



Clinicopathologic and molecular analysis of a *BCOR–CCNB3* undifferentiated sarcoma of the kidney reveals significant epigenetic alterations

Tamara J. Hagoel,^{1,2,9} Eduardo Cortes Gomez,^{3,9} Ajay Gupta,^{1,2} Clare J. Twist,^{1,2} Rafal Kozielski,⁴ Jeffrey C. Martin,⁵ Lingui Gao,⁵ Joseph Kuechle,⁶ Prashant K. Singh,⁷ Miranda Lynch,⁸ Lei Wei,⁸ Song Liu,³ Jianmin Wang,³ and Joyce E. Ohm⁵

¹Department of Pediatrics, University at Buffalo Jacobs School of Medicine and Biomedical Sciences, Buffalo, New York 14203, USA; ²Department of Pediatric Hematology-Oncology, ³Department of Biostatistics & Bioinformatics, Roswell Park Comprehensive Cancer Center, Buffalo, New York 14203, USA; ⁴Department of Pathology, Oishei Children's Hospital, Buffalo, New York 14203, USA; ⁵Department of Cancer Genetics and Genomics, ⁶Department of Medicine, ⁷Center for Personalized Medicine, Roswell Park Comprehensive Cancer Center, Buffalo, New York 14203, USA; ⁸Hauptman-Woodward Structural Institute, Buffalo, New York 14203, USA

Abstract Undifferentiated soft tissue sarcomas (UDSTSs) are a group of mesenchymal tumors that remain a diagnostic challenge because of their morphologic heterogeneity and unclear histologic origin (Peters et al., *Mod Pathol* **28**: 575 [2015]). In this case report, we present the first multiomics molecular signature for a *BCOR–CCNB3* sarcoma (BCS) that includes mutation analysis, gene expression, DNA methylation, and micro RNA (miRNA) expression. We identify a paucity of additional mutations in this tumor and detail that there is significant dysregulation of gene expression of epigenetic remodeling agents including key members of the PRC, Sin3A/3b, NuRD, and NcoR/SMRT complexes and the DNA methyltransferases DNMT1, DNMT3a, and DNMT3b. This is accompanied by significant DNA methylation changes and dysregulation of multiple miRNAs with known links to tumorigenesis. This study significantly increases our understanding of the *BCOR* effects on fusion-positive undifferentiated sarcomas at both the genomic and epigenomic level and suggests that as better-tailored and more refined treatment algorithms continue to evolve, epigenetic modifying agents should be further evaluated for their efficacy against these tumors.

Corresponding author:
joyce.ohm@roswellpark.org

© 2022 Hagoel et al. This article is distributed under the terms of the Creative Commons Attribution-NonCommercial License, which permits reuse and redistribution, except for commercial purposes, provided that the original author and source are credited.

Ontology term: renal sarcoma

Published by Cold Spring Harbor Laboratory Press

doi:10.1101/mcs.a005942

[Supplemental material is available for this article.]

INTRODUCTION

Undifferentiated soft tissue sarcomas (UDSTSs) are a group of mesenchymal tumors that remain a diagnostic challenge because of their morphologic heterogeneity and unclear

⁹These authors contributed equally to this work.

histologic origin (Peters et al. 2015b). Recurrent gene fusions have been implicated in many UDSTSs and their presence is often used to refine diagnosis (Mertens et al. 2016; Laetsch et al. 2018). In 2012, Pierron et al. discovered a subset of round-cell undifferentiated sarcomas that share morphologic and immunohistochemical properties with Ewing sarcoma (ES), yet lack the defining ES chimeric gene fusions between *EWSR1* and *FUS* to the ETS family of transcription factors. These rare mimickers acquired the title “Ewing-like” sarcomas and result from an intrachromosomal paracentric inversion on the short arm of Chromosome X fusing the *BCOR* and *CCNB3* genes together. Multiple breakpoints have been identified in the 3′ untranslated region (UTR) or the *MID1IP1* intergenic region of *BCOR* and intron 4 of *CCNB3*. A splice acceptor site is embedded in the stop codon of *BCOR*’s exon 15 that splices with exon 5 of *CCNB3*. The fusion transcript thus is the product of *BCOR* exons 1–15 fusing to *CCNB3* exons 5–12, and its expression relies on the *BCOR* promoter (Pierron et al. 2012; Puls et al. 2014; Peters et al. 2015b; Machado et al. 2016; Le Loarer et al. 2017; Kao et al. 2018; Renzi et al. 2019). Gene expression profiling suggests that these tumors are biologically distinct from ES and other pediatric round-cell sarcomas (Pierron et al. 2012; Kao et al. 2018; Renzi et al. 2019).

Since their discovery, *BCOR–CCNB3* sarcomas (BCSs) continue to be a subject of investigation and clinical interest. Several additional case series have highlighted the clinical, morphologic, and genomic similarities and differences between BCS and ES (Puls et al. 2014; Peters et al. 2015b; Argani et al. 2017; Kao et al. 2018; Koelsche et al. 2018). Other reports suggest an overlap between clear cell sarcoma of the kidney (CCSK) and other *BCOR*-related sarcomas of the kidney and soft tissues including BCS and primitive myxoid mesenchymal tumor of infancy (PMMTI) (Argani et al. 2017; Kao et al. 2018), many of which contain *BCOR* internal tandem duplications (*BCOR*-ITDs) (Astolfi et al. 2019). A distinctive *BCOR*-driven transcriptional profile has been identified, suggesting that all of these tumors are genetically related (Argani et al. 2017), although additional research is needed at the molecular level to determine whether these tumors should be considered the same entity for clinical purposes. We now present the first case report detailing a multiomics molecular signature for BCS that includes mutation analysis, gene expression, DNA methylation, and micro RNA (miRNA) expression.

BCS is most often detected in bone with equal distribution in the axial skeleton and long bones and less commonly present in soft tissue and rarely the viscera (most often the kidney) of primarily adolescent males (Pierron et al. 2012; Puls et al. 2014; Peters et al. 2015b; Le Loarer et al. 2017; Kao et al. 2018). Histologically, BCS has a round or ovoid and sometimes spindle-shaped appearance. Molecularly, they exhibit patchy cytoplasmic expression of CD99 and diffuse intranuclear expression of *CCNB3* and *BCOR* (Kao et al. 2016; Le Loarer et al. 2017; Kao et al. 2018). Currently, a definitive diagnosis of BCS relies on the detection of the *BCOR–CCNB3* translocation by reverse transcriptase-polymerase chain reaction (RT-PCR) and/or fluorescence in situ hybridization (FISH) (Pierron et al. 2012; Puls et al. 2014; Matsuyama et al. 2017; Kao et al. 2018; Renzi et al. 2019).

Because of the rarity and novelty of BCS, prospective clinical trial investigation of children with such diagnoses is lacking. Therefore, the clinical course and therapy-related response of patients with BCS can only be gleaned from case studies and retrospective reviews. As no uniform standardized treatment algorithm exists for BCS, patients identified with BCS sarcomas have been managed according to a variety of more established sarcoma regimens; most commonly utilized are regimens designed to treat ES. Treatment modalities have included a combination of surgical resection, focal radiation, and/or multiagent systemic neoadjuvant versus adjuvant chemotherapy with agents such as doxorubicin, etoposide, vincristine, cyclophosphamide, and ifosfamide (Pierron et al. 2012; Cohen-Gogo et al. 2014; Puls et al. 2014; Peters et al. 2015b; Kao et al. 2018). Retrospective data suggest that patients with BCS have a 5-yr survival rate equal to that of the localized ES (Cohen-Gogo et al. 2014; Puls et al. 2014; Peters et al. 2015b; Machado et al. 2016; Laetsch et al. 2018).

Here we present the clinical presentation, morphologic characteristics, treatment course and outcome, and associated multiomics molecular profile of a single adolescent male patient diagnosed with a *BCOR-CCNB3* undifferentiated round-cell sarcoma of the kidney. Treatment of our patient according to a regimen used to treat CCSK resulted in complete remission and no evidence of disease >36 mo following completion of therapy.

RESULTS

Clinical Presentation and Family History

A previously healthy 14-yr-old male with a history of a well-controlled seizure disorder presented to the emergency room with a 2-d history of left flank pain and a 3-d history of brown-tinged urine progressing to frank hematuria. There was no notable history of trauma, night sweats, chills, weight loss, or fevers. Upon initial evaluation, blood chemistries showed no evidence of tumor lysis syndrome, and urinalysis and urine cultures were negative for a urinary tract infection. His hemoglobin was 13.6 mg/dL and creatinine was 0.82 mg/dL. Family history was notable for maternal history of kidney stones, but no family history of cancer was documented.

Renal ultrasound with Doppler computed tomography (CT) scan of the torso and magnetic resonance imaging (MRI) scan of the abdomen and pelvis revealed a well-circumscribed enhancing focal left posterior upper pole 3.2 × 2.0 × 2.2-cm hypoechoic intrarenal lesion concerning a primary renal neoplasm with secondary hemorrhage into the upper pole calyces and resultant central collecting system dilation with hazy perinephric fat stranding adjacent to the renal hilum. No renal calculi or vascular thromboses were identified and no other lesions concerning pathologic lymphadenopathy or metastatic disease were noted in the abdomen or chest.

Urine cytology was sent for pathologic examination with no evidence of malignancy. The patient underwent an open left nephrectomy with regional lymph node sampling. Our patient did well postoperatively with no complications. The mass exhibited focal extension from the left renal pelvicalyceal system into the renal sinus soft tissue with no lymph node involvement. The preliminary pathology report morphologically identified a spindle cell neoplasm of the kidney. The following immunohistochemical stains were performed: vimentin, CD99, CD10, ER, PR, CD34, EMA, CK, desmin, Myod1, actin, Bcl-2, PAX8, WT1, and TLE1. Vimentin and Bcl-2 stains were strongly positive and suggestive of a CCSK, but nondiagnostic. Consultation with the Children's Oncology Group (COG) renal tumor pathologist facilitated further immunohistochemical testing which identified strong nuclear *BCOR* positivity and a *BCOR-CCNB3* fusion transcript was confirmed by PCR; our patient was therefore diagnosed with a stage II *BCOR-CCNB3* undifferentiated round-cell sarcoma of the kidney.

Baseline brain MRI and positron emission tomography scans were performed prior to initiation of treatment with no evidence of residual or metastatic disease. The patient was treated as per a COG protocol designed for the treatment of CCSK (AREN0322, Regimen I) given the *BCOR* abnormalities associated with CCSK. Our patient received a total of six fractions of radiation to the left flank totaling 1080 cGy followed by a multi-agent chemotherapy regimen consisting of alternating cycles of cyclophosphamide, doxorubicin, and vincristine with cyclophosphamide and etoposide over a course of 24 wk. His cumulative doses equaled the following: 14,000 mg/m² of cyclophosphamide, 225 mg/m² of doxorubicin, and 2000 mg/m² of etoposide and vincristine 22 mg/m².

Treatment Outcomes

Our patient suffered a total of three generalized tonic–clonic seizures during his treatment course. The seizures were managed by his neurologist and they subsided following adjustments to his anti-epileptic medications. He suffered no additional adverse events during the course of treatment. He completed his therapy following 6 mo of multimodal therapy. After therapy completion, our patient was evaluated for relapsed disease and secondary complications every 3 mo in our pediatric clinic with surveillance scans, a physical exam, and routine laboratory screening. He is presently 36 mo out from completion of therapy without any evidence of recurrent or relapsed disease.

Genomic Analyses

To gain a better understanding of the molecular events driving tumorigenesis in this patient, we performed whole-exome sequencing (WES), RNA-sequencing (RNA-seq), global methylation analysis (Illumina EPIC array), and miRNA-sequencing (miRNA-seq) on pretreatment formalin-fixed paraffin-embedded (FFPE) abnormal and adjacent matched normal kidney tissue belonging to our patient that was taken at the time of surgical resection.

Mutation Analysis

To identify possible additional mutations in our pretreatment tumor sample, WES was performed to identify somatic mutations in intronic, exonic, and protein-coding regions. Somatic DNA mutations were determined based on a comparison of the patient's tumor to germline from his matched normal tissue using multiple independent mutations' calling algorithms. The somatic variants were called using a multiple-caller ensembled method, including MuTect2 (v4.1.2.0) (Cibulskis et al. 2013), MuSE (v1.0) (Fan et al. 2016), Strelka2 (v2.9.10) (Saunders et al. 2012), and VarDict (v1.8) (Lai et al. 2016) and manually reviewed to remove false-positive calls. The analysis pipelines are reproducible and publicly available through the Rcw package (Hu et al. 2021). A total of 32 genes were identified to potentially harbor mutations, five of which were common among both mutations' calling packages and with sufficient read support (>20 reads from each of the tumor and normal samples spanning the mutation region) for consideration as potential mutations in this patient's tumor (Table 1). A sixth mutation in the NOX4 was identified to be in a repetitive region and likely attributed to being caused by a mapping artifact. Silent mutations were detected in two genes (*IGFBP3* and *FAM134C*), whereas missense mutations were identified in the three other genes (*ACP2*, *MLC1*, and *MPHOSPH9*).

The *ACP2* gene encodes lysosomal acid phosphatase 2, an enzyme found in the lysosome organelle. Mutations in the *ACP2* gene are associated with bone structure abnormalities, lysosomal storage defects, metabolic disorders, and abnormalities in the central nervous system. *ACP2* overexpression has been implicated in human colorectal cancer (Lee et al. 2017). *MLC1* mutations have been associated with megalencephalic leukoencephalopathy with subcortical cysts, a neurodegenerative disorder, as well as seizures (Dubey et al. 2018). *MPHOSPH9* polymorphisms are associated with multiple sclerosis susceptibility (Mowry et al. 2013). None of the above genes have been previously implicated in sarcomagenesis. The functional effects of our patient's specific mutations and their correlations to any clinical phenotypes are unknown and remain an area for future investigation. Overall, the mutation rate was very low compared to other cancer types, and after all quality filters and manual review, it was determined that this cancer appears to have a very quiet genome.

Table 1. Genetic mutations detected in tumor specimen

Gene	Chromosome	HGVS DNA reference	HGVS protein reference	Variant type	Predicted effect	dbSNP/dbVarID	Genotype	Position ^a	Mutant reads	Coverage	Reference allele	Mutant allele
IGFBP3	7	NM_001013398	L16L	Silent	N/A	N/A	N/A	45960694	10	58	G	A
ACP2	11	NM_001610	T215M	Missense	N/A	N/A	N/A	47266414	41	127	G	A
MPHOSPH9	12	NM_022782	S562L	Missense	N/A	N/A	N/A	123687267	14	69	G	A
FAM134C	17	NM_178126	E443E	Silent	N/A	N/A	N/A	40733903	16	73	C	T
MLC1	22	NM_015166	F76L	Missense	N/A	N/A	?N/A	50521554	15	75	A	G

(N/A) Not available.

^aReference assembly GRCh37.

Transcriptome Analysis

To identify differentially expressed genes between tumor and normal kidney and to identify genes and pathways that may be driving tumorigenesis in this patient, RNA-seq was performed comparing tumor to the adjacent normal kidney. Importantly, RNA-seq confirmed the presence of a *BCOR–CCNB3* fusion in the patient tumor sample that was absent from the normal kidney. Several other gene fusion events were identified using the STAR-Fusion program from STAR package (Haas et al. 2019). However, only two of these gene fusion calls were supported by a significant number of read counts in the RNA-seq data. First, the *BCOR–CCNB3* fusion was present as expected. The fusion-specific breakpoint was identified in *BCOR* exon 15 at Chr X: 39,911,366 and exon 5 of *CCNB3* at Chr X: 50,051,505 (Fig. 1). A fusion between Chr Y and Chr 16 resulting in an *RNA5 8SP6–RNA5 8SP2* fusion was also identified (data not shown), but was deemed to be a low confidence observation, as no spanning fragment reads were called. *RNA5-8SP6* (RNA, 5.8S Ribosomal Pseudogene 6) and *RNA5-8SP2* (RNA, 5.8S Ribosomal Pseudogene 2) are both pseudogenes, nonfunctional segments of DNA that resemble functional genes that arise as extra copies of functional genes, either directly by DNA duplication or indirectly by reverse transcription of an mRNA transcript. The biological significance of the presence of this potential second fusion is unclear and may represent a sequencing artifact.

Recognizing the absence of significant mutations beyond the *BCOR–CCNB3* fusion and in an attempt to increase our understanding of the genes and pathways driving tumorigenesis in our patient's tumor, differential gene expression (DGE) analysis was performed on aligned reads. Fold-change (FC) values and differential expression analyses for the two samples were estimated using EBSeq (Leng et al. 2013) (v1.18.1), which utilizes an empirical Bayesian model and can adequately handle data without replicates. Pathway gene expression analysis was performed using gene set enrichment analysis (GSEA). The top 10 up-regulated pathways in our tumor sample in comparison to paired normal kidney were analyzed and are shown in Table 2 with corresponding enrichment plots in Figure 2. Our analysis demonstrates up-regulated pathways related to cell cycle processes, DNA damage repair, transcriptional regulation, epigenetic remodeling, and activation of HOX genes.

The top 10 down-regulated pathways (Table 3; Fig. 3) identified in our tumor sample compared to paired normal kidney were primarily related to cellular and drug metabolism, cholesterol and electrolyte homeostasis, and biologic processes related to the kidney.

This is not unexpected with normal kidney serving as the control sample. The cell of origin of this cancer is unknown, and the normal adjacent kidney sample taken at the time of surgical resection was the best available control for this study. Many of the pathways identified as being enriched in the tumor contained overlapping lists of genes with known roles in epigenetic regulation. Multiple genes encoding key players of the PRC2, PRC4, NuRD, NCoR, and mSIN3A epigenetic remodeling complexes were found to be differentially expressed in this tumor, suggesting an extensive dysregulation of the epigenome in this patient's tumor. These results are summarized in Figure 4.

DNA Methylation

Because of the known roles of BCOR in the regulation of the epigenome and results of the transcriptomic data suggesting a dramatic dysregulation of multiple epigenetic complexes, we next focused our attention on the epigenome. DNA methylation analysis was performed using the Illumina EPIC DNA methylation array to identify differentially methylated CpG dinucleotides in tumor versus normal and to determine if this differential methylation had an impact on gene expression. Differentially methylated CpG dinucleotides were identified as those having a loss or gain of methylation >20% in tumor versus normal (corresponding to a Delta beta value of <−0.2 or >0.2, respectively). Because of the single patient sample

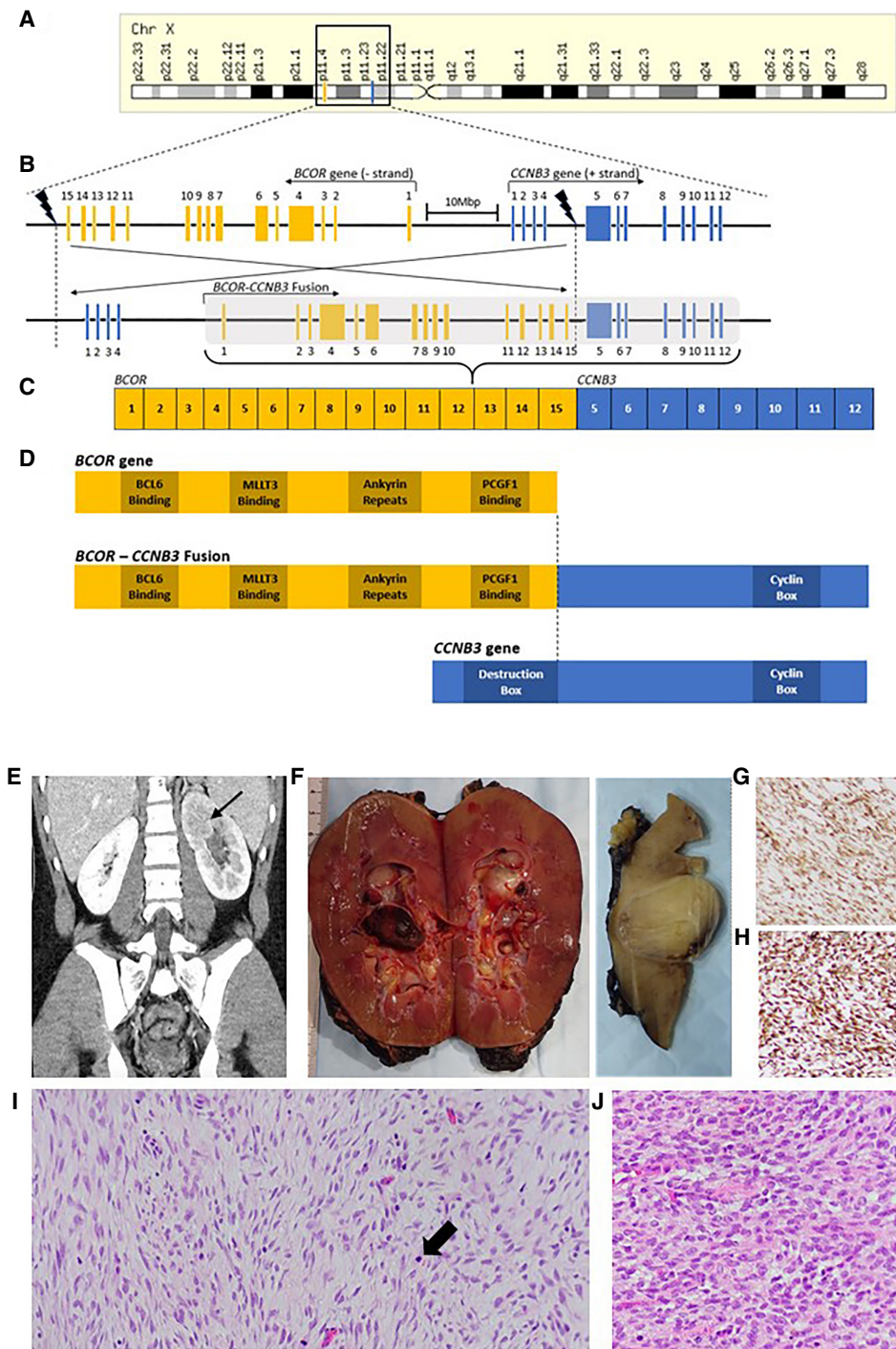


Figure 1. Clinical evaluation. (A) Genomic locations for *BCOR* (Xp11.4, Chr X:40,049,815–40,177,390) and *CCNB3* (Xp11.22, Chr X:50,202,713–50,351,914) genes on Chromosome X. (B) A schematic of the X chromosome with strand orientation indicated for the *BCOR* and *CCNB3* genes. Exonic regions for the *BCOR* and *CCNB3* genes are shown. The breakpoints are indicated by the thunderbolt with the resultant paracentric inversion. (C) Schematic depicting the resulting fusion between *BCOR* exons 1–15 and *CCNB3* exons 5–12. (D) Depiction of *BCOR* and *CCNB3* gene protein domains. (Legend continues on next page.)

Table 2. Gene set enrichment analysis: top 10 up-regulated pathways in tumor versus normal

Name	Size	ES	NES	NOM p-val	FDR q-val
REACTOME_CONDENSATION_OF_PROPHASE_C	65	0.694	2.179	0	0.0000000
REACTOME_NUCLEOSOME_ASSEMBLY	66	0.696	2.164	0	0.0009501
REACTOME_G2_M_DNA_DAMAGE_CHECKPOINT	88	0.654	2.132	0	0.0009510
REACTOME_PRC2_METHYLATES_HISTONES_A	65	0.670	2.105	0	0.0009509
REACTOME_MEIOTIC_RECOMBINATION	75	0.662	2.104	0	0.0009498
REACTOME_HDACS_DEACETYLATE_HISTONES	87	0.649	2.101	0	0.0007915
REACTOME_CHROMOSOME_MAINTENANCE	99	0.630	2.090	0	0.0009506
REACTOME_CELL_CYCLE_CHECKPOINTS	277	0.548	2.070	0	0.0008318
REACTOME_DNA_METHYLATION	56	0.690	2.065	0	0.0008468
REACTOME_CLASS_C_3_METABOTROPIC_GLU	16	0.851	2.046	0	0.0016249

(ES) Enrichment score, (NES) normalized enrichment score, (NOM) nominal p-value, (FDR) false discovery rate.

included in this analysis, it is difficult to draw statistically relevant conclusions, but we identified a total of 99,720 CpG sites as being potentially differentially methylated in this tumor. Consistent with what is frequently observed in other cancers, the majority of these sites exhibited a loss of methylation, with 55,779 sites with a Delta beta of <-0.2 and 43,824 sites showing a gain of methylation with a Delta beta of >0.2 . A closer look at this data also reveals that CpG with a loss of methylation in the tumor were more likely to be found outside of and not associated with CpG islands (Fig. 5A), whereas gains in methylation were more likely to occur within or adjacent to a CpG island (Fig. 5B). This is true independently of the location of the CpG island in proximity to known gene features (Fig. 5C), an observation that is also consistent with what is known about global methylation changes in cancer. Particularly striking in the disproportionate gain in methylation in CpG islands in the 3' UTR of genes, within gene bodies, and at exon boundaries within genes (Fig. 5C). This suggests that altered methylation may particularly affect enhancer/transcriptional regulatory regions and isoform expression. Next, we used the aggregate methylation analysis and the DMRcate (Lent et al. 2021) package to identify differentially methylated regions (DMRs; multiple CpG sites showing differential methylation that are close to each other on genomic positions). A total of 4785 potential DMRs were identified including regions spanning multiple genes with known links to cancer including *GATA3*, *PTEN*, *HIC1*, etc. (Supplemental Table 1). An example DMR associated with the *GATA3* gene is shown in Figure 5D.

To determine whether highly differentially methylated CpG dinucleotides were associated with corresponding changes in gene expression, we also compared the methylation data

Figure 1. (Continued.) (E) Computed tomography (CT) scan of pelvis with IV contrast. The left kidney demonstrates diffusely abnormal contrast uptake, which is delayed in comparison to the right side. There is moderate left hydronephrosis with dilatation of the major calyces. At the upper pole of the left kidney, there is some segmental hypoperfusion with a round enhancing structure projecting into one of the upper pole major calyces. (F) Gross pathology of a BCOR–CCNB3 sarcoma of the kidney. A cut section of the solid neoplasm with areas of hemorrhage and necrosis, bulging into the renal pelvis. Although the tumor appears well-circumscribed, the lesion infiltrates soft tissue of the renal sinus. (G) Tumor cells show positivity for bcl-2 (Bcl-2, IHC, 10× objective). (H) Strong vimentin expression is also present (vimentin, immunohistochemistry [IHC], 10× objective). (I) Hematoxylin and eosin (H&E) stain, 20× objective. Histologic findings include monomorphic spindle cell morphology focally arranged in a vague fascicular pattern intermixed with a few small round blue cells (arrow) and alternating hypocellular and hypercellular areas (not shown). (J) H&E stain, 20× objective. Histologic findings show infiltration of tumor cells.

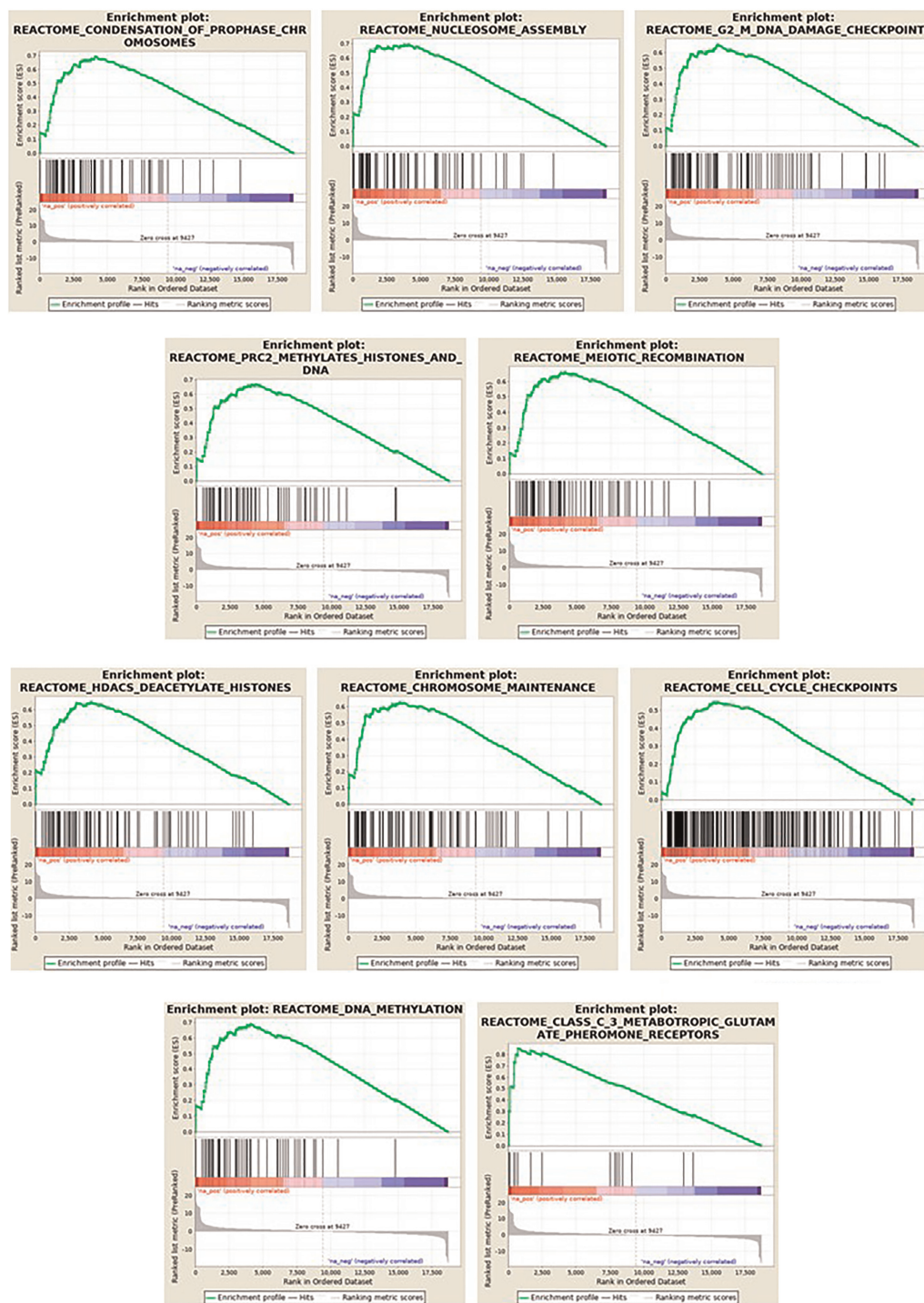


Figure 2. Gene set enrichment analysis (GSEA) shows the top 10 up-regulated pathways in tumor versus normal. Enrichment plots are shown for GSEA top 10 pathways up-regulated in *BCOR–CCNB3* sarcoma versus normal kidney control. Enrichment values can be seen in Table 2. Pathways are enriched for cell cycle processes, DNA damage repair, transcriptional regulation, epigenetic remodeling, and activation of HOX genes.

Table 3. GSEA: top 10 down-regulated pathways in tumor versus normal

Name	Size	ES	NES	NOM p-val	FDR q-val
KEGG_DRUG_METABOLISM_CYTOCHROME_P45	53	-0.615	-2.112	0.0000000	0.0257956
KEGG_RETINOL_METABOLISM	44	-0.647	-2.072	0.0000000	0.0258688
KEGG_METABOLISM_OF_XENOBIOTICS_BY_C	52	-0.620	-2.070	0.0000000	0.0175883
REACTOME_CYTOCHROME_P450_ARRANGED_B	54	-0.614	-2.056	0.0000000	0.0170372
BIOCARTA_NUCLEARRS_PATHWAY	31	-0.661	-2.033	0.0000000	0.0233554
REACTOME_BIOLOGICAL_OXIDATIONS	186	-0.486	-2.024	0.0000000	0.0219967
REACTOME_GLUCURONIDATION	16	-0.768	-2.023	0.0000000	0.0194713
REACTOME_PHASE_I_FUNCTIONALIZATION_	91	-0.540	-1.993	0.0000000	0.0269419
REACTOME_SYNTHESIS_OF_LEUKOTRIENES_	19	-0.732	-1.911	0.0000000	0.0741014
REACTOME_PEROXISOMAL_LIPID_METABOLI	28	-0.643	-1.907	0.0000000	0.0700744

(ES) Enrichment score, (NES) normalized enrichment score, (NOM) nominal p-value, (FDR) false discovery rate.

to our transcriptomic data. Each methylation array probe represents specific location related to corresponding genes. For the methylation array, a \log_2 ratio of M-values between tumor and the normal sample was estimated for each probe for which \log_2 FC expression of normalized tumor and normal sample is estimated for each gene. Methylation and expression data were matched according to gene symbols and plotted only selecting the most extreme points based on the corresponding distributions. Notice that this matching may entail a many-to-one mapping (e.g., a methylation site is related to many different genes and would be matched to the corresponding expression entry). Our analysis identified a total of 221 genes with increased expression and at least one CpG with increased methylation >20%; 39 genes with decreased expression and at least one CpG with increased methylation; 78 genes with increased expression and at least one CpG with decreased methylation; and only 20 genes with decreased expression and at least one CpG with increased methylation (data not shown). This relatively low correlation between gene expression and DNA methylation changes suggests that DNA methylation may be a relatively minor contributor to the observed gene expression changes in our *BCOR-CCNB3* tumor.

miRNA Analysis

Next, we wanted to determine whether the *BCOR-CCNB3* fusion may be altering the expression of miRNA, thus contributing to the altered transcriptional/epigenetic programming in these cells. To look at global miRNA expression levels, we performed miRNA-seq according to established protocols. Raw counts were obtained through the miRDeep2 (Friedländer et al. 2012) pipeline. and then the anamiR (Wang et al. 2019) pipeline was used to integrate miRNA and RNA data in order to detect (inversely) correlated features. We identified six miRNAs that were detected as differentially expressed. These are shown in Figure 6 (gray circles), and several of these miRNA have previously been implicated in tumorigenesis. For example, the expression of miRNA-199a-3p has been shown to suppress tumor growth, migration, invasion, and angiogenesis through negative regulation of various target genes in hepatocellular carcinoma, multiple myeloma, papillary thyroid carcinoma, and ovarian cancer (Minna et al. 2014; Raimondi et al. 2014; Deng et al. 2017; Ghosh et al. 2017). The under expression of miR-199a-3p in osteosarcoma plays an important role in the development of metastasis, recurrence, and drug resistance (Gao et al. 2015). Therefore, miR-199a-3p appears to have many important tumor suppressor roles. The miRNA-199b-3p is an isoform of miRNA-199a-3p with an identical sequence and similar functions (Dabbah et al. 2017). miR-196a-5p has been shown to promote metastasis of colorectal cancer (Xin et al. 2019). One study found

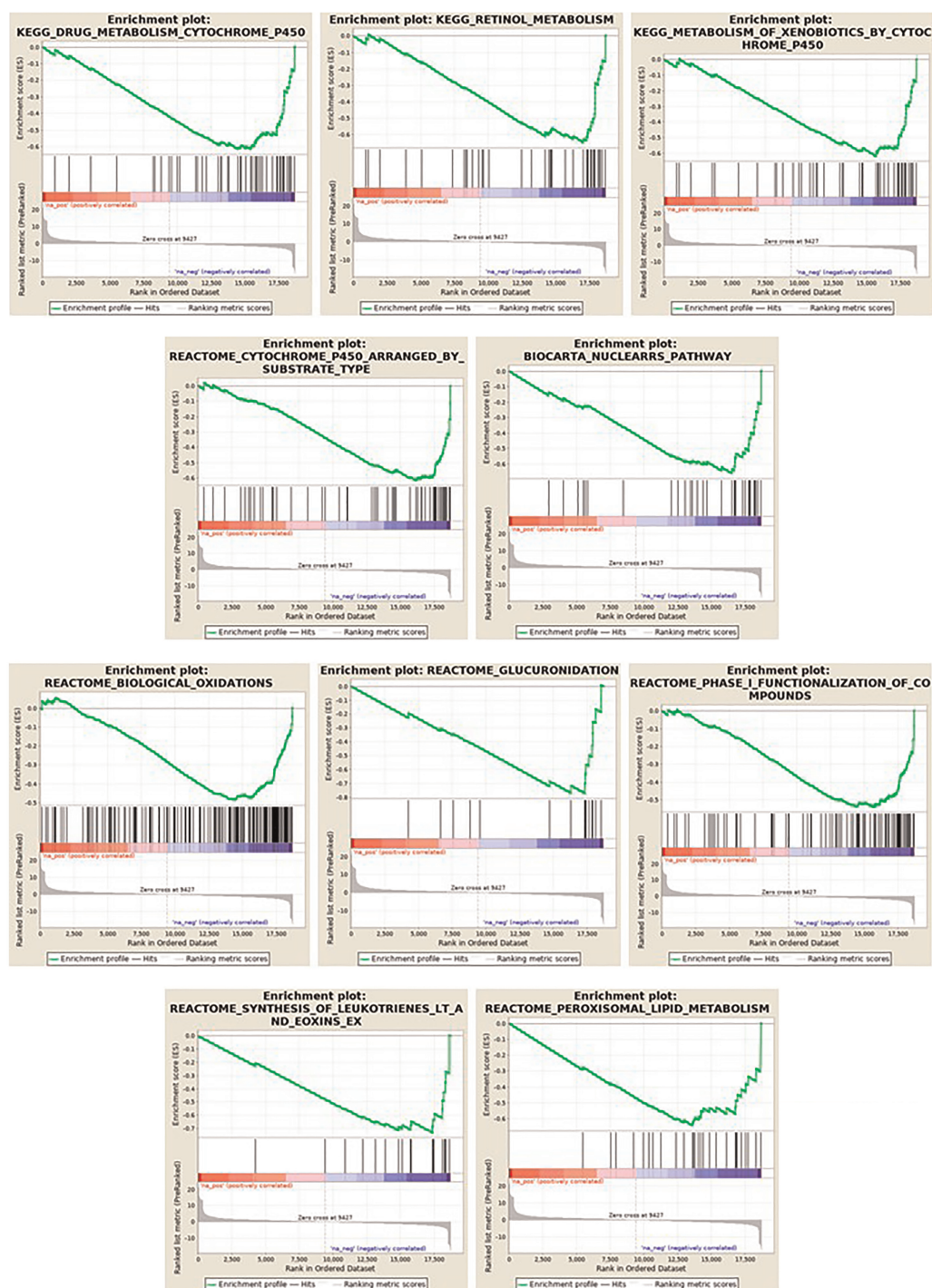


Figure 3. Gene set enrichment analysis (GSEA) shows the top 10 down-regulated pathways in tumor versus normal. Enrichment plots are shown for GSEA top 10 pathways down-regulated in the *BCOR–CCNB3* sarcoma versus normal kidney control. Enrichment values can be seen in Table 3. Pathways are primarily focused on cellular and drug metabolism, cholesterol and electrolyte homeostasis, and biologic processes related to normal kidney function.

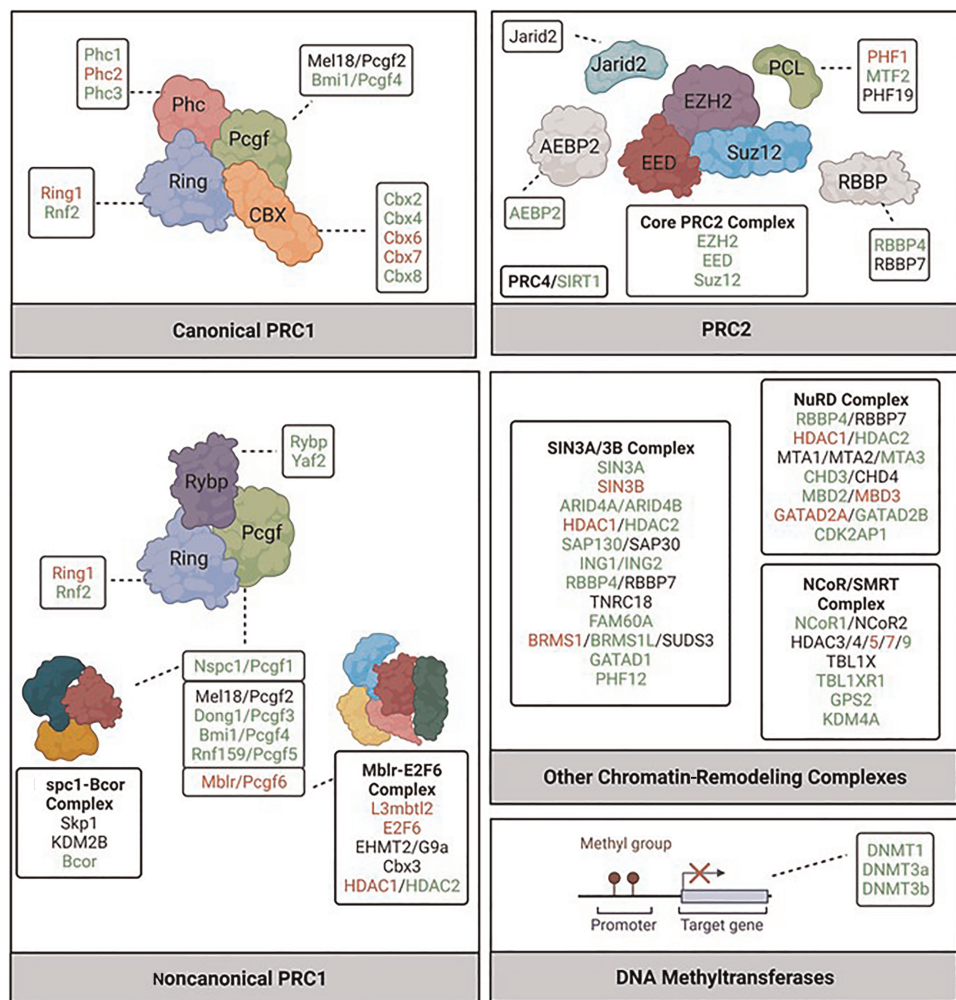


Figure 4. Differential gene expression in epigenetic remodeling complexes. Messenger RNA sequencing (mRNA-seq) identified numerous epigenetic regulatory genes that are differentially expressed in the *BCOR–CCNB3* tumor versus normal. Genes associated with the canonical and noncanonical Polycomb repressive complex 1 (PRC1) and PRC2, SIN3A/3B, NuRD, and NCoR/SMRT chromatin-remodeling complexes and DNA methyltransferases are shown. Genes with greater than twofold up-regulation are shown in green text. Genes with greater than twofold down-regulation in tumor versus normal are shown in red text. Genes that are not differentially expressed are shown in black text.

that miR-23a-3p is significantly up-regulated in renal cell carcinoma tissue samples and that low expression of miR-23a-3p in patients with renal cell carcinoma was associated with longer overall survival (Quan et al. 2019). miR-204-5p is known to suppress the proliferation, migration, and invasion of endometrial carcinoma cells (Bao et al. 2013) and colorectal carcinomas cells (Yin et al. 2014). To help elucidate the role that differential miRNA expression may play in regulating gene expression, we also performed an integrated miRNA/mRNA analysis. Genes and miRNAs with small read counts were considered as unexpressed and removed from the analysis. Posterior fold-change values were estimated on miRNA and mRNA using EBseq. Two thousand four hundred and ninety-seven genes identified as differentially expressed with cutoff values of false discovery rate (FDR) < 0.05 and absolute fold change > 2 were compared with the six miRNAs detected as significantly differentially expressed. We then used

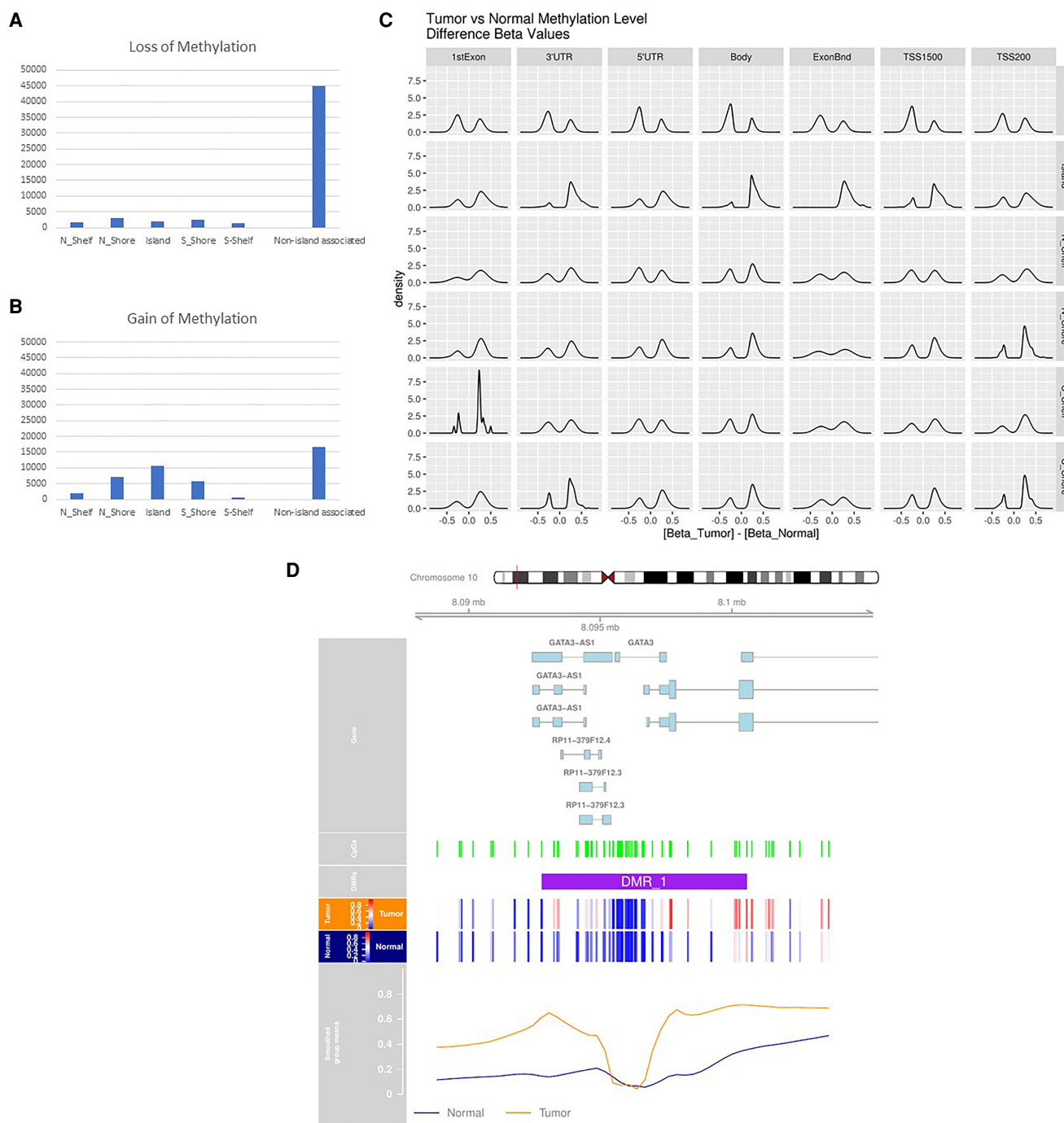


Figure 5. Differential methylation analysis. Illumina EPIC arrays were used to identify differentially methylated CpG loci in tumor versus normal. (A) Total number of CpG dinucleotides shows a >20% ($\Delta\beta > 0.2$) loss of methylation associated with CpG islands. (B) Total number of CpG dinucleotides showing a >20% ($\Delta\beta > 0.2$) gain of methylation in relation to position with regard to CpG islands. (C) Density of gain and loss of methylation in relation to gene features and CpG island features. (D) Differentially methylated region associated with the GATA3 gene is shown. CpG dinucleotides are shown in green. Methylation from 0.0 (blue bars) to 1.0 (red bars) shown for individual CpG sites (middle panel) and across the region (lower panel) b-value for tumor (orange) and normal (blue) are shown.

Detected miRNA–mRNA associations
miRNAs: 6; genes: 86

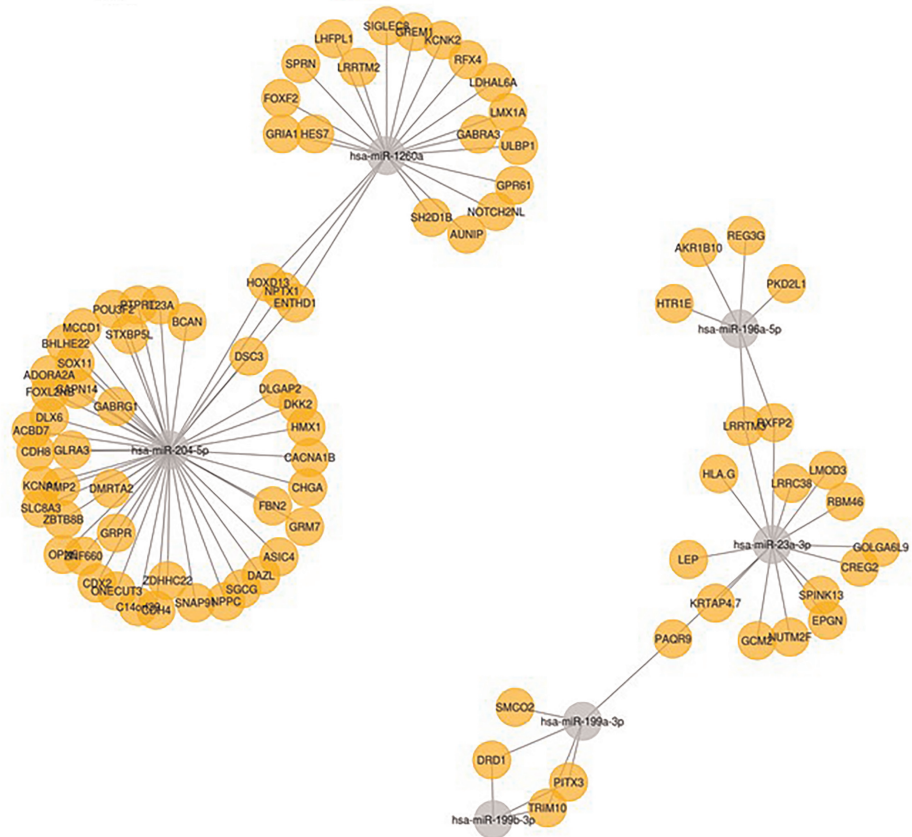


Figure 6. Detected micro RNA (miRNA)–messenger RNA (mRNA) associations. Plot depicts detected miRNA–mRNA associations for six miRNA (gray circles) identified by miRNA-seq and are differentially expressed in tumor versus normal and 86 negatively correlated, differentially expressed genes (orange circles) by mRNA-seq, which are known targets of these miRNA.

anamiR (Wang et al. 2019) (v1.13.0), an integrated analysis package of miRNA and mRNA expression data, to find the negatively correlated points and also to query these into 10 different known databases (such as miRDB, miRanda, etc.). We identified 86 genes that are known targets and were inversely correlated with the differentially expressed miRNA, further suggesting that these miRNAs are altering downstream target gene expression in this tumor. A summary of these results is shown in Figure 6.

DISCUSSION

BCOR–CCNB3 fusion sarcomas are a rare and newly described subtype of soft-tissue sarcomas, primarily presenting in young male patients. Because of the rarity of this disease, little is known about the molecular characteristics of this tumor. In this study, WES and RNA-seq were combined with DNA methylation and miRNA analysis and performed on our patient sample with the goal of increasing our understanding the genomic and epigenomic landscape of this disease. Because of the paucity of other mutations in this tumor, it is probable that our patient’s tumor is primarily driven by the in-frame fusion between *BCOR* and

CCNB3, as expected, which is identical to those described in other case reports (Pierron et al. 2012; Peters et al. 2015b).

The combined data suggest few somatic mutations and a high frequency of deregulated gene expression involving both chromatin-modifying enzymes and *HOX* genes. Epigenetic alterations due to the up-regulation of *HOX* transcription factors and proteins involved in epigenetic regulation appear to lead to the dysregulation of genes encoding several developmental programs and stem cell differentiation, which is similarly seen across a spectrum of pediatric sarcomas (Lawlor and Thiele 2012). This is not surprising given that *BCOR* is known to epigenetically regulate the transcription of specific genes, especially those of mesenchymal stem cell function. The *Bcl6* co-repressor, or *BCOR* gene, encodes a transcriptional co-repressor that binds the *Bcl6* oncoprotein as part of a large Polycomb repressive complex, the PRC1.1 complex that functions to repress transcriptional start sites. Interestingly, none of the common components of PRC1 were differentially expressed in our tumor compared to normal (Fig. 4). Of note, exon 15 of the *BCOR* gene is implicated in the development of sarcoma across a spectrum of *BCOR*-related diseases including BCS, CCSK, and PMMTI (Kao et al. 2018; Santiago et al. 2018). *BCOR* internal tandem duplications (ITDs) involving exon 15 of *BCOR* are well-known genetic alterations seen in CCSK and PMMTI (Roy et al. 2015). Exon 15 encodes the PCGF ubiquitin-like fold discriminator (PUFD) domain, a region that binds selectively to another domain, the ubiquitin-like RAWUL domain. This interactive binding between PUFD and ubiquitin-like RAWUL is crucial for assembly of the Polycomb group RING finger complex, a distinct Polycomb repressive complex necessary to mediate epigenetic modifications of histones (Junco et al. 2013; Kao et al. 2018). Disruption of the PRC1.1/*BCOR* complex and up-regulation of PRC2 targets has been described in ITD-driven CCSK and PMMTI, leading to overall Polycomb complex dysregulation. BCS tumors similarly have abnormalities involving the PUFD domain of the *BCOR* gene increasing our understanding of the similarities between these *BCOR* mutant tumors (Roy et al. 2015).

Numerous epigenetically regulated genes and pathways are identified in our analysis, which may serve as the foundation for future studies focused on *BCOR*-associated malignancies, but only a small percentage of these DNA methylation changes is associated with corresponding changes in gene expression (data not shown), suggesting that abnormal DNA methylation is likely a consequence, not a leading cause of tumorigenesis in this patient. We did observe multiple changes in miRNA expression with known links to tumorigenesis. The clinical significance of the key miRNAs related to *BCOR–CCNB3* fusion sarcomas is still unclear, but this raises important questions with regard to the role of miRNA in *BCOR–CCNB3* fusion-driven sarcomas. In the absence of additional mutations in this tumor, it appears that the *BCOR–CCNB3* fusion drives a global epigenetic reprogramming event that is likely to be the primary driver of transformation and tumorigenesis.

As we continue to increase our understanding of the *BCOR* effects on fusion-positive undifferentiated sarcomas, better-tailored and more refined treatment algorithms will continue to evolve. Targeted inhibition of the *BCOR–CCNB3* fusion transcript may be an efficacious treatment strategy for this disease, possibly in combination with epigenetic modifying agents. Our molecular analysis suggests the possibility that a trial of epigenetic modifying agents in patients with BCS may restore a functional epigenome and, in the absence of a high number of secondary mutations in these patients, may prove an efficacious therapeutic strategy, particularly where epigenome targeted therapeutic agents are already being tested in the clinic (i.e., HDAC inhibitors and EZH2 inhibitors). We recognize this study includes only one BCS patient analysis and a further in-depth analysis of additional tumor samples at the molecular level may lend support to this approach. Furthermore, *in vitro* and *in vivo* modeling of the fusion will be necessary to increase our understanding of how the *BCOR–CCNB3* fusion may promote epigenetic remodeling and whether targeted therapies may be efficacious for patients with BCS tumors.

Future Research Plans

A cellular model will facilitate the search for a better understanding of BCSs, their driver mutations, cellular effects, and epigenetic alterations. As BCSs appear to be fusion-driven sarcomas, a search for targeted and refined treatments is important. With additional functional studies of a BCS-driven cellular model, we may be able to (1) clarify the role of BCOR epigenetic dysregulation, (2) test/develop novel therapeutics (epigenetic modifiers vs. oncogene-targeted therapies), and (3) identify a distinct epigenetic signature crafted by BCSs and establish a reference methylation standard to identify tumors in cases that are diagnostically challenging.

METHODS

Tissue Source and Processing

Acquisition of specimen from a 14-yr-old pediatric male patient who was diagnosed with a BCS of the kidney was done following IRB approval and patient consent. FFPE samples of the tumor with adjacent paired normal (germline) from the patient's kidney specimen were obtained from the Department of Pathology at Oishei Children's Hospital. The tumor sample was dissected from an FFPE block to a depth of 5 μm and fixed on slides. Paired normal was provided as shaved curls from an FFPE¹³ block to a depth of 20 μm . Our tumor and matched normal specimens were sent to the Genomics Shared Resource of Roswell Park Comprehensive Cancer Center for DNA and RNA extraction, WES, RNA-seq, miRNA-seq, and EPIC arrays analyses. Briefly, WES libraries were prepared using Agilent, SureSelect^{XTHS} Human All Exon V7 capture library and sequenced on Illumina NovaSeq using 100PE sequencing. For RNA-seq, libraries were prepared using SureSelect XT RNA Direct kit and sequenced on Illumina NovaSeq using 100PE sequencing. The small RNA libraries were prepared using SMARTer smRNA-Seq kit and sequenced on Illumina NovaSeq for single-end sequencing.

Clinical Case Data Review

The patient presentation and follow-up information are obtained from review of the electronic health records and direct conversation with the treating physician and pathologist.

Genomic Analysis

Detection of Somatic Mutations

Exome sequencing reads were aligned to the human reference genome (GRCh37 b37) using BWA-MEM (v0.7.12) (Li and Durbin 2009). Sequencing coverage is shown in Supplemental Figure 1. PCR duplicates were marked by Picard Mark Duplicates (v2.18.2) and base quality recalibration was performed by GATK (v4.1.2.0) (McKenna et al. 2010). The somatic variants were called using a multiple-caller ensembled method, including MuTect2 (v4.1.2.0) (Cibulskis et al. 2013) for sensitive detection of somatic point mutations in impure and heterogeneous cancer samples, MuSE (v1.0) (Fan et al. 2016), which accounts for intertumor heterogeneity using a sample-specific error model and improves sensitivity and specificity in mutation calling for sequencing data, Strelka2 (v2.9.10) (Saunders et al. 2012), and VarDict (v1.8) (Lai et al. 2016) and manually reviewed to remove false-positive calls. The analysis pipelines are reproducible and publicly available through the RcwI package (Hu et al. 2021). Putative mutations are identified by running NeuSomatic (Sahraeian et al. 2019) with default parameters. All putative mutations are manually inspected to filter out potential false-positives introduced by sequencing and mapping artifacts. The identified somatic mutations are compared to the public human germline databases including dbSNP (Sherry et al.

2001), 1000 Genomes Project (Genomes Project Consortium et al. 2012), and the National Heart, Lung, and Blood Institute's Exome Sequencing Project (NHLBI 2013) to further exclude remaining germline polymorphisms. All mutations are annotated using ANNOVAR (Wang et al. 2010) with the National Center for Biotechnology Information (NCBI) RefSeq database.

RNA-seq Data Analysis

Paired-end reads are generated by Illumina sequencer and QCed using fastqc (Andrews 2010) (v0.10.1) and mapped to the GRCh38 human reference genome and ENSEMBLE (v25) annotation database using STAR [2] (v2.7.0f). Aligned reads are further checked with RSeQC (Wang et al. 2012) (v2.6.3) in order to identify potential library preparation-related problems. Gene level expression is estimated with featureCounts from Subread (Liao et al. 2014) (v1.6.0) using the fracOverlap 1 option. FC estimation and differential expression analyses for the two samples are carried out using EBSeq (Leng et al. 2013) (v1.18.1), which utilizes the empirical Bayesian model and can handle data without replicates. Heatmaps are generated using pheatmap (Kolde and Kolde 2015) (v1.0.8) R library using data generated by DESeq2's (Love et al. 2014) regularized-log transformation. The fusion results are generated from STAR-Fusion (Haas et al. 2019) program from STAR package.

miRNA-seq

Raw single-end reads are examined using fastqc to check for low-quality within-read position trends. Reads are filtered by removing the first 16 bp and last 23 bp to keep the miRNA sequence only. Resulting reads are then aligned and quantified using the miRDeep2 (Friedländer et al. 2012) (v2.0.1.2) pipeline. Briefly, miRDeep2 removes reads smaller than 18 bp, collapses duplicated reads, and maps them against the reference genome (GRCh38) using Bowtie (Langmead 2010) (v1.2.3). Mapped reads are then identified and quantified to known miRNAs (mature and hairpin) accounting for secondary structures.

Methylation Array

EPIC array (Illumina Human Methylation 850k) files are read, processed, filtered, and normalized using minfi (Aryee et al. 2014) (v1.32) R package. M and beta values are then extracted for further analysis. Delta beta values (tumor-normal beta values) are estimated, removing observations with absolute values of <0.2 to identify candidate methylation sites with different methylation levels between tumor and normal samples. Dmrcate (Peters et al. 2015a) Bioconductor package is used to identify differentially methylated regions, and sites are annotated using both the UCSC refGene and CpG island databases to identify the relative position to each gene and CpG islands.

Integrated Analysis

mRNA and miRNA integration: Genes and miRNAs with small read count are considered as unexpressed and removed from the analysis. Posterior FC values were estimated on miRNA and mRNA using EBseq. Two thousand four hundred and ninety-seven genes are differentially expressed with cutoff values of FDR < 0.05 and absolute fold change > 2 and six miRNAs are detected as significant according to FDR < 0.25 (hsa-miR-199a-3p, hsa-miR-199b-3p, hsa-miR-196a-5p, hsa-miR-23a-3p, hsa-miR-204-5p, hsa-miR-1260a). Then anamiR (Wang et al. 2019) (v1.13.0), an integrated analysis package of miRNA and mRNA expression data, was used to find the negatively correlated points and also to query these into 10 different known databases (such as miRDB, miRanda, etc.).

mRNA and methylation integration: mRNAs are filtered and processed as described in the miRNA integration analysis, but with a log₂FC cutoff value of 6. The log₂ ratio of M-values

forms tumor and normal samples are calculated and methylation sites with absolute $\log_2 > 3$ are kept as candidate differential methylation sites. A final list is obtained matching overlapped genes found in the methylation filtered data annotation with the filtered mRNA data.

ADDITIONAL INFORMATION

Data Deposition and Access

Sequencing data within this paper are deposited in the GEO (<https://www.ncbi.nlm.nih.gov/geo/>) repository and are now publicly available. The reference series GEO accession number is GSE188555. The data set accession numbers are GSE188552, GSE188553, and GSE188554.

Ethics Statement

The research project used tissue specimens that had been collected as part of an ongoing biorepository. The biorepositories obtain written informed consent from participants and/or their legal guardians for the collection and research use of the specimens. All aspects of the study were approved by the Roswell Park Comprehensive Cancer Center Institutional Review Board.

Acknowledgments

The work herein was supported by the laboratory of Joyce E. Ohm, PhD. We also acknowledge the members of the Ohm lab—Joyce E. Ohm, Kristen M. Humphrey, Lyn Gao, and Jeffrey Martin—for their continued support and mentorship throughout this project. The tissue donation provided by the patient in order to conduct this research is also greatly appreciated. We thank our colleagues at Oishei Children’s Hospital, Victoria Frano and Dr. Rafal Kozielski, MD, who provided additional pathologic insight for this case and allowed for the acquisition of our patient’s tumor. We thank Dr. Elizabeth Perlman, MD, and Dr. Mariana Cajaiba, MD, at the Division of Pathology at Lurie Children’s Hospital in consultation with the Children’s Oncology Group Renal Tumor Pathology Center.

Author Contributions

All of the molecular experiments presented here were performed by T.J.H. under the guidance of J.E.O. and with the assistance of J.C.M. and L.G. Bioinformatics and biostatistical analysis of all data were performed by the team of E.C.G., M.L., L.W., S.L., and J.W. P.K.S. led all of the multiomics analysis in the Genomics Shared Resource. T.J.H., A.G., C.J.T., R.K., and J.K. contributed clinical and pathological input. All authors have contributed to writing and/or editing components of this manuscript.

Competing Interest Statement

The authors have declared no competing interest.

Received August 17, 2021;
accepted in revised form
November 10, 2021.

Funding

This work was funded by a grant from the Roswell Park Alliance Foundation to J.E.O. This work was partly supported by National Cancer Institute (NCI) grant P30CA016056 involving the use of Roswell Park Comprehensive Cancer Center’s Genomic Shared Resource and Bioinformatics Shared Resource.

REFERENCES

Andrews S. 2010. *FastQC: a quality control tool for high throughput sequence data*. Babraham Bioinformatics, Babraham Institute, Cambridge.

- Argani P, Kao YC, Zhang L, Bacchi C, Matoso A, Alaggio R, Epstein JI, Antonescu CR. 2017. Primary renal sarcomas with BCOR-CCNB3 gene fusion: a report of 2 cases showing histologic overlap with clear cell sarcoma of kidney, suggesting further link between BCOR-related sarcomas of the kidney and soft tissues. *Am J Surg Pathol* **41**: 1702–1712. doi:10.1097/PAS.0000000000000926
- Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA. 2014. minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**: 1363–1369. doi:10.1093/bioinformatics/btu049
- Astolfi A, Fiore M, Melchionda F, Indio V, Bertuccio SN, Pession A. 2019. BCOR involvement in cancer. *Epigenomics* **11**: 835–855.
- Bao W, Wang HH, Tian FJ, He XY, Qiu MT, Wang JY, Zhang HJ, Wang LH, Wan XP. 2013. A TrkB-STAT3-miR-204-5p regulatory circuitry controls proliferation and invasion of endometrial carcinoma cells. *Mol Cancer* **12**: 155. doi:10.1186/1476-4598-12-155
- Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, Getz G. 2013. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* **31**: 213–219.
- Cohen-Gogo S, Cellier C, Coindre JM, Mosseri V, Pierron G, Guillemet C, Italiano A, Brugieres L, Orbach D, Laurence V, et al. 2014. Ewing-like sarcomas with BCOR-CCNB3 fusion transcript: a clinical, radiological and pathological retrospective study from the Societe Francaise des Cancers de L’Enfant. *Pediatr Blood Cancer* **61**: 2191–2198. doi:10.1002/pbc.25210
- Dabbah M, Attar-Schneider O, Zismanov V, Tartakover Matalon S, Lishner M, Drucker L. 2017. Letter to the Editor: miR-199b-3p and miR-199a-3p are isoforms with identical sequence and established function as tumor and metastasis suppressors. *J Leukoc Biol* **101**: 1069. doi:10.1189/jlb.3LT0117-038
- Deng Y, Zhao F, Hui L, Li X, Zhang D, Lin W, Chen Z, Ning Y. 2017. Suppressing miR-199a-3p by promoter methylation contributes to tumor aggressiveness and cisplatin resistance of ovarian cancer through promoting DDR1 expression. *J Ovarian Res* **10**: 50. doi:10.1186/s13048-017-0333-4
- Dubey M, Brouwers E, Hamilton EMC, Stiedl O, Bugiani M, Koch H, Kole MHP, Boschert U, Wykes RC, Mansvelter HD, et al. 2018. Seizures and disturbed brain potassium dynamics in the leukodystrophy megalencephalic leukoencephalopathy with subcortical cysts. *Ann Neurol* **83**: 636–649. doi:10.1002/ana.25190
- Fan Y, Xi L, Hughes DS, Zhang J, Zhang J, Futreal PA, Wheeler DA, Wang W. 2016. MuSE: accounting for tumor heterogeneity using a sample-specific error model improves sensitivity and specificity in mutation calling from sequencing data. *Genome Biol* **17**: 178.
- Friedländer MR, Mackowiak SD, Li N, Chen W, Rajewsky N. 2012. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucl Acids Res* **40**: 37–52. doi:10.1093/nar/gkr688
- Gao Y, Feng Y, Shen JK, Lin M, Choy E, Cote GM, Harmon DC, Mankin HJ, Hornicek FJ, Duan Z. 2015. CD44 is a direct target of miR-199a-3p and contributes to aggressive progression in osteosarcoma. *Sci Rep* **5**: 11365. doi:10.1038/srep11365
- Genomes Project Consortium, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA. 2012. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**: 56–65. doi:10.1038/nature11632
- Ghosh A, Dasgupta D, Ghosh A, Roychoudhury S, Kumar D, Gorain M, Butti R, Datta S, Agarwal S, Gupta S, et al. 2017. MiRNA199a-3p suppresses tumor growth, migration, invasion and angiogenesis in hepatocellular carcinoma by targeting VEGFA, VEGFR1, VEGFR2, HGF and MMP2. *Cell Death Dis* **8**: e2706. doi:10.1038/cddis.2017.123
- Haas BJ, Dobin A, Li B, Stransky N, Pochet N, Regev A. 2019. Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. *Genome Biol* **20**: 213. doi:10.1186/s13059-019-1842-9
- Hu Q, Hutson A, Liu S, Morgan M, Liu Q. 2021. Bioconductor toolchain for reproducible bioinformatics pipelines using Rowl and RowLPipelines. *Bioinformatics* **37**: 3351–3352. doi:10.1093/bioinformatics/btab208
- Junco SE, Wang R, Gaipa JC, Taylor AB, Schirf V, Gearhart MD, Bardwell VJ, Demeler B, Hart PJ, Kim CA. 2013. Structure of the Polycomb group protein PCGF1 in complex with BCOR reveals basis for binding selectivity of PCGF homologs. *Structure* **21**: 665–671.
- Kao YC, Sung YS, Zhang L, Jungbluth AA, Huang SC, Argani P, Agaram NP, Zin A, Alaggio R, Antonescu CR. 2016. BCOR overexpression is a highly sensitive marker in round cell sarcomas with BCOR genetic abnormalities. *Am J Surg Pathol* **40**: 1670–1678. doi:10.1097/PAS.0000000000000697
- Kao YC, Owosho AA, Sung YS, Zhang L, Fujisawa Y, Lee JC, Wexler L, Argani P, Swanson D, Dickson BC, et al. 2018. BCOR-CCNB3 fusion positive sarcomas: a clinicopathologic and molecular analysis of 36 cases with comparison to morphologic spectrum and clinical behavior of other round cell sarcomas. *Am J Surg Pathol* **42**: 604–615. doi:10.1097/PAS.0000000000000965

- Koelsche C, Hartmann W, Schimpf D, Stichel D, Jabar S, Ranft A, Reuss DE, Sahm F, Jones DTW, Bewerunge-Hudler M, et al. 2018. Array-based DNA-methylation profiling in sarcomas with small blue round cell histology provides valuable diagnostic information. *Mod Pathol* **31**: 1246–1256. doi:10.1038/s41379-018-0045-3
- Kolde R, Kolde MR. 2015. Package ‘pheatmap’. *R Package* **1**: 790.
- Laetsch TW, Roy A, Xu L, Black JO, Coffin CM, Chi YY, Tian J, Spunt SL, Hawkins DS, Bridge JA, et al. 2018. Undifferentiated sarcomas in children harbor clinically relevant oncogenic fusions and gene copy-number alterations: a report from the children’s oncology group. *Clin Cancer Res* **24**: 3888–3897. doi:10.1158/1078-0432.CCR-18-0672
- Lai Z, Markovets A, Ahdesmaki M, Chapman B, Hofmann O, McEwen R, Johnson J, Dougherty B, Barrett JC, Dry JR. 2016. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. *Nucl Acids Res* **44**: e108.
- Langmead B. 2010. Aligning short sequencing reads with Bowtie. *Curr Protoc Bioinformatics* **32**: 11.17. 11–11.17. 14. doi:10.1002/0471250953.bi1107s32
- Lawlor ER, Thiele CJ. 2012. Epigenetic changes in pediatric solid tumors: promising new targets. *Clin Cancer Res* **18**: 2768–2779. doi:10.1158/1078-0432.CCR-11-1921
- Lee YC, Su CY, Lin YF, Lin CM, Fang CY, Lin YK, Hsiao M, Chen CL. 2017. Lysosomal acid phosphatase 2 is an unfavorable prognostic factor but is associated with better survival in stage II colorectal cancer patients receiving chemotherapy. *Oncotarget* **8**: 12120–12132. doi:10.18632/oncotarget.14552
- Le Loarer F, Pissaloux D, Coindre JM, Tirode F, Vince DR. 2017. Update on families of round cell sarcomas other than classical Ewing sarcomas. *Surg Pathol Clin* **10**: 587–620. doi:10.1016/j.path.2017.04.002
- Leng N, Dawson JA, Thomson JA, Ruotti V, Rissman AI, Smits BMG, Haag JD, Gould MN, Stewart RM, Kendzioriski C. 2013. EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* **29**: 1035–1043. doi:10.1093/bioinformatics/btt087
- Lent S, Cardenas A, Rifas-Shiman SL, Perron P, Bouchard L, Liu CT, Hivert MF, Dupuis J. 2021. Detecting differentially methylated regions with multiple distinct associations. *Epigenomics* **13**: 451–464. doi:10.2217/epi-2020-0344
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**: 1754–1760. doi:10.1093/bioinformatics/btp324
- Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923–930. doi:10.1093/bioinformatics/btt656
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550. doi:10.1186/s13059-014-0550-8
- Machado I, Navarro S, Lombart-Bosch A. 2016. Ewing sarcoma and the new emerging Ewing-like sarcomas: (CIC and BCOR-rearranged-sarcomas). A systematic review. *Histol Histopathol* **31**: 1169–1181.
- Matsuyama A, Shiba E, Umekita Y, Nosaka K, Kamio T, Yanai H, Miyasaka C, Watanabe R, Ito I, Tamaki T, et al. 2017. Clinicopathologic diversity of undifferentiated sarcoma with BCOR–CCNB3 fusion: analysis of 11 cases with a reappraisal of the utility of immunohistochemistry for BCOR and CCNB3. *Am J Surg Pathol* **41**: 1713–1721. doi:10.1097/PAS.0000000000000934
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**: 1297–1303
- Mertens F, Antonescu CR, Mitelman F. 2016. Gene fusions in soft tissue tumors: recurrent and overlapping pathogenetic themes. *Genes Chromosomes Cancer* **55**: 291–310. doi:10.1002/gcc.22335
- Minna E, Romeo P, De Cecco L, Dugo M, Cassinelli G, Pilotti S, Degl’Innocenti D, Lanzi C, Casalini P, Pierotti MA, et al. 2014. miR-199a-3p displays tumor suppressor functions in papillary thyroid carcinoma. *Oncotarget* **5**: 2513–2528. doi:10.18632/oncotarget.1830
- Mowry EM, Carey RF, Blasco MR, Pelletier J, Duquette P, Villoslada P, Malikova I, Roger E, Kinkel RP, McDonald J, et al. 2013. Multiple sclerosis susceptibility genes: associations with relapse severity and recovery. *PLoS ONE* **8**: e75416. doi:10.1371/journal.pone.0075416
- NHLBI. 2013. Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (<http://evs.gs.washington.edu/EVS/>) [June 04, 2013 accessed].
- Peters TJ, Buckley MJ, Statham AL, Pidsley R, Samaras K RVL, Clark SJ, Molloy PL. 2015a. De novo identification of differentially methylated regions in the human genome. *Epigenetics Chromatin* **8**: 6. doi:10.1186/1756-8935-8-6
- Peters TL, Kumar V, Polikepahad S, Lin FY, Sarabia SF, Liang Y, Wang WL, Lazar AJ, Doddapaneni H, Chao H, et al. 2015b. BCOR–CCNB3 fusions are frequent in undifferentiated sarcomas of male children. *Mod Pathol* **28**: 575–586. doi:10.1038/modpathol.2014.139
- Pierron G, Tirode F, Lucchesi C, Reynaud S, Ballet S, Cohen-Gogo S, Perrin V, Coindre JM, Delattre O. 2012. A new subtype of bone sarcoma defined by BCOR–CCNB3 gene fusion. *Nat Genet* **44**: 461–466. doi:10.1038/ng.1107

- Puls F, Niblett A, Marland G, Gaston CL, Douis H, Mangham DC, Sumathi VP, Kindblom LG. 2014. BCOR–CCNB3 (Ewing-like) sarcoma: a clinicopathologic analysis of 10 cases, in comparison with conventional Ewing sarcoma. *Am J Surg Pathol* **38**: 1307–1318. doi:10.1097/PAS.0000000000000223
- Quan J, Pan X, Li Y, Hu Y, Tao L, Li Z, Zhao L, Wang J, Li H, Lai Y, et al. 2019. MiR-23a-3p acts as an oncogene and potential prognostic biomarker by targeting PNRC2 in RCC. *Biomed Pharmacother* **110**: 656–666. doi:10.1016/j.biopha.2018.11.065
- Raimondi L, Amodio N, Di Martino MT, Altomare E, Leotta M, Caracciolo D, Gullà A, Neri A, Taverna S, D'Aquila P, et al. 2014. Targeting of multiple myeloma–related angiogenesis by miR-199a-5p mimics: in vitro and in vivo anti-tumor activity. *Oncotarget* **5**: 3039–3054. doi:10.18632/oncotarget.1747
- Renzi S, Anderson ND, Light N, Gupta A. 2019. Ewing-like sarcoma: an emerging family of round cell sarcomas. *J Cell Physiol* **234**: 7999–8007. doi:10.1002/jcp.27558
- Roy A, Kumar V, Zorman B, Fang E, Haines KM, Doddapaneni H, Hampton OA, White S, Bavle AA, Patel NR, et al. 2015. Recurrent internal tandem duplications of BCOR in clear cell sarcoma of the kidney. *Nat Commun* **6**: 8891.
- Sahraeian SME, Liu R, Lau B, Podesta K, Mohiyuddin M, Lam HYK. 2019. Deep convolutional neural networks for accurate somatic mutation detection. *Nat Commun* **10**: 1041. doi:10.1038/s41467-019-09027-x
- Santiago T, Clay MR, Allen SJ, Orr BA. 2018. Recurrent BCOR internal tandem duplication and BCOR or BCL6 expression distinguish primitive myxoid mesenchymal tumor of infancy from congenital infantile fibrosarcoma. *Mod Pathol* **30**: 884–891.
- Saunders CT, Wong WS, Swamy S, Becq J, Murray LJ, Cheetham RK. 2012. Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics* **28**: 1811–1817.
- Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. 2001. dbSNP: the NCBI database of genetic variation. *Nucl Acids Res* **29**: 308–311. doi:10.1093/nar/29.1.308
- Wang K, Li M, Hakonarson H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucl Acids Res* **38**: e164. doi:10.1093/nar/gkq603
- Wang L, Wang S, Li W. 2012. RSeQC: quality control of RNA-seq experiments. *Bioinformatics* **28**: 2184–2185. doi:10.1093/bioinformatics/bts356
- Wang T-T, Lee C-Y, Lai L-C, Tsai M-H, Lu T-P, Chuang EY. 2019. anamiR: integrated analysis of microRNA and gene expression profiling. *BMC Bioinformatics* **20**: 239. doi:10.1186/s12859-019-2870-x
- Xin H, Wang C, Liu Z. 2019. miR-196a-5p promotes metastasis of colorectal cancer via targeting IκBα. *BMC Cancer* **19**: 30. doi:10.1186/s12885-018-5245-1
- Yin Y, Zhang B, Wang W, Fei B, Quan C, Zhang J, Song M, Bian Z, Wang Q, Ni S, et al. 2014. miR-204-5p inhibits proliferation and invasion and enhances chemotherapeutic sensitivity of colorectal cancer cells by downregulating RAB22A. *Clin Cancer Res* **20**: 6187–6199. doi:10.1158/1078-0432.CCR-14-1030