



**UNIL** | Université de Lausanne

Unicentre

CH-1015 Lausanne

<http://serval.unil.ch>

---

*Year : 2021*

## Identification of Epigenetic Drivers in DSRC Tumors

Broye Liliane Celeste

Broye Liliane Celeste, 2021, Identification of Epigenetic Drivers in DSRC Tumors

Originally published at : Thesis, University of Lausanne

Posted at the University of Lausanne Open Archive <http://serval.unil.ch>

Document URN : urn:nbn:ch:serval-BIB\_242B011989E75

### **Droits d'auteur**

L'Université de Lausanne attire expressément l'attention des utilisateurs sur le fait que tous les documents publiés dans l'Archive SERVAL sont protégés par le droit d'auteur, conformément à la loi fédérale sur le droit d'auteur et les droits voisins (LDA). A ce titre, il est indispensable d'obtenir le consentement préalable de l'auteur et/ou de l'éditeur avant toute utilisation d'une oeuvre ou d'une partie d'une oeuvre ne relevant pas d'une utilisation à des fins personnelles au sens de la LDA (art. 19, al. 1 lettre a). A défaut, tout contrevenant s'expose aux sanctions prévues par cette loi. Nous déclinons toute responsabilité en la matière.

### **Copyright**

The University of Lausanne expressly draws the attention of users to the fact that all documents published in the SERVAL Archive are protected by copyright in accordance with federal law on copyright and similar rights (LDA). Accordingly it is indispensable to obtain prior consent from the author and/or publisher before any use of a work or part of a work for purposes other than personal use within the meaning of LDA (art. 19, para. 1 letter a). Failure to do so will expose offenders to the sanctions laid down by this law. We accept no liability in this respect.



**UNIL** | Université de Lausanne

Faculté de biologie  
et de médecine

Département de la Formation et de la Recherche

# Identification of Epigenetic Drivers in DSRC Tumors

Thèse de doctorat ès sciences de la vie (Ph.D.)

Présentée à la

Faculté de biologie et de médecine  
de l'Université de Lausanne

par

**Liliane Celeste BROYE**

Master en Sciences et Technologies du Vivant  
de l'Ecole Polytechnique de Lausanne (EPFL)

## **Jury**

Prof. François Bochud, Président du comité de thèse

Prof. Nicolo Riggi, Directeur de thèse

Prof. Ivan Stamenkovic, Co-directeur de thèse

Prof. Maja Beck-Popovic, Experte

Prof. Douglas Hanahan, Expert

Lausanne 2021

# Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

<b>Président·e</b>	Monsieur	Prof.	François	<b>Bochud</b>
<b>Directeur·trice de thèse</b>	Monsieur	Prof.	Nicolo	<b>Riggi</b>
<b>Co-directeur·trice</b>	Monsieur	Prof.	Ivan	<b>Stamenkovic</b>
<b>Expert·e·s</b>	Madame	Prof.	Maja	<b>Beck-Popovic</b>
	Monsieur	Prof.	Douglas	<b>Hanahan</b>

le Conseil de Faculté autorise l'impression de la thèse de

**Madame Liliane Celeste Broye**

Master en Sciences et technologies du vivant, EPFL, Suisse

intitulée

**Identification of epigenetic  
drivers in DSRC tumors**

Date de l'examen : 19 janvier 2021

Date d'émission de l'imprimatur : Lausanne, le 28 mai 2021

pour le Doyen  
de la Faculté de biologie et de médecine



Prof. Niko GELDNER  
Directeur de l'Ecole Doctorale

# Table of Contents

<b>Résumé</b>	<b>viii</b>
<b>Summary</b>	<b>ix</b>
<b>1. Introduction</b>	<b>1</b>
1.1. The hallmarks of cancer	1
1.1.1. Genome instability and mutation	2
1.1.2. Evading growth suppressors	2
1.2. Rare cancers	5
1.3. Sarcomas	6
1.4. Desmoplastic Small Round-Cell Tumor	9
1.4.1. EWSR1: Ewing's sarcoma breakpoint region 1 gene	11
1.4.2. WT1: Wilms' Tumor suppressor gene 1	12
1.4.3. Known downstream targets of EWS-WT1	15
1.5 Epigenetics	18
<b>2. Aim of my research</b>	<b>22</b>
2.1. Specific aims	22
2.1.1. Define the chromatin signature and transcriptional landscape of DSRCT	22
2.1.2. Elucidate the EWS-WT1 mediated oncogenic mechanisms in DSRCT	23
2.1.3. Exploring the therapeutic potential of target genes/pathways in DSRCT	23
<b>3. Results</b>	<b>24</b>
3.1. Definition of the chromatin landscape and transcriptional profile of DSRCT	24
3.1.1. Identification of EWS-WT1 target genes in DSRCT	24
3.1.2. EWS-WT1 confers an active chromatin structure and regulates transcription at specific loci	27
3.2. EWS-WT1 mediated oncogenic mechanisms in DSRCT	29
3.2.1. Derivation and characterization of human pediatric mesenchymal stem-cells (hpMSC)	29



3.2.2. EWS-WT1 expression recapitulates early stages of DSRCT oncogenesis	33
3.2.3. EWS-WT1 isoforms regulate unique target set of genes in DSRCT	36
3.2.4. DSRCT cells express both EWS-WT1 isoforms concomitantly	49
3.2.5. Expression of both EWS-WT1 isoforms in MeT-5A cells induce tumorigenesis in vivo	53
3.3. Identification of EWS-WT1-specific direct targets genes for potential therapeutic application in DSRCT	61
3.3.1. 3D chromatin looping analysis identified a cohort of EWS-WT1 direct targets in DSRCT	61
3.4. CDK4/6 inhibitors impair DSRCT tumor growth	63
3.4.1. Therapeutic potential of CCND1 -CDK4/6	64
3.4.2. DSRCT is highly sensitive to CDK4/6 small molecule inhibitors	65
3.4.3. The CDK4/6 inhibitor palbociclib impairs DSRCT tumor growth in vivo	66
<b>4. Discussion</b>	<b>72</b>
<b>5. Future directions</b>	<b>80</b>
5.1. Study hpMSCs transcriptome	80
5.2. Study DSRCT tissue heterogeneity	80
5.3. Identify EWS-WT1 collaborative protein complexes	81
5.4. Impede EWS-WT1 expression in xenograft tumors	81
5.5. Identify the roles of EWS-WT1 isoforms	82
5.6. Effect of palbociclib during early stages of DSRCT treatment	82
5.7. Publication of the results	82
<b>6. Collaboration to publications</b>	<b>83</b>
6.1. Ewing's Sarcoma	83
6.1.1. EWSR1 and FLI1 genes	84
6.2. BRG1/BRM-associated factor (BAF) complexes	86
6.3. Synovial sarcoma	87
6.3.1. SS18 and SSX genes	88

6.3.2. Synovial sarcoma target genes	88
<b>7. Materials and methods</b>	<b>96</b>
7.1. Cell culture	96
7.1.1. JN-DSRCT cell-line	96
7.1.2. MeT-5A cell-line	96
7.1.3. LentiX HEK 293T cell-line	97
7.3. Cryopreservation of cells	97
7.5. hpMSC differentiation	
7.5.1. MSC differentiation towards an adipogenic phenotype	98
7.5.2. MSC differentiation towards an osteogenic phenotype	99
7.5.3. MSC differentiation towards a chondrogenic phenotype	99
7.6. Lentiviral Generation	99
7.7. Real-Time Quantitative PCR (qPCR)	100
7.8. Western Blot Analysis	101
7.9. Proliferation Assays	102
7.9.1. Absolute IC50 Calculation	103
7.10. ChIP-seq	103
7.10.1. ChIP-seq Bioinformatic Processing	104
7.10.2. Peakcall	104
7.10.3. Peak intersections	104
7.10.4. Peak annotation	104
7.10.5. Heatmaps and composite plots	104
7.10.6. DNA Motif enrichment analysis	104
7.11. RNA-Seq	105
7.11.1. RNA-seq bioinformatic Processing	105
7.11.2. Identification of differentially expressed genes	105
7.12. Biotinylated isoxazole-Mediated Precipitation	105
7.13. In vivo studies	106
7.13.1. JN-DSRCT1 and MeT-5A xenografts	106
7.13.2. PDX generation from St. Jude Children's Research Hospital	106

7.13.3. In vivo palbociclib experiment	107
7.14. Staining	107
7.14.1. Hematoxylin and Erythrosine (H&E) staining	107
7.14.2. Ki67 staining	108
7.14.3. Cleaved Caspase 3 (CC3) staining	108
7.15. BaseScope	108
7.16. Imaging	110
<b>Acknowledgments</b>	<b>111</b>
<b>List of Abbreviations</b>	<b>114</b>
<b>List of Figures</b>	<b>123</b>
<b>List of Tables</b>	<b>128</b>
<b>My Participation to Figures</b>	<b>130</b>
<b>My Participation to Tables</b>	<b>134</b>
<b>References</b>	<b>135</b>

*Je dédie cette thèse à ma famille où qu'elle se trouve à travers le monde.*

## Résumé

Les cancers pédiatriques représentent une classe hétérogène de tumeurs présentant une variété de formes et biologies. Les sarcomes englobent 20% de tous les cancers pédiatriques touchant les os et les tissus mous. Ces tumeurs sont très agressives avec de fréquentes rechutes et un haut potentiel métastatique. Cette étude se concentre sur des tumeurs avec des altérations génétiques spécifiques qui résultent en l'expression aberrante d'une protéine de fusion à la base de la pathogénèse de ces sarcomes. Les tumeurs desmoplastiques à petites cellules rondes (DSRCTs) sont des néoplasmes mésenchymateux rares et extrêmement agressifs. Ces sarcomes d'origine inconnue touchent les enfants, adolescents et jeunes adultes. Malgré des thérapies multimodales intensives, les DSRCTs ont un faible taux de survie. Les DSRCTs sont invariablement caractérisés par une translocation chromosomique  $t(11;22)(p13;q12)$ , ce qui conduit à l'expression du facteur de transcription aberrant EWS-WT1. Cette protéine possède deux isoformes d'épissage alternatives qui se différencient par l'insertion/absence de trois acides aminés (KTS).

La signature génétique globale des DSRCTs n'a pas été étudiée en profondeur. À notre connaissance, cette étude est la première à définir le panorama épigénétique et la signature transcriptionnelle de tumeurs DSRCT humaines. Nous avons profilé la signature chromatinienne de quatre tumeurs primaires humaines de DSRCT et une lignée cellulaire, appelée JN-DSRCT1. Nous avons analysé les changements d'état de la chromatine après la régulation à la baisse de la protéine EWS-WT1 dans des cellules JN-DSRCT1 et avons identifié une signature chromatinique spécifique à EWS-WT1 qui diffère des autres sarcomes. Notre étude a démontré que chaque isoforme de EWS-WT1 se lie à des régions spécifiques de l'ADN avec une sensibilité différente. De plus, lors de l'expression simultanée des deux isoformes dans des cellules mésothéliales MeT-5a, une potentielle cellule d'origine du DSRCT, on y voit l'apparition de tumeurs viables in vivo.

Nous avons identifié des gènes cibles avec une implication thérapeutique potentielle spécifiques aux DSRCTs. Nous nous sommes focalisés sur le complexe CCND1-CDK4/6 qui peut être inhibé par un médicament tel que le palbociclib. Nous avons démontré que le palbociclib diminuait in vivo la taille des tumeurs DSRCT dérivée de patients. Nous pensons que ce médicament pourrait être utilisé dans les essais précliniques pour traiter les DSRCTs, en combinaison d'autres traitements.

## Summary

Pediatric cancers represent a heterogeneous class of malignancies displaying various clinical presentations, outcomes, and biology. Sarcomas encompass 20% of all pediatric cancers arising from soft tissues and bones, and display very aggressive behaviors with frequent local relapse and high metastatic potential. This study focuses on tumors with specific genetic alterations that result in the expression of a fusion oncoprotein underlying its molecular pathogenesis and clinical behavior. DSRCTs are relatively rare and extremely aggressive malignant mesenchymal neoplasms. This sarcoma of unknown origin affects children, adolescents, and young adults. Despite intensive multimodal therapies, DSRCTs display a low survival rate (less than 15% after five years). DSRCTs are invariably characterized by the specific reciprocal chromosomal translocation  $t(11;22)(p13;q12)$ , which leads to the expression of the aberrant transcription factor EWS-WT1. This protein has two alternative splicing isoforms differing by the insertion/absence of three amino acids (KTS).

The global chromatin signature and epigenetic landscape of DSRCTs are not well studied. To our knowledge, this is the first study defining the chromatin landscape and transcriptional signature of primary DSRCT tumors. We profiled the chromatin signature of EWS-WT1 in four primary human DSRCT tumors and a DSRCT cell-line, JN-DSRCT1. We analyzed the chromatin perturbation after the down-regulation of EWS-WT1 in JN-DSRCT1 cells and identified a specific chromatin signature of EWS-WT1 different from other sarcomas. We used this signature to identify promoters, enhancers, and repressive chromatin domains in JN-DSRCT1. Our study demonstrated that each EWS-WT1 isoform has a specific binding motif with different DNA accessibility potential. The ectopic expression of both EWS-WT1 in the mesothelial cell-line MeT-5A cells, a potential DSRCT cell of origin, generated viable tumors *in vivo*. We identified specific target genes with a potential therapeutic implication in DSRCTs and focused on the CCND1-CDK4/6 complex. We demonstrated that palbociclib, a CDK4/6 inhibitor, decreased tumor burden in PDX DSRCT tumors *in vivo*. We think that palbociclib could be used in preclinical trials to treat DSRCTs in combination with current treatments.

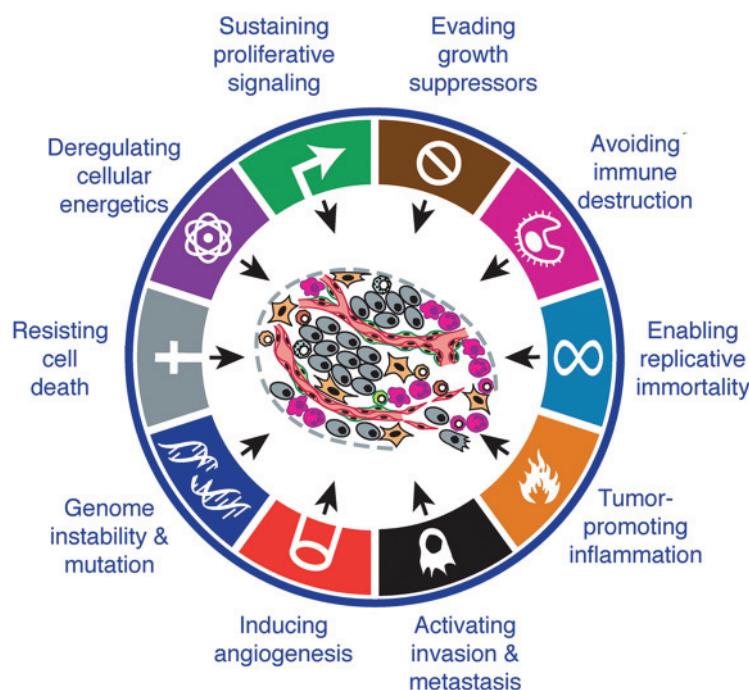
# 1. Introduction

Although it seems to have followed our species throughout its history, cancer remains a disease that science still has a lot to learn about, and researchers worldwide invest their time and knowledge to understand its mechanisms to cure it. Despite this collective effort, cancer remains a formidable opponent that claims, without distinction, millions of lives around the globe, and most individuals, myself included, lost cherished close ones and family members. Unlike any other affliction, cancer is indifferent to ethnicity, gender, age, wealth, or ideological opinions. Although the task of fighting this disease might appear Sisyphean, I still wanted to partake in the effort of understanding some of its mechanisms to find new or better ways to defeat some of its manifestations.

Cancer is one of the leading causes of death worldwide, and millions of people lose their lives every year. Despite extensive research and treatment strategies, cancer incidence and death rates are increasing at alarming rates. Therefore, understanding the complex biology of cancer and formulating new therapeutic approaches are crucial. It will substantially improve the life expectancy and quality of life of patients and their families worldwide. More than 150 different types of cancers can originate from different cells of the human body. Depending on the cell of origin, major oncogenic driver events, and supporting mechanisms, each cancer type possesses unique characteristics. In 2000, D. Hanahan and R. Weinberg described these characteristics and, thereafter, referred to as the “hallmarks of cancers”. They proposed six essential and most common acquired capabilities by normal cells leading to tumorigenesis<sup>1</sup>.

## 1.1. The hallmarks of cancer

These hallmarks are: “sustaining proliferative signaling, evading growth suppressors, enabling replicative immortality, activating tissue invasion and metastasis, inducing angiogenesis and resisting cell death”<sup>1</sup>. Eleven years later, the same authors proposed two additional emerging hallmarks: “avoiding immune destruction and deregulating cellular energetics”, and two enabling characteristics “tumor-promoting inflammation and genome instability and mutation”<sup>2</sup> shared by cancers (Figure 1). Hereafter, I will focus on some of them, as they are relevant in the scope of this study.



**Figure 1: The hallmarks of cancer.** From D. Hanahan and R. Weinberg<sup>2</sup>. List of the eight hallmarks: sustaining proliferative signaling, evading growth suppressors, enabling replicative immortality, activating tissue invasion and metastasis, inducing angiogenesis, resisting cell death, avoiding immune destruction, and deregulating cellular energetics. Here are included two enabling characteristics: tumor-promoting inflammation and genome instability and mutation. All these capabilities can give rise to cancers.

### 1.1.1. Genome instability and mutation

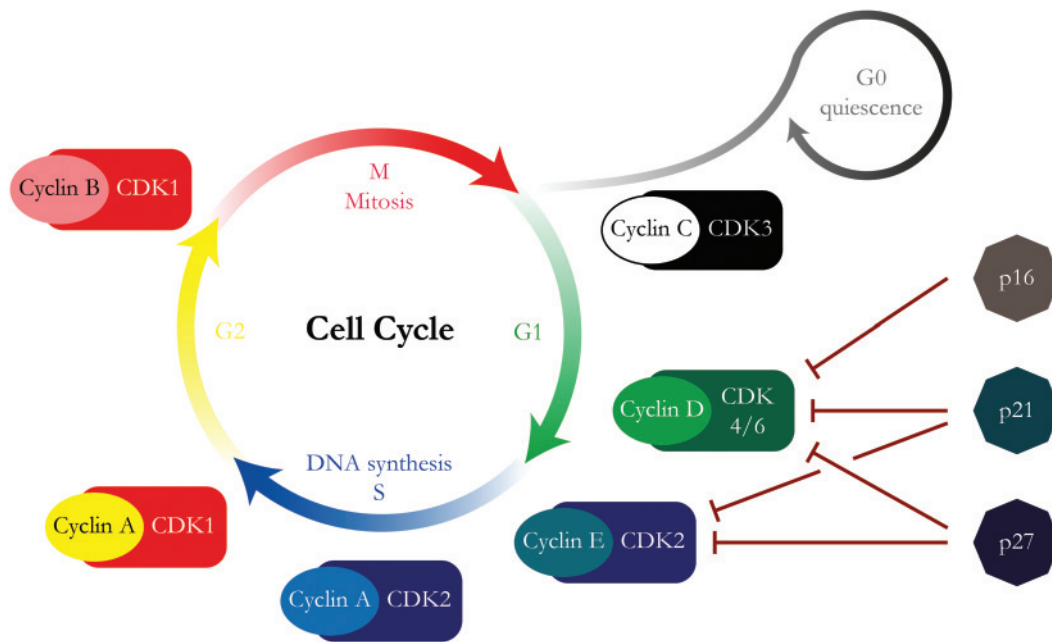
Mutations are rare events occurring through time, and our genome has the ability to maintain and repair DNA<sup>1,2</sup>. However, most cancers are known to be initiated by a genomic aberration, leading to the deregulation of other processes driving cancer development. Therefore, neoplastic cells can acquire several mutations leading to the initiation and progression of cancer. Epigenetics mechanisms may also contribute to genomic instability. Alteration of neoplastic cells' genome can generate random mutations and chromosomal rearrangements that can give rise to the acquisition of several hallmarks of cancers<sup>2</sup>.

### 1.1.2. Evading growth suppressors

Cancer cells have to escape cell-cycle regulation<sup>1,2</sup>. Aberrant regulation of the cell cycle leads to uncontrolled cell growth and proliferation<sup>1,3</sup>. It may be relevant, in the scope of this thesis, to go deeper into describing proliferation in the cell. Proliferation is achieved by the orderly progression through four distinct phases of the cell-cycle (G1, S-DNA synthesis, G2, and M-mitosis). These phases are

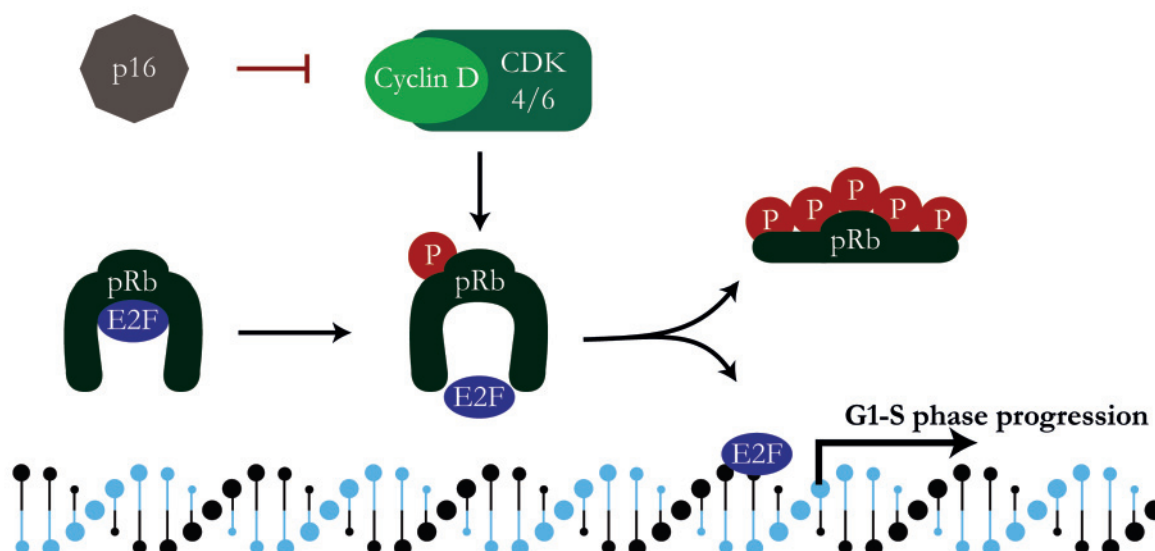


tightly regulated by protein complexes composed of cyclins and cyclin-dependent kinases (CDKs)<sup>4</sup>, as shown in Figure 2. Cyclins are part of a very diverse protein family subdivided into several classes (A to E), have no enzymatic activity, and regulate CDKs<sup>4</sup>. CDKs are members of the well-conserved serine/threonine kinases family, playing a central role in cancer pathogenesis by controlling transcription, RNA splicing, and cell-cycle progression<sup>3,4</sup>. CDKs activity is regulated at multiple levels during the cells-cycle, mainly by a family of CDK-inhibitory factors, such as p16, p21, and p27<sup>3</sup> (Figure 2).



**Figure 2: Cell-cycle and cyclin-CDK complexes involved.** More than 20 years ago, cyclins and CDK's role in the cell-cycle progression were discovered 5: CDK1 participates in the S-G2 and G2-M transition by binding, consecutively, cyclin A and cyclin B, respectively. Cyclin E-CDK2 complex is implicated in G1-S transition. CDK2 is associated with cyclin A during the S phase. Cyclin C-CDK3 complex regulates the G0-G1 transition. Lastly, cyclin D-CDK4/6 complex regulates the G1-S transition. CDK inhibitors also participate in the cell cycle, such as p16, p21, and p27 that can inhibit CDK2, CDK 4, and CDK6 activities.

Progression into the S phase of the cell-cycle is initiated after the formation of a complex between Cyclin D1 and CDK4/6<sup>4</sup>. Normal tissues can develop in the absence of cyclin D1-CDK4/6 complex<sup>6,7</sup>. In contrast, cell-cycle regulation is crucial for many cancers, where cyclin or CDK activity is deregulated. Cyclin D1-CDK4/6 complex phosphorylates the tumor-suppressor protein retinoblastoma (Rb) and inactivates its growth inhibition function leading to cell tumor growth<sup>3,4</sup>. When Rb protein is active, Rb traps the transcription factor E2F preventing advancement in the G1 phase<sup>4,8</sup>. Subsequently, after Rb phosphorylation by Cyclin D1-CDK4/6 complex, E2F is released and activates target genes that support DNA synthesis and facilitate the transition from G1 to S<sup>4,9,10</sup> (Figure 3).



**Figure 3: Regulation and function of Cyclin D-CDK4/6 complex.** Cyclin D1 and CDK4/6 form a complex during the G1 phase. This complex can be inactivated by p16 binding. The activated Cyclin D-CDK4/6 complex initiates Rb phosphorylation. An E2F transcription factor is then released and drives the expression of genes required to enter the S phase. Notably, CDK4/6 regulates an essential checkpoint for the G1-S transition and the commitment of the cell-to-cell division.

The tumor suppressor protein 53 (TP53) pathway can sense intracellular stress signals that impair DNA replication and cell-division process fidelity. This stress is transmitted to TP53 by post-translational modifications. Once TP53 is activated, it initiates the cell-cycle arrest that can lead to cellular senescence or apoptosis. Interestingly, the TP53 pathway is inactivated in the majority of human cancers, hence escaping cell-cycle regulation. Cancer cells can evade the cell-cycle by altering the Rb or the TP53 pathway function, leading to escape critical gatekeepers of the cell-cycle progression and allowing persistent cell proliferation<sup>1,2</sup>.

The hallmarks of cancer are functional capabilities that a cell may need to acquire to initiate cancer development. They are useful concepts for the global understanding of the complex biology of cancer development. Nevertheless, several cancers develop without the acquisition of all these capabilities. Additionally, these hallmarks can be obtained by different mechanisms, whereas some specific genetic events can contribute to acquiring one capability in one cancer, in another, one event may lead to the acquisition of several of them. Moreover, the acquisition of these hallmarks is not linear and can happen at different moments during the tumorigenesis<sup>2</sup>. The hallmarks of cancer show the complexity of tumor development, maintenance, and dissemination. Intensive research on these described acquired

capabilities lead to the development of new therapeutic strategies and drugs. Some are currently in clinical trials<sup>1</sup>. These new treatments hold promise as cancer therapy to increase the survival of cancer patients<sup>2</sup>. Nevertheless, a substantial number of patients are refractory to these treatments, prompting researchers to explore new approaches.

Hereafter, I would like to focus on rare cancers, particularly pediatric sarcomas. Afterwards, I will describe Desmoplastic Small Round-Cell Tumors (DSRCTs) in greater detail to summarize the current knowledge on this disease, and the rationale of my research project.

## 1.2. Rare cancers

The Office of Rare Diseases from the National Institute of Health defines rare diseases as those having an incidence of less than 200'000 people affected in the USA. In Europe, rare diseases are defined by 1 case per 2'000 persons<sup>11</sup>. Although rare as individual tumors, in aggregate, rare tumors have various clinical presentations and diverse biology<sup>11</sup>. There are several challenges related to rare tumors. First, their management remains difficult, lacking specific standard protocols for treatment. Second, there is a small number of tumors available for research. Therefore, knowledge of these tumors is based on limited data<sup>11</sup>.

All childhood cancers are considered as rare. Pediatric cancers comprise more than 60 different types of cancers originating from diverse tissues<sup>12</sup>. Nevertheless, most of them are considered to arise from embryonal cells<sup>12</sup>. Childhood cancers diverge from adult tumors by their very low number of genetic mutations<sup>12</sup>. Thus, they use other mechanisms to develop. Due to limited data and samples available on pediatric cancers, current treatments are based on aggressive therapies with long-term side effects impacting the patients' quality of life<sup>12</sup>.

Bone and soft-tissue sarcomas are a rare form of cancer affecting people of all ages and representing less than 1% of all adult malignant tumors<sup>13-15</sup>. In children, this percentage rises to 20%<sup>14</sup>, with high mortality and morbidity. Consequently, it is fundamental to investigate it further in order to discover new druggable targets for pediatric sarcomas.

### 1.3. Sarcomas

Similar to epithelial cancer, sarcoma include a collection of malignancies with diverse biological characteristics. Although the cell of origin is unknown for most of them, they arise in tissues derived from the mesenchymal lineage, in contrast to carcinomas that arise from epithelial lineage tissues. Sarcomas are very aggressive soft-tissue and bone tumors with high invasive, metastatic potential<sup>12,16,17</sup>, and intra-tumoral heterogeneity. Soft-tissue sarcomas represent 20% of all pediatric cancers with genetically heterogeneous backgrounds, as they are morphologically and clinically diverse<sup>18</sup>. These types of cancers arise with a low frequency and have a poor prognosis. Currently, treatment consists of a combination of surgery (when possible) and chemo- or radiotherapy<sup>16,19</sup>. However, the frequency of relapse is high, despite this multimodal therapeutic approach<sup>17</sup>.

Genetics of sarcoma are highly variable. Nonetheless, genomic analysis distinguishes two groups of sarcomas. The first class includes tumors with complex genetic abnormalities lacking any specific pattern that can be driven by copy number alterations or by mutations in canonical drivers<sup>20</sup>. On the contrary, the second class encompasses tumors with defined simple genetic alterations, leading to sequential and well-orchestrated events. I will focus on this second type of sarcoma, which comprises malignancies such as Ewing's sarcoma and DSRCTs, among others. The development of these tumors is driven by unique and well-characterized balanced chromosomal translocations, leading to the expression of aberrant fusion genes that directly underlie the oncogenic transcriptional programs of these malignancies<sup>17,21</sup>. These specific fusion proteins are also used as diagnostic biomarkers to determine the type of sarcoma<sup>12</sup>.

Translocated sarcomas harbor unique mutations that give rise to FET (for FUS, EWSR1, and TAF15)-associated or FET-independent fusion proteins (Table 1). The FET-fusion proteins often give rise to aberrant transcription factors or transcriptional regulators, whereas FET-independent fusion proteins frequently lead to constitutively active growth factor or receptor tyrosine kinases<sup>21</sup>.

Tumors with defined simple genetic alterations, such as translocated sarcomas, produce aberrant fusion proteins, which transcriptional activity is often higher than each of its components. This newly expressed protein leads to *de novo* activation of specific genes implicated in the development of the

corresponding cancer type<sup>17</sup>. For sarcomas displaying a very low number of genetic mutations<sup>12</sup>, it has been demonstrated that epigenetic alterations are involved in the development of these tumors<sup>22</sup>.

In pediatric sarcomas, chromosomal abnormalities can create oncogenic fusion proteins with novel functions, as detailed in Table 1. These oncogenic fusion proteins can induce epigenetic modifications that often synergize with cell-cycle regulators to promote oncogenic processes. For example, the introduction of the EWS-FLI1 protein in primary human mesenchymal stem cells (hpMSCs) can promote a differentiation program reminiscent of Ewing's sarcoma<sup>23</sup>. In this project, we hypothesize that expression of the aberrant EWS-WT1 protein in a permissive cellular background may induce an oncogenic expression program resulting in the development of tumors resembling DSRCT.

Sarcoma Name	Translocation	Fusion transcript	Proportion
Desmoplastic small round-cell tumor	t(11;22)(p13;q12)	EWSR1-WT1	>95%
Ewing's sarcoma	t(11;22)(q24;q12)	EWSR1-FLI1	~85%
	t(21;22)(q22;q12)	EWSR1-ERG	~10%
	t(7;22)(p22;q12)	EWSR1-ETV1	
	t(17;22)(q12;q12)	EWSR1-ETV4	
	t(2;22)(q35;q12)	EWSR1-FEV	
	t(16;21)(p11;q22)	FUS-ERG	
Synovial sarcoma	t(x;18)(p11.2;q11.2)	SS18-SSX	>90%
Clear cell sarcoma	t(12;22)(q13;q12)	EWSR1-ATF1	
	t(2;22)(q33;q12)	EWSR1-CREB1	
Alveolar rhabdomyosarcoma	t(2;13)(q35;q14)	PAX3-FOXO1A	70%
	t(1;13)(q36;q14)	PAX7-FOXO1A	10%
Myxoid liposarcoma	t(12;16)(q13;p11)	FUS-DDIT3	95%
	t(12;22)(q13;p12)	EWSR1-DDIT3	2%
Low grade fibromyxoid sarcoma	t(7;16)(q33-34;p11)	FUS-CREB3L2	
	t(11;16)(p11;p11)	FUS-CREB3L1	
Ewing's-like sarcoma	t(4;19)(q35;q13)	CIC-DUX4	
	t(10;19)(q26;q13)	CIC-DUX4	
	t(x;19)(q13;q13.3)	CIC-FOXO4	
	inv(X)(p11.4;p11.2)	BCOR-CCNB3	
Extraskeletal myxoid chondrosarcoma	t(9;22)(q22;q12)	EWSR1-NR4A3,	
	t(9;17)(q22;q11)	TAF2N-NR4A3	
	t(9;15)(q22;q21)	TCF12-NR4A3	
	t(3;9)(q11;q22)	TGF-NR4A3	

**Table 1: Major gene fusions of bone and soft tissue tumors with poor prognosis.** The translocated sarcomas can harbor FET-fusion proteins or FET-independent fusion proteins. Here is a list of sarcomas with name, translocation related, and fusion transcript with their corresponding percentage of detection in each sarcoma.

## 1.4. Desmoplastic Small Round-Cell Tumor

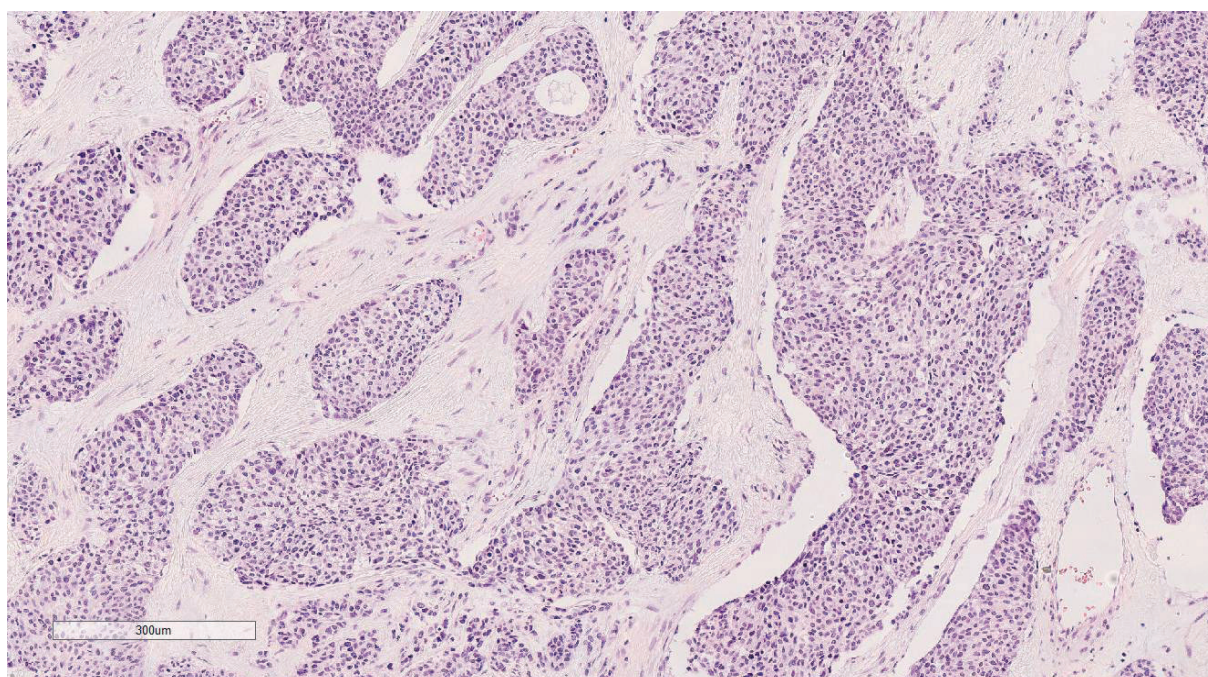
DSRCT is a rare and extremely aggressive malignant mesenchymal neoplasm<sup>24,25</sup> that affects children, adolescents, and young adults. It was first described in 1989 by W. Gerald and J. Rosai<sup>24</sup>. DSRCT has a peak incidence of 0.74 cases/million people/year worldwide in individuals between 20–24 years of age<sup>26</sup>. In African-Americans, the peak incidence is higher than in Caucasians or other ethnic groups (0.5 vs. 0.2 cases/million, respectively) and arises later between 25 and 29 years old<sup>26</sup>. DSRCT arise predominantly in males, with a male to female ratio of 4:1<sup>25-27</sup>. The origin of DSRCT remains unclear, but preliminary evidences point to undifferentiated cells with multilineage potentials, such as primitive mesenchymal cells, lining the coelomic cavities, and gonads<sup>28,29</sup>.

Clinically, DSRCT is a soft-tissue sarcoma that arises mainly in the abdominal cavity and pelvis, often as a multicentric abdominal mass, but can also arise less frequently in the head, limbs, neck, kidney, ovary/fallopian tube, and brain<sup>25-27</sup>. The symptoms are pain, abdominal distention, palpable mass, and organ obstruction<sup>25</sup>. In most cases, patients present advanced diseases because most of them remain asymptomatic for an extended period of time<sup>29</sup>. For diagnosis, magnetic resonance imaging (MRI) or positron emission tomography (PET) scans are used as well as histopathologic and cytogenetic analysis<sup>30</sup>. Unfortunately, DSRCT has a low survival rate (44% after three years, less than 15% after five years)<sup>31</sup>, and 40% of patients already have metastasis at the time of diagnosis. The most common metastasis sites are in the liver, distant lymph nodes, lungs, and bones<sup>27,29,30</sup>. Due to the rarity of this tumor, DSRCT patients often join clinical studies designed for other sarcomas<sup>28</sup>. Currently, treatment of DSRCT is a combination of cytoreductive surgery, radiotherapy, and intense high-dose alkylation-based<sup>30</sup> combination chemotherapy, with or without hyperthermic intraperitoneal chemotherapy<sup>29,32</sup>. Other approaches involve tyrosine kinases inhibitors<sup>33</sup>, mTOR (mammalian target of rapamycin) inhibitors, or immunotherapies with the human B7 homolog 3 (B7-H3) or single-agent anti programmed death-1 (PD1) antibodies<sup>29</sup>, but these trials demonstrated only limited clinical benefit<sup>29</sup>. Very few druggable targets have been identified, so far, to treat DSRCT and showed limited therapeutic potential<sup>26,28</sup>, underscoring the need to identify new therapeutic strategies for these tumors.

Histologically, DSRCT is composed by nests of small round-cells with poly-phenotypic differentiation (neural, epithelial, and mesenchymal features) surrounded by dense desmoplastic stroma<sup>27</sup>



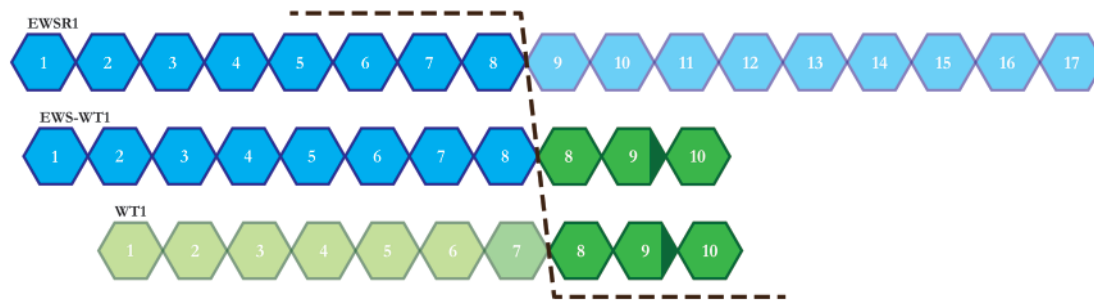
(Figure 4). Immunohistochemical stainings show the expression of divergent differentiation markers by the neoplastic cells. DSRCT express epithelial membrane antigen (EMA) and keratin (epithelial features), desmin and vimentin (mesenchymal features), and neuron-specific enolase (NSE) and NCAM (neural features)<sup>25,27,28,30</sup>. Nevertheless, all these markers are expressed in subpopulations of the tumors. Therefore, they are not uniformly expressed across the tumor, and they are not all present simultaneously in the same parts of the tumor. Additionally, DSRCT also expresses other antigens, such as GD2, GD3, and B7-H3<sup>28</sup>. GD2 is a glycosphingolipid, whose functional role remains poorly understood but may help tumor cells to attach to the extracellular matrix<sup>28</sup>. GD3 is a ganglioside that is associated with proliferation, adhesion, and invasiveness in melanoma. B7-H3 (also called CD276) antigens are expressed in more than 90% of DSRCT, where they contribute to tumor cell-immune evasion and play a prominent role in tumor growth and metastasis<sup>25,27,28,30</sup>. The tumor microenvironment is composed mainly of scant fibroblasts embedded in dense collagen matrix<sup>25</sup> with increased expression of CCN2 (Cellular Communication Network Factor 2), a connective tissue growth factor associated with abundant extracellular matrix<sup>28</sup>. Furthermore, the stromal microenvironment is positive for  $\alpha$  Smooth-Muscle Actin ( $\alpha$ -SMA)<sup>25</sup>. Finally, Vascular Endothelial Growth Factor (VEGF)-dependent angiogenesis seems to be important in DSRCT with elevated expression of VEGFA and VEGFR-2 in DSRCT xenografts<sup>27</sup>.



**Figure 4: Representative histological image of a human DSRCT stained with hematoxylin and eosin (H&E).** The pathological examination of morphology shows several nests of small round-cells surrounded by desmoplastic stroma. Scale bar 300µm.



The molecular hallmark of DSRCT is the reciprocal chromosomal translocation  $t(11;22)(p13;q12)^{27}$ , which leads to the expression of the aberrant transcription factor *EWSR1-WT1*. This translocation consists of an in-frame fusion of the 3' end of the Ewing's sarcoma gene (*EWSR1* exon 1 to 7) with the DNA binding domain of the Wilms' tumor gene, containing three zinc fingers (*WT1* exon 8 to 10)<sup>25,34</sup> (see Figure 5).



**Figure 5: Schematic representation of EWSR1, WT1, and the resulting EWSR1-WT1 fusion gene.**

EWSR1 is located in chromosome 22q12, has 17 exons, and the average protein size is ~ 68kDa. WT1 is located in chromosome 11p13, has ten exons, and the average protein size is ~52 kDa. In WT1, exons 7 to 10 encode for zinc fingers 1 to 4, in green. WT1 has two significant isoforms differentiated by the insertion of nine nucleotides at the end of exon 9 (dark green). The fusion gene EWSR1-WT1, in the middle, fuses the first seven exons of the EWSR1 gene (depicted in blue) with the last three exons of the WT1 gene (depicted in green). All DSRCT tumors express both EWSR1-WT1 isoforms.

#### 1.4.1. EWSR1: Ewing's sarcoma breakpoint region 1 gene

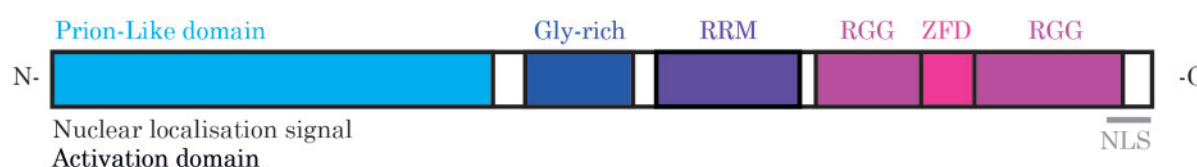
Fused in liposarcoma (FUS), Ewing's Sarcoma (EWS) and TATA-binding associated factor 15 (TAF15) belong to the FET family of DNA and RNA binding proteins. FET proteins are highly conserved and ubiquitously expressed. The Ewing's sarcoma breakpoint region 1 or EWS RNA binding protein 1 (*EWSR1*) gene is expressed ubiquitously in almost all cells-types<sup>35</sup>. It plays epigenetic roles in gene expression and RNA processing. *EWSR1* was firstly identified in 1992 at the chromosomal breakpoint  $t(11;22)(q24;q12)$  characterizing Ewing's sarcoma<sup>19</sup>. This translocation forms the aberrant fusion protein EWS-FLI1, where the N-terminal part of EWS fuses to the C-terminal part of FLI1<sup>16</sup>. EWS-FLI1 can activate target genes at GGAA microsatellite repeats and induce *de novo* activation of an enhancer from a closed chromatin<sup>36,37</sup>.

Ewing's sarcoma represents 1.5% of all childhood cancers<sup>16</sup>. This cancer is highly malignant, with poorly differentiated small round blue cell tumors, with an overall poor survival rate. It affects 1-3:1'000'000 children per year, with a male to female ratio of 6 to 4<sup>38,39</sup>. It affects mainly white Cauca-

sians compared to other populations (66% in White, 24% in Hispanic, 6% in Asian, and 3% in Black population)<sup>38-40</sup>. Mutations or deficiencies in *EWSR1* can lead to several disorders or diseases<sup>35</sup>. For example, EWS deficiency contributes to premature senescence in hematopoietic stem-cell, and *EWSR1* missense mutations are associated with central nervous system disorders such as amyotrophic lateral sclerosis and frontotemporal dementia<sup>35</sup>. Nevertheless, despite known roles in cellular functions, the exact mechanisms of *EWSR1* are not yet fully understood<sup>35</sup>.

*EWSR1* is located in chromosome 22q12, has 17 exons, and the average protein size is ~ 68kDa. EWS protein has a prion-like domain (PrLD) at its N-terminal, enriched in particular amino-acids (glutamine, tyrosine, asparagine, and serine)<sup>41,42</sup>, embedded in the DNA activation domain. EWS also has an RNA-recognition motif (RRM), two arginine-glycine-glycine-rich (RGG) domains separated by one zinc finger domain (ZFD), and a C-terminal nuclear localization sequence that is also involved in nucleic-acid binding (Figure 6).

## EWS



**Figure 6: Schematic representation of EWS protein and its corresponding domains.** EWS has a Prion-Like domain, a Glycine-rich domain (Gly-rich), an RNA-recognition motif (RRM), two arginine-glycine-glycine-rich (RGG) domains separated by one zinc finger domain (ZFD), and a C-terminal nuclear localization sequence (NLS).

### 1.4.2. WT1: Wilms' Tumor suppressor gene 1

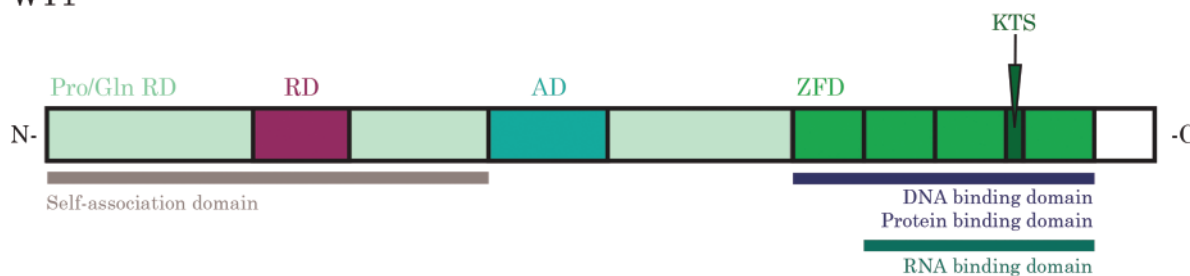
The Wilms' Tumor suppressor gene 1 (*WT1*) is a zinc finger transcription factor, described as playing a role in developing the pediatric kidney Wilms' tumor<sup>43-45</sup>. WT1 is crucial for normal fetal development, kidney and gonad embryogenesis during normal development, and expressed in a tissue-specific manner in adult tissues<sup>46,47</sup>. WT1 can function as a tumor suppressor or an oncogene under different cellular conditions<sup>46,48</sup>, and its expression in tumors is associated with worse overall survival<sup>49</sup>. 10 to 15% of sporadic Wilms' tumor patients have inactivating mutations in the *WT1* gene. Wilms' tumor is the most common type of kidney cancer and genito-urinary malignancy in children.

This tumor affects 1:10'000 children, 75% of them are younger than four years old. This cancer can be sporadic or hereditary<sup>19,43,46</sup>.

*WT1* is located in chromosome 11p13, contains ten exons, and the average protein size is ~52 kDa. It has a self-association domain, an activation domain, and a repression domain on the N-terminal part of the protein, functioning independently<sup>44</sup>. At the C-terminal, WT1 has a DNA/RNA/protein-binding domain composed of four zinc fingers<sup>44,45</sup>. *WT1* has four main alternatively RNA splicing variants that encode for four functionally different products<sup>50</sup>. Furthermore, RNA editing and alternative initiation codons increase the number of potential isoforms to 16. The two major splicing events occur on exon 5 and 9. Splicing at the first site results in the inclusion or exclusion of 17 amino acids encoding for exon 5 at the N-terminus of WT1 that is present only in mammals<sup>51</sup>. Splicing at the second site, results in the inclusion or exclusion of 3 amino acids in exon nine at the C-terminus of WT1<sup>46</sup>. This second alternative splice event is highly conserved throughout evolution. Therefore, it can be found in all vertebrates, suggesting that each isoform may display distinct but important biological functions<sup>44,52</sup>. This alternative splice event on exon 9 of WT1, leads to a protein that differs by three amino acids: lysine, threonine, and serine (KTS) between the zinc fingers 3 and 4<sup>34,53</sup> (Figure 7). The expression ratio between the WT1 +KTS and -KTS isoforms is 2:1, suggesting a higher biological significant role for the +KTS isoform<sup>46</sup>.

Nevertheless, it has been shown that the KTS insertion decreases the DNA-binding affinity of WT1<sup>54,55</sup>, and supports the evidence that the +KTS variant could have different transcriptional targets than the -KTS isoform<sup>53</sup>. Thus, they can bind different DNA motif<sup>54</sup>. It is thought that the ability of WT1 to bind RNA<sup>55</sup> and to interact with proteins is provided by the +KTS isoform, but it is still unclear how the KTS insertion can affect the RNA/DNA binding capacity of WT1.

## WT1



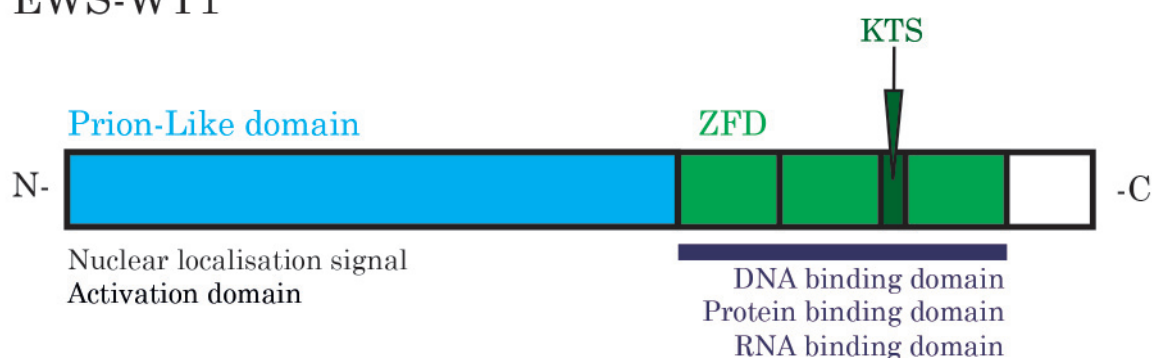
**Figure 7: Schematic representation of WT1 protein and its corresponding domains.** The N-terminal part of WT1 is rich in prolines and glutamines (Pro/Gln RD) and has a self-association domain, a repression domain (RD), and an activation domain (AD)<sup>45</sup>. The repression and activation domains can function independently<sup>44</sup>. A DNA/RNA/protein binding domain at the C-terminal is composed of four Krüppel-like zinc fingers (ZFD). Each zinc finger is formed by two  $\alpha$ -helix (with two histidines) and one  $\beta$ -sheet (with two cysteines) that are stabilized by a zinc ion. WT1 has a major alternative splicing event at its C-Terminal, involving nine nucleotides that can be either included or excluded between zinc finger 3 and 4 (KTS)<sup>44</sup>. The ratio of WT1 +KTS and –KTS isoform is 2:1<sup>46</sup>.

Heterozygous mutations of WT1 give rise to WT1-related syndromes called: WAGR syndrome (Wilms tumor, Aniridia, Genitourinary malformation, and mental Retardation), Denys-Drash syndrome, and Frasier syndrome<sup>19,44-46</sup>. The WAGR syndrome is a rare genetic disorder characterized by the deletion of 11p13 region, where *WT1* and neighboring genes can be deleted, including paired box-protein 6 (*PAX6*) and brain-derived neurotrophic factor (*BDNF*)<sup>56</sup>. A mutation, rather than a deletion, of *WT1* is found in Denys-Drash syndrome and Frasier syndrome<sup>57</sup>. Both syndromes are inherited diseases and are characterized by renal failure, gonadal tumor, and male pseudohermaphroditism<sup>57</sup>. In the Denys-Drash syndrome, a mutation in exons 8 or 9 of *WT1* leads to abnormal WT1 protein production. In Frasier syndrome, a point mutation in intron 9 (intervening sequence IVS9) of *WT1* disrupt splicing at the second alternative splice donor site<sup>58</sup>. This point mutation leads to the downregulation of the WT1+KTS isoform, inverting the WT1 +KTS/-KTS ratio to 1:2<sup>58</sup>, which, in turn, diminishes sex-determining region Y (SRY) expression<sup>56,57</sup>. Patients presenting WAGR and Denys-Drash syndromes display a 50% and 90% risk, respectively, to develop Wilms' Tumor. In most of these tumors, the remaining normal *WT1* allele is lost, consistent with the “two-hit hypothesis”<sup>243</sup>, where one copy of a tumor suppressor gene is inactivated by a mutation, and the other copy is inactivated by loss of heterozygosity<sup>59</sup>. Nevertheless, this event does not happen in all patients with WT1-related syndromes, suggesting that a second hit's window opportunity is temporally limited<sup>58</sup>.

### 1.4.3. Known downstream targets of EWS-WT1

In addition to its inactivation in Wilms' tumor, WT1 plays an essential role in DSRCT, wherein EWS fuses to the C-Terminal part of WT1. The fusion to EWS does not alter the KTS site's integrity, leading to the generation of two alternative chimeric proteins: EWS-WT1 +KTS and EWS-WT1 -KTS<sup>34</sup>, as shown in Figure 8. DSRCT is the third known tumor-type associated with a translocation of EWSR1, after Ewing's sarcoma and clear-cell sarcoma, and it is the first tumor type associated with the translocation of WT1<sup>34</sup>.

#### EWS-WT1



**Figure 8: Schematic representation of EWS-WT1 protein and its corresponding domains.** The size of the EWS-WT1 protein is ~ 64kDa. At the N-Terminal, there is a nuclear localization signal, an activation domain, and a prion-like domain. At the C-Terminal, there are three zinc finger domains (ZFD) with an RNA/DNA/protein-binding domain. There is a tripeptide (KTS) alternative splice variant, indicated in dark green.

The molecular mechanisms by which the two chimeric EWS-WT1 isoforms contribute to DSRCT transformation are still poorly understood. Previous studies suggest that these proteins play critical and diverse roles in various DSRCT oncogenesis stages through aberrant transcription of different target genes. So far, 18 downstream target genes have been linked to DSRCT<sup>29</sup>, among them, there are three main categories: genes encoding for growth factors, transcriptional regulators, and genes involved in cell adherence and extracellular matrix deposition. The first category includes targets encoding for growth factors ligands and receptors, such as *PDGFR $\alpha$* <sup>60</sup>, *IGF-1R*<sup>61,62</sup>, *IL2RB*<sup>63</sup>, *EGFR*<sup>29,64</sup>, *IGF2*<sup>65</sup>, *FGFR4*<sup>65</sup>, and *CCN2*<sup>66</sup>. The second category includes regulators of transcription, such as *c-MYC*, *N-MYC*, *PAX2-2*<sup>29</sup>, *ENT4*<sup>67</sup>, *BALAP3*<sup>68</sup>, *ASCL1*<sup>69</sup>, and *MLF-1*<sup>65,70,71</sup>. The last category involves genes, such as *LRRC15*<sup>72</sup>, *Syndecan-1*, *TALLA1*<sup>73</sup>, and *E-cadherin*<sup>29</sup> (see Table 2 for further details). It is still unclear how these proteins contribute to the pathophysiology of DSRCT, but they can be used as potential therapeutic targets.

Target	Name	Description
IGF-1R	Insulin-like growth factor 1 receptor	Receptor tyrosine kinase, crucial for tumor transformation and survival of malignant cells
PDGF $\alpha$	Platelet-Derived Growth Factor Subunit A	Growth factor, involved in the regulation of embryonic development, cell proliferation, cell migration, survival, and chemotaxis
IL2RB	Interleukin 2/15 receptor $\beta$	Receptor involved in the cytokine signaling pathway
IGF2	Insulin-like growth factor 2	Growth factor that plays an crucial role in the proliferation
CCN2	Cellular Communication Network Factor 2	Plays a role in cell adhesion in many cell types, and it usually is a mitogen secreted by vascular endothelial cells
FGFR4	Fibroblast Growth Factor Receptor 4	Tyrosine kinase and cell-surface receptor for fibroblast growth factors. It is involved in the regulation of several pathways, among them cell proliferation, differentiation, and migration
EGFR	Epidermal growth factor receptor	Receptor tyrosine kinase binding ligands of the EGF family
ASCL1	Achaete-scute homolog 1	Neuronal reprogramming factor
BAIAP3	BAI1-associated protein 3	TP53-target gene that encodes a Brain-specific Angiogenesis Inhibitor
ENT4	Equilibrative nucleoside transporter member 4	Polyspecific organic cation transporter
c-MYC	Myc proto-oncogene protein	Transcription factor regulating growth-related genes binds to the VEGFA promoter
N-MYC	N-myc proto-oncogene protein	Transcription factor
PAX2-2	Paired box-protein 2	Transcription factor with a critical role in the development of the urogenital tract, the eyes, and the central nervous system
MLF-1	Myeloid leukemia factor 1	Plays a role in proliferation and cell survival
TALLA1	T-cell acute lymphoblastic leukemia-associated antigen 1	Cell-surface glycoprotein, regulating cell adhesion, migration, and metastasis, also known as TSPAN7
LRRC15	Leucine-rich repeat-containing protein 15	Transmembrane and matricellular protein
Syndecan-1		Cell-surface proteoglycan, links the cytoskeleton to the interstitial matrix
E-cadherin	Epithelial cadherin	Transmembrane and cell adhesion protein

**Table 2: List of identified EWS-WT1 targets in the literature.**

Several genes were identified as specific targets of the EWS-WT1 –KTS isoform<sup>74</sup>, such as *PDGFA*<sup>60</sup>, *IGF-1R*<sup>61,62</sup>, *BALAP3*<sup>68</sup>, *MLF-1*<sup>71</sup>, and *TALLA1*<sup>73</sup>. In the literature, several consensus sequences associated with the EWS-WT1 –KTS isoforms can be found and are listed in Table 3 for further information. On the contrary, the only known downstream transcriptional target of the EWS-WT1 +KTS isoform is the Leucine-Rich Repeat Containing 15 (LRRC15)<sup>72</sup>, a cell-surface protein involved in cellular invasion<sup>27,45</sup>. Furthermore, the only consensus sequences associated with the EWS-WT1 +KTS isoform is the 5'-GGA-GG(A/G)-3' DNA motif<sup>72</sup>, confirming specific DNA binding sites of each EWS-WT1 isoform. This difference in binding motif suggests that the insertion of the tripeptide KTS can alter the DNA binding capacity of EWS-WT1<sup>54</sup>. Depending on the presence or absence of the KTS tripeptide in the DNA-binding domain of EWS-WT1, different target genes can be activated. These substantial changes in the transcriptome of transformed cells by the fusion oncogene contribute to the broad inter-tumor heterogeneity of DSRCTs<sup>28</sup>.

Specificity	Motif sequence
EWS-WT1 (-KTS)	TCCTCCTCCTCCTCTCC <sup>75</sup>
EWS-WT1 (-KTS)	GCGGGGGCG <sup>62</sup>
EWS-WT1 (-KTS)	(G/C)(G/C)(G/C)TGGGGG <sup>54,63</sup>
EWS-WT1 (+KTS)	GGAGG(A/G) <sup>54,72</sup>
EWS-WT1	(C/A)CTCCC(A/C)C(A/G)C <sup>65</sup>

**Table 3: Summary of DNA binding motifs of EWS-WT1 in the literature.**

Recent genomic analysis of DSRCT primary tumors revealed a low frequency of additional recurrent secondary genomic alterations<sup>76</sup> and deregulation in genes related to the epithelial-mesenchymal transition (EMT), mesenchymal-epithelial reverse transition (MET), immune response, and DNA damage response, such as *ATR*, *SLFN11*<sup>32</sup>, *PARP1*<sup>32</sup>, *RRM2* and *USP9X*<sup>28,29,33</sup>.

One investigative path of rare diseases leads to epigenetics. Epigenetics mechanisms are one of the key processes for the initiation and progression of cancer<sup>77</sup>. In DSRCT, epigenetic modifications are thought to drive the development of this tumor. Recently, advances in research technologies have opened the epigenetic field. Nowadays, these techniques are used to characterize epigenetic changes in cancer cells or to decipher molecular epigenetics mechanisms that will be discussed later in this chapter.



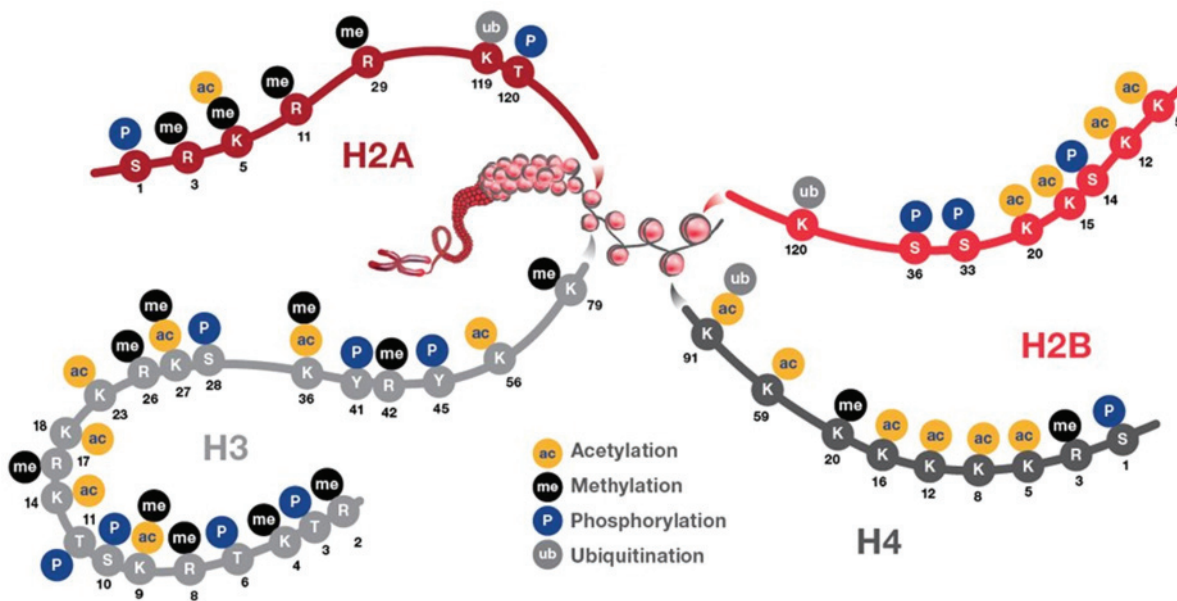
## 1.5 Epigenetics

The term “epigenetics” has evolved over the last 80 years to become “the study of heritable changes in gene expression caused by mechanisms other than changes in the DNA sequence”<sup>78</sup>. In other words, epigenetic mechanisms are collective dynamic processes that regulate gene expression<sup>77</sup>. Epigenetic mechanisms include DNA methylation, chromatin remodeling, histone modification, non-coding regulatory RNA (lncRNAs, miRNAs, and eRNAs), and chromatin three-dimensional (3D) architecture<sup>77-79</sup>. Therefore, epigenetic mechanisms are critical for controlling the essential processes of the cells.

The earliest epigenetic modification found in humans is DNA methylation<sup>79</sup>. DNA methylation at cytosine bases in the genome is tightly linked to gene expression<sup>80</sup>. Methylated cytosines are considered to be an inheritable epigenetic mark<sup>77,80</sup>, and 5-methylcytosine (5mC) is the most abundant covalent modification found on genomic DNA<sup>81</sup>. 5mC is found in all tissues<sup>80</sup>, corresponding to 5% of all cytosines. In cytosine–phosphate–guanine (CpG) regions, the 5mC percentage rises to 80%<sup>80</sup>. 5mC can recruit CpG binding proteins along with their associated histone-modifying and remodeling complexes<sup>81</sup>. Ten-Eleven Translocation (TET) enzymes are methylcytosine dioxygenases that play an essential role in maintaining the fidelity of DNA methylation patterns<sup>82</sup>. TET can catalyze DNA oxidation of 5mC in 5-hydroxymethylcytosine (5hmC), in a Fe(II) and  $\alpha$ -ketoglutarate ( $\alpha$ -KG) dependent manner<sup>79,80,82-84</sup>.

The most common covalent post-translational histone modifications comprise mono-, di-, and tri-methylation (me) of Lysines (K), acetylation (ac) of Lysines, phosphorylation (P) of Serines (S), and ubiquitination (ub) of Lysines (Figure 9). In addition to these modifications, histones can also be altered by citrullination, ADP-ribosylation, deamination, formylation, O-GlcNAcylation, butyrylations, propionylation, crotonylations, and proline isomeration<sup>83,85</sup>.

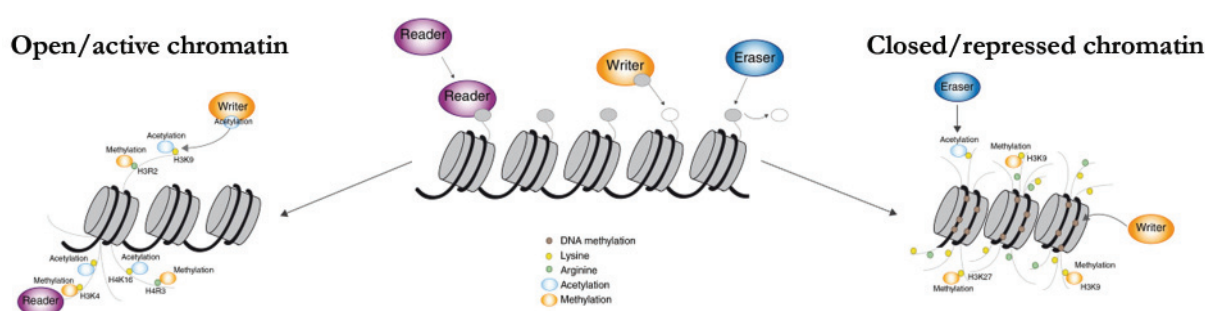




**Figure 9: Possible histone modifications (modified from Thermo Fisher Scientific).** DNA is wrapped around an octamer of four histone proteins (H2A, H2B, H3, and H4)<sup>83,85</sup>. Each histone has a tail, which is subject to post-translational modifications to govern the chromatin state<sup>85</sup>. Here, each histone's tail is represented by a color (dark red for H2A, red for H2B, grey for H3, and dark grey for H4). On each tail, only some amino acids (S - Serine, R - Arginine, K - Lysine, T - Threonine) can be modified post-transcriptionally. The four main possible modifications are acetylation (yellow), methylation (black), phosphorylation (blue) or ubiquitination (grey). Most studied histone modifications occur on Lysines of H3, as they can be linked to active or repressed genes.

Dynamic chromatin states guide cells through developmental trajectories and provides an “epigenetic memory” to maintain the cell-type-specific transcriptional programs<sup>80,85</sup>. In differentiated cells, repressed chromatin states are represented by heterochromatin marked by tri-methylation of Lysine 9 of Histone 3 (H3K9me3)<sup>22,86</sup>, and DNA methylation within polycomb-group target regions enriched in tri-methylation of Lysine 27 of Histone 3 (H3K27me3)<sup>86,87</sup>. These repressive marks play a crucial role in regulating cell-type-specific transcription programs, helping to maintain cell identity<sup>22</sup>. In activated chromatin states, open chromatin is characterized by several marks, among which mono- or tri-methylation of Lysine 4, or tri-methylation of Lysine 36 of Histone 3 (H3K4me1 or H3K4me3 or H3K36me3) and acetylation of Lysine 9 or 27 of Histone 3 (H3K9ac or H3K27ac)<sup>86,87</sup>. H3K4me1 is a mark for enhancers, whereas H3K4me3 is a mark associated with promoters, and H3K36me3 is observed in the body of actively transcribed genes<sup>86,87</sup>. Besides these, H3K9ac is a mark for active promoters, while H3K27ac is a mark for active enhancers.

Histone marks are deposited by chromatin regulators that play vital roles in cellular reprogramming and oncogenesis<sup>79</sup>. There are three main categories of epigenetic regulators: readers, writers, and erasers (Figure 10). Writers are enzymes that add chemical groups to chromatin, such as histone acetyltransferases or DNA methyltransferases (DNMT)<sup>22,88</sup>. Erasers remove these epigenetic marks, as, for example, histone deacetylases (HDAC) or histone demethylases<sup>22,88</sup>. Readers recognize epigenetic modifications, are recruited to specific DNA regions, and act as effector and scaffold proteins, such as bromodomain proteins<sup>88</sup>.



**Figure 10: Mechanisms of histone modification by epigenetic chromatin regulators (modified from Panzeri et al.89).** Erasers (HDAC and histone demethylases) remove preexisting marks deposited on the chromatin; they work in collaboration with writers such as DNMT and histone acetyltransferases to maintain the chromatin state. Writers and readers interact with open chromatin, wherein active genes are associated with H3K4me1, H3K4me3, H3K36me3, H3K9ac, and H3K27ac marks. Erasers and writers interact with repressed chromatin, where repressed genes are associated with H3K9me3 and H3K27me3 marks<sup>83,86,87</sup>. Aberrant transcription factors can recruit epigenetic chromatin regulators in sarcomas to modulate target genes' accessibility and change chromatin state.

Epigenetic changes can be a prerequisite for transforming normal cells into cancer cells<sup>77</sup>. Perturbation of the epigenetic mechanisms can be achieved by deregulated DNA methylation, changes in histone post-translational modifications, or modification of chromatin remodeling complexes<sup>20</sup>. For example, hypermethylation of CpG islands in the promoters of tumor suppressor genes can lead to gene silencing and tumorigenesis<sup>90</sup>. Furthermore, aberrant DNA methylation patterns and overexpression of histone demethylases were observed in several tumor tissues<sup>90</sup>. Epigenetic mechanisms being reversible processes, they have a high therapeutic potential<sup>90</sup>. Epigenetic drugs, such as DNMT inhibitors, HDAC inhibitors, and histone demethylase inhibitors, have been recently identified and are currently under clinical trials<sup>90</sup>. Nevertheless, several challenges still need to be overcome, such as strong side-effects and choice of mono- or combination-therapy<sup>90</sup>.

In summary, DSRCT is a rare sarcoma of unknown origin that mainly arises in adolescent males. The aberrant expression of EWS-WT1 triggers the development of DSRCT. This aberrant protein's transcriptional activity broadly disrupts the expression profile of the precursor cell, leading to tumor initiation through mechanisms not yet fully understood. EWS-WT1 has two isoforms varying by the insertion of three amino acid (KTS) at its N-terminus. These isoforms may have different properties, but their distinct roles warrant further investigations. The whole panel of downstream effectors of EWS-WT1 is still unknown, therefore, their identification could help to understand DSRCT pathophysiology and ultimately identify new potential therapeutic targets.

## 2. Aim of my research

The Desmoplastic Small Round-Cell Tumor (DSRCT) is a sarcoma lacking a targeted treatment, and the oncogenic mechanisms induced by EWS-WT1 are still not fully understood. Based on the existing knowledge, EWS-WT1 could reconfigure the cells' epigenome and promote tumorigenicity. In keeping with this, in the present study, we aim to characterize the epigenetic mechanisms that contribute to the development of DSRCT. This could provide a better understanding of the pathogenic mechanisms and regulatory networks underlying the emergence of this tumor. A detailed understanding of the function of EWS-WT1 would provide important insight into DSRCT oncogenesis, and the identification of its direct target genes may lead to new therapeutic strategies.

### 2.1. Specific aims

#### 2.1.1. Define the chromatin signature and transcriptional landscape of DSRCT

Although DSRCT pathophysiology and its driver mutation, EWS-WT1, has been described, its global chromatin signature and epigenetic landscape are not well characterized. It is crucial to understand how the chromatin and the epigenome become rewired in DSRCT primary tumors providing an altered target gene expression program. Therefore, a comprehensive study on target gene expression and chromatin states in patient-derived primary tumor samples is highly valuable. One of our primary objectives is to profile and characterize DSRCT primary tumors for transcriptional and chromatin signatures. Additionally, a cellular model would provide functional opportunities to address questions on direct target genes and epigenetic mechanisms that drive EWS-WT1 tumorigenesis. Utilizing the DSRCT cell line, we can ask how EWS-WT1 changes the global transcriptome and epigenome by loss-of-function studies. This system would also help us to identify target genes associated with EWS-WT1. Further, 3D chromatin architecture studies would allow us to identify EWS-WT1 direct targets through elucidating EWS-WT1 chromatin loops. Thus, this objective would provide a complete picture of DSRCT chromatin signature, EWS-WT1 transcriptional mechanisms, and information on any potential druggable targets.

### 2.1.2. Elucidate the EWS-WT1 mediated oncogenic mechanisms in DSRCT

DSRCT cells can express two EWS-WT1 isoforms displaying distinct physiological properties. Therefore, it is essential to delineate the contribution of each isoform to the pathogenesis of DSRCT. As these isoforms are nearly identical in their amino acid composition, it is challenging to study endogenous genes other than by qualitative detection using RNA probes. Therefore, an ectopic expression system with distinct molecular tags would address fundamental questions on its role and understand molecular events that instigate DSRCT. We would be able to study the chromatin occupancy and epigenetic or transcriptional mechanisms associated with each isoform separately or together in these permissive cells.

### 2.1.3. Exploring the therapeutic potential of target genes/pathways in DSRCT

DSRCT is a rare sarcoma with a poor survival rate and lacking specific targeted treatment. Our previous two objectives will help elucidating EWS-WT1-mediated transcriptional mechanisms and direct target genes in DSRCT. Careful screening of those genes and pathways for potential druggable targets would be advisable. Any DSRCT-specific pathway or EWS-WT1 direct target gene that has been targeted in other cancer types with specific drugs would be potentially interesting candidates. Those candidate drugs could be tested in our *in vitro* cell-line model and further in *in vivo* xenograft models to validate the efficacy of those drugs for DSRCT. We would like to identify any potential targets that could be tested *in vitro* and *in vivo*. This would confirm a potential therapeutic benefit in the treatment of DSRCT to be proposed for preclinical investigation.

## 3. Results

### 3.1. Definition of the chromatin landscape and transcriptional profile of DSRCT

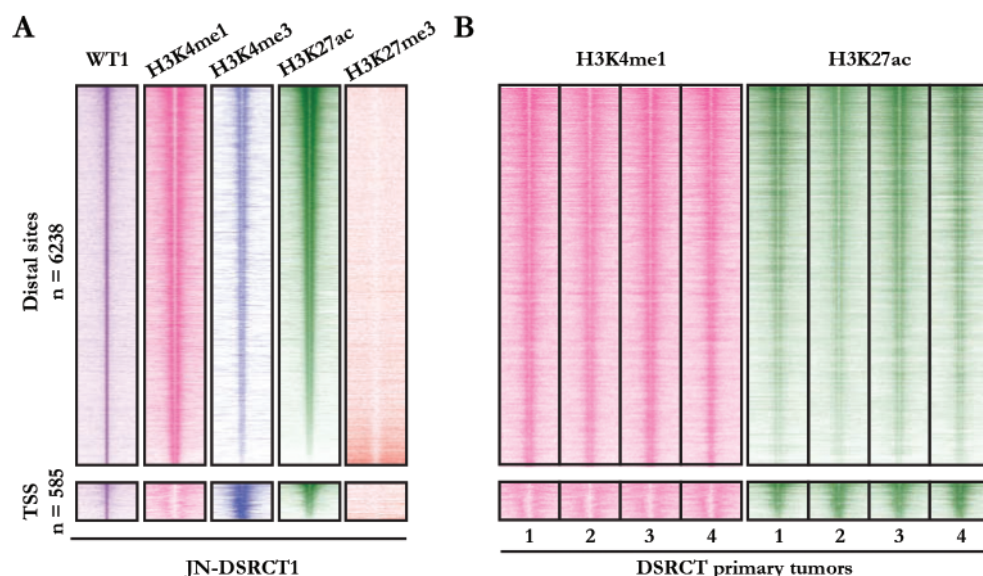
#### 3.1.1. Identification of EWS-WT1 target genes in DSRCT

The mechanisms by which EWS-WT1 regulates target genes is not fully understood. Some of the EWS-WT1 direct targets have been described previously, however the chromatin regulatory mechanisms that modulate target gene expression need to be elucidated.

##### 3.1.1.1. EWS-WT1 binding sites are enriched with active chromatin states

We followed a coordinated analysis of chromatin states at EWS-WT1 binding sites in four DSRCT primary tumors and the only available DSRCT cell-line, JN-DSRCT1, to characterize the mechanisms by which EWS-WT1 controls differentiation and oncogenic pathways. We profiled the genome-wide occupancy of EWS-WT1 and associated epigenetic marks to have an overview of the chromatin state. We first identified direct binding sites of endogenous EWS-WT1 in JN-DSRCT1, an established human DSRCT cell-line<sup>91</sup> lacking the expression of endogenous WT1<sup>65</sup>. This was realized by chromatin immunoprecipitation sequencing (ChIP-seq) with an antibody directed against the C-terminal part of WT1 contained in EWS-WT1. 6823 EWS-WT1 peaks were identified, and among them, 6238 were present at distal regulatory elements, whereas 585 were found at transcriptional start sites (TSS) regions in the JN-DSRCT1 cell-line (Figure 11.A).

Next, we mapped key histone modifications to link these binding sites to epigenetic states, including histone 3 lysine 4 monomethylation (H3K4me1), H3K4me3, H3K27ac, and H3K27me3. We found that the majority of EWS-WT1 binding sites were enriched for H3K4me1, a ubiquitous marker of cis-regulatory elements, and for H3K27ac, a more specific marker of enhancer activity (Figure 11.A). We also observed that when EWS-WT1 was bound to TSS regions, these sites were highly enriched for H3K4me3, a mark of transcriptional initiation (see the bottom part of Figure 11). Furthermore, we detected no overlap between EWS-WT1 binding sites and H3K27me3, a repressive modification mark deposited by polycomb repressive complexes in the JN-DSRCT1 cell-line.

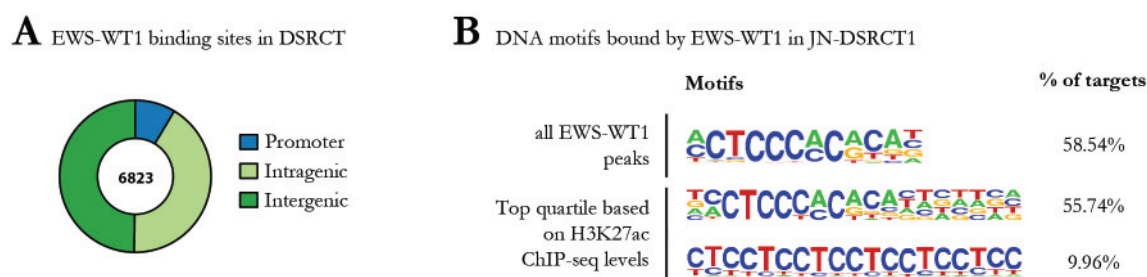


**Figure 11: Chromatin profiling highlight a subset of EWS-WT1 binding sites associated with higher levels of activity in DSRCT.** Heatmaps from ChIP-seq signal densities: (A) Heatmaps depict EWS-WT1, H3K4me1, H3K4me3, H3K27ac, and H3K27me3 signal intensities for 6238 EWS-WT1 bound distal regulatory elements in JN-DSRCT1 cell-line. (B) Heatmaps of H3K4me1 and H3K27ac in four DSRCT primary tumors. On the bottom part, heatmaps depicting signals for 585 EWS-WT1 peaks overlapping with TSS. The rows show a 10-kb window in each panel. They are centered on EWS-WT1 binding sites, ranked by overall signal intensities of H3K4me1 and H3K27ac.

To link these binding sites to an active epigenetic state, we mapped our EWS-WT1 peaks to H3K4me1 and H3K27ac modification marks in the four primary DSRCT tumors. Chromatin patterns over EWS-WT1 binding sites were highly concordant for both marks between the JN-DSRCT1 cell-line and the four primary tumors (Figure 11.B). We found that the majority of EWS-WT1 binding sites, previously identified, were also enriched for H3K4me1 and H3K27ac in the primary DSRCT tumors.

The majority of EWS-WT1 bound sites were located in intergenic and intragenic regions of the genome of JN-DSRCT1 (Figure 12.A). Our motif enrichment analysis indicated that the DNA binding motifs of EWS-WT1 in JN-DSRCT1 cells was the CTCCA/CC consensus sequence in 58.54% of sites (see Figure 12.B). We analyzed the DNA binding motifs based on the top quartile peaks of the enhancer H3K27ac histone mark and found two other binding motifs (a CTCCC cluster and TCC repeats).



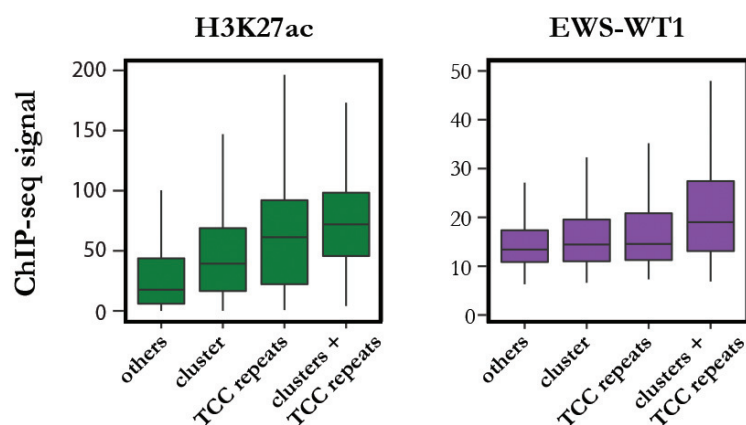


**Figure 12: Analysis of binding sites and motif of EWS-WT1 in the JN-DSRCT1 cells.** (A) The pie chart shows the genomic location of EWS-WT1 in the JN-DSRCT1 cells. In blue, the EWS-WT1 protein binds to promoter regions. In dark-green, the EWS-WT1 protein binds directly to the genes. (B) Model-based Analysis of ChIP-Seq (MACS)-called peaks of EWS-WT1 in the JN-DSRCT1 cells. Top panel: Motifs analysis of all EWS-WT1 peaks. Bottom panel: top quartile peaks based on the H3K27ac ChIP-seq levels.

20 years ago, Liu *et al.*<sup>5</sup> analyzed 14 DSRCT tumor samples by gel electrophoresis mobility shift assay (EMSA) and showed that EWS-WT1 had the ability to bind the GC and TC elements of the transcription factor Early Growth Response protein 1 (EGR-1) as well as WT1 consensus sequences. More precisely, they demonstrated that EWS-WT1 protein bound to the “TCCTCCTCCTCCTCTCC” sequence. This sequence is similar to our “CTCCTCCTCCTCCTCCTCC” sequence found in 9.96% of targets. Recently, Hingorani *et al.*<sup>65</sup> demonstrated by ChIP-seq that EWS-WT1 peaks were enriched particularly at intergenic and intronic regions of the genome of JN-DSRCT1 cell-line. Furthermore, their motif analysis showed that the EWS-WT1 peaks were enriched for known WT1 binding sequences: “(A/C)CTCCC(A/C)C(A/G)C” sequence and “G(C/T)GTGGG(T/C)G” sequence. These results are in concordance with our results in the JN-DSRCT1 cell-line, as we also showed particular binding of EWS-WT1 to genomic regions in these cells and our first CTCCC cluster was similar to the top motif identified by Hingorani *et al.*

Our ChIP-seq signals obtained from JN-DSRCT1 demonstrated a positive correlation between the EWS-WT1 signal and H3K27ac. CTCC clusters and TCC regions where EWS-WT1 bound were highly enriched in H3K27ac marks (Figure 13.A). Further, we asked whether EWS-WT1 is indispensable for the tumor-associated chromatin state by EWS-WT1 loss of function studies.

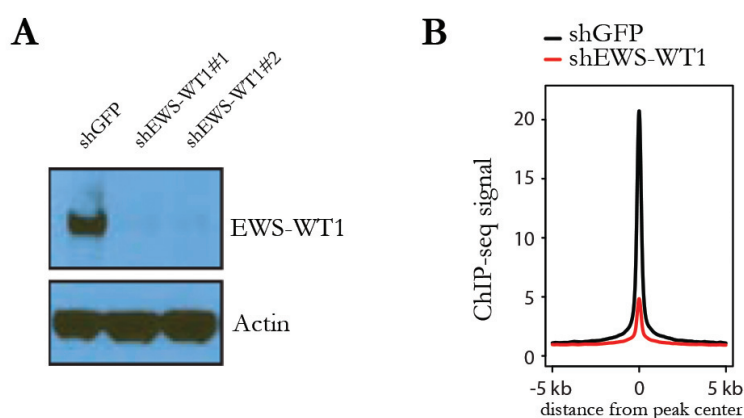




**Figure 13 : Positive correlation between EWS-WT1 and H3K27ac in JN-DSRCT1.** (A) Boxplots show ChIP-seq signal intensity for H3K27ac (left) and EWS-WT1 (right) at the four subsets of EWS-WT1 binding sites identified. Each category contained 5871, 609, 181, and 162 sites, respectively.

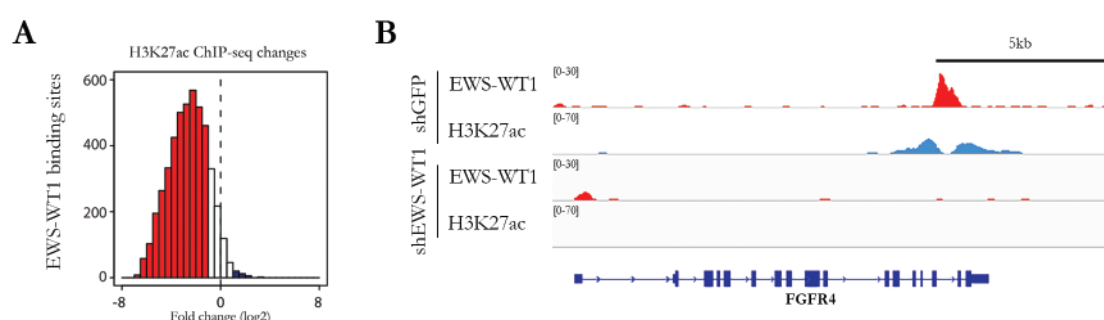
### 3.1.2. EWS-WT1 confers an active chromatin structure and regulates transcription at specific loci

To test the functional impact of EWS-WT1 on the enhancer activity of DSRCT, we depleted the fusion protein by small hairpin RNA (shRNA) and measured resulting changes in the chromatin state. We knocked-down the EWS-WT1 fusion gene in the JN-DSRCT1 cell-line and confirmed a reduction in EWS-WT1 protein levels by western blot and a decrease in EWS-WT1 ChIP-seq signals at its binding sites (Figure 14).



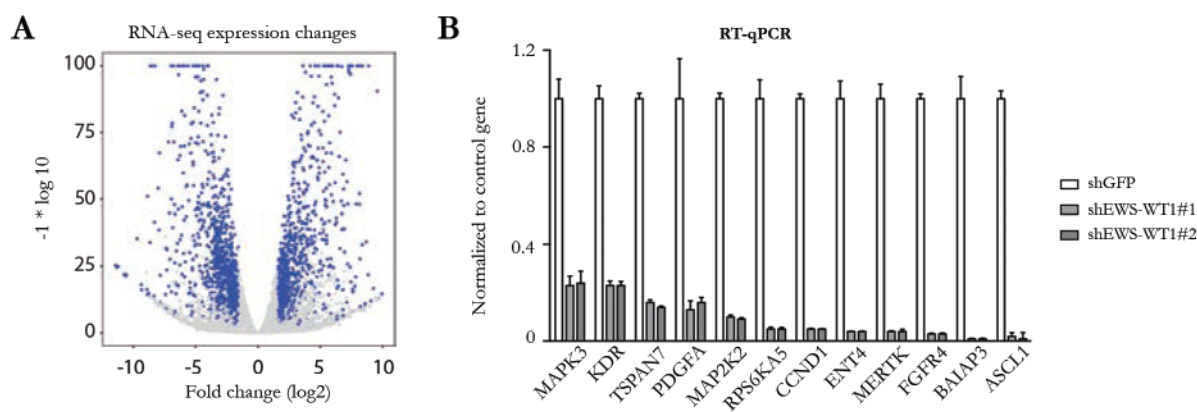
**Figure 14: Decrease of EWS-WT1 expression in JN-DSRCT1 cells upon knockdown.** (A) Western blot of JN-DSRCT1 cells lentivirally induced with a shRNA specifically targeting EWS-WT1. (B) ChIP-seq signal in JN-DSRCT1 cells upon EWS-WT1 knockdown. ChIP-seq signal shows a 5kb window centered on EWS-WT1 binding sites.

Moreover, our ChIP-seq analysis revealed that H3K27ac was decreased upon EWS-WT1 removal at the translocation binding sites. More specifically, loss of EWS-WT1 resulted in a significant reduction of H3K27ac marks. Classification of binding sites by changes in H3K27ac levels revealed that 4711 EWS-WT1 activated loci were downregulated by 2-fold or more upon EWS-WT1 knockdown (Figure 15.A). Furthermore, we also confirmed the EWS-WT1 knockdown and the specificity of the ChIP-seq signal for EWS-WT1 at the Fibroblast Growth Factor Receptor 4 (*FGFR4*) locus (Figure 15.B). We did not observe any EWS-WT1 or H3K27ac signal at *FGFR4* locus after EWS-WT1 knockdown. We reasoned that EWS-WT1 binding is required for the establishment of active enhancer state and H3K27acetylation.



**Figure 15: Acetylation changes in JN-DSRCT1 cells upon EWS-WT1 knockdown.** (A) Acetylation changes in JN-DSRCT1 cells upon EWS-WT1 knockdown. H3K27ac ChIP-seq changes at EWS-WT1 binding sites (4711 sites are downregulated upon KD, red=down 2-fold, blue =up 2-fold). (B) Example of ChIP tracks for JN-DSRCT1 cells lentivirally induced with a shRNA specifically targeting EWS-WT1 and its corresponding acetylation pattern around the *FGFR4* locus. In blue are tracks of H3K27ac for each condition. The last two tracks are the result of EWS-WT1 knockdown in JN-DSRCT1 cells. Upon EWS-WT1 removal, we observed a robust silencing of enhancers activity. These tracks demonstrated a decrease in binding to *FGFR4* at the EWS-WT1-bound site in JN-DSRCT1 cells when treated with the shEWS-WT1 knockdown.

Our RNA-sequencing analysis (RNA-seq) revealed that several genes were both down- or up-regulated upon EWS-WT1 knockdown (Figure 16.A). We observed a decrease in the expression of selected DSRCT targets by qPCR upon EWS-WT1 knockdown (Figure 16.B). Globally, the majority of EWS-WT1 targets were downregulated upon EWS-WT1 knockdown.



**Figure 16: RNA expression changes in JN-DSRCT1 cells upon EWS-WT1 knockdown.** (A) Volcano plot shows gene expression changes. Y-axis is capped at a minimum adjusted p-value of – 100. Genes selected for further analyses are shown in blue. (B) qPCR for selected EWS-WT1 target genes upon EWS-WT1 knockdown in JN-DSRCT1 cells.

To summarize, we evaluated the regulatory networks of DSRCT by epigenetically profiling four human primary DSRCT tumors and the JN-DSRCT1 cell-line, which showed high similarity in chromatin states at the translocation binding sites. We elucidated the role of EWS-WT1 in DSRCT as a key protein that establishment an active transcription state. Furthermore, we identified EWS-WT1specific chromatin signature. Now it will be interesting to analyze the early stages of oncogenesis in DSRCT, as this is still unknown.

### 3.2. EWS-WT1 mediated oncogenic mechanisms in DSRCT

A primary cell model is highly desirable to understand the early changes in regulatory networks underlying the pathogenesis of DSRCT. Several studies have demonstrated that ectopic expression of EWS-FLI1, FUS-CHOP or SS18-SSX in Mesenchymal Stem-Cells (MSCs) initiated the development of Ewing's sarcoma<sup>92-94</sup>, myxoid liposarcoma<sup>95</sup> or synovial sarcoma<sup>96</sup>, respectively. Therefore, we reasoned that a similar approach could be used to develop a primary cell model for DSRCT by ectopic expression of the EWS-WT1 fusion protein in MSCs, leading to the development of this malignancy.

#### 3.2.1. Derivation and characterization of human pediatric mesenchymal stem-cells (hpMSC)

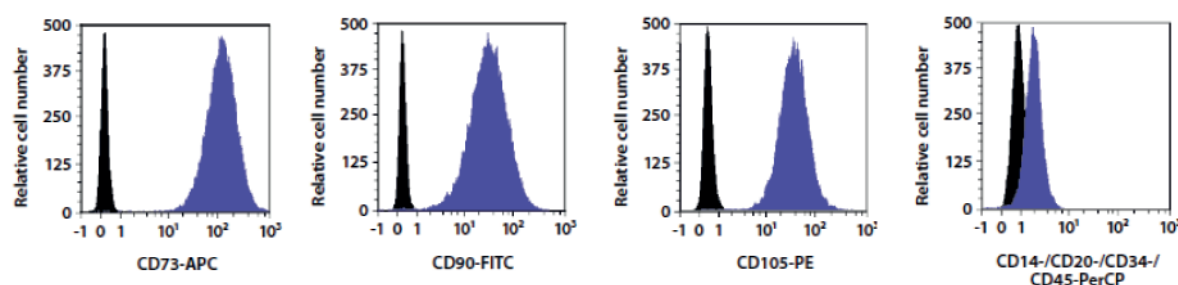
Mesenchymal stem-cells (MSCs) are bone marrow-derived fibroblast-like cells and are the progenitors of all cells constituting marrow stroma, bone, cartilage, tendons, and ligaments<sup>92,97-99</sup>. The multipotent MSCs can be easily isolated from donors, and expanded *in vitro*.

Primary human pediatric cells were isolated from healthy donors undergoing corrective bone surgery. We chose child donors since studies demonstrated an age-dependent decline of the self-renewal capacity of MSCs<sup>100</sup>. The cells were derived from donors of different sexes and ages (from 1 to 18 years old) at the Children's Hospital – Centre Hospitalier Universitaire Vaudois (CHUV) – Lausanne, according to the guidelines of the Cantonal commission of ethics on human research of the Canton de Vaud (Switzerland).

MSCs can be characterized using three minimal factors proposed by the International Society for Cellular Therapy (ISCT): adherence to the plastic surface under standard culture conditions, expression of cell-surface antigens, and *in vitro* differentiation capacity<sup>101-103</sup>.

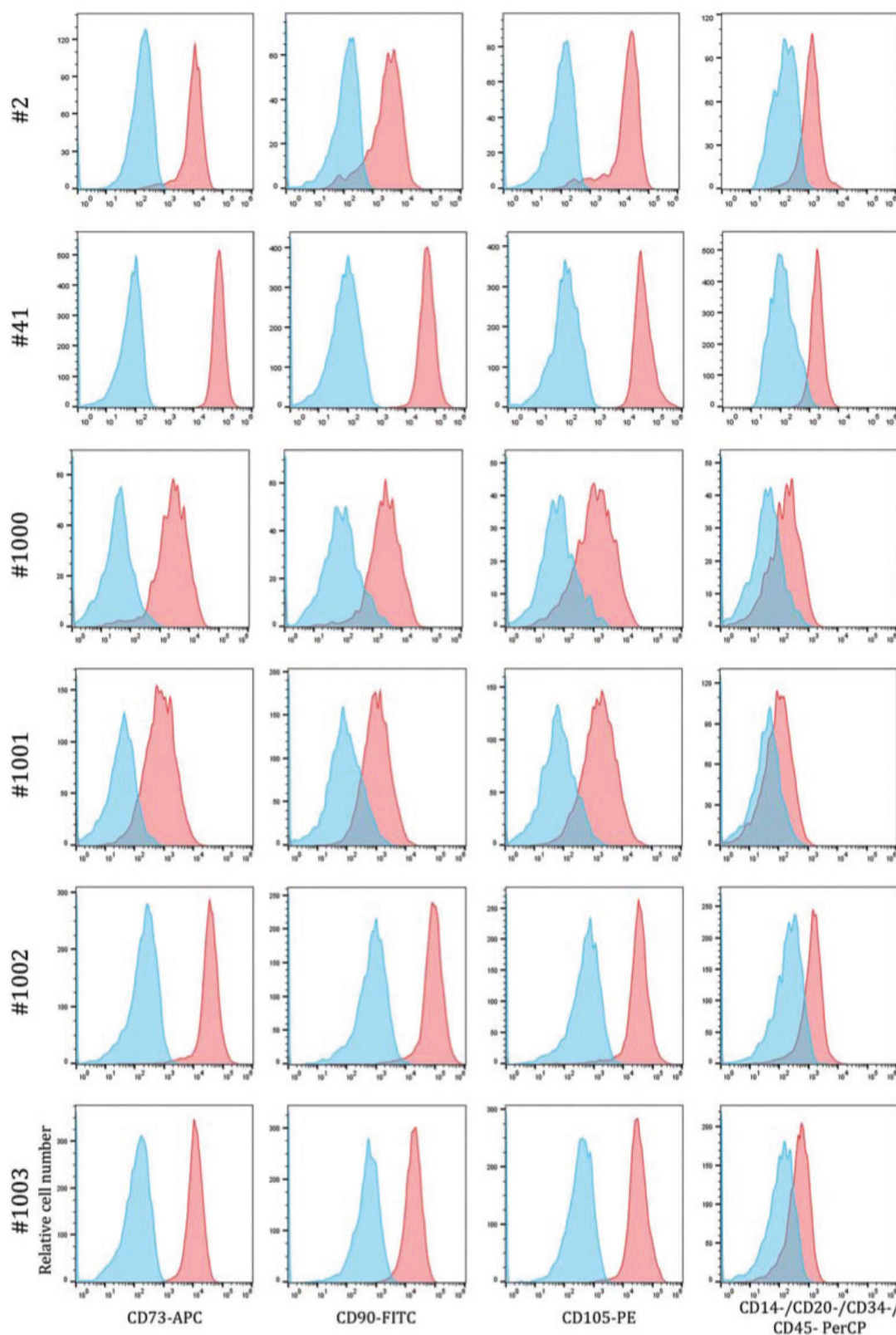
### 3.2.1.1. The cells generated from bone samples exhibit the three minimal factors characterizing MSCs

The primary cells generated from bone fragments were cultured *in vitro* and attached to the plates confirming the first ISCT requirement to be defined as MSC. MSCs express cell-surface markers such as Cluster of Differentiation 73 (CD73), CD90, and CD105 markers, and do not express hematopoietic lineage markers (such as CD14, CD20, CD34, and CD45 antigens)<sup>92,97,101</sup> (Figure 17).



**Figure 17: Representative immune profile of MSCs, by Fluorescence-Activated Cell Sorting (FACS - adapted from Miltenyi Biotec).** MSCs are positive for CD73, CD90, and CD105 markers and negative for the hematopoietic markers CD14, CD20, CD34, and CD45. Markers are coupled with APC (Allophycocyanin), FITC (fluorescein isothiocyanate), PE (phycoerythrin), or PerCP (Peridinin Chlorophyll Protein Complex). Therefore, it is easy to differentiate between unstained control in black and stained MSCs in blue.

All the derived primary cell cultures stained positively for CD73, CD90, and CD105 surface antigens and negatively for the hematopoietic markers. Some representative results are shown in Figure 18. Small differences can be observed between cell batches, these differences may be due to the age of the donor or site of MSCs extraction. The expression of these specific cell-surface antigens on the cells confirmed that these latter were immune-phenotypically similar to MSCs.

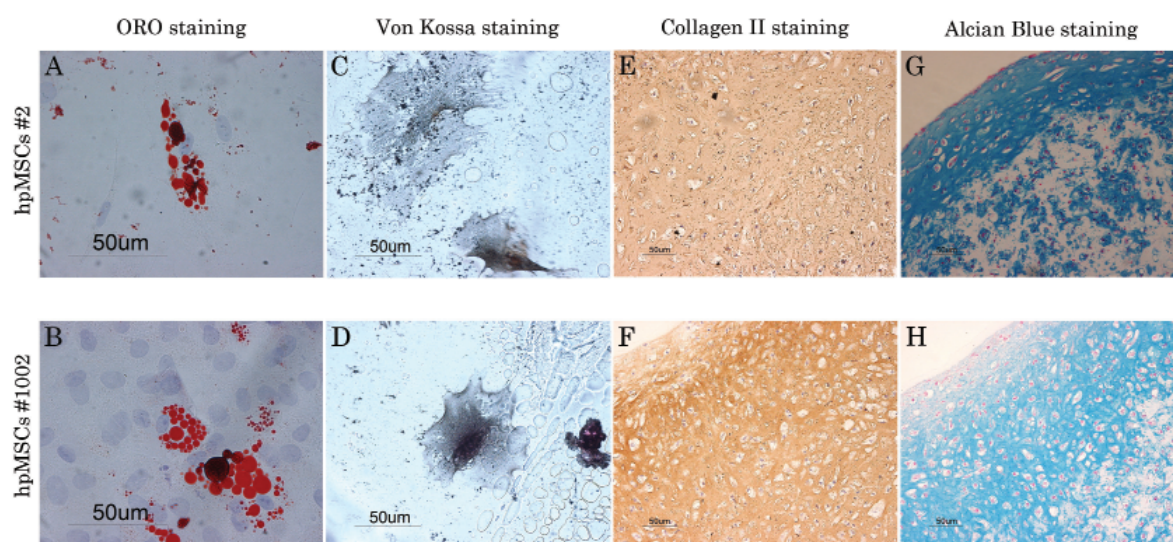


**Figure 18: Primary cells generated are immunophenotypically similar to MSCs.** Representative results of the FACS immunophenotyping. All batches (#2, #41, #1000, #1001, #1002 and #1003) demonstrated positive staining for the canonical MSCs markers CD73, CD90, and CD105, and negative staining for the hematopoietic markers CD14, CD20, CD34 and CD45. This confirms that the cells extracted are immune-phenotypically similar and share a common phenotype with normal MSCs. Sample: red shaded area. Control un-stained: blue shaded area.



The differentiation capacity of MSCs can be assessed by their tri-lineages mesodermal differentiation capacities, which include adipocytes, osteoblasts, and chondrocytes differentiation<sup>92,99,101</sup>. Hereafter, we investigated the *in vitro* differentiation capacities of the isolated cells. To assess the cells' differentiation potential towards the adipogenic, osteogenic, or chondrogenic lineages, cells were grown under specific conditions.

All cell batches tested were able to differentiate *in vitro* into these three different mesodermal lineages. Two representative examples of these results are shown in Figure 19. In these assays, not all cells within the batches displayed a complete tri-lineage differentiation capacity. As shown in Figure 19, following adipogenic differentiation, the presence of lipids was confirmed by Oil Red O (ORO) staining, which stained lipids in red. Then, following osteogenic differentiation, the presence of calcium deposits was confirmed by Von Kossa staining. Finally, positive collagen II and alcian blue staining demonstrated the expression of chondrogenic proteins, following chondrogenic differentiation, including type II collagen and aggrecan, respectively. The cells' *in vitro* differentiation capacity confirmed the last minimal ISCT factor to characterize our derived cells as MSCs. Therefore, hereafter those cells will be called human pediatric MSCs (hpMSCs).

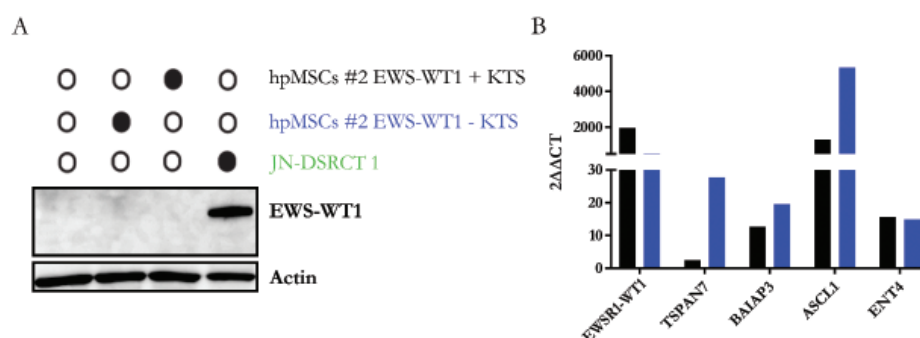


**Figure 19: Primary cells generated exhibit similar differentiation potential than MSCs *in vitro*.** Batches #2 (A, C, E, and G) and #1002 (B, D, F, and F). (A-B) Adipogenic differentiation confirmed with ORO staining. Lipids were stained in red, and the nuclei counterstained with Hematoxylin. 60x magnification. (C-D) Osteogenic differentiation confirmed with Von Kossa staining revealed the presence of calcium deposits in grey. Morphology changes were observed, differentiated cells exhibited a cuboidal morphology. 40x magnification. Chondrogenic differentiation confirmed by positive collagen II (E-F) and alcian blue (G-H) staining. Expression of chondrogenic proteins (type II collagen and aggrecan) following chondrogenic differentiation. 20x magnification. Scale bar 50 µm.

### 3.2.2. EWS-WT1 expression recapitulates early stages of DSRCT oncogenesis

It was suggested that MSCs could be the cellular origin of some translocation-associated solid tumors, such as the Ewing's sarcoma and DSRCT. Studies demonstrated that EWS-FLI1 expression could transform MSCs. This expression induced several Ewing's sarcoma hallmarks, including selected EWS-FLI1 target genes<sup>92-94</sup>. Therefore, we reasoned that a similar approach could be used to develop a primary cell model for DSRCT by ectopic expression of the EWS-WT1 fusion protein in hpMSCs, and recapitulate some of the most important DSRCT molecular hallmarks, leading to the development of this malignancy. Since EWS-WT1 has two isoforms, it is essential to delineate the contribution of each isoforms to the pathogenesis of DSRCT. As these isoforms differ by three amino acid, we transduced hpMSCs with each EWS-WT1 isoform constructs in parallel. We checked for the expression of the *EWSR1-WT1* transcripts and *de-novo* activation of their specific target genes as some genes are known to be direct EWS-WT1 targets and be implicated in DSRCT development<sup>27</sup>.

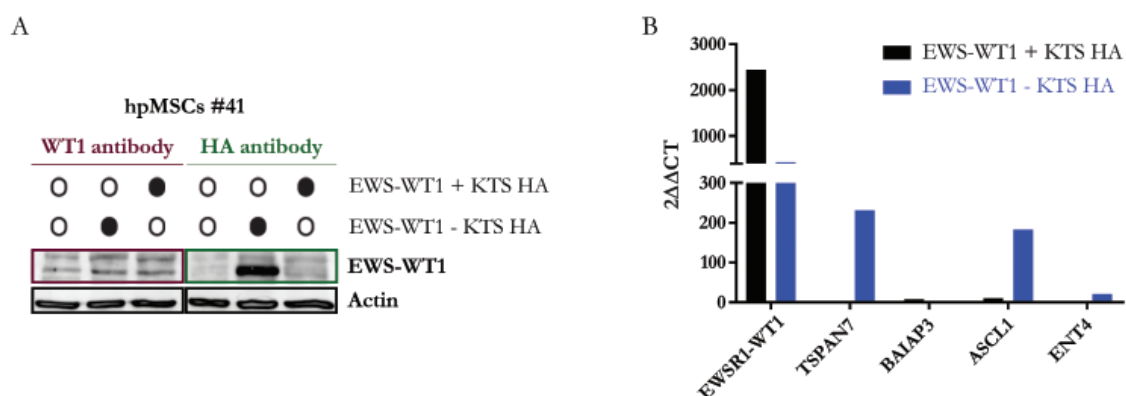
hpMSCs expressing EWS-WT1 exhibit a different morphology compared to the fibroblast-like morphology of control cells. We confirmed high mRNA levels for both *EWSR1-WT1* isoforms, as well as a clear induction in the expression of select DSRCT target genes (*TSPAN7* or *TALLA1*, *BALAP3*, *ASCL1*, and *ENT4*, see Figure 20.B). Unfortunately, despite the high *EWSR1-WT1* transcript levels detected by qPCR, we were not able to detect EWS-WT1 at the protein level in hpMSCs (Figure 20.A). However, EWS-WT1 was detected in the positive control, JN-DSRCT1, by western blot.



**Figure 20: hpMSCs #2 lentivirally transduced with EWS-WT1 fail to express the aberrant protein, but induce transcription of some DSRCT targets.** (A) Western blot of hpMSCs #2 pLIV Control, hpMSCs #2 pLIV EWS-WT1 + KTS (black), hpMSCs #2 pLIV EWS-WT1 - KTS (blue) and JN-DSRCT1 cells (green). Equivalent protein loading was confirmed by actin (~42kDa). The EWS-WT1 fusion protein (~64kDa) was detected in the JN-DSRCT1 cell-line (internal positive control) but not in hpMSCs #2 samples, confirming the capacity of the WT1 antibody to recognize the EWS-WT1 protein. The WT1 antibody targeting the C-terminal part of the protein was used, as there is no endogenous WT1 in hpMSCs nor in JN-DSRCT1 cells<sup>65</sup>. (B) Relative mRNA expression levels of hpMSCs #2 pLIV EWS-WT1 + KTS (black) and hpMSCs #2 pLIV EWS-WT1 - KTS (blue) were analyzed using the  $2^{-\Delta\Delta C_t}$  method; the data were first normalized to the corresponding GAPDH levels and then calculated in reference to hpMSCs #2 pLIV Control infected with a lentivirus containing an empty vector.



The lack of EWS-WT1 protein detection may be linked to a protection mechanism from hpMSCs that stops producing the aberrant protein, or a 3D protein conformation artefact in these cells, and thereby masking the epitope for the WT1 antibody. Therefore, we added an epitope tag, either the parainfluenza virus 5 tag (V5) or the hemagglutinin (HA) tag to the C-terminus of the EWS-WT1  $\pm$  KTS fusion proteins, and repeated the experiment. We detected the HA-tag in the EWS-WT1+KTS isoform, but not its  $-$ KTS counterpart (Figure 21.A). The discrepancy in the detection of two different isoforms is likely due to the differences in their protein folding as described before<sup>104</sup>. Moreover, the WT1 antibody failed to detect EWS-WT1 fusion protein in hpMSC samples by western blot (Figure 21.A). As before, high mRNA levels for both *EWSR1-WT1* isoforms and selected target genes were detected (Figure 21.B). We also noticed a variation in the induction of DSRCT targets between hpMSCs batches, we assumed that this may be due to the epigenetic background among the hpMSCs isolated.

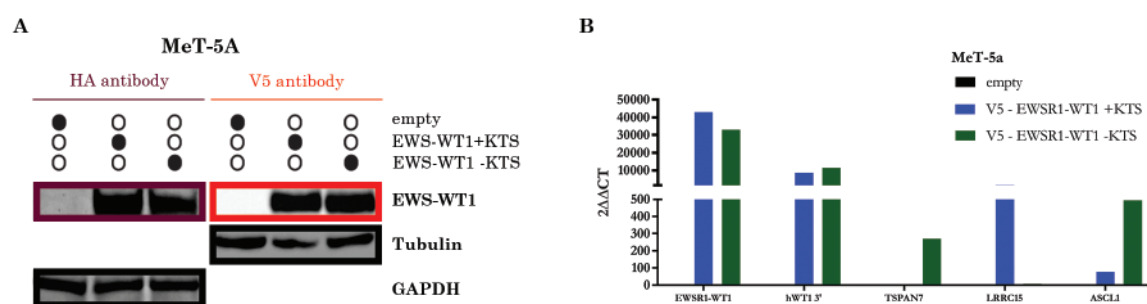


**Figure 21: HA-tagged EWS-WT1 induce transcription of some DSRCT targets, but fail to expression the EWS-WT1  $-$ KTS isoform in hpMSCs #41.** (A) Western blot was performed to analyze HA and EWS-WT1 protein expression in hpMSCs #41. hpMSCs #41 pLIV Control, hpMSCs #41 pLIV EWS-WT1 + KTS (black), and hpMSCs #41 pLIV EWS-WT1 - KTS (blue). The EWS-WT1 fusion protein was not detected with the WT1 antibody in any of the hpMSCs samples. HA tag was detected at roughly 64kDa in hpMSCs #41 pLIV EWS-WT1+KTS, but not in  $-$ KTS, corresponding to the canonical size of the EWS-WT1 isoform. Equivalent protein loading was confirmed by actin (42kDa). As actin was detected in all samples, the EWS-WT1 protein's absence is not linked to a technical problem. (B) Relative mRNA expression levels of hpMSCs #41 pLIV EWS-WT1 + KTS (black), and hpMSCs #41 pLIV EWS-WT1 - KTS (blue) HA-tagged were analyzed using the  $2^{-\Delta\Delta C_t}$  method; the data were first normalized to the corresponding GAPDH levels, and then calculated in reference to hpMSCs #41 pLIV Control infected with a lentivirus containing an empty vector.

Though ectopic expression of EWS-WT1 in hpMSCs induced selected target gene activation, epigenetic background among the hpMSCs isolated may change the results. Furthermore, difficulties in isoform-specific protein detection in hpMSCs are probably due to altered folding and associated morphological changes. Therefore, we reasoned hpMSCs might not represent an ideal primary cellular

model for DSRCT. We decided to change our model to a primary SV40 immortalized mesothelial cell-line (MeT-5A), which could be closer to the DSRCT cell of origin. The MeT-5A cell-line is commercially available (ATCC® CRL-9444™). ATCC ensures that the cells are not tumorigenic when injected into immunosuppressed mice and we confirmed it by injecting these cells into immunodeficient NOD/SCID c Gamma (NSG) mice.

To address how EWS-WT1 isoforms contribute to the initiation of DSRCT, we analyzed the properties of EWS-WT1 isoforms in the mesothelial cell-line MeT-5A, which we anticipate may recapitulate the early stages of DSRCT development<sup>27</sup>. Additionally, similar to DSRCT tumors, the MeT-5a cell-line does not express wild-type endogenous WT1. In MeT-5A cells, we adopted an inducible strategy, where expression of V5 or HA-tagged EWS-WT1 +KTS or EWS-WT1 -KTS proteins were modulated by the addition of doxycycline. This approach allowed us to successfully circumvent the toxic effect of the translocation expression in MeT-5A cells. We confirmed high mRNA levels for both *EWSR1-WT1* isoforms, as well as *de novo* activation of selected known target genes (Figure 22.B). By western blot, we were able to detect EWS-WT1 only with the V5- or HA-tags but not with the WT1 antibody (Figure 22.A).



**Figure 22: HA or V5-tagged EWS-WT1 proteins were detected in MeT-5A cells and corresponding EWS-WT1 targets were transcript.** (A) Western blot of MeT-5A cells infected with pIND20 empty, pIND20 EWS-WT1 + KTS and pIND20 EWS-WT1 – KTS virus, either HA- or V5-tagged. Equivalent protein loading was confirmed by tubulin or GAPDH. The EWS-WT1 fusion protein was detected in the MeT-5A cells with a HA (violet) and V5 antibody (orange). (B) Relative mRNA expression levels of MeT-5A pIND20 V5 EWSR1-WT1 + KTS (blue) and MeT-5A pIND20 V5 EWSR1-WT1 - KTS (green) were analyzed using the  $2^{-\Delta\Delta C_t}$  method; the data were first normalized to the corresponding GAPDH levels, and then calculated in reference to MeT-5A infected with a lentivirus containing an empty vector. TSPAN7 and ASCL1 are known targets of the EWS-WT1 –KTS isoform, and they were only highly expressed in cells where this isoform was induced. LRRC15 is the only known target specific for the EWS-WT1 +KTS isoform, therefore it was expected to be only induced in cells expressing this isoform.

We concluded that MeT-5A cells could produce EWS-WT1 protein in sufficient amounts to activate known DSRCT targets and the isoform-specific proteins could be detected using the HA- or V5-tags. Therefore, we selected this model for further analyses on early stages of DSRCT oncogenesis.

### **3.2.3. EWS-WT1 isoforms regulate unique target set of genes in DSRCT**








Expression of EWS-WT1 in MeT-5A cells recapitulate expression of known targets of DSRCT, therefore it was logical to pursue this investigation with the epigenetic profiling of EWS-WT1 in order to establish the corresponding chromatin remodeling events genome-wide, but also other properties of the EWS-WT1 isoforms.

#### **3.2.3.1. EWS-WT1 isoforms display different binding affinities to the chromatin**

We planned to assess whether the two isoforms share DNA binding-sites or tend to occupy unique genomic regions and, therefore, regulate specific direct targets. We followed a coordinated analysis of chromatin states at EWS-WT1 binding sites in our DSRCT cell model to characterize the mechanisms by which each EWS-WT1 isoform controls the transformation of a non-tumorigenic cell. To identify the corresponding underlying DNA motifs of each isoform, we profiled the genome-wide occupancy of HA- or V5-tagged EWS-WT1 isoforms in MeT-5A cells.

We first identified direct binding sites of EWS-WT1 isoforms in MeT-5A cells, stably expressing each isoform. This was realized by ChIP-seq with an antibody directed against the V5 or HA- tags linked to EWS-WT1. The motif enrichment analyses indicated that the DNA binding motifs of both EWS-WT1 isoforms shared sequence similarities that were represented by a CTCCC clusters and TCC repeats (more details in Table 4). We observed that the top motif for the EWS-WT1 –KTS isoform was the “CTCCC(A/C)C” sequence found in 82.94% of targets. Our previous DNA binding motifs analysis revealed that the top DNA binding motif for EWS-WT1 in JN-DSRCT1 was somewhat similar to the top DNA binding motifs of the EWS-WT1 –KTS isoform. The second top motif for the EWS-WT1 –KTS isoform identified was the “CCNCNCACCCCCAC(A/C)” sequence found in 45% of targets. The last sequence identified is a TCC repeats (5x) motif in 12% of targets.

The top motif identified for the EWS-WT1 +KTS isoform was the “(A/C)CTCCC” sequence found in 60.84% of targets. The second top motif identified for this isoform was a “AAGTAAACA” sequence in 38.22% of targets, followed by a C/T repeat sequence in 27.78% of targets. Further, we also identified a TTCC repeat motif in 35.93% of targets. In 2003, Reynolds *et al.*<sup>72</sup> expressed the EWS-WT1 +KTS isoform in U2OS osteosarcoma cells and identified, by EMSA, the GGAGG(A/G) consensus sequence as the optimal binding site for EWS-WT1 +KTS isoform. Our results did not confirm the identification of this EWS-WT1+KTS responsive element sequence.

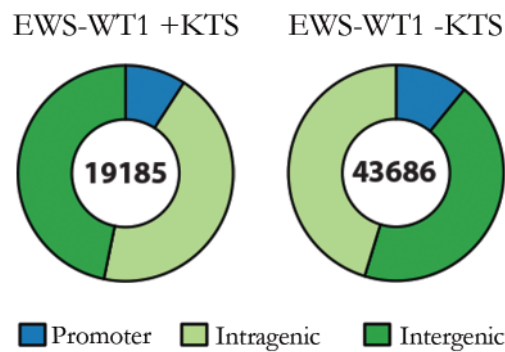
EWS-WT1 -KTS		EWS-WT1 +KTS	
Motif	% targets	Motif	% targets
	82.94		60.84
	44.93		38.22
	12.03		35.93
			27.78

**Table 4: Motif enrichment analysis of EWS-WT1 isoforms in MeT-5A cells show similar but specific binding motifs.** MACS-called peaks results were displayed as a sequence logo, representing the position weight matrix (PWM) of the relative preference of EWS-WT1 for each base in the binding site. Motifs analysis of EWS-WT1 -KTS isoform pointed binding at 82.94% to a CTCCC(A/C)C sequence. Motifs analysis of EWS-WT1 +KTS isoform indicated binding at 60.84% to a CTCCCT sequence. The CTCCC motif is similar to the one previously identified in the JN-DSRCT1 cell-line.

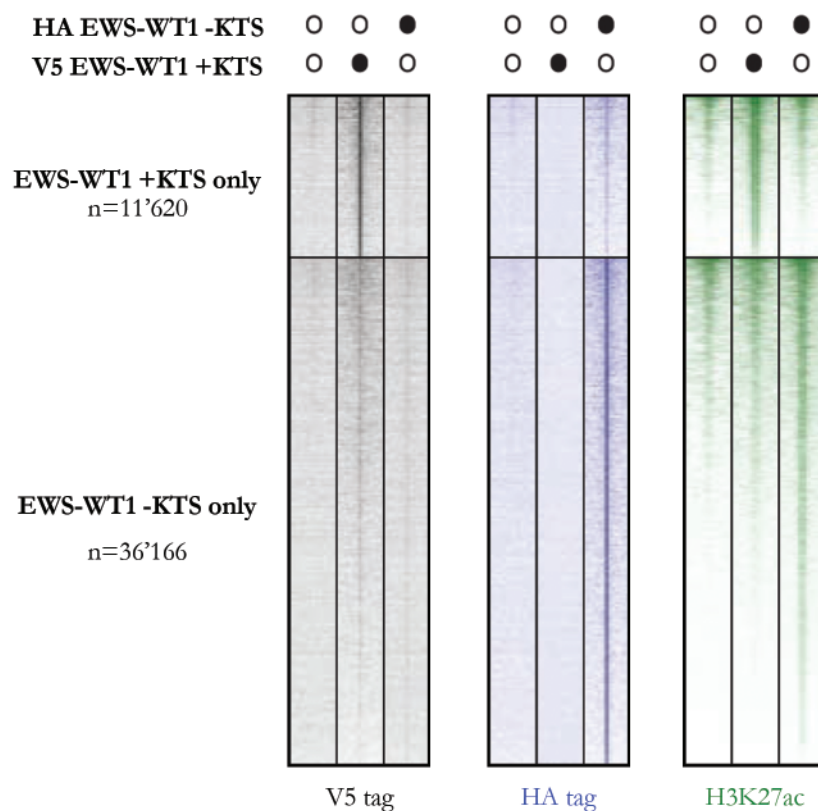
We can conclude that the top motif identified in both EWS-WT1 isoforms is similar to the top motif we identified earlier in the JN-DSRCT1 cell-line. Furthermore, we demonstrated that each EWS-WT1 isoform has specific DNA binding motifs giving us a hint to specific regulation of different targets by each EWS-WT1 isoform.

To understand further the role of EWS-WT1 isoforms, we assessed the genomic distribution of them. 19'185 EWS-WT1 peaks were present in MeT-5A cells stably expressing EWS-WT1 +KTS isoform, whereas 43'686 EWS-WT1 peaks were present in MeT-5A cells stably expressing EWS-WT1 -KTS isoform (Figure 23). Therefore, the results indicated that the EWS-WT1 -KTS isoform binds two times more to the DNA than the EWS-WT1 +KTS isoform. Although both EWS-WT1 isoforms bound peaks were located mainly at intergenic and intragenic regions (Figure 24), the EWS-WT1 -KTS isoform tended to bind more to promoter regions than the other isoform. Our results indicated that the distribution of these

binding sites displayed a high degree of similarity to the EWS-WT1 binding profile observed JN-DSRCT1 cells. Furthermore, following our hypothesis, we also observed that each EWS-WT1 isoform displayed distinct genomic target sites with a minority of binding-sites (7'565) shared by both isoforms.



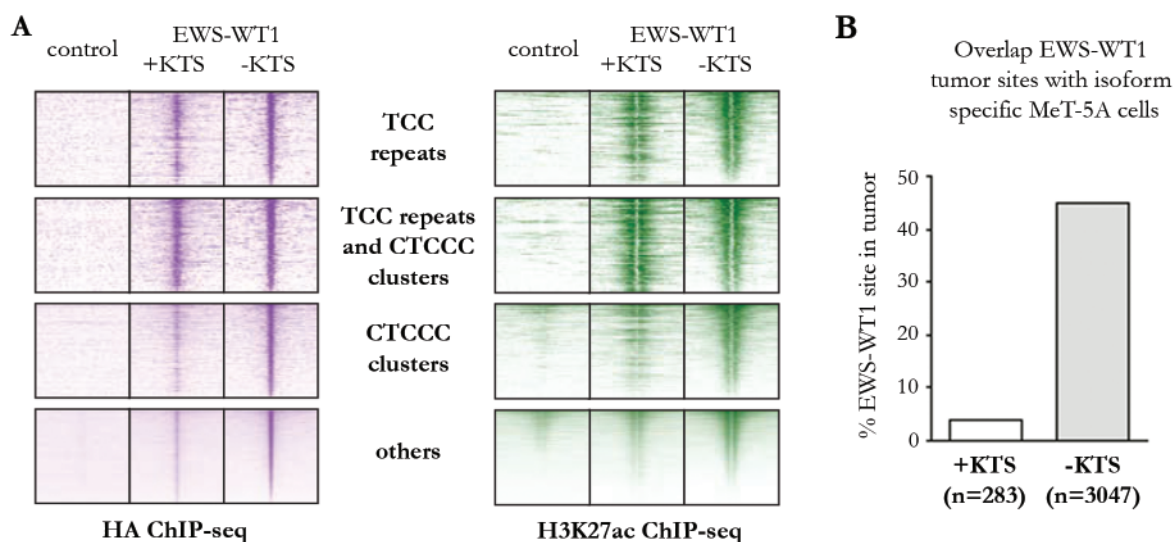
**Figure 23: Genomic location of EWS-WT1 isoforms in the MeT-5A cells.** In blue, the EWS-WT1 protein binds to promotor regions. In green, the EWS-WT1 protein binds directly to the genes. The EWS-WT1 – KTS isoform bound to 43'686 genomic regions, whereas the EWS-WT1 +KTS isoform recognized only 19'185 binding sites (2.2 times less).



**Figure 24: Chromatin profiling highlights a subset of EWS-WT1 binding sites are associated with high levels of acetylation in MeT-5A cells expressing EWS-WT1 isoforms.** ChIP-seq signal heatmaps of EWS-WT1 (black and blue) and H3K27ac (green). ChIP-seq signal densities in MeT-5A cells expressing HA EWS-WT1 –KTS, V5 EWS-WT1 –KTS and empty control. 5-kb windows in each panel are centered on EWS-WT1 binding-sites.

Following the motif enrichment results of EWS-WT1, we decided to separate EWS-WT1 peaks based on their chromatin's motif location (see Figure 25.A). Unsurprisingly, both EWS-WT1 isoform bound strongly to the DNA at TCC repeats and CTCCC clusters. The EWS-WT1 -KTS isoform seemed to bound more at these chromatin regions than the EWS-WT1 +KTS isoform, but these results could be influenced by the difference in the number of binding-sites detected for each isoform.

To link these binding sites to an active epigenetic state, we mapped the EWS-WT1 peaks to H3K27ac. We found that the majority of EWS-WT1 binding sites were enriched in H3K27ac, a marker of enhancer activity (Figure 24). Further, we decided to separate the EWS-WT1 peaks, based on the CTCCC clusters and TCC repeats motif identified earlier in JN-DSRCT1 cells. We noticed that the majority of EWS-WT1 binding sites at these regions were enriched in H3K27ac (Figure 25.A).



**Figure 25: Chromatin profile of EWS-WT1 isoforms in our cellular DSRCT model separated by motifs identified in JN-DSRCT1 and positive correlation between the primary DSRCT tumors and our model.** (A) Heatmaps of EWS-WT1 (violet) and H3K27ac (green). ChIP-seq signal densities in MeT-5A cells expressing HA-tagged EWS-WT1 isoforms. 10-kb windows in each panel are centered on EWS-WT1 binding-sites, and separated into TCC repeats and CTCCC clusters. (B) Overlap of EWS-WT1 tumor binding sites with EWS-WT1 +KTS (283) and EWS-WT1 -KTS isoform (3047) specific sites from MeT-5A cells.

Next, we compared the epigenetic landscape of the two EWS-WT1 isoforms in MeT-5A with the primary DSRCT tumors, and observed an overlap of 45% of EWS-WT1-KTS binding sites and a smaller portion for the EWS-WT1 +KTS isoform (Figure 25.B). The low overlap between MeT-5A cells and primary DSRCT tumors may be due to the high number of changes occurring within the cell at the early stages of development that may not appear in fully formed tumors. Another possibility for



the low overlap, particularly with the EWS-WT1 +KTS isoform, is linked to the low number of peaks identified previously for this isoform.

Based on these results, we conclude that the EWS-WT1 -KTS isoform displays a higher DNA accessibility potential than the EWS-WT1 +KTS isoform. In wild-type WT1, studies have shown that the WT1 +KTS isoform is a RNA binding protein<sup>54,55</sup>. It is therefore interesting to analyze how each isoform contribute to the development of DSRCT, based on the functional properties of EWS-WT1 isoforms.

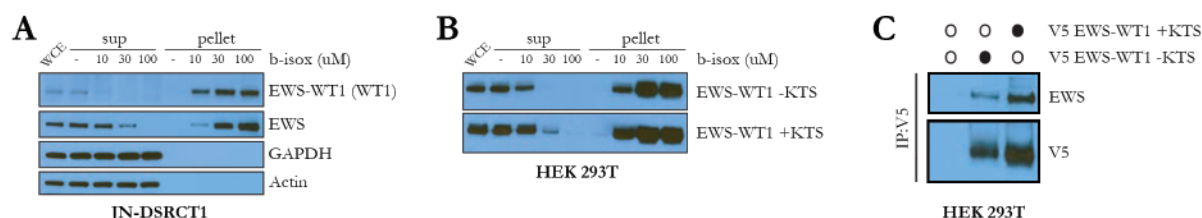
### 3.2.3.2. EWS-WT1 isoforms multimerize and exhibit phase-transition properties

Membrane-less compartments are formed through forces pushing two liquids apart, like oil on top of water, called liquid-liquid phase separation (LLPS)<sup>41,105-108</sup>. Some of these compartments, exhibiting droplet-like properties, are present in the nucleus (e.g., nucleoli, Cajal bodies, and nuclear speckles) and in the cytoplasm (e.g., stress granules, centrioles, germ granules and processing bodies or P-granules)<sup>41,106,109</sup>. The aggregation of supramolecules into membrane-less organelles can be driven by intrinsically disordered regions with low-complexity sequences<sup>41,108,109</sup>. These low-complexity sequences, found in prion-like domains (PrLDs), are sufficient for LLPS processes<sup>41,109-111</sup>. EWS belongs to the FET family of proteins that have been shown to mediate LLPS, multimerization, and pathological protein aggregation<sup>112-115</sup>. Recently, Boulay *et al.* confirmed the phase separation potential of EWS-FLI1 in the setting of Ewing's sarcoma cells<sup>116</sup>. Therefore, we would like to investigate the LLPS properties of EWS-WT1 in the setting of DSRCT.

In collaboration with Prof. Rivera's laboratory from the Massachusetts General Hospital (MGH, Boston – USA), we assessed the phase-transition potential of EWS-WT1, by testing its ability to precipitate in the presence of biotinylated isoxazole (b-isox). The b-isox chemical property is used to selectively precipitate RNA and proteins with particular conformations in cell-free reactions<sup>42</sup>. Thus, b-isox precipitates proteins containing low complexity domains, such as PrLD, known to form membrane-less compartments<sup>42</sup>. The results revealed that EWS-WT1 isoforms exhibited robust precipitation in HEK 293T cells stably expressing EWS-WT1 +KTS or EWS-WT1 -KTS (Figure 26.B). Furthermore, we confirmed the phase transition potential of EWS-WT1 in JN-DSRCT1, as the b-isox experiments attested that EWS-WT1 exhibits robust, concentration-dependent precipitation, higher than wild-type



EWS in JN-DSRCT1 (see Figure 26.A). Immunoprecipitation assay of V5-tagged EWS-WT1 isoforms displayed a strong interaction with endogenous EWS (Figure 26.C). We observed that EWS-WT1 +KTS has a stronger interaction with EWS compared to the -KTS isoform, although the +KTS isoform displays a lower binding to the DNA.



**Figure 26: EWS-WT1 can precipitate and recruit endogenous EWSR1.** (A) EWS-WT1 precipitated upon treatment with b-isox in JN-DSRCT1 cell lysates. Endogenous wild-type EWS strongly interacted with EWS-WT1 in JN DSRCT1 cells. (B) EWS-WT1 +KTS and EWS-WT1 -KTS isoforms precipitated upon treatment with b-isox in HEK 293T cell lysates. (C) Immunoblots of anti-V5 immunoprecipitates from HEK 293T cells transfected with either control vector, V5 EWS-WT1 +KTS, or V5 EWS-WT1 -KTS isoforms. The V5 EWS-WT1 -KTS isoform seems to bind more endogenous EWS.

Boulay *et al.* showed the multimerization potential of EWS-FLI1. Therefore, we investigated this property in EWS-WT1. To assess the dimerization potential of EWS-WT1, we performed a reciprocal immunoprecipitation experiments using overexpressed HA- and V5-tagged EWS-WT1 isoforms in HEK 293T cells (Figure 27). The results attested that both EWS-WT1 isoforms were able to homodimerize, but also that the EWS-WT1 isoforms could heterodimerize.



**Figure 27: The fusion protein EWS-WT1 can homo and hetero-dimerize.** Co-immunoprecipitation experiments using anti-V5 and HA antibodies in transiently transfected HEK 293T cells co-expressing HA-tagged and V5-tagged mutant isoforms. (A) HA-tagged and V5-tagged EWS-WT1 -KTS co-expressed in HEK 293T cells. (B) HA-tagged and V5-tagged EWS-WT1 +KTS co-expressed in HEK 293T cells. (C) V5-tagged EWS-WT1 +KTS and HA-tagged EWS-WT1 -KTS isoforms co-expressed in HEK 293T cells.

### 3.2.3.3. EWS-WT1 regulate unique target set of genes in DSRCT

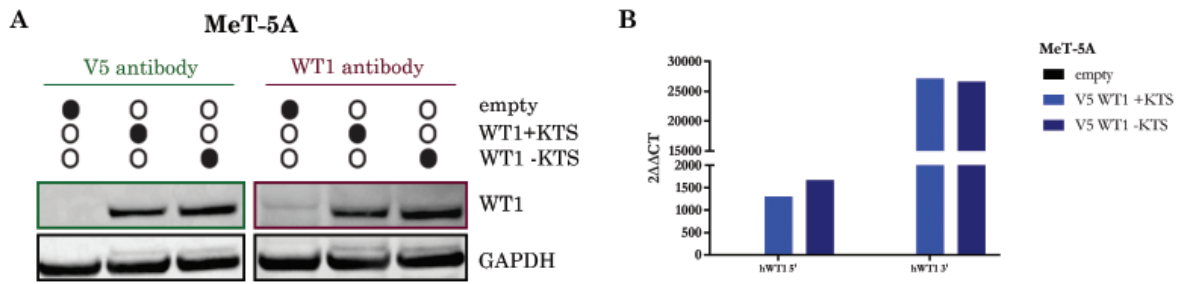
Boulay *et al.* reported the neomorphic properties of EWS-FLI1 compared to FLI1 wild-type in Ewing's sarcoma<sup>116</sup>. We therefore asked whether wild-type WT1 isoforms share binding sites with EWS-WT1 or regulate common targets. WT1 contains four Cys(2)His(2)-type zinc fingers that specif-

ically binds GC-rich sequences in the promoter regions of its target genes<sup>50</sup>. Thus, several studies have identified DNA binding motifs<sup>7</sup> that are summarized in Table 5. However, most of these studies do not discriminate between the WT1 +KTS and WT1 -KTS isoforms<sup>44</sup>. Lack of discrimination between these isoforms is partially due to the lack of antibodies available to distinguish between those isoforms, therefore ChIP profiling is expected to precipitate a mixture of both isoforms.

To determine whether wild-type WT1 isoforms share binding motifs with EWS-WT1 isoforms, we profiled genome-wide the occupancy pattern of V5-tagged WT1 isoforms stably expressed in the MeT-5A cell-line. Before proceeding to the ChIP-seq analysis, we confirmed the presence of *WT1* RNA transcripts by qPCR and confirmed the WT1 protein expression by western blot with a V5 and a WT1 antibody (Figure 28).

Reference	Year	Motif name	Motif sequence
Rauscher <i>et al.</i>	1990	EGR-1 motif	CGC-CCC-CGC
Bickmore <i>et al.</i>	1992		GTG-AGG-CTG
			CTC-CCT-CCC
			GAG-AGG-GAG-GAT
Wang <i>et al.</i>	1993	(TCC) <sub>n</sub>	TCC-TCC-TCC-TCC-TCT-CC
Rupprecht <i>et al.</i>	1994		AGC-AGG-GGG-AGG-CT
			AGA-GGA-GGG-TGT-CT
Nakagama <i>et al.</i>	1995	WTE motif	GCG-TGG-GAG-T
Hamilton <i>et al.</i>	1995		GCG-TGG-GCG-(T/G)(T/A/G)(T/G)
Duarte <i>et al.</i>	1998		GAG-CCG-GAC
Stoll <i>et al.</i>	2007		GNG-NGG-GNG
Miyamoto <i>et al.</i>	2008		GGA-GGA-GGG-A
Wells <i>et al.</i>	2010	WKE motif	ACC-AAG-CGG-GAT-GCG-GAG-CCG-CCGCCG-CCG-CCG
Yengo <i>et al.</i>	2018		GGC-GTG-GGC-GG

**Table 5: Summary of DNA binding motifs of WT1 in the literature.** Adapted from Ullmark *et al.*<sup>44</sup>. Authors suggested that nearly all these identified motifs match WT1 -KTS isoforms as it has a stronger DNA binding ability compared to WT1 +KTS isoform<sup>44</sup>. The WKE motif is the only validated WT1 +KTS binding-site<sup>54</sup>.



**Figure 28: MeT-5A lentivirally transduced with V5-tagged WT1 isoforms induce WT1 transcription and WT1 protein expression.** (A) Western blot of MeT-5A pLIV empty, MeT-5A pLIV V5 WT1 + KTS, and MeT-5A pLIV V5 WT1 - KTS. GAPDH confirmed equivalent protein loading. The WT1 protein was detected in the MeT-5A cells with the V5 and WT1 antibody. (B) Relative mRNA expression levels in MeT-5A pLIV V5 WT1 ± KTS were analyzed using the 2<sup>-ΔΔCt</sup> method; the data were first normalized to the corresponding GAPDH levels, and then calculated in reference to MeT-5A infected with a lentivirus containing an empty vector.

The motif enrichment analysis of these cells indicated that the top DNA binding motifs for WT1 were a CTCCC cluster and a TGACTCA consensus sequence for both isoforms (see Table 6). We observed that WT1 +KTS has very few binding-sites compared to WT1 – KTS (84 times more).

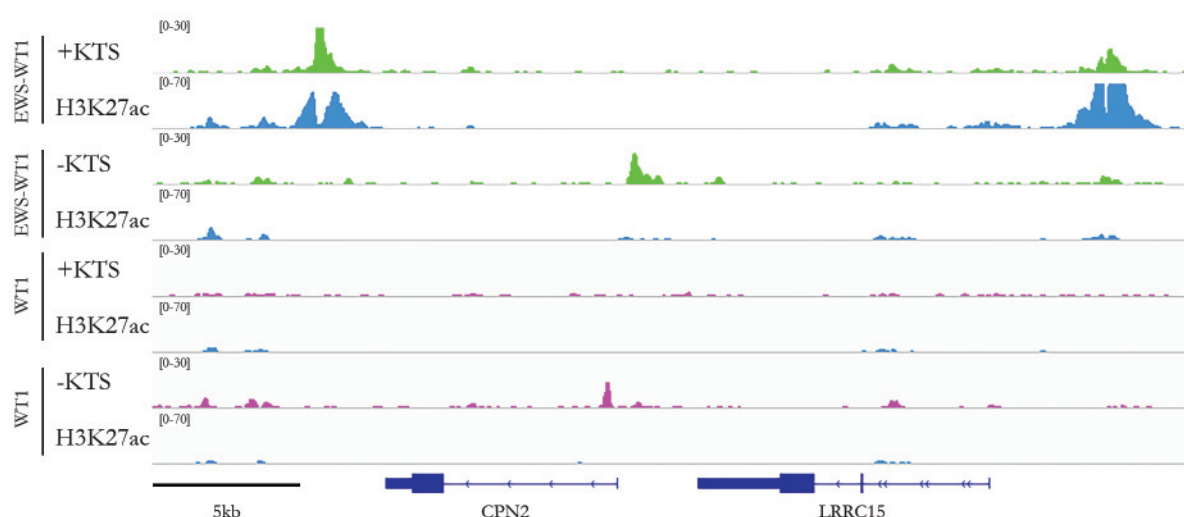
WT1 +KTS	(737 peaks)	WT1 -KTS	(62243 peaks)
Motifs	% targets	Motifs	% targets
	66.62		55.58
	46.95		25.15

**Table 6: MACS-called peaks analysis of WT1 isoforms in MeT-5A cells are different from EWS-WT1 motifs.** Left panel: Motifs analysis of WT1 +KTS isoform peaks. WT1 +KTS binds a CTCCCT sequence in 66% of the times and to TGAG/CTCA sequence 46.95% of the time. Right panel: Motifs analysis of WT1 -KTS isoform peaks. WT1 –KTS bind mainly to a CTCCCA/CC sequence. WT1 –KTS binds more DNA sequences (62243 peaks).

Zinc finger proteins can bind to DNA, RNA, and DNA–RNA hybrids<sup>117-119</sup>. WT1 has four zinc finger domains, and it has been shown that only the last three domains are indispensable to bind to the DNA<sup>118</sup>. In tissues and tumors expressing WT1, the most abundant WT1 form is the +KTS isoform<sup>34</sup>. The expression ratio between the WT1 +KTS and –KTS isoforms is 2:1, suggesting a higher biological significant role for the +KTS isoform<sup>46</sup>. Studies have shown that the WT1 –KTS isoform recognizes and binds DNA and regulates transcription through binding to the EGR-1 consensus binding site<sup>118</sup>. It was demonstrated that the WT1 +KTS isoform has a reduced ability to bind to DNA<sup>44,118</sup> and that this latter can associate with components of the RNA splicing machinery<sup>120</sup>. A study revealed that the WT1 +KTS isoform

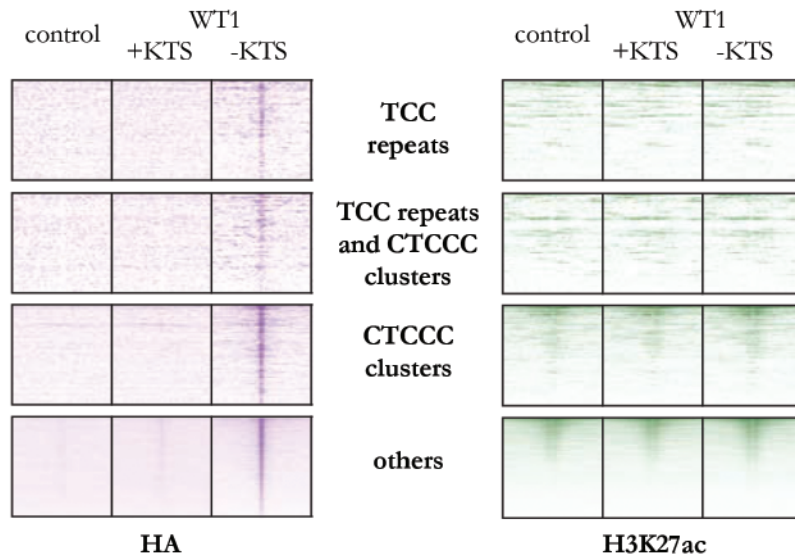
recognizes the DNA with the same specificity as the –KTS isoform<sup>104</sup>. However, due to the KTS insert, the WT1 +KTS isoform loses the C-capping interactions stabilizing the complex with the DNA<sup>104</sup>. It was also suggested that the WT1 +KTS isoform might be an RNA binding protein involved in nuclear RNA metabolism and posttranscriptional mechanisms.<sup>118,120,121</sup>. Other studies have shown that WT1 isoforms localize at different compartments within the cell. WT1 –KTS localizes diffusely in the nucleus, allowing more interaction with the chromatin, whether WT1+KTS localizes primarily in nuclear speckles<sup>45</sup>, together with RNA binding proteins like Sm antigens<sup>44</sup>. Our results are in concordance with these observations and could explain the low number of peaks found in the analysis of the WT1 +KTS isoform.

Our results demonstrated that wild-type WT1 did not bind the same target genes of EWS-WT1 (Figure 29 and 30). These results were in concordance with studies demonstrating that the DNA binding domains of WT1 and EWS-WT1 are structurally different<sup>122</sup>.



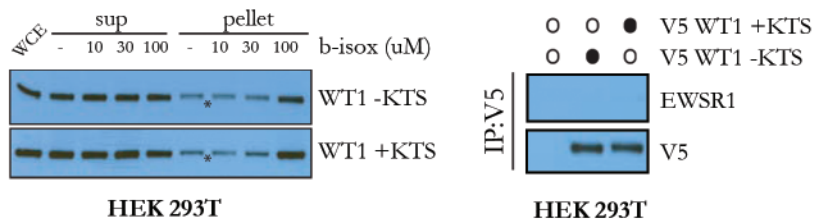
**Figure 29: EWS-WT1 display a different chromatin profile when compared to WT1 around the LRRC15 gene.** ChIP tracks for EWS-WT1 (in green) and WT1 (in violet) isoforms were expressed in Met-5A cells. In blue are tracks of H3K27ac for each condition.

To complete the comparison between EWS-WT1 and WT1, we decided to separate the WT1 ChIP peaks, based on the CTCCC clusters and TCC repeats motif identified for EWS-WT1. We confirmed that WT1 +KTS does not bind to the DNA at the EWS-WT1 binding sites. The WT1 –KTS bound the DNA to some of the CTCCC cluster peaks of EWS-WT1, nevertheless, they do not activate or change in the chromatin, as confirmed by the lack of H3K27ac signal at these regions (Figure 30). These results were in concordance with studies demonstrating that the EWS-WT1 protein has a higher binding affinity for a given recognition target than WT1<sup>122</sup>.



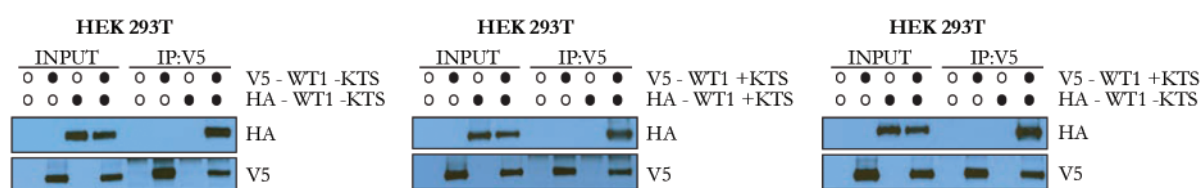
**Figure 30: WT1 isoforms cannot change the chromatin state of MeT-5A cells at EWS-WT1 binding sites.** ChIP-seq signal densities in MeT-5A cells expressing HA-tagged WT1 isoforms. Heatmaps of WT1 (violet) and H3K27ac (green). 10-kb windows in each panel are centered on EWS-WT1 binding sites and separated based on motifs identified in JN-DSRCT1.

We previously demonstrated that EWS-WT1 can bind to EWS wild-type, had phase transition properties, and can homo- and heterodimerize. Therefore, we decided to assess whether these properties could be shared with the wild-type WT1 protein. In order to assess the phase transition potential of WT1, we tested its ability to precipitate in the presence of b-isox. The experiments demonstrated that WT1 isoforms exhibit weak precipitation in HEK 293T cells stably expressing WT1 +KTS and WT1 -KTS isoform (Figure 31.A). This precipitation may be due to the N-terminal part of WT1 that is rich in prolines and glutamines. Furthermore, the immunoprecipitation assay of V5-tagged WT1 isoforms indicated no interaction with endogenous EWS (Figure 31.B). Therefore, WT1 cannot recruit the EWS protein, as does EWS-WT1.



**Figure 31: WT1 isoforms precipitate upon high doses of b-isox in HEK 293T cell lysates but fail to recruit endogenous EWSR1.** (A) Transfected WT1 +KTS and WT1 -KTS isoforms precipitate upon treatment high dose of b-isox in HEK 293T cell lysates. (B) Immunoblots of anti-V5 immunoprecipitates from HEK 293T cells transfected with either control vector, V5 WT1 +KTS, or V5 WT1 -KTS isoforms. Endogenous wild-type EWS does not interact with WT1.

To assess the dimerization of WT1, we performed a reciprocal immunoprecipitation experiments using overexpressed HA- and V5-tagged WT1 isoforms in HEK 293T cells. These results showed that both WT1 isoforms can homodimerize but also that the WT1 isoforms can heterodimerize (Figure 32), as does EWS-WT1. Our results are in concordance with studies that demonstrated that the N-terminal part of WT1 has a known self-association domain and that WT1 zinc fingers harbor a dimerization motif<sup>122,123</sup>. Our data suggested that a major difference between WT1 and the fusion protein EWS-WT1 is the ability to interact with wild-type EWS.



**Figure 32: The WT1 protein can homo and heterodimerize.** Co-immunoprecipitation experiments using anti-V5 and HA antibodies in transiently transfected HEK 293T cells co-expressing HA-tagged and V5-tagged WT1 isoforms. (A) HA-tagged and V5-tagged WT1 -KTS co-expressed in HEK 293T cells. (B) HA-tagged and V5-tagged WT1 +KTS co-expressed in HEK 293T cells. (C) V5-tagged WT1 +KTS and HA-tagged WT1 -KTS isoforms co-expressed in HEK 293T cells.

In summary, our results demonstrated that EWS-WT1 bound to both CTCCC clusters and TCC repeats, and that EWS-WT1 bound regions exhibited features of active enhancer elements. On the contrary, WT1 wild-type was observed to bind to CTCCC and TGANTCA clusters. WT1 -KTS bound to the chromatin, but this binding was not sufficient to change the chromatin activity at EWS-WT1 sites. Our results proved that EWS-WT1 targets were specific for the DSRICT tumors and not shared by wild-type WT1 protein. This also suggested that the fusion of EWS with WT1 may allow this latter to access the chromatin more widely.

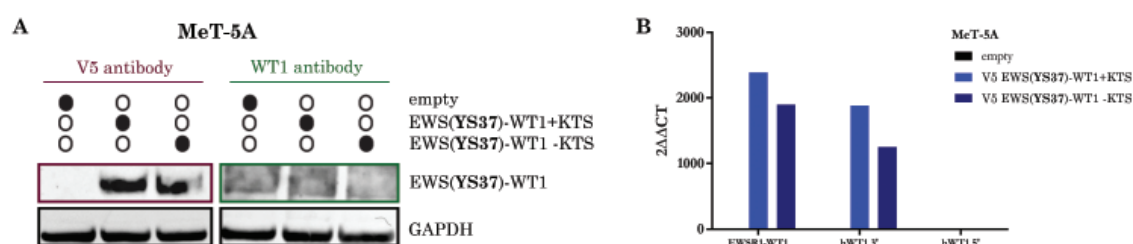
### 3.2.3.4. Tyrosine residues in the EWS-WT1 PrLD are necessary for DNA binding.

To deepen our understanding on the EWS-WT1 fusion protein, we then investigated the contribution of EWS to the function of the translocation. In Ewing's sarcoma, EWS-FLI1 neomorphic properties rely on the tyrosine residues present in the prion-like domain (PrLD) of EWS<sup>116</sup>. Therefore, it is pertinent to investigate this in the DSRICT context. We decided to test whether a similar mechanism may also underlie the functional differences we observed between EWS-WT1 and WT1 wild-type. We hypothesize that the EWS part of EWS-WT1 may confer pioneering properties to the translo-



cation, as compared to wild-type WT1. To test this postulate, we generated a HA- and V5-tagged EWS-WT1 mutant proteins with point mutations where all 37 tyrosines of the PrLD were mutated into serine (namely EWS(Y337)-WT1). We mutated all tyrosines of the PrLD into serines to retain its phosphorylation potential and only impair its LLPS properties.

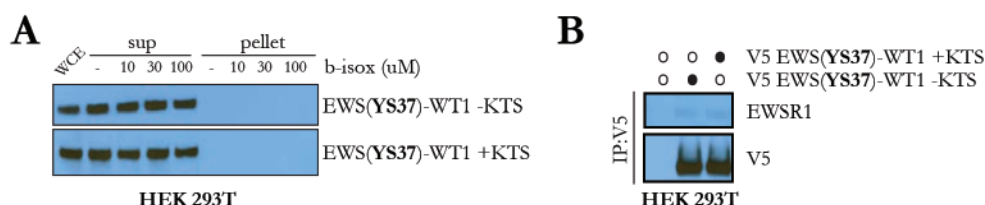
Before proceeding to the ChIP-seq analysis, we confirmed the presence of the aberrant *EWS(Y337)-WT1* RNA transcripts by qPCR and confirmed expression of the mutated protein with a western blot with a V5 antibody (Figure 33).



**Figure 33: MeT-5A lentivirally transduced with V5-tagged EWS(Y337)-WT1 isoforms induce EWS(Y337)-WT1 transcription and expression of the mutated protein.** (A) Western blot of MeT-5A pLIV empty, MeT-5A pLIV V5 EWS(Y337)-WT1 +KTS and MeT-5A pLIV V5 EWS(Y337)-WT1 -KTS. GAPDH confirmed equivalent protein loading. The EWS(Y337)-WT1 modified protein was detected in the MeT-5A cells with the V5 antibody but not the WT1 antibody. (B) Relative mRNA expression levels of MeT-5A pLIV V5 EWSR1(Y337)-WT1  $\pm$  KTS were analyzed using the  $2^{-\Delta\Delta C_t}$  method; the data were first normalized to the corresponding GAPDH levels, and then calculated in reference to MeT-5A infected with a lentivirus containing an empty vector.

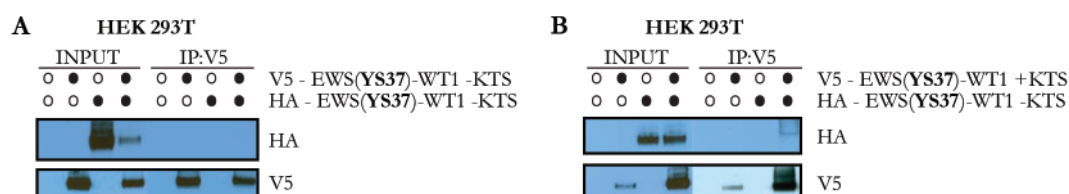
To assess the phase-transition potential of EWS(Y337)-WT1 isoforms, we tested its ability to precipitate in the presence of b-isox. The experiments revealed that EWS(Y337)-WT1 did not precipitate in the presence of b-isox (Figure 34.A). Additionally, immunoprecipitation of V5-tagged EWS(Y337)-WT1 displayed no interactions with wild-type EWS for both isoforms, when compared to EWS-WT1 (Figure 34.B). Therefore, EWS(Y337)-WT1 cannot recruit the EWS protein as does EWS-WT1. The normal PrLD of EWS-WT1 is the only responsible for the recruitment of endogenous EWS.





**Figure 34: The mutated EWS(YS37)-WT1 do not precipitate upon treatment with b-isox and fails to recruit endogenous EWSR1.** (A) Transfected EWS(YS37)-WT1 +KTS and EWS(YS37)-WT1 -KTS isoforms did not precipitate upon treatment with b-isox in HEK 293T cell lysates. (B) Immunoblots of anti-V5 immunoprecipitates from HEK 293T cells transfected with either control vector or EWS(YS37)-WT1 +KTS and -KTS isoforms. Endogenous wild-type EWS does not interact with EWS(YS37)-WT1.

Furthermore, the EWS(YS37)-WT1 mutant isoform proteins lost the ability to heterodimerize, as assessed by reciprocal immunoprecipitation experiments using overexpressed HA- and V5-tagged isoforms in HEK 293T cells (Figure 35). Therefore, the capacity of EWS-WT1 to homo- and heterodimerize is provided by the PrLD domain of EWS.



**Figure 35: The mutated EWS(YS37)-WT1 protein cannot dimerize.** Co-immunoprecipitation experiments using anti-V5 and HA antibodies in transiently transfected HEK 293T cells co-expressing HA-tagged and V5-tagged mutated isoforms. (A) HA-tagged and V5-tagged EWS(YS37)-WT1 -KTS co-expressed in HEK 293T cells. (B) V5-tagged EWS(YS37)-WT1 +KTS and HA-tagged EWS(YS37)-WT1 -KTS isoforms co-expressed in HEK 293T cells.

Based on these findings, we further tested the ability of the EWS(YS37)-WT1 mutant protein to bind DNA. ChIP-seq experiments clearly demonstrated that both EWS(YS37)-WT1 isoforms lost their ability to bind DNA. In line with this observation, no change in the activity was detected (figure not shown). Taken together, these results demonstrate that the tyrosine residues present in the EWS-WT1 prion-like domain are necessary to mediate phase transitions, recruit endogenous EWS, and to allow DNA binding.

In summary, EWS-WT1 is able to bind and activate the chromatin. Each EWS-WT1 isoform has a specific binding motif and, therefore, a unique set of targets. EWS-WT1 can bind to the DNA due to its PrLD domain and zinc finger domains. Furthermore, the PrLD domain of EWS-WT1 is required

to mediate phase transitions and dimerization properties along with the recruitment of endogenous EWS. These data suggest that, in DSRCT, EWS-WT1 has acquired neomorphic properties partly given by the EWS part of the fusion protein as compared to the wild-type WT1 protein. Therefore, these neomorphic properties can modify the cells and regulate unique target set of genes in DSRCT.

### 3.2.4. DSRCT cells express both EWS-WT1 isoforms concomitantly

It is still unknown whether DSRCT tumors express both EWS-WT1 isoforms concomitantly or if the tumor cells express only one or the other isoform. Unfortunately, commercially available WT1 antibodies cannot discriminate between the +KTS or -KTS isoforms. To circumvent this limitation, we decided to use a new RNA *in situ* hybridization technology called BaseScope to define the expression of the two EWS-WT1 isoforms within the tumor cells.

Existing data concerning the oncogenic properties of each isoform is contradictory<sup>74</sup>, and the distribution of each EWS-WT1 isoform is still unknown. Therefore, it becomes relevant to determine the location of the EWS-WT1 isoforms within the tumor, as their expression can influence the patients' survival outcome.

We used an assay based on the RNA in situ hybridization technology to determine the distribution of EWS-WT1 +KTS or EWS-WT1 -KTS isoforms within the DSRCT tumors called BaseScope that allows the detection of a targeted mRNA within cells or tissues<sup>124</sup>. The probes used were custom-designed for this study to target either the insertion or the absence of the KTS tripeptide of the WT1 part of EWSR1-WT1, called, respectively, plusKTS and noKTS (Table 7).

BaseScope probe	Target gene	Channel	Chromogenic Dye
<b><u>Positive control:</u></b> BA-Hs-1ZZ PPIB-C1, POLR2A-C2	<i>PPIB</i> (Cyclophilin B) and <i>POLR2A</i> (DNA-directed RNA polymerase II subunit RPB1)	C1 and C2	HRP-based Green (C1) and AP-based Fast Red (C2)
<b><u>Negative control:</u></b> BA-DapB-1ZZ	<i>DapB</i> (Nonspecific bacterial gene)	C1 and C2	HRP-based Green (C1) and AP-based Fast Red (C2)
BA-hWT1-plusKTS-Junc	+KTS insert	C1	HRP-based Green
BA-hWT1-noKTS-Junc-C2	-KTS insert	C2	AP-based Fast Red

**Table 7: List of 1ZZ BaseScope probes, target genes, and its respective detection channel and dye.**

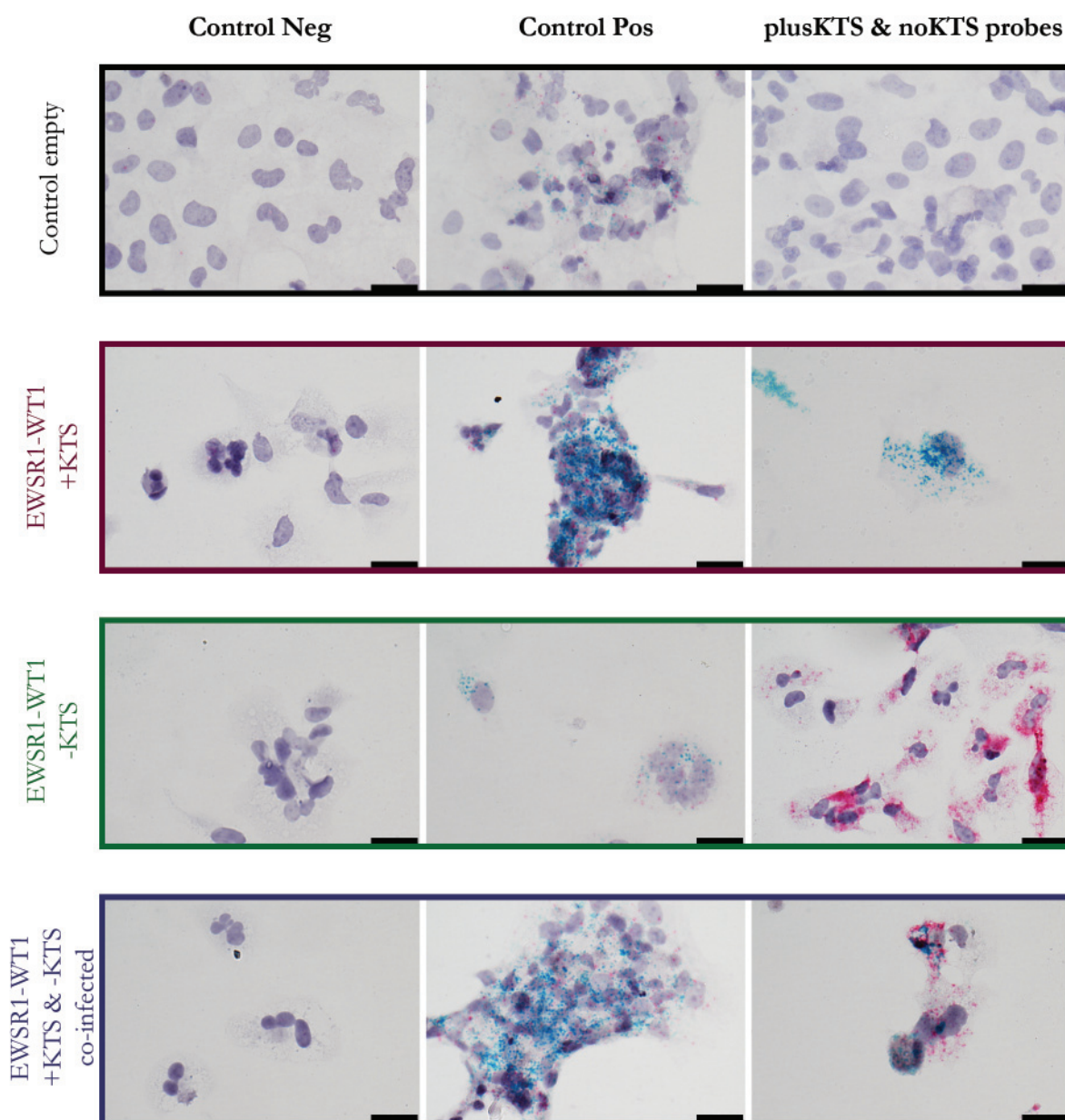
The manufacture provided positive control probes that target species-specific housekeeping genes (*PPIB* and *POLR2A*) and negative control probe targeting the bacterial *DapB* gene. There is only one ZZ oligo recognizing each RNA transcript. Alkaline phosphatase (AP) and horseradish peroxidase (HRP)-based enzymatic reactions.

We confirmed the specificity of each KTS probe using the previously mentioned MeT-5A cellular model, expressing either *EWSR1-WT1+KTS* or *-KTS*. Before testing the isoform-specific probes within these samples, we validated the RNA quality of the cells using the BaseScope positive controls probes *PPIB* and *POLR2A* (Figure 36). For the custom-made BaseScope 1ZZ probes, four different conditions were used to address the specificity of the plusKTS and noKTS probes (Figure 36).

In the first condition, MeT-5A cells infected with an empty vector, no RNA was recognized by either KTS probes, demonstrating no off-target effect. In the second condition, MeT-5A cells expressing only the *EWSR1-WT1 +KTS* isoform, only the plusKTS probe detected RNA (green dots), and the noKTS probe did not capture any signal (no red dots). In the third condition, MeT-5A cells expressing only the *EWSR1-WT1 -KTS* isoform, only the noKTS probe identified the correct RNA (red dots), and the plusKTS probe did not detect any RNA (no green dots). Therefore, each specific custom-made probe identified the correct RNA transcript uniquely. With this, we confirmed that the plusKTS probe recognized specifically the *EWSR1-WT1 +KTS* transcript, and noKTS probe recognized specifically the *EWSR1-WT1 -KTS* transcript with no off-target effect nor unspecific background signal.

Furthermore, we wanted to investigate if the 1ZZ probes could recognize each isoform specifically when both were expressed concomitantly (last condition). Therefore, in MeT-5A cells expressing both

EWSR1-WT1 isoforms simultaneously, the plusKTS probe recognized *EWSR1-WT1* +KTS (green dots) and the noKTS probe detected the *EWSR1-WT1* -KTS (red dots) transcript. Thus, these probes can be used further to detect EWSR1-WT1 isoforms transcripts in DSRCT cells or tissues.

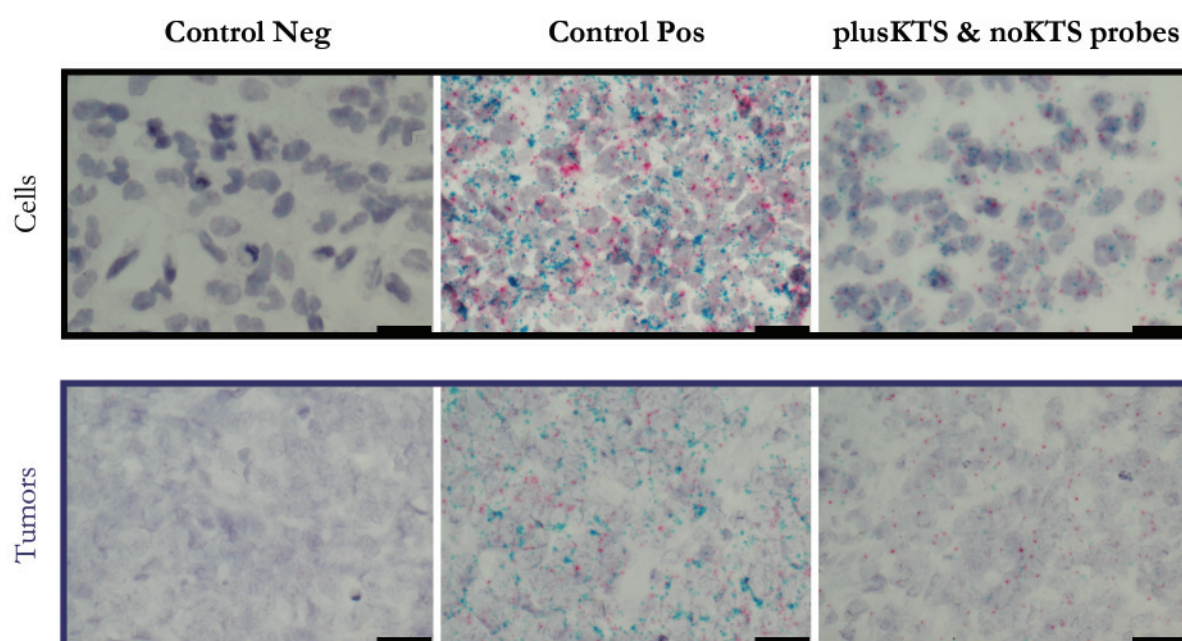


**Figure 36: 1ZZ KTS BaseScope probes recognize EWSR1-WT1 isoforms.** Representative images of the BaseScope assay with MeT-5A cells, scale bar 300µm. Top panel (black), MeT-5A cells expressing control empty. Second panel (violet) MeT-5A EWSR1-WT1 +KTS. Third panel (green) MeT-5A EWSR1-WT1 -KTS. Bottom panel (blue), MeT-5A EWSR1-WT1 +KTS and EWSR1-WT1 -KTS co-infected. On the left column, the BaseScope negative control probe does not recognize any target in Met-5A cells, as the probe used targets a nonspecific bacterial gene DapB. Therefore, there is no unspecific signal from the ZZ probes. Middle column shows positive control (PPIB and POLR2A). PPIB (green dots) and POLR2A (red dots) were recognized heterogeneously within the cells, confirming an overall good RNA quality of all the samples. On the right column shows plusKTS (Green) and noKTS (Red) probes. All the MeT-5A cells were plated on glass slides and permeabilized for this assay.



Since the 1ZZ probes can specifically detect *EWSR1-WT1 +KTS* and *EWSR1-WT1 -KTS* isoforms transcripts, we used them to determine the distribution of each *EWSR1-WT1* isoforms in the only available DSRCT cell-line, JN-DSRCT1. The plusKTS and noKTS probes identified the *EWSR1-WT1 +KTS* (green dots) and *EWSR1-WT1 -KTS* (red dots) isoforms, respectively, within the same cells. This result confirmed that both isoforms are concomitantly expressed within the same cell.

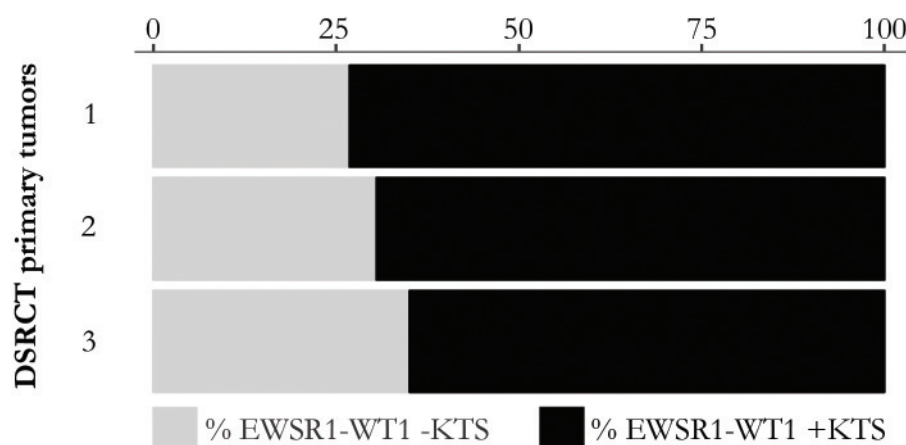
Next, we investigated the location of each *EWSR1-WT1* isoform in a JN-DSRCT1-derived tumor. A formalin-fixed paraffin-embedded (FFPE) tissue from JN-DSRCT1 xenograft was tested, and we were able to detect both *EWSR1-WT1* isoforms within the tumor (Figure 37).



**Figure 37 : 1ZZ KTS BaseScope probes recognize *EWSR1-WT1* isoforms expressed concomitantly in JN-DSRCT1 cells and derived tumors.** Representative images of the BaseScope assay, scale bar 300µm. Top panel (black), JN-DSRCT1 cells plated on glass slides. Bottom panel (blue), FFPE JN-DSRCT1 tumors. Left column is the negative control (*DapB*), where the probe did not recognize any target. Middle column is the positive control, where *PPIB* (Green) and *POLR2A* (Red) were detected heterogeneously, therefore both samples have a good quality of RNA. Right column is plusKTS (green) and noKTS (red) probes. We noticed very little detection of *EWSR1-WT1 +KTS* transcript in JN-DSRCT1 xenograft tumors.

With the BaseScope KTS probes, we detected the *EWSR1-WT1 +KTS* and *EWSR1-WT1 -KTS* isoforms transcripts in MeT-5A cells expressing only one or both isoforms. Furthermore, we confirmed that both *EWSR1-WT1* isoforms are simultaneously expressed within the same cell in JN-DSRCT1. Nevertheless, we encountered some challenges in detecting equal amounts of both *EWSR1-WT1* isoforms within the same cells in JN-DSRCT1-derived tumors for undetermined reasons. Interestingly,

we detected differential amounts of *EWSR1-WT1* RNA isoform transcripts in three primary human DSRCT tissues. In these tissues, we detected between 26% and 30% *EWS-WT1 -KTS* transcripts (Figure 38). This is in agreement with studies showing that the *EWS-WT1 +KTS* isoform was more abundant in the DSRCT tumors, with a ratio of 2 (+KTS):1(-KTS)<sup>34</sup>. Our results are consistent with the observation that both *EWS-WT1* isoforms are co-expressed in JN-DSRCT1 cells.

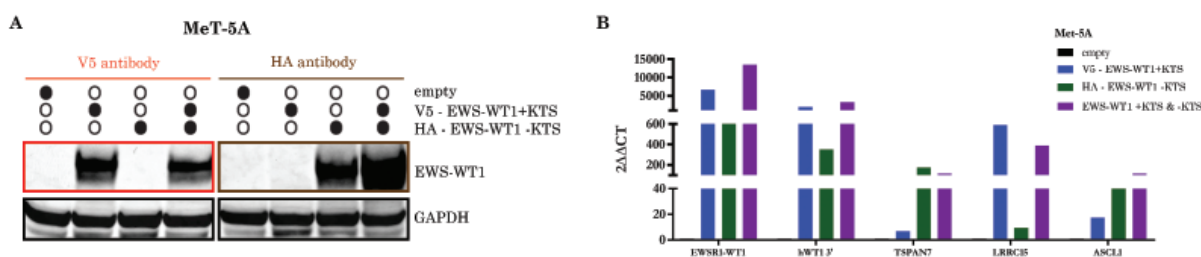


**Figure 38: Primary human DSRCT tumors displays higher *EWSR1-WT1 +KTS* transcripts.** RNA-seq analysis of three different human DSRCT primary tumors showed up to 30% of *EWSR1-WT1 -KTS* transcripts. Therefore primary tumors transcript more *EWSR1-WT1 +KTS*.

### 3.2.5. Expression of both *EWS-WT1* isoforms in MeT-5A cells induce tumorigenesis in vivo

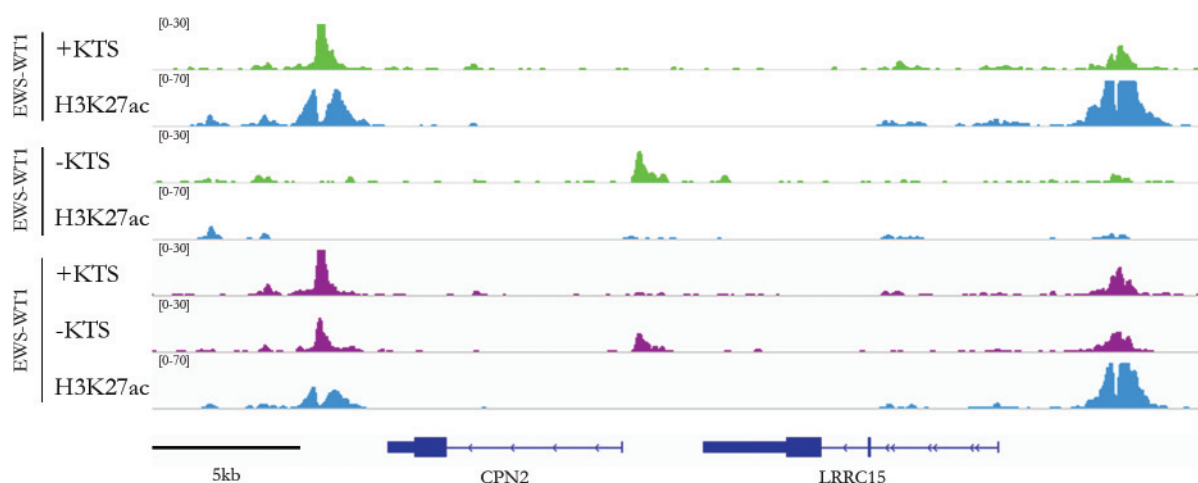
#### 3.2.5.1. Concomitant expression of *EWS-WT1* isoforms in MeT-5A cells confers unique target activation

Based on our results and the distinct genome occupancy profiles of *EWS-WT1* isoforms in MeT-5A, we decided to investigate the epigenetic landscape and transcriptional profiles of MeT-5A cells co-expressing both isoforms. To this end, V5 and HA-tagged *EWS-WT1 +KTS* and *-KTS* isoforms were co-expressed in MeT-5A cells. We validated the protein expression of *EWS-WT1* by western blot analysis using anti-HA and -V5 antibodies (Figure 39.A). Relative quantification by qPCR further confirmed high mRNA levels for both *EWSR1-WT1* isoforms and a selected panel of known direct DSRCT target genes (Figure 39.B).



**Figure 39: MeT-5A lentivirally transduced with V5 or HA-tagged EWS-WT1 express the aberrant protein and induce transcription of DSRCT targets.** (A) Western blot of MeT-5A pIND20 empty, MeT-5A pIND20 V5 EWS-WT1 + KTS, MeT-5A pIND20 HA EWS-WT1 – KTS and MeT-5A pIND20 V5 EWS-WT1 + KTS & HA EWS-WT1 - KTS. GAPDH confirmed equivalent protein loading. The EWS-WT1 fusion protein was detected in the MeT-5A cells with a HA or V5 antibody. (B) Relative mRNA expression levels of MeT-5A pIND20 EWSR1-WT1  $\pm$  KTS were analyzed using the  $2^{-\Delta\Delta Ct}$  method; the data were first normalized to the corresponding GAPDH levels, and then calculated in reference to MeT-5A infected with a lentivirus containing an empty vector.




These results prompted us to assess the relative contribution of each EWS-WT1 isoform to the chromatin landscape of MeT-5A cells. Interestingly, we observed a distinct pattern of genome occupancy when both isoforms were co-expressed. At some genomic regions normally bound only by one isoform, we observed the recruitment of both EWS-WT1 isoforms upon their co-expression (Figure 40). The heterodimerization capacity of EWS-WT1 can explain the recruitment of both isoforms at the same binding site.



**Figure 40: When expressed together, EWS-WT1 +KTS binds to EWS-WT1–KTS and recruit this latter at the LRRC15 locus.** In blue are tracks of H3K27ac for each condition. The first two tracks are the result of EWS-WT1 +KTS expression in MeT-5A cells. Then the two following tracks represent only EWS-WT1 -KTS expression in MeT-5A cells. The last three tracks are the result of the concomitant expression of both EWS-WT1 isoforms in MeT-5A cells.

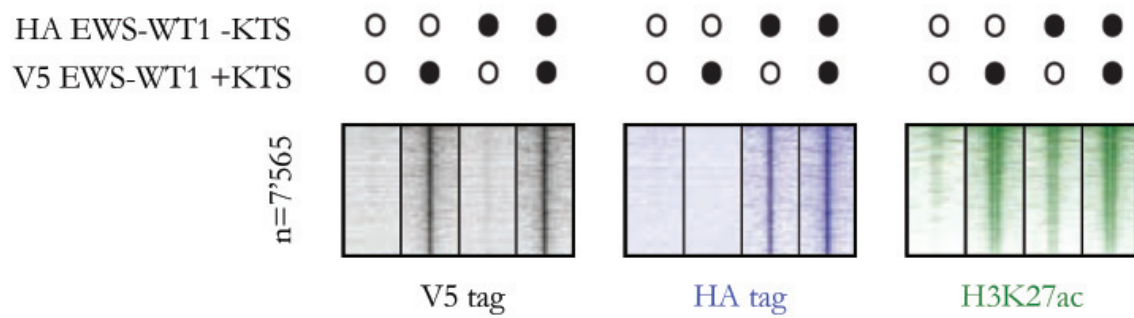


We identified direct binding sites of EWS-WT1 heterodimers in MeT-5A cells, stably expressing both isoforms. The ChIP-seq was realized with an antibody directed against the V5 or HA-tags linked to EWS-WT1. The motif enrichment analyses indicated that the DNA binding motifs of both EWS-WT1 isoforms shared sequence similarities represented by CTCCC clusters and TCC repeats (more details in Table 8). We observed that the top motif for EWS-WT1 heterodimers was a “CTCCC” cluster sequence found in 67.34% of targets. This motif shared similarities with the motifs we identified in the JN-DSRCT1 cell-line and the MeT-5A cells expressing only one of the EWS-WT1 isoforms. The second top motif identified was the TCC repeat motif in 38.76% of targets. This motif is similar to the motif we identified for the EWS-WT1 –KTS isoform (TCCTCCTCCTCCTCCT). The last motif identified when both EWS-WT1 isoforms were concomitantly expressed was a “TCCCTC-CCTTCCTTCC” sequence in 24.08% of targets.

EWS-WT1 +KTS and -KTS	
Motif	% targets
	67.39
	38.76
	24.08

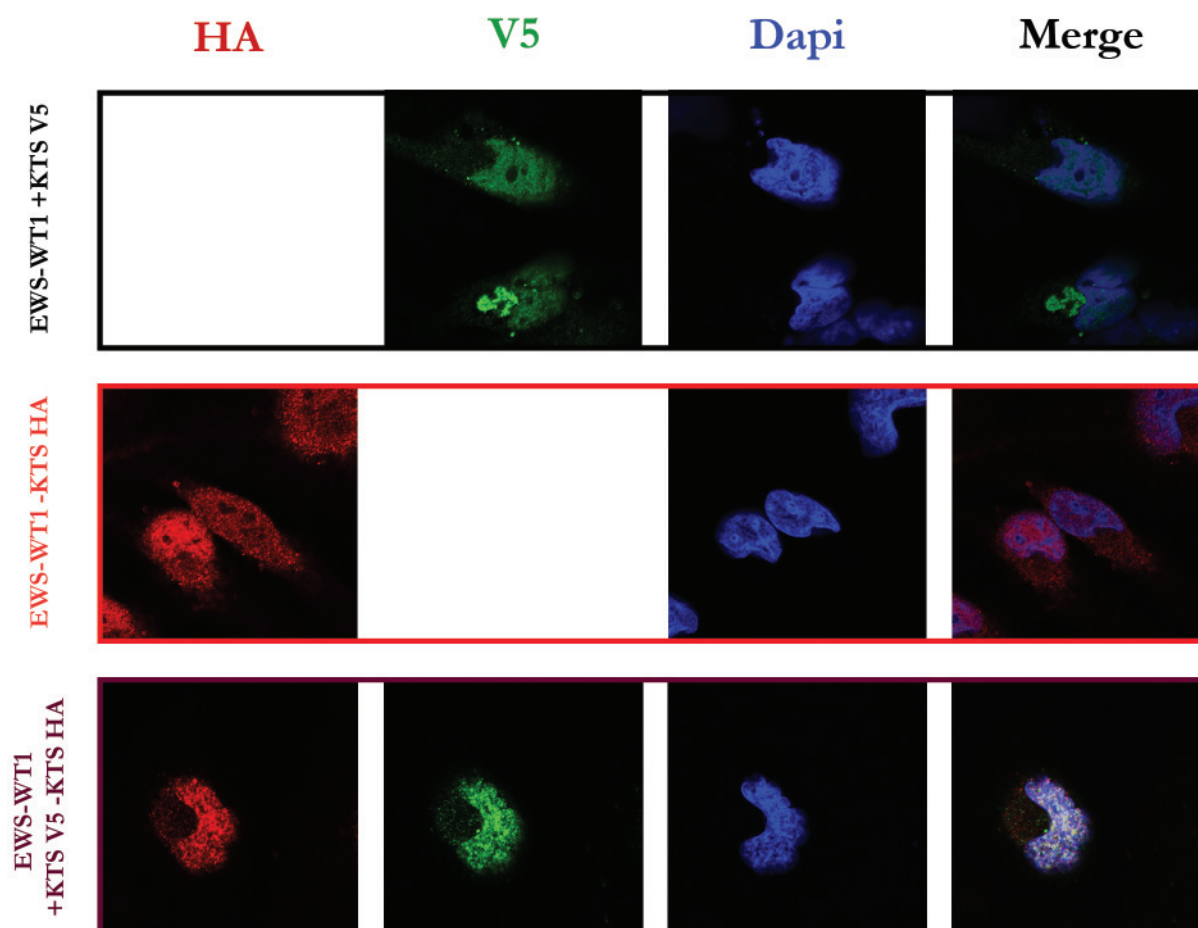
**Table 8: Motif enrichment analysis of EWS-WT1 isoforms expressed simultaneously within the same MeT-5A cell share similarities the motifs identified in JN-DSRCT1 cell line.** MACS-called peaks after expression of both EWS-WT1 isoforms in MeT-5A cells. Motifs analysis of concomitant expression of both EWS-WT1 isoform pointed binding at 67.39% to a (C/A)CTCCC sequence. CTCCC motif similar to the one identified in JN-DSRCT1 and EWS-WT1 isoforms. The second motif is a TCC repeat in 38.76% of targets. This motif matches the identified EWS-WT1 –KTS motif.

We profiled a panel of signature histone marks to link these binding sites to the corresponding chromatin state (Figure 41). We found that the majority of EWS-WT1 heterodimer binding sites were enriched in H3K27ac as similarly to the JN-DSRCT1 cell-line and MeT-5A cells expressing EWS-WT1 isoforms alone.



**Figure 41: EWS-WT1 isoforms can be recruited at the same binding sites and change the chromatin state of MeT-5A cells.** Heatmaps of EWS-WT1 and H3K27ac (green) ChIP-seq signal densities in MeT-5A cells expressing both EWS-WT1 isoforms. 5-kb windows in each panel are centered on EWS-WT1 binding sites.

We confirmed the isoform-specific expression of the EWS-WT1 protein alone or concomitant by immunocytofluorescence (IF) using the corresponding V5- and HA-tags (Figure 42). In MeT-5A cells, EWS-WT1 +KTS was concentrated within the nucleus, whereas EWS-WT1 –KTS was diffused between the nucleus and the cytoplasm. This result was surprising because we demonstrated that the EWS-WT1 –KTS isoform displayed a high DNA accessibility potential. Therefore, we expected to find the EWS-WT1 –KTS isoform mainly within the nucleus. We supposed that the HA protein tag detached from EWS-WT1 and diffused within the cell as a side effect of the protein's overexpression. This results could also be linked to a technical problem related to the fluorescent antibodies used that might have exhibited a background noise signal or accumulated in the cytoplasm. When both EWS-WT1 isoforms were expressed concomitant in MeT-5A cells, both proteins co-localized within the nucleus.



**Figure 42: Concomitant expression of EWS-WT1 isoforms colocalize them within the nucleus of the cell.** Immunocytofluorescence of HA and V5-tags in MeT-5A cells expressing EWS-WT1. Top panel (black), MeT-5A V5 EWS-WT1 +KTS. The second panel (orange), MeT-5A HA EWS-WT1 –KTS. Bottom panel (violet) MeT-5A EWS-WT1 +KTS –KTS co-infected. Cells in the first column were stained for HA (in red). Cells in the second column were stained for V5 (in green). Third column stained with Dapi (in blue) to color the nucleus of cells. The last column shows a merge of the channels. Only tumors expressing both EWS-WT1 isoforms are positively for V5 and HA-tag.

### 3.2.5.2. Concomitant expression EWS-WT1 isoforms in MeT-5A cells generated viable tumors in vivo

There is no genetically engineered mouse model to study DSRCT initiation. The reason for the absence of a model may probably lie in the histology of the tumor. DSRCT is composed of a dense desmoplastic stroma with small round blue cells. These cells are the component of what makes the cell-line. The lack of mouse models may also be related to the unknown cell of origin of DSRCT. An animal model that can accurately present the cellular and molecular changes associated with the initiation and progression of human DSRCT can significantly impact the understanding and treatment of this sarcoma. Therefore, to understand the early stages of DSRCT oncogenesis, we created a primary cellular model, where we expressed EWS-WT1 in MeT-5A cells to recreate the original event supporting the development of DSRCT. Additionally, we assessed

whether the expression of EWS-WT1 in this model might trigger tumorigenesis *in vivo*.

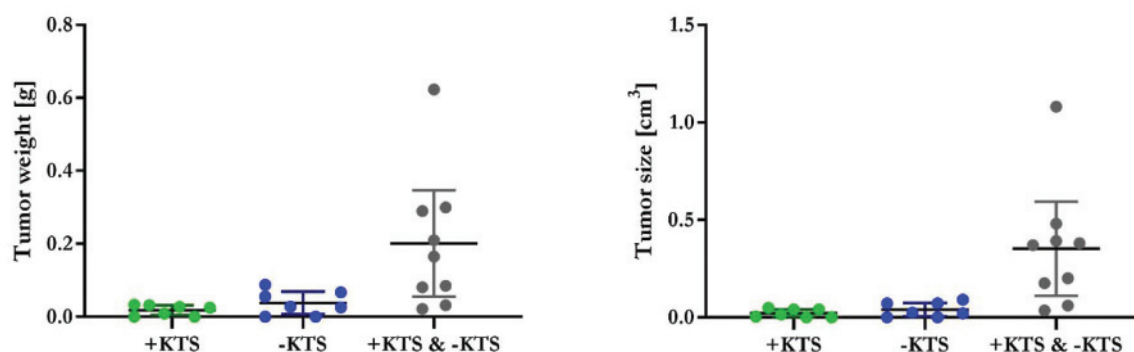
EWS-WT1 regulates specific targets in DSRCT, but it is still unknown the importance of each EWS-WT1 isoform during DSRCT tumorigenesis. To address this, we assessed the tumorigenic capacity of MeT-5A cells expressing HA- or V5-tagged EWS-WT1 proteins described previously. We used four conditions: MeT-5A expressing an empty vector (control), MeT-5A cells expressing EWS-WT1 +KTS, MeT-5A cells expressing EWS-WT1 -KTS, and MeT-5A cells expressing both EWS-WT1 isoforms concomitantly. We injected  $2 \times 10^6$  cells from each condition subcutaneously in the flank of a 5-7 weeks old immunodeficient NOD/SCID c Gamma (NSG) mouse. We observed that cells expressing both EWS-WT1 isoforms gave rise to tumors *in vivo*. The first and second tumor samples were collected six and eight months after injection in these mice, respectively. Accordingly, the experience was stopped nine months after subcutaneous cell injection into the mice.

The results confirmed that 100% of mice bearing MeT-5A cells expressing both EWS-WT1 isoform developed tumors, whereas only a minority of MeT-5A expressing only one of the isoforms developed small lesions that did not progress over time.

To ensure the continuous expression of EWS-WT1 protein within MeT-5A cells injected into the mice, we added doxycycline (0.625 mg/kg doxycycline) to the food ingested by the mice. None of the mice fed with doxycycline-free food developed any tumor. Furthermore, none of the control mice bearing MeT-5A cells with an empty vector developed tumors, confirming once more that MeT-5A cells are not tumorigenic.

One of the tumors derived from MeT-5a cells expressing both EWS-WT1 isoforms was further subcutaneously engrafted in a new mouse and gave rise to a new viable tumor. This serial xenograft experiment confirmed the malignant capacity acquired by our MeT-5A cells concomitantly expressing both EWS-WT1 isoforms.

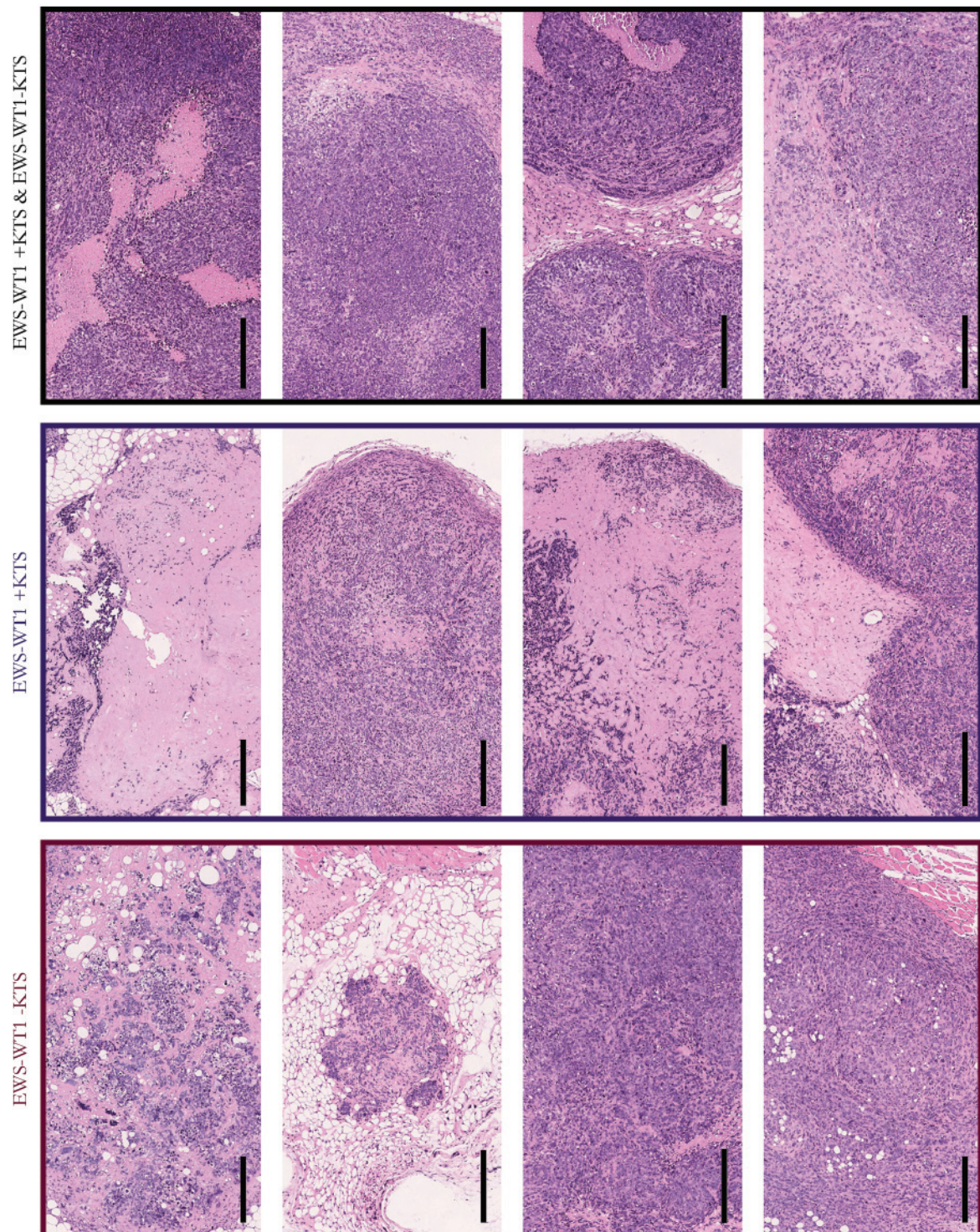
Tumors derived from MeT-5A cells expressing both EWS-WT1 isoforms were bigger than tumors from the other conditions (Figure 43). The tumors expressing only one of the EWS-WT1 isoforms were tiny, rigid, and white.



**Figure 43: MeT-5A cells expressing both EWS-WT1 isoform are capable of develop tumors.** Scatter plot of tumor weight (left) and size (right) of MeT-5A cells stably expressing EWS-WT1.  $2 \times 10^6$  MeT-5A cells were subcutaneously injected into a flank of NSG mouse. There were four conditions: MeT-5A pIND20 empty (not shown), MeT-5A pIND20 V5 EWS-WT1 +KTS, MeT-5A pIND20 HA EWS-WT1 -KTS, and MeT-5A expressing V5 EWS-WT1 +KTS and HA EWS-WT1 -KTS concomitantly. For each condition, we had an internal negative control that did not receive doxycycline. 100% of mice bearing Met-5A cells expressing both EWS-WT1 isoform developed tumors. None of the mice eating food without doxycycline developed any tumor. Furthermore, none of the control mice bearing MeT-5A cells with an empty vector developed tumors, confirming once more that MeT-5A cells are not tumorigenic. Tumor dimensions were measured using a caliper at the end of the experiment. Plots show all data, with mean value and 95% CI (confidence interval).

The morphology of the tumors expressing both EWS-WT1 isoforms was reminiscent of human DSRCT tumor histology, composed of islands of small round cells surrounded by desmoplastic stroma (Figure 44). Tumors expressing only the EWS-WT1 +KTS isoform seemed more desmoplastic, displaying larger stroma parts around the tumor, whereas tumors expressing the EWS-WT1 -KTS isoform were very cellular with scant stroma. No metastasis were observed in any of the mice bearing tumors.





**Figure 44 : Tumors expressing concomitantly both EWS-WT1 isoforms are similar to DSRCTs.** Histological view by H&E. Black scale bar = 300µm. Top panel (black), MeT-5A expressing EWS-WT1 +KTS and EWS-WT1 –KTS concomitantly. Middle panel (blue) MeT-5A cells expressing EWS-WT1 +KTS. Bottom panel (violet) MeT-5A expressing EWS-WT1 –KTS.

The results demonstrated that only the expression of both EWS-WT1 isoforms could give rise to healthy tumors that can continue to develop, as normal cancers do. Furthermore, it appears to be a



complementary effect of the expression of both EWS-WT1 isoforms in the tumorigenesis of DSRCT, and each isoform may have a different influence on the microenvironment surrounding them.

To summarize, we analyzed the early stages of DSRCT oncogenesis with our MeT-5A model, where stable expression of EWS-WT1 recapitulated and supported the early events driving DSRCT development. We demonstrated that EWS-WT1 isoforms had the ability to homo- and hetero-dimerize and to recruit wild-type EWS. Further, we unveiled that each EWS-WT1 isoform had different binding affinities to the DNA, but both induced a H3K27ac chromatin activation. We also determined that each isoform regulated a unique target set of genes in DSRCT. We demonstrated that when both EWS-WT1 isoforms heterodimerize, a distinct pattern of genomic occupancy emerged. Finally, we demonstrated that the concomitant expression of both EWS-WT1 isoforms was crucial to trigger the early stages of DSRCT tumorigenesis and maintain its development.

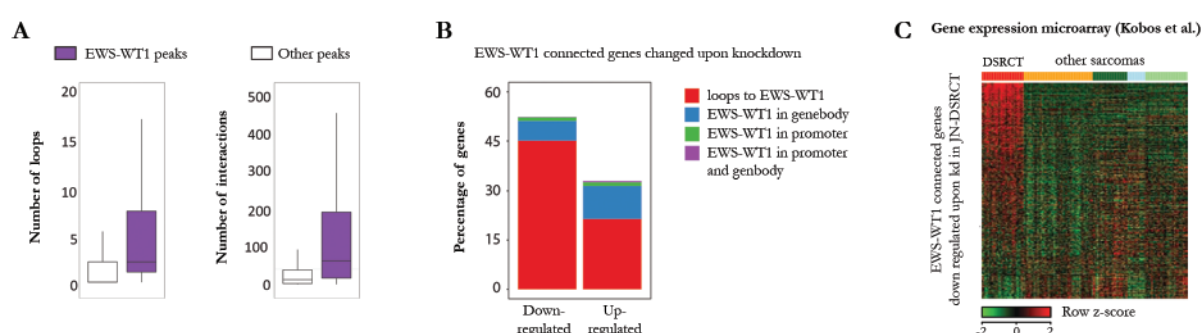
### **3.3. Identification of EWS-WT1-specific direct targets genes for potential therapeutic application in DSRCT**

Generally, the cells' oncogenic transformation involves global changes in the chromatin structure to reset the epigenetic landscape. Therefore, we would like to investigate how tumor cells establish *de novo* enhancers that drive oncogenic processes in DSRCT.

#### **3.3.1. 3D chromatin looping analysis identified a cohort of EWS-WT1 direct targets in DSRCT**

An essential layer of DNA regulation is chromatin's spatial organization, which is crucial for regulating gene transcription. Therefore, mapping all interactions between active enhancers and gene promoters in DSRCT will provide useful information to understand its development. To this aim, we used high-throughput chromosome conformation capture (3C) techniques coupled with chromatin immunoprecipitation (HiChIP). This technology allows the profiling of all chromatin interactions genome-wide. This approach is a protein-centric chromatin conformation method centered on a specific histone mark, such as H3K27ac. Typically, enhancer-promoter crosstalk is partitioned by the genome's organization into topologically associated domains (TADs). Therefore genomic interactions tend to be stronger within a TAD, where distal regulatory elements can directly contact their target genes by chromatin loops<sup>125</sup>.

We performed HiChIP chromatin looping analysis using H3K27ac mark as a surrogate to identify the active enhancer repertoire, and EWS-WT1 mediated enhancer looping in DSRCT. Our epigenetic analysis in the primary tumors and DSRCT cell model revealed that EWS-WT1 conferred active enhancers by establishing the H3K27ac mark at the translocation binding sites. HiChIP provided all H3K27ac-mediated looping, including to the enhancer-promoter and enhancer-enhancer interactome of JN-DSRCT1 cells. Integrative analysis of EWS-WT1 occupancy and H3K27ac modification, identified EWS-WT1-mediated loops to its targets. The interaction score of these EWS-WT1 loops was stronger than non-EWS-WT1 loops within JN-DSRCT1 cells, suggesting a strong potential of EWS-WT1 for target activation (Figure 45.A). The direct target genes identified in this study showed a significant overlap with the genes deregulated after the EWS-WT1 knockdown in the JN-DSRCT1 cell-line. Upon EWS-WT1 knockdown, loops connected to EWS-WT1 changed (Figure 45.B). Most of the changes occurred within loops connecting EWS-WT1 to enhancers, while only a few changes were observed in loops connecting EWS-WT1 to promoters. Next, we assessed the specificity of the EWS-WT1 loops and transcriptionally activated genes in DSRCT. To this aim, we compared our gene set to a published microarray data set comprising information from other sarcomas, including Synovial Sarcoma, Alveolar RhabdoMyoSarcoma, Alveolar Soft Part Sarcomas, and Ewing's Sarcoma (Figure 45.C). The comparative analysis revealed the unique pattern of the EWS-WT1 target genes in DSRCT, providing the opportunity to work with a cohort of target genes under the direct control of EWS-WT1 for potential therapeutic targets.



**Figure 45: EWS-WT1 governs regulatory networks in DSRCT.** (A) Boxplots showing the number of loops and number of interactions between anchor sites in H3K27ac Hi-ChIP in the absence or presence of EWS-WT1 in JN-DSRCT1 cells. (B) Bar plot shows the percentage of genes up- or down-regulated upon EWS-WT1 knockdown directly associated with EWS-WT1 through enhancer looping, direct binding at the promoter, or within the gene body. (C) Heatmap shows gene expression levels in primary tumors for the genes connected to EWS-WT1 and down-regulated upon its knockdown identified in B (304 genes (corresponding to 477 probes) out of 520 down-regulated connected genes were present in the microarray data set). Other sarcomas are Synovial Sarcoma, Alveolar RhabdoMyoSarcoma, Alveolar Soft Part Sarcomas, and Ewing's Sarcoma.

### 3.4. CDK4/6 inhibitors impair DSRCT tumor growth

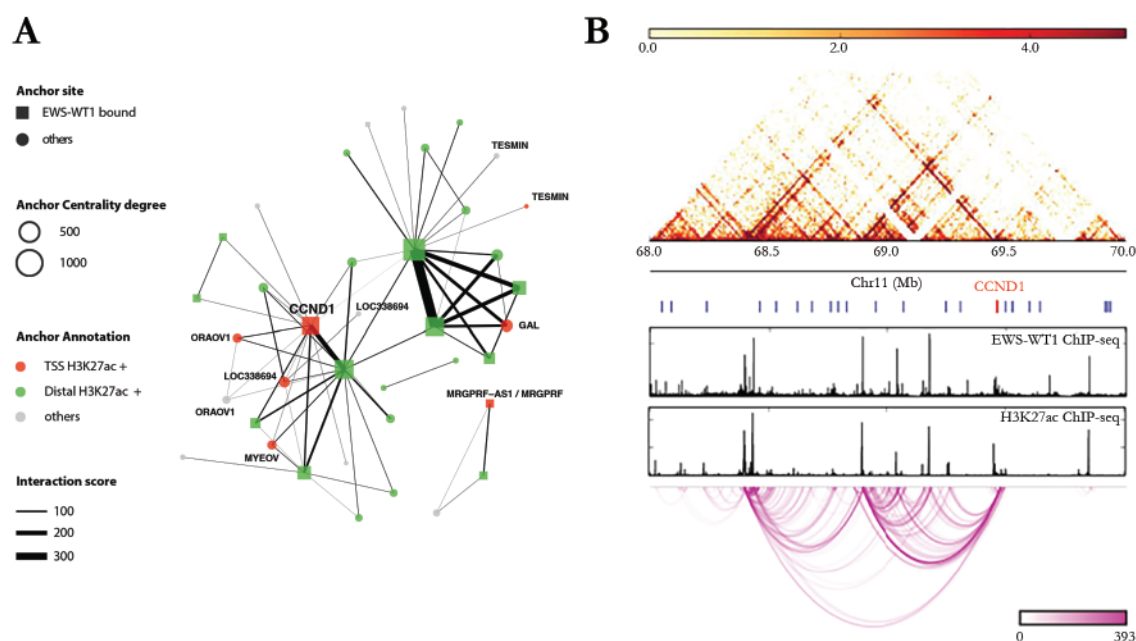
To identify a potential therapeutic candidate for DSRCT, we analyzed the EWS-WT1 direct target gene list. Among the 100 most expressed genes connected to EWS-WT1 in the JN-DSRCT1 cell-line, we identified 12 potential therapeutic targets linked with at least one known pharmacological inhibitor commercially available (Figure 46).

Cyclin D1 gene (*CCND1*) is one of the top hits based on our prioritization criteria (Figure 47). Several EWS-WT1 bound enhancers are in physical contact with the *CCND1* promoter and regulate its expression. *CCND1* is an important regulator of the cell-cycle, where progression from G1 to S phase is initiated after forming a complex between *CCND1* and CDK4/6<sup>4</sup>. Deregulation of the cell-cycle checkpoint proteins is a key hallmark of cancer. Therefore, *CCND1* being a specific target for DSRCT, we investigated *CCND1* as a potential therapeutic strategy for DSRCT patients.



Target gene	Compound
<i>CACNA2D2</i>	isradipine
<i>CCND1</i>	palbociclib
<i>VEGFA</i>	muparfostat
<i>TH</i>	metyrosine
<i>CALM1</i>	dibucaine
<i>MAP2K2</i>	bosutinib
<i>KDR</i>	ramucirumab
<i>FGFR4</i>	azd-4547
<i>MERTK</i>	bms-777607
<i>MAPK3</i>	sulindac
<i>SLC19A1</i>	methotrexate
<i>PIK3C2B</i>	buparlisib

**Figure 46: List of 12 potential therapeutic targets specific for DSRCTs, linked with at least one known pharmacological inhibitor.** Top genes among the 100 most highly regulated EWS-WT1 direct target genes in DSRCT.



**Figure 47: EWS-WT1 highly interacts with CCND1.** (A) Hi-ChIP enhancer network at chr11:28Mb-70Mb. (B) Integrative Genomics Viewer (IGV) plots of ChIP-Seq showing co-occupancy of EWS-WT1 and H3K27ac at the super-enhancer and promoter of the CCND1 gene locus. Hi-C interactions were analyzed from the JN-DSRCT1 cell-line and are shown in violet.

### 3.4.1. Therapeutic potential of CCND1 -CDK4/6

Pharmacological agents can impair the activity of the CCND1-CDK4/6 complex. Three small molecules have been identified as CDK4/6 inhibitors: Palbociclib, Abemaciclib, and Ribociclib<sup>3</sup> (Figure 48). They all selectively inhibit the formation of the CCND1-CDK4/6 complex by binding to the adenosine triphosphate (ATP)-binding pocket of CDKs<sup>3,4</sup>. These ATP-competitive molecules can be administered orally and are reported to induce little or no suppression of other CDK activities at clinical doses<sup>8</sup>.

These three inhibitors were approved by the Food and Drug Administration (FDA) for the treatment of metastatic breast cancer<sup>4,5,126,127</sup>. Pharmacokinetics of these inhibitors demonstrated high tissue penetration with low absorption and elimination, resulting in an elimination half-life of 29h, 32h, and 18h for palbociclib, ribociclib, and abemaciclib, respectively<sup>4</sup> (Figure 48). They are administered orally once (palbociclib and ribociclib)<sup>128</sup> or twice a day (abemaciclib)<sup>5</sup>. Therefore, they inhibit Retinoblastoma tumor suppressor protein (Rb) phosphorylation, reducing the expression of E2F targets, thus blocking the cell at the G1 phase, and inhibiting cancer cell proliferation<sup>4</sup> (Figure 49). Studies reported that inhibition of CDK4/6 decreases sarcoma cell growth<sup>3</sup>. Furthermore, these inhibitors are currently under clinical evaluation for Ewing and Synovial sarcoma treatment.

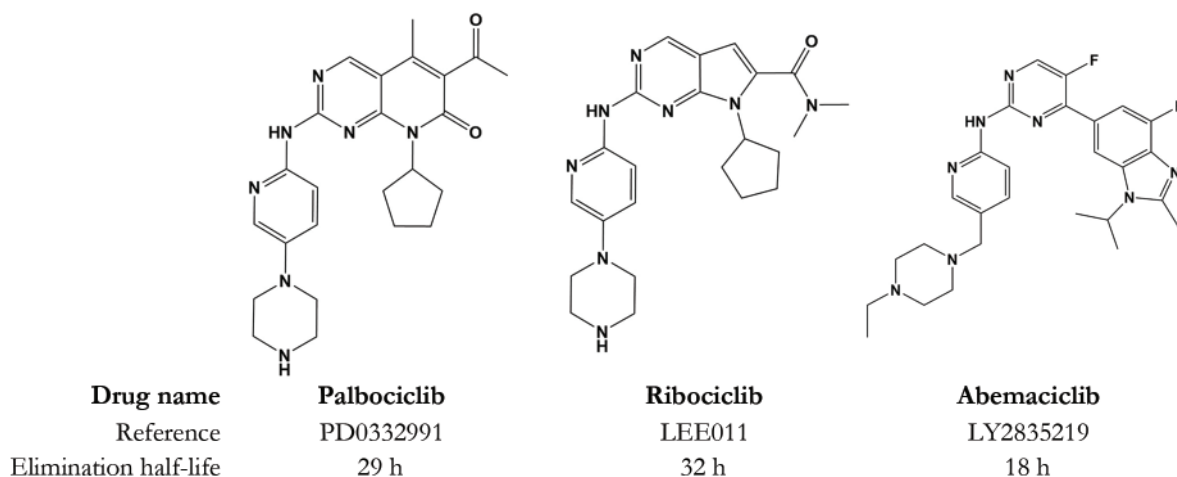


Figure 48: Molecular structure of CDK4/6 inhibitors approved by the FDA with its corresponding name, reference, and elimination half-life.

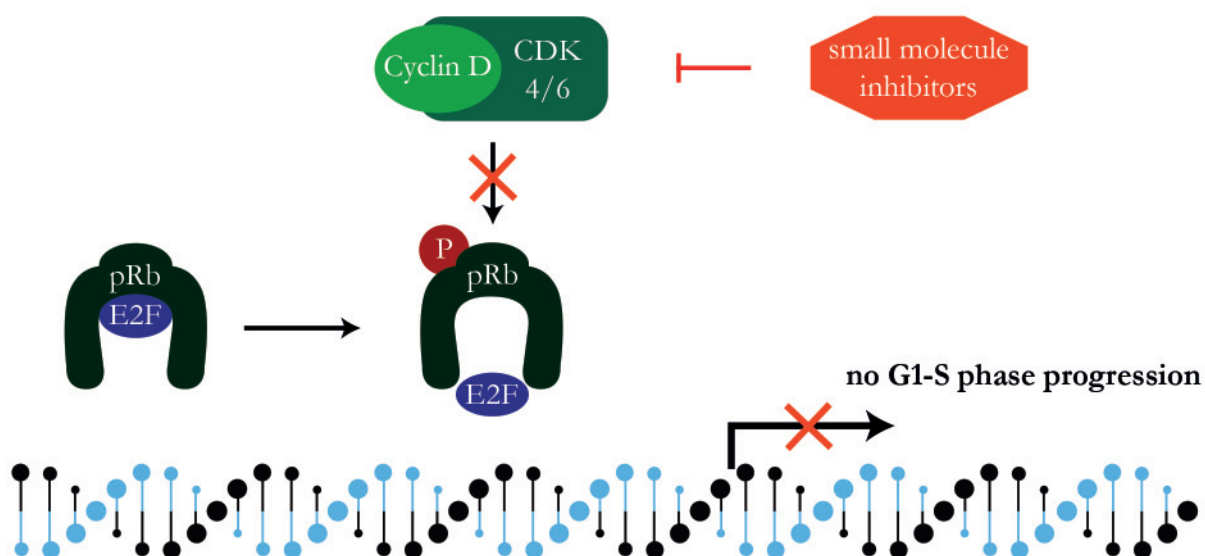
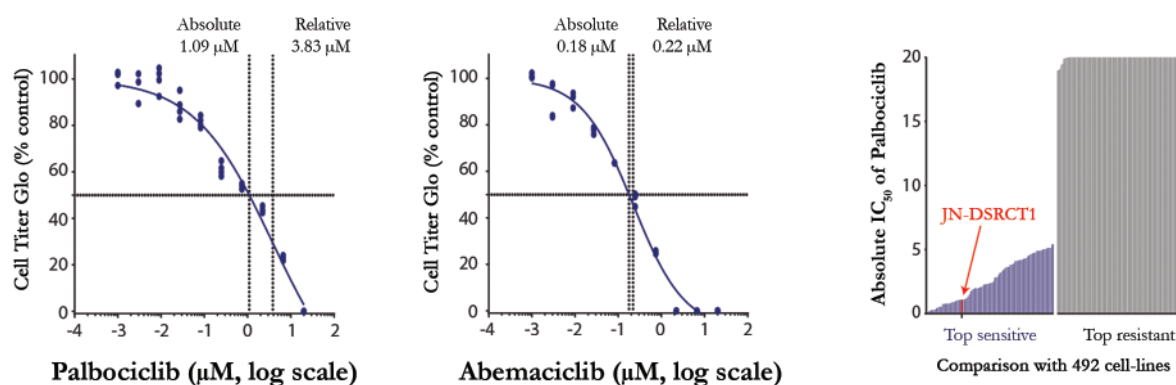


Figure 49: Effect of palbociclib on the Cyclin D-CDK4/6 complex. Cyclin D and CDK4/6 form a complex during the G1 phase. This complex can be blocked by Cyclin D1 small molecule inhibitors, such as palbociclib. Therefore, the Cyclin D-CDK4/6 complex can not phosphorylate the Rb-E2F complex, and the cell stays in the G1 phase.

### 3.4.2. DSRCT is highly sensitive to CDK4/6 small molecule inhibitors

A study comparing relative sensitivity across different tumor types was recently performed by Gong *and al.*<sup>129</sup>, where they screened more than four hundred tumor cell lines for their sensitivity toward abemaciclib and palbociclib. CDK4/6 small molecule inhibitors have never been studied as a potential treatment for DSRCT tumors. Therefore, we decided to assess palbociclib and abemaciclib as potential treatments for DSRCT patients. Based on Gong's protocol, we assessed the sensitivity of

the JN-DSRCT1 cell-line to these compounds *in vitro*. Our results confirmed that JN-DSRCT1 was highly sensitive to both abemaciclib and palbociclib, with an absolute  $IC_{50}$  (half maximal inhibitory concentration) value of 0.18  $\mu M$  and 1.09  $\mu M$ , respectively (Figure 50.B and 50.A). Gong *and al.*<sup>129</sup> proposed that highly sensitive cells are those with an absolute  $IC_{50}$  of less than 1  $\mu M$  or 1.2  $\mu M$ , based on abemaciclib or palbociclib concentration, respectively, and its active metabolites measured in breast cancer patients under treatment<sup>129-131</sup>. Therefore, when we compared our results to Gong's cell line panel, JN-DSRCT1 ranked in the top 10 highly sensitive cells (Figure 50.C).



**Figure 50: Cyclin D1 is a key regulator of proliferation in JN-DSRCT1.** Cell viability for palbociclib (A) and abemaciclib (B) with its corresponding absolute and relative  $IC_{50}$ , after two doubling times after adding the compound. (C) Comparison of the absolute  $IC_{50}$  of Palbociclib in JN-DSRCT1 vs. a panel of other cell-lines. Top 100 most sensitives and top 100 most resistant cells are displayed.

Our results suggested that the JN-DSRCT1 cell-line was sensitive to the functional disruption of the CCND1-CDK4/6 by small molecule inhibitors, such as palbociclib.

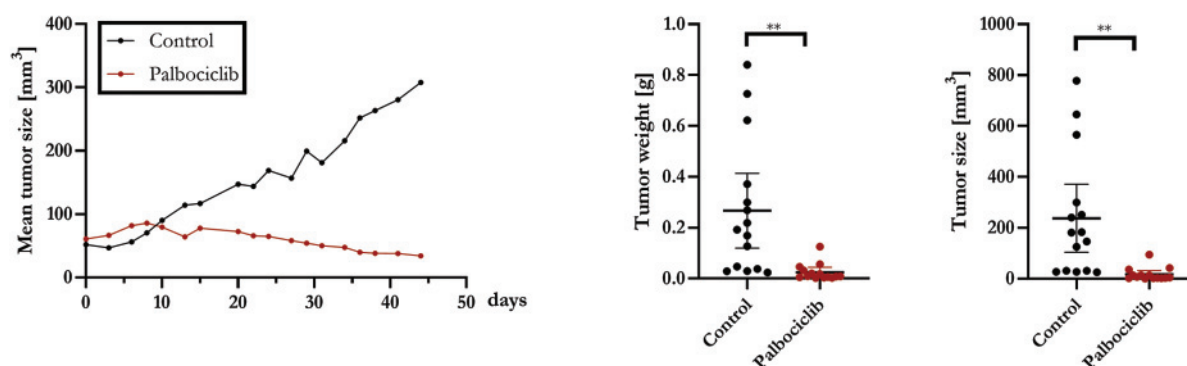
### 3.4.3. The CDK4/6 inhibitor palbociclib impairs DSRCT tumor growth *in vivo*

To expand our *in vitro* results, we investigated the effect of CDK4/6 inhibitors in a more physiological setting of DSRCT *in vivo*. We performed an *in vivo* assessment of palbociclib in JN-DSRCT1 derived tumors in NSG mice. In the clinic, palbociclib doses vary between 125mg/day to 200mg/day for 14 to 28 days of treatment<sup>5</sup>. Therefore, we decided to give 100mg/kg mice daily for 26 days, with a two days break during the weekends. Treatment was overall well-tolerated by the mice, with no obvious clinical signs of distress and no weight loss after administering the drug.

During the experiment, tumor growth was significantly impaired by the palbociclib treatment, as demonstrated by the mean tumor size curve shown in Figure 51. At the end of the experiment, tumor

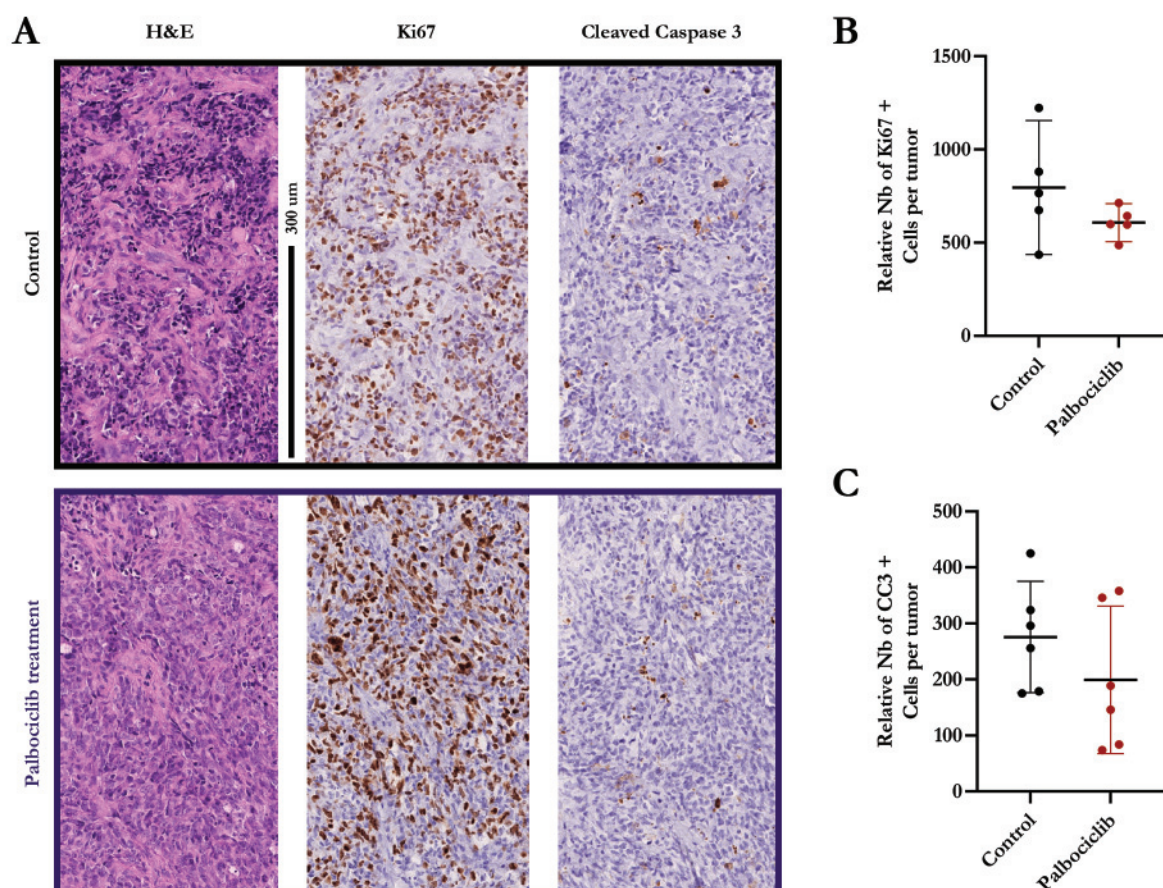


size and weight measurements confirmed a significant reduction of tumor growth upon the palbociclib treatment compared to the control group (Figure 51). Therefore, we can confirm that palbociclib potently inhibits the growth of JN-DSRCT1 derived tumors *in vivo*.



**Figure 51: Palbociclib treatment decrease tumor size of JN-DSRCT1 derived tumors.**  $2 \times 10^6$  cells JN-DSRCT1 cells were subcutaneously engrafted into a flank of NSG mouse, and tumors grew until they reach a mean of  $50 \text{ mm}^3$  in volume. Then, the rodents were treated with palbociclib  $100 \text{ mg/kg}$  mice for five weeks (5 days per week, a total of 26 days of treatment), via oral administration. Left panel: Graph showing the mean tumor volume of palbociclib treatment (in red) and control (in black) through time. Tumors were measured with an electronic caliper three times a week. Right panels are scattered plot graphs showing all data with mean value and 95% CI for the tumor volume [ $\text{mm}^3$ ] and the tumor weight [g] at the end of the experiment. \*\* pValue < 0.0021 (t-test).

All tumors were analyzed for expression of Ki67, a marker of proliferation, and cleaved caspase 3 (CC3), a marker of apoptosis, by immunohistochemistry (IHC). Surprisingly, no significant differences were observed between control and treated tumors (Figure 52). These results suggest that the tumor growth impairment was neither due to cell-cycle arrest, as tumor cells continue proliferating, nor apoptosis.



**Figure 52: Palbociclib treatment does not affect proliferation nor induce apoptosis in JN-DSRCT1-derived tumor xenografts.** A. Representative images of hematoxylin and eosin staining (H&E, left panel), and IHC for Ki67 (middle panel) and Cleaved Caspase 3 (CC3, right panel) in JN-DSRCT1 tumors two days after last palbociclib treatment. For a total of 26 days of treatment. No significant difference in either Ki67 or CC3 IHC signal was observed between palbociclib-treated (bottom panel) and control tumors (top panel). Scale bar = 300  $\mu$ m. B and C. Scattered plots showing the relative number of positive cells in palbociclib-treated (in red) vs. control tumors (in black), at the end of the experiment with mean value and 95% CI. B. Ki67 positive cells were counted in five different tumors and three different areas at 400x. C. CC3 positive cells were counted in seven different tumors and ten different areas at 400x.

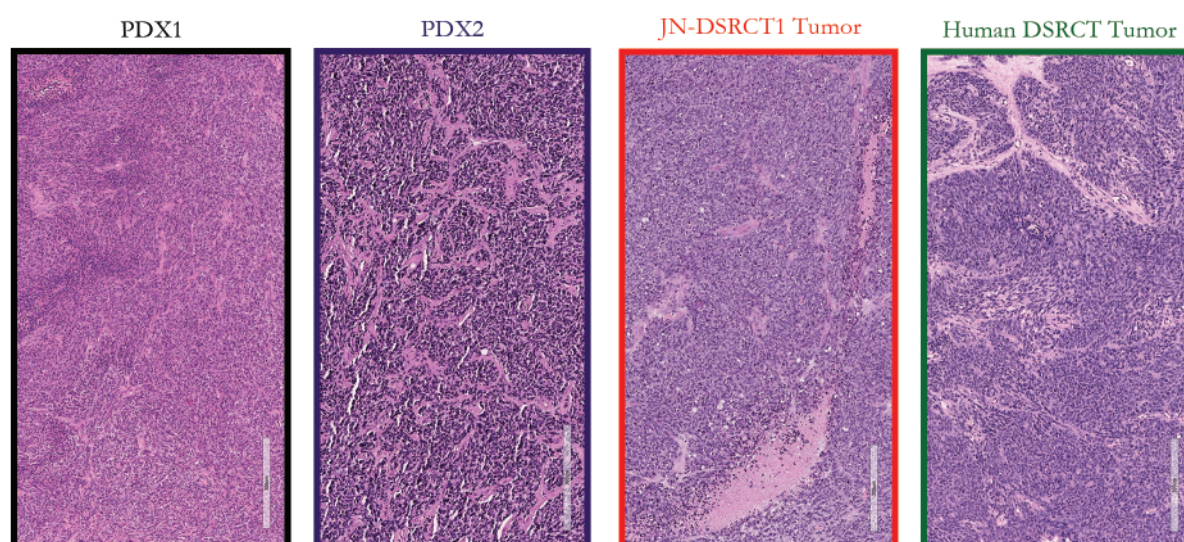
### 3.4.3.1. Primary DSRCT derived xenografts

There are only a few primary DSRCT tumor models available for basic research. In 2013, the St. Jude Children's Research Hospital of Philadelphia (USA) launched "The Childhood Solid Tumor Network" allowing researchers worldwide to access and use material and data from pediatric solid tumors to foster the study of these entities. Therefore, we requested DSRCT orthotopic patient-derived xenografts (PDX) and received three of them (see Table 9).

PDX number	SJ Number	Age	Sex	Xenograft Site	Location
PDX1	SJDSRCT046151_X1	14 yr	M	Primary	Abdomen/pelvis
PDX2	SJDSRCT049192_X1	11 yr	M	Primary	Abdomen/pelvis
PDX3	SJDSRCT046155_X1	23 yr	M	Primary	Abdomen

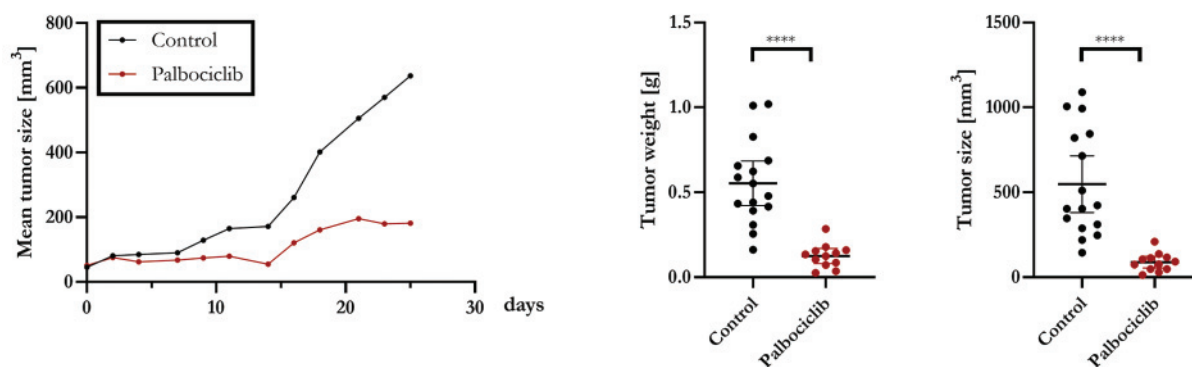
**Table 9: List of the orthotopic xenografts derived from DSRCT patients received from the St. Jude Children's Research Hospital of Philadelphia (USA).** PDX number with its corresponding reference, age, and sex of the donor along with the location of the sample collection.

The PDXs received were all engrafted in NSG mice, and after several months, two of them gave rise to tumors, with no sign of metastasis. H&E confirmed that these PDXs shared a similar morphology to primary DSRCT and JN-DSRCT1-derived tumors (see Figure 53).



**Figure 53: Histological view of St. Jude DSRCT tumors are similar to JN-DSRCT1 derived tumors and primary human DSRCT tumors.** Representative images of H&E of PDX1 (SJ DSRCT 046151 in black), PDX2 (SJ DSRCT 049192 in blue), JN-DSRCT1 derived tumor (in orange), and a human DSRCT tumor (in green). Scale bar 300µm.

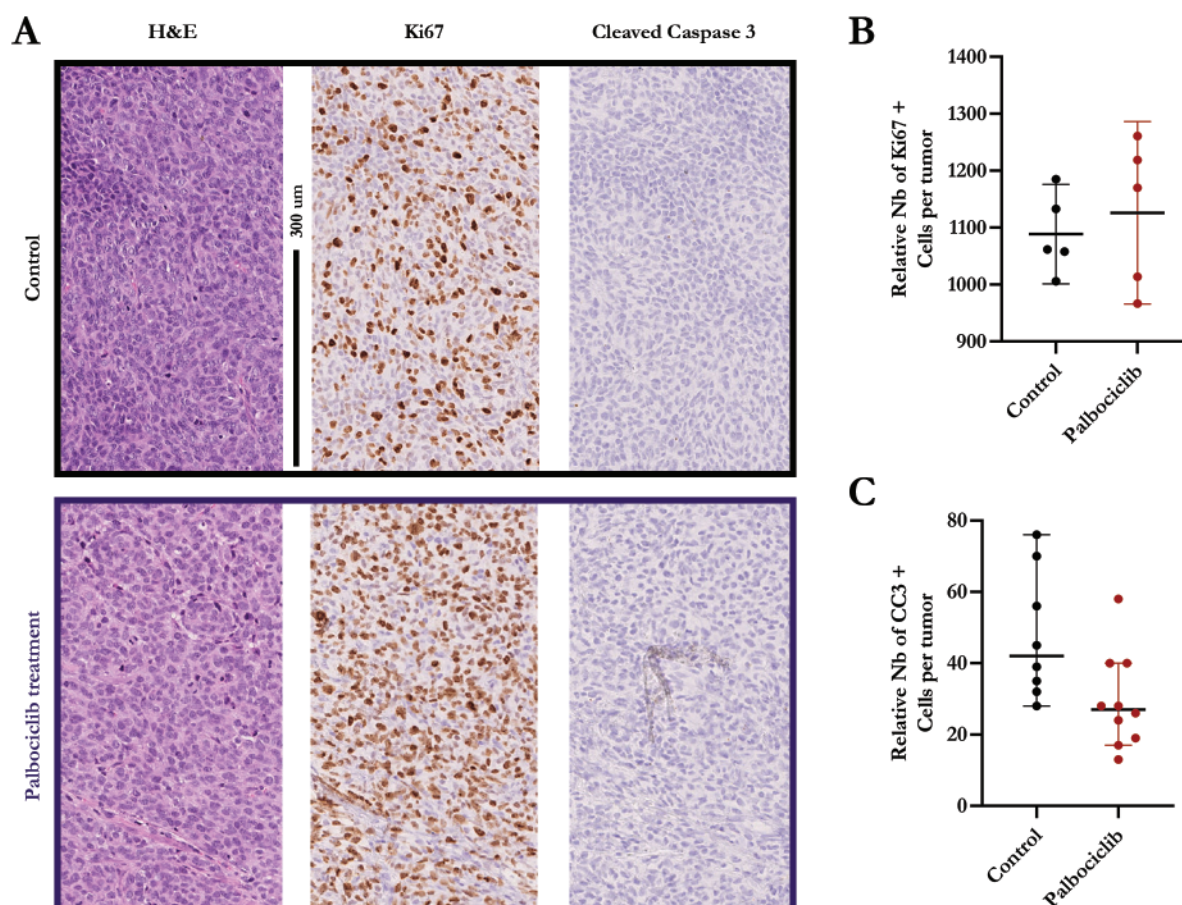
We assessed palbociclib's effect in one SJ DSRCT PDX mouse model (PDX1) treated for 19 days. We observed that tumor growth was rapidly significantly impaired upon palbociclib treatment (Figure 54 left). At the end of the experiment, tumor size and weight were significantly reduced compared to the control group (Figure 54 right). Our results with the SJ DSRCT PDX tumors were more striking than for the JN-DSRCT1 cell line, confirming that palbociclib also potently inhibits the tumor growth, a more primary DSRCT tumor model.



**Figure 54: Palbociclib treatment highly impairs DSRCT proliferation in vivo.** PDX1 tumors were subcutaneously engrafted into a flank of NSG mouse, and tumors grew until they reach a mean of 50 mm<sup>3</sup> in volume. Palbociclib treatment of 100 mg/kg mice for four weeks (5 days per week, a total of 19 days of treatment), via oral administration. Left panel: Graph showing the mean tumor volume of palbociclib treatment (in red) and control (in black) through time. Tumors were measured with an electronic caliper three times a week. Right panels are scattered plot graphs showing all data with mean value and 95% CI for the tumor volume [mm<sup>3</sup>] and the tumor weight [g] at the end of the experiment. \*\*\*\* pValue < 0.0001 (t-test).

All tumors were analyzed for expression of Ki67 and Cleaved-Caspase 3. Unfortunately, as before, no difference in proliferation nor apoptosis was observed between tumors of control and palbociclib treated mice (Figure 55).





**Figure 55: Palbociclib treatment does not affect proliferation nor induce apoptosis in human primary tumor xenografts.** A. Representative images of hematoxylin and eosin staining (H&E, left panel), and IHC for Ki67 (middle panel) and Cleaved Caspase 3 (CC3, right panel) in SJ-DSRCT PDX1 tumors two days after last palbociclib treatment. For a total of 19 days of treatment. No significant difference in either Ki67 or CC3 IHC signal was observed between palbociclib-treated (bottom panel) and control tumors (top panel). Scale bar = 300  $\mu$ m. B and C. Scattered plots showing the relative number of positive cells in palbociclib-treated (in red) vs. control tumors (in black), at the end of the experiment with mean value and 95% CI. B. Ki67 positive cells were counted in five different tumors and three different areas at 400x. C. CC3 positive cells were counted in seven different tumors and ten different areas at 400x.

## 4. Discussion

In this study, we described the epigenetic regulatory networks of the EWS-WT1 driven sarcoma. We profiled DSRCT tumor samples (both primary and a cell-line) to identify specific chromatin signature associated with the tumor. The tumor-specific chromatin signature defined distinct regulatory modules including active genes, enhancer elements, and repressive chromatin domains in DSRCTs. Further, loss-of-function studies allowed us to identify and characterize specific chromatin features of EWS-WT1 in JN-DSRCT1 cells.

The early stages of DSRCT development were studied using a cell-line model system through the ectopic expression of EWS-WT1. Further, we demonstrated that the concomitant expression of EWS-WT1 isoforms in this cell-line model system generated tumors *in vivo*. We identified specific binding motifs for each EWS-WT1 isoform. The PrDL domain of EWS-WT1 allowed its binding to the chromatin and the recruitment of EWS wild-type. Therefore, EWS-WT1 expression can modify the cells and regulate a unique target set of genes in DSRCT.

We elucidated enhancer-promoter interactions and EWS-WT1 specific chromatin loops in JN-DSRCT1 cells. Further, we identified a specific cohort of DSRCT target genes including some potential therapeutic candidates for DSRCT. We focused on the CCND1-CDK4/6 complex. Our data supported the use of CDK4/6 inhibitors in DSRCT treatment. Particularly, palbociclib treatment efficiently reduced the proliferation of JN-DSRCT1 cells *in vitro* and decreased tumor burden in PDX DSRCT tumors *in vivo*. Palbociclib could be a potential therapeutic drug for the care of DSRCTs.

The global chromatin signature and epigenetic landscape of DSRCT are not well studied. To our knowledge, this is the first study defining the chromatin landscape and transcriptional signature of DSRCT primary tumors. We profiled and characterized four primary human DSRCT tumors for its transcriptional signature and chromatin landscape. We defined the chromatin binding profile of EWS-WT1 in the JN-DSRCT1 cell-line and its functional link to active chromatin states. Our results demonstrated that most



of the EWS-WT1 peaks colocalized with H3K4me1 and H3K27ac marks at enhancer sites. H3K4me3 was significantly enriched at TSS regions upon EWS-WT1 binding, suggesting that a small portion of EWS-WT1 colocalized with H3K4me3 at promoter regions. We confirmed that EWS-WT1 bind mainly intergenic and intronic regions, and we identified two binding motifs for EWS-WT1 in JN-DSRCT1, represented by a CTCCA/CC consensus sequence and a TCC repeat motif. Our results are in concordance with Hingorani *et al.*<sup>65</sup>, which demonstrated that in the JN-DSRCT1 cell-line, EWS-WT1 is located particularly at intergenic and intronic regions of the genome and that EWS-WT1 recognized a similar CTCC cluster sequence. Using the JN-DSRCT1 cell-line, we analyzed the global changes in DSRCT transcriptome and epigenome through loss-of-function studies. The EWS-WT1 depletion led to a significant reduction of H3K27ac marks at the translocation binding sites and decreased the transcription of EWS-WT1 target genes.

The aberrant transcriptional activity of EWS-WT1 broadly disrupted the chromatin and the expression profile of tumor precursor cells leading to DSRCT development by mechanisms not yet fully understood. Studies have shown that the fusion of EWS to novel DNA binding domain creates aberrant proteins with altered DNA specificity and transactivation potential<sup>122</sup>. The fusion of WT1 with EWS generated an aberrant protein with DNA binding properties distinct from those of the WT1 tumor suppressor gene<sup>122</sup>. Furthermore, the same study revealed a clear functional difference in DNA binding between EWS-WT1 isoforms, similar to what has been reported for WT1 +KTS isoforms<sup>122</sup>.

In this study, we generated a cellular system recapitulating the early stages of DSRCT oncogenesis. Specifically, we expressed EWS-WT1 in MeT-5a cells and observed the expression of known DSRCT target genes. Furthermore, our results provided new essential information about EWS-WT1 target genes and epigenetic mechanisms that may initiate DSRCT tumorigenesis. We observed an increase in the H3K27 acetylation state of the chromatin upon EWS-WT1 binding in the MeT-5A cells. This increase was observed at the same binding sites identified earlier in the JN-DSRCT1 cell-line, revealing an active transcription of these genes in the MeT-5A cells expressing EWS-WT1.

EWS-WT1 has two isoforms varying by three amino acids (KTS) at its C-terminus. In this study, we confirmed that each EWS-WT1 isoforms bound mostly to intergenic and intragenic regions and fewer promoter regions of the chromatin, similar to EWS-WT1 in JN-DSRCT1 cells. We demonstrated

that the EWS-WT1 isoforms displayed different binding affinities to the chromatin. Particularly, we revealed that the EWS-WT1 -KTS isoform displayed a higher chromatin accessibility potential than the EWS-WT1 +KTS isoform. This binding difference might be due to the KTS insert, as in WT1 wild-type, the +KTS isoform loses the C-capping interactions stabilizing the complex with the DNA<sup>104</sup>. Studies have shown that in wild-type WT1, the WT1 +KTS isoform is an RNA binding protein<sup>54,55</sup> involved in nuclear RNA metabolism and posttranscriptional mechanisms<sup>118,120,121</sup>. We considered the possibility that the EWS-WT1 +KTS isoform might have the capacity to bind to the RNA. However, the role of this isoform needs more investigation.

We identified EWS-WT1 binding motifs for each isoform. For the EWS-WT1 -KTS isoform, the top motifs observed were the “CTCCC(A/C)C” sequence, the “CCNCNCACCCCCAC(A/C)” sequence, and a TCC repeats (5x) motif. This TCC repeated sequence was similar to the TCCTCCTCCTCCTCTCC” sequence identified 20 years ago by Liu *et al.*<sup>75</sup> in DSRCT tumor samples. For the EWS-WT1 +KTS isoform, the top motifs observed were the “(A/C)CTCCC” sequence, the “AAGTAAACA” sequence, followed by a C/T repeat sequence, and a TCC repeat motif. The binding sequences recognized in our study were not in concordance with the only EWS-WT1 +KTS binding sequence identified by Reynolds *et al.*<sup>72</sup> in 2003. Reynolds expressed the EWS-WT1 +KTS isoform in U2OS osteosarcoma cells and identified a “GGAGG(A/G)” consensus sequence. Osteosarcoma is a bone pediatric cancer displaying genetic abnormalities in their complex karyotype with no specific pattern. The genetic background of cell models could explain the difference in motifs recovered. In this study, we used MeT-5A cells, a potential DSRICT cell of origin that is more appropriate than U2OS cells.

The top binding motif identified for both EWS-WT1 isoforms was a CTCCC cluster in MeT-5A cells. These motifs were similar to the top motif we identified in the JN-DSRICT1 cell-line and primary DSRICT tumors. Furthermore, as each EWS-WT1 isoform demonstrated specific DNA binding motifs, we revealed that each EWS-WT1 isoform has a unique specific set of target genes.

Several studies have shown that the low complexity domains present in a subgroup of proteins, as the prion-like domain (PrLD) of EWS, can promote liquid-liquid phase separation and create droplet-like compartments. In eukaryotes, organelles require a lipid bilayer to carry specific functions and provide spatiotemporal control of molecular materials and biological process<sup>41,109</sup>. Membrane-less compartments

can act as organizational hubs. Therefore chromatin compartments, formed by liquid-liquid phase separation (LLPS), provide a distinct 3D structure for the constitutive heterochromatin compartmentalization to organize repetitive sequences of the DNA into compacted, silenced assemblies, usually found around nucleoli<sup>41,106</sup>. In diseases, accumulation of such proteins was observed in amyotrophic lateral sclerosis. In previous reports on Ewing sarcoma, EWS-FLI1 neomorphic properties relied on the tyrosine residues present in the PrLD of EWS<sup>116</sup>. These reports demonstrated that similar LLPS mechanisms supported protein accumulation and stabilized the binding of EWS-FLI1 to the chromatin. In this study, we investigated this LLPS property in the DSRCT context. We revealed that the tyrosine residues on the PrLD part of EWS-WT1 are necessary to mediate phase transition properties, recruit endogenous EWS, and confer the fusion protein with the ability to homo- or heterodimerize. Therefore, we also demonstrated that EWS-WT1 can bind to the chromatin in virtue of its PrLD and zinc finger domains. However, when we investigated whether the WT1 could provide shared properties linked to EWS-WT1, we observed that WT1 +KTS could not access the chromatin, unlike the EWS-WT1 +KTS isoform that can bind to the DNA. These results suggest that EWS-WT1 may have acquired neomorphic properties partly given by the EWS part of the fusion protein. The EWS-WT1 binding is required for the *de novo* establishment of active enhancers. These neomorphic properties can modify tumor precursor cells and regulate a unique set of target genes elucidated in this study.

We still need to investigate which other partners EWS-WT1 may be recruit to regulate its target genes. Based on Ewing's sarcoma studies, we hypothesize that EWS-WT1 may collaborate with BRG1/BRM-associated factor (BAF) complexes or with P300, as it was shown that EWS-FLI1 could recruit these proteins to specific loci of the DNA. The BAF complexes are known to play a critical role in diseases<sup>132,133</sup>. They are significantly involved in Ewing's and Synovial Sarcoma, where BAF complexes are hijacked to alter the transcriptome of the precursor cells during oncogenesis. To date, there is no report showing a link between BAF complexes and DSRCT. This link may need a further investigation that may lead to finding other therapeutic axis for the treatment of DSRCT. Altogether, our results provide additional data supporting the notion that EWS-WT1 plays an essential role in DSRCTs and acts as pioneer factor inducing *de novo* activation of genes that can modify tumor precursor cells and triggering DSRCT development.

To our knowledge, this is the first study showing that DSRCT tumor cells express both EWS-WT1 isoforms concomitantly. We showed that the two EWS-WT1 isoforms colocalize in the same tumor cell using the BaseScope *in situ* RNA hybridization technology. We also confirmed the presence of both EWS-WT1 isoforms in JN-DSRCT1-derived tumors and PDX xenografts received from the St. Jude Children's Research Hospital of Philadelphia (data not shown). The use of these custom-made 1ZZ KTS probes might help us detecting tumorigenic cells in new fresh DSRCT samples with an additional EWS-WT1-specific BaseScope 1ZZ probe spanning the breakpoint. Simultaneous detection of isoform-specific KTS RNA and breakpoint of EWS-WT1 could help us determine the area where tumor cells are located.

Our cellular precursor model allowed us to recapitulate the early stages of DSRCT initiation when engrafted into mice. We demonstrated that the expression of both EWS-WT1 isoforms was necessary to generate DSRCT-like tumors *in vivo*. Furthermore, a serial xenograft experiment confirmed the malignant capacity acquired by these tumors. We hypothesized that the interaction and expression of both EWS-WT1 isoforms may be required to initiate DSRCT development and that each isoform may have a different influence on the surrounding microenvironment. In this study, we observed that the tumors derived from MeT-5A cells expressing only one of the EWS-WT1 isoforms were significantly smaller and displayed a different morphology.

In sarcomas, the tumor microenvironment (TME) plays an important role in oncogenesis, not only in tumor initiation but also in metastasis. Studies have shown that tumor development strongly depends on changes in the interaction between the tumor and its direct microenvironment. It is tempting to speculate that DSRCT tumors may be able to generate their own growth signals to reduce their dependence on the surrounding microenvironment. Therefore, it would be interesting to investigate further these aspects, and to analyze whether the cells present in the stroma are also malignant. If further studies confirm that stroma-like cells surrounding DSRCTs are tumorigenic, this will have important clinical implications in DSRCT treatment. For the clinicians, identification of DSRCT tumor cells within the supposed healthy stroma might change the strategy for DSRCTs resection. Clinicians need to ensure safe tumor-free margin removal of primary tumors in patients. The high mortality of DSRCT patients might be linked to tumor cell presence in the stroma, as it is known that if the tumorigenic cells are not entirely removed, the tumor continues to proliferate, diminishing

long-term patient survival. Therefore, safe tumor-free margin removal in DSRCT patients is critical to provide a better outcome.

Despite initial favorable response to multimodal therapies, nearly all patients relapse. The rarity of DSRCT samples makes the research difficult despite the urgent need for new strategies. We identified specific EWS-WT1-regulated target genes, depicting a more defined picture of DSRCT chromatin profile and EWS-WT1 transcriptional programs. We established a list of 12 EWS-WT1 downstream target genes, which are potentially druggable. Particularly, we demonstrated a high direct regulation of the *CCND1* gene by EWS-WT1 in DSRCT. *CCND1* can be indirectly targeted by inhibition of CDK4/6, as it is known that the *CCND1*-CDK4/6 complex regulates the G1-S transition during the cell-cycle. Recently, a comparative study analyzed the sensitivity of two CDK4/6 inhibition (palbociclib and abemaciclib) in hundreds of tumor cell-lines<sup>129</sup>, including breast cancer and Ewing's sarcoma cell-lines, but lacking DSRCTs.

Many studies have shown high CDKs mRNA levels and protein expression in sarcomas, resulting in a selective advantage on growth, proliferation, cell migration, and invasion<sup>3</sup>. CDK4/6 inhibitors are used routinely to treat estrogen-receptor-positive breast cancer<sup>134</sup>, where they induce cell-cycle arrest of tumor cells. Other studies reported that inhibition of CDKs decreases sarcoma cell growth and may induce apoptosis<sup>3</sup>. It has also been reported that CDK4/6 inhibitors are a promising strategy to treat sarcomas harboring an altered version of the retinoblastoma gene<sup>135</sup>. Today, many specific CDKs inhibitors are available and some are currently under clinical evaluation for Ewing's and Synovial sarcoma treatment, providing an additional promising therapeutic strategy. All these studies and our results encouraged us to investigate further the impact of CDK4/6 inhibition in DSRCT treatment. In this study, we demonstrated that the JN-DSRCT1 cell-line was very sensitive to the functional disruption of the *CCND1*-CDK4/6 axis by these small molecule inhibitors. Specifically, palbociclib impairs *CCND1* functions, and we proved its efficacy *in vitro*. Furthermore, we demonstrated the antitumor effect of palbociclib on DSRCT *in vivo* at low doses in two tumor-derived models. Though palbociclib is known to have anti-proliferative effects, we did observe no proliferation changes nor induced apoptosis in DSRCT.

Studies revealed that CDK4/6 inhibitors target the proliferative function of Cyclin D associated kinases to induce cell-cycle exit<sup>136</sup>. CDK4/6 inhibitors can reinforce cytostasis induced by signaling pathway inhibitors. In some cases, the cells can enter quiescence or go to senescence after CDK4/6 inhibition<sup>136</sup>. In

our study, the treated tumors were still proliferating and therefore not in a quiescence or senescence state. In August 2020, EMJ published online the “Latest Developments In Advanced Breast Cancer Treatment,” where they compared treatment by CDK4/6 inhibitors. In this review, “Dr. Llombart-Cussac advocated administration of abemaciclib to patients with more aggressive scenarios and explained this was justified to avoid non-exposure of the tumor to a CDK4/6 inhibitor for one week, particularly because increases in Ki67 protein measurements have been observed when treatment is stopped.” Our results are in agreement with this statement, as xenografted tumors were collected two days after the last Palbociclib dose administration. Therefore, this may explain why we do not observe a difference in the proliferation of treated vs. control tumors. Some studies have shown that CDK4/6 inhibition can affect the maturation of the immune system’s sentinel cells and the expansion of regulatory T cells<sup>136</sup>. Since we used immunodeficient NOD/SCID c Gamma mice, the tumor growth impairment by the palbociclib treatment cannot be linked to the immune system. Studies have shown that CDK4/6 inhibitors can also alter cellular metabolism, depleting antioxidants, increasing reactive oxygen species (ROS), and triggering apoptosis<sup>136</sup>. This might be an explanation, but we did not confirm this possibility in the impairment of DSRCT tumor growth upon palbociclib treatment. In a recent publication, palbociclib treatment induced remodeling of cancer cell chromatin, leading to luminal differentiation and apoptotic evasion of breast cancer tumors<sup>134</sup>. We suppose that the effect of CDK4/6 inhibition in DSRCT may be due to chromatin remodeling changes in the cancer cells epigenome, as we observed no proliferation changes nor induced apoptosis but a dramatic change in tumor burden upon palbociclib treatment. This effect need further investigation, however, CDK inhibitors may represent a promising candidate to add to current multimodal therapies for DSRCT management.

Find new potential drugs for the treatment of DSRCT is crucial, however ensure delivery of the drug to the tumor is also essential in the care of DSRCT patients. Most of the time, drug delivery relies on the enhanced permeability and retention effects, but often this strategy fails to deliver a uniform distribution of the drug across desmoplastic tumors due to the deregulated extracellular matrix<sup>137</sup>. Furthermore, tumors often display abnormal tumor vasculature with an irregular shape and a disordered structure<sup>138</sup>, diminishing drug delivery. The TME may help the tumor to progress, invade and metastasize, but also to acquire drug resistance, including drug delivery-related and sensitivity-related resistance. In DSRCT, suboptimal treatment efficacy may be linked to the dense stroma and the relative absence of vessels in the TME. It would also be relevant in the fight against DSRCT to find other ways



to deliver the drug directly and uniformly into the tumors, using nanoscale drug carriers or organic polymer particles for example, thus decreasing drug-related side effects and increasing the efficiency of the treatment. Since DSRCT displays low response rates to conventional chemotherapeutic agents and classical delivery strategies, new approaches are needed to improve DSRCT patients' lives.

## 5. Future directions

Our current findings lay down a solid background on the molecular determinants of DSRCT tumorigenesis. However, additional mechanistic studies are necessary to have a deeper understanding of DSRCT biology and regulatory mechanisms. Therefore, we propose a few of our immediate future goals to provide additional insights into DSRCT pathogenesis and potential therapeutic strategies.

### 5.1. Study hpMSCs transcriptome

Sarcomas are rare cancers of mesenchymal origin that give rise to more than 80 distinct histologic subtypes. It would be interesting to determine the epigenetic background of mesenchymal stem-cell (MSC) to understand better the features that provide a permissive environment for developing such cancers. MSCs can be extracted from healthy donors without significant ethical issues, compared to hematopoietic bone marrow or embryonic stem-cells, which usages are highly debated. In Stamenkovic's and Riggi's laboratories, hpMSCs are used in other studies investigating sarcomas, particularly Ewing's sarcoma and Synovial sarcoma. This study reported that the hpMSCs were immune-phenotypically similar but exhibit background differences in gene expression. Our laboratory is interested in the heterogeneous features of hpMSCs, which can be promoted under *in vitro* culture conditions or cryopreservation<sup>99,100</sup>. Cryopreservation on MSCs is not well-documented, but a study revealed that prolonged storage could affect specific properties of MSCs<sup>99</sup>. Cold shock or toxicity of DMSO can reduce cell viability and change the original gene expression profile and MSC morphology<sup>99</sup>. In the future, we would like to investigate potential genetic expression heterogeneity among hpMSCs. To this aim, we are planning to do a single-cell gene expression analysis in collaboration with Prof. Suvà's laboratory at the Broad Institute in Boston, Massachusetts – USA. We hope this study will provide essential information about the standard gene expression profile of primary MSC cells and help understand healthy cells' mechanisms and cellular background and determine the cellular genetic niche for sarcomas development.

## 5.2. Study DSRCT tissue heterogeneity

Sarcomas are very aggressive soft-tissue and bone tumors with high intratumoral heterogeneity, invasiveness, and metastatic potential<sup>12,16,17</sup>. In DSRCT, cells express EWS-WT1 +KTS, and EWS-WT1 -KTS isoforms activating a set of different target genes. These substantial changes in the transcriptome may contribute to the marked inter-tumoral heterogeneity observed in DSRCTs<sup>28</sup>. Therefore, we would like to investigate the epigenetic determinants of primary DSRCT tumors' heterogeneity by single-cell RNA sequencing. We hope that this future study will provide essential information about the different subpopulations composing DSRCT tumors and their gene expression signatures and differentiation states. This could help to identify novel biomarkers to improve DSRCT prognosis.

## 5.3. Identify EWS-WT1 collaborative protein complexes

Some studies have shown that oncogenic proteins form complexes with other proteins to target specific chromatin regions in sarcoma. EWS-WT1 may probably recruit other binding partners to form protein complexes and ensure their constant expression and action. To our knowledge, no study has identified protein partners that mediate EWS-WT1 stabilization and function. It would be relevant to analyze EWS-WT1 isoform binding partners separately to determine whether they could influence the tumor's behavior. It would be interesting to characterize these potentially distinct complexes to define the overall picture of EWS-WT1 activity. Identification of proteins interacting directly with EWS-WT1 could be assessed by mass spectrometry. Therefore, we may be able to disrupt EWS-WT1 binding to partners as a possible new therapy.

## 5.4. Impede EWS-WT1 expression in xenograft tumors

There is no genetically engineered mouse model to study DSRCT development. The reason of this absence may probably lie on the unknown cell of origin of DSRCT, the histology of the tumor, and the lack of commercially available cell-lines. For this reason, we generated a cellular model recapitulating the early stages of DSRCT. However, it would be interesting to investigate whether the inhibition of EWS-WT1 may impede tumor growth and proliferation. To this aim, we would like to stop the expression of EWS-WT1 in fully formed tumors. Therefore, we could observe chromatin perturbation or other epigenetic changes after EWS-WT1 removal. This may provide us useful information on pathways that may be interesting in the treatment of DSRCT.

### 5.5. Identify the roles of EWS-WT1 isoforms

In this study, we demonstrated that DSRCT cells express both EWS-WT1 isoforms concomitantly and that each isoform has its unique set of target genes. It would be interesting to investigate the epigenetic landscape of DSRCTs when only one of the EWS-WT1 isoforms is knockdown. This approach could inform us about the impact of each isoform in DSRCT. It could also be interesting to ask whether the absence of one isoform could impair cell growth. Identification of specific pathways for each isoform will help us determine each EWS-WT1 isoforms' role more precisely and lead to a combinational treatment to improve the DSRCT survival rate.

### 5.6. Effect of palbociclib during early stages of DSRCT treatment

Our results showed that EWS-WT1 strongly regulates CCND1 expression in DSRCTs, and that treatment by palbociclib (a CDK4/6 inhibitor) significantly reduces tumor growth. Nevertheless, no difference in proliferation or apoptosis was observed in DSRCT xenografts at the end of palbociclib treatment, as assessed by Ki67 and CC3 IHC. It would be interesting to investigate whether palbociclib affects tumor cell proliferation and apoptosis during the early stages, by assessing these parameters after five, ten, and twenty days of treatment. Furthermore, xenografts should be collected immediately after the last palbociclib dose.

### 5.7. Publication of the results

Most of the content of this thesis is scheduled to be published in a peer-review journal over the next 6 months. In addition, this work will also be presented during future conferences, as proposed talk or at poster sessions.

Cancer remains a challenging research field, and every discovery of new oncogenic mechanisms often raise more questions than they answer. However, those questions might also pave the way for a better global understanding of cancer and identify innovative strategies to fight and ultimately cure it. In the case of the DSRCT, our findings have suggested a new therapeutic target that could, with a strong researchers' and clinicians' collaboration, evolve towards a possible new orientation for patients' treatment. Although all the therapeutic targets proposed still need to go through clinical assays to test their efficacy in DSRCT patients, they might be beneficial for treating this disease.

## 6. Collaboration to publications

In the course of my PhD, I actively collaborated with Prof. Rivera's laboratory and Prof. Suvà's laboratory at the Massachusetts General Hospital (MGH, Boston - USA), and Prof. Stamenkovic's laboratory at the CHUV (Lausanne, Switzerland), and significantly contributed to some of their projects. This scientific collaboration resulted in the publication of important articles in the field of sarcomas. In this chapter, I would like to introduce two sarcomas: Ewing's sarcoma and Synovial Sarcoma. Furthermore, I will focus briefly on BRG1/BRM-associated factor (BAF) complexes as these complexes are hijacked in these two sarcomas during their oncogenesis. At the end of the chapter, you will find information about the publications, such as authors' names, abstract and my contribution to each of them.

### 6.1. Ewing's Sarcoma

Ewing's sarcoma is one of the most studied pediatric sarcomas in the world. This cancer is highly malignant, with poorly differentiated small round blue cell tumors caused by a chromosomal translocation driving the development of the cancer<sup>16</sup>.

Ewing's sarcoma is a malignancy originally described by Dr. James Ewing's in 1921, differentiating it from osteosarcoma. Nowadays, this rare sarcoma represent 1.5% of all childhood cancers and is the second most common primary bone malignancy worldwide in young adults and children<sup>16</sup>. Around the world, the incidence of Ewing's sarcoma is 1-3 cases per million persons per year<sup>16</sup>, with a male to female ratio of 6 to 4<sup>38,39</sup>. Another factor influencing the incidence of Ewing's sarcoma is the population background. For instance, there are tenfold less cases in the African American population compared to Europeans (66% in White, 24% in Hispanic, 6% in Asian and 3% in Black population)<sup>38-40</sup>.

The primary site of Ewing's Sarcoma is the pelvis, but it can also arise in bone extremities, and outside the bones, in regions such as cartilage, trunk, head and neck. Symptoms of patients are: localized bone pain, swelling and fever. Additionally, affected children can break their bones with no apparent reason.

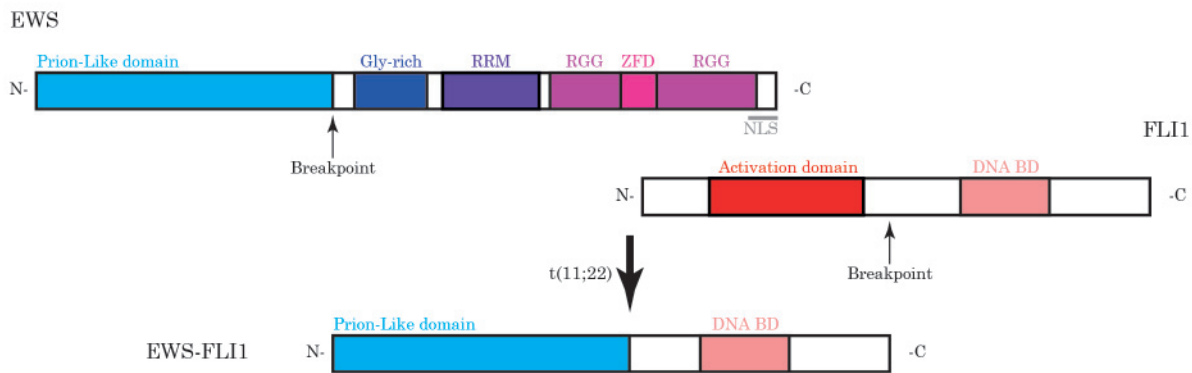
Diagnosis of Ewing's Sarcoma is achieved using magnetic resonance imaging (MRI), computed tomography (CT) or positron emission tomography (PET) scans, but also by doing a biopsy, X-rays or DNA analysis. Unfortunately, by time of diagnosis, 25% of patients have metastasis, usually in the lungs, other bones or bone marrow<sup>19,139,140</sup>. Therefore overall survival rate is poor: 50% after 5 years and it drops to 30% after 10 years<sup>16</sup>. Unfortunately, the cell from which originates Ewing's sarcoma is still unknown, but evidences show bone-marrow derived Mesenchymal Stem Cell (MSCs) as a potential cell of origin<sup>16,23,94</sup>.

One of the main advances in understanding the molecular basis behind Ewing's sarcoma was the discovery in 1992 of the characteristic chromosomal translocation  $t(11;22)(q24;q12)$ , occurring in 85% of cases<sup>19</sup>. This translocation forms the aberrant fusion protein EWS-FLI1, where the N-terminal part of EWS fuses to the C-terminal part of FLI1<sup>16</sup>. EWS-FLI1 can activate target genes at GGAA micro-satellite repeats and induce de novo activation of enhancer from a closed chromatin<sup>36,37</sup>.

#### 6.1.1. EWSR1 and FLI1 genes

As described before, EWS belongs to the FET family of proteins. EWSR1 has a prion-like domain (PrLD) at its N-terminal, which is enriched in particular amino acids (glutamine, tyrosine, asparagine and serine)<sup>41,42</sup>. The *FLI1* (Friend Leukemia virus Integration site 1) gene belongs to the Erythroblastosis-virus-associated Transforming Sequences (ETS) family of DNA-binding transcription factors<sup>16,19</sup>. FLI1 is expressed in hematopoietic cells<sup>16</sup>, but it is also known to control growth and development of cells by regulating transcription of specific genes. A schematic illustration of EWS, FLI1 and EWS-FLI1 fusion protein domains are detailed in Figure 56.





**Figure 56: Schematic representation of EWS, FLI1 and type 1 EWS-FLI1 fusion protein and its corresponding domains.** EWS has a Prion-Like domain, a Glycine-rich domain (Gly-rich), a RNA-recognition motif (RRM), two arginine-glycine-glycine-rich (RGG) domains separated by one zinc finger domain (ZFD), and a C-terminal nuclear localization sequence (NLS). FLI1 has an activation domain and a DNA binding domain (DNA BD). In the EWS-FLI1 fusion protein type 1, the first 1–264 amino acids of the EWSR1 are fused with the 219–452 amino acids of FLI1, leading to the expression of the aberrant transcription factor EWSR1-FLI1. Smooth arrows indicate the region of in frame chromosome translocation. Bold arrow indicates the t(11;22) translocation result.

Two forms of *EWSR1-FLI1* exist with varying chromosomal break points, in the EWSR1 part<sup>19</sup>. In 10-15% of Ewing's sarcoma cases, other fusion proteins have been observed, where a t(21;12)(22;12) generates the EWSR1-ERG fusion protein<sup>16</sup>. Nevertheless, other fusion proteins were also reported<sup>141</sup> and they are summarized in Table 10.

Repeat elements represent a large portion of human genome with unknown biological functions. In Ewing's sarcoma, EWS-FLI1 behaves as a pioneer factor at GGAA microsatellite repeats. In one of our publications, we analyzed the effect of the PrLD of EWS-FLI1, because it recruits BAF complexes to GGAA microsatellite sequences. This recruitment changes the DNA conformation and represents a neomorphic properties of EWSR1-FLI1 compared to FLI1 wild-type, functional differences rely on the tyrosine residues on the PrLD of EWSR1, therefore allowing de novo target activation promoting and maintaining Ewing's Sarcoma.

FET Family Partner	Fusion With ETS-like Oncogene Partner	Translocation	Proportion
<i>EWSR1</i>	<i>EWSR1-FLI1</i>	t(11;22)(q24;q12)	~85%
	<i>EWSR1-ERG</i>	t(21;22)(q22;q12)	~10%
	<i>EWSR1-ETV1</i>	t(7;22)(p22;q12)	Rare
	<i>EWSR1-ETV4</i>	t(17;22)(q12;q12)	Rare
	<i>EWSR1-FEV</i>	t(2;22)(q35;q12)	Rare
	<i>EWSR1-NEAT2a</i>	t(20;22)(q13;q12)	Rare
	<i>EWSR1-POU5F1a</i>	t(6;22)(p21;q12)	
	<i>EWSR1-SMARCA5a</i>	t(4;22)(q31;q12)	Rare
	<i>EWSR1-ZSGa</i>	t(6;22)(p21;q12)	
	<i>EWSR1-SP3a</i>	t(2;22)(q31;q12)	Rare
<i>FUS</i>	<i>FUS-ERG</i>	t(16;21)(p11;q22)	Rare
	<i>FUS-FEV</i>	t(2;16)(q35;p11)	Rare

**Table 10: List of translocations detected in Ewing's Sarcoma.**

We further investigate the functional implication of this specific gene regulation, in another of our publications, where we demonstrated that silencing these repeat elements impairs local nascent transcription and decrease putative Ewing's sarcoma genes. Particularly, disrupting EWS-FLI1 binding to GGAA repeat near the sex-determining region Y-box (SOX) 2 locus that reduces growth of Ewing's sarcoma xenografts. Therefore, repetitive sequences in the genome can make dangerous contribution to oncogenic programs.

## 6.2. BRG1/BRM-associated factor (BAF) complexes

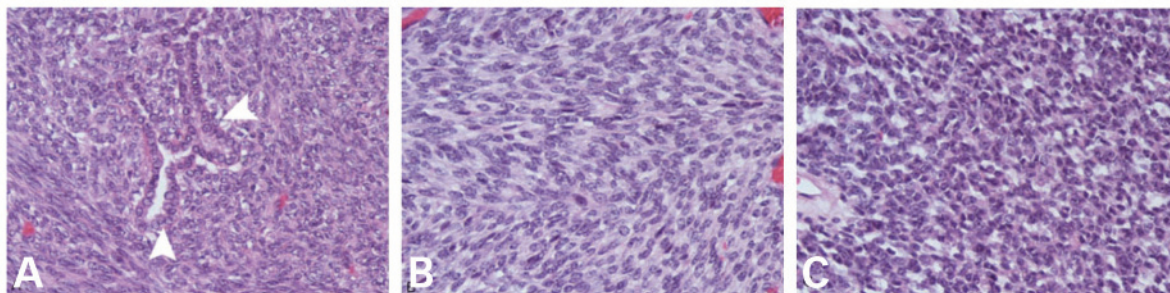
BAF complexes are known to play a critical role in diseases<sup>132,133</sup>. They are particularly involved in Ewing's and Synovial Sarcoma, where BAF complexes are hijacked to change the transcriptome of the cells during oncogenesis. Therefore, it is appropriate to understand their normal function within the cell.

The BAF complex is an adenosine triphosphate (ATP)-dependent chromatin remodeler controlling part of the epigenetic mechanisms occurring in the cell<sup>133,142,143</sup>. It allows the decompaction and compaction of DNA in the nucleus, whereas maintaining the capacity for selective gene expression, replication, and DNA repair and recombination<sup>133,144</sup>. Therefore, BAF can drive the transition from inaccessible higher order chromatin structure towards accessibility due to the direct eviction of Polycomb repressive complexes (PRC)<sup>132,142,144,145</sup>.

BAF complexes are mutated in 20% of human cancers and their mutations are tissue-specific<sup>132,133,142-144,146</sup>. In tumors, BAF complex can behave as oncogenes or tumor suppressors<sup>132,142,144,146</sup>. One well-studied mutation of chromatin remodelers that cause malignancy is the SS18-SSX fusion oncogene in Synovial Sarcoma.

### 6.3. Synovial sarcoma

Synovial sarcoma is an aggressive soft tissue tumor that represents 10% of all human soft tissue malignancies and commonly arises in the extremities of adolescents and young adults<sup>17,143,147-151</sup>. The peak incidence of this disease happens at 15-29 years and the male to female ration is 1.2 to 1<sup>152</sup>. Unfortunately, despite surgical resection, radiotherapy, and chemotherapy, the 5-year survival rate is 60%, but drops to 20% after 10 years<sup>152</sup>. This low survival rate may be linked to an immune evasion mechanism as this neoplasm is characterized by low T cell infiltration. This cold tumor has three histologic subtypes: monophasic, biphasic, and poorly differentiated<sup>18,148,149,151</sup> (Figure 57).



**Figure 57 : Histology subtypes of synovial sarcoma.** A: biphasic displaying glandular structures (arrow-heads). B: monophasic composed predominantly of densely packed spindle cells. C: poorly differentiated composed of round cells with scant cytoplasm that falls into the morphologic category of small round blue cell tumors. Modified from the Swiss Medical Weekly, <https://smw.ch/article/doi/smw.2018.14667>

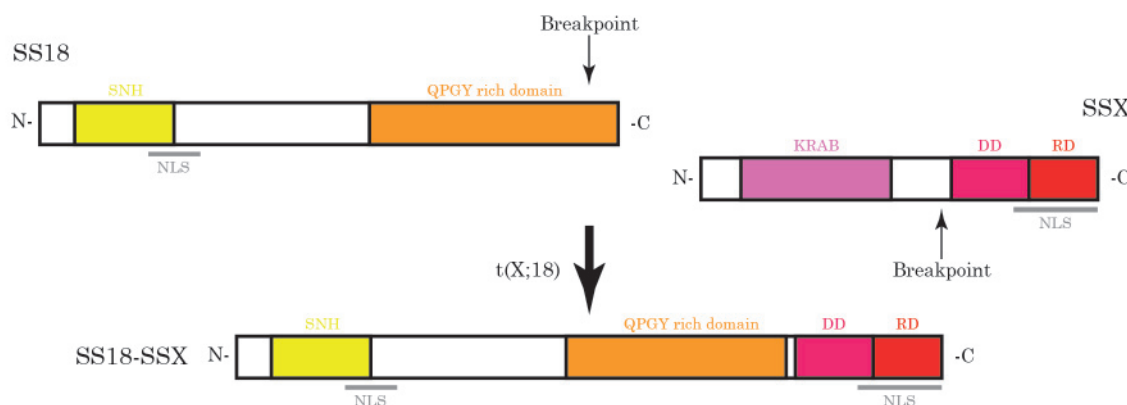
The tissue or cell of origin for Synovial sarcoma is still unknown, but there are several hypotheses. One of them is that Synovial sarcoma may arise from an unidentified multipotent stem-cell capable of mesenchymal and neuroectodermal differentiation<sup>147,153</sup>. Another hypothesis is that this sarcoma may arise from a skeletal muscle lineage with broad PRC2 marked domains, related to activation of neural genes implicated in oncogenesis<sup>133,148,154</sup>.

All Synovial sarcoma patients have a specific chromosomal translocation t(X;18)(p11.2;q11.2) that fuses exactly the last 78 amino acids of the protein SSX (SSX1, SSX2 and rarely SSX4) replacing the

8 C-terminal amino acids of the dedicated BAF complex subunit SS18<sup>21,132,133,142-144,146,150,151,155</sup>. Interestingly, the incorporation of the in-frame SS18-SSX fusion protein into the BAF complex removes its BAF47 tumor suppressor subunit<sup>133,143</sup>, therefore, dysregulates its normal function. Despite the characteristic t(X;18) translocation, synovial sarcoma lacks recurrent secondary mutations, meaning that only the SS18-SSX fusion protein initiates and drives its oncogenesis<sup>132,133,143,147</sup>.

### 6.3.1. SS18 and SSX genes

The SS18 subunit of BAF chromatin remodeling complex (SS18) is ubiquitously expressed in all human tissue, it is localized in the nucleus exhibiting a speckled nuclear distribution<sup>150</sup>. SS18 normally interacts with BRM and BRG1 ATPases and P300 histone acetyl transferase, promoting cell adhesion and fibronectin matrix formation<sup>21,147,148,155</sup>. The synovial sarcoma X (SSX) expression is limited to testis, it is located in the nucleus where it binds core histones and associates with mitotic chromosomes<sup>155</sup>. SS18 and SSX lack binding domains, but they are assigned to distinct nuclear domains. Therefore, binding partners may play a crucial role in the location of these proteins<sup>148,150,155</sup>. For a schematic illustration of SS18, SSX and SS18-SSX fusion protein domains (Figure 58).



**Figure 58 : Schematic representation of the domains of SS18, SSX and SS18-SSX fusion protein and its corresponding domains.** SS18 has a SS18 N-terminal homology (SNH), a nuclear localization signal (NLS) and a glutamine-proline-glycine-tyrosine (QPGY) rich domain. SSX has a Kruppel-associated box domain (KRAB), a divergent domain (DD), a repression domain (RD) and a NLS at its C-terminus. Smooth arrows indicate the region of in frame chromosome translocation. Bold arrow indicates the t(X;18) translocation result.

### 6.3.2. Synovial sarcoma target genes

In Synovial sarcoma, the expression of the chimeric fusion protein causes genome-wide changes in BAF targeting genes, with the emergence of broad peaks of around 2236 bp wide<sup>133</sup>. During this gain of chromatin occupancy, SS18-SSX retargets the BAF complex from enhancers to broad polycomb

domains in order to oppose the PRC2 mediated repression and activate bivalent genes<sup>132,133,142,144</sup>. One example of these retargeted oncogenes locus, is the *SRY (sex determining region Y)-box 2 (SOX2)* locus, which upon BAF relocation, SOX2 expression is activated. SOX2 expression allows reprogramming of mature cells in pluripotent stem cells, giving the stem-cell like gene expression pattern of Synovial Sarcoma<sup>133,142,143</sup>. Multiple developmental pathways are upregulated in Synovial sarcoma, among them are EGF, FGF, TGF- $\beta$ , Notch, Wnt, Hedgehog and retinoic acid receptor signaling pathway, but also genes involved in neural development<sup>156</sup>.

In our recent publication, we explore the cellular plasticity and the core oncogenic program of 12 primary synovial sarcomas by single cell sequencing. Here, we demonstrate that SS18-SSX sustain the core oncogenic program, blocks differentiation and dysregulates key transcriptional cofactors like histone deacetylases and cell cycle kinases promoting tumor growth. Furthermore, Tumor necrosis factor (TNF) and interferon (IFN) expression can repress the SS18-SSX program. We confirmed that modulation of epigenetic and cell cycle processes can impair immune evasion mechanism of this malignancy. More precisely, in this cold tumor, inhibition of histone deacetylases (HDAC) and CDK4/6 simultaneously, sensitizes the tumor to T-cell mediated killing and improves T cell reactivity. Therefore, this inhibition helps the recruitment of immune cells in the tumor microenvironment and therefore help fight the cancer and improve overall survival of patients.

In our most recent publication, we developed Synovial sarcoma models of primary tumor organoids. We used these models and mesenchymal precursors to define specific chromatin remodeling mechanisms and dependencies in Synovial sarcoma. Our genome-wide epigenomic profile revealed that SS18-SSX induced broad BAF domains at its binding sites coupled with H3K27me3 eviction and ubiquitin deposition at lysine 119 on histone 2A (H2AK119ub). These epigenomic changes lead to the establishment of *de novo* active regulatory elements that drive Synovial sarcoma. This work helped to define the mechanisms of epigenetic dysregulation on which Synovial sarcoma cells are dependent.

## Publication 1:

### Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain.

**Authors:** Gaylor Boulay,<sup>1,3,5,6</sup> Gabriel J. Sandoval,<sup>2,3,6</sup> Nicolo Riggi,<sup>4</sup> Sowmya Iyer,<sup>1</sup> Rémi Buisson,<sup>5</sup> Beverly Naigles,<sup>1,5</sup> Mary, E. Awad,<sup>1,5</sup> Shruthi Rengarajan,<sup>1,5</sup> Angela Volorio,<sup>1,4,5</sup> Matthew J. McBride,<sup>2,3</sup> Liliane C. Broye,<sup>4</sup> Lee Zou,<sup>5</sup> Ivan Stamenkovic,<sup>4</sup> Cigall Kadoch,<sup>2,3,7,\*</sup> and Miguel N. Rivera<sup>1,3,5,7,8,\*</sup>

<sup>1</sup> Department of Pathology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA,

<sup>2</sup> Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA 02215, USA,

<sup>3</sup> Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA,

<sup>4</sup> Institute of Pathology, Centre Hospitalier Universitaire Vaudois, Faculty of Biology and Medicine, University of Lausanne, 1011 Lausanne, Switzerland,

<sup>5</sup> Center for Cancer Research, Massachusetts General Hospital and Harvard Medical School, Charlestown, MA 02129, USA,

<sup>6</sup> These authors contributed equally,

<sup>7</sup> Senior author,

<sup>8</sup> Lead Contact.

Corresponding author e-mail address: MNRIVERA@mgh.harvard.edu

#### ABSTRACT

Alterations in transcriptional regulators can orchestrate oncogenic gene expression programs in cancer. Here, we show that the BRG1/BRM-associated factor (BAF) chromatin remodeling complex, which is mutated in over 20% of human tumors, interacts with EWSR1, a member of a family of proteins with prion-like domains (PrLD) that are frequent partners in oncogenic fusions with transcription factors. In Ewing's sarcoma, we find that the BAF complex is recruited by the EWSR1-FLI1 fusion protein to tumor-specific enhancers and contributes to target gene activation. This process is a neomorphic property of EWSR1-FLI1 compared to wild-type FLI1 and depends on tyrosine residues that are necessary for phase transitions of the EWSR1 prion-like domain. Furthermore, fusion of short fragments of EWSR1 to FLI1 is sufficient to recapitulate BAF complex retargeting and EWSR1-FLI1 activities. Our studies thus demonstrate that the physical properties of prion-like domains can retarget critical chromatin regulatory complexes to establish and maintain oncogenic gene expression programs.

**My contribution:** extracted, cultured, expanded and tested human pediatric Mesenchymal Stem Cells (hpMSCs) derived from healthy donors, used for the experiments. Provided also conceptual advice on the project.

**Cell. VOLUME 171, ISSUE 1, P163-178.E19, SEPTEMBER 21, 2017. Published online 2017 Aug 24. DOI: 10.1016/j.cell.2017.07.036**



## Publication 2:

### Epigenome editing of microsatellite repeats defines tumor-specific enhancer functions and dependencies.

**Authors:** Gaylor Boulay,<sup>1,2,3,5</sup> Angela Volorio,<sup>1,2,4,5</sup> Sowmya Iyer,<sup>1,2</sup> Liliane C. Broye,<sup>4</sup> Ivan Stamenkovic,<sup>4,6</sup> Nicolo Riggi,<sup>4,6</sup> and Miguel N. Rivera<sup>1,2,3,6</sup>.

<sup>1</sup> Department of Pathology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114, USA;

<sup>2</sup> Center for Cancer Research, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114, USA;

<sup>3</sup> Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA;

<sup>4</sup> Institute of Pathology, Department of Experimental Pathology, Centre Hospitalier Universitaire Vaudois, University of Lausanne, 1011 Lausanne, Switzerland;

<sup>5</sup> These authors contributed equally to this work;

<sup>6</sup> Senior authors.

Corresponding author e-mail address: MNRIVERA@mgh.harvard.edu

#### ABSTRACT

Various types of repetitive sequences are dysregulated in cancer. In Ewing's sarcoma, the oncogenic fusion protein EWS-FLI1 induces chromatin features typical of active enhancers at GGAA microsatellite repeats, but the function of these sites has not been directly demonstrated. Here, by combining nascent transcription profiling with epigenome editing, we found that a subset of GGAA microsatellite repeats is transcriptionally active in Ewing's sarcoma and that silencing individual repeats abolishes local nascent transcription and leads to markedly reduced expression of putative target genes. Epigenome silencing of these repeat sites does not affect gene expression in unrelated cells, can prevent the induction of gene expression by EWSR1-FLI1, and, in the case of a GGAA repeat that controls SOX2 expression from a distance of 470 kb, is sufficient to impair the growth of Ewing's sarcoma xenografts. Using an experimental approach that is broadly applicable to testing different types of repetitive genomic elements, our study directly demonstrates that specific repeat microsatellites can have critical gene regulation functions in cancer and thus represent tumor-specific vulnerabilities that may be exploited to develop new therapies.

**My contribution:** extracted, cultured, expanded and tested human pediatric Mesenchymal Stem Cells (hpMSCs) derived from healthy donors, used for the experiments. Performed the *in vivo* tumorigenic assays with Ewing's cell lines A673 and SK-N-MC (figure 6B, and 6C).

**Genes Dev. 2018 Aug 1; 32(15-16): 1008-1019. DOI: 10.1101/gad.315192.118**

## Publication 3:

### Opposing immune and genetic forces shape oncogenic programs in synovial sarcoma

**Authors:** Livnat Jerby-Arnon<sup>1,2,†</sup>, Cyril Neftel<sup>1,2,3,4,†</sup>, Marni E. Shore<sup>1,2,3</sup>, Matthew J. McBride<sup>5</sup>, Brian Haas<sup>1</sup>, Benjamin Izar<sup>1,6</sup>, Hannah R. Weissman<sup>1,2,3</sup>, Angela Volorio<sup>1,3</sup>, Gaylor Boulay<sup>1,3</sup>, Luisa Cironi<sup>4</sup>, Alyssa R. Richman<sup>1,2,3</sup>, Liliane C. Broye<sup>4</sup>, Joseph M. Gorski<sup>7</sup>, Christina C. Luo<sup>3</sup>, Ravindra Mylvaganam<sup>3</sup>, Lan Nguyen<sup>1</sup>, Shaolin Mei<sup>8</sup>, Johannes c. Melms<sup>6</sup>, Christophe Georgescu<sup>1</sup>, Ofir Cohen<sup>1</sup>, Jorge E. Buendia-Buendia<sup>6</sup>, Michael S. Cuoco<sup>1</sup>, Danny Labes<sup>9</sup>, Daniel R. Zollinger<sup>10</sup>, Joseph M. Beechem<sup>10</sup>, G. Petur Nielsen<sup>3</sup>, Ivan Chebib<sup>7</sup>, Gregory Cote<sup>11</sup>, Edwin Choy<sup>11</sup>, Igor Letovanec<sup>4</sup>, Stéphane Chérix<sup>12</sup>, Nikhil Wagle<sup>6</sup>, Peter K. Sorger<sup>8</sup>, Alex B. Haynes<sup>13</sup>, John T. Mullen<sup>13</sup>, Ivan Stamenkovic<sup>4</sup>, Miguel N. Rivera<sup>1,3</sup>, Cigall Kadoch<sup>1,5</sup>, Orit Rozenblatt-Rosen<sup>1,2</sup>, Mario L. Suvà<sup>1,2,3,†,\*</sup>, Nicolò Riggi<sup>4,†,\*</sup> and Aviv Regev<sup>1,2,14,†,\*</sup>.

<sup>1</sup> Broad Institute of Harvard and MIT, Cambridge, MA, 02142, USA;

<sup>2</sup> Klarman Cell Observatory, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA;

<sup>3</sup> Department of Pathology and Center for Cancer Research, Massachusetts General Hospital and Harvard Medical School, Boston, MA, 02114, USA;

<sup>4</sup> Institute of Pathology, Faculty of Biology and Medicine, Centre Hospitalier Universitaire Vaudois, Lausanne, 1011, Switzerland;

<sup>5</sup> Department of Pediatric Oncology, Dana-Farber Cancer Institute and Harvard Medical School, 450 Brookline Avenue, Boston, MA, 02215, USA;

<sup>6</sup> Department of Medical Oncology, Dana-Farber Cancer Institute and Harvard Medical School, 450 Brookline Avenue, Boston, MA, 02215, USA;

<sup>7</sup> Massachusetts General Hospital Cancer Center, 55 Fruit Street, Boston, MA, 02114, USA;

<sup>8</sup> Laboratory for Systems Pharmacology, Harvard Medical School, Boston, MA, 02115, USA;

<sup>9</sup> Flow Cytometry Facility, Department of Biology and Medicine, University of Lausanne, Lausanne, 1011, Switzerland;

<sup>10</sup> NanoString Technologies Inc., 530 Fairview Avenue North, Seattle, WA 98109, USA;

<sup>11</sup> Department of Medicine, Division of Hematology and Oncology, Massachusetts General Hospital, Boston, MA, 02114, USA;

<sup>12</sup> Department of Orthopedics, Faculty of Biology and Medicine, Centre Hospitalier Universitaire Vaudois, Lausanne, 1011, Switzerland;

<sup>13</sup> Department of Surgery, Massachusetts General Hospital, Boston, MA, 02114, USA;

<sup>14</sup> Howard Hughes Medical Institute, Koch Institute for Integrative Cancer Research, Department of Biology, MIT, Cambridge, MA, 02139, USA

\* Corresponding authors: aregev@broadinstitute.org; nicolo.riggi@chuv.ch; Suva.Mario@mgh.harvard.edu.

† These authors contributed equally to this work.

**ABSTRACT**

Synovial sarcoma is an aggressive mesenchymal neoplasm, driven by the SS18-SSX fusion, and characterized by immunogenic antigens expression and exceptionally low T cell infiltration levels. To study the cancer-immune interplay in this disease, we profiled 16,872 cells from 12 human synovial sarcoma tumors using single-cell RNA-sequencing (scRNA-Seq). Synovial sarcoma manifests antitumor immunity, high cellular plasticity and a core oncogenic program, which is predictive of low immune levels and poor clinical outcomes. Using genetic and pharmacological perturbations, we demonstrate that the program is controlled by the SS18-SSX driver and repressed by cytokines secreted by macrophages and T cells in the tumor microenvironment. Network modeling predicted that SS18-SSX promotes the program through HDAC1 and CDK6. Indeed, the combination of HDAC and CDK4/6 inhibitors represses the program, induces immunogenic cell states, and selectively targets synovial sarcoma cells. Our study demonstrates that immune evasion, cellular plasticity, and cell cycle are co-regulated and can be co-targeted in synovial sarcoma and potentially in other malignancies.

**My contribution:** extracted, cultured, expanded and tested human pediatric Mesenchymal Stem Cells (hpMSCs) derived from healthy donors, used for the experiments. Helped for the single cell sorting leading to supplemental figure 6.

**Nature Medicine, 27, 289–300 (2021). DOI: [10.1038/s41591-020-01212-6](https://doi.org/10.1038/s41591-020-01212-6)**

## Publication 4:

### The chromatin landscape of primary synovial sarcoma organoids is linked to specific epigenetic mechanisms and dependencies

**Authors:** Gaylor Boulay<sup>3,4,11</sup>, Luisa Cironi<sup>1,2,11</sup>, Sara P. Garcia<sup>3,12</sup>, Shruthi Rengarajan<sup>3,12</sup>, Yu-Hang Xing<sup>3</sup>, Lukuo Lee<sup>3</sup>, Mary E. Awad<sup>3</sup>, Beverly Naigles<sup>3</sup>, Sowmya Iyer<sup>3</sup>, Liliane C. Broyc<sup>1,2</sup>, Tugba Keskin<sup>1,2</sup>, Alexandra Cauderay<sup>1,3</sup>, Carlo Fusco<sup>1,2</sup>, Igor Letovanec<sup>1</sup>, Ivan Chebib<sup>3</sup>, Petur Gunnalugur Nielsen<sup>3</sup>, Stéphane Tercier<sup>6</sup>, Stéphane Cherix<sup>5</sup>, Nguyen-Ngoc Tu<sup>7</sup>, Gregory Cote<sup>8</sup>, Edwin Choy<sup>8</sup>, Paolo Provero<sup>9,10</sup>, Mario-Luca Suvà<sup>3,4</sup>, Miguel N. Rivera<sup>3,4,13</sup>, Ivan Stamenkovic<sup>1,2,13,\*</sup> and Nicolò Riggi<sup>1,2,13,\*</sup>.

<sup>1</sup> Institute of Pathology, Centre Hospitalier Universitaire Vaudois, Faculty of Biology and Medicine, University of Lausanne, 1011 Lausanne, Switzerland

<sup>2</sup> Swiss Cancer Center Leman (SCCL), Centre Hospitalier Universitaire Vaudois, Faculty of Biology and Medicine, University of Lausanne, 1011 Lausanne, Switzerland

<sup>3</sup> Department of Pathology and Center for Cancer Research, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA

<sup>4</sup> Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA

<sup>5</sup> Department of Orthopedics, Faculty of Biology and Medicine, Centre Hospitalier Universitaire Vaudois, Lausanne, 1011, Switzerland

<sup>6</sup> Department of Woman-Mother-Child, Centre Hospitalier Universitaire Vaudois, Faculty of Biology and Medicine, University of Lausanne, 1011 Lausanne, Switzerland

<sup>7</sup> Department of Oncology, Centre Hospitalier Universitaire Vaudois, Faculty of Biology and Medicine, University of Lausanne, 1011 Lausanne, Switzerland

<sup>8</sup> Department of Medicine, Division of Hematology and Oncology, Massachusetts General Hospital, Boston, MA, 02114, USA

<sup>9</sup> Center for Translational Genomics and Bioinformatics San Raffaele Scientific Institute, Milan, Italy

<sup>10</sup> Department of Molecular Biotechnology and Health Sciences University of Turin

<sup>11</sup> These authors contributed equally, <sup>12</sup> These authors contributed equally, <sup>13</sup> Senior authors

\*Corresponding authors

#### ABSTRACT

Synovial sarcoma (SyS) is an aggressive mesenchymal malignancy invariably associated with the chromosomal translocation t(X:18; p11:q11), which results in the in-frame fusion of the BAF complex gene SS18 to one of three SSX genes. Fusion of SS18 to SSX generates an aberrant transcriptional regulator, which, in permissive cells, drives tumor development by initiating major chromatin remodeling events that disrupt the balance between BAF-mediated gene activation and Polycomb dependent repression. Here, we developed SyS organoids and performed genome-wide epigenomic profiling of these models and mesenchymal precursors to define SyS-specific chromatin remodeling

mechanisms and dependencies. We show that SS18-SSX induces broad BAF domains at its binding sites, which oppose Polycomb Repressor Complex (PRC) 2 activity while facilitating recruitment of a non-canonical (nc)PRC1 variant. Along with the uncoupling of Polycomb complexes, we observed H3K27me3 eviction, H2AK119ub deposition and the establishment of de novo active regulatory elements that drive SyS identity. These alterations are completely reversible upon SS18-SSX depletion and are associated with vulnerability to USP7 loss, a core member of ncPRC1.1. Using the power of primary tumor organoids, our work helps define the mechanisms of epigenetic dysregulation on which SyS cells are dependent.

**My contribution:** extracted, cultured, expanded and tested human pediatric Mesenchymal Stem Cells (hpMSCs) derived from healthy donors, used for the experiments and performed part of the investigation.

**Life Science Alliance Dec 2020, 4 (2) e202000808; DOI: 10.26508/lsa.202000808**

## 7. Materials and methods

### 7.1. Cell culture

All media were supplemented with 10% Fetal Bovine Serum (FBS, Corning, 35-079-CV) and 1% Penicillin/Streptavidin (P/S, Gibco, 15140-122). All cells were cultured at 37°C with 5% CO<sub>2</sub>. Cells were maintained and split every 3-4 days (at approximately 80% confluence) according to manufacturer recommendations. When needed, FBS was replaced with Tet-Free FBS from Pan Biotech (P30-3602).

#### 7.1.1. JN-DSRCT cell-line

JN-DSRCT1 cell line was a gift from Daniel Haber. This cell-line has been derived from a seven-year-old boy's pleural effusion, with typical intra-abdominal DSRCT and pulmonary metastasis<sup>91</sup>. JN-DSRCT1 cells retained the morphologic characteristics of DSRCT with positive expression of EMA, NSE, vimentin, desmin, and keratins<sup>91</sup>. JN-DSRCT1 cells have been maintained in vitro for more than 190 passages, and the population doubling time of logarithmic growth-phase is 38 hours<sup>91</sup>. In this cell-line, the chimeric EWS-WT1 fuses the exon 10 of EWS to the exon 8 of WT1, leading to a protein of higher molecular weight compared to the canonical isoform<sup>91</sup> (see Figure 59). JN-DSRCT1 cells were grown in Dulbeccos Modified Eagle Medium (DMEM)/F12.



**Figure 59: Schematic representation of the EWSR1-WT1 fusion gene in JN-DSRCT1.** The first ten exons of the EWSR1 gene are depicted in blue, and fused to the last three exons of the WT1 gene (in green). In other DSRCT tumors, only the first seven exons of the EWSR1 gene fused to the last three exons of the WT1 gene. The two isoforms differ by the insertion of nine nucleotides at the end of exon 9 (dark green). Both isoforms are expressed in JN-DSRCT1 cells.

#### 7.1.2. MeT-5A cell-line

Primary immortalized mesothelial MeT-5A were obtained from ATCC (ATCC® CRL-9444™). These cells were isolated from a healthy individual, immortalized by the transfection of a plasmid



containing the Rous sarcoma virus long terminal repeat and the SV40 early region. Met-5A cells were grown in Medium 199.

### 7.1.3. LentiX HEK 293T cell-line

Human embryonic kidney (LentiX HEK 293T) were obtained from Clontech (cat# 632180). All media were purchased from Life Technologies (Invitrogen). LentiX HEK 293T cells were grown in DMEM.

## 7.2. Human pediatric MSC isolation

Human pediatric MSCs (hpMSCs) were isolated from healthy donors undergoing bone corrective surgery (age: between 1 and 18 years old) at the Children's Hospital, CHUV, Lausanne, with the approval and according to the guidelines of the cantonal commission of ethics on human research of the Canton de Vaud (project authorization no. 131/12). Samples were de-identified prior to our analysis. In order to maximize MSCs retrieval, bone fragments were mechanically disaggregated and plated on a 10 cm<sup>2</sup> culture dish in presence of the appropriate MSC medium, composed of Iscove's Modified Dulbecco's Medium (IMDM) with GlutaMAX and 25 mM HEPES (Invitrogen), supplemented with 10% FBS, 1% MEM Non Essential Amino Acids without L-Glut (NEAA, Invitrogen), 1% P/S, 1/1000 Recombinant Human Platelet Derived Growth Factor (PDGF, ProSpec-Tany TechnoGene Ltd, Cat# cyt-501-b). All cells in culture exhibit a fibroblast-like shape. After 2 days of culture, residual bone fragments were removed, and the attached cells were washed with 1X PBS (pH 7.1-7.3, Grosse Apotheke Dr Bichsel) and allowed to grow at 37°C, 5% CO<sub>2</sub>, until they reached 80% confluence.

## 7.3. Cryopreservation of cells

Adherent cell cultures were detached using trypsin/EDTA (Lonza Verviers), and centrifuged at 1300rpm for 5 min. The supernatant was aspirated and cells were re-suspended in cryogenic medium [90% FBS with 10% dimethyl sulfoxide (DMSO)] and transferred into cryogenic vials. Cells were stored at -80°C overnight, and then transferred into liquid nitrogen for long-term storage. Cryopreserved cells were recovered from long term storage by thawing at room temp for 5 min. Cryogenic medium was removed prior to culture, by transferring the contents of each vial in 15 mL Falcon® tubes containing 9 mL of corresponding media and centrifuged at 1300rpm for 5 min.

## 7.4. hpMSCs immunophenotyping

Immunophenotyping of hpMSCs was performed by flow cytometry, using the MSC phenotyping kit (130-095-198, Miltenyi Biotec). This kit includes the following panel of antibodies: APC-CD73, FITC-CD90, PE-CD105 and PerCP CD14/CD20/CD34/CD45. To determine MSCs phenotype 0.5-1.0  $\times 10^6$  cells were incubated in separated tubes for 10 min in 1X PBS at 4°C with Fc Receptor Blocking reagent (Miltenyi Biotec), then washed in 1X PBS and stained for 25 min, protected from light at 4°C, with the phenotyping antibodies and a Live and Dead marker (LIVE/DEAD®, Life Technology). Then, cells were washed and fixed in 1% PFA (paraformaldehyde). Data was acquired the same day or within 5 days after fixation with the SH800S Cell Sorter (Sony Biotechnology) or MOFLO Astrios EQ and at least 100'000 events were collected for each marker. The strategy for the gating begins with cell size selection, exclusion of cell doublets and dead cells elimination. Compensation beads (VersaComp Antibody Capture Bead kit, Beckman Coulter) were used to set up the compensation for the flow cytometry experiments. Analysis was performed using the FloJo software, version 10.

## 7.5. hpMSC differentiation

For the differentiation assay, media was changed every 2 days and cells were kept at 37°C, 5% CO<sub>2</sub>.

### 7.5.1. MSC differentiation towards an adipogenic phenotype

To assess adipogenic differentiation potential, hpMSCs of 4<sup>th</sup> to 6<sup>th</sup> passage were seeded at 40'000 cells/well density in a 12-well plate, on a glass slide. For 3 weeks, hMSC were exposed to a adipogenic medium composed of 10 ml IMDM, 10% FBS, 1% P/S, 1% NEAA, 10  $\mu$ l ITSS (10  $\mu$ g/ml Insulin Transferin Sodium Selenite, Roche), 10  $\mu$ l Dexamethasone ( $10^{-3}$  M, Sigma), 10  $\mu$ l IBMX (100  $\mu$ M 3-Isobutyl-1-Methylxanthine, Sigma) and 20  $\mu$ l Indo (100  $\mu$ M Indomethacin, Fluka). For the control cells (non-induced hpMSC), normal MSC media was used. Following differentiation, the corresponding media was removed and all cells were fixed with 4% PFA for 1h at room temperature (RT). The day of the staining, cells were treated with 70% EtOH (ethanol) for 30 min and stained with Oil Red O (ORO, Sigma) fresh working solution for 1h at RT. Cells were then washed 3 times with 70% EtOH, followed by two consecutive washes in dd water to remove any unbound ORO. Then the nucleus was stained with Mayer's Hematoxyline (Histolab) for 1min. Cells were then washed 3 times in dd water and fixed with Aquamount for long storage.

### 7.5.2. MSC differentiation towards an osteogenic phenotype

To assess osteogenic differentiation potential, hpMSCs of 4<sup>th</sup> to 6<sup>th</sup> passage were seeded at 40'000 cells/well density in a 12-well plate, on a glass slide. For 3 weeks, hMSC were exposed to an osteogenic medium composed of 10 ml DMEM, 10% FBS, 1% P/S, 1% NEAA, 100 µl APP (50 µg/ml Ascorbic Acid P, Sigma), 100 µl Dexamethasone (10<sup>-5</sup> M, Sigma) and 100 µl βGP (5 mM beta Glycerophosphate, Sigma). For the control cells (non-induced hpMSC), normal MSC media was used. Following differentiation, the corresponding media was removed, the cells were fixed with 4% PFA and calcium in the cells was stained with NovaUltra Von Kossa kit (IHC World). Cells were incubated with silver nitrate solution under ultraviolet (UV) light for 60 min at RT. Cells were then washed 3 times with dd water, un-reacted silver was removed with Sodium thiosulfate for 3 minutes, followed by two consecutive washes in dd water. Then the nucleus was counterstained with nuclear fast red for 10 minutes. Cells were then washed 3 times in dd water and dehydrated with EtOH to be fixed with Eukitt for long storage.

### 7.5.3. MSC differentiation towards a chondrogenic phenotype

To assess chondrogenic differentiation potential, hpMSCs of 4<sup>th</sup> to 6<sup>th</sup> passage were seeded at 0.5M cells/pellet, in 1ml media in a partially open 15ml polypropylene canonical Flacon tube. For 3 weeks, hMSC were exposed to a chondrogenic medium, composed of 10 ml DMEM + 1% P/S, 10 µl ITSS, 100 µl AAP, 100 µl linoleic acid (1 mg/ml, Sigma), 100 µl Dexamethasone (10<sup>-5</sup> M) and 10 µl TGF-β1 (Transforming growth factor beta 1, Miltenyi Biotec). For the control cell pellet the media was made of DMEM, P/S, ITSS and linoleic acid. Following differentiation, the medium was aspirated and the cell pellet fixed with 4% PFA and then kept in 70% EtOH. Alcian blue staining was performed after cut and fixation. Detection of Collagen II (II6B3, Developmental Studies Hybridoma Bank), with pre-treatment of Chondroitinase ABC, was performed at the EPFL by Dr. Jessica Sordet-Dessimoz from the histology core facility.

## 7.6. Lentiviral Generation

Lentivirus was produced in HEK 293T LentiX cells by LT1 (Mirus Bio LLC, cat#MIR2305) transfection with gene delivery vector (listed in Tables 11 and 12) and packaging vectors pMD2G and pCMV plasmids, in a 3:2:1 ratio using FuGene6 transfection reagent (Promega, cat# E2692).

The growth medium was changed after 7 hours from the initial transfection to remove malformed virus. Viral supernatants were collected 72h after transfection and concentrated using LentiX concentrator (Clontech-Takara), following the manufacturer's instructions. Virus containing pellets were resuspended in DMEM and added dropwise on cells, in presence of media supplemented with 6 µg/ml polybrene (Hexadimethrine bromide, Sigma, H9268). Selection of lentivirally-infected cells was achieved with puromycin (InvivoGen, ant-pr-1) used at 1-1-0.75 µg/ml (JN-DSRCT1, Met-5A and hpMSCs respectively) or with neomycin (Geneticin – G418, Gibco, 10131-027) used at 0.25 µg/ml (JN-DSRCT1 and Met-5A). Induction of genes in lentivirally-infected cells was achieved with doxycycline hyclate (Sigma, D9891) used at 250 µg/ml. Overexpression or knockdown efficiency was determined by Western blot analysis and qPCR.

A	pLIV empty	B	pIND20 empty
	pLIV EWS-WTS +KTS		pIN20 EWS-WTS +KTS V5 tag cTerm
	pLIV EWS-WTS -KTS		pIN20 EWS-WTS -KTS V5 tag cTerm
	pLIV EWS-WTS +KTS 3xHA tag cTerm		pIN20 V5 EWS-WTS +KTS NTerm tag
	pLIV EWS-WTS -KTS 3xHA tag cTerm		pIN20 V5 EWS-WTS -KTS NTerm tag
	pLIV EWS-WTS +KTS V5 tag cTerm		pIN20 EWS-WTS +KTS 3xHA tag cTerm
	pLIV EWS-WTS -KTS V5 tag cTerm		pIN20 EWS-WTS -KTS 3xHA tag cTerm

**Table 11: List of plasmids used in this study.** A: expression of pLIV plasmids vectors were selected with puromycine. B: expression of pIND20 plasmids vectors were selected with neomycine.

shRNA	Target sequence
shGFP	GCAAGCTGACCCTGAAGTTCAT
shEWS-WT1 #1	GTGTCTGCTAATGTAAACTTT
shEWS-WT1 #2	GTGTCTGCTAATGTAAACTTT

**Table 12: List of shRNA targeting sequences in pKLO.1 plasmids used in this study.**

## 7.7. Real-Time Quantitative PCR (qPCR)

For gene expression assays, total RNA was isolated from cells using RNeasy minikit (Qiagen, cat#74104), for all samples, DNase treatment was performed following the manufacturer's instructions. cDNA was obtained using a high-capacity RNA to-cDNA kit (ThermoFisher, cat#4368814), according to the manufacturer's instructions. 1µg of template total RNA and random hexamers were

used for each reaction. Real-time qPCR amplification was performed using PowerUp SYBR® Green PCR Master Mix (Applied Biosystems Europe BV, cat# A25742) and specific PCR primers in a QuantStudio 5 Real-Time PCR System (ThermoFisher Scientific) for 40 cycles after an initial denaturation at 95°C. Relative quantification of each target, normalized to an endogenous control (housekeeping gene, GAPDH), was performed using the comparative Ct method. Error bars indicate SD of three technical replicates and represent at least two independent biological experiments. Statistical analyses were performed by Student's t-test. Oligonucleotides used were manufactured commercially by Microsynth AG and details regarding sequence and annealing temperature are provided in **Table 13**.

Gene name	Primer sequence (5' - 3')	Annealing temp (°C)
GAPDH	Fw - GGT CTC CTC TGA CTT CAA CA	60
	Rv - GTG AGG GTC TCT CTC TTC CT	61
EWS-WT1	Fw - AGC CAA CAG AGC AGC AGC TAC	67
	Rv - TGA GTC CTG GTG TGG GTC TTC	65
TSPAN7	Fw - CTC ATC GGA ACT GGC ACC ACT A	66
	Rv - CCT GAA ATG CCA GCT ACG AGC T	66
BAIAP3	Fw - CTC AGC CAC ATC CAG GAG TTG T	66
	Rv - CTT CCG AAG CAG CTC CGT ATA	63
ASCL1	Fw - TCT CAT CCT ACT CGT CGG ACG A	66
	Rv - CTG CTT CCA AAG TCC ATT CGC AC	66
ENT4	Fw - CCG ATG ACC GTT ATC ACG CCA T	66
	Rv - CCA AGA TGT AGG TGA GGC TCA TG	64

**Table 13: Oligonucleotides sequence used in this study for the qPCR and corresponding annealing temperature.**

## 7.8. Western Blot Analysis

Cell lysis, SDS-PAGE and blotting were performed by standard procedures, and protein bands were detected using the western blotting detection kit WesternBright Sirius (Witec) according to the manufacturer's instructions. Primary antibodies used for Western blotting are listed in Table 14. Secondary antibodies were HRP-conjugated goat anti-mouse (GE Healthcare UK, NA93IV, 1:10000 dilution), goat anti-rabbit (Dako, P0448, 5:20000 dilution) and goat anti-rat (Cell Signalling, 7077S, 3:10000 dilution). Protein signals were revealed by SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific, Cat# 34080) and captured with the Fusion FX device (Vilber Lourmat) and analyzed with the FusionCapt Advance FX7 software.

Antibody name	Dilution	Provider	Catalogue number	Developed in	Used for
anti-V5 (SV5-Pk1)	1/1000	Abcam	ab27671	Mouse	WB
anti HA (clone 3F10)	2/1000	Roche	11867423001	Rat	WB
$\beta$ -actin	5/10000	Abcam		Mouse	WB
$\alpha$ -tubulin	5/10000	Merck Millipore	DM1A, CP06	Mouse	WB
anti-WT1	1/1000	Proteintech	12609-1-AP	Rabbit	WB
anti-WT1 (C19)	1/1000	Santa Cruz	sc-192	Rabbit	WB
anti-GAPDH-HRP	2/35000	Sigma	G9295		WB
anti-Ki67	1/100	Abcam	Ab16667	Rabbit	IHC
anti-Cleaved Caspase-3 (Asp175)	1/200	Cell Signalling	9661	Rabbit	IHC
Collagen II (II6B3)	1/100	DBSH	07-473	Mouse	IHC
anti-V5	1/100	Abcam	ab9137	Rabbit	IF
anti-HA (12CA5)	1/100	Roche	11583816001	Mouse	IF
Alexa Fluor 647	1/1000				IF
Alexa Fluor 568	1/1000				IF
H3K4me3		EMD millipore	07-473	Rabbit	ChIP
H3K27ac		Active Motif	39133	Rabbit	ChIP
H3K4me1		Abcam	ab8895	Rabbit	ChIP
H3K27me3		Cell signalling	9733	Rabbit	ChIP

**Table 14: List of antibodies used in this study, with its corresponding dilution, provider and other useful information. IHC = Immunohistochemistry, IF = Immunofluorescence, WB = Western Blot and ChIP.**

## 7.9. Proliferation Assays

Prior to drug challenge, we determined the optimal density of the JN-DSRCT1 cell line in 96-well white-walled, white flat-bottom plates (ThermoFisher Scientific) by seeding the cells at different concentrations in 100  $\mu$ L full JN-DSRCT1 medium. Cell viability was measured twice: 16 hours after plating and 16 hours + 2 doubling times (JN-DSRCT1 doubling time is 38 hours<sup>91</sup>) by CellTiter-Glo (Promega, Cat# G7571) according to the manufacturer's protocol. We assessed the optimal seeding density at 2000 cells per well to keep cells in a logarithmic phase of growth during the 92 hours experiment.



Then, we challenged the JN-DSRCT1 cell line with abemaciclib (Selleckchem, LY2835219) or palbociclib (Selleckchem, PD0332991) in concentrations from 0.001  $\mu$ M up to 20  $\mu$ M, according to Gong's protocol<sup>129</sup>. All cells were seeded at 2000 cells per 50  $\mu$ L/well full JN-DSRCT1 medium. Cells were allowed to adhere for 16h and treated with the 50  $\mu$ L drug, for a final volume of 100  $\mu$ L per well. 76 hours after addition of the drug (corresponding to two doubling times), cell viability was measured by CellTiter-Glo.

CellTiter-Glo reagents were prepared according to the manufacturer's protocol. For the CellTiter-Glo assay, plates were incubated at RT for 10 min before recording with a luminescent reader (Synergy Mx, BioTek Instruments, Cat# 7191000). IC<sub>50</sub> values were calculated using GraphPad Prism software.

### 7.9.1. Absolute IC<sub>50</sub> Calculation

To determine the absolute IC<sub>50</sub> (half maximal inhibitory concentration) of abemaciclib and palbociclib in the JN-DSRCT1 cell line, we adapted a GraphPad Prism Software file ([https://www.graphpad.com/guides/prism/7/curve-fitting/index.htm?reg\\_fitting\\_the\\_absolute\\_ic50.htm](https://www.graphpad.com/guides/prism/7/curve-fitting/index.htm?reg_fitting_the_absolute_ic50.htm)) to our experiment. We normalized our data by defining 100% cell viability as the signal from cells in medium only (no drug) and 0% cell viability as the signal from medium only (no cells). The absolute IC<sub>50</sub> of the drug is the concentration of the drug that reduces the viability of the cells by 50%, compared to the cells in full medium with no drug. The details of the advantages of the absolute IC<sub>50</sub> over the relative IC<sub>50</sub> are explained by Gong *and al*<sup>129</sup>.

### 7.10. ChIP-seq

ChIP assays were carried out on approximately 2-5 million cells per sample and per epitope, following the procedures described previously<sup>157</sup>. In brief, chromatin from formaldehyde-fixed cells was fragmented to a size range of 200–700 bases with a Branson 250 sonifier. Solubilized chromatin was immunoprecipitated with the indicated antibodies (listed in Table 14) overnight at 4°C. Antibody-chromatin complexes were pulled down with protein G-Dynabeads (Invitrogen, 10004D), washed, and then eluted. After crosslink reversal, RNase A, and proteinase K (Biolabs, P8107S) treatment, immunoprecipitated DNA was extracted with AMP Pure beads (Agencourt, A63881). ChIP DNA was quantified with Qubit 3.0

Fluorometer (Life technologies). 1-5 ng ChIP DNA samples were used to prepare sequencing libraries, and ChIP DNA and input controls were sequenced with the Nextseq 500 Illumina genome analyzer.

### 7.10.1. ChIP-seq Bioinformatic Processing

Reads were aligned to hg19 (human samples) using bwa<sup>158</sup>. Aligned reads were then filtered to exclude PCR duplicates and were extended to 200 bp to approximate fragment sizes. Density maps were generated by counting the number of fragments overlapping each position using igvtools, and normalized to 10 million reads. Average ChIP-seq signals across intervals were calculated using bwtool<sup>159</sup>.

### 7.10.2. Peakcall

We used macs2<sup>160</sup> to call peaks using matching input controls with a q-value threshold of  $10^{-4}$ . Peaks were filtered to exclude blacklisted regions as defined by the ENCODE consortium<sup>161</sup> and peaks within 200 bp of each other were merged.

### 7.10.3. Peak intersections

Peak intersections were identified using bedtools<sup>162</sup>. Consensus peaks were defined using the Bioconductor genomic ranges package<sup>163</sup>.

### 7.10.4. Peak annotation

Peaks within 1 kb of RefSeq transcription start sites or with strong H3K4me3 ChIP-seq signals (average normalized signal above 8) were considered as promoters, peaks overlapping a RefSeq gene body were considered as distal intragenic and the remaining peaks were considered distal extragenic.

### 7.10.5. Heatmaps and composite plots

Signals shown in heatmaps (100 bp windows) and composite plots (10 bp window) were calculated using bwtool<sup>159</sup>. Heatmap signals are in log2 scale, centered on the indicated peaks and are capped at the 99th percentile.

### 7.10.6. DNA Motif enrichment analysis

For motif discovery, we used the Homer suite of tools<sup>164</sup>. findMotifsGenome.pl was used to identify de novo motifs within indicated peaks.

## 7.11. RNA-Seq

Total RNA was isolated from cells using NucleoSpin RNA Plus (Clontech). 0.5-1 µg of total RNA was treated with Ribogold zero to remove ribosomal RNA. Illumina sequencing libraries were constructed using random primers according to the manufacturer's instructions using the Truseq Stranded RNA LT Kit.

### 7.11.1. RNA-seq bioinformatic Processing

Reads were aligned to hg19 using STAR<sup>165</sup>. Mapped reads were filtered to exclude PCR duplicates and reads mapping to known ribosomal RNA coordinates, obtained from rmsk table in the UCSC database (<http://genome.ucsc.edu>). Gene expression was calculated using featureCounts<sup>166</sup>. Only primary alignments with mapping quality of 10 or more were counted. Counts were then normalized to 1 million reads. Signal tracks were generated using bedtools<sup>162</sup>. Differential expression was calculated using DESeq2<sup>167</sup> and transcript abundance using Cufflinks<sup>168</sup>.

### 7.11.2. Identification of differentially expressed genes

The genes differentially expressed upon EWS-WT1 knock-down were selected based on a DESeq2 3-fold change and an adjusted p-value lower than 0.001. Down-regulated genes had initial expression levels > 1 FPKM and up-regulated genes had final expression levels > 1 FPKM.

## 7.12. Biotinylated isoxazole-Mediated Precipitation

These assays were performed as previously described<sup>113</sup> with slight modifications. Biotinylated isoxazole (b-isox, Sigma-Aldrich) was reconstituted in DMSO. Briefly, 5-10 million cells were resuspended in 1 mL lysis buffer (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 5 mM MgCl<sub>2</sub>, 0.5% NP-40 and 10% glycerol supplemented with 1X Protease/Phosphatase inhibitors (Pierce), 0.1 mM PMSF and 20 mM beta-mercaptoethanol) and incubated for 30 minutes with rotation at +4°C. Protein supernatant was then collected after centrifugation for 15 min at 14000 rpm at 4°C. A 5 % whole cell

extract (WCE) control was collected and the remaining proteins were divided into four aliquots before addition of biotinylated isoxazole at 0, 10, 30 100  $\mu$ M final concentrations. The Reaction solutions were incubated at +4°C for 1h with rotation and centrifuged for 15 min at 14000 rpm +4°C. Supernatant was saved for further analysis and pellets were washed twice in supplemented Lysis buffer before resuspension in Laemmli buffer. WCE, pellets and supernatants were analyzed by 4%-12% Tris-Glycine gradient gels (Life Technologies) and Western blotting was performed using standard protocols.

### **7.13. In vivo studies**

All mice experiments were approved by the Cantonal Veterinary Office, authorization number VD3021. For the experiments, we used 5-7 weeks old NOD/SCID c Gamma (NSG) mice (purchased from Jackson Laboratory, USA – stock number 005557). All mice were monitored three times a week for tumor development, and sacrificed using CO<sub>2</sub> inhalation and cervical dislocation, when the tumor reached 1000-1500 mm<sup>3</sup> volume or at the end of the experiment. Every time, an accurate necropsy of the mice was performed to assess metastatic spread.

#### **7.13.1. JN-DSRCT1 and MeT-5A xenografts**

We used  $2 \times 10^6$  cells for the JN-DSRCT1 cell-line and  $4 \times 10^6$  cells for the Met-5A cells. The indicated amounts of cells were resuspended in 100  $\mu$ L of 1X PBS, and injected subcutaneously into NSG mice. Mice were then monitored, and sacrificed as described above. At the end of the experiment, the collected tumors were minced with scissors and a tumor fragments was dissociated, when needed, using the Tumor Dissociation Kit human (Miltenyi Biotec) in combination with the Gentle-MACS device, following the manufacturer's instructions. The remaining fragments were storage at -80°C or used for histopathological and molecular analyses. When needed, mice were fed with doxycycline food red colored provided by SAFE (U8200P Version 0115, A03 at 0.625 g/kg Doxycycline, irradiated)

#### **7.13.2. PDX generation from St. Jude Children's Research Hospital**

Primary DSRCT derived xenografts were obtained from the St. Jude Children's Research Hospital of Philadelphia that launched in 2013 "The Childhood Solid Tumor Network" allowing researchers around the world to access and use material and data from pediatric solid tumors to study them and

theirs related biology. The samples received were defrosted and washed in 1X PBS before immediate implantation in the flank tumor of a NSG mice. Mice were then monitored, and sacrificed as described above. Upon collection, each tissue was maintained less than one hour in DMEM with GlutaMax at RT before implantation. The collected tumors were minced with scissors and a tumor fragments was paraffined and the remaining parts were cut into small fragments of 80-100 mg and used for subcutaneous implantation into new NSG mice. Again, mice were then monitored, and sacrificed as described above. Then, the tumors were regrown in a similar way to perform the indicated experiments.

### 7.13.3. In vivo palbociclib experiment

For PDX1 tumors were harvested and minced with scissors into small fragments of 80-100 mg. For the JN-DSRCT1 cell-line,  $2 \times 10^6$  cells in 100uL 1X PBS was used. Then these cells or tumor fragments were subcutaneous implanted into NSG mice. The experiments were performed using cohorts of eight rodents. Mice were randomly allocated to the control-treated (n=8) or the drug-treated (n=8) group. Once the sarcoma reach a mean of 50 mm<sup>3</sup> in volume, the rodents were treated with 100mg/kg mice dose of palbociclib for 4 weeks (5 days per week, with 2 days rest), by gavage as a solution of sodium lactate buffer (Aldrich, 71718, 50mM Sodium lactate, pH 4.0) based on mean group body weight. All mice were monitored daily for signs of distress and were weighed three times a week. Tumor size was measured three times a week, with a caliper, according to the following equation:  $(\text{length} \times \text{width}^2)/2 = \text{tumor volume [mm}^3\text{]}$ . The palbociclib was purchased from Selleckchem (PD0332991), freshly prepared and orally administered daily. We used a daily dose of palbociclib 100mg/kg mouse in 200  $\mu$ l sodium lactate buffer. Control treated mice received the buffer used for oral administration. Mice were sacrificed, as described above, the day after the last dose of palbociclib, H&E, Ki97 and Cleaved Caspase 3 staining were performed to assess tissue profile.

### 7.14. Staining

H&E, Ki97 and CC3 stainings were performed by Jean-Christophe Stehle, head of the mouse pathology facility of the UNIL-CHUV, using the following protocols. All tumors paraffin bloc were cut at 5 $\mu$ m, mounted and dried on glass slides. All tissues were deparaffined in xylol, followed by rehydration in EtOH (100%, 95%, 80%) and water.

#### 7.14.1. Hematoxylin and Erythrosine (H&E) staining

The tissues were stained with Harris hematoxylin (J.T. Baker, cat# 3873) for 4 min and destained with 1% acid ethanol (1% HCl 37% in 70% EtOH) for 3 sec. Stained with Erythrosine (Merk, cat# 15936, 0,2% in dd water and 1/1000 formaldehyde) for 30 sec, then tissues were dehydrated and mounted on coverslips.

#### 7.14.2. Ki67 staining

Endogenous peroxidase activity was inhibited using 1%  $\text{H}_2\text{O}_2$  in methanol for 15 minutes at RT. Then antigen retrieval is achieved using TRIS/EDTA pH 9 for 2min. Nonspecific binding was saturated using 1X PBS solution with 1% goat serum for 10 min at RT. A rabbit anti-human Ki67 primary antibody (Abcam, ab 16667, 1/100) was used. The primary antibody was applied to tissue sections and slides were incubated one hour at RT. A secondary HRP antibody against rabbit (Dako Envision anti Rabbit HRP) was applied for 30 min at RT. Staining was revealed using DAB+ kit (DAKO K3467, 1/50) for 5 min at RT. The tissues were counterstained with Harris hematoxylin, rehydrated and mounted using Eukitt on coverslips. All washes, between every step were performed using 1X PBS.

#### 7.14.3. Cleaved Caspase 3 (CC3) staining

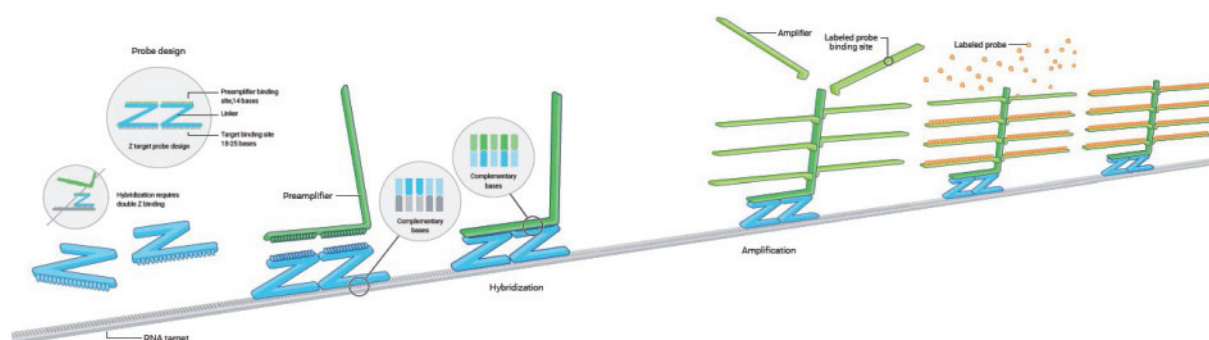
Endogenous peroxidase activity was inhibited using 1%  $\text{H}_2\text{O}_2$  in methanol for 15 minutes at RT. Then antigen retrieval is achieved using TRIS/EDTA pH 9 for 2min. Nonspecific binding was saturated using 1X PBS solution with 1% goat serum for 10 min at RT. A rabbit anti-Cleaved-Caspase3 primary antibody (Cell Signaling #9661) 1/200 was used. The primary antibody was applied to tissue sections and slides were incubated one hour at RT. A secondary HRP antibody against rabbit was applied for 30 min at RT. Staining was revealed using DAB+ kit for 5 min at RT. The tissues were counterstained with Harris hematoxylin, rehydrated and mounted using Eukitt on coverslips. All washes, between every step were performed using 1X PBS.

### 7.15. BaseScope

BaseScope is an *in situ* hybridization technology that allows the detection of a targeted mRNA within cells or tissues. It uses probes designed with a Z shape allowing detection of an mRNA target, amplifying the signal and suppressing the background noise at the same time<sup>169</sup> (see a summary of this



assay in Figure 60). Two transcribed genes can be visualized at the same time using two chromogenic dyes (alkaline phosphatase (AP) and horseradish peroxidase (HRP)-based enzymatic reactions). This technology enables the analysis of RNA expression within cells or tissues with high sensitivity and specificity<sup>124,169</sup>. The BaseScope paired oligo-probes are designed to detect particular regions of the mRNA (like exon junctions of 50bp, mutations of 10bp, and short target sequences of 50-300 bp).



**Figure 60: Summary of the BaseScope assay (from the ACDbio website).** The assay begins with recognizing and hybridizing the targeted RNA (in grey) by the ZZ paired oligo probe (in blue). Then, a preamplifier binds specifically the upper part of the ZZ probe (in green) before amplification of the signal. Finally, the ZZ probe is labeled with HRP-based Green or AP-based Fast Red dye to be visualized with a microscope.

BaseScope assays were performed and assisted following guidelines (BaseScope™ Detection Reagent Kit-RED User Manual) provided by the supplier (Advanced Cell Diagnostics, Newark, CA). Briefly, sections were cut at 5 µm thickness onto Superfrost plus slides (Thermo Scientific, New Hampshire, USA) and allowed to dry overnight at room temperature (RT). Sections were then baked at 60°C for two hours before being deparaffinized in xylene (2 x 10 min) and EtOH (2 x 5 min), then dried by baking at 60°C for 5 min. Pretreatment with hydrogen peroxide was applied on sample for 10 min at RT, followed by target retrieval for 30 min at 100°C in a steaming bowl, with two rinses in distilled water between pretreatments. Then transfer slides to 100% EtOH and dried by baking at 60°C for 5 min. Tissue sections were incubated with protease IV for 30 min at 40°C. BaseScope ZZ probes were then applied for 2 h at 40°C in a HybEZ oven before slides were kept overnight at 4°C in SSC. The next day, samples were hybridized in a HybEZ oven with reagents AMP1 (30 min at 40°C), AMP2 (30 min at 40°C), AMP3 (15 min at 40°C), AMP4 (30 min at 40°C), AMP5 (30 min at 40°C), AMP6 (15 min at 40°C), AMP7 (30 min at RT), and AMP8 (15 min at RT). Slides were rinsed with wash buffer (2 x 2 min) between AMP incubations. Then, slides were incubated with Fast Red for 10 min at RT in the dark. Then samples were hybridized in a HybEZ oven with reagents AMP9 (15

min at 40°C), AMP10 (15 min at 40°C), AMP11 (30 min at RT), and AMP12 (15 min at RT). Slides were rinsed with wash buffer (2 x 2 min) between AMP incubations. Then, slides were incubated with Green for 10 min at RT in the dark. Then, slides were counterstained with hematoxylin for 2 min before drying for 30 min at 60°C. Finally, slides were mounted and kept at 4°C protected from light until analysis.

## 7.16. Imaging

Some of the images were taken with a Nikon Eclipse E800 microscope and a digital camera DXM1200, at 4x, 20x, 40x or 60x magnification. The image resolution was 1'280x1'024. Images were analyzed with the ACT-1 (v.2) software.

Other images were acquired by Carlo Fusco, using a differential Interference contrast microscope in the bright light field using Zeiss ImagerZ1 Upright microscope with an objective EC Plan Neofluar 100x magnification with N.A. (numerical Aperture) of 1.3 in Oil. The color camera used for the acquisition was a Zeiss AxioCam MRc5. Software used to control the camera was Axiovision SE64 rel 4.9.1.

Slides were scanned using the NanoZoomer S60 Digital slide scanner (C13210-01) from Hamamatsu, with a 40x objective for Brightfield and Fluorescence. Software used to control the camera was Color CMOS camera or sCMOS Orca Flash LT Plus, software NDP.view2.

## Acknowledgments

This thesis was supported by the Swiss National Science Foundation (Switzerland), the Emma Muschamp Foundation (Switzerland), and the University of Lausanne (Switzerland). I am grateful for these supports, which allowed me to complete my Ph.D. in excellent conditions.

I would like to thank Prof. Nicolo Riggi. I was fortunate to find the way to his recently opened laboratory at CHUV. I cannot find the words to express my gratitude as he offered me a diamond opportunity, and I am honored to be his student and colleague. I appreciated his enthusiasm, kindness, guidance, and support. Prof. Riggi was always available to discuss and give useful advice, even under pressure. He is an excellent mentor, and I hope to continue collaborating with him in the future wherever my career will take me. In Prof. Riggi's laboratory, I would like to express my gratitude to Sanalkumar Rajendran. Since his arrival, he has always been of excellent advice and support, particularly during my thesis manuscript's redaction. His companionship in the laboratory was always delightful, and I know that I can rely on him. Similarly, I would like to express my gratitude to other members of Prof. Riggi's laboratory, Arnaud Bakaric, with whom I shared a bit of the Ph.D. work-life, Viviane Praz, and Santosh Sundaramoorthy.

I am also grateful to Prof. Ivan Stamenkovic for his guidance, time, and support. In Prof. Stamenkovic's laboratory, my thanks go to Emely Moller, who showed me how to work in this laboratory and always gave good advice. I would like to express my appreciation to Patricia Waszyk, who became a friend through challenging Ph.D. life, ADAS presidency, and as my Disney movie mate. I would like to show my gratitude to Tugba Keskin, who was always available and supportive, even when overloaded by work. Since it would double the wordcount of this thesis to express the whole extent of my thankfulness towards each member, I would like to thank all of Stamenkovic's laboratory's team as a group for their kindness and time. Both Riggi and Stamenkovic laboratories work as one team, and it has always been a pleasure to come to work every day.

I would like to thank the members of my thesis committee for their time. Prof. Maja Beck-Popovic and Prof. Douglas Hanahan offered me valuable feedbacks throughout this study that enriched my work. I would like to acknowledge Prof. François Bochud as president of my jury.

This study was possible through many collaborations. I would like to express my gratitude to Prof. Miguel Rivera and Prof. Mazar Adli to have welcomed me in their laboratories in the USA to learn new techniques for my Ph.D. thesis. It was interesting to exchange scientific ideas with them, and I am glad that I was given the opportunity to learn directly from them. I would like to thank Gaylor Boulay for many reasons, particularly for his time, advice, input, support, and collaboration during my Ph.D. I hope we can soon finish our paper and wish him the best for his new position in France. I would also like to acknowledge Sowmya Iyer for her explanations, time, bioinformatics analysis and help with the figures. I would like to thank all the people in the different facilities at the UNIL and EPFL: the Mouse Pathology Facility, the Animal Facility, and the Cellular Imaging Facility at Epalinges, the Flow Cytometry Facility at Agora, and the Histology Core Facility at EPFL. Their support and advice were highly appreciated during my projects.

There are people not related to my work life that I would like to acknowledge. My friends were tremendously supportive during my studies and understanding of my occasional mood fluctuations. I enjoyed explaining my work to them, even if the topic can somehow be challenging to understand in detail. We always had crazy talks and shared innovative ideas that might lead to something one day. I would like to express gratitude to the ADAS, which made me discover the backstage of doctoral life at UNIL. We shared unforgettable memories through the years. I was given the opportunity to organize conferences and social events, as well as meeting fascinating people within the FBM and during university commissions. This knowledge allowed me to understand how politics and networks are essential in a career, but most importantly, these rich encounters made me realize how lucky I was to be in such a great team lead by a great PI.

I take this opportunity to thank my family. They always believed in me and encouraged me all my life. Querida familia Gonzalez-Ostos (extendida), gracias por todo, están siempre en mi corazón. Mama, Papa, merci de m'avoir poussé à entreprendre des études, à explorer de nouveaux horizons et à voyager par-delà l'inconnu. Mes très chers Paul et Mathilde aussi, merci pour tout. Rires, pleurs et

soutien dans toutes les situations. Unidos siempre. Merci à tous d'avoir été présents pour moi, même si parfois c'était à longue distance. I want to deeply thank my husband for everything, without him it would have been difficult to finish my PhD. Je sais que tu n'aimes pas que j'étaie notre vie, mais ces petits rien de tous les jours m'ont permis de persévérer et d'avancer dans mes études, mon doctorat et dans la vie. Tu as toujours du temps pour moi et surtout tu me comprends, même quand je ne dis rien. J'adore nos discussions toujours enrichissantes et drôles. Je suis très heureuse d'arriver au bout de ce manuscrit en sachant que toi aussi tu l'as lu du début à la fin. TQM.

## List of Abbreviations

<b>3C</b>	Chromosome Conformation Capture
<b>3D</b>	Three-dimensional
<b>5hmC</b>	5-hydroxymethylcytosine
<b>5mC</b>	5-methylcytosine
<b>5mC</b>	5-methylcytosine
<b>ac</b>	Acetylation
<b>AD</b>	Activation Domain
<b>AP</b>	Alkaline Phosphatase
<b>APC</b>	Allophycocyanin
<b>APP</b>	Ascorbic Acid P
<b>ASCL1</b>	Achaete-scute homolog 1
<b>ATF1</b>	Activating transcription factor 1
<b>ATP</b>	Adenosine TriPhosphate
<b>ATR</b>	Ataxia Telangiectasia and Rad3-related
<b>B7-H3</b>	Human B7 homolog 3
<b>BAF</b>	BRG1/BRM-Associated Factor
<b>BAIAP3</b>	Brain-specific Angiogenesis Inhibitor 1-Associated Protein 3
<b>BDNF</b>	Brain-Derived Neurotrophic Factor
<b>b-isox</b>	Biotinylated isoxazole
<b>CC3</b>	Cleaved Caspase 3
<b>CCN2</b>	Cellular Communication Network Factor 2



<b>CCND1</b>	Cyclin D1
<b>CD14, CD20, CD34, and CD45</b>	Cluster of Differentiation 14, 20, 34 and 45
<b>CD73, CD90, and CD105</b>	Cluster of Differentiation 73, 90 and 105
<b>CDK</b>	Cyclin-Dependent Kinases
<b>ChIP</b>	Chromatin ImmunoPrecipitation
<b>ChIP-seq</b>	Chromatin ImmunoPrecipitation sequencing
<b>CHUV</b>	Centre Hospitalier Universitaire Vaudois
<b>CI</b>	Confidence Interval
<b>c-MYC</b>	c-Myc proto-oncogene protein
<b>CpGs regions</b>	Cytosine–phosphate–Guanine regions
<b>CSS</b>	Clear Cell Sarcoma
<b>CT</b>	Computed Tomography
<b>Cys(2)His(2)-type</b>	C2H2-type, Cysteine(2)-Histidine(2) type
<b>DapB</b>	4-hydroxy-tetrahydrodipicolinate reductase
<b>DD</b>	Divergent Domain
<b>dd water</b>	Double distilled water
<b>DMEM</b>	Dulbeccos Modified Eagle Medium
<b>DMSO</b>	Dimethyl Sulfoxide
<b>DNA</b>	Deoxyribonucleic Acid
<b>DNA BD</b>	DNA Binding Domain
<b>DNMT</b>	DNA methyltransferases
<b>DSRCT</b>	Desmoplastic small round-cell tumor

<b>E-cadherin</b>	Epithelial cadherin
<b>EDTA</b>	Ethylenediaminetetraacetic Acid
<b>EGF</b>	Human Epidermal Growth Factor
<b>EGFR</b>	Epidermal Growth Factor Receptor
<b>EGR-1</b>	Early Growth Response protein 1
<b>EMA</b>	Epithelial Membrane Antigen
<b>EMSA</b>	Electrophoresis Mobility Shift Assay
<b>EMT</b>	Epithelial-Mesenchymal Transition
<b>ENT4</b>	Equilibrative Nucleoside Transporter member 4
<b>EPFL</b>	École Polytechnique Fédérale de Lausanne
<b>ERG</b>	ETS-Related Gene
<b>eRNAs</b>	Enhancer RNAs
<b>EtOH</b>	Ethanol
<b>ETS</b>	Erythroblastosis-virus-associated Transforming Sequences
<b>EWS</b>	Ewing's Sarcoma protein
<b>EWS(YS37)-WT1</b>	EWS-WT1 mutant where all 37 tyrosines in the PrLD of EWS were mutated into serine
<b>EWSR1</b>	Ewing's sarcoma gene
<b>FBS</b>	Fetal Bovine Serum
<b>FDA</b>	Food and Drug Administration
<b>Fe(II)</b>	Iron in its +2 oxidation state (Fe <sup>2+</sup> )
<b>FET</b>	for FUS, EWSR1, and TAF15
<b>FFPE</b>	Formalin-Fixed Paraffin-Embedded

<b>FGF</b>	Human Fibroblast Growth Factor
<b>FGFR4</b>	Fibroblast Growth Factor Receptor 4
<b>FITC</b>	Fluorescein isothiocyanate
<b>FLI1</b>	Friend Leukemia virus Integration site 1
<b>FUS</b>	Fused in liposarcoma
<b>GAPDH</b>	Glyceraldehyde 3-phosphate dehydrogenase
<b>Gly-rich</b>	Glycine-rich domain
<b>H&amp;E</b>	Hematoxylin and Eosin
<b>H2A, H2B, H3, and H4</b>	Histone proteins 2A, 2B, 3 and 4
<b>H3K27ac</b>	Acetylation of Lysine 27 of Histone 3
<b>H3K27me3</b>	Trimethylation of Lysine 27 of Histone 3
<b>H3K36me3</b>	Trimethylation of Lysine 36 of Histone 3
<b>H3K4me1</b>	Monomethylation of Lysine 4
<b>H3K4me3</b>	Trimethylation of Lysine 4
<b>H3K9ac</b>	Acetylation of Lysine 9 of Histone 3
<b>H3K9me3</b>	Trimethylation of Lysine 9 of Histone 3
<b>HA tag</b>	Hemagglutinin tag
<b>HDAC</b>	Histone deacetylases
<b>HEK-293T</b>	Human Embryonic Kidney 293T
<b>HiChIP</b>	High-throughput 3C techniques coupled with ChIP
<b>hpMSC</b>	Human pediatric Mesenchymal Stem Cell
<b>HRP</b>	Horseradish Peroxidase

<b>IBMX</b>	3-Isobutyl-1-Methylxanthine
<b>IC<sub>50</sub></b>	Half maximal Inhibitory Concentration
<b>IGF-1R</b>	Insulin-like Growth Factor 1 Receptor
<b>IGF2</b>	Insulin-like Growth Factor 2
<b>IL2RB</b>	Interleukin 2/15 receptor $\beta$
<b>IMDM</b>	Iscoe's Modified Dulbecco's Medium
<b>Indo</b>	Indomethacin
<b>INF</b>	Interferon
<b>ISCT</b>	International Society for Cellular Therapy
<b>ITSS</b>	Insulin Transferin Sodium Selenite
<b>IVS9</b>	Intervening Sequence 9
<b>K</b>	Lysine
<b>KRAB</b>	Kruppel-Associated Box Domain
<b>KTS</b>	Lysine, threonine and serine
<b>LLPS</b>	Liquid-Liquid Phase Separation
<b>lncRNAs</b>	Long non-coding RNAs
<b>LRRC15</b>	Leucine-Rich Repeat-Containing protein 15
<b>MACS</b>	Model-based Analysis of ChIP-Seq
<b>MET</b>	Mesenchymal-Epithelial reverse Transition
<b>met</b>	Methylation
<b>MeT-5A</b>	Mesothelial cell-line 5A
<b>MGH</b>	Massachusetts General Hospital

<b>miRNAs</b>	Micro RNAs
<b>MLF-1</b>	Myeloid Leukemia Factor 1
<b>MRI</b>	Magnetic Resonance Imaging
<b>mRNA</b>	Messenger RNA
<b>MSC</b>	Mesenchymal Stem Cell
<b>mTOR</b>	Mammalian Target Of Rapamycin
<b>N.A.</b>	Numerical Aperture
<b>NEAA</b>	MEM Non Essential Amino Acids without L-Glut
<b>NLS</b>	Nuclear Localization Sequence
<b>N-MYC</b>	N-myc proto-oncogene protein
<b>NSE</b>	Neuron-Specific Enolase
<b>NSG mouse</b>	NOD/SCID c Gamma mouse
<b>ORO staining</b>	Oil Red O staining
<b>P</b>	Phosphorylation
<b>P/S</b>	Penicillin/streptomycin
<b>PARP1</b>	Poly ADP-Ribose Polymerase 1
<b>PAX2, PAX6</b>	Paired box-protein 2, 6
<b>PBS</b>	Phosphate Buffered Saline
<b>PD1</b>	Programmed Death-1
<b>PDGF<math>\alpha</math></b>	Human Platelet-Derived Growth Factor Subunit A
<b>PDX</b>	Patient Derived Xenograft
<b>PE</b>	Phycoerythrin

<b>PerCP</b>	Peridinin Chlorophyll Protein Complex
<b>PET</b>	Positron Emission Tomography
<b>PFA</b>	Paraformaldehyde
<b>POLAR2A</b>	DNA-directed RNA polymerase II subunit RPB1
<b>PPIB</b>	Cyclophilin B
<b>PRC</b>	Polycomb Repressive Complexes
<b>PrLD</b>	Prion-Like Domain
<b>Pro/Gln RD</b>	Prolines and glutamines rich domain
<b>PWM</b>	Position Weight Matrix
<b>qPCR</b>	Quantitative real-time Polymerase Chain Reaction
<b>QPGY rich domain</b>	Glutamine-proline-glycine-tyrosine rich domain
<b>R</b>	Arginine
<b>Rb</b>	Retinoblastoma tumor suppressor protein
<b>RD</b>	Repression Domain
<b>RGG-rich domain</b>	Arginine-glycine-glycine-rich domain
<b>RNA</b>	Ribonucleic acid
<b>RNA-seq</b>	RNA sequencing
<b>RRM</b>	RNA-Recognition Motif
<b>RRM2</b>	Component of the Ribonucleotide Reductase complex
<b>RT</b>	Room Temperature
<b>S</b>	Serine
<b>shRNA</b>	Small hairpin RNA



<b>SLFN11</b>	Schlafen Family Member 11
<b>SNH</b>	SS18 N-terminal Homology
<b>SOX2</b>	SRY-box 2
<b>SRY</b>	Sex-determining Region Y
<b>T</b>	Threonine
<b>TAD</b>	Topologically Associated Domain
<b>TAF15</b>	TATA-binding Associated Factor 15
<b>TALLA1</b>	T-cell acute lymphoblastic leukemia-associated antigen 1, also called TSPAN7
<b>TET</b>	Ten-Eleven Translocation
<b>TGF-<math>\beta</math></b>	Transforming Growth Factor-beta
<b>TNF</b>	Tumor Necrosis Factor
<b>TP53</b>	Tumor suppressor protein 53
<b>TSS</b>	Transcriptional Start Sites
<b>U2OS</b>	Osteosarcoma cell line
<b>ub</b>	Ubiquitination
<b>UNIL</b>	Université de Lausanne
<b>USP9X</b>	Ubiquitin Specific Peptidase 9 X-Linked
<b>UV</b>	Ultraviolet
<b>V5 tag</b>	Parainfluenza Virus 5 tag
<b>VEGF</b>	Vascular Endothelial Growth Factor
<b>WAGR syndrome</b>	Wilms tumor, Aniridia, Genitourinary malformation, and mental Retardation syndrome

<b>WB</b>	Western blot
<b>WCE</b>	Whole Cell Extract
<b>WKE motif</b>	WT1 +KTS element binding motif
<b>Wnt</b>	Wingless and Int-1
<b>WT1</b>	Wilms' tumor suppressor gene 1
<b>WTE motif</b>	WT1 element binding motif
<b>Y</b>	Tyrosine
<b>ZFD</b>	Zinc Finger Domain
<b><math>\alpha</math>-KG</b>	Alpha-Ketoglutarate
<b><math>\alpha</math>-SMA</b>	Alpha Smooth-Muscle Actin
<b><math>\beta</math>GP</b>	Beta Glycerophosphate

# List of Figures

	pg
<b>Figure 1:</b> The hallmarks of cancer	2
<b>Figure 2:</b> Cell-cycle and cyclin-CDK complexes involved	3
<b>Figure 3:</b> Regulation and function of Cyclin D-CDK4/6 complex.	4
<b>Figure 4:</b> Representative histological image of a human DSRCT stained with hematoxylin and eosin (H&E).	10
<b>Figure 5:</b> Schematic representation of EWSR1, WT1, and the resulting EWSR1-WT1 fusion gene	11
<b>Figure 6:</b> Schematic representation of EWS protein and its corresponding domains	12
<b>Figure 7:</b> Schematic representation of WT1 protein and its corresponding domains	14
<b>Figure 8:</b> Schematic representation of EWS-WT1 protein and its corresponding domains	15
<b>Figure 9:</b> Possible histone modifications (modified from Thermo Fisher Scientific)	19
<b>Figure 10:</b> Mechanisms of histone modification by epigenetic chromatin regulators	20
<b>Figure 11:</b> Chromatin profiling highlight a subset of EWS-WT1 binding sites associated with higher levels of activity in DSRCT	25
<b>Figure 12:</b> Analysis of binding sites and motif of EWS-WT1 in the JN-DSRCT1 cells	26
<b>Figure 13:</b> Positive correlation between EWS-WT1 and H3K27ac in JN-DSRCT1	27
<b>Figure 14:</b> Decrease of EWS-WT1 expression in JN-DSRCT1 cells upon knockdown	27
<b>Figure 15:</b> Acetylation changes in JN-DSRCT1 cells upon EWS-WT1 knockdown	28
<b>Figure 16:</b> RNA expression changes in JN-DSRCT1 cells upon EWS-WT1 knockdown	29

<b>Figure 17:</b>	Representative immune profile of MSCs, by Fluorescence-Activated Cell Sorting	30
<b>Figure 18:</b>	Primary cells generated are immunophenotypically similar to MSCs	31
<b>Figure 19:</b>	Primary cells generated exhibit similar differentiation potential than MSCs in vitro.	32
<b>Figure 20:</b>	hpMSCs #2 lentivirally transduced with EWS-WT1 fail to express the aberrant protein, but induce transcription of some DSRCT targets	33
<b>Figure 21:</b>	HA-tagged EWS-WT1 induce transcription of some DSRCT targets, but fail to expression the EWS-WT1 –KTS isoform in hpMSCs #41	34
<b>Figure 22:</b>	HA or V5-tagged EWS-WT1 proteins were detected in MeT-5A cells and corresponding EWS-WT1 targets were transcript	35
<b>Figure 23:</b>	Genomic location of EWS-WT1 isoforms in the MeT-5A cells	38
<b>Figure 24:</b>	Chromatin profiling highlights a subset of EWS-WT1 binding sites are associated with high levels of acetylation in MeT-5A cells expressing EWS-WT1 isoforms	38
<b>Figure 25:</b>	Chromatin profile of EWS-WT1 isoforms in our cellular DSRCT model separated by motifs identified in JN-DSRCT1 and positive correlation between the primary DSRCT tumors and our model	39
<b>Figure 26:</b>	EWS-WT1 can precipitate and recruit endogenous EWSR1	41
<b>Figure 27:</b>	The fusion protein EWS-WT1 can homo and hetero-dimerize	41
<b>Figure 28:</b>	MeT-5A lentivirally transduced with V5-tagged WT1 isoforms induce WT1 transcription and WT1 protein expression	43
<b>Figure 29:</b>	EWS-WT1 display a different chromatin profile when compared to WT1 around the LRRC15 gene.	44

<b>Figure 30:</b> WT1 isoforms cannot change the chromatin state of MeT-5A cells at EWS-WT1 binding sites	45
<b>Figure 31:</b> WT1 isoforms precipitate upon high doses of b-isox in HEK 293T cell lysates but fail to recruit endogenous EWSR1	45
<b>Figure 32:</b> The WT1 protein can homo and heterodimerize	46
<b>Figure 33:</b> MeT-5A lentivirally transduced with V5-tagged EWS(YS37)-WT1 isoforms induce EWS(YS37)-WT1 transcription and expression of the mutated protein	47
<b>Figure 34:</b> The mutated EWS(YS37)-WT1 do not precipitate upon treatment with b-isox and fails to recruit endogenous EWSR1	48
<b>Figure 35:</b> The mutated EWS(YS37)-WT1 protein cannot dimerize	48
<b>Figure 36:</b> 1ZZ KTS BaseScope probes recognize EWSR1-WT1 isoforms	51
<b>Figure 37:</b> 1ZZ KTS BaseScope probes recognize EWSR1-WT1 isoforms expressed concomitantly in JN-DSRCT1 cells and derived tumors	52
<b>Figure 38:</b> Primary human DSRCT tumors displays higher EWSR1-WT1 +KTS transcripts	53
<b>Figure 39:</b> MeT-5A lentivirally transduced with V5 or HA-tagged EWS-WT1 express the aberrant protein and induce transcription of DSRCT targets	54
<b>Figure 40:</b> When expressed together, EWS-WT1 +KTS binds to EWS-WT1 –KTS and recruit this latter at the LRRC15 locus	54
<b>Figure 41:</b> EWS-WT1 isoforms can be recruited at the same binding sites and change the chromatin state of MeT-5A cells	56
<b>Figure 42:</b> Concomitant expression of EWS-WT1 isoforms colocalize them within the nucleus of the cell	57

<b>Figure 43:</b> MeT-5A cells expressing both EWS-WT1 isoform are capable of develop tumors	59
<b>Figure 44:</b> Tumors expressing concomitantly both EWS-WT1 isoforms are similar to DSRCTs	60
<b>Figure 45:</b> EWS-WT1 governs regulatory networks in DSRCT	62
<b>Figure 46:</b> List of 12 potential therapeutic targets specific for DSRCTs, linked with at least one known pharmacological inhibitor	63
<b>Figure 47:</b> EWS-WT1 highly interacts with CCND1	64
<b>Figure 48:</b> Molecular structure of CDK4/6 inhibitors approved by the FDA with its corresponding name, reference, and elimination half-life	65
<b>Figure 49:</b> Effect of palbociclib on the Cyclin D-CDK4/6 complex	65
<b>Figure 50:</b> Cyclin D1 is a key regulator of proliferation in JN-DSRCT1	66
<b>Figure 51:</b> Palbociclib treatment decrease tumor size of JN-DSRCT1 derived tumors	67
<b>Figure 52:</b> Palbociclib treatment does not affect proliferation nor induce apoptosis in JN-DSRCT1 -derived tumors xenografts	68
<b>Figure 53:</b> Histological view of St. Jude DSRCT tumors are similar to JN-DSRCT1 derived tumors and primary human DSRCT tumors	69
<b>Figure 54:</b> Palbociclib treatment highly impairs DSRCT proliferation in vivo	70
<b>Figure 55:</b> Palbociclib treatment does not affect proliferation nor induce apoptosis in human primary tumor xenografts	71
<b>Figure 56:</b> Schematic representation of EWS, FL1 and type 1 EWS-FLI1 fusion protein and its corresponding domains	85
<b>Figure 57:</b> Histology subtypes of synovial sarcoma	87
<b>Figure 58:</b> Schematic representation of the domains of SS18, SSX and SS18-SSX fusion protein and its corresponding domains	88



<b>Figure 59:</b>	Schematic representation of the EWSR1-WT1 fusion gene in JN-DSRCT1	96
<b>Figure 60:</b>	Summary of the BaseScope assay (from the ACDbio website)	109

# List of Tables

	<b>pg</b>
<b>Table 1:</b> Major gene fusions of bone and soft tissue tumors with poor prognosis	8
<b>Table 2:</b> List of identified EWS-WT1 targets in the literature	16
<b>Table 3:</b> Summary of DNA binding motifs of EWS-WT1 in the literature	17
<b>Table 4:</b> Motif enrichment analysis of EWS-WT1 isoforms in MeT-5A cells show similar but specific binding motifs	37
<b>Table 5:</b> Summary of DNA binding motifs of WT1 in the literature (Adapted from Ullmark et al.)	42
<b>Table 6:</b> MACS-called peaks analysis of WT1 isoforms in MeT-5A cells are different from EWS-WT1 motifs	43
<b>Table 7:</b> List of 1ZZ BaseScope probes, target genes, and its respective detection channel and dye	50
<b>Table 8:</b> Motif enrichment analysis of EWS-WT1 isoforms expressed simultaneously within the same MeT-5A cell share similarities the motifs identified in JN-DSRCT1 cell line	55
<b>Table 9:</b> List of the orthotopic xenografts derived from DSRCT patients received from the St. Jude Children's Research Hospital of Philadelphia (USA)	69
<b>Table 10:</b> List of translocations detected in Ewing's Sarcoma	86
<b>Table 11:</b> List of plasmids used in this study	100
<b>Table 12:</b> List of shRNA targeting sequences in pKLO.1 plasmids used in this study	100
<b>Table 13:</b> Oligonucleotides sequence used in this study for the qPCR and corresponding annealing temperature	101

<b>Table 14:</b>	List of antibodies used in this study, with its corresponding dilution, provider and other useful information	102
------------------	---	-----

## My Participation to Figures

I performed most of the experiment described in this manuscript. However, I also collaborated with researchers in Switzerland and the USA, and I would like to thank them for their contribution and help. Hereafter is a detailed list of my participation for figures and tables.

**Figure 1:** From Hallmarks of cancer: the next generation<sup>2</sup>

**Figure 2:** Graphic design made by myself

**Figure 3:** Graphic design made by myself

**Figure 4:** H&E picture taken by myself

**Figure 5:** Graphic design made by myself

**Figure 6:** Graphic design made by myself

**Figure 7:** Graphic design made by myself

**Figure 8:** Graphic design made by myself

**Figure 9:** Modified from Thermo Fisher Scientific

**Figure 10:** Modified from Panzeri *et al.*

**Figure 11:** Samples collected and experiences performed by G. Boulay, analysis done by S. Iyer

**Figure 12:** Samples collected by myself and experiences performed by G. Boulay, analysis done by S. Iyer

**Figure 13:** Samples collected by myself and experiences performed by G. Boulay, analysis done by S. Iyer

**Figure 14:** Samples collected and experiences performed by G. Boulay, analysis done by S. Iyer

**Figure 15:** Samples collected and experiences performed by G. Boulay, and myself, analysis done by S. Iyer

**Figure 16:** Samples collected and experiences performed by G. Boulay, and myself, analysis done by S. Iyer

**Figure 17:** Modified from Miltenyi Biotec

**Figure 18:** Samples collected, experience performed and analyzed by myself

**Figure 19:** Samples collected, experience performed and analyzed by myself

**Figure 20:** Samples collected, experience performed and analyzed by myself

**Figure 21:** Samples collected, experience performed and analyzed by myself

**Figure 22:** Samples collected, experience performed and analyzed by myself

**Figure 23:** Samples collected by myself and analyzed by S. Iyer

**Figure 24:** Samples collected and experiences performed by myself, analysis done by S. Iyer

**Figure 25:** Samples collected and experiences performed by myself, analysis done by S. Iyer

**Figure 26:** Samples collected and experiences performed by G. Boulay

**Figure 27:** Samples collected and experiences performed by G. Boulay

**Figure 28:** Samples collected, experience performed and analyzed by myself

**Figure 29:** Samples collected, experience performed and analyzed by myself

**Figure 30:** Samples collected and experiences performed by myself, analysis done by S. Iyer

**Figure 31:** Samples collected and experiences performed by G. Boulay

**Figure 32:** Samples collected and experiences performed by G. Boulay

**Figure 33:** Samples collected, experience performed and analyzed by myself

**Figure 34:** Samples collected and experiences performed by G. Boulay

**Figure 35:** Samples collected and experiences performed by G. Boulay

- Figure 36:** Samples collected, experience performed and analyzed by myself
- Figure 37:** Samples collected, experience performed and analyzed by myself
- Figure 38:** Samples collected by G. Boulay and analyzed S. Iyer
- Figure 39:** Samples collected, experience performed and analyzed by myself
- Figure 40:** Samples collected, experience performed and analyzed by myself
- Figure 41:** Samples collected and experiences performed by myself, analysis done by Iyer
- Figure 42:** Samples collected, experience performed and analyzed by myself
- Figure 43:** Samples collected, experience performed and analyzed by myself
- Figure 44:** Samples collected, experience performed and analyzed by myself
- Figure 45:** Samples collected by myself and G. Boulay, experiences performed by S. Rajandran, analysis done by Iyer
- Figure 46:** Statistical analysis done by S. Iyer
- Figure 47:** Samples collected by myself and G. Boulay, experiences performed by S. Rajandran, analysis done by Iyer
- Figure 48:** Graphic design made by myself
- Figure 49:** Graphic design made by myself
- Figure 50:** Experience supervised by myself, but samples collected, experience performed and analyzed by B. Duc.
- Figure 51:** Experience designed, prepared and analyzed by myself. Daily gavage and measures done by the animal caretakers in Epalinges.
- Figure 52:** Samples collected by myself, staining done by the mouse facility
- Figure 53:** PDX1, PDX2 and JN-DSRCT1 tumors samples collected by myself
- Figure 54:** Experience designed, prepared and analyzed by myself. Daily gavage and measures done by the animal caretakers in Epalinges.



**Figure 55:** Samples collected by myself, staining done by the mouse facility

**Figure 56:** Graphic design made by myself

**Figure 57:** H&E from the Swiss medical weekly

**Figure 58:** Graphic design made by myself

**Figure 59:** Graphic design made by myself

**Figure 60:** Modified from the ACDbio website

## My Participation to Tables

**Table 1:** I recapitulated all the information in this table

**Table 2:** I recapitulated all the information in this table

**Table 3:** I recapitulated all the information in this table

**Table 4:** Samples collected and experiences performed by myself, analysis done by S. Iyer

**Table 5:** Adapted from Ullmark *et al.*

**Table 6:** Samples collected and experiences performed by myself, analysis done by S. Iyer

**Table 7:** I recapitulated all the information in this table

**Table 8:** Samples collected and experiences performed by myself, analysis done by S. Iyer

**Table 9:** Information received from the St. Jude Children's Research Hospital of Philadelphia  
(USA)

**Table 10:** I recapitulated all the information in this table

**Table 11:** I recapitulated all the information in this table

**Table 12:** I recapitulated all the information in this table

**Table 13:** I recapitulated all the information in this table

**Table 14:** I recapitulated all the information in this table

## References

- 1 Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57-70, doi:10.1016/s0092-8674(00)81683-9 (2000).
- 2 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).
- 3 Liao, Y., Feng, Y., Shen, J., Hornicek, F. J. & Duan, Z. The roles and therapeutic potential of cyclin-dependent kinases (CDKs) in sarcoma. *Cancer and Metastasis Reviews* **35**, 151-163, doi:10.1007/s10555-015-9601-1 (2016).
- 4 de Dueñas, E. M. *et al.* Preclinical and clinical development of palbociclib and future perspectives. *Clinical and Translational Oncology* **20**, 1136-1144, doi:10.1007/s12094-018-1850-3 (2018).
- 5 Schettini, F. *et al.* CDK 4/6 Inhibitors as Single Agent in Advanced Solid Tumors. *Frontiers in Oncology* **8**, doi:10.3389/fonc.2018.00608 (2018).
- 6 Perez, M., Muñoz-Galván, S., Jiménez-García, M. P., Marín, J. J. & Carnero, A. Efficacy of CDK4 inhibition against sarcomas depends on their levels of CDK4 and p16ink4 mRNA. *Oncotarget* **6** (2015).
- 7 Rauscher, F., Morris, J., Tournay, O., Cook, D. & Curran, T. Binding of the Wilms' tumor locus zinc finger protein to the EGR-1 consensus sequence. *Science* **250**, 1259-1262, doi:10.1126/science.2244209 (1990).
- 8 VanArsdale, T., Boshoff, C., Arndt, K. T. & Abraham, R. T. Molecular Pathways: Targeting the Cyclin D–CDK4/6 Axis for Cancer Treatment. *Clinical Cancer Research* **21**, 2905-2910, doi:10.1158/1078-0432.ccr-14-0816 (2015).
- 9 Topacio, B. R. *et al.* Cyclin D-Cdk4,6 Drives Cell-Cycle Progression via the Retinoblastoma Protein's C-Terminal Helix. *Molecular cell* **74**, 758-770.e754, doi:10.1016/j.molcel.2019.03.020 (2019).
- 10 Pandey, K. *et al.* Molecular mechanisms of resistance to CDK4/6 inhibitors in breast cancer: A review. *Int J Cancer* **145**, 1179-1188, doi:10.1002/ijc.32020 (2019).
- 11 Pashankar, F. D., Rodriguez-Galindo, C. & Pappo, A. Development of a Therapeutic Approach to Rare Cancers in Children: The Children's Oncology Group Experience. *Journal of Pediatric Hematology/Oncology* **34**, S37-S38, doi:10.1097/MPH.0b013e31824e378e (2012).

- 12 Dupain, C., Harttrampf, A. C., Urbinati, G., Geoerger, B. & Massaad-Massade, L. Relevance of Fusion Genes in Pediatric Cancers: Toward Precision Medicine. *Mol Ther Nucleic Acids* **6**, 315-326, doi:10.1016/j.omtn.2017.01.005 (2017).
- 13 Smith, G. M., Johnson, G. D., Grimer, R. J. & Wilson, S. Trends in presentation of bone and soft tissue sarcomas over 25 years: little evidence of earlier diagnosis. *Ann R Coll Surg Engl* **93**, 542-547, doi:10.1308/147870811X13137608455055 (2011).
- 14 Burningham, Z., Hashibe, M., Spector, L. & Schiffman, J. D. The Epidemiology of Sarcoma. *Clinical Sarcoma Research* **2**, 14, doi:10.1186/2045-3329-2-14 (2012).
- 15 Sbaraglia, M., Bellan, E. & Dei Tos, A. P. The 2020 WHO Classification of Soft Tissue Tumours: news and perspectives. *Pathologica*, doi:10.32074/1591-951x-213 (2020).
- 16 Riggi, N. & Stamenkovic, I. The Biology of Ewing sarcoma. *Cancer Letters* **254**, 1-10, doi:<http://dx.doi.org/10.1016/j.canlet.2006.12.009> (2007).
- 17 Riggi, N., Cironi, L., Suvà, M. L. & Stamenkovic, I. Sarcomas: genetics, signalling, and cellular origins. Part 1: The fellowship of TET. *The Journal of Pathology* **213**, 4-20, doi:10.1002/path.2209 (2007).
- 18 Mertens, F., Panagopoulos, I. & Mandahl, N. Genomic characteristics of soft tissue sarcomas. *Virchows Archiv* **456**, 129-139, doi:10.1007/s00428-009-0736-8 (2010).
- 19 Messahel, B., Nash, R., Jeffrey, I., Pritchard-Jones, K. & Hing, S. Clinical features of molecular pathology of solid tumours in childhood. *The Lancet Oncology* **6**, 421-430, doi:[https://doi.org/10.1016/S1470-2045\(05\)70209-6](https://doi.org/10.1016/S1470-2045(05)70209-6) (2005).
- 20 Nacev, B. A. *et al.* The epigenomics of sarcoma. *Nature Reviews Cancer*, doi:10.1038/s41568-020-0288-4 (2020).
- 21 Suvà, M.-L., Cironi, L., Riggi, N. & Stamenkovic, I. Sarcomas: genetics, signalling, and cellular origins. Part 2: TET-independent fusion proteins and receptor tyrosine kinase mutations. *The Journal of Pathology* **213**, 117-130, doi:10.1002/path.2208 (2007).
- 22 Feinberg, A. P., Koldobskiy, M. A. & Göndör, A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nat Rev Genet* **17**, 284-299, doi:10.1038/nrg.2016.13 (2016).
- 23 Riggi, N. *et al.* EWS-FLI-1 Expression Triggers a Ewing's Sarcoma Initiation Program in Primary Human Mesenchymal Stem Cells. *Cancer Research* **68**, 2176-2185, doi:10.1158/0008-5472.can-07-1761 (2008).

- 24 Gerald, W. L. & Rosai, J. Case 2. Desmoplastic small round cell tumor with divergent differentiation. *Pediatr Pathol.* **9**, 177-183, doi:10.3109/15513818909022347 (1989).
- 25 Fletcher CDM, Bridge JA, Hogendoorn PCW & F, M. Vol. Volume 5 468 (Lyon, 2013).
- 26 Lettieri, C. K., Garcia-Filion, P. & Hingorani, P. Incidence and Outcomes of Desmoplastic Small Round Cell Tumor: Results from the Surveillance, Epidemiology, and End Results Database. *Journal of Cancer Epidemiology* **2014**, 5, doi:10.1155/2014/680126 (2014).
- 27 Mora, J. *et al.* Desmoplastic small round cell tumor 20 years after its discovery. *Future Oncology* **11**, 1071-1081, doi:10.2217/fon.15.32 (2015).
- 28 Loktev, A. & Shipley, J. M. Desmoplastic small round cell tumor (DSRCT): emerging therapeutic targets and future directions for potential therapies. *Expert Opinion on Therapeutic Targets* **24**, 281-285, doi:10.1080/14728222.2020.1738392 (2020).
- 29 Bulbul, A. *et al.* Desmoplastic Small Round Blue Cell Tumor: A Review of Treatment and Potential Therapeutic Genomic Alterations. *Sarcoma* **2017**, 1278268, doi:10.1155/2017/1278268 (2017).
- 30 Jordan, A. H. & Pappo, A. Management of Desmoplastic Small Round-cell Tumors in Children and Young Adults. *Journal of Pediatric Hematology/Oncology* **34**, S73-S75, doi:10.1097/MPH.0b013e31824e38ad (2012).
- 31 Davis, J. L. & Rudzinski, E. R. Small Round Blue Cell Sarcoma Other Than Ewing Sarcoma: What Should an Oncologist Know? *Current treatment options in oncology* **21**, 90, doi:10.1007/s11864-020-00785-1 (2020).
- 32 van Erp, A. E. M. *et al.* Olaparib and temozolomide in desmoplastic small round cell tumors: a promising combination in vitro and in vivo. *Journal of cancer research and clinical oncology* **146**, 1659-1670, doi:10.1007/s00432-020-03211-z (2020).
- 33 Bexelius, T. S., Wasti, A. & Chisholm, J. C. Mini-Review on Targeted Treatment of Desmoplastic Small Round Cell Tumor. *Front Oncol* **10**, 518, doi:10.3389/fonc.2020.00518 (2020).
- 34 Gerald, W. L., Rosai, J. & Ladanyi, M. Characterization of the genomic breakpoint and chimeric transcripts in the EWS-WT1 gene fusion of desmoplastic small round cell tumor. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 1028-1032 (1995).
- 35 Lee, J. *et al.* EWSR1, a multifunctional protein, regulates cellular function and aging via genetic and epigenetic pathways. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1865**, 1938-1945, doi:<https://doi.org/10.1016/j.bbadis.2018.10.042> (2019).

- 36 Riggi, N. *et al.* EWS-FLI1 utilizes divergent chromatin remodeling mechanisms to directly activate or repress enhancer elements in Ewing sarcoma. *Cancer cell* **26**, 668-681, doi:10.1016/j.ccell.2014.10.004 (2014).
- 37 Gangwal, K. *et al.* Microsatellites as EWS/FLI response elements in Ewing's sarcoma. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 10149-10154, doi:10.1073/pnas.0801073105 (2008).
- 38 Williams, L. A. & Spector, L. G. Survival Differences Between Males and Females Diagnosed With Childhood Cancer. *JNCI Cancer Spectr* **3**, pkz032-pkz032, doi:10.1093/jncics/pkz032 (2019).
- 39 Jawad, M. U. *et al.* Ewing sarcoma demonstrates racial disparities in incidence-related and sex-related differences in outcome. *Cancer* **115**, 3526-3536, doi:10.1002/cncr.24388 (2009).
- 40 Beck, R. *et al.* EWS/FLI-responsive GGAA microsatellites exhibit polymorphic differences between European and African populations. *Cancer Genet* **205**, 304-312, doi:10.1016/j.cancergen.2012.04.004 (2012).
- 41 Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382, doi:10.1126/science.aaf4382 (2017).
- 42 Han, Tina W. *et al.* Cell-free Formation of RNA Granules: Bound RNAs Identify Features and Components of Cellular Assemblies. *Cell* **149**, 768-779, doi:10.1016/j.cell.2012.04.016 (2012).
- 43 Call, K. M. *et al.* Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumor locus. *Cell* **60**, 509-520, doi:10.1016/0092-8674(90)90601-a (1990).
- 44 Ullmark, T. *et al.* Distinct global binding patterns of the Wilms' tumor gene 1 (WT1) -KTS and +KTS isoforms in leukemic cells. *Haematologica*, haematol.2016.149815, doi:10.3324/haematol.2016.149815 (2016).
- 45 Rivera, M. N. & Haber, D. A. Wilms' tumour: connecting tumorigenesis and organ development in the kidney. *Nat Rev Cancer* **5**, 699-712, doi:10.1038/nrc1696 (2005).
- 46 Yang, L., Han, Y., Saurez Saiz, F. & Minden, M. D. A tumor suppressor and oncogene: the WT1 story. *Leukemia* **21**, 868-876, doi:10.1038/sj.leu.2404624 (2007).
- 47 Chau, Y.-Y. & Hastie, N. D. The role of Wt1 in regulating mesenchyme in cancer, development, and tissue homeostasis. *Trends in Genetics* **28**, 515-524, doi:<https://doi.org/10.1016/j.tig.2012.04.004> (2012).

- 48 Han, Y., San-Marina, S., Liu, J. & Minden, M. D. Transcriptional activation of c-myc proto-oncogene by WT1 protein. *Oncogene* **23**, 6933-6941, doi:10.1038/sj.onc.1207609 (2004).
- 49 Qi, X.-w. *et al.* Wilms' tumor 1 (WT1) expression and prognosis in solid cancer patients: a systematic review and meta-analysis. *Sci Rep* **5**, 8924-8924, doi:10.1038/srep08924 (2015).
- 50 Nurmemmedov, E., Yengo, R. K., Uysal, H., Karlsson, R. & Thunnissen, M. M. New insights into DNA-binding behavior of Wilms tumor protein (WT1)--a dual study. *Biophysical chemistry* **145**, 116-125, doi:10.1016/j.bpc.2009.09.009 (2009).
- 51 Kent, J., Coriat, A. M., Sharpe, P. T., Hastie, N. D. & van Heyningen, V. The evolution of WT1 sequence and expression pattern in the vertebrates. *Oncogene* **11**, 1781-1792 (1995).
- 52 Haber, D. A. *et al.* Alternative splicing and genomic structure of the Wilms tumor gene WT1. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 9618-9622, doi:10.1073/pnas.88.21.9618 (1991).
- 53 Stetefeld, J. & Ruegg, M. A. Structural and functional diversity generated by alternative mRNA splicing. *Trends in Biochemical Sciences* **30**, 515-521, doi:10.1016/j.tibs.2005.07.001 (2005).
- 54 Wells, J., Rivera, M. N., Kim, W. J., Starbuck, K. & Haber, D. A. The Predominant WT1 Isoform (+KTS) Encodes a DNA-Binding Protein Targeting the Planar Cell Polarity Gene Scribble in Renal Podocytes. *Molecular Cancer Research* **8**, 975-985, doi:10.1158/1541-7786.mcr-10-0033 (2010).
- 55 Laity, J. H., Dyson, H. J. & Wright, P. E. Molecular basis for modulation of biological function by alternate splicing of the Wilms' tumor suppressor protein. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 11932-11935, doi:10.1073/pnas.97.22.11932 (2000).
- 56 Fischbach, B. V., Trout, K. L., Lewis, J., Luis, C. A. & Sika, M. WAGR syndrome: a clinical review of 54 cases. *Pediatrics* **116**, 984-988, doi:10.1542/peds.2004-0467 (2005).
- 57 Ezaki, J. *et al.* Gonadal tumor in Frasier syndrome: a review and classification. *Cancer prevention research (Philadelphia, Pa.)* **8**, 271-276, doi:10.1158/1940-6207.capr-14-0415 (2015).
- 58 Koziell, A. & Grundy, R. Frasier and Denys-Drash syndromes: different disorders or part of a spectrum? *Arch Dis Child* **81**, 365-369, doi:10.1136/ad.81.4.365 (1999).
- 59 Knudson, A. G. Two genetic hits (more or less) to cancer. *Nature Reviews Cancer* **1**, 157-162, doi:10.1038/35101031 (2001).
- 60 Lee, S. B. *et al.* The EWS-WT1 translocation product induces PDGFA in desmoplastic small round-cell tumour. *Nature Genetics* **17**, 309-313, doi:10.1038/ng1197-309 (1997).



- 61 Finkeltoy, I. *et al.* Transcriptional regulation of IGF-I receptor gene expression by novel isoforms of the EWS-WT1 fusion protein. *Oncogene* **21**, 1890-1898, doi:10.1038/sj.onc.1205042 (2002).
- 62 Karnieli, E., Werner, H., Rauscher, F. J., 3rd, Benjamin, L. E. & LeRoith, D. The IGF-I receptor gene promoter is a molecular target for the Ewing's sarcoma-Wilms' tumor 1 fusion protein. *J Biol Chem* **271**, 19304-19309, doi:10.1074/jbc.271.32.19304 (1996).
- 63 Wong, J. C. *et al.* Induction of the interleukin-2/15 receptor beta-chain by the EWS-WT1 translocation product. *Oncogene* **21**, 2009-2019, doi:10.1038/sj.onc.1205262 (2002).
- 64 Bulbul, A., Shen, J. P., Xiu, J., Tamayo, P. & Husain, H. Genomic and Proteomic Alterations in Desmoplastic Small Round Blue-Cell Tumors. *JCO Precision Oncology*, 1-9, doi:10.1200/PO.17.00170 (2018).
- 65 Hingorani, P. *et al.* Transcriptome analysis of desmoplastic small round cell tumors identifies actionable therapeutic targets: a report from the Children's Oncology Group. *Sci Rep* **10**, 12318, doi:10.1038/s41598-020-69015-w (2020).
- 66 Rachfal, A. W., Luquette, M. H. & Brigstock, D. R. Expression of connective tissue growth factor (CCN2) in desmoplastic small round cell tumour. *Journal of Clinical Pathology* **57**, 422, doi:10.1136/jcp.2003.012344 (2004).
- 67 Li, H. *et al.* Adenosine Transporter ENT4 Is a Direct Target of EWS/WT1 Translocation Product and Is Highly Expressed in Desmoplastic Small Round Cell Tumor. *PLOS ONE* **3**, e2353, doi:10.1371/journal.pone.0002353 (2008).
- 68 Palmer, R. E. *et al.* Induction of BAIAP3 by the EWS-WT1 chimeric fusion implicates regulated exocytosis in tumorigenesis. *Cancer Cell* **2**, 497-505, doi:[https://doi.org/10.1016/S1535-6108\(02\)00205-2](https://doi.org/10.1016/S1535-6108(02)00205-2) (2002).
- 69 Kang, H.-J. *et al.* EWS-WT1 Oncoprotein Activates Neuronal Reprogramming Factor ASCL1 and Promotes Neural Differentiation. *Cancer Research* **74**, 4526-4535, doi:10.1158/0008-5472.can-13-3663 (2014).
- 70 Dufresne, A. *et al.* Desmoplastic small round cell tumor: current management and recent findings. *Sarcoma* **2012**, 714986-714986, doi:10.1155/2012/714986 (2012).
- 71 Gerald, W. L. & Haber, D. A. The EWS-WT1 gene fusion in desmoplastic small round cell tumor. *Seminars in Cancer Biology* **15**, 197-205, doi:<https://doi.org/10.1016/j.semcancer.2005.01.005> (2005).
- 72 Reynolds, P. A. *et al.* Identification of a DNA-binding site and transcriptional target for the

- EWS–WT1(+KTS) oncoprotein. *Genes & Development* **17**, 2094-2107, doi:10.1101/gad.1110703 (2003).
- 73 Ito, E. *et al.* A tetraspanin-family protein, 'T-cell acute lymphoblastic leukemia-associated antigen 1, is induced by the Ewing's sarcoma-Wilms' tumor 1 fusion protein of desmoplastic small round-cell tumor. *Am J Pathol* **163**, 2165-2172, doi:10.1016/s0002-9440(10)63573-0 (2003).
  - 74 Bandopadhyay, P. *et al.* The oncogenic properties of EWS/WT1 of desmoplastic small round cell tumors are unmasked by loss of p53 in murine embryonic fibroblasts. *BMC Cancer* **13**, 585-585, doi:10.1186/1471-2407-13-585 (2013).
  - 75 Liu, J. *et al.* Molecular Heterogeneity and Function of EWS-WT1 Fusion Transcripts in Desmoplastic Small Round Cell Tumors. *Clinical Cancer Research* **6**, 3522-3529 (2000).
  - 76 Chow, W. A. *et al.* Recurrent secondary genomic alterations in desmoplastic small round cell tumors. *BMC medical genetics* **21**, 101, doi:10.1186/s12881-020-01034-w (2020).
  - 77 Widschwendter, M. *et al.* Epigenome-based cancer risk prediction: rationale, opportunities and challenges. *Nature reviews. Clinical oncology* **15**, 292-309, doi:10.1038/nrclinonc.2018.30 (2018).
  - 78 Popova, E. & Barnstable, C. J. Epigenetics rules. *J Ocul Biol Dis Infor* **4**, 93-94, doi:10.1007/s12177-012-9088-8 (2012).
  - 79 Jin, Z. & Liu, Y. DNA methylation in human diseases. *Genes & Diseases* **5**, 1-8, doi:<https://doi.org/10.1016/j.gendis.2018.01.002> (2018).
  - 80 Rasmussen, K. D. & Helin, K. Role of TET enzymes in DNA methylation, development, and cancer. *Genes & development* **30**, 733-750, doi:10.1101/gad.276568.115 (2016).
  - 81 Kong, L. *et al.* A primary role of TET proteins in establishment and maintenance of De Novo bivalency at CpG islands. *Nucleic acids research* **44**, 8682-8692, doi:10.1093/nar/gkw529 (2016).
  - 82 Melamed, P., Yosefzon, Y., David, C., Tsukerman, A. & Pnueli, L. Tet Enzymes, Variants, and Differential Effects on Function. *Front Cell Dev Biol* **6**, 22-22, doi:10.3389/fcell.2018.00022 (2018).
  - 83 Dawson, M. A. & Kouzarides, T. Cancer epigenetics: from mechanism to therapy. *Cell* **150**, 12-27, doi:10.1016/j.cell.2012.06.013 (2012).
  - 84 Ito, S. *et al.* Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129-1133, doi:10.1038/nature09303 (2010).
  - 85 Audia, J. E. & Campbell, R. M. Histone Modifications and Cancer. *Cold Spring Harbor Perspec-*

- tives in Biology* **8**, doi:10.1101/cshperspect.a019521 (2016).
- 86 Suvà, M. L., Riggi, N. & Bernstein, B. E. Epigenetic Reprogramming in Cancer. *Science* **339**, 1567-1570, doi:10.1126/science.1230184 (2013).
- 87 Jones, P. A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* **13**, 484-492, doi:10.1038/nrg3230 (2012).
- 88 Cossío, F. P., Esteller, M. & Berdasco, M. Towards a more precise therapy in cancer: Exploring epigenetic complexity. *Current opinion in chemical biology* **57**, 41-49, doi:10.1016/j.cbpa.2020.04.008 (2020).
- 89 Panzeri, I., Rossetti, G. & Pagani, M. in *Medical Epigenetics* (ed T. O. Tollefsbol) 47-63 (Academic Press, 2016).
- 90 Ghasemi, S. Cancer's epigenetic drugs: where are they in the cancer medicines? *The Pharmacogenomics Journal* **20**, 367-379, doi:10.1038/s41397-019-0138-5 (2020).
- 91 Nishio, J. *et al.* Establishment and Characterization of a Novel Human Desmoplastic Small Round Cell Tumor Cell Line, JN-DSRCT-1. *Lab Invest* **82**, 1175-1182 (2002).
- 92 Lee, H.-Y. & Hong, I.-S. Double-edged sword of mesenchymal stem cells: Cancer-promoting versus therapeutic potential. *Cancer Science*, n/a-n/a, doi:10.1111/cas.13334 (2017).
- 93 Kauer, M. *et al.* A Molecular Function Map of Ewing's Sarcoma. *PLOS ONE* **4**, e5415, doi:10.1371/journal.pone.0005415 (2009).
- 94 Riggi, N. *et al.* Development of Ewing's Sarcoma from Primary Bone Marrow-Derived Mesenchymal Progenitor Cells. *Cancer Research* **65**, 11459-11468, doi:10.1158/0008-5472.can-05-1696 (2005).
- 95 Riggi, N. *et al.* Expression of the FUS-CHOP Fusion Protein in Primary Mesenchymal Progenitor Cells Gives Rise to a Model of Myxoid Liposarcoma. *Cancer Research* **66**, 7016-7023, doi:10.1158/0008-5472.can-05-3979 (2006).
- 96 Cironi, L. *et al.* Epigenetic Features of Human Mesenchymal Stem Cells Determine Their Permissiveness for Induction of Relevant Transcriptional Changes by SYT-SSX1. *PLoS ONE* **4**, e7904, doi:10.1371/journal.pone.0007904 (2009).
- 97 Suva, D. *et al.* Non-hematopoietic human bone marrow contains long-lasting, pluripotential mesenchymal stem cells. *Journal of Cellular Physiology* **198**, 110-118, doi:10.1002/jcp.10396 (2004).
- 98 Caplan, A. I. Mesenchymal stem cells. *Journal of Orthopaedic Research* **9**, 641-650, doi:10.1002/jor.1100090504 (1991).

- 99 Davies, O. G., Smith, A. J., Cooper, P. R., Shelton, R. M. & Scheven, B. A. The effects of cryopreservation on cells isolated from adipose, bone marrow and dental pulp tissues. *Cryobiology* **69**, 342-347, doi:<https://doi.org/10.1016/j.cryobiol.2014.08.003> (2014).
- 100 Heo, J. S., Choi, Y., Kim, H.-S. & Kim, H. O. Comparison of molecular profiles of human mesenchymal stem cells derived from bone marrow, umbilical cord blood, placenta and adipose tissue. *Int J Mol Med* **37**, 115-125, doi:10.3892/ijmm.2015.2413 (2016).
- 101 Wagner, W. *et al.* The heterogeneity of human mesenchymal stem cell preparations—Evidence from simultaneous analysis of proteomes and transcriptomes. *Experimental Hematology* **34**, 536-548, doi:<https://doi.org/10.1016/j.exphem.2006.01.002> (2006).
- 102 Kfoury, Y. & Scadden, David T. Mesenchymal Cell Contributions to the Stem Cell Niche. *Cell Stem Cell* **16**, 239-253, doi:10.1016/j.stem.2015.02.019 (2015).
- 103 Dominici, M. *et al.* Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy* **8**, 315-317, doi:10.1080/14653240600855905 (2006).
- 104 Yengo, R. K., Nurmemmedov, E. & Thunnissen, M. M. G. M. Structure of WT1 zinc fingers bound to its cognate DNA: Implications of the KTS insert. *bioRxiv*, 271577, doi:10.1101/271577 (2018).
- 105 Mohammadi, P. *et al.* Phase transitions as intermediate steps in the formation of molecularly engineered protein fibers. *Communications Biology* **1**, 86, doi:10.1038/s42003-018-0090-y (2018).
- 106 Strom, A. R. & Brangwynne, C. P. The liquid nucleome – phase transitions in the nucleus at a glance. *Journal of Cell Science* **132**, jcs235093, doi:10.1242/jcs.235093 (2019).
- 107 Dolgin, E. What lava lamps and vinaigrette can teach us about cell biology. *Nature* **555**, 300-302, doi:10.1038/d41586-018-03070-2 (2018).
- 108 Aguzzi, A. & Altmeyer, M. Phase Separation: Linking Cellular Compartmentalization to Disease. *Trends Cell Biol* **26**, 547-558, doi:10.1016/j.tcb.2016.03.004 (2016).
- 109 Boeynaems, S. *et al.* Protein Phase Separation: A New Phase in Cell Biology. *Trends Cell Biol* **28**, 420-435, doi:10.1016/j.tcb.2018.02.004 (2018).
- 110 Franzmann, T. M. & Alberti, S. Prion-like low-complexity sequences: Key regulators of protein solubility and phase behavior. *J Biol Chem* **294**, 7128-7136, doi:10.1074/jbc.TM118.001190 (2019).
- 111 Bolognesi, B. *et al.* The mutational landscape of a prion-like domain. *Nature Communications* **10**, 4162, doi:10.1038/s41467-019-12101-z (2019).

- 112 Couthouis, J. *et al.* A yeast functional screen predicts new candidate ALS disease genes. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 20881-20890, doi:10.1073/pnas.1109434108 (2011).
- 113 Kato, M. *et al.* Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753-767, doi:10.1016/j.cell.2012.04.017 (2012).
- 114 Kwon, I. *et al.* Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. *Cell* **155**, 1049-1060, doi:10.1016/j.cell.2013.10.033 (2013).
- 115 Schwartz, J. C., Wang, X., Podell, E. R. & Cech, T. R. RNA seeds higher-order assembly of FUS protein. *Cell reports* **5**, 918-925, doi:10.1016/j.celrep.2013.11.017 (2013).
- 116 Boulay, G. *et al.* Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain. *Cell* **171**, 163-178.e119, doi:<https://doi.org/10.1016/j.cell.2017.07.036> (2017).
- 117 Berg, J. M. Zinc finger domains: hypotheses and current knowledge. *Annual review of biophysics and biophysical chemistry* **19**, 405-421, doi:10.1146/annurev.bb.19.060190.002201 (1990).
- 118 Bardeesy, N. & Pelletier, J. Overlapping RNA and DNA binding domains of the wt1 tumor suppressor gene product. *Nucleic acids research* **26**, 1784-1792, doi:10.1093/nar/26.7.1784 (1998).
- 119 Shi, Y. & Berg, J. M. Specific DNA-RNA hybrid binding by zinc finger proteins. *Science* **268**, 282-284, doi:10.1126/science.7536342 (1995).
- 120 Larsson, S. H. *et al.* Subnuclear localization of WT1 in splicing or transcription factor domains is regulated by alternative splicing. *Cell* **81**, 391-401, doi:10.1016/0092-8674(95)90392-5 (1995).
- 121 Caricasole, A. *et al.* RNA binding by the Wilms tumor suppressor zinc finger proteins. *Proc Natl Acad Sci U S A* **93**, 7562-7566, doi:10.1073/pnas.93.15.7562 (1996).
- 122 Kim, J., Lee, K. & Pelletier, J. The DNA binding domains of the WT1 tumor suppressor gene product and chimeric EWS/WT1 oncoprotein are functionally distinct. *Oncogene* **16**, 1021-1030, doi:10.1038/sj.onc.1201616 (1998).
- 123 Johnstone, R. W. *et al.* A novel repressor, par-4, modulates transcription and growth suppression functions of the Wilms' tumor suppressor WT1. *Molecular and cellular biology* **16**, 6945-6956, doi:10.1128/mcb.16.12.6945 (1996).
- 124 Wang, F. *et al.* RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J Mol Diagn* **14**, 22-29, doi:10.1016/j.jmoldx.2011.08.002 (2012).
- 125 Sexton, T. & Cavalli, G. The Role of Chromosome Domains in Shaping the Functional Genome. *Cell* **160**, 1049-1059, doi:<http://dx.doi.org/10.1016/j.cell.2015.02.040> (2015).

- 126 Cook Sangar, M. L. *et al.* Inhibition of CDK4/6 by Palbociclib Significantly Extends Survival in Medulloblastoma Patient-Derived Xenograft Mouse Models. *Clinical Cancer Research* **23**, 5802-5813, doi:10.1158/1078-0432.ccr-16-2943 (2017).
- 127 Vijayaraghavan, S. *et al.* CDK4/6 and autophagy inhibitors synergistically induce senescence in Rb positive cytoplasmic cyclin E negative cancers. *Nature Communications* **8**, 15916, doi:10.1038/ncomms15916 (2017).
- 128 Fry, D. W. *et al.* Specific inhibition of cyclin-dependent kinase 4/6 by PD 0332991 and associated antitumor activity in human tumor xenografts. *Molecular cancer therapeutics* **3**, 1427-1438 (2004).
- 129 Gong, X. *et al.* Genomic Aberrations that Activate D-type Cyclins Are Associated with Enhanced Sensitivity to the CDK4 and CDK6 Inhibitor Abemaciclib. *Cancer Cell* **32**, 761-776. e766, doi:<https://doi.org/10.1016/j.ccell.2017.11.006> (2017).
- 130 Vlenterie, M. *et al.* Targeting Cyclin-Dependent Kinases in Synovial Sarcoma: Palbociclib as a Potential Treatment for Synovial Sarcoma Patients. *Ann Surg Oncol* **23**, 2745-2752, doi:10.1245/s10434-016-5341-x (2016).
- 131 Konecny, G. E. *et al.* Expression of p16 and retinoblastoma determines response to CDK4/6 inhibition in ovarian cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **17**, 1591-1602, doi:10.1158/1078-0432.ccr-10-2307 (2011).
- 132 Kadoch, C. *et al.* Dynamics of BAF–Polycomb complex opposition on heterochromatin in normal and oncogenic states. *Nature Genetics* **49**, 213-222, doi:10.1038/ng.3734 (2017).
- 133 McBride, M. J. *et al.* The SS18-SSX Fusion Oncoprotein Hijacks BAF Complex Targeting and Function to Drive Synovial Sarcoma. *Cancer Cell* **33**, 1128-1141.e1127, doi:<https://doi.org/10.1016/j.ccell.2018.05.002> (2018).
- 134 Watt, A. C. *et al.* CDK4/6 inhibition reprograms the breast cancer enhancer landscape by stimulating AP-1 transcriptional activity. *Nature Cancer*, doi:10.1038/s43018-020-00135-y (2020).
- 135 Grünewald, T. G. *et al.* Sarcoma treatment in the era of molecular medicine. *EMBO Molecular Medicine* **12**, e11131, doi:<https://doi.org/10.15252/emmm.201911131> (2020).
- 136 Klein, M. E., Kovatcheva, M., Davis, L. E., Tap, W. D. & Koff, A. CDK4/6 Inhibitors: The Mechanism of Action May Not Be as Simple as Once Thought. *Cancer Cell* **34**, 9-20, doi:10.1016/j.ccell.2018.03.023 (2018).
- 137 Miao, L., Lin, C. M. & Huang, L. Stromal barriers and strategies for the delivery of nanomedicine to desmoplastic tumors. *J Control Release* **219**, 192-204, doi:10.1016/j.jconrel.2015.08.017



- (2015).
- 138 Kutova, O. M., Guryev, E. L., Sokolova, E. A., Alzeibak, R. & Balalaeva, I. V. Targeted Delivery to Tumors: Multidirectional Strategies to Improve Treatment Efficiency. *Cancers* **11**, 68, doi:10.3390/cancers11010068 (2019).
  - 139 Meshram, G. G., Kaur, N. & Hura, K. S. Ewing's sarcoma with distant metastasis: A brief note on management and emerging therapies. *Clin Pract* **9**, 1111-1111, doi:10.4081/cp.2019.1111 (2019).
  - 140 Group, T. E. E. S. N. W. Bone sarcomas: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology* **25**, iii113-iii123, doi:10.1093/annonc/mdu256 (2014).
  - 141 Sbaraglia, M., Righi, A., Gambarotti, M. & Dei Tos, A. P. Ewing sarcoma and Ewing-like tumors. *Virchows Archiv : an international journal of pathology* **476**, 109-119, doi:10.1007/s00428-019-02720-8 (2020).
  - 142 Alfert, A., Moreno, N. & Kerl, K. The BAF complex in development and disease. *Epigenetics Chromatin* **12**, 19-19, doi:10.1186/s13072-019-0264-y (2019).
  - 143 Kadoch, C. & Crabtree, G. R. Reversible disruption of mSWI/SNF (BAF) complexes by the SS18-SSX oncogenic fusion in synovial sarcoma. *Cell* **153**, 71-85, doi:10.1016/j.cell.2013.02.036 (2013).
  - 144 Kadoch, C. & Crabtree, G. R. Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci Adv* **1**, e1500447-e1500447, doi:10.1126/sciadv.1500447 (2015).
  - 145 Lee, T. I. *et al.* Control of Developmental Regulators by Polycomb in Human Embryonic Stem Cells. *Cell* **125**, 301-313, doi:<https://doi.org/10.1016/j.cell.2006.02.043> (2006).
  - 146 Kadoch, C. *et al.* Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nature Genetics* **45**, 592-601, doi:10.1038/ng.2628 (2013).
  - 147 Cironi, L. *et al.* Epigenetic features of human mesenchymal stem cells determine their permissiveness for induction of relevant transcriptional changes by SYT-SSX1. *PLoS one* **4**, e7904-e7904, doi:10.1371/journal.pone.0007904 (2009).
  - 148 Haldar, M., Randall, R. L. & Capecchi, M. R. Synovial sarcoma: from genetics to genetic-based animal modeling. *Clin Orthop Relat Res* **466**, 2156-2167, doi:10.1007/s11999-008-0340-2 (2008).
  - 149 Haldar, M., Hancock, J. D., Coffin, C. M., Lessnick, S. L. & Capecchi, M. R. A Conditional



- Mouse Model of Synovial Sarcoma: Insights into a Myogenic Origin. *Cancer Cell* **11**, 375-388, doi:<https://doi.org/10.1016/j.ccr.2007.01.016> (2007).
- 150 Brett, D. *et al.* The SYT Protein Involved in the t(X;18) Synovial Sarcoma Translocation is a Transcriptional Activator Localised in Nuclear Bodies. *Human Molecular Genetics* **6**, 1559-1564, doi:10.1093/hmg/6.9.1559 (1997).
- 151 Clark, J. *et al.* Identification of novel genes, SYT and SSX, involved in the t(X;18)(p11.2;q11.2) translocation found in human synovial sarcoma. *Nature Genetics* **7**, 502-508, doi:10.1038/ng0894-502 (1994).
- 152 Wang, S. *et al.* Survival changes in Patients with Synovial Sarcoma, 1983-2012. *J Cancer* **8**, 1759-1768, doi:10.7150/jca.17349 (2017).
- 153 Kawai, A., Kondo, T., Suehara, Y., Kikuta, K. & Hirohashi, S. Global Protein-expression Analysis of Bone and Soft Tissue Sarcomas. *A Publication of The Association of Bone and Joint Surgeons® | CORR®* **466** (2008).
- 154 Haldar, M., Hedberg, M. L., Hockin, M. F. & Capecchi, M. R. A CreER based random induction strategy for modeling translocation-associated sarcomas in mice. *Cancer research* **69**, 3657-3664, doi:10.1158/0008-5472.CAN-08-4127 (2009).
- 155 Kato, H. *et al.* SYT Associates with Human SNF/SWI Complexes and the C-terminal Region of Its Fusion Partner SSX1 Targets Histones. *Journal of Biological Chemistry* **277**, 5498-5505, doi:10.1074/jbc.M108702200 (2002).
- 156 Francis, P. *et al.* Diagnostic and prognostic gene expression signatures in 177 soft tissue sarcomas: hypoxia-induced transcription profile signifies metastatic potential. *BMC Genomics* **8**, 73, doi:10.1186/1471-2164-8-73 (2007).
- 157 Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553-560, doi:10.1038/nature06008 (2007).
- 158 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760, doi:10.1093/bioinformatics/btp324 (2009).
- 159 Pohl, A. & Beato, M. bwtool: a tool for bigWig files. *Bioinformatics* **30**, 1618-1619, doi:10.1093/bioinformatics/btu056 (2014).
- 160 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9**, R137, doi:10.1186/gb-2008-9-9-r137 (2008).
- 161 An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74,

- doi:10.1038/nature11247 (2012).
- 162 Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842, doi:10.1093/bioinformatics/btq033 (2010).
- 163 Lawrence, M. S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214-218, doi:10.1038/nature12213 (2013).
- 164 Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell* **38**, 576-589, doi:10.1016/j.molcel.2010.05.004 (2010).
- 165 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21, doi:10.1093/bioinformatics/bts635 (2013).
- 166 Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930, doi:10.1093/bioinformatics/btt656 (2014).
- 167 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
- 168 Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28**, 511-515, doi:10.1038/nbt.1621 (2010).
- 169 Baker, A.-M. *et al.* Robust RNA-based in situ mutation detection delineates colorectal cancer subclonal evolution. *Nature Communications* **8**, 1998, doi:10.1038/s41467-017-02295-5 (2017).