clin Oncol. 16, 3634-3640 (1998).
52. Hanks S, Perdeaux ER, Seal S, Ruark E, Mahamdallie SS, Murray A, Ramsay E, Del Vecchio Duarte S, Zachariou A, de Souza B, Warren-Perry M, Elliott A, Davidson A, Price H, Stiller C, Pritchard-Jones K, Rahman N. Germline mutations in the PAF1 complex gene CTR9 predispose to Wilms tumour. Nat Commun. 7, 4398 (2014).
53. Leclercq T.M., Pitson S.M. Cellular signalling by sphingosine kinase and sphingosine 1-phosphate. IUBMB Life. 58, 467-472 (2006).
54. Pyne N.J., Pyne S. Sphingosine 1-phosphate and cancer. Nat Rev Cancer. 10, 489–503 (2010).
55. Plano D., Amin S., Sharma A.K. Importance of sphingosine kinase as a target in developing cancer therapeutics and recent developments in the synthesis of novel SPHK inhibitors. J Med Chem. 57, 5509-5524 (2014).
56. Liu H., Sugiura M., Nava V.E., Edsall L.C., Kono K., Poulton S., Milstien S., Kohama T., Spiegel S. Molecular cloning and functional characterization of a novel mammalian sphingosine kinase type 2 isoform. J Biol Chem. 275, 19513-19520 (2000).
57. Melendez A.J., Carlos-Dias E., Gosink M., Allen J.M., Takacs L. Human sphingosine kinase: molecular cloning, functional characterization and tissue distribution. Gene. 251, 19-26 (2000).
58. Maceyka M., Sankala H., Hait N.C., Le Stunff H., Liu H., Toman R., Collier C.,

Zhang M., Satin L.S., Merrill A.H. Jr, Milstien S., Spiegel S. SphK1 and SphK2, sphingosine kinase isoenzymes with opposing functions in sphingolipid metabolism. J Biol Chem. 280, 37118-37129 (2005).

59. Igarashi N., Okada T., Hayashi S., Fujita T., Jahangeer S., Nakamura S. Sphingosine kinase 2 is a nuclear protein and inhibits DNA synthesis. J Biol Chem. 278, 46832-46839 (2003).
60. Ding G., Sonoda H., Yu H., Kajimoto T., Goparaju S.K., Jahangeer S., Okada T., Nakamura S. Protein kinase D-mediated phosphorylation and nuclear export of sphingosine kinase 2. J Biol Chem. 282, 27493-27502 (2007).
61. Okada T., Ding G., Sonoda H., Kajimoto T., Haga Y., Khosrowbeygi A., Gao S., Miwa N., Jahangeer S., Nakamura S. Involvement of N-terminal-extended form of sphingosine kinase 2 in serum-dependent regulation of cell proliferation and apoptosis. J Biol Chem. 280, 36318-36325 (2005).
62. Hait N.C., Sarkar S., Le Stunff H., Mikami A., Maceyka M., Milstien S., Spiegel S. Role of sphingosine kinase 2 in cell migration toward epidermal growth factor. J Biol Chem. 280, 29462-29469 (2005).
63. Hait N.C., Bellamy A., Milstien S., Kordula T., Spiegel S. Sphingosine kinase type 2 activation by ERK-mediated phosphorylation. J Biol Chem. 282, 12058-12065 (2007).
64. Mizugishi K., Yamashita T., Olivera A., Miller G.F., Spiegel S., Proia R.L. Essential role of sphingosine kinases in neural and vascular development. Mol Cell Biol. 25, 11113-11121 (2005).

65. Hait NC, Allegood J, Maceyka M, Strub GM, Harikumar KB, Singh SK, Luo C, Marmorstein R, Kordula T, Milstien S, Spiegel S. Regulation of histone acetylation in the nucleus by sphingosine 1-phosphate. Science. 325, 1254-1257 (2009).
66. Wallington-Beddoe CT, Powell JA, Tong D, Pitson SM, Bradstock KF, Bendall LJ. Sphingosine kinase 2 promotes acute lymphoblastic leukemia by enhancing MYC expression. Cancer Res. 74, 2803-2815 (2014).
67. Kirby R.J., Jin Y., Fu J., Cubillos J., Swertfeger D., Arend L.J. Dynamic regulation of sphingosine-1-phosphate homeostasis during development of mouse metanephric kidney. Am J Physiol Renal Physiol. 296, 634-641 (2009).
68. Rahman N., Arbour L., Tonin P., Baruchel S., Pritchard-Jones K., Narod S.A., Stratton M.R. The familial Wilms' tumour susceptibility gene, FWT1, may not be a tumour suppressor gene. Oncogene. 14, 3099-3102 (1997).
69. Ha M., Kim V.N. Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol. 15, 509-524 (2014).
70. Chu J.Y., Sims-Lucas S., Bushnell D.S., Bodnar A.J., Kreidberg J.A., Ho J. Dicer function is required in the metanephric mesenchyme for early kidney development. Am J Physiol Renal Physiol. 306, 764-772 (2014).
71. Carleton M., Cleary M.A., Linsley P.S. MicroRNAs and cell cycle regulation. Cell Cycle. 6, 2127–2132 (2007).
72. Harfe B.D. MicroRNAs in vertebrate development. Curr Opin Genet Dev. 15, 410–415 (2005).
73. Lau N.C., Lim L.P., Weinstein E.G., Bartel D.P. An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science. 294, 858–862 (2001).

74. Garzon R., Calin G.A., Croce C.M. MicroRNAs in cancer. Annu Rev Med. 60, 167-179 (2009).
75. Nagalakshmi V.K., Ren Q., Pugh M.M., Valerius M.T., McMahon A.P., Yu J. Dicer regulates the development of nephrogenic and ureteric compartments in the mammalian kidney. Kidney Int. 79, 317–330 (2011).
76. Sequeira-Lopez M.L., Weatherford E.T., Borges G.R., Monteagudo M.C., Pentz E.S., Harfe B.D., Carretero O., Sigmund C.D., Gomez R.A. The microRNA-processing enzyme dicer maintains juxtaglomerular cells. J Am Soc Nephrol. 21, 460–467 (2010).
77. Shi S., Yu L., Chiu C., Sun Y., Chen J., Khitrov G., Merkenschlager M., Holzman L.B., Zhang W., Mundel P., Bottinger E.P. Podocyte-selective deletion of dicer induces proteinuria and glomerulosclerosis. J Am Soc Nephrol 19, 2159–2169 (2008).
78. Harvey S.J., Jarad G., Cunningham J., Goldberg S., Schermer B., Harfe B.D., McManus M.T., Benzing T., Miner J.H. Podocyte-specific deletion of dicer alters cytoskeletal dynamics and causes glomerular disease. J Am Soc Nephrol. 19, 2150–2158 (2008).
79. Ho J., Ng K.H., Rosen S., Dostal A., Gregory R.I., Kreidberg J.A. Podocyte-specific loss of functional microRNAs leads to rapid glomerular and tubular injury. J Am Soc Nephrol. 19, 2069–2075 (2008).
80. Gregory R.I., Yan K.P., Amuthan G., Chendrimada T., Doratotaj B., Cooch N., Shiekhattar R.. The microprocessor complex mediates the genesis of microRNAs. Nature. 432, 235-240 (2004).
81. Han J., Lee Y., Yeom K.H., Kim Y.K., Jin H., Kim V.N. The Drosha-DGCR8

complex in primary miRNA processing. Genes Dev. 18, 3016-3027 (2004).

82. Landthaler M., Yalcin A., Tuschl T. The human DGCR8 and its D. melanogaster homolog are required for miRNA biogenesis. Curr Biol. 14, 2162-2167 (2004).
83. Lee Y., Ahn C., Han J., Choi H., Kim J., Yim J., Lee J., Provost P., Radmark O., Kim S., Kim V.N.. The nuclear Rnase III Drosha initiate miRNA processing. Nature. 425, 415-419 (2003).
84. Yi R, Qin Y, Macara IG, Cullen BR. Exportin-5 mediates the nuclear export of pre-miRNAs and short hairpin RNAs. Genes Dev. 17, 3011-3016 (2003).
85. Bohnsack MT, Czaplinski K, Gorlich D. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. RNA. 10, 185-191 (2004).
86. Lund E., Güttinger S., Calado A., Dahlberg J.E., Kutay U. Nuclear export of miRNA precursors. Science. 303, 95-98 (2004).
87. Fukunaga R., Han B.W., Hung J.H., Xu J., Weng Z., Zamore P.D.. Dicer partner proteins tune the length of mature miRNAs in flies and mammals. Cell. 151, 533-546 (2012).
88. Foulkes W.D., Priest J.R., Duchaine T.F. DICER1: mutations, microRNAs and mechanisms. Nat Rev Can. 14, 662-672 (2014).
89. Slade I., Bacchelli C., Davies H., Murray A., Abbaszadeh F., Hanks S., Barfoot R., Burke A., Chisholm J., Hewitt M., Jenkinson H., King D., Morland B., Pizer B., Prescott K., Saggar A., Side L., Traunecker H., Vaidya S., Ward P., Futreal P.A., Vujanic G., Nicholson A.G., Sebire N., Turnbull C., Priest J.R., Pritchard-Jones K., Houlston R., Stiller C., Stratton M.R., Douglas J., Rahman N. DICER1 syndrome: clarifying the diagnosis, clinical features and management implications

of a pleiotropic tumour predisposition syndrome. J Med Genet. 48, 273-278 (2011).

90. Foulkes W.D., Bahubeshi A., Hamel N., Pasini B., Asioli S., Baynam G., Choong C.S., Charles A., Frieder R.P., Dishop M.K., Graf N., Ekim M., Bouron-Dal Soglio D., Arseneau K., Young R.H., Sabbaghian N., Srivastava A., Tischkowitz M.D., Priest J.R. Extending the phenotypes associated with DICER1 mutations. Hum Mutat. 32, 1381-1384 (2011)
91. Wu M.K., Sabbaghian N., Xu B., Addidou-Kalucki S., Bernard C., Zou D., Reeve A.E., Eccles M.R., Cole C., Choong C.S., Charles A., Tan T.Y., Iglesias D.M., Goodyer P.R., Foulkes W.D. Biallelic DICER1 mutations occur in Wilms tumours. J Pathol. 230, 154-164 (2013).
92. Hill D.A., Ivanovich J., Priest J.R., Gurnett C.A., Dehner L.P., Desruisseau D., Jarzembowski J.A., Wikenheiser-Brokamp K.A., Suarez B.K., Whelan A.J., Williams G., Bracamontes D., Messinger Y., Goodfellow P.J. DICER1 mutations in familial pleuropulmonary blastoma. Science. 325, 965 (2009).
93. Natrajan R., Little S.E., Sodha N., Reis-Filho J.S., Mackay A., Fenwick K., Ashworth A., Perlman E.J., Dome J.S., Grundy P.E., Pritchard-Jones K., Jones C. Analysis by array CGH of genomic changes associated with the progression or relapse of Wilms' tumour. J Pathol. 211, 52–59 (2007).
94. Erdösová B., Wagner F., Kylarová D. The detection of myc proteins in the developing human kidney. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub. 148, 205-207 (2004).
95. Fievet A., Belaud-Rotureau M.A., Dugay F., Abadie C., Henry C., Taque S., Andrieux J., Guyetant S., Robert M., Dubourg C., Edan C., Rioux-Leclercq N.,

Odent S., Jaillard S. Involvement of germline DDX1-MYCN duplication in inherited nephroblastoma. Eur J Med Genet. 56, 643-647 (2013).

96. Noguera R., Villamón E., Berbegall A., Machado I., Giner F., Tadeo I., Navarro S., Llombart-Bosch A. Gain of MYCN region in a Wilms tumor-derived xenotransplanted cell line. Diagn Mol Pathol. 19, 33-39 (2010).
97. Turnbull C., Perdeaux E.R., Pernet D., Naranjo A., Renwick A., Seal S., Munoz-Xicola R.M., Hanks S., Slade I., Zachariou A., Warren-Perry M., Ruark E., Gerrard M., Hale J., Hewitt M., Kohler J., Lane S., Levitt G., Madi M., Morland B., Neefjes V., Nicholson J., Picton S., Pizer B., Ronghe M., Stevens M., Traunecker H., Stiller C.A., Pritchard-Jones K., Dome J., Grundy P., Rahman N. A genome-wide association study identified susceptibility loci for Wilms tumor. Nat Genet. 44, 681-684 (2012).
98. Han C., Liu Y., Wan G., Choi H.J., Zhao L., Ivan C., He X., Sood A.K., Zhang X., Lu X. The RNA-binding protein DDX1 promotes primary microRNA maturation and inhibits ovarian tumor progression. Cell Rep. 8, 1447-1460 (2014).
99. de Kock L, Sabbaghian N, Druker H, Weber E, Hamel N, Miller S, Choong CS, Gottardo NG, Kees UR, Rednam SP, van Hest LP, Jongmans MC, Jhangiani S, Lupski JR, Zacharin M, Bouron-Dal Soglio D, Huang A, Priest JR, Perry A, Mueller S, Albrecht S, Malkin D, Grundy RG, Foulkes WD. Germline and somatic DICER1 mutations in pineoblastoma. Acta Neuropathol. 4, 583-595 (2014).
100. de Kock L, Sabbaghian N, Plourde F, Srivastava A, Weber E, Bouron-Dal Soglio D, Hamel N, Choi JH, Park SH, Deal CL, Kelsey MM, Dishop MK, Esbenshade A, Kuttesch JF, Jacques TS, Perry A, Leichter H, Maeder P, Brundler MA, Warner J, Neal J, Zacharin M, Korbonits M, Cole T, Traunecker H,

McLean TW, Rotondo F, Lepage P, Albrecht S, Horvath E, Kovacs K, Priest JR, Foulkes WD. Pituitary blastoma: a pathognomonic feature of germline DICER1 mutations. Acta Neuropathol. 1, 111-122 (2014).

101. Xu J., Wong E.Y.M., Cheng C., Li J., Sharkar M.T.K., Xu C.Y., Chen B., Sun J., Jung D., Xu P.X. Eya1 interacts with Six2 and Myc to regulate expansion of the nephron progenitor pool during nephrogenesis. Dev Cell. 31, 434-447 (2014).

102. Chalfant C.E., Szulc Z., Roddy P., Bielawska A., Hannun Y.A. (2004) The structural requirements for ceramide activation of serine-threonine protein phosphatases. J Lipid Res 45, 496-506 (2004).

103. Perry D.M., Kitatani K., Roddy P., El-Osta M., Hannun Y.A. Identification and characterization of protein phosphatase 2C activation by ceramide. J Lipid Res 53, 1513–1521 (2012).

104. Woodcock J.M., Ma Y., Coolen C., Pham D., Jones C., Lopez A.F., Pitson S.M. Sphingosine and FTY720 directly bind pro-survival 14–3-3 proteins to regulate their function. Cell Signal 22, 1291–1299 (2010).

105. Pyne N.J., Pyne S. Sphingosine 1-phosphate and cancer. Nat Rev Cancer 10, 489–503 (2010).

106. Meyer zu Heringdorf D., Jakobs K.H. Lysophospholipid receptors: signalling, pharmacology and regulation by lysophospholipid metabolism. Biochim Biophys Acta 1768, 923 – 940 (2007).

107. Rosen H., Gonzalez-Cabrera P.J., Sanna M.G., Brown S. Sphingosine 1-phosphate receptor signaling. Annu Rev Biochem 78, 743–768 (2009).

VITA

Blake Palculict was born in El Dorado, AR in 1984. He attended West Brook High School in Beaumont, TX, and went on to matriculate at Texas A&M University where he studied Biology and was active in research at the Texas A&M College of Veterinary Medicine. He then joined the University of Arkansas for Medical Sciences and earned his Master of Science degree in physiology. After working at the Human Genome Sequencing Center (Baylor College of Medicine) and the Experimental Therapeutics Department at MD Anderson Cancer Center, he joined Dr. Vicki Huff's lab at MD Anderson, where he completed his doctoral research training focusing on familial Wilms tumor.