HHS Public Access

Author manuscript

Cancer Cell. Author manuscript; available in PMC 2016 March 21.

Published in final edited form as:

Cancer Cell. 2015 February 9; 27(2): 286–297. doi:10.1016/j.ccell.2015.01.003.

Recurrent *DGCR8*, *DROSHA*, and *SIX* Homeodomain Mutations in Favorable Histology Wilms Tumors

Amy L. Walz[1], **Ariadne Ooms**[2], **Samantha Gadd**[3], **Daniela S. Gerhard**[4], **Malcolm A. Smith**[5], **Jamie M. Guidry Auvil**[4], **Daoud Meerzaman**[6], **Qing-Rong Chen**[6], **Chih Hao Hsu**[6], **Chunhua Yan**[6], **Cu Nguyen**[6], **Ying Hu**[6], **Reanne Bowlby**[7], **Denise Brooks**[7], **Yussanne Ma**[7], **Andrew J. Mungall**[7], **Richard A. Moore**[7], **Jacqueline Schein**[7], **Marco A. Marra**[7,8], **Vicki Huff**[9], **Jeffrey S. Dome**[10], **Yueh-Yun Chi**[11], **Charles G. Mullighan**[12], **Jing Ma**[12], **David A. Wheeler**[13], **Oliver A. Hampton**[13], **Nadereh Jafari**[14], **Nicole Ross**[15], **Julie M. Gastier-Foster**[15], and **Elizabeth J. Perlman**[3]

[1]Division of Hematology-Oncology and Transplantation, Ann and Robert H. Lurie Children's Hospital of Chicago, Northwestern University's Feinberg School of Medicine, Chicago, IL, 60611, USA [2]Department of Pathology, Josephine Nefkens Institute, Erasmus Medical Center, Rotterdam, CA, 3000, The Netherlands [3]Department of Pathology and Laboratory Medicine, Lurie Children's Hospital, Northwestern University's Feinberg School of Medicine and Robert H. Lurie Cancer Center, Chicago, IL, 60611, USA [4]Office of Cancer Genomics, National Cancer Institute, Bethesda, MD, 20892, USA [5]Cancer Therapy Evaluation Program, National Cancer Institute, Bethesda, MD, 20892, USA [6]Center for Biomedical Informatics and Information Technology, National Cancer Institute, Bethesda, MD, 20892, USA [7]Canada's Michael Smith Genome Sciences Centre, British Columbia Cancer Agency (BCCA), Vancouver, BC, V5Z 4S6, Canada [8]Department of Medical Genetics, University of British Columbia, Vancouver, BC, V6H 3N1, Canada [9]Department of Genetics, The University of Texas MD Anderson Cancer Center, Houston, TX, 77030, USA [10]Division of Pediatric Hematology/Oncology, Children's National Medical Center, Washington, DC, 20010, USA [11]Department of Biostatistics, University of Florida, Gainesville, FL, 32610, USA [12]Department of Pathology, St. Jude Children's Research Hospital, Memphis, TN, 38105, USA [13]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas 77030, USA [14]Center for Genetic Medicine, Northwestern University, Chicago, IL 60611, USA [15]Department of Pathology and Laboratory Medicine, Nationwide Children's Hospital, Ohio State University College of Medicine, Columbus, OH, 43205, USA

Correspondence: E.J. Perlman, 225 East Chicago Avenue, Box 17, Chicago, Illinois 60611, eperlman@luriechildrens.org.

AUTHOR CONTRIBUTIONS: ALW and AO performed the miRNAPG and *SIX1/2* mutation analyses, respectively; both contributed equally to the writing of the manuscript. DSG, JMGA, MAS oversaw the administrative and data management of the TARGET project; DM, CHH, CY, CN, Q-RC, YH provided the bioinformatic analysis for WXS, copy number and methylation-expression correlation; Y-YC provided the statistical analysis; JMGF, NR performed the specimen processing and quality control; JD, VH provided the samples and the clinical data; CGM, JM performed the copy number analysis; SG performed the gene expression analysis and bioinformatics for gene expression, miRNA expression, and protein analysis; YM, AJM, RAM, JS, DB, RB, MAM performed the mRNA-seq, miRNA-seq, and target capture sequencing; DAW, OAH performed the WXS; NJ performed the methylation analysis; EJP designed and oversaw all aspects of the study; all authors contributed to the manuscript.

SUMMARY

We report the most common single nucleotide substitution/deletion mutations in Favorable Histology Wilms Tumors (FHWT) to occur within *SIX1/2* (7% of 534 tumors) and microRNA processing genes (miRNAPG) *DGCR8* and *DROSHA* (15% of 534 tumors). Comprehensive analysis of 77 FHWTs indicates that tumors with *SIX1/2* and/or miRNAPG mutations show a pre-induction metanephric mesenchyme gene expression pattern and are significantly associated with both perilobar nephrogenic rests and 11p15 imprinting aberrations. Significantly decreased expression of mature *Let-7a* and the miR-200 family (responsible for mesenchymal-to-epithelial transition) in miRNAPG-mutant tumors is associated with an undifferentiated blastemal histology. The combination of *SIX* and miRNAPG mutations in the same tumor is associated with evidence of RAS activation and a higher rate of relapse and death.

INTRODUCTION

Wilms tumor (WT) represents the most common pediatric renal malignancy, with an estimated annual incidence of 500 cases in the United States [Howlander et al., 2013]. WTs commonly display epithelial, stromal, and undifferentiated (blastemal) components in varying proportions and often closely resemble the different stages of renal development [Rivera and Haber, 2005]. They often arise within precursor lesions known as perilobar and intralobar nephrogenic rests (PLNR, ILNR) [Beckwith et al., 1990]. Evidence suggests that WT development depends not only on the nature of specific genetic events, but also on the timing of their occurrence within early renal development [Gadd et al., 2012]. The developmental window begins with the early intermediate mesoderm, which contains progenitor cells of both the urinary collecting system and pre-induction metanephric mesenchyme. The metanephric mesenchyme undergoes induction, including mesenchymal-to-epithelial transition (MET), resulting in nephron development [Kobayashi et al., 2008].

Mutations in *WT1*, *WTX*, and *CTNNB1* contribute to WT development; in addition, loss of imprinting (LOI) or loss of heterozygosity (LOH) at 11p15 (resulting in biallelic expression of *IGF2)* is present in the majority of WTs [Gadd et al., 2012]. However, the identification of 11p15 LOH in normal tissue from some WT patients [Chao et al., 1993], and the absence of tumor development in mutant mice with LOI of 11p15 [Hu et al., 2011] suggest that biallelic expression of *IGF2* alone is insufficient for tumor development. Five subsets of WT identified based on gene expression patterns differ in their histology, nephrogenic rest status, clinical outcome, and show evidence of arrest at different stages of renal development [Gadd et al., 2012]. Two subsets express high levels of *WT1*: Subset 1 (~5% of FHWT) is comprised exclusively of epithelial tumors lacking nephrogenic rests in infants that do not relapse and show a post-induction metanephric mesenchyme gene expression pattern; Subset 5 (S5; ~70% of FHWT) is exemplified by tumors with 11p15 LOI or LOH that arise within PLNRs and have the gene expression pattern of pre-induction metanephric mesenchyme. The remaining three subsets (S2-4) are defined by a low *WT1* expression pattern, are often accompanied by *WT1* mutations and/or deletions, and arise within ILNRs. Subset 2 (~15% of FHWT) arises in young infants, shows high expression of muscle-related genes, has an excellent prognosis, and has a gene expression pattern of the intermediate mesoderm. Subset 3 (S3; ~10% of FHWT) arises in older children, has a higher relapse rate, and a pre-

induction metanephric mesenchyme gene expression pattern without high expression of muscle genes. Subset 4 (S4; ~5% of FHWT) has a gene expression profile similar to that of S2 tumors (that of the intermediate mesoderm), but occurs in older children, and has the highest relapse rate.

In the US, WTs are treated with primary resection (if possible), followed by stage-specific adjuvant chemotherapy, whereas in Europe, neoadjuvant chemotherapy followed by resection is the preferred treatment [Dome et. al, 2013]. Over 95% of WTs are classified as Favorable Histology because they lack evidence of anaplasia (presence of nuclear hyperchromasia and enlargement with atypical mitoses, often accompanied by *TP53* mutations) [Beckwith and Palmer, 1978; Bardeesy et al., 1994]. Patients with FHWTs, the subject of this study, overall have an excellent survival (~90%); however, over 15% relapse and approximately 40% of these patients eventually die from their disease [Dome et al., 2002]. The National Cancer Institute's (NCI) "Therapeutically Applicable Research to Generate Effective Treatments" (TARGET) initiative seeks to identify driver mutations and therapeutic targets for high-risk pediatric tumors through comprehensive integrative genomics (http://ocg.cancer.gov/program/target). We report the mutations commonly identified in FHWT and place these in their clinical, pathologic, and developmental context.

RESULTS

A discovery set of 77 pre-therapy FHWTs that subsequently relapsed was analyzed by whole genome (WGS, n=58) or whole exome (WXS, n=19) sequencing. Bioinformatic analysis identified 825 high-quality somatic, non-synonymous variants, with an average of 11 candidate mutations/case (range 2–42, Figure S1). Consistent with previous reports [Ruteshouser et al., 2008], somatic SNVs or small deletions were identified in *WT1* (3 patients, 4%), in *WTX* (5 patients, 6.5%), and in *CTNNB1* (5 patients, 6.5%). Unexpectedly, 12 somatic variants were identified in miRNAPG in 11 patients (14%). Lastly, eight tumors (10%) had variants in either the *SIX1* or *SIX2* homeodomain; strikingly, 5/8 *SIX1/2*-mutant tumors also had mutations in miRNAPG. None of the miRNAPG or *SIX1/2* variants were annotated as polymorphisms in dbSNP versions 134 and 135 [Sherry et al., 2001] or in the 1000 Genomes Pilot Projects 1, 2, and 3 [Abecasis et al., 2012], and none had been previously identified in COSMIC Version 69 [Forbes et al., 2010]. All variants were verified and expressed by mRNA-sequencing (mRNAseq). All were predicted to be deleterious by PolyPhen Version 2 [Adzhubei et al., 2010].

Recurrent *SIX1/2* homeodomain hotspot mutations

In the discovery set, 4 *SIX1* and 4 *SIX2* mutations involved the same location (p.Q177R) in the SIX homeodomain responsible for DNA binding and protein interaction (Table S1) [Christenson et al., 2008]. Q177 resides in a region conserved in 95% of 100 homologous proteins by a UniProt sequence similarity search (http://www.ncbi.nlm.nih/gov/pubmed/24253303); by Protein Homology/Analogy Recognition Engine Version 2 (Phyre 2) [Kelly et al., 2009], this glutamine residue was predicted to specifically interact with DNA (Figure 1A). All mutations except one (a tumor with copy neutral LOH for chromosome 2) were heterozygous, with both alleles highly expressed by mRNAseq (Figure 1B). In the validation

set of 534 FHWT, the same *SIX1* and *SIX2* p.Q177R missense mutations were identified in 23 and 13 patients, respectively, for an overall frequency of 6.7%. An additional *SIX2* variant involved p.Y129N (Table S2, Figure 1C).

Gene expression characteristics of *SIX1/2* Q177R mutant tumors

Global gene expression analysis was performed with 75 discovery samples that passed quality control. Unsupervised analysis with Non-negative Matrix Factorization (NMF) [Brunet et al., 2004], resulted in k=6 clusters having the highest cophenetic correlation (0.95) after k=2. NMF Cluster 2 contained all 7 evaluable *SIX1/2* mutant tumors (Figure 1D). Gene Set Enrichment Analysis (GSEA) [Subramanian et al., 2005] comparing the 7 evaluable *SIX1/2*-mutant tumors with all *SIX1/2* wild-type tumors identified no significantly enriched canonical pathways, GO biologic processes, or oncogenic signatures. Hierarchical analysis of the 100 top ranked differentially expressed genes in the *SIX1/2*-mutant tumors by GSEA (Table S3) reveals a similar distinctive expression pattern for both *SIX1*- and *SIX2*-mutant tumors (Figure 1E), suggesting that they share a common function. Tumors showing copy number change at the *SIX1/2* loci did not cluster with the mutant tumors, indicating that the mutations may have a neomorph function. The NMF clusters were then compared with the Subsets previously outlined in the introduction. All tumors in NMF Cluster 2 were members of Subset 5, previously characterized by their similarity to the pre-induction metanephric mesenchyme, including high expression of *SIX1*, *PAX2*, *EYA1*, *SALL1*, *MEOX1*, *MEIS2*, *WASF*, and *CCND2* [Gadd et al., 2012] (Figure 1F). To identify genes characterizing *SIX* mutations within NMF Cluster 2, the *SIX1/2*-mutant tumors were compared to the *SIX1/2* wild-type tumors in this cluster, identifying 100 top ranked genes (Table S3), two of which (*CCND2*, p= 0.0001 and *MEIS2*, p=0.001) are illustrated in Figure 1G.

Recurrent miRNAPG hotspot mutations

We identified recurrent somatic mutations in *DROSHA* (recently reported [Rakheja et al., 2014; Torrezan et al., 2014]) and in *DGCR8*. Lastly, a somatic variant in *XPO5* was identified (Table S1). The frequency of each mutation was established within the validation cohort (Table S2, Figure 2A).

DROSHA—In the discovery set, 8 somatic *DROSHA* mutations were identified in 7 patients. Six were missense mutations involving exon 29 in the RNase IIIB domain responsible for cleaving the 5′ end of primary miRNAs to form precursor miRNAs [Winter et al., 2009], including p.E1147K (4), p.D1151A (1), and D1151G (1). These residues reside within a 95% conserved region of the protein by UniProt. The two remaining variants were nonsense mutations in one patient (PAKZHF) resulting in loss of both RNase III domains. Therefore, all mutations affected the RNase IIIB domain of DROSHA. All *DROSHA* mutations were heterozygous, and mRNAseq confirmed equivalent levels of the mutant and wild-type genes with the exception of PAKZHF, which had discordant transcript ratios of 85% (p.Q46*) and 14% (p.R414*) (Figure 2B). These findings support the reported evidence of a dominant-negative mechanism for p.E1147 mutations [Rakheja et al., 2014; Wegert et al., 2015]. Within the validation set (534 tumors), 59 *DROSHA* variants were identified in 58 patients for a frequency of 11%; 42/59 variants were either p.E1147 (38) or

p.D1151 (4). Two nonsense variants occurred in one patient. The remaining 15 were missense with recurrent mutations in the RNase IIIA domain (p.E969, 4 tumors) and the RNase IIIB domain (p.E1222, 3 tumors).

DGCR8—In the discovery set, 3 p.E518K mutations were identified in the double-stranded RNA binding domain. E518 is in a region conserved in 95% of the 100 top sequences by UniProt. The second *DGCR8* allele was deleted in all mutant tumors, and mRNAseq confirmed an allelic fraction of >90% (Figure 2B). These observations are likewise reported by Wegert et al. [2015]. In the validation set of 534 tumors, 20 *DGCR8* variants (17 p.E518K mutations, a missense variant in exon 12, and two nonsense variants in exon 2) resulted in an overall frequency of 4% in FHWT.

XPO5—One heterozygous somatic nonsense *XPO5* mutation was identified in the discovery set, resulting in loss of the C-terminus required for binding pre-miRNAs [Okada et al., 2009]. Within the validation set, 10/534 non-recurring, damaging *XPO5* variants spanning the length of the transcript were identified in 7 patients, for a frequency of 1%.

miRNAPG mutant tumors have reduced expression of critical miRNAs

In the NMF analysis previously described (Figure 1D), cluster 2 contained 8/11 tumors with somatic miRNAPG mutations as well as all *SIX1/2*-mutant tumors. Clusters 4 and 5 contained the remaining somatic miRNAPG-mutant tumors, one each in *DROSHA* p.D1151G, *DGCR8* p.E518K, and *XPO5*. GSEA analysis comparing miRNAPG-mutant tumors (n=11) with the remainder (n=64) revealed significant negative enrichment of three gene lists, two of which contain genes up-regulated in breast cancer cell cultures over-expressing either MYC-C or E2F3 (Table S4), suggesting that MYC and E2F3 are relatively inactivated in miRNAPG-mutant tumors compared with other WTs. Hierarchical analysis using the 100 top ranked genes differentially expressed in tumors with miRNAPG somatic mutations by GSEA (Table S3) shows clustering of all somatic miRNAPG variants (Figure 2C), supporting a similar underlying mechanism of action. Further, tumors showing copy number loss at any of the miRNAPG loci did not cluster with the mutant tumors unless they also contained miRNAPG mutations. Intriguingly, *DICER1* was expressed at significantly higher levels in tumors with somatic miRNAPG mutations ($p<0.001$, Table S3).

Mutations in miRNAPGs are expected to result in decreased mature miRNAs and increased primary miRNAs [Winter et al., 2009]; this has recently been documented in 3 mutant and 5 wild-type tumors [Rakheja, et al., 2014]. To confirm this, we analyzed mature and primary *Let-7a* miRNA expression within 77 discovery tumors. Given that haploinsufficient miRNAPGs may effect function [Lambertz et al., 2010], the samples were analyzed as three groups: those with somatic miRNAPG mutations (n=11), those without mutations but with copy number loss of miRNAPG loci (n=10), and those without either miRNAPG somatic mutations or copy number loss (n=56). The expression of mature *Let-7a* was significantly lower in both tumors with somatic miRNAPG mutations ($p= 0.004$) and those with miRNAPG copy number loss ($p= 0.047$), compared with those tumors lacking either (Table S5, Figure 2D). While the expression of the primary *Let-7a* transcript (PRI-*Let-7a*) was

higher in the miRNAPG-mutant group compared with those lacking either mutations or copy number loss, this did not achieve statistical significance.

The effect of miRNAPG mutations on the global miRNA landscape was evaluated by miRNAseq. Two-class Significance Analysis of Microarray Sequencing [Tusher et al., 2001] comparing the 11 miRNAPG-mutant tumors with the 56 tumors lacking both mutations and copy number loss identified 43 differentially expressed miRNAs (FDR<1% and BH-corrected p<0.05) (Table S6). Hierarchical analysis of all 77 tumors using these 43 miRNAs shows localization of 10/11 miRNAPG-mutant tumors within a single cluster (cluster 4 in Figure 2E). Tumors with only *SIX1/2* mutations (lacking miRNAPG mutations) and tumors with copy number loss of the miRNAPG loci without concomitant miRNAPG mutations did not cluster with the mutant tumors. Cluster 4 is characterized by decreased expression of the entire miR-200 family (miR-200a, -200b, -141, and -429) and miR-181b, all of which are involved in MET and stem cell maintenance [Hua et al., 2013; Ceppi et al., 2010; Park et al., 2008; Ceppi and Peter, 2014]. Decreased expression of the miR-200 family is predicted to reduce MYC and E2F expression [Hua et al., 2013], as was observed in our GSEA analysis of the miRNAPG-mutant tumors.

Correlating *SIX1/2* and miRNAPG Mutations with Clinicopathologic Features

Analysis of the validation set revealed a significant female predominance in tumors with *DGCR8* E518K and *DROSHA* exon 29 (miRNAPG-HS) mutations and a greater prevalence of tumors with blastemal predominant histology in patients with miRNAPG-HS and/or *SIX1/2* Q177R mutations (Table 1). There was also a significantly higher association with PLNRs and a lower association with ILNRs in those tumors with *SIX1/2* Q177R and miRNAPG-HS mutations (Table 1). Since PLNRs are associated with loss of the normal imprinting pattern at 11p15 [Ravenel et al., 2001], 11p15 methylation was analyzed within the 77 discovery set tumors. LOH, LOI, and retention of imprinting (ROI) were identified in 29/77 (38%), 30/77 (39%), and 18/77 (23%), respectively in the entire group. LOI was significantly more frequent in both those tumors containing miRNAPG-HS mutations (7/9 patients, 78%, p= 0.011), and in those with *SIX* mutations (7/8 patients, 87.5%, p= 0.003).

Integration of the above clinicopathologic features with mutation, copy number, 11p15 imprinting status, and membership in gene expression subsets and in miRNA expression categories is provided in Figure 3, arranged by NMF cluster. The NMF cluster (Figure 3, first row) correlates closely with the previously reported gene expression subsets (Figure 3, second row, assigned as shown in Figure S2) with the exception that S5 is represented most prominently within two NMF clusters, clusters 1 and 2. NMF cluster 2, which includes the majority of the miRNAPG mutations and all *SIX1/2* mutations, demonstrates a predominance of tumors in miRNA cluster 4 (defined largely by low miR-200 family expression), a high prevalence of both blastemal histology and PLNRs, and a high frequency of 11p15 LOI. The second large S5 predominant group, NMF Cluster 1, lacks miRNAPG mutations, does not show reduction of the miR-200 family, and is associated with ILNRs rather than PLNRs. This cluster shows a relatively high frequency of *DICER1* loss (through loss of chromosome 14). Loss of one *DICER1* allele results in partial impairment of miRNA processing [Gurtan et al., 2012] and promotes tumorigenesis [Kumar et al., 2007; Lambetz

et al., 2010]. These findings suggest that almost half of the large S5 group may be driven by miRNAPG and/or *SIX* mutations and the possibility remains that *DICER1* loss may contribute to pathogenesis in some of the remaining S5 tumors.

NMF clusters 3 and 4 contain the majority of the *WT1*, *WTX*, and *CTNNB1* mutations, show a high frequency of membership in S3 and S4 (Figure 3, second row), and are characterized by mixed histology and an association with ILNRs. The presence of two tumors in NMF cluster 4 with somatic miRNAPG mutations associated with membership in S4 (characterized by a gene expression pattern of the intermediate mesoderm and a high relapse rate) suggests that miRNAPG mutations may also be pathogenic when they occur earlier in renal development, within the intermediate mesoderm. The overlap of miRNAPG-HS and *SIX* mutations with *WT1*, *WTX*, or *CTNNB1* mutations was evaluated in the validation set. Of 36 patients (7%) with *WT1* variants, 1 *DGCR8* E518K and no *DROSHA* or *SIX* variants were present. *WTX* variants were identified in 31 patients (6%); in these 1 *DGCR8* E518K, 1 *DROSHA* exon 29, and 3 *SIX* Q177R variants were present. Of 62 patients with *CTNNB1* variants (12%), 3 *DROSHA* exon 29 variants, 1 *SIX* Q177R variant, and no *DGCR8* variants were also present; in contrast, 18 also had *WT1* variants (a recognized association [Maiti et al., 2000]). Of note, we were unable to evaluate exonic deletions of *WT1*and *WTX* in this data, which represent ~70% of the genetic aberrations that occur at these two loci [Ruteshouser et al., 2008; Gadd et al., 2012].

The combination of *SIX1/2* and miRNAPG mutations results in poor outcome and RAS activation

Within the discovery set containing FHWT that subsequently relapsed, 5/8 (63%) tumors with *SIX1/2* mutations also had somatic miRNAPG-HS mutations. In the validation set of 534 tumors, of 36 tumors with *SIX1/2* Q177R variants, 10 (28%) also had a *DGCR8* E518K (1) or *DROSHA* exon 29 (9) variants (p= 0.0015). There was no significant difference in either the rate of relapse (30%) or the number of deaths (14%) in the entire validation set compared with patients with *DGCR8* E518K or *DROSHA* exon 29 variants without *SIX* Q177R variants (31% and 14%, respectively), or in those with *SIX* Q177R variants without associated miRNAPG-HS variants (31% and 15%, respectively). However, the 10 patients whose tumors contained both miRNAPG-HS and *SIX* Q177R variants had a significantly higher relapse rate (8/10, 80%, p= 0.0001, Figure 4) and a higher rate of death (40%). Hence, while the miRNAPG-HS and *SIX* Q177R variants alone do not portend a worse outcome, the combination of these mutations, while rare, appears to result in a worse outcome.

To identify possible therapeutic targets for this group, we identified previously published Affymetrix U133A data on 291/534 validation tumors and deposited these in the TARGET Data Matrix [Gadd et al., 2012]. GSEA analysis of 22 tumors with miRNAPG-HS variants and without *SIX1/2* variants and the analysis of 12 tumors with *SIX1/2* Q177R variants and without miRNAPG-HS variants did not identify significant enrichment for canonical pathways, GO biologic processes, or oncogenic signatures, similar to our experience with the discovery set. In contrast, analysis of the 6 tumors with available gene expression data with miRNAPG-HS variants in combination with *SIX1/2* Q177R variants demonstrated a

large number of highly significantly (FDR<5%) enriched gene lists (Table S4). While the number of samples is small, of interest is the positive enrichment of 7 gene sets differentially expressed in a variety of tumor types following over-expression of an oncogenic KRAS mutation, 4 genes sets differentially expressed in a medulloblastoma cell line following knock-down of PCGF2 (a polycomb group protein that functions by transcription repression), and two gene sets up-regulated during embryoid body differentiation.

Germline variants in miRNAPG are identified in FHWT

Constitutional *DICER1* mutations result in development of pleuropulmonary blastoma syndrome (PPBS), which includes cystic nephroma, and extremely rarely, WT [Doros et al., 2014; Hill et al., 2009; Slade et al., 2011]. Given the recent documentation of rare *DICER1* and *DROSHA* germline variants in patients with WT [Rakheja et al., 2014], we examined the discovery set for germline exonic variants in the miRNAPGs and *SIX1/2* and identified 4 variants (Table S1, Figure 2A). These were verified with Sanger sequencing, expressed by mRNAseq, predicted to be damaging by PolyPhen Version 2 [Adzhubei et al., 2010], not annotated as polymorphisms in dbSNP versions 134 and 135 [Sherry et al., 2001], or present in the 1000 Genomes Pilot Projects 1, 2, and 3 [Abecasis et al., 2012]. Further, these variants were not identified in a dataset of over 200,000 individuals in the NHLBI GO Exome Sequencing Project (http://evs.gs.washington.edu/EVS/). A *DROSHA* variant (p.P82T) involved the proline-rich domain implicated in protein-protein and protein-nucleic acid interactions [Nicholson, 2014]. The remaining allele was retained in the tumor sample. A germline *XPO5* nonsense variant p.R159*, located in the Exportin-1 domain, was identified in the same patient who had the somatic nonsense *XPO5* mutation. Two missense germline *DICER1* variants were identified, one (p.R1368C) within the RNase IIIA domain, and one (p.I85M) in the helicase ATP-binding domain. Neither of the tumor samples showed loss of the remaining *DICER1* allele. Analysis of the validation set revealed 9/534 damaging *DICER1* variants in 8 patients, for an overall frequency of 1.5%. Eight variants affected the RNase IIIB domain and 6/8 were at previously identified hotspot locations [Doros et al., 2014]. Analysis of the mRNA and miRNA expression patterns of the 3 germline mutant tumors that lacked a somatic mutation revealed patterns distinct from that of the somatic miRNAPG mutations (Figures 1D, 2C, 2E, and 3).

DISCUSSION

Wilms tumor (WT) is an embryonal tumor of the kidney remarkable for its replication of early renal development. While mutations or deletions in *WT1*, *WTX*, and/or *CTNNB1* are found in approximately 30% of WTs [Ruteshouser et al., 2008], the underlying pathogenesis of most WTs remains unknown. Hence, in the last few years several groups of investigators simultaneously embarked on in-depth molecular characterization studies to further elucidate the genetic landscape of WTs. Two groups recently reported their findings [Rakheja et al., 2014; Torrezan et al., 2014] and another group is reporting their findings in this journal [Wegert et al., 2015]. All point to the importance of the *DROSHA* E1147K missense mutation in the development of WTs. Torrezan et al. evaluated a family trio by WXS, revealing a *DROSHA* E1147K mutation and prompted sequencing of the *DROSHA* RNase

IIIB domain in a validation set of 221 FHWTs (including a mixture of pre- and post-therapy samples from patients treated on different protocols). Rakheja et al. performed WXS in 15 patients followed by validation in 29 tumors, identifying three somatic *DROSHA* mutations (2 E1147K and 1 D1151Y). Our study offers the benefits of a much larger discovery set (77 tumors) and validation set (534 tumors) comprised exclusively of pre-therapy samples. This allowed for the identification of 1) recurrent somatic mutations in *DGCR8* E518K, 2) recurrent *DROSHA* mutations other than E1147K, 3) recurrent *SIX1/2* homeodomain mutations, 4) the association between 11p15 LOI and miRNAPG and *SIX* mutations, and 5) decreased expression of the miR-200 family in miRNAPG-HS mutant tumors, supporting the role of MET arrest in the function of these mutations. Lastly, the study of a large number of patients treated on a cooperative group protocol allowed documentation of the impact of mutations on clinical and pathologic features, including the association with blastemal histology, nephrogenic rest status, timing in renal development, the female predominance in miRNAPG-mutant tumors, and the poor clinical outcome of patients with both miRNAPG-HS and *SIX1/2* mutations. The identification of RAS activation in such tumors suggests they may be treatable in the future with precision medicine.

Synthesis of mature miRNA requires normal function of *DGCR8*, *DROSHA*, *XPO5*, and *DICER1*. In brief, primary miRNAs (pri-miRNA) are cleaved in the nucleus by the DROSHA-DGCR8 microprocessor complex to form precursor miRNA (pre-miRNA), which are exported from the nucleus by XPO5. Within the cytoplasm, DICER1 cleaves the pre-miRNA to form mature miRNAs [Winters et al., 2009]. Given the multitude of cellular pathways miRNAs are known to affect, combined with their interactions and feedback loops, the range of effects associated with miRNAPG mutations is likely to be heterogeneous and complex [Hua et al., 2013]. Impaired miRNA synthesis has been shown to accelerate oncogenic transformation by deregulating target oncogenes and globally reducing mature miRNA levels [Kumar et al., 2007]. Further, miRNAs have an essential and unique role during mammalian kidney development [Bartram et al., 2013; Ho et al., 2013]. Recent studies have shown that *DROSHA* RNase IIIB mutations result in global impairment of miRNA processing, with specific impairment in tumor-suppressing miRNAs [Rakheja et al., 2014]. We demonstrate that miRNAPG mutations are associated with down-regulation of all members of the miR-200 family (miR-200a, -200b, -200c, -141, and -429), which are key regulators of MET [Hua et al., 2013]. Reduction of miR-200 results in a mesenchymal, highly motile, and aggressive phenotype of cancer cells [Ceppi et al., 2014; Park et al., 2008; Ceppi et al., 2010]. MET is a critical step in early renal development during which the capacity to form nephrons occurs [Kobayashi et al., 2008]. Therefore, decreased expression of these miRNAs in the pre-induction metanephric mesenchyme would prevent MET, resulting in failure of epithelial differentiation and a predominance of undifferentiated cells, as was seen in the miRNAPG-mutant tumors we report.

Another miRNA, *Let-7a*, has long been linked to tumor development [reviewed in Garzon et al., 2009] and decreased *Let-7a* expression has been implicated in the development of WT via up-regulation of *LIN28* [Urbach et al., 2014]. *LIN28* is a RNA-binding protein that specifically binds to *PRI/PRE-Let-7* miRNAs, preventing maturation [Viswanathan et al., 2009]. In the murine embryonic kidney, over-expression of *LIN28* within the pre-induction

metanephric mesenchyme results in sustained proliferation, failure of MET, and tumor formation, a process that is rescued by *Let-7* over-expression [Urbach et al., 2014]. Of particular interest, over-expression of *LIN28* in the post-induction metanephric mesenchyme failed to result in tumor development suggesting that the effects of miRNAPG mutations depend on the cellular context in which they arise. Rakheja et al. functionally confirmed that *DROSHA* E1147K and D1151Y mutations result in decreased expression of the *Let-7* family within an *in vitro* model; in this study we now confirm decreased *Let-7a* in a large population of miRNAPG-mutant WTs.

Non-recurrent germline variants were also identified in miRNAPG, although these tumors were outliers by gene expression and miRNA expression compared with those containing somatic mutations. Therefore, germline miRNAPG variants do not appear to function in the same manner as the somatic miRNAPG mutations, and there is no direct evidence that they are pathogenic. Germline *DICER1* variants constitute the greatest clinical concern due to their association with familial PPBS. However, available data suggest that this syndrome follows a classic two-hit model of tumorigenesis, with germline truncating *DICER1* mutations followed by deleterious somatic missense mutations involving the RNase IIIB domain [Doros et al., 2011]. Both germline and somatic *DICER1* mutations were observed in 3 WTs by Wu et al. [2013] and in one WT by Rakheja et al. [2014], who also describe an additional patient with a germline *DICER1* mutation only, similar to our two discovery cases. The contribution of germline *DICER1* variants (as well as germline *DROSHA* and *XPO5* variants) in such patients is not clear. However, their repeated identification in patients with WT suggests they may result in a predisposition to WT; our study documents this risk to be present in approximately 1.3% of FHWT.

Recurrent mutations involving a specific residue of the homeodomain of transcription factors *SIX1* and *SIX2* were identified in 7% of FHWT. The highly homologous *SIX1* and *SIX2* genes have a critical role in renal development [Christensen et al. 2008; Zu et al. 2003]. *SIX1*-deficient mice exhibit renal hypoplasia or agenesis [Li et al., 2003], and *SIX1* mutations have been reported in the Branchio-Oto-Renal syndrome (BOR), although at positions distant from Q177R; BOR syndrome is not associated with WT [Patrick et al., 2009]. *SIX2* maintains a population of undifferentiated renal blastemal cells, and loss of *SIX2* results in premature differentiation of mesenchymal cells into epithelia [Kobayashi et al., 2008; Self et al. 2006]. Within a renal cell line, over-expression of *SIX2* results in an increased percentage of cells in the S-phase and increased migration [Senanayake, et al., 2013]. Given the known function of *SIX1* and *SIX2*, the localization of mutations within the *SIX* homeodomain, and the high expression of both mutant and wild-type alleles, it is probable that these *SIX1/2* mutations in FHWT are activating, resulting in failure of MET and continued proliferation of the metanephric mesenchyme. Indeed, we demonstrate significant up-regulation of *CCND2* in *SIX*-mutant tumors, which may be an important underlying cause of the continued proliferation. These findings are supported by those of Wegert et al. [2015], who report increased expression of both cell cycle genes and genes highly expressed in the pre-induction metanephric mesenchyme.

In summary, mutations in miRNAPG and/or *SIX1/2* genes are identified in approximately 20% of FHWT. Clinical, pathologic, gene and miRNA expression data support disruption of

MET at the time of induction as the underlying mechanism of tumorigenesis in this group of WTs, although no direct functional data is presented to confirm this hypothesis. Lastly, the very high prevalence of 11p15 LOI in WTs harboring both miRNAPG-HS and *SIX* Q177R mutations provides further evidence that multiple genetic events may be involved in the development and progression of WTs.

EXPERIMENTAL PROCEDURES

The TARGET initiative maintains public availability of the gene expression, chromosome copy number, DNA methylation, sequence analysis (i.e. MAF and summary files), and clinical information for the cases studied (available through the TARGET Data Matrix;http://target.nci.nih.gov/dataMatrix/TARGET_DataMatrix.html) in fully annotated MIAME compliant MAGE-TAB files describing the methods, specimen processing details, and quality control parameters. The aligned sequencing data (BAM files) are deposited in the Sequence Read Archive (SRA) at the National Center for Biotechnology Information and are accessible through the database of genotypes and phenotypes (dbGAP, http://www.ncbi.nlm.nih.gov/gap) under the accession number phs000471. See Supplemental Experimental Procedures for details.

Specimens

Pre-therapy tumor and normal DNA from peripheral blood or kidney from 77 FHWTs banked by COG with parental informed consent were included in the discovery set. A validation set of patients registered on the NWTS-5 protocol included all patients with available primary tumor DNA who subsequently relapsed and a random selection of all patients irrespective of relapse, resulting in 534 tumors enriched for relapse. Studies were performed with the approval of the Lurie Children's Hospital Institutional Review Board.

DNA Sequencing

WGS libraries were sequenced using Complete Genomics Inc. (CGI) technology [Drmanac et al., 2010]; alignment of reads to the NCBI Build 37 reference human genome assembly was performed by the CGI Cancer Sequencing service analytic pipeline version 2 [Carnevali et al., 2012]. WXS was performed on the Illumina HiSeq platform. Variant calling from the aligned BAM files was performed using both ATLAS and SAMtools and annotation and filtering was performed using the SACBE annotation pipeline [Bainbridge et al., 2013; Lupski et al., 2013] as well as Bambino Version 1.05 [Edmonson et al., 2011].

mRNA and miRNA Sequencing

Libraries were prepared following a paired-end protocol and sequenced on the Illumina HiSeq 2000/2500 platform using HiSeq Control Software version 2.0.10, aligned to GRCh37-lite genome-plus-junctions reference [Morin et al., 2008] using BWA Version 0.5.7 [Li et al., 2009]. For mRNA analysis, variants were detected on positive- and negative-split BAMs separately and annotated with SnpEff [Cingolani et al., 2012a] (Ensembl 66) and SnpSift [Cingolani et al., 2012b] (dbSNP137 and COSMIC64). For miRNA analysis, reads aligning to known miRNAs in miRBase v20 were summed and normalized to a million miRNA-aligned reads to generate the quantification files.

Target Capture Sequence Analysis

Probes were designed using Agilent's SureDesign (https://earray.chem.agilent.com/suredesign/) and probe density was specified at 2x with 98.7144% coverage of the target region (Agilent SureSelect XT Custom 0.5–2.9Mb probes). Genomic DNA libraries were constructed as described above and hybridized to the RNA probes. Post-capture material was enriched with 10 cycles of PCR. Paired-end 100 base reads were sequenced on the Illumina HiSeq2500 instrument. SNVs were filtered out if not predicted to be damaging by at least 2/3 of the following: SIFT [McLaren et al., 2010], PolyPhen Version 2 [Adzhubei et al., 2010], or Provean Version 1.1.3 [Choi et al., 2012].

Copy Number Analysis was performed on tumor and normal pairs according to the manufacturer's protocol for the AffyMetrix 6.0 SNP array and processed using AffyMetrix Genotyping Console 4.0 software. Reference normalization utilizing a diploid chromosome for each sample [Pounds et al., 2009] was performed in R using the DNAcopy BioConductor package. Segmented regions were identified by Circular binary segmentation (CBS) and those containing at least 8 markers in which the log2 value was ≥ +0.5 or ≤ −0.5 were considered gained or lost, respectively.

Gene Expression Analysis was analyzed with the Affymetrix U133+2 chip, according to the manufacturer's protocol using the Gene-Chip Operating Software and normalized using robust Multichip Average normalization. Unsupervised analysis was performed using Non-negative Matrix Factorization Consensus Version 5 [Brunet et al., 2004]. Gene Set Enrichment Analyses, version 2.0.14, (http://www.broadinstitute.org/gsea) [Subramanian et al., 2005] were run using 1000 permutations and phenotype permutation. Significant enrichment was defined as those lists with >50 genes, a FDR <10%, and a p-value <5%.

Methylation Analysis was performed with Illumina Infinium Human Methylation 450K Bead Chips (Illumina, San Diego, CA), according to the manufacturer's protocol. Methylation levels for all probes in imprint control regions ICR1 (*IGF2/H19*) and ICR2 (*KCNO1/CDKN1C*) were averaged. ROI was defined as 0.3–0.7 for ICR1 and ICR2, LOI as 0.8–1 for ICR1 and 0.3–0.7 for ICR2, and LOH as 0.8–1 for ICR1 and 0–0.2 for ICR2. Tumors outside of these ranges were not classified.

Let-7a Mature and Primary miRNA Expression

Reverse transcription, amplification, and real time PCR were performed per the manufacturer's protocol (Life Technologies Corporation, Carlsbad, CA). Samples were run in triplicate and analyzed using the Applied Biosystems® 7500 Fast SDS Software (Life Technologies Corporation, Carlsbad, CA). Nuclear and cytoplasmic fractions from HEK293 cells (ATCC, Manassas, VA) were evaluated for mature and primary *Let-7a* expression to confirm probe specificity.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The TARGET initiative is supported by NCI Grant U10 CA98543. Work performed under contracts from the NCI, US NIH within HHSN261200800001E includes specimen processing (the COG Biopathology Center), WGS (CGI, Inc.), WXS (Baylor College of Medicine), miRNAseq, RNAseq, and target capture sequencing (BCCA Genome Sciences Center). The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does the mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. The authors thank Patee Gesuwan and Leandro Hermida and the Data Coordinating Center for their support. We also thank the Clinical Applications of Core Technology Laboratory of the Hartwell Center for Bioinformatics and Biotechnology of St Jude Children's Research Hospital for performing the copy number analysis, and the Northwestern University Genomic Core facility for performing the methylation analysis. This work is also supported by NCI T32 CA079447 (ALW); NIH U10CA42326 (EJP); U10CA98543 (JSD, EJP); U24 CA114766; UO1CA88131 (EJP), and the American and Lebanese Syrian Associated Charities of St Jude (JM, CM). The authors are grateful for the expertise of Karen Novik and Laura Monovich, Patricia Beezhold, Donna Kersey, Debbie-Payne Turner, Mary McNulty, and Yvonne Moyer. This work would not be possible without the dedication of all the experts within the many clinical disciplines at the local institutions and centrally within the COG and NWTSG.

References

Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491:56–65. [PubMed: 23128226]

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. Nat Methods. 2010; 7:248–249. [PubMed: 20354512]

Bainbridge MN, Hu H, Muzny DM, Musante L, Lupski JR, Graham BH, Chen W, Gripp KW, Jenny K, Wienker TF, et al. *De novo* truncating mutations in *ASXL3* are associated with a novel clinical phenotype with similarities to Bohring-Opitz syndrome. Genome Med. 2013; 5:11. [PubMed: 23383720]

Bardeesy N, Falkoff D, Patruzzi M, Nowak N, Zabel B, Adam M, Aquiar MC, Grundy P, Shows T, Pelletier J. Anaplastic Wilms' tumour, a subtype displaying poor prognosis, harbours p53 gene mutations. Nat Genet. 1994; 7:91–97. [PubMed: 8075648]

Bartram MP, Höhne M, Dafinger C, VÖlker LA, Albersmeyer M, Heiss J, GÖbel H, BrÖnneke H, Burst V, Liebau MC, et al. Conditional loss of kidney microRNAs results in congenital anomalies of the kidney and urinary tract (CAKUT). J Mol Med. 2013; 91:739–748. [PubMed: 23344677]

Beckwith JB, Palmer NF. Histopathology and Prognosis of Wilms Tumor. Results from the First National Wilms' Tumor Study. Cancer. 1978; 41:1937–1948. [PubMed: 206343]

Beckwith JB, Kiviat NB, Bonadio JF. Nephrogenic Rests, nephroblastomatosis, and the pathogenesis of Wilms' Tumor. Pediatr Pathol. 1990; 10:1–36. [PubMed: 2156243]

Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. Proc Natl Acad Sci. 2004; 101:4164–4169. [PubMed: 15016911]

Carnevali P, Baccash J, Halpern AL, Nazarenko I, Nilsen GB, Pant KP, Ebert JC, Brownley A, Morenzoni M, Karpinchyk V, et al. Computational techniques for human genome resequencing using mated gapped reads. J Comput Biol. 2012; 19:279–292. [PubMed: 22175250]

Ceppi P, Mudduluru G, Kumarswamy R, Rapa I, Scagliotti GV, Papotti M, Allgayer H. Loss of miR-200c expression induces an aggressive, invasive, and chemoresistant phenotype in non-small cell lung cancer. Mol Cancer Res. 2010; 8:1207–1216. [PubMed: 20696752]

Ceppi P, Peter ME. MicroRNAs regulate both epithelial-to-mesenchymal transition and cancer stem cells. Oncogene. 2014; 33:269–278. [PubMed: 23455327]

Chao LY, Huff V, Tomlinson G, Riccardi VM, Strong LC, Saunders GF. Genetic mosaicism in normal tissues of Wilms' tumour patients. Nat Genet. 1993; 3:127–131. [PubMed: 8388768]

Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the Functional Effect of Amino Acid Substitutions and Indels. PLoS ONE. 2012; 7:e46688. [PubMed: 23056405]

Christensen KL, Patrick AN, McCoy EL, Ford HL. The Six Family of Homeobox Genes in Development and Cancer. Adv Cancer Res. 2008; 101:93–126. [PubMed: 19055944]

Cingolani P, Platts A, Wangle L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin). 2012a; 6:80–92. [PubMed: 22728672]

Cingolani P, Patel VM, Coon M, Nguyen T, Land SJ, Ruden DM, Lu X. Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. Front Genet. 2012b; 3:35. [PubMed: 22435069]

Dome JS, Liu T, Krasin M, Lott L, Shearer P, Daw NC, Billups CA, Wilimas JA. Improved Survival for Patients With Recurrent Wilms Tumor: The Experience at St. Jude Children's Research Hospital. J Pediatr Hematol Onc. 2002; 24:192–198.

Dome JS, Fernandez CV, Mullen EA, Kalapurakal JA, Geller JI, Huff V, Gratias EJ, Dix DB, Ehrlich PF, Khanna G, et al. Children's Oncology Group's 2013 Blueprint for Research: Renal Tumors. Pediatr Blood Cancer. 2013; 60:994–1000. [PubMed: 23255438]

Doros LA, Rossi CT, Yang J, Field A, Williams GM, Messinger Y, Cajaiba MM, Perlman EJ, Schultz KA, Cathro HP, et al. *DICER1* mutations in childhood cystic nephroma and its relationship to *DICER1*-renal sarcoma. Mod Pathol. 2014:1–14.

Drmanac R, Sparks AB, Callow MJ, Halpern AL, Burns NL, Kermani BG, Carnevali P, Nazarenko I, Nilsen GB, Yeung G, et al. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. Science. 2010; 327:78–81. [PubMed: 19892942]

Edmonson MN, Zhang J, Yan C, Finney RP, Meerzaman DM, Buetow KH. Bambino: a variant detector and alignment viewer for next-generation sequencing data in the SAM/BAM format. Bioinformatics. 2011; 27:865–866. [PubMed: 21278191]

Forbes SA, Bindal N, Bamford S, Cole C, Kok CY, Beare D, Jia M, Shepherd R, Leung K, Menzies A, et al. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucl Acids Res. 2010; 39:D945–D950. [PubMed: 20952405]

Gadd S, Huff V, Huang CC, Ruteshouser EC, Dome JS, Grundy PE, Breslow N, Jennings L, Green DM, Beckwith JB, et al. Clinically Relevant Subsets Identified by Gene Expression Patterns Support a Revised Ontogenic Model of Wilms Tumor: A Children's Oncology Group Study. Neoplasia. 2012; 14:742–756. [PubMed: 22952427]

Garzon R, Calin GA, Croce CM. MicroRNAs in Cancer. Ann Rev Med. 2009; 60:167–179. [PubMed: 19630570]

Gurtan A, Lu V, Bhutkar A, Sharp PA. In vivo structure-function analysis of human Dicer reveals directional processing of precursor miRNAs. RNA. 2012; 18:1116–1122. [PubMed: 22546613]

Hill DA, Ivanovich J, Priest JR, Gurnett CA, Dehner LP, Desruisseau D, Jarzembowski JA, Wikenheiser-Brokamp KA, Suarez BK, Whelan AJ, et al. DICER1 mutations in familial pleuropulmonary blastoma. Science. 2009; 325:965. [PubMed: 19556464]

Ho J, Kreidberg JA. MicroRNAs in renal development. Pediatr Nephrol. 2013; 28:219–225. [PubMed: 22660936]

Howlader, N.; Noone, AM.; Krapcho, M.; Garshell, J.; Neyman, N.; Altekruse, SF.; Kosary, CL.; Yu, M.; Ruhl, J.; Tatalovich, Z., et al. SEER Cancer Statistics Review, 1975–2010. Bethesda, MD: National Cancer Institute; 2013. http://seer.cancer.gov/csr/1975_2010/

Hua Y, Larsen N, Kalyana-Sundaram S, Kjems J, Chinnaiyan AM, Peter ME. miRConnect 2.0: identification of oncogenic, antagonistic miRNA families in three human cancers. BCM Genomics. 2013; 14:179. (http://www.biomedcentral.com/1471-2164/14/179).

Hu Q, Gao F, Tian W, Ruteshouser EC, Wang Y, Lazar A, Stewart J, Strong LC, Behringer RR, Huff V. Wt1 ablation and Igf2 upregulation in mice result in Wilms tumors with elevated ERK1/2 phosphorylation. J Clin Invest. 2011; 121:174–183. [PubMed: 21123950]

Kelly LA, Sternberg MJ. Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc. 2009; 4:363–371. [PubMed: 19247286]

Kobayashi A, Valerius MT, Mugford JW, Carroll TJ, Self M, Oliver G, McMahon AP. Six2 defines and regulates a multipotent self-renewing nephron progenitor population throughout mammalian kidney development. Cell Stem Cell. 2008; 3:169–181. [PubMed: 18682239]

Kumar MS, Lu J, Mercer KL, Golub TR, Jacks T. Impaired microRNA processing enhances cellular transformation and tumorigenesis. Nature Genet. 2007; 39:673–677. [PubMed: 17401365]

Lambertz I, Nittner D, Mestdagh P, Denecker G, Vandesompele J, Dyer MA, Marine JC. Monoallelic but not biallelic loss of Dicer1 promotes tumorigenesis in vivo. Cell Death Differ. 2010; 4:633–641. [PubMed: 20019750]

Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25:1754–1760. [PubMed: 19451168]

Li X, Ohgi KA, Zhang J, Krones A, Bush KT, Glass CK, Nigam SK, Aggarwal AK, Maas R, Rose DW, et al. Eya protein phosphatase activity regulates Six1-Dach-Eya transcriptional effects in mammalian organogenesis. Nature. 2003; 426:247–254. [PubMed: 14628042]

Lupski JR, Gonzaga-Jauregui C, Yang Y, Bainbridge MN, Jhangiani S, Buhay C, Kovar CL, Wang M, Hawes AC, Reid JG, et al. Exome sequencing resolves apparent incidental findings and reveals further complexity of SH3TC2 variant alleles causing CMT neuropathy. Genome Med. 2013; 5:57–70. [PubMed: 23806086]

Maiti S, Alam R, Amos CI, Huff V. Frequent Association of *β-Catenin* and *WT1* Mutations in Wilms Tumors. Cancer Res. 2000; 60:6288–6292. [PubMed: 11103785]

McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with Ensembl API and SNP Effect Predictor. Bioinformatics. 2010; 26:2069–2070. [PubMed: 20562413]

Morin R, Bainbridge M, Fejes A, Hirst M, Krzywinski M, Pugh T, McDonald H, Varhol R, Jones S, Marra M. Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. Biotechniques. 2008; 45:81–94. [PubMed: 18611170]

Nicholson AW. Ribonuclease III mechanisms of double-stranded RNA cleavage. Wires RNA. 2014; 5:31–48. [PubMed: 24124076]

Okada C, Yamashita E, Lee SJ, Shibata S, Katahira J, Nakagawa A, Yoneda Y, Tsukihara T. A High-Resolution Structure of the Pre-microRNA Nuclear Export Machinery. Science. 2009; 326:1275–1279. [PubMed: 19965479]

Patrick AN, Schiemann BJ, Yang K, Zhao R, Ford HL. Biochemical and Functional Characterization of Six *SIX1* Branchio-oto-renal Syndrome Mutations. J Biol Chem. 2009; 284:20781–20790. [PubMed: 19497856]

Pounds S, Cheng C, Mullighan C, Rimondi SC, Shurtleff S, Downing JR. Reference Alignment of SNP Microarray Signals for Copy Number Analysis of Tumors. Bioinformatics. 2009; 25:315–321. [PubMed: 19052058]

Rakheja D, Chen KS, Liu Y, Shukla AA, Schmid V, Chang T, Khokhar S, Wickiser JE, Karandikar NJ, Malter JS, et al. Somatic mutations in *DROSHA* and *DICER1* impair microRNA biogenesis through distinct mechanisms in Wilms tumours. Nat Commun. 2014; 5:4802. [PubMed: 25190313]

Ravenel JD, Broman KW, Perlman EJ, Niemitz EL, Jayawardena TM, Bell DW, Haber DA, Uejima H, Feinberg AP. Loss of Imprinting of Insulin-Like Growth Factor-II (IGF2) Gene in Distinguishing Specific Biologic Subtypes of Wilms Tumor. J Natl Cancer Inst. 2001; 93:1698–1703. [PubMed: 11717330]

Rivera MN, Haber DA. Wilms tumour: connecting tumorigenesis and organ development in the kidney. Nat Rev Cancer. 2005; 9:699–712. [PubMed: 16110318]

Ruteshouser EC, Robinson SM, Huff V. Wilms tumor genetics: mutations in *WT1*, *WTX*, and *CTNNB1* account for only about one-third of tumors. Gene Chromosome Canc. 2008; 47:461–470.

Self M, Lagutin OV, Bowling B, Hendrix J, Cai Y, Dressler DR, Oliver G. Six2 is required for suppression of nephrogenesis and progenitor renewal in the developing kidney. The EMBO J. 2006; 25:5214–5228. [PubMed: 17036046]

Senanayake U, Koller K, Puchler M, Leuschner I, Strohmaier H, Hadler U, Das S, Hoefler G, Guertl B. The pluripotent renal stem cell regulator *SIX2* is activated in renal neoplasms and influences cellular proliferation and migration. Human Pathology. 2013; 44:336–345. [PubMed: 22995329]

Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. dbSNP: the NCBI database of genetic variation. Nucl Acids Res. 2001; 29:308–311. [PubMed: 11125122]

Slade I, Bacchelli C, Davies H, Murray A, Abbaszadeh F, Hanks S, Barfoot R, Burke A, Chisholm J, Hewitt M, et al. *DICER1* syndrome: clarifying the diagnosis, clinical features and management

implications of a pleiotropic tumour predisposition syndrome. J Med Genet. 2011; 48:273–278. [PubMed: 21266384]

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci. 2005; 102:15545–15550. [PubMed: 16199517]

Torrezan GT, Ferreira EN, Nakahata AM, Barros BF, Castro MT, Correa BR, Krepischi AC, Olivieri EH, Cunha IW, Tabori U, et al. Recurrent somatic mutation in *DROSHA* induces microRNA profile changes in Wilms tumour. Nat Commun. 2014; 5:4039. [PubMed: 24909261]

Tusher V, Tibshirani R, Chu G. Significance analysis of microarrays applied to transcriptional responses to ionizing radiation. Proc Natl Acad Sci. 2001; 98:5116–5121. [PubMed: 11309499]

Urbach A, Yermalovich A, Zhang J, Spina CS, Zhu H, Perez-Atayde AR, Shukrun R, Charlton J, Sebire N, Mifsud W, et al. Lin28 sustains early renal progenitors and induces Wilms tumor. Genes Dev. 2014; 28:971–982. [PubMed: 24732380]

Viswanathan SR, Powers JT, Einhorn W, Hoshida Y, Ng TL, Toffanin S, O'Sullivan M, Lu J, Phillips LA, Lockhart VL, et al. Lin28 promotes transformation and is associated with advanced human malignancies. Nat Genet. 2009; 41:843–848. [PubMed: 19483683]

Wang G, Guo X, Hong W, Liu Q, Wei T, Lu C, Gao L, Ye D, Zhou Y, Chen J, et al. Critical regulation of miR-200/ZEB2 pathway in Oct4/Sox2-induced mesenchymal-to-epithelial transition and induced pluripotent stem cell generation. Proc Natl Acad Sci. 2013; 110:2858–2863. [PubMed: 23386720]

Wegert J, Ishaque N, Vardapour R, Geörg C, Gu Z, Bieg M, Ziegler B, Bausenwein S, Nourkami N, Ludwig N, et al. Mutations in the SIX1/2 pathway and the DROSHA/DGCR8 miRNA microprocessor complex underlie high-risk blastemal type Wilms tumors. Cancer Cell. 2015 in press.

Winter J, Jung S, Keller S, Gregory RI, Diederichs S. Many roads to maturity: microRNA biogenesis pathways and their regulation. Nat Cell Biol. 2009; 11:228–234. [PubMed: 19255566]

Wu MK, Sabbaghian N, Xu B, Addidou-Kalucki S, Bernard C, Zou D, Reeve AE, Eccles MR, Cole C, Choong CS, et al. Biallelic DICER1 mutations occur in Wilms tumours. J Pathol. 2013; 230:154–164. [PubMed: 23620094]

Zu PX, Zheng W, Huang L, Marie P, Laclef C, Silvius D. *Six1* is required for the early organogenesis of mammalian kidney. Development. 2003; 103:3085–3094.

SIGNIFICANCE

We establish recurrence of *DGCR8* E518K mutations, confirm the high frequency of recurrent *DROSHA* exon 29 mutations, identify recurrent hotspot mutations in the *SIX1/2* homeodomain, and identify high frequencies of 11p15 LOI in both miRNAPG and *SIX1/2* mutant FHWT. Mutations in miRNAPG were associated with dysregulation of microRNAs involved in oncogenesis and mesenchymal-to-epithelial transition, and increased frequency of undifferentiated histology. The combination of multiple genetic events in some FHWT, including mutations in both miRNAPG and *SIX1/2*, and 11p15 imprinting abnormalities, provides evidence of a complex, multi-step process resulting in failure of normal differentiation, maintenance of progenitor cells, and support of proliferation. The critical combination of such genetic events is shown to result in an adverse outcome and may be targetable.

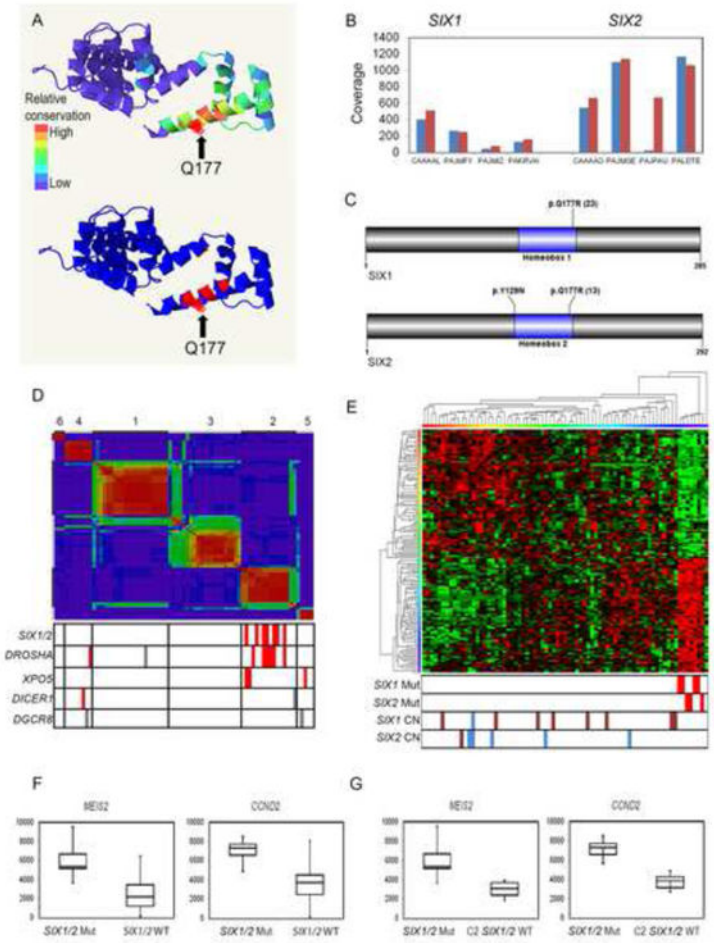

Figure 1. Recurrent *SIX1/2* Q177 mutations in FHWT

(A) Phyre2 images showing the Q177 residue of SIX1 to be located within a conserved region (top panel) and is predicted to be a specific DNA-contact base (bottom panel; blue and red indicate regions of low and high likelihood of DNA contact, respectively).

(B) Coverage of the reference allele (blue bar) and variant allele (red bar) as determined by mRNA-sequencing for *SIX1* and *SIX2* mutant tumors.

(C) Location of validation set variants within the SIX1 and SIX2 proteins; number of variants detected are provided in parenthesis.

(D) Unsupervised NMF clustering of 75 FHWT with annotation of mutations identified (bottom, Red= somatic, Gray = germline)

(E) Supervised hierarchical clustering of 75 FHWT according to the top 100 genes differentially expressed in *SIX1/2*-mutant tumors with annotation of tumors with *SIX1/2* mutations and copy number changes (blue = gain; dark red = loss) shown at the bottom.

(F) Boxplots of *MEIS2* and *CCND2* in *SIX1/2*-mutants versus wild-type FHWTs. The bottom and top of the box represent the first and third quartiles, respectively, the band inside the box represents the median, and the whiskers represent the maximum and minimum values.

(G) Boxplots of *MEIS2* and *CCND2* in NMF cluster 2 *SIX1/2*-mutants versus wild-type tumors.

See also Figure S1 and Tables S1–S3.

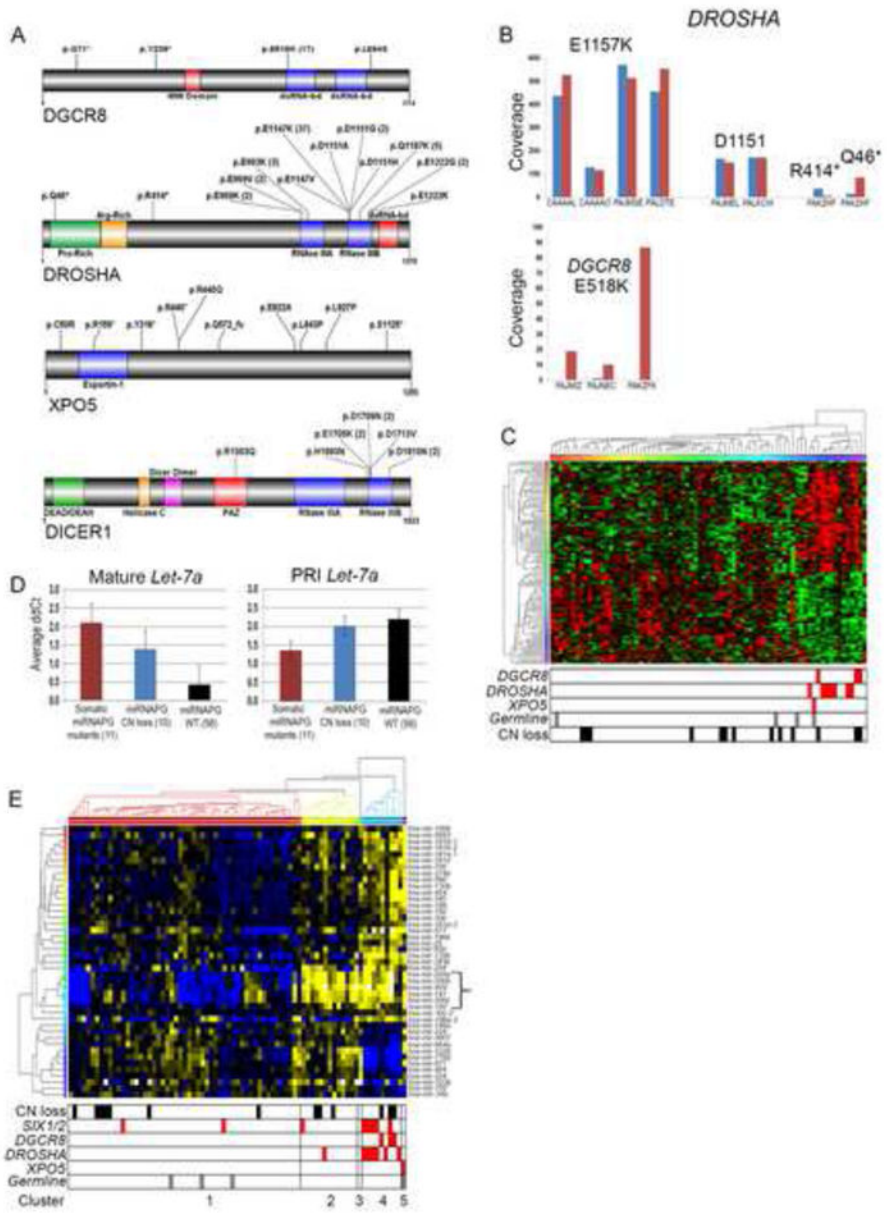

Figure 2. Recurrent miRNAPG hotspot mutations in FHWT

(A) Location of validation set variants within the DGCR8, DROSHA, XPO5, and DICER1 proteins; number of variants detected are provided in parenthesis.

(B) Coverage of the reference allele (blue bar) and variant allele (red bar) as determined by mRNA-sequencing for *DROSHA* (top panel) and *DGCR8* (bottom panel).

(C) Supervised hierarchical clustering of 75 FHWT according to the top 100 genes differentially expressed in miRNAPG-mutant tumors with annotation of miRNAPG mutations (red= somatic, gray = germline) and copy number loss.

(D) Mature *Let-7a* average ddCt (left panel) and primary *Let-7a* average ddCt (right panel) in FHWT with miRNAPG mutations (red bar), copy number loss (blue bar), and lacking both miRNAPG mutations and copy number loss (black bar). Error bars = +SEM.

(E) Hierarchical analysis of the 43 miRNAs significantly differentially expressed in somatic miRNAPG mutant FHWTs compared with those lacking both miRNAPG mutations and copy number loss with annotation of miRNAPG mutations and copy number loss. Blue and yellow represent relatively high and low expression, respectively. Five clusters were observed, as indicated at the bottom.
See also Tables S4–S6.

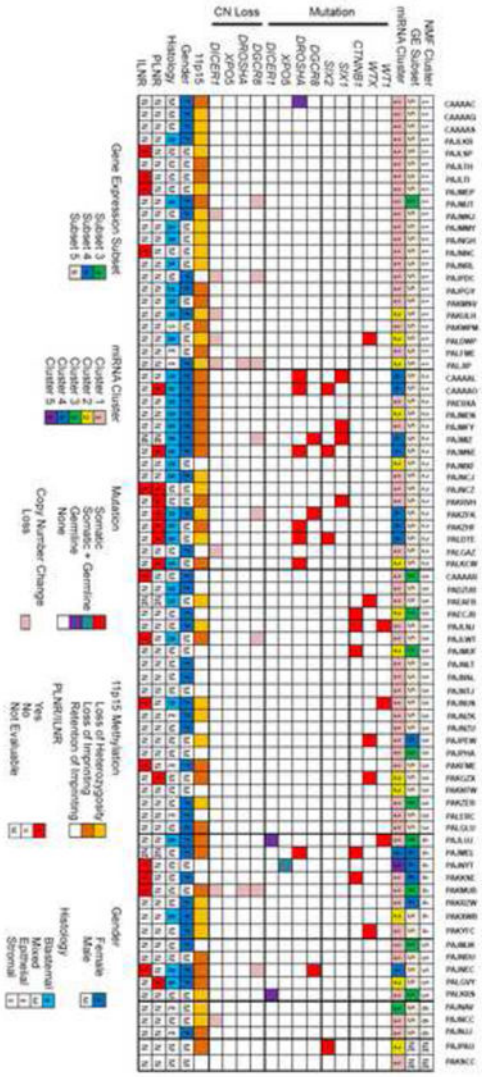

Figure 3. Integrative Analysis of Non-negative Matrix Factorization Clusters

Clinical, pathologic, and genetic features of FHWT arranged according to the NMF identified in Figure 1D. The key is illustrated at the bottom. See also Figure S2.

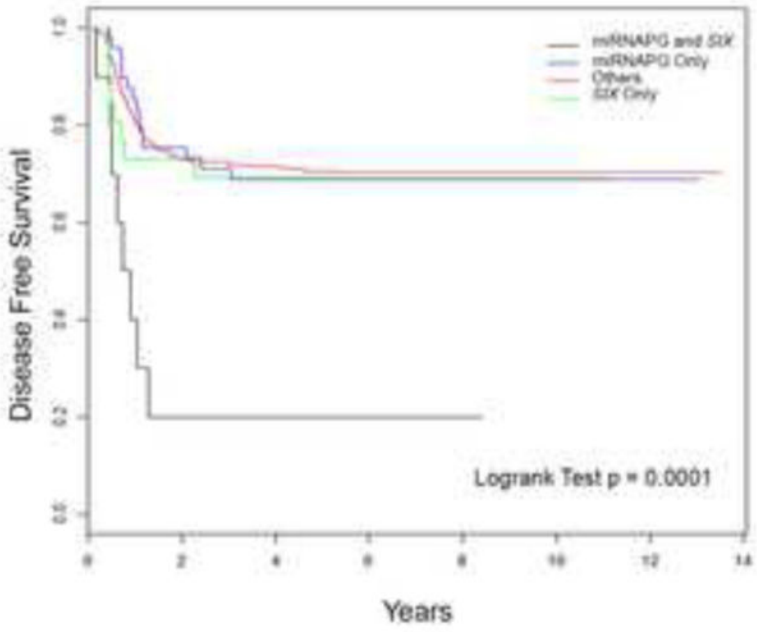

Figure 4. Disease Free Survival

Kaplan-Meier curve of disease free survival in the following four validation set groups: (1) tumors with *SIX1/2* and miRNAPG-HS variants (black line), (2) tumors with miRNAPG-HS variants without *SIX1/2* variants (blue line), (3) *SIX1/2* variants without miRNAPG-HS variants (green line), and (4) all other validation set tumors (red line).

Table 1

Significant Validation Set Patient Characteristics

Group	Total Number	Age at Diagnosis (months)	Gender (F:M)	Blastemal Histology	PLNR	ILNR
DGCR8 E518K Variants	17	57	15:2 (p=0.004)	32/59 (p=0.003)	27/55 (p=<0.001)	8/55 (p=0.079)
DROSHA exon 29 Variants	42	51	31:11 (p=0.009)			
SIX Q177R Variants	36	52	22:14 (not significant)	20/36 (p=0.008)	16/35 (p=0.001)	3/35 (p=0.026)
Entire Validation Set	534	44	290:244	189	119	125