SCUOLA DI DOTTORATO
UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Department of

Medicine and Surgery

PhD program **Molecular and Translational Medicine** Cycle **XXX**

DEVELOPMENT OF AN IN VITRO DISEASE MODEL FOR DISSECTING THE EPIGENETIC MECHANISMS UNDERLYING PATHOGENESIS OF KABUKI SYNDROME

Surname **Fasciani** Name **Alessandra**

Registration number **798701**

Tutor: **Dott. Alessio Zippo**

Coordinator: **Prof. Andrea Biondi**

ACADEMIC YEAR 2016/2017

Table of contents

1. General introduction

1.1Kabuki Syndrome

Kabuki syndrome (KS; OMIM 147920) is a rare autosomal dominant disease that arises from *de novo* mutation in majority of cases. Estimates of birth prevalence vary between 1:32,000 for the Japanese population and 1:86,000 for the populations of New Zealand and Australia. It was first described by the group of Niikawa and Kuroki that described a syndrome of intellectual disability, short stature and a peculiar facial gestalt, which reminded the authors of the make-up of actors in the traditional Japanese Kabuki theater. The molecular cause of KS had long been unknown until 2010. Ng et al. reported the identification of heterozygous mutations in *KMT2D* (MLL4) as the major genetic cause of KS. A second gene for KS was described by Lederer et al. showing heterozygous whole-gene or larger partial-gene deletions of *KDM6A* (UTX). For a large number of patients (approximately 30%), however, the underlying genetic cause remains unidentified. [1]

1.1.1 Clinical features

The groups that first identified the disease described five cardinal manifestations, used for the diagnosis of the disease: typical facial features, skeletal anomalies, postnatal growth deficiency, persistence of fetal fingertip pads and mild to moderate intellectual disability. However, these features may not at all be present even in cases with genetically confirmed KS.

1.1.1.1 Craniofacial anomalies

The most remarkable aspects of KS are the atypical facial features, which usually prompt the clinician to consider this diagnosis. The long palpebral fissures with an eversion of the lower eyelid, long, dense eyelashes and arched eye-brows are the common characteristic features of KS. The ears are prominent, with simply formed or hypoplastic helices, and pre-auricular pits can be observed in some patients. The tip of the nose is typically depressed because of a short columella. The mouth has a thin upper and full lower lip, and the corners of the mouth slant downwards. Prognathism and symmetrical nodules of the lower lip may be observed. It has become evident that this classical facial appearance is observed more often in patients carrying a *KMT2D* mutation. The craniofacial phenotype of KS can include cleft lip/palate and abnormal dentition (including widely spaced teeth and/or hypodontia) [1].

1.1.1.2 Body weight and growth

Children with KS are typically born with normal growth parameters, but fail to thrive during infancy and show postnatal growth retardation because of feeding problems caused by poorly coordinated sucking and swallowing. More than half the patients fail to thrive in infancy, and 57% have a body mass index in the overweight or obese range in middle childhood or adolescence[2].

1.1.1.3 Skeletal anomalies

Post-natal proportionate short stature is a common finding in KS reported in approximately 58% of cases. Moreover, genotype–

phenotype comparisons revealed a considerably higher frequency of short stature in patients with a *KMT2D* mutation[1].

Patients with KS show some skeletal abnormality in about 80% of cases. These anomalies include rib anomalies, malformations of the vertebrae, scoliosis and abnormalities of the fingers (brachydactyly and/or clinodactyly)[3].

KS patients show also cranial abnormalities, including coronal synostosis, metopic synostosis, incomplete development of the frontal and/or maxillary sinuses, digital impression of the skull, and underdevelopment of the mastoid processes.

Joint hypermobility is seen in one-half to three quarters of patients, and joint dislocations are not uncommon, particularly of the hips, patellae, and shoulders[4].

1.1.1.4 Skin and connective tissues

A distinctive symptom of KS is the presence of persistent fetal finger pads. This characteristic is present in up to 92% of patients who carry a mutation in *KMT2D*. Another common finding that has been reported in up to 96% of patients is an abnormality of the dermatoglyphic pattern, such as absence of the digital triradius ‘c’ and ‘d’ and presence of interdigital triradii, as well as hypothenar and interdigital ulnar loop patterns[4].

1.1.1.5 Development and behavior

Intellectual disability of a varying degree has been reported for approximately 90% of KS patients. The degree of intellectual disability ranges from mild to severe with the majority of patients

being mildly affected. However, there have been multiple reports of patients with KS and normal intelligence[3]. However, there have also been reports of more severely affected individuals, including individuals with IQs of <35. No consistent cognitive pattern has been described in KS but a decline in IQ over time has been rarely reported [5].

Apart from the intellectual disability, the psychomotor development of KS patients may be complicated by hypotonia and seizures. Muscular hypotonia is common, described in 51–98% of patients[6,7]. Hypotonia appears to be most prominent in the neonatal period and to improve over time. Schrander-Stumpel et al. reported in a literature review that 24% of cases of KS are complicated by seizures[7].

1.1.1.6 Other symptoms

Involvement of other body apparatus has been found in KS patients.

Congenital heart defects are common in patients with KS. The reported frequencies vary between approximately 40% and 50% in different literature reviews. The most frequent malformations are atrial septal defects, ventricular septal defects and aortic coarctation [3,8].

Recurrent otitis media, upper respiratory tract infections, and/ or pneumonia suggest that patients have susceptibility to infections. Although the underlying mechanism is unknown, certain immunodeficiencies may exist[3]. Also autoimmune disease may be present, in particular idiopathic thrombocytopenic purpura and/or hemolytic anemia in rare cases[1].

1.1.2 Molecular aspects of disease

The genetic etiology of KS was not elucidated until 2010 when Ng et al. reported that heterozygous mutations in *KMT2D* (MIM #602113) are the major genetic cause of KS[9]. Other mutation screenings in KS cohorts found a *KMT2D* mutation in 55.8% to 80% of patients [10]. In 2012, Lederer et al. reported the association of heterozygous whole or large partial deletion of *KDM6A* (#MIM 300128) (observed in three patients) with a broad phenotypic spectrum[11]. Then, other works also reported point mutation or splice-site mutations of *KDM6A* suggesting that *KDM6A* sequencing should be considered in all *KMT2D* mutation-negative KS patients[12,13,14].

1.1.2.1 Type of mutations and distribution

KMT2D was the first gene found to be associated with KS. Considering yet published mutation screening studies, the cumulative mutation detection rate is approximately 67%[1]. All mutations identified so far are heterozygous mutations, most of them truncating, leading most likely to haploinsufficiency of the *KMT2D* protein (named MLL4). Mutations show no significant clustering in a certain exon or a known protein-domain in terms of a mutational hot spot although mutations in *KMT2D* appear to be overrepresented toward the C-terminus (Fig 1).

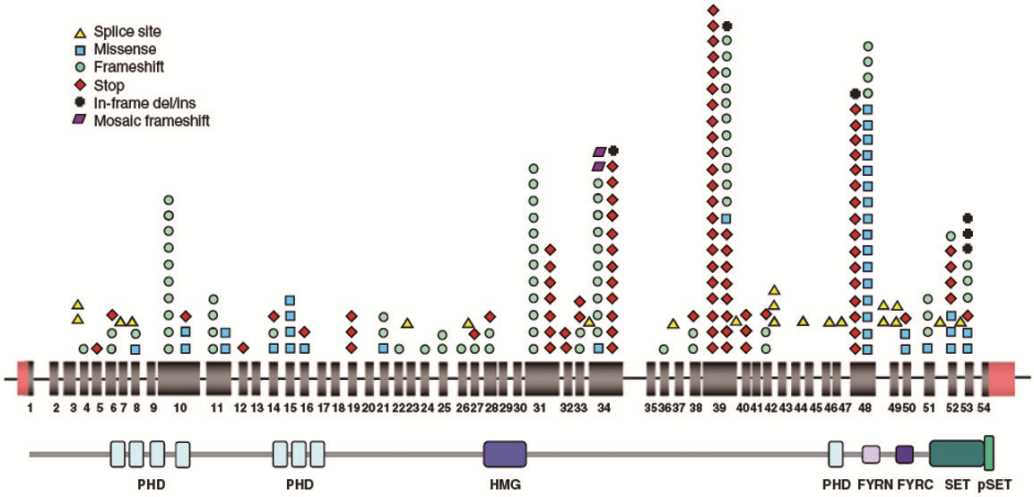


Figure 1. Distribution of *KMT2D* mutations. Schematic view of the *KMT2D* protein domains, all coding exons and localization of all described *KMT2D* mutations. *Unmasking Kabuki syndrome, Clin Genet 2013.*

Although *KDM6A* mutations seem to be a very rare cause of KS, recent works indicate a need to sequence KDM6A in routine diagnostics. Lederer et al. identified *KDM6A* deletions in 3 of 22 *KMT2D* mutation-negative patients, showing a detection rate of 12.5% [11]. A recent report showed two nonsense mutations and one 3-basepair deletion of *KDM6A* in 3 of 32 *KMT2D* mutation-negative KS patients[14].

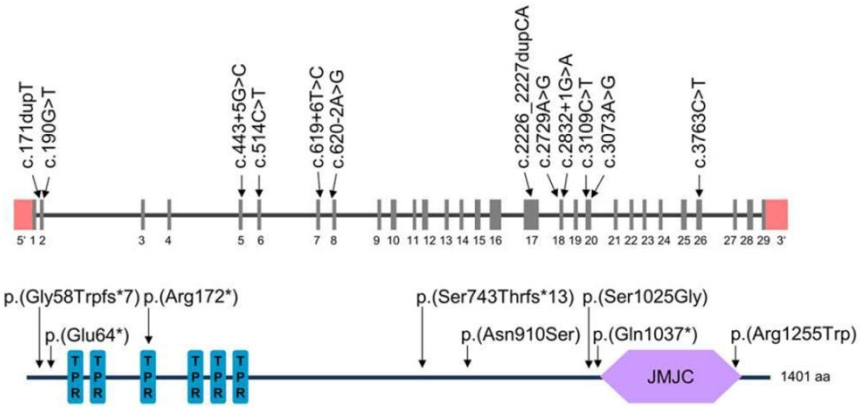


Figure 2. Overview of identified *KDM6A* mutations relative to a schematic representation of the *KDM6A* gene and KDM6A protein structure. *Mutation Update for Kabuki Syndrome Genes KMT2D and KDM6A and Further Delineation of X-Linked Kabuki Syndrome Subtype 2. HUMAN MUTATION, 2016*

1.1.2.2 *KMT2D* and *KDM6A*: molecular pathogenesis

KMT2D gene product (MLL4) is a histone H3 lysine 4 (H3K4)-specific methyltransferase needed for H3K4 monomethylation, a hallmark of an active transcription state. As other H3K4 methyltransferases, MLL4 acts in stable multiprotein complex, the ASCOM complex. Among other proteins, KDM6A gene product is also part of this complex. *KDM6A*, the homologous of Drosophila UTX, is a H3K27 demethylase. Agger et al. put forth a model where ASCOM complex acts through the removal of repressive epigenetic marks, Polycomb displacement and positioning of activating methylation marks[15].

Very little is known about the pathogenesis of KS. Considering *KMT2D*, expression studies showed that it is expressed in a variety of human and murine tissues, and a space and time-specific expression during embryonic development has been shown in mouse embryos[16]. Currently, our knowledge on genes transcriptionally regulated by the

KMT2D complex during development is limited and, therefore, the underlying pathogenesis of KS remains unclear. However, the identification of KMT2D as the causative gene for a multiple congenital malformation syndrome underscored the importance of time and space-specific transcriptional regulation through the orchestrated interplay of different epigenetic marks during embryonic development in humans.

In their work, Ang et al. show how Mll4 deficiency affects heart development. In their mouse model, Kmt2d is required in cardiac precursors and cardiomyocytes during cardiogenesis. In particular, they found that Kmt2d regulate a specific cardiac transcriptional program for ion homeostasis and heart development. Although KS is a multi-organ syndrome in humans, cardiac defects are described in 40-50% of patients and this study shed light on the primary nature of cardiac defects in the absence of Mll4 and uncovers possible etiologies for Kabuki syndrome patients[17].

Another work use zebrafish as animal model to understand the pathogenesis of KS. Indeed, the authors of this work found that knockdown of the zebrafish homologs of *KMT2D* and *KDM6A* causes defects in tissues commonly affected in patients with KS. Knockdown of kmt2d and kdm6a lead to craniofacial defects, most notably hypoplasia of cartilaginous structures of the viscerocranium. Similarly, both kmt2d and kdm6al were required for proper cardiac looping morphogenesis. Both paralogs appear to be involved in neuronal development, as knockdown of either kmt2d or kdm6a is sufficient to induce a reduction in brain size and a reduced population of post-mitotic neurons (Fig 3)[18].

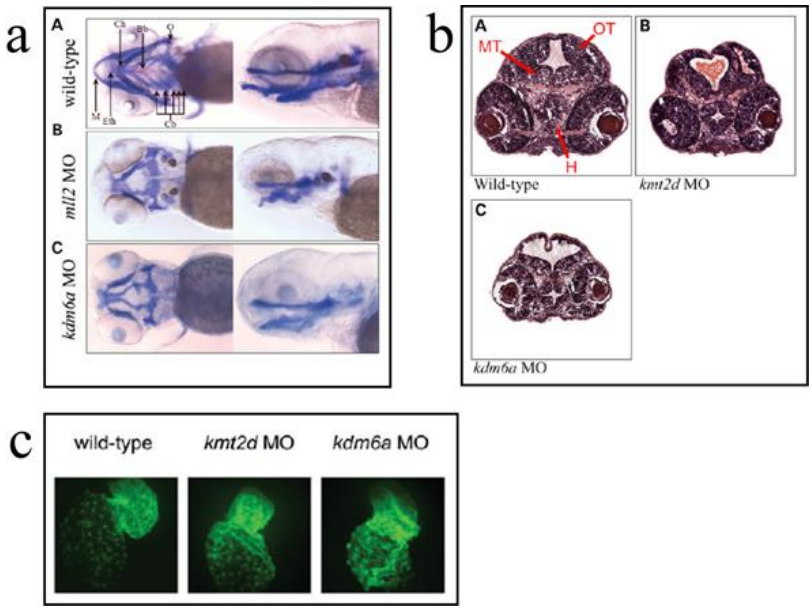


Figure 3. Critical roles of KMT2D and KDM6A in craniofacial, heart and brain development. a, Defects in craniofacial development observed in morphant zebrafish embryos. Five day post-fertilization embryos were treated with alcian and alizarin dyes. Both ventral and left sagittal views focused on the visceral cranial cartilages at 10× magnification are shown. Bb, basibranchial; Cb, ceratobranchial arches 3–7; Ch, ceratohyal; O, opercle; Eth, ethmoid plate; M, Meckel's cartilage. b, Defects in brain morphology observed in morphant zebrafish embryos. Hematoxylin and eosin (H&E) stained transverse sections of the zebrafish containing elements of both fore- and mid-brain at 48 h post-fertilizat. H, hypothalamus; MT, midbrain tegmentum; OT, optic tectum. Defects in heart looping observed in morphant zebrafish embryos. Images shown of hearts from 48 h post-fertilization zebrafish embryos in transgenic line Tg (szhand2:mCherry; cmlc2:gfp)co10 which expresses green fluorescent protein (GFP) under the cmlc2 promoter were analyzed. Modified from *Kabuki syndrome genes KMT2D and KDM6A: Functional analysis demonstrate critical roles incraniofacial, heart and brain development. Human Molecular Genetics, 2015*

Some aspects of the KS phenotype appear to be related to a global growth and developmental delay, including small stature and microcephaly. However, the other defects also observed in zebrafish model do not appear to be due to global underdevelopment but are

likely due to perturbations in the development of specific tissues, resulting from cellular dysfunction.

1.1.2.3 Mesenchymal stem cells

Despite the phenotypic variability of KS patients, craniofacial abnormalities, short stature and joint laxity are the most frequent symptoms in patients. At the cellular level, all these tissues belong from a common precursor cell, the mesenchymal stem cell (MSCs).

MSCs are precursors of connective tissue cells and can be isolated from many adult organs. Alexander Friedenstein was the first to isolate fibroblastic cells with the capacity to differentiate into osteocytes, to establish colonies from a single cell and to generate multiple skeletal tissues in vivo and in vitro (bone, fat, cartilage, muscle, tendon, ligaments, and fibroblasts). Bone-marrow derived MSCs are additionally able to provide the stromal support system for hematopoietic stem cells. MSCs have been isolated and characterized from many human sources, including adipose tissue, skeletal muscle, umbilical cord blood and Wharton's jelly[19].

Currently, bone marrow aspirate is considered to be the most accessible and enriched source of MSCs. MSCs isolated directly from bone marrow are positive for CD34, the hallmark antigen for positive immunoselection of the hematopoietic stem cell (but lose this antigen upon in vitro culture) and negative for CD50, which is common only to the hematopoietic stem cell. Isolation and enrichment of the MSCs population has been greatly facilitated by the identification of Stro-1 as a marker of a population of bone-marrow-derived cells capable of

differentiating into multiple mesenchymal lineages, including hematopoiesis supportive stromal cells, adipocytes, osteoblasts, and chondrocytes. Other markers specific of MSCs have been identified as surface proteins present only in undifferentiated cells and lost during differentiation. These markers are endoglin CD105, the receptor for TGF-β3, and CD73, the ecto-5′-nucleotidase[20]. As mentioned before, MSCs are able to differentiate into multiple cell lineages both *in vitro* and *in vivo.* Osteocytes and chondrocytes are the two main components of skeletal tissue (one of the most affected in KS). Skeletal tissues have three distinct cellular origins and formed through two different modes of action. In the head, neural crest cells originating from the dorsal neural tube form the facial bones and the cranium. Parietal bone, base of the skull, and the axial vertebral skeleton and associated ribs originate from the paraxial, presomitic, and somitic mesoderm. Then, the lateral plate mesoderm gives rise to the skeletal elements of the limbs. Two mechanisms give rise to skeletal cells: intramembranous and endochondral ossification (Fig 4). Skull and facial bones develop through intramembranous ossification, where condensed mesenchymal cells directly differentiate into bone-forming osteoblasts. Most of the remainder of the skeleton develops through endochondral ossification, in which an initial cartilage model is replaced by mineralized bone tissue with the exception of the joints that are capped by articular cartilage[21].

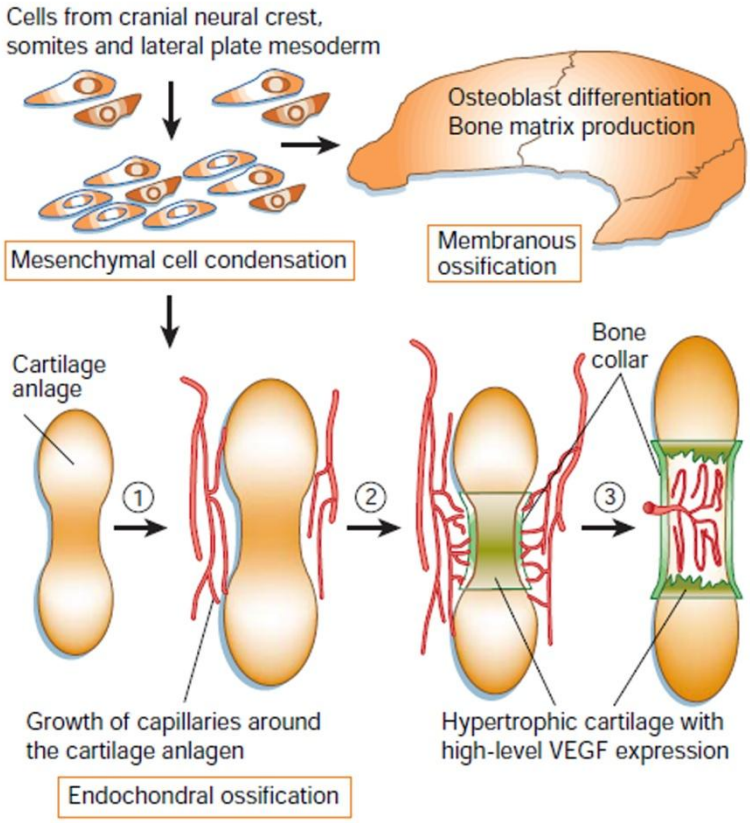


Figure 4. Skeletal tissues development. In intramembranous ossification, differentiation of mesenchymal cells to osteoblasts and production of bone matrix occurs directly, whereas in endochondral ossification, differentiation to chondrocytes and formation of cartilage models (anlagen) occurs first, followed by replacement of the models by bone. *The genetic basis for skeletal diseases. Nature 2003.*

1.1.2.3.1 Chondrogenesis

The process of endochondral ossification starts with the condensation of mesenchymal cells, which differentiate into chondrocytes. These cells initially proliferate and differentiate into two subpopulations of proliferating chondrocytes: round, low proliferating chondrocytes at

the distal ends of the condensation (RP chondrocytes) and high proliferating chondrocytes that are organized in columns (CP chondrocytes) towards the center. In the center of the growing cartilage anlagen, chondrocytes exit the cell cycle and differentiate into prehypertrophic and hypertrophic chondrocytes. These produce a mineralized extracellular cartilage matrix that serves as a template for the subsequent deposition of bone matrix. The cartilage elements are surrounded by a layer of fibroblast- like cells, the perichondrium, that will differentiate into osteoblasts forming the periost, which is highly vascularized. Blood vessels from the periost invade the calcified matrix of terminal hypertrophic chondrocytes, which undergo apoptosis. During postnatal development, secondary ossification centers are formed within the zone of the RP chondrocytes. Between the two ossification centers chondrocytes remain forming the postnatal growth plate. During embryonic and postnatal development, the balance between chondrocyte proliferation and differentiation has to be tightly controlled to ensure normal longitudinal bone growth and endochondral bone formation.

As mentioned before, the first step of chondrogenesis is the mesenchymal condensation. At the molecular level, the transcription factor Sox9 is the earliest marker of condensing mesenchymal stem cells that contributes to both chondrocytes and osteoblasts, but later Sox9 expression is restricted to the chondrocyte lineage. The expression of Sox9 starts in chondroprogenitor cells and is maintained in proliferating chondrocytes. In chimeric mice carrying Sox9 mutant cells in a wild type background, Sox9- deficient cells are excluded from the cartilage condensations indicating that SOX9 function is

required to establish the chondrogenic lineage. Moreover, in humans, mutations in *SOX9* lead to campomelic dysplasia syndrome, which is characterized by severe malformations of the endochondral skeleton. In mice, heterozygous loss of *Sox9* results in hypoplastic skeletal elements, while homozygous deletion completely inhibits chondrogenesis[22]. Sox9 is a high mobility group domain containing transcription factor. Sox9 acts in conjunction with two other Sox family members, Sox5 and Sox6, which serve as cofactors for Sox9 dependent transcription. Their expression is dependent on Sox9, and their expression domain is similar to that of Sox9 in the developing limb. Prior to recent genome-scale ChIP-seq analyses, various approaches revealed Sox9- associated enhancers around Sox9 itself plus several genes encoding cartilage growth factors or extracellular matrix proteins[21]. Three groups recently reported ChIP-seq analyses of Sox9–DNA interactions[23,24,25]. As expected, all of these studies documented a strong association of Sox9 binding around the enhancers of genes that are involved in skeletal development and chondrocyte differentiation. These Sox9 targets have been termed Class II targets to distinguish them from non- enhancer mediated interactions (Class I targets). These Class I Sox9 targets show a significant interaction around the TSSs of any gene that was highly expressed within the chondrocyte. Sox9 association was correlated with the level of that gene's expression and likely reflected protein–protein interactions between Sox9 and components of the basal transcriptional complex, such as p300[21].

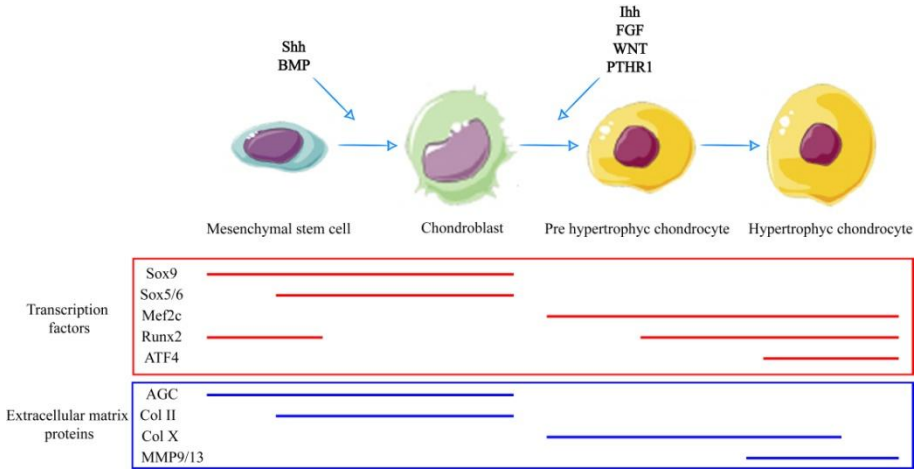


Figure 5. Molecular aspects of chondrogenesis. Expression pattern of chondrogenic transcription factors and extracellular matrix proteins during differentiation of MSCs into chondrocytes. Modified from: *Role and regulation of runx2 in osteogenesis. European Cells and Material, Vol. 28, 2014.*

During embryogenesis, the Sonic Hedgeog pathway is fundamental for chondrogenesis. Indeed, Shh directly induce the expression of Sox9. In addition, bone morphogenetic protein (BMP) signals maintain the expression of this gene after their initial induction by Shh. Sox9 directly activates chondrocyte differentiation markers and induces the expression of Sox5 and Sox6, which work together with Sox9 to activate the chondrocyte differentiation program. Sox9 expression is also necessary to sustain chondrocyte survival through a PI3K-AKT pathway and has been noted to block chondrocyte maturation into hypertrophic cells. This activated transcriptional program lead to the expression of collagen type II and aggrecan,

secreted matrix proteins that are markers of immature proliferative chondrocytes[26]. Other two transcription factors regulate the differentiation of proliferative chondrocyte into pre-hypertrophic chondrocyte. Mef2c and Runx2 are both expressed in pre-hypertrophic and hypertrophic chondrocyte. In particular, Mef2c functions upstream of Runx2, and is necessary to either induce or maintain Runx2 expression. Several studies demonstrate that Runx2 can directly bind to the regulatory sequences that drive the expression of Ihh, Vegfa, Col10a1 and Mmp13, typical markers of hypertrophic chondrocytes. The activation of canonical Wnt signaling in chicken chondrocytes increases Runx2 expression, which may in part explain how this signaling pathway promotes chondrocyte hypertrophy. Other signaling pathways are crucial for the induction and maintenance of hypertrophy. These pathways (FGF signaling, PTH receptor-1 signaling, Ihh signaling and the C-type natriuretic peptide) induce the expression of these transcription factors needed for chondrocytes maturation. Initiation of the terminal phase of chondrocyte hypertrophy is marked by a loss of Col10a1 expression, coupled with the induction of VEGFA and the metalloproteinases, Mmp9 and Mmp13. Signaling via the EGF receptor (Egfr) plays a crucial role in orchestrating this transition[27]. For many years it has been appreciated that apoptosis is detectable at the chondro-osseous front in the growth plate, suggesting that initiation of an apoptotic program of cell death is the ultimate fate of late hypertrophic chondrocytes. However, by employing mice in which Cre recombinase has been knocked into one allele of the Col10a1 locus, Cheah and colleagues demonstrated that a considerable number of hypertrophic chondrocytes that previously

expressed Col10a1 do indeed become osteocytes located in the trabeculae of bones[28].

1.1.2.3.2 Osteoblastogenesis

The differentiation process of osteoblasts is often divided into stages of mesenchymal progenitors, pre-osteoblasts and osteoblasts. Osteoblasts are characterized by the expression of osteocalcin and the transcription factor Runx2. A subset of osteoblasts can become osteocytes upon being entombed within the bone matrix. The rest of the osteoblasts are thought to either undergo apoptosis or become inactive bone-lining cells. As mentioned before, during vertebrate embryogenesis, osteoblasts are generated from mesenchymal cells through two distinct processes: intramembranous or endochondral ossification. Osteoblasts produced by either process are thought to be similar but not identical. During intramembranous ossification, mesenchymal progenitors condense and directly differentiate into osteoblasts. On the other hand, during endochondral ossification, mesenchymal progenitors condense to form chondrocytes and perichondrial cells (that can be defined as pre-osteoblasts). Only when chondrocyte become hypertrophic, perichondrial cells start to differentiate into osteoblasts. Hypertrophic cartilage is invaded by blood vessels. This process is fundamental both because blood vessels form the nascent bone marrow cavity (where osteoblasts differentiate) and because the invading blood vessels seem to provide a vehicle for the perichondrial cells to enter the nascent bone marrow.

As mentioned before, Sox9 is a key transcription factor for the initial mesenchymal stem cells condensation. Sox9 is expressed in “chondro-

osteo progenitors" but then it is no more expressed in mature osteoblasts. Sox9 is needed for the expression of another transcription factors, Runx2, the master regulator of osteoblastogenesis. During development, Runx2 is expressed within the osteo-chondro progenitors. Following the formation of cartilage, Runx2 becomes more restricted to the perichondrial cells and osteoblasts, but it is re-expressed in the early hypertrophic chondrocytes. Runx2 is not only indispensable for osteoblasts induction but also for the proper function of mature osteoblasts, including the synthesis of bone matrix[29]. Homozygous deletion of Runx2 in mice resulted in a complete lack of osteoblasts, whereas haploinsufficiency of RUNX2 in humans led to cleidocranial dysplasia[30,31]. Runx2 directly induces the expression of major bone matrix protein genes (Col1A, Osteopontin, Ibsp, and osteocalcin) in osteoblast progenitors, allowing the cells to acquire the osteoblastic phenotype while keeping the osteoblastic cells in an immature stage. Runx2 expression has to be down regulated for the acquisition of phenotype of fully mature osteoblasts, which form mature bone[32]. Runx2 also induces the expression of another transcription factor, osterix (OSX), which has important roles at multiple stages within the osteoblast lineage, both during embryonic development and in postnatal life. Activating transcription factor 4 (ATF4) has an important roles in the more mature osteoblast lineage cells. It directly regulates the expression of the osteoblast-derived molecules osteocalcin (a bone matrix protein) but also ATF4 promotes efficient amino acid import to ensure proper protein synthesis by osteoblasts. Moreover, ATF4 was shown to act in chondrocytes to stimulate the expression of IHH (an inducer of osteoblast

differentiation). Indeed, IHH signaling pathway is needed for the expression of Runx2 and OSX in pre-osteoblasts[29]. Mice deficient of IHH completely lack osteoblasts within the endochondral skeleton, although the intramembranous osteoblasts form, indicating that IHH signaling is dispensable for intramembranous osteoblast differentiation[33]. Another signaling pathway important for osteoblast differentiation is the β-catenin WNT signaling. Specifically, β‑catenin is required for the progression from the Runx2+ stage to the Runx2+OSX+ stage and from Runx2+OSX+ cells to mature osteoblasts. β-catenin independent WNT signaling has also been implicated in regulating the osteoblast lineage. PKCδ induce the progression from RUNX2+ cells to RUNX2+OSX+ cells. Also Wnt5a and Wnt10 promote osteoblastogenesis inhibiting adipogenesis[29].

1.1.2.3.3 Neural crest stem cells (NCSCs) as precursor of MSCs

As mentioned before, in the head region, skeletal tissues originate from neural crest stem cells (NCSCs). NCSCs originate during vertebrate embryogenesis within the dorsal margins of the closing neural folds. Initially, they are integrated within the neuroepithelium where they are morphologically indistinguishable from the other neural epithelial cells. Upon induction by signals that come from contact-mediated tissue interactions between the neural plate and the surface ectoderm, NCSCs delaminate through an epithelial-to-mesenchymal transition and start migrate extensively to several different locations in the embryo where they contribute to a remarkably diverse array of different tissue types ranging from the peripheral nervous system (PNS) to the craniofacial skeleton. NCSCs

derivatives originate from four different segments of the neuraxis: cranial, cardiac, vagal, and trunk. The cranial NCSCs may be interesting considering KS. Indeed these cells give rise to the majority of the bone and cartilage of the head and face, as well as to nerve ganglia, smooth muscle, connective tissue and pigment cells. Neural crest cells from the dorsal margins of the closing neural tube migrate into the anterior region of the skull. Here NCSCs differentiate into MSCs[34]. The initiation of skeletogenesis starts with migration of mesenchymal cells derived from these embryonic lineages to the sites of the future bones. Here they form condensations of high cellular density. Within the condensations, the mesenchymal cells either differentiate into chondrocytes and form cartilage models of the future bones (endochondral bone formation) or differentiate into osteoblasts to form directly bone (intramembranous bone formation). Endochondral bone formation occurs in the skull base and the posterior part of the skull while intramembranous bone formation takes place in the membranous neuro- and viscerocranium[35].

1.2 KMT2D: role and function

KMT2D was identified and characterized in 1997. It encodes a 5537 amino acid histone H3 lysine 4 (H3K4)-specific methyltransferase (MLL4). It belongs from the KMT2 (MLL) family proteins, a group of enzymes characterized by the presence of a SET domain, which is responsible for the methyltransferase activity.

1.2.1 KMT2 family

The lysine methyltransferase 2 family (KMT2, also known as mixed lineage leukemia (MLL) methylates histone H3 lysine 4 (H3K4). All the family members' proteins share homology in their C-terminal SET domains, which is the histone methyltransferase domain.

The MLL family is highly conserved, having evolved from an ancient lineage of proteins present in unicellular eukaryotes[36]. Indeed, the MLL first member, Set1, was isolated in yeast Saccharomyces cerevisiae within the Set1/COMPASS protein complex. Then, three independent genes were identified as yeast Set1 homologous in Drosophila: dSet1, trithorax (trx) and trithorax-related (Trr). The dSet1 is the direct descendent of yeast Set1, while trx and Trr are more distantly related to yeast Set1[37]. The trithorax group (trxG), together with the Polycomb group (PcG), plays a central role in the regulation of gene expression throughout development. The clustered homeotic (Hox) genes in the Bithorax and Antennapedia gene complexes in Drosophila are some of the known targets of the trxG and PcG proteins. The Drosophila trx gene was discovered based on the identification of several mutations in this gene that resulted in the partial transformation of the halteres into wings[38,39].

The first mammalian homolog of Drosophila trx, the Mixed Lineage Leukemia (MLL) gene, was cloned based on the identification of its random translocations found in patients suffering from hematological malignancies. Chromosomal duplication during mammalian evolution resulted in two paralogs in each MLL Drosophila subgroup:

- trx-related *KMT2A* (MLL1 protein) and *KMT2B* (MLL2 protein);

- trr-related *KMT2C* (MLL3 protein) and *KMT2D* (MLL4 or ALR protein)
- dSET1 related KMT2F (SETD1A protein) and KMT2G (SETD1B protein)[40];
- KMT2E (MLL5 protein) is more homologous to the SET3 family and has no intrinsic methyltransferase activity[41].

The studies to date clearly demonstrate that, despite their structural similarities, the MLL family proteins in mammals have distinct and non-redundant functions in development[42].

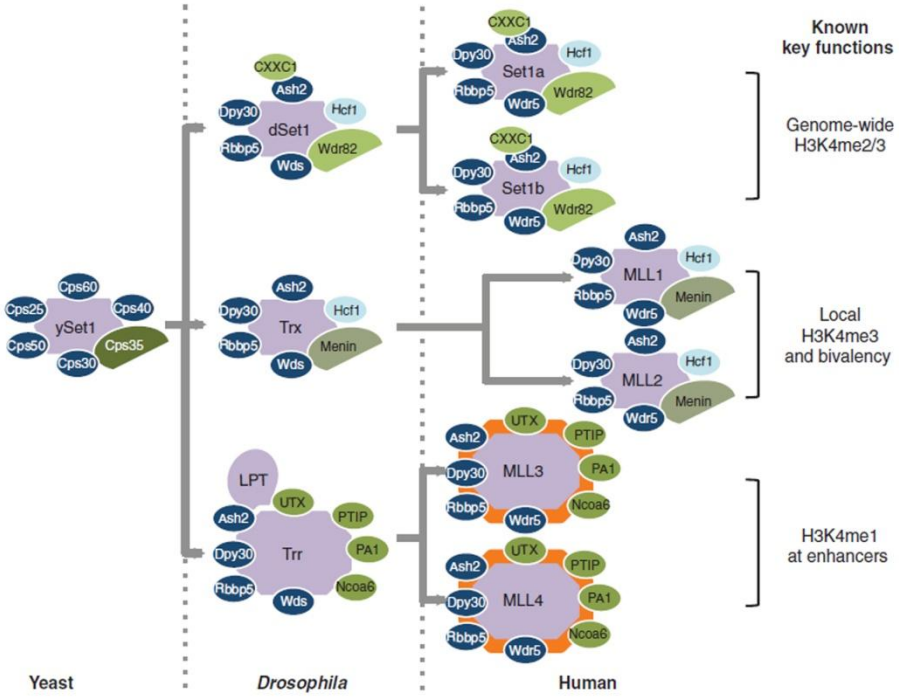


Figure 6. The COMPASS family of histone H3 lysine 4 (H3K4) methyltransferases in yeast, flies, and humans. *MLL3/MLL4/COMPASS Family on Epigenetic Regulation of Enhancer Function and Cancer. Cold Spring Harb Perspect Med 2016*

MLL proteins reside in large, multi-subunit complexes composed of unique sets of interacting proteins that is conserved from yeast to mammalian. This complex was named COMPASS for Complex of Proteins Associated with Set1. The formation of the complex is fundamental for the methyltransferase activity. Indeed, in yeast, Set1 alone is not active as a KMTase, as Set1 within COMPASS is the active form of the enzyme[43,44]. In mammalian, each KMT2 proteins can associate with different partner forming different complexes. Despite the diversity, four subunits are commonly found in all KMT2 complexes: WD repeat protein 5 (WDR5), retinoblastoma binding protein 5 (RbBP5), ASH2L and DPY30. Biochemical reconstitution shows that WDR5, RbBP5 and ASH2L form a steichiometric core entity which stably interact with the MLL enzymes and enhances their otherwise weak activities (~50–500 fold). Disruption of WDR5-MLL1 binding by the WIN peptides or their derivatives leads to disintegration of the MLL1 complex and inhibition of MLL1 catalytic activity in vitro. Apart from these 4 common subunits, MLL enzymes also associate with a unique set of proteins that confers diversity and specificity to all the different MLL family members.

Generally, H3K4 methylation is associated with transcription. Although H3K4 methylation is a universal mark for nucleosomes near the transcription start sites (TSS) of expressed genes in all eukaryotes, different studies suggest that H3K4me1, H3K4me2 and H3K4me3 have specific roles in chromatin and might be targeted by different effectors and associated modifiers. H3K4me3 is found in nucleosomes near the TSS of expressed genes in all eukaryotes. In addition,

H3K4me2 is found in nucleosomes further downstream in the body of genes, while H3K4me1 is prevalent in their 3' regions. Works on fly and mammalian genome reveal that developmental enhancers of these organisms are in general enriched for H3K4me1, while H3K4me2 or H3K4me3 are underrepresented or lacking[45].

A growing body of evidence suggests that the responsibilities of H3K4 methylation are divided among the COMPASS family members to ensure proper transcriptional modulation. Indeed, different studies have shown that SETD1A/B (the Drosophila dSet1) are responsible for bulk H3K4 di- and trimethylation across the genome[42,46]. MLL1/MLL2 (the Drosophila Trx) are necessary for gene-specific H3K4 trimethylation, including Hox gene promoters and bivalent promoters (promoters marked by concurrent trimethylation of H3K4 and H3K27 are poised to express developmental genes) in mouse embryonic stem (mES) cells[47,48]. Trr and its mammalian homologs MLL3/MLL4 have been accredited as key H3K4 mono methyltransferases at enhancers. Herz et al. and Hu et al. have shown how depletion of Trr and MLL3/MLL4 resulted in a striking genome-wide reduction of H3K4 monomethylation, primarily occurring at enhancer regions[48,49].

1.2.2 KMT2D

In human, *KMT2D* is located on chromosome 12q13.12. The transcript is 19,419 base pairs long and contains 54 exons. The gene encode for a 5537 amminoacid protein that is widely expressed in adult tissues[50]. The MLL4 protein contains two clusters of plant homeotic domains (PHDs) in the N-terminus region (three PHDs per

cluster) and an enzymatically active C-terminal SET domain. The PHDs in the second cluster (PHD4–6) recognize H4 tails on nucleosomes in vitro and could be critical for KMT2D-catalyzed nucleosome methylation[51,52]. The protein also contain other domains: there are a PHD and FY-rich N/Cterminal (FYRN and FYRC) domains, a high mobility group (HMG-I) and nine nuclear receptor interacting motifs (LXXLLs)[36].

KMT2D is part of the protein complex called ASCOM complex. The proteins WDR5, RbBP5, ASH2L, and DPY30 (the common proteins among all the MLL family members) form the four subunit sub-complex WRAD, which is critical for H3K4 methyltransferase activity. Then, a unique set of proteins was found associated to MLL4: PTIP (Pax transactivation domain-interacting protein),PA1, NCOA6, UTX (KDM6A, a H3K27 demethylases)[53]. Interestingly, absence of the MLL4 protein results in the collapse of the MLL4 complex and the destabilization of UTX in cells. Indeed, Lee et al. has shown that Kmt2d KO led to destruction of MLL4 complex. Moreover, Jang et al. demonstrate that the methyltransferase activity is fundamental for the stability of MLL4 and for the stability of the complex. Indeed, enzyme-dead Mll4 is unstable in ES cells, since the protein level decrease in mutated cells. Interestingly, also the protein level of KDM6A was found reduced in these cells, reinforcing the role of functioning MLL4 in the complex stability [54,55].

As mentioned before, MLL4 is associated with monomethylation of enhancer regions. Enhancers are genomic regulatory elements, often

bound by cell type specific transcription factors (TFs) that are essential for cell-type specific gene expression in eukaryotic cells. Typically, enhancers are identified as region rich in H3K4me1, H3K27ac and are characterized by a hypersensitivity to DNA nucleases. The presence of H3K4me1 precedes the enhancer activation, thus participating in enhancer priming. Indeed H3K27ac is acquired in the context of pre-existing H3K4me1[56]. In stem cells, primed enhancer can be present in a poised state that is also characterized by the H3K27me3 mark. The switch from poised (H3K4me1/H3K27me3) to active (H3K4me1/H3K27ac) enhancer is mediated by tissue-specific TFs binding and the subsequent recruitment of chromatin modifiers. Among them, the ASCOM complex has been show to play a key role in modulating both H3K4 mono methylation and H3K27 de-methylation. Indeed, MLL4 introduces the mono-methyl group on Lysine 4 (H3K4me1), thus priming enhancers and, within the same complex, UTX participates in modulating enhancer transition by removing methyl group from K27, thus facilitating its acetylation (Fig 7)[57].

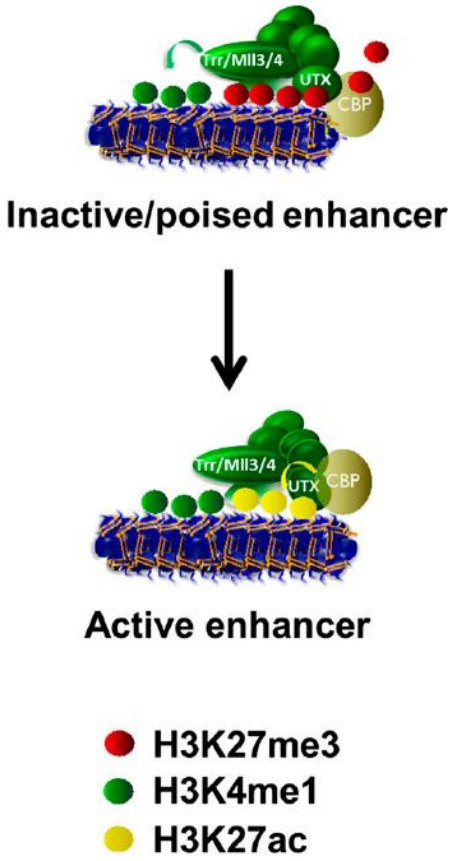


Figure 7. Model describing the role of KMT2C/D COMPASS-like complexes in the transitioning of enhancer activation. *Enhancer-associated H3K4 monomethylation by Trithorax-related, the Drosophila homolog of mammalian Mll3/Mll4. Genes and Dev. 2012*

MLL4's function as a major enhancer regulator in mammalian cells translates to various biological processes, including regulation of development and differentiation. In the work of Lee et al., Kmt2D whole-body knockout (KO) mice exhibit embryonic lethality around E9.5[55]. On the other hand, Wang et al. demonstrate how KMT2D is dispensable for maintaining ESCs and somatic cell identity. Indeed, using different cellular model (ESCs, iPSCs, fibroblasts and adipocytes) they shown that KMT2D is dispensable for the maintenance of cell identity but it is essential for the induction of differentiation. KMT2D has a dichotomous role in enhancer regulation: it is dispensable for maintaining active enhancers but it is

essential for activating enhancers of cell-identity genes. In particular, it is dispensable for maintaining binding of p300 on active enhancers of cell identity genes but is required for p300 binding on enhancers activated during cell differentiation. Moreover, with the depletion of MLL4 in ESCs the authors observed a markedly decrease of H3K27ac, eRNA, and enhancer–promoter interaction, suggesting that KMT2D is required for establishing the looping between active enhancers and the corresponding promoters[58]. The role of Kmt2D during differentiation was described for adipogenesis, myogenesis, B cells and hearth development, neurons and osteoblasts differentiation[55,17,59,60,61].

In their work, Dorighi et al. separate the enzymatic activity of MLL4 from its effect on transcription. Indeed, they observed that loss of MLL4 enzymatic activity, and consequently the loss of H3K4me1, leads to a partial depletion of H3K27ac from enhancers, but has only a minor effect on transcription. In contrast, loss of MLL4 protein strongly reduces transcription activity, at both enhancers and target genes. Interestingly, MLL4-dependent loss of RNA Polymerase II (Pol II) and nascent transcription from enhancers is not associated with changes in Pol II occupancy at nearby promoters, but instead impacts Pol II density across gene bodies[51].

Recently, it has been described a role of MLL4 (and MLL3) also in genomic stability. MLL3- and MLL4 dependent H3K4 methylation was found to be induced at replication forks upon replication stress[62]. Different works show how MLL4 physically interacts with the helicase RECQL5 to prevent collisions between transcription and replication machineries[63,64]. MLL4-deficient cells exhibited

transcription stress associated with RNAPII pausing, stalling, and backtracking, leading to chromosomal aberrations primarily within ERFSs[63]. These data indicate that MLL4 (and MLL3) could contribute to cancer development patenting DNA breaks and chromosomal rearrangements preferentially in early replicating regions.

In addition to KS, KMT2D mutations have been found in different type of cancer (medulloblastoma, pheochromocytoma, non-Hodgkin lymphoma, diffuse large B-cell lymphoma, pancreatic and prostate cancer). Frameshift and non-sense mutation are the mutations found in KMT2D related cancers and it occur mostly in the SET and PHD domains. The mechanisms of action for these types of cancer are unknown. One possibility could be that the mutations result in defective enhancer regulation and altered cell-type specific gene expression. Akhtar-Zaidi et al. described a different enrichment of H3K4me1 between colon cancer and normal crypts[65]. Also genomic instability could explain why KMT2D mutations are so common in these cancers. Indeed, KMT2D deficiency could results in transcription stress that could cause major transcriptional alterations and DNA breaks and lead to formation of cancerous tumors[66].

1.3 KDM6A: role and functions

KDM6A is the second gene found mutated in KS patients. It is located on chromosome X and encode for a 1453 amminoacid protein. It catalyzes the removal of methyl group from lysine 27 of histone 3.

In 2007, several groups identified Ubiquitously Transcribed Tetratricopeptide Repeat on chromosome X (UTX or KDM6A) and Jumonji D3 (JMJD3 or KDM6B) as novel histone demethylases that catalyze the removal of di- and trimethyl groups on histone H3 lysine 27. Ubiquitously Transcribed Tetratricopeptide Repeat on chromosome Y (UTY) is a closely related homolog of UTX on the Y-chromosome but no enzymatic H3K27me2/3 demethylase activity has been reported for UTY[15].

KDM6A, *UTY* and *KDM6B* are evolutionary conserved from Caenorhabditis elegans (C. elegans) to mammals. Indeed, mammal genomes contains the three H3K27me2/3 demethylase family members whereas the genome of Drosophila melanogaster harbors only one ortholog of the mammalian H3K27me2/3 demethylases, called dUTX[67].

KDM6A and *KDM6B* preferentially demethylate H3K27me3 followed by H3K27me2 in vitro and in vivo. This demethylase activity is dependent on the catalytic JmjC domain[68]. Moreover, a newly identified zinc-binding domain within these H3K27me2/3 erasers provides specificity toward the histone lysine H3K27 and excludes interaction with the near-cognate histone lysine H3K9, whose demethylation is catalyzed by LSD1[69].

Methylation of H3K27 is a critical mediator of transcriptional gene repression and contributes to important biological processes including X-inactivation, genomic imprinting, stem cell maintenance, animal body patterning, circadian rhythms and cancer[70]. Regulation of cellular H3K27me3 levels is mainly mediated by the H3K27

methyltransferase Polycomb Repressor Complex 2 (PRC2) and the H3K27me2/3 demethylases KDM6A and KDM6B. One of the most studied KDM6A targets are the HOX genes. The HOX genes are highly conserved transcriptional regulators essential during development of the anterior-posterior axis. Aberrant expression of HOX genes leads to developmental defects and disease. HOX expression is tightly regulated by opposing activities of the H3K27me3 writer PRC2 and the H3K27me2/3 eraser KDM6A in order to enable context dependent transcriptional regulation[71]. H3K27 methylation is important in the maintenance of self-renewing ESCs by repressing tissue-specific developmental genes. During ESC differentiation, the re-activation of developmental genes is associated with loss of H3K27me3. Loss of KDM6A, like MLL4, does not influence ESC self-renewal and proliferation, but seems to provoke an effect on the differentiation capacity of ESCs[72]. Contradictory studies reported on whether the effect of KDM6A loss on ESC differentiation depends on its H3K27me3 demethylase activity[73,74].

As *KMT2D*, also *KDM6A* is involved in different cancers, including multiple myeloma, esophageal and renal cancer. In particular, bladder cancer has one of the highest frequencies of KDM6A mutations[75]. Inactivating mutations of KDM6A are found in most cancer types, suggesting that it acts as a tumor suppressor, possibly by regulating the expression of cell cycle regulators, such as retinoblastoma-binding proteins[76].

1.4 Scope of the thesis

Kabuki syndrome is an autosomal dominant condition caused by heterozygous loss-of-function mutations in either of two genes: lysine-specific methyltransferase 2D (*KMT2D*) on human chromosome 12 or lysine-specific demethylase 6A (*KDM6A*) on human chromosome X.

Nothing or very little is known about the pathogenesis of the disease. First of all, KS is a complex disease in which different body apparatus are affected but there is not a particular loss of function for any of them. This make the understanding of the pathology even more complicated and suggest that the mutation occurs early during embryogenesis affecting the normal development of the whole embryo. The types of mutations complicate the picture of the disease. In most cases the mutation lead to a loss of function of the protein but there are also other kind of mutations that may not cause the lost of the activity of the protein, suggesting that the disease could not be associated to a simple loss of function disease.

The most frequent gene found mutated in KS patients is *KMT2D*. The precise role of MLL4 protein has not been yet elucidated. It is known that complete KO of *Kmt2D* is not compatible with life but it has been demonstrated that it is dispensable for self renewal maintaining of embryonic stem cells. Indeed MLL4 is required for the establishment of lineage-specific enhancers thus regulating differentiation. Finally, there is also a lot of debate concerning the H3K4me1 function. It is well accepted that this histone modification marks active or poised enhancers. However very little is known about who drives the deposition of monomethyl and about the “readers” of this histone

marks. Moreover, it is not known if the H3K4me1 serves solely as a marker of active enhancers or if it has a specific and functional role.

Giving this premise, the scope of this thesis is the development of an *in vitro* disease model to dissecting the molecular mechanisms behind the Kabuki syndrome. We decide to focus our attention on *KMT2D*, the gene found mutated in the majority of cases. We choose as experimental model the mesenchymal stem cells. This is because, as mentioned before, MSCs give rise to chondrocyte and bone tissues, two of the most affected tissues in disease. Through a CRISPR/Cas genome editing approaches, we want to reproduce the mutations of Kabuki syndrome patients in MSCs. These cells may be used as a model to understand the molecular mechanisms of the disease. Understanding the molecular aspect of KS is fundamental in order to design and develop therapeutic approaches. In this view, the *in vitro* model would be the first step for the development of drug screening approaches aimed to find a treatment that could rescue, at least in part, the phenotype. Furthermore, this model could also be useful to deepen the molecular functions of *KMT2D*.

In the results chapter it is described the approach that we used to develop the in vitro model of the disease together with the obtained results. In the last chapter it was developed a discussion about the obtained results and the future perspectives.

Bibliography

1. Bogershausen N, W. B. Unmasking Kabuki syndrome. *Clin Genet* 201–211 (2013). doi:10.1111/cge.12051
2. Niikawa, N. *et al.* Kabuki make-up (Niikawa-Kuroki) syndrome: a study of 62 patients. *Am. J. Med. Genet.* **31,** 565–89 (1988).
3. Matsumoto, N. & Niikawa, N. Kabuki make-up syndrome: A review. *Am. J. Med. Genet.* **117C,** 57–65 (2003).
4. Adam, M. P. & Hudgins, L. Kabuki syndrome: A review. *Clin. Genet.* **67,** 209–219 (2005).
5. Ho, H. H. & Eaves, L. C. Kabuki make-up (Niikawa–Kuroki) syndrome: cognitive abilities and autistic features.pdf. *Dev. Med. Child Neurol.* 487–490 (1997).
6. Schrander-Stumpel, C. T. R. M., Spruyt, L., Curfs, L. M. G., Defloor, T. & Schrander, J. J. P. Kabuki syndrome: Clinical data in 20 patients, literature review, and further guidelines for preventive management. *Am. J. Med. Genet.* **132 A,** 234–243 (2005).
7. Paulussen, A. D. C. *et al.* MLL2 mutation spectrum in 45 patients with Kabuki syndrome. *Hum. Mutat.* **32,** 2018–2025 (2011).
8. Ito, N. *et al.* Hypothalamic pituitary complications in Kabuki syndrome. *Pituitary* **16,** 133–138 (2013).
9. Ng, S. B. *et al.* Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat. Genet.* **42,** 790–793 (2010).
10. Li, Y. *et al.* A mutation screen in patients with Kabuki syndrome. *Hum. Genet.* **130,** 715–724 (2011).
11. Crespin, M. *et al.* Deletion of KDM6A , a Histone Demethylase Interacting with MLL2 , in Three Patients with Kabuki Syndrome. 119–124 (2012). doi:10.1016/j.ajhg.2011.11.021
12. Banka, S. *et al.* Novel KDM6A (UTX) mutations and a clinical and molecular review of the X-linked Kabuki syndrome (KS2). *Clin. Genet.* **87,** 252–258 (2015).
13. Cheon, C. K. *et al.* Identification of KMT2D and KDM6A mutations by exome sequencing in Korean patients with Kabuki syndrome. *J. Hum.*

Genet. **59,** 321–325 (2014).

14. Miyake, N. *et al.* KDM6A Point Mutations Cause Kabuki Syndrome. *Hum. Mutat.* **34,** 108–110 (2013).
15. Agger, K. *et al.* UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **449,** (2007).
16. Prasad, R. *et al.* Structure and expression pattern of human ALR, a novel gene with strong homology to ALL-1 involved in acute leukemia and to Drosophila trithorax. *Oncogene* **15,** 549–60 (1997).
17. Ang, S. *et al.* KMT2D regulates specific programs in heart development via histone H3 lysine 4 di-methylation. 810–821 (2016). doi:10.1242/dev.132688
18. Quintana, A. M. *et al.* Kabuki syndrome genes KMT2D and KDM6A : Functional analyses demonstrate critical roles in craniofacial , heart and brain development analyses demonstrate critical roles in craniofacial , heart and brain development. (2015). doi:10.1093/hmg/ddv180
19. Barry, F. & Murphy, M. Mesenchymal stem cells in joint disease and repair. *Nat. Rev. Rheumatol.* **9,** 584–594 (2013).
20. Tuan, R. S., Boland, G. & Tuli, R. Adult mesenchymal stem cells and cell-based tissue engineering. **5,** (2003).
21. Hojo, H., McMahon, A. P. & Ohba, S. An Emerging Regulatory Landscape for Skeletal Development. *Trends Genet.* **32,** 774–787 (2016).
22. Wuelling, M. & Vortkamp, A. Chondrocyte Proliferation and Differentiation. *Cartil. bone Dev. its Disord.* **21,** 1–11 (2011).
23. Oh, C. Do *et al.* SOX9 regulates multiple genes in chondrocytes, including genes encoding ECM proteins, ECM modification enzymes, receptors, and transporters. *PLoS One* **9,** (2014).
24. Ohba, S. *et al.* Distinct Transcriptional Programs Underlie Sox9 Regulation of the Mammalian Chondrocyte Article Distinct Transcriptional Programs Underlie Sox9 Regulation of the Mammalian Chondrocyte. *CellReports* **12,** 229–243 (2015).
25. Liu, C. F. & Lefebvre, V. The transcription factors SOX9 and SOX5/SOX6 cooperate genome-wide through super-enhancers to drive chondrogenesis. *Nucleic Acids Res.* **43,** 8183–8203 (2015).

26. U, B. D. C. *et al.* Transcriptional mechanisms of chondrocyte differentiation ☙. 389–394 (2000).
27. Kozhemyakina, E., Lassar, A. B. & Zelzer, E. A pathway to bone : signaling molecules and transcription factors involved in chondrocyte development and maturation. 817–831 (2015). doi:10.1242/dev.105536
28. Yang, L., Tsang, K. Y., Tang, H. C., Chan, D. & Cheah, K. S. E. Hypertrophic chondrocytes can become osteoblasts and osteocytes in endochondral bone formation. *PNAS* **111,** 13 (2014).
29. Long, F. Building strong bones: molecular regulation of the osteoblast lineage. *Nat. Rev. Mol. Cell Biol.* **13,** 27–38 (2011).
30. Komori, T. *et al.* Targeted Disruption of Cbfa1 Results in a Complete Lack of Bone Formation owing to Maturational Arrest of Osteoblasts. *Cell* **89,** 755–764 (2017).
31. Mundlos, S. *et al.* Mutations involving the trascription factor CBFA1 caused cleidocranial dysplasia. *Cell* **89,** 773–779 (1997).
32. Komori, T. Regulation of bone development and extracellular matrix protein genes by RUNX2. *Cell Tissue Res.* **339,** 189–195 (2010).
33. St-Jacques, B., Hammerschmidt, M. & McMahon, A. P. Indian hedgehog signaling regulates proliferation and differentiation of chondrocytes and is essential for bone formation. *Genes Dev.* **13,** 2072–2086 (1999).
34. Achilleos, A. & Trainor, P. A. Neural crest stem cells : discovery , properties and potential for therapy. *Cell Res.* **22,** 288–304 (2012).
35. Berendsen, A. D. & Olsen, B. R. Bone development. *Bone* **80,** 14–18 (2016).
36. Rao, R. C. & Dou, Y. Hijacked in cancer: the MLL/KMT2 family of methyltransferases. *Nat Rev Cancer* **15,** 334–346 (2016).
37. Mohan, M., Lin, C., Guest, E. & Shilatifard, A. Licensed to elongate: a molecular mechanism for MLL-based leukaemogenesis. *Nat. Rev. Cancer* **10,** 721–728 (2010).
38. Ringrose, L. & Paro, R. Epigenetic Regulation of Cellular Memory by the Polycomb and Trithorax Group Proteins. *Annu. Rev. Genet.* **38,** 413–443 (2004).

39. Ingham, P. & Whittle, R. Trithorax: A new homoeotic mutation of Drosophila melanogaster causing transformations of abdominal and thoracic imaginal segments - I. Putative role during embryogenesis. *MGG Mol. Gen. Genet.* **179,** 607–614 (1980).
40. FitzGerald, K. T. & Diaz, M. O. MLL2: A new mammalian member of the trx/MLL family of genes. *Genomics* **59,** 187–192 (1999).
41. Schaft, D. *et al.* The S. cerevisiae SET3 complex includes two histone deacetylases, Hos2 and Hst1, and is a meiotic-specific repressor of the sporulation gene program. *Genes Dev.* **1,** 2991–3004 (2001).
42. Min Wu, *et al.* Molecular Regulation of H3K4 Trimethylation by Wdr82, a Component of Human Set1/COMPASS. *Mol. Cell. Biol.* **28,** 7337–7344 (2008).
43. Krogan, N. J. *et al.* COMPASS, a histone H3 (lysine 4) methyltransferase required for telomeric silencing of gene expression. *J. Biol. Chem.* **277,** 10753–10755 (2002).
44. Roguev, A. *et al.* The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *EMBO J.* **20,** 7137–7148 (2001).
45. Kusch, T. Histone H3 lysine 4 methylation revisited. **3,** 310–314 (2012).
46. Hallson, G. *et al.* dSet1 is the main H3K4 Di- and tri-methyltransferase throughout Drosophila development. *Genetics* **190,** 91–100 (2012).
47. Wang, P. *et al.* Global Analysis of H3K4 Methylation Defines MLL Family Member Targets and Points to a Role for MLL1-Mediated H3K4 Methylation in the Regulation of Transcriptional Initiation by RNA Polymerase II. *Mol. Cell. Biol.* **29,** 6074–6085 (2009).
48. Hu, D. *et al.* The MLL3 / MLL4 Branches of the COMPASS Family Function as Major Histone H3K4 Monomethylases at Enhancers. **33,** 4745–4754 (2013).
49. Herz, H. *et al.* Enhancer-associated H3K4 monomethylation by Trithorax-related , the Drosophila homolog of mammalian Mll3 / Mll4. **2,** 2604–2620 (2012).
50. Froimchuk, E., Jang, Y. & Ge, K. Histone H3 lysine 4 methyltransferase KMT2D. *Gene* **627,** 337–342 (2017).

51. Dorighi, K. M. *et al.* Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from Promoters Independently of H3K4 Monomethylation. *Mol. Cell* **66,** 568–576.e4 (2017).
52. Ruthenburg, A. J., Allis, C. D. & Wysocka, J. Methylation of Lysine 4 on Histone H3: Intricacy of Writing and Reading a Single Epigenetic Mark. *Mol. Cell* **25,** 15–30 (2007).
53. Cho, Y. W. *et al.* PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. *J. Biol. Chem.* **282,** 20395–20406 (2007).
54. Jang, Y., Wang, C., Zhuang, L., Liu, C. & Ge, K. H3K4 Methyltransferase Activity Is Required for MLL4 Protein Stability. *J. Mol. Biol.* **429,** 2046–2054 (2017).
55. Lee, J. *et al.* H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. *Elife* **2,** 1–25 (2013).
56. Creyghton, M. P. *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci.* **107,** 21931–21936 (2010).
57. Calo, E. & Wysocka, J. Modification of Enhancer Chromatin: What,How,andWhy. *Mol. Cell* **49,** 825–837 (2013).
58. Wang, C. *et al.* Enhancer priming by H3K4 methyltransferase MLL4 controls cell fate transition. 1–6 (2016). doi:10.1073/pnas.1606857113
59. Zhang, J. *et al.* Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. *Nat. Med.* 1–13 (2015). doi:10.1038/nm.3940
60. Dhar, S. S. *et al.* Trans-tail regulation of MLL4-catalyzed H3K4 methylation by H4R3 symmetric dimethylation is mediated by a tandem PHD of MLL4. *Genes Dev.* **26,** 2749–2762 (2012).
61. Munehira, Y., Yang, Z. & Gozani, O. Systematic Analysis of Known and Candidate Lysine Demethylases in the Regulation of Myoblast Differentiation. *J. Mol. Biol.* **429,** 2055–2065 (2017).
62. Chaudhuri, A. R. *et al.* Replication fork stability confers chemoresistance in BRCA-deficient cells. *Nat. Publ. Gr.* (2016). doi:10.1038/nature18325
63. Kantidakis, T. *et al.* Mutation of cancer driver MLL2 results in transcription

stress and genome instability. **4,** 408–420 (2016).

64. Saponaro, M. *et al.* RECQL5 Controls Transcript Elongation and Suppresses Genome Instability Associated with Transcription Stress. *Cell* 1037–1049 (2014). doi:10.1016/j.cell.2014.03.048
65. Akhtar-zaidi, B. *et al.* Epigenomic Enhancer Profiling Defines a Signature of Colon Cancer. *Science (80-.).* **336,** 736–740 (2012).
66. Tubbs A, N. A. Endogenous DNA Damage as a Source of Genomic Instability in Cancer. *Cell* 644–656 (2017). doi:10.1016/j.cell.2017.01.002
67. Smith, E. R. *et al.* Drosophila UTX Is a Histone H3 Lys27 Demethylase That Colocalizes with the Elongating Form of RNA Polymerase II. *Mol. Cell. Biol.* **28,** 1041–1046 (2008).
68. Hong, S. *et al.* Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases. *Proc. Natl. Acad. Sci.* **104,** 18439–18444 (2007).
69. Kim, E. & Song, J. J. Diverse ways to be specific: A novel zn-binding domain confers substrate specificity to UTX/KDM6a histone H3 lys 27 demethylase. *Genes Dev.* **25,** 223–2226 (2011).
70. Lee, M. G. *et al.* Demethylation of H3K27 Regulates Polycomb Recruitment and H2A Ubiquitination. *Science (80-.).* **318,** 447–450 (2007).
71. Mallo, M. & Alonso, C. R. The regulation of Hox gene expression during animal development. *Development* **140,** 3951–3963 (2013).
72. Welstead, G. G. *et al.* X-linked H3K27me3 demethylase Utx is required for embryonic development in a sex-specific manner. *Proc. Natl. Acad. Sci.* **109,** 13004–13009 (2012).
73. Wang, C. *et al.* UTX regulates mesoderm differentiation of embryonic stem cells independent of H3K27 demethylase activity. *Proc. Natl. Acad. Sci.* **109,** 15324–15329 (2012).
74. Mansour, A. A. *et al.* The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. *Nature* **488,** 409–413 (2012).
75. van Haaften, G. *et al.* Somatic mutations of the histone H3K27 demethylase gene UTX in human cancer. *Nat. Genet.* **41,** 521–523 (2009).
76. Herz, H. M. *et al.* The H3K27me3 Demethylase dUTX Is a Suppressor of Notch- and Rb-Dependent Tumors in Drosophila. *Mol. Cell. Biol.* **30,** 2485–

2497 (2010).

2. Development of an in vitro disease model for dissecting the epigenetic mechanisms underlying pathogenesis of Kabuki syndrome

Fasciani A., Zippo A.

2.1 Results

2.1.1 Immortalized mesenchymal stem cells as experimental model for Kabuki Syndrome

Adult mesenchymal stem cells (MSCs) are among the most easy to recover type of stem cells. Bone marrow aspirate represents a widely used source for MSCs. However, even if great effort has been made to maintain primary MSCs in culture, these cells show senescence and stop to divide after several passages. This implies the need to continuously isolate and purify fresh primary MSCs during the study, adding a further level of complexity, already intrinsic in primary cells cultures that could make the results difficult to be interpreted. For these reasons, we adopted a different strategy which could permit to overcome the limited cell division capacity of MSCs, yet preserving their multipotency. In particular we used MSCs that have been immortalized with human telomerase (hTert) described in the work of Tátrai, P. *et al*[1]. At first, we asked whether these cells (thereafter named iMSCs) retained the typical characteristics of MSCs in terms of self-renewing capacity, expression of phenotypic markers and differentiation capacity. We decided to analyze through flow citometry the expression of typical specific markers used for the characterization of primary MSCs. We therefore analyzed CD90 (a

cell surface glycoprotein involved in cell adhesion and cell communication) and CD105 (a transmembrane glycoprotein, component of the transforming growth factor beta receptor complex). Figure 1a shows flow citometry analysis of CD90 and CD105 in primary MSCs and iMSCs. We noticed that iMSCs maintain the expression of CD90 and CD105. Importantly, the expression of these surface markers resulted heterogeneous among the primary mesenchymal stem cells. In contrast, the iMSCs showed more homogeneous expression of both CD90 and CD105, indicating a decrease in the cell heterogeneity that arises from primary cells. Thereafter, we tested the differentiation capacity of the iMSCs towards adipocytes, osteocytes and chondrocytes, thus determining their multipotency. Morphological and histochemical analyses show that iMSCs are able to differentiate into the three principal lineages (Fig 1b). Indeed, in both the cell line and the primary cells, Oil red O staining allowed us to observe lipid droplets typical of differentiated adipocytes (Fig. 1b, left panel). Likewise, Alizarin red revealed the calcium depositions, distinctive of osteoblast differentiation, in both the cell types (Fig. 1b, central panel). Finally, differentiation into the chondrocyte lineage (Fig. 1b, right panel) is also comparable between iMSCs and primary cells, as revealed by the morphology and the Alcian blue staining. The differentiation capacity of iMSCs was also confirmed by analyzing the expression level of lineage-specific genes, which are induced in differentiated adipocytes (Fig. 1c left panel), osteocytes (Fig. 1c center panel) and chondrocytes (Fig. 1c right panel). qRT-PCR analyses showed that immortalized cells upregulate markers of specific lineages upon stimulation, thus indicating the

activation of the specific transcriptional programs, responsible for cell identity. Of importance, the gene expression pattern retrieved in differentiated iMSCs is comparable to those obtained by differentiating primary mesenchymal stem cells (Fig 1C). In conclusion, the molecular and cellular characterizations indicated that the iMSCs maintained the major features of mesenchymal stem cells resulting as a suitable *in vitro* model for studying Kabuki Syndrome.

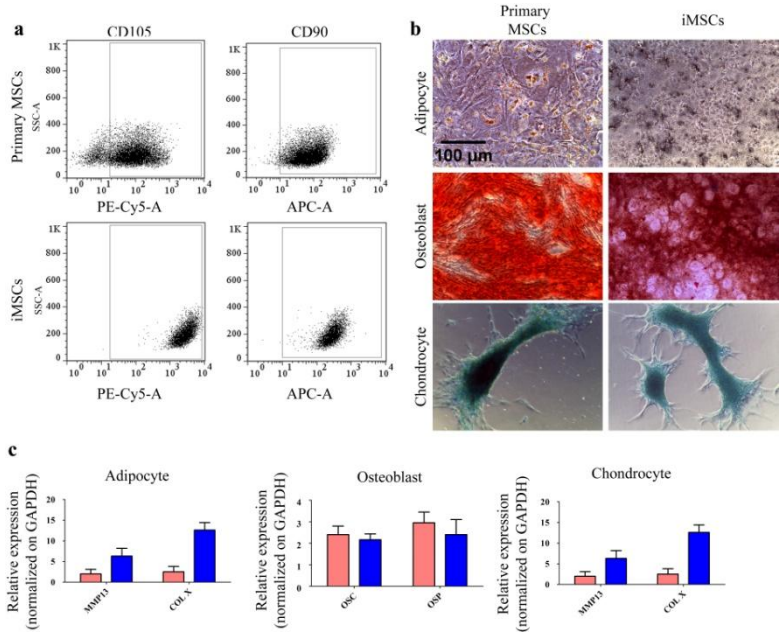


Figure 8 iMSCs preserve the same properties of primary MSCs.
Analysis of cell identity and differentiation capacity of MSCs. **a**, flow citometry analysis of MSCs surface markers CD105 (left) and CD90 (right). **b**, terminal differentiation of primary and immortalized MSCs into adipocyte (left), osteoblasts (centre) and chondrocytes (right) lineages. Adipocytes were stained with Oil red, osteoblasts with Alizarin red and chondrocyte with Alcian blue. Scale bar: 100µm. **c**, q-RTPCR analysis of markers of the three lineages at the end of differentiation. Data were normalized on expression values from undifferentiated cells and are represented as mean + SEM of three independent replicates. LPL, lipoprotein lipase; FABP4, fatty acid binding protein; OSC, osteocalcin; OSP, osteopontin; MMP13, matrix metallopeptidase; COLX, collagen type 10.

2.1.2 Genome editing approach to mimic Kabuki syndrome (KS) patients mutations

Kabuki syndrome (KS) is a monogenic disorder whose causative mutations may result in loss-of-function of KMT2D or KDM6A genes, leading to multiple abnormalities affecting different tissues. In order to mimic the acquisition of those *de novo* mutations that target KMT2D in KS patients, we developed a genome editing approach using the CRISPR/Cas9 technology. The CRISPR/Cas system is a prokaryotic immune system that confers resistance to foreign genetic elements. Invading DNA from viruses or plasmids is cut into small fragments and incorporated into a CRISPR locus. The loci are transcribed and transcripts are then processed to generate small RNAs which are used to guide effectors endonucleases that target invading DNA relying on sequence complementarity. The system has been engineered in order to cut the genomic DNA (thanks to the activity of the nuclease, the Cas9 protein) with a sequence-specific precision (thanks to the 20bp guide RNA sequence, called sgRNA). In the cell, the damaged DNA will be repaired through the endogenous DNA repair mechanisms. One of this is the non-homologous end joining (NHEJ) system. NHEJ typically utilizes short homologous DNA sequences, called microhomologies, to guide repair. Because the overhang between the microhomologies and the DNA is not always perfect, an imprecise repair can occur, leading to loss or addiction of nucleotides. In 2 out 3 cases, this repair system leads to frame-shift mutation. Typically, the frameshift mutations alter the first stop codon encountered in the sequence. The resulting protein could be abnormally short or abnormally long, and will most likely not be

functional. We decided to take advantage of this mechanism to reproduce the nonsense mutations found in KS patients.

Figure 2a shows the different kind and the distribution of all known mutations affecting KMT2D among the KS patients. We decided to reproduce nonsense mutations that are the most frequent occurring on KMT2D. Specifically, these nonsense mutations cause the insertion of a premature stop codon that result in the production of a truncated form of the protein that lacks the enzymatic domain, the SET domain. Moreover, we focused our attention on exon 39 that seems to be one of the most frequently affected sites, representing a mutational hotspot.

We planned to introduce nonsense mutation that have been identified in at least two individuals and that give rise to truncated forms of KMT2D protein. Specifically, we focused on the Q4081X mutation and we designed a CRISPR/Cas9 genome editing approach to insert it into iMSCs. Considering that we did not know the efficacy and specificity of the CRISPR/Cas9 in iMSCs, we designed four distinct sgRNAs surrounding the Q4081 amminoacid (Fig 2b). In order to reach these objectives, iMSCs have been transduced with a lentiviral vector containing a Cas9 cassette, whose expression is under the control of a doxycicline inducible promoter (iMSC-Cas9). We selected Cas9 over expressing iMSCs by treating cells with puromycin. Thereafter, iMSCs-Cas9 have been transduced with lentiviral vectors expressing each of the four designed sgRNAs. After four days of Cas9 induction, cells have been harvested and their genome has been analyzed with the surveyor assay, which permits the evaluation of cutting efficiency guided by each sgRNAs on the

selected locus. Briefly, the surveyor assay exploits the ability of T7 endonucleases to cut mismatched DNA. The NHEJ of the Cas9-induced double-strand break will result in different resolution of the Cas9 induced double-strand break obtaining a heterogeneous cell population carrying different DNA insertions or deletions nearby the cutting site. By performing surveyor assay, the occurrence of the Cas9-mediated genome editing can be visualized and quantified on agarose gel. We performed the surveyor assay on the genomic region around the Q4081 amminoacid, retrieved from iMSC-Cas9 edited with the four sgRNAs. The obtained data showed an efficiency of the sgRNAs-induced genome editing up to 30% (Fig 2c and 2d). Considering the results obtained from the four independent sgRNAs, we decided to continue our study focusing on sgRNA 1.

Since we wanted to obtain edited iMSCs carrying the specific mutation, we isolated single cell clones and performed a modified surveyor assay on their genome (Fig 2e). The CRISPR/Cas9 system can lead to mutation in one or both alleles, but the standard surveyor assay results in the exclusion of homozygous edited clones, which do not cause mismatch and are therefore associated with a "non-cutted" pattern. To avoid this, we perform surveyor analysis mixing 1:1 the fragments of edited cells with those of untreated cells, giving us the possibility to visualize both homozygous and heterozygous mutants. The obtained results showed that three out of the four analyzed clones have been edited by the CRISPR/Cas9 on the specific locus. In order to determine whether the genome editing resulted in a frame-shift of the coding sequence and to know exactly the type of mutation that we obtained, we performed DNA sequencing of each individual clones

(Fig. 2f). The sequencing allowed us to identify three different frameshift, one of which resulted in homozygosis. We decided to continue our studies using a heterozygous and homozygous clone. In particular, the heterozygous clone carried a 10 bp deletion leading to the insertion of a premature STOP codon instead of the 4092 glutamine (Q4092X) (Fig 2f, top panel). The homozygous clone has a deletion of 31 bp which causes the substitution of the glutamine 4085 with a STOP codon (Q4085X) (Fig 2f, bottom panel). Even though the mutations that we obtained are not the Q4081X mutation, they are in its close proximity, suggesting that the resulting protein have the same properties of the KMT2D Q4081X mutant. For this reason we considered that the obtained clones harbor nonsense mutations on KMT2D gene, resembling the *de novo* mutations retrieved in KS patients.

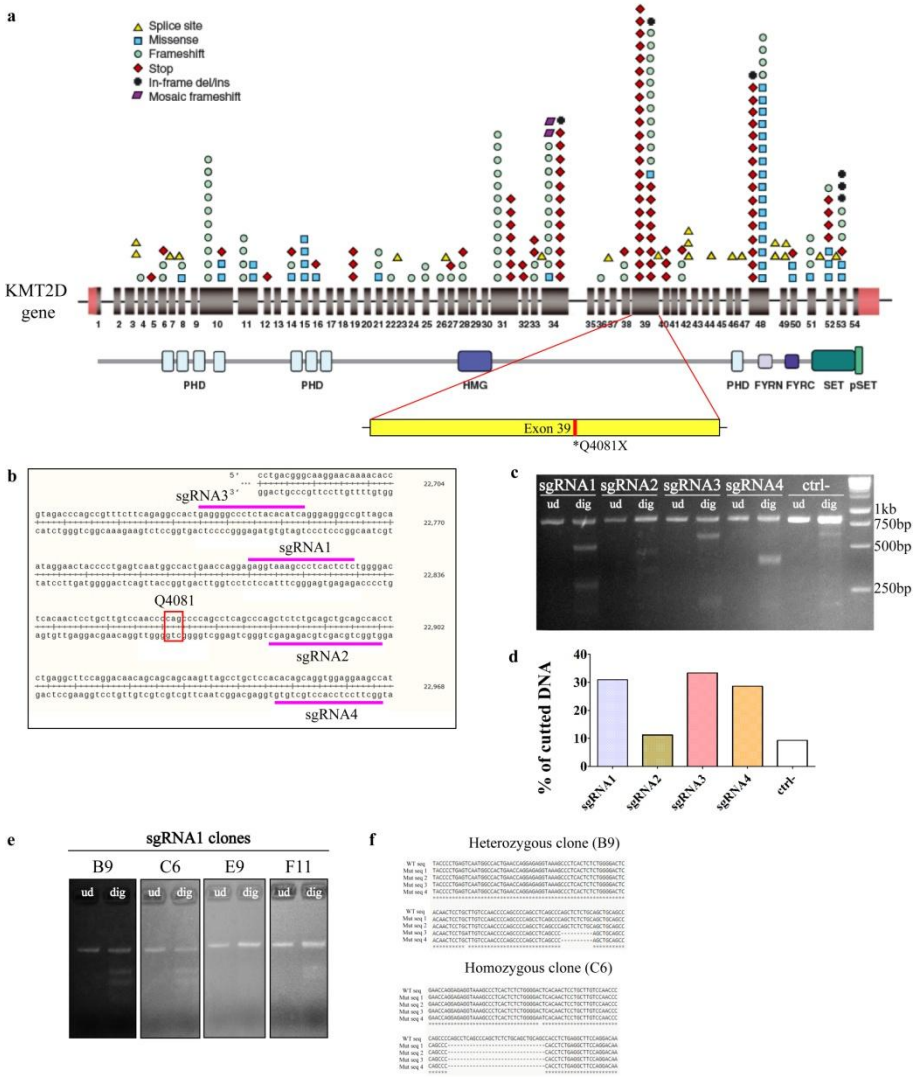


Figure 2. Genome editing approach to introduce nonsense mutations on *KMT2D* gene. **a**, schematic representation of distribution and type of known mutations found in KS patients with highlight on the Q4081X nonsense mutation on exon 39. **b**, localization of the four sgRNAs designed in order to reproduce the Q4081X mutation. Two sgRNAs (sgRNA1 and sgRNA3) are on the 5'→3' strand while the other two (sgRNA2 and sgRNA4) are on the opposite strand). **c**, genomic DNA electrophoresis resulting from the surveyor assay performed on iMSCs transduced with the four different sgRNAs. For each sgRNA two lanes are shown: the first one is the undigested (ud) PCR fragment and the second one is the digested (dig) PCR fragment. The negative control (ctrl-) is genomic DNA obtained from untransduced iMSC-Cas9. **d**, quantification of surveyor assay. The percentage was obtained dividing the intensity of cut band with the intensity of uncut band. **e**, genomic DNA electrophoresis resulting from surveyor assay of different cells clones derived from iMSCs population treated with sgRNA1. **f**, sequence alignment of WT clone (first lane of the alignment) and mutated clones.

2.1.3 Molecular and phenotypic characterization of mutated iMSCs

Through the genome editing approach we obtained two different clones carrying nonsense mutations of *KMT2D,* thus resulting in a premature STOP codon which should give rise to a truncated protein In particular, we obtained a heterozygous clone (named $KMT2D^{+/-}$) carrying the Q4092X mutation and a homozygous clone (named $KMT2D^{-/-}$) carrying the Q4085X mutation. We next aimed to molecularly and phenotypically characterize $KMT2D^{+/-}$ and $KMT2D^{-/-}$ iMSCs.

It has been previously demonstrated that KMT2D truncating mutations result in mRNA degradation through the nonsense-mediated mRNA decay[2]. Considering that, we first asked whether the introduced mutations could alter the stability and the enzymatic activity of the codified protein MLL4. For this purpose, we performed western blot analysis on protein extracts obtained from wild type (WT) and mutated iMSCs. In both mutated clones we detected a decrease of MLL4 protein level respect to WT clone (Fig. 3a), indicating that the introduced nonsense mutations affect the protein stability. Subsequently, as MLL4 is a histone methyltransferase, we asked whether its reduction could impact on the H3K4 methylation state. To this end, we enriched for bulk histones by performing acidic-based protein extraction from both control and KMT2D-mutated iMSCs. Western blot analysis revealed that, the global level of H3K4 mono-methylation (H3K4me1) was decreased respect to control cells (Fig. 3b). Of note, the tri-methylation of H3K4 (H3K4me3) resulted unaffected confirming the mono-

methyltransferase activity of MLL4. Considering the possible histone cross-talk between H3K4me1 and H3K27 acetylation (H3K27ac)[3], we also measured the level of H3K27ac on isolated histones. Interestingly, we detected a global decrease in H3K27 acetylation in iMSCs carrying *KMT2D* mutations (Fig 3b). These results demonstrate a decrease in MLL4 protein levels in mutated iMSCs, accompanied by its decreased enzymatic activity, as indicated by the reduction in the global level of H3K4me1. Furthermore, the reduction of H3K27ac levels also suggests a wider deregulation of mechanisms responsible for histone modifications deposition.

Afterwards, we asked whether the introduced mutations alter the features and properties of iMSCs. We performed cell cycle analysis to assess whether (I) there is alteration in cell cycle progression, (II) eventual genomic instability induced by MLL4 impairment, visualized as cells with higher DNA content, and (III) the level of apoptosis, visualized in the sub-G1 phase as cells with less DNA content. The proportion of cells in sub-G1 phase did not change between WT and mutated clones, suggesting that there is not an increase in apoptotic cells. Moreover, we could not detect cells with higher DNA content (>4n), indicating that MLL4 impairment do not induce polyploidy. Together these analyses showed that *KMT2D* mutations do not cause genomic alteration nor increased apoptosis (Fig. 3c). However, we observed slight differences in cell cycle phases between WT and mutated iMSCs. In particular there is an increase in G2-phase cell population (14% of WT cells versus 20% of both heterozygous and homozygous clone) and in S-phase cell population (16% of WT cells versus 19% of both heterozygous and homozygous clone) (Fig. 3c).

These data suggest a slight increase in proliferation of mutated clones. To better understand if this change affects cell growth, we performed a growth curve analysis. We did not detect any significant difference between the three cell populations, indicating that there is not a growth advantage due to MLL4 loss or haploinsufficiency (Fig. 3d).

We next analyzed cellular morphology. Mutated iMSCs showed a slight different shape when compared to WT iMSCs. In particular, they appear more rounded and with less cytoplasmatic protrusion (Fig 3e).

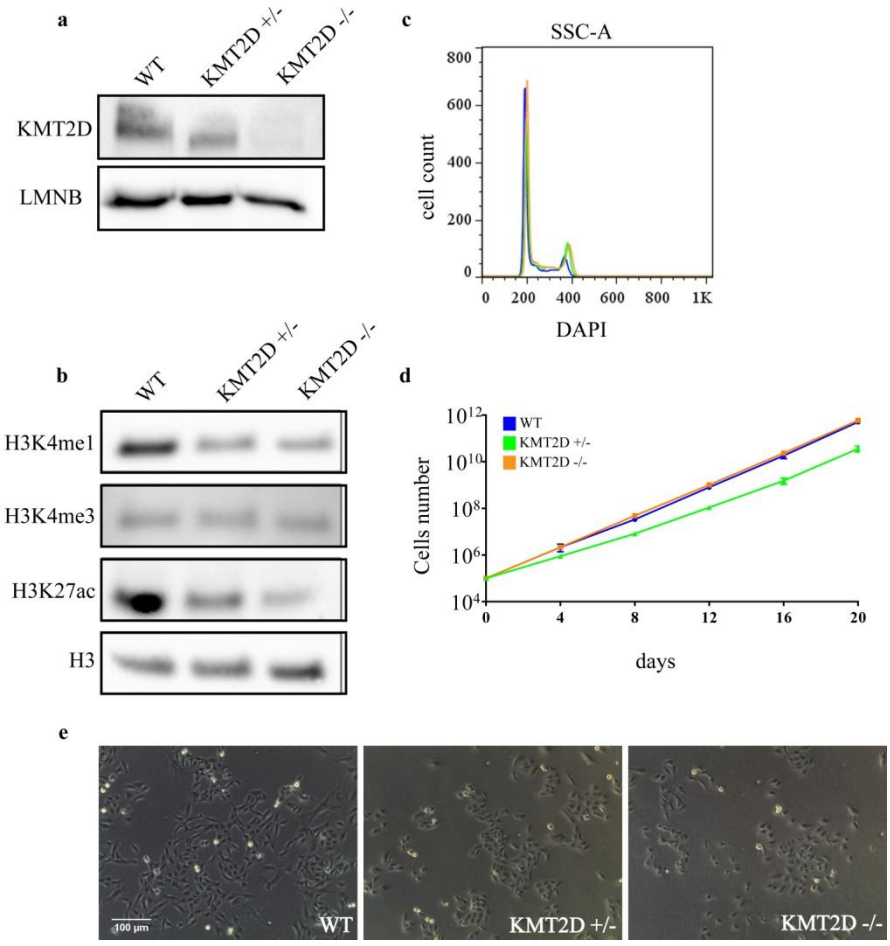


Figure 3. Molecular and phenotypic analysis of mutated clones
a, western blot analysis of KMT2D protein levels in wild type clone (WT) heterozygous clone (KMT2D+/-) and homozygous clone (KMT2D-/-). Lamin b (LMNB) was used as loading control. **b**, western blot analysis of histone modifications H3K4me1, H3K4me3 and H3K27ac in the three different indicated clones. Histone 3 (H3) was used as loading control. **c**, flow citometry analysis of cell cycle. X axis represents DAPI intensity values while Y axis represents cells number. **d**, growth curve of WT, KMT2D+/- and KMT2D-/- iMSCs. Values are reported as mean of three independent experiments ± SEM. Two way Anova analysis was performed, showing any statistical difference. **e**, bright field images of WT, KMT2D+/- and KMT2D-/- iMSCs. Scale bar: 100µm.

In order to deepen the morphological analysis, we performed actin staining with phalloidin conjugated with TRITC fluorescent dye. Actin staining confirmed our primary observations. WT cells have a well oriented structure formed by the actin stress fibers, while in mutated cells we observed a different organization of actin cytoskeleton (Fig 4a). Both heterozygous and homozygous clone appeared more "rounded", as we observed in bright field images. Furthermore, KMT2D$^{+/-}$ and KMT2D $^{-/-}$ cells are less oriented, having the stress fiber pointing in multiple directions. This phenotype is even more evident at higher magnification (Fig 4a, inserts). Taken together, these observations revealed that partial or total loss of MLL4 alters the organization of actin cytoskeleton in iMSCs.

The actin staining allowed us to observe differences in cell spreading between WT and mutated iMSCs. To measure these differences, we quantified the surface occupied by actin fibers. We observed a decreased surface occupation of the two mutated clones respect to WT cells (Fig. 4b). To corroborate this result, we also measured cell size by flow citometry. Compared to WT cells, KMT2D$^{+/-}$ and KMT2D $^{-/-}$ cells appeared smaller, as revealed by the different distribution of the forward scatter (FSC-A) (Fig. 4c). Altogether, these results indicate that, in iMSCs, MLL4 haploinsufficiency causes changes in cell shape, disorientation of actin fibers impairing cytoskeleton organization and cell size. We further investigate the mechanisms insights that could explain a causal link between *KMT2D* mutation and cell shape alteration.

It has been demonstrated that MSCs should maintain a well oriented and a functional actin cytoskeleton in order to rapidly respond to

extracellular cues that induce cell lineage commitment. Among other factors, actin cytoskeleton serves to respond to mechanical stimuli induced by cell-cell and cell-matrix contact[4]. The capacity of a cell to convert mechanical signals conveyed by the microenvironment into biochemical signals to adapt their behavior to the environments is called mechanotransduction[5]. YAP/TAZ are the most characterized factors that transduce the external mechanical stimuli into the nucleus. The activation of YAP and TAZ entails their accumulation in the nucleus. This cellular distribution is controlled by cell shape, by rigidity and topology of the extra cellular matrix (ECM) substrate. YAP and TAZ are nuclear localized in those cells perceiving high levels of mechanical signaling, such as cells cultured on rigid substrates, occupying large adhesive area or experiencing deformation and cytoskeletal tension[5]. Considering the actin cytoskeleton alteration that we observed in our mutated cells, we asked whether this structural changing has a functional impact in altering the cellular mechanotransduction. To this aim, we measured the level of nuclear and cytoplasmatic YAP/TAZ by quantifying their signal from immunofluorescence. As expected, WT iMSCs, showing a large spreading of actin cytoskeleton in culture condition, have a YAP/TAZ signal mostly nuclear. Both mutated clones, instead, showed a decrease in the nuclear fraction of YAP/TAZ (Fig 4d and 4e). As WT and mutated cells are grown under the same culture conditions, these results indicate that both KMT2D$^{+/-}$ and KMT2D $^{-/-}$ cells are affected in their ability to correctly sense the external mechanical forces.

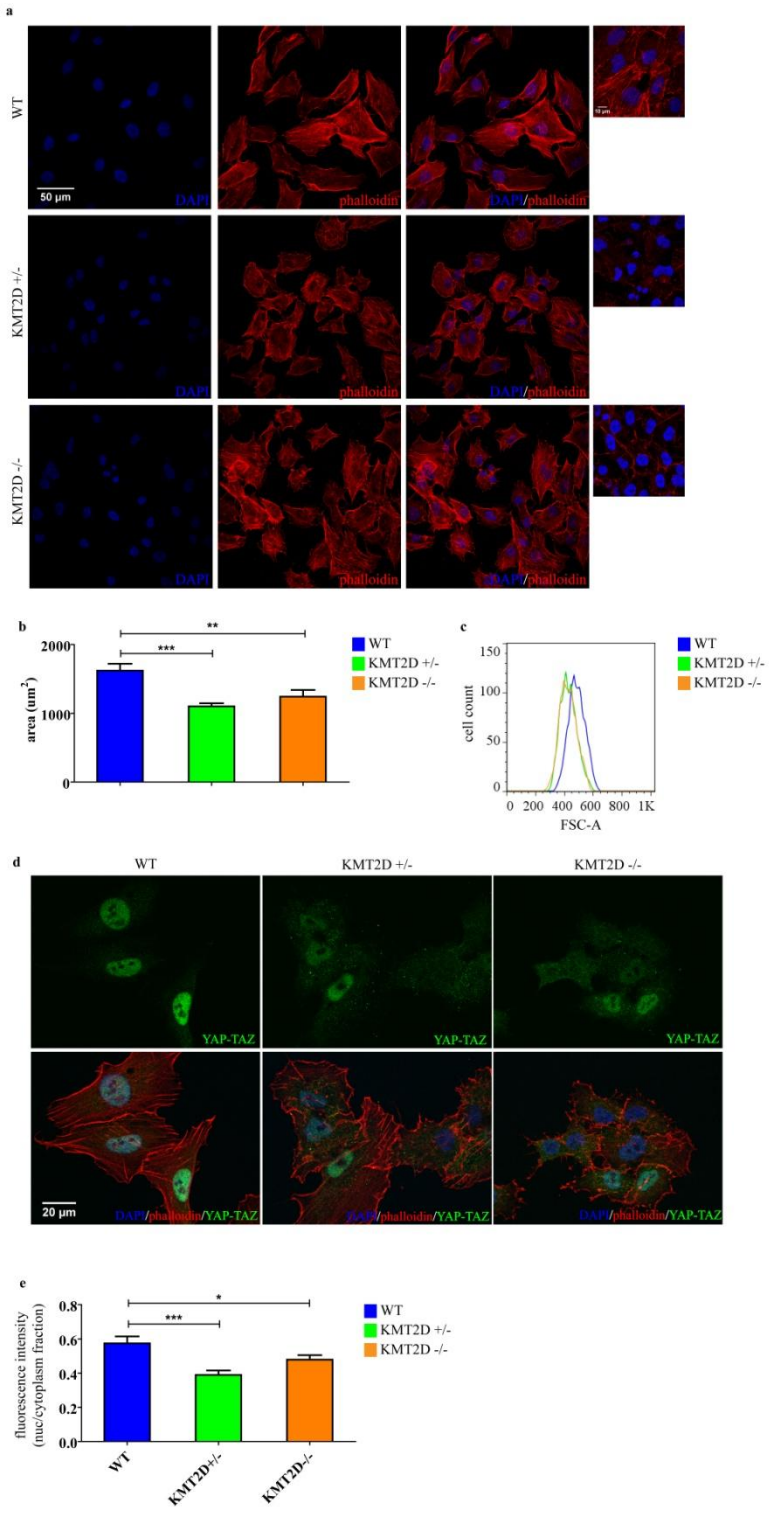

a
WT
KMT2D +/-
KMT2D -/-
50 μm
DAPI
phalloidin
DAPI/phalloidin
b
area (um²)
2000
1000
0
**
***
WT
KMT2D +/-
KMT2D -/-
c
cell count
150
100
50
0
0 200 400 600 800 1K
FSC-A
WT
KMT2D +/-
KMT2D -/-
d
WT
KMT2D +/-
KMT2D -/-
YAP-TAZ
20 μm
DAPI/phalloidin/YAP-TAZ
e
fluorescence intensity
(nuc/cytoplasm fraction)
0.8
0.6
0.4
0.2
0.0
*
***
WT
KMT2D +/-
KMT2D -/-
WT
KMT2D+/-
KMT2D-/-

Figure 9. Analysis of structural changes in mutated iMSCs **a**, confocal images of WT, KMT2D$^{+/-}$ and KMT2D$^{-/-}$ iMSCs. Nuclei were stained with DAPI and actin was stained with phalloidin-TRITC. Scale bar: 50µm. Inserts represent confocal images at higher magnification. Scale bar: 10 µm. **b**, quantification of cell surface occupied by WT, KMT2D$^{+/-}$ and KMT2D$^{-/-}$ iMSCs. Phalloidin-TRITC signal was used to identify and measure cell surface. t-test analysis was performed. **: p-value < 0,001; ***: p-value < 0,0001. **c**, flow citometry analysis of cell dimension of the three different clones. X axis represents forward scatter values while Y axis represents cells number. **d**, confocal images of WT, KMT2D$^{+/-}$ and KMT2D$^{-/-}$ iMSCs. Nuclei were stained with DAPI and actin was stained with phalloidin-TRITC. Green signal represents YAP/TAZ. Scale bar: 20µm. **e**, quantification of the nuclear YAP/TAZ signal. Data are showed as ratio between nuclear and cytoplasmic YAP/TAZ signal. t-test analysis was performed. *: p-value< 0,05 **: p-value < 0,01; ***: p-value < 0,001.

2.1.4 Mutated iMSCs show alteration during chondrocyte differentiation

Analysis performed in undifferentiated KMT2D$^{+/-}$ and KMT2D$^{-/-}$ iMSCs revealed differences mainly in cell shape and cell size.

One of the major characteristics of MSCs is the ability to differentiate into adipocytes, osteoblasts and chondrocytes. We wander if MLL4 loss or haploinsufficiency causes defects in the maintenance of multipotency. To investigate this aspect, we induced WT or mutated iMSCs to differentiate by culturing cells with the lineage-specific differentiation medium for 14 days (Fig 5a). Morphological analyses of differentiated cells showed that *KMT2D* mutation did not affect the cell lineage commitments towards adipogenesis and osteogenesis. Nevertheless, we noticed significant differences in chondrogenic differentiation. Indeed, at the end of the differentiation process, WT iMSCs form typical chondrogenic nodules organized by precisely oriented cells. Instead, KMT2D$^{-/-}$ cells appeared as cellular agglomerates, without any specific orientation or organization. Differentiated KMT2D$^{+/-}$ cells showed an intermediate phenotype with just some organized chondrogenic nodules. Histological stainings specific for each lineages confirmed morphological results. Indeed, when we compared WT to mutated cells, we did not notice major differences in the production of lipid droplets (as revealed by Oil Red staining) (Fig 5b, left panel) or for the production of calcified extracellular matrix (as indicated by Alizarin red staining) (Fig 5b, central panel). Conversely, Alcian blue staining (that reveal glycosaminoglycans, typical of chondrocytes extracellular matrix) showed a significant reduction of ECM production by mutated MSCs

with respect to WT MSCs (Fig 5b, right). This data may indicate that cell lineage commitment and the fully acquisition of chondrogenic phenotype is not reached in mutated cells. To rule out the possibility that the observed phenotype could be partially due to culture conditions, we adopted a 3D culture system, which better mimic the *in vivo* conditions of differentiation. For these reasons, we plated iMSCs as micromass culture, combined with pro-chondrogenic differentiation medium. Also in this 3D differentiation setting we observed a remarkable difference between WT and mutated cells (Fig 5d). Homozygous and heterozygous clones appear as aggregates of rounded cells with apparently reduced cell-matrix or cell-cell contacts. Of note, the micromass of $KMT2D^{+/-}$ and of $KMT2D^{-/-}$ partially or completely flake apart. Conversely, WT cells appear as a full micromass with well oriented cells at the edge and with a compact structure in the center. These results demonstrate how mutated clones fail to fully commit towards chondrocytes and form proper 3D micromass structures. Of importance, these findings are in agreement with the results obtained from the 2D culture system, in which mutated cells did not assume the same chondrogenic phenotype as the WT cells. Considering that one of the key features of chondrocytes is the production and deposition of extracellular matrix[6], we next asked whether loss of MLL4 also affected this chondrogenic-specific cellular function in the 3D setting. To this end, we performed Alcian blue staining to detect glycosaminoglycans in the ECM. Both clones showed a decreased deposition of extracellular matrix (Fig 5d), with respect to control cells. Of note, heterozygous clone revealed a less severe phenotype than homozygous clone, suggesting a gene-dosage

effect. Taken together, these results indicate a role of MLL4 in chondrogenesis, as its total or partial loss cause impairments in organization and maintenance of chondrocyte nodules and determine a reduction in the production of extracellular matrix proteins.

Figure 5. Chondrogenesis is impaired upon MLL4 loss
a, differentiation of WT, KMT2D$^{+/-}$ and KMT2D$^{-/-}$ clones into adipocytes (left), osteoblasts (center) and chondrocytes (right). Scale bar: 100µm. **b**, staining of differentiated cells with Oil red O for adipogenesis (left), Alizarin Red for osteoblastogenesis (center) and Alcian Blue for chondrogenesis (right). **c**, differentiation of WT, KMT2D$^{+/-}$ and KMT2D$^{-/-}$ clones with the micromass culture system into chondrocytes lineage. Scale bar: 500 µm. On the right inserts images at higher magnification. Scale bar: 100µm. **d**, Alcian blue staining of differentiated WT, KMT2D$^{+/-}$ and KMT2D$^{-/-}$ clones into chondrocyte lineages with micromass culture system. Scale bar: 500µm

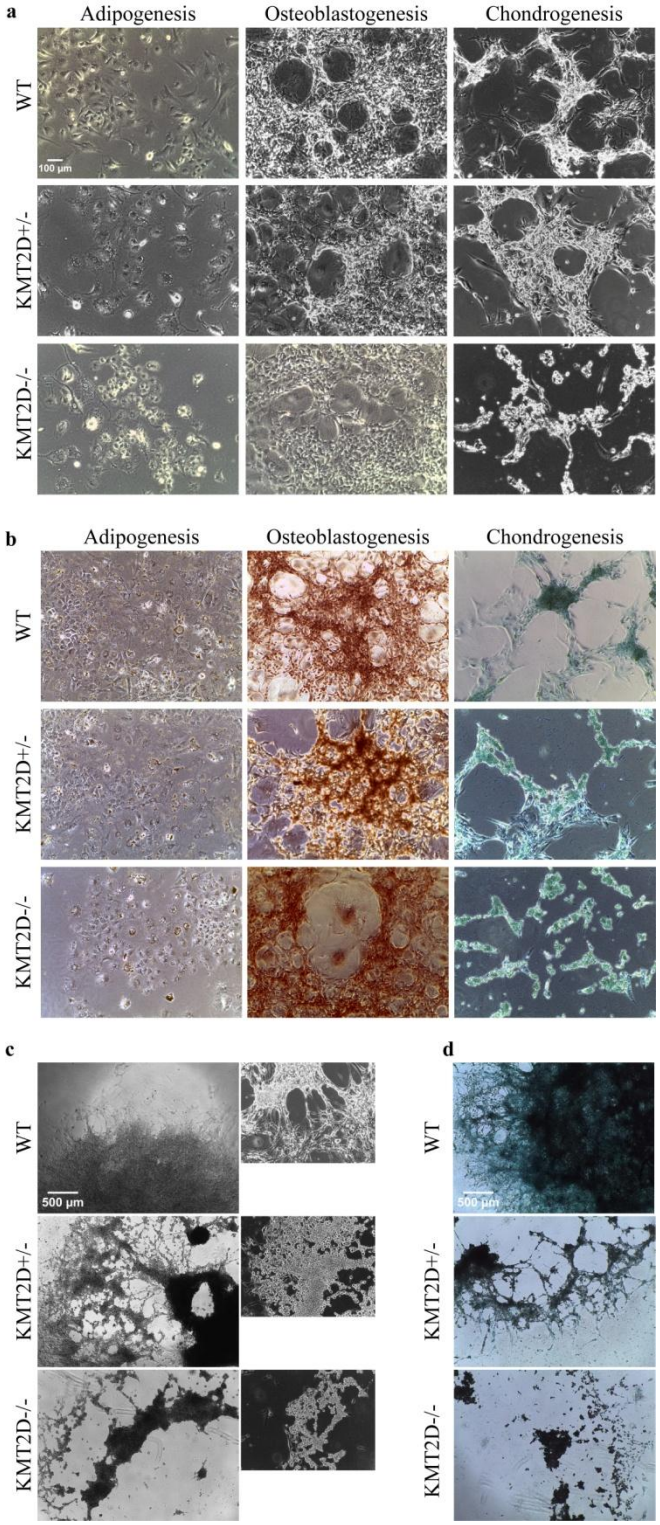


Figure 5. Chondrogenesis is impaired upon MLL4 loss

To understand the differentiation stage when the MLL4 impairment affects chondrogenesis of mutated clones, we analyze the process at different time points: 3, 6, 9 and 13 days post induction (PI) (Fig 6a). When induced to differentiate, mutated clones showed early differences, compared to WT cells. Indeed, already at 3 days PI, disparity in cells shape and micromass configuration are detectable. Mutated cells appear more rounded and isolated, creating holes into the micromass. This suggests that the MLL4 loss influences the first steps of chondrocyte differentiation and the terminal phenotype could be the consequences of these initial issues.

As for undifferentiated iMSCs, we performed actin staining to better visualize cell structure and, eventually, confirm the morphological phenotype. At 3 days of differentiation, in WT cells the F-actin is mostly cortical, defining a cuboidal shape of these cells (Fig 6b). This is in line with different previous reports describing how chondrocyte assume a cuboidal shape with cortical actin filaments during chondrocyte differentiation[7]. This organization is not achieved by mutated cells that have a disorganized actin structure without the precise definition of a cell shape (Fig 6b).

Together with changes in cell shape, another mark of chondrocytes differentiation is the increase of cell size, following the acquisition of hypertrophic phenotype[8]. To assess if KMT2D mutations also alter the expected size increment, we measured the cell size of differentiated chondrocytes by flow cytometry. As previously showed (Fig 4c), heterozygous and homozygous cells are smaller than WT cells. While an increase in cell size is detectable during differentiation of WT cells, both mutants retained their dimensions upon cell lineage commitment

(Fig 6c). These primary finding begin to shed light on *KMT2D* function during chondrogenesis. Partial or total loss of MLL4 prevents the correct organization of the actin cytoskeleton and the loss of micromass structure. Moreover, mutated cells decreased their capacity in depositing extracellular matrix and there is not an increase of cell size, indicating a failure in chondrocyte differentiation process. Considering that differences between WT and mutated cells are observed as soon as 3 days PI, it may be reasonable that MLL4 plays a major role in the induction of chondrocyte specification.

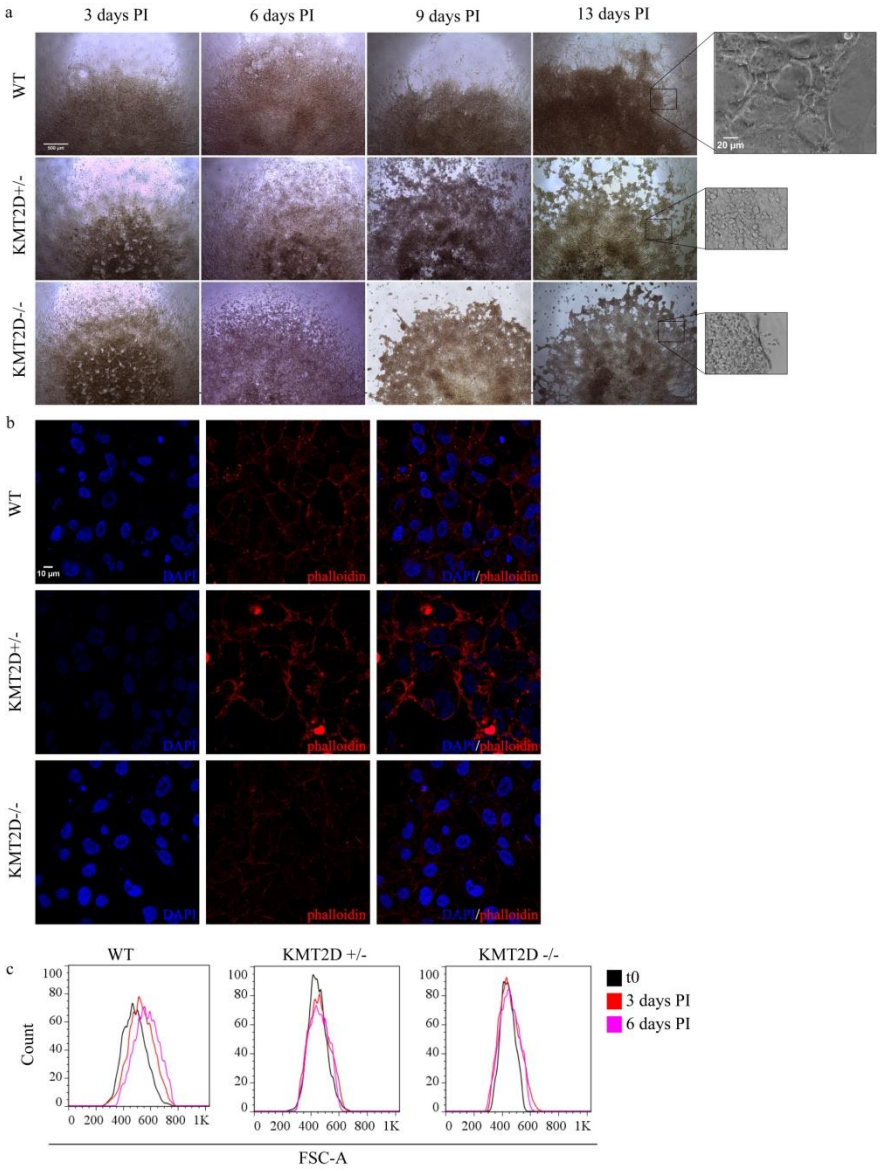


Figure 6 Chondrogenesis defects occur early during the differentiation process of mutated iMSCs. **a**, bright field images of chondrocyte differentiation of WT, KMT2D$^{+/-}$ and KMT2D$^{-/-}$ clones at different time points. PI: post induction. Scale bar: 500 µm. Inserts represent higher magnification images of 13 days PI. Scale bar: 50µm. **b**, confocal images of WT, KMT2D$^{+/-}$ and KMT2D$^{-/-}$ clones at 3rd day of chondrocyte differentiation. Nuclei are stained with DAPI and actin is stained with phalloidin-TRITC. **c**, flow citometry analysis of cell size of WT, KMT2D$^{+/-}$ and KMT2D$^{-/-}$ clones at t0 (undifferentiated cells), 3 and 6 days PI.

2.1.5 Cell cycle of mutated iMSCs is altered during chondrocyte differentiation

Chondrogenesis begins with the initial condensation of MSCs that forms primary chondrogenic nodules. The first immature chondrocytes are actively proliferating cells and the exit from cell cycle is a tightly controlled step that is fundamental for the subsequent differentiation into hypertrophic chondrocytes. Deregulation of these processes can result in many different diseases such as chondrodysplasias and chondrogenic tumors[9].

Taking into consideration this important aspect, we wandered whether one of the causes of the differentiation failure of $KMT2D^{+/-}$ and $KMT2D^{-/-}$ cells could depend on impairment of cell cycle exit upon differentiation. For this reason, we analyzed cell cycle profile of differentiating iMSCs. Cell cycle consists of four distinct phases. In G1 phase cell synthesizes mRNAs and proteins in preparation for subsequent steps leading to mitosis. If the cell decides to divide, it passes in the S phase, during which DNA is replicated. Then, cells enter the G2 phase, characterized by protein synthesis and rapid cell growth to prepare the cell for mitosis, which represents the last phase when cellular energy is focused on the orderly division into two daughter cells. In G1 phase, cells can also exit from cell cycle in response to differentiation cues, reaching quiescence (named G0 phase)[10]. Through flow citometry, cell cycle phases can be detected by measuring their DNA content. Indeed, when cells terminate DNA replication, they possess a double content of DNA respect to G1 phase.

Regarding WT cells (Fig.7, upper panels), we observed a decrease of cell population in S phase at 3 days PI, with the consequent increase of G1 cell population. This indicates that WT cells exit from cell cycle starting from 3 days PI. Interesting, there is also a progressive increase of the cell population in sub-G1 phase, indicative of apoptotic cells. Therefore, a small proportion of WT cells undergo apoptosis during differentiation. This is not surprising, because later stages of chondrogenesis are characterized by apoptotic chondrocytes[9].

Analyzing heterozygous and homozygous clones at 3 days PI, S-phase cell population slightly decreases respect to undifferentiated MSCs but it remains higher respect to WT cells (Fig 7, central and bottom panel), indicating that there is a fraction of cells that keep on dividing. Strikingly, we did not observe an increase of cell proportion in G1-phase, but there is an increase of cells in G2/M-phase. Of note, the G2/M check-point blocks cells entering mitosis until they repair damaged DNA, which otherwise leads to apoptosis[10]. These data suggest that heterozygous and homozygous clones do not exit from cells cycle but they remain blocked in the G2/M checkpoint indicating the presence of some impediments to the cell cycle progression. These results were also in agreement with the detection of an increase of apoptotic cells in mutated clones during differentiation (Fig 7, central and bottom panel).Together these results prompt us to hypothesize that total or partial loss of MLL4 impairs the timing of cell cycle progression during chondrocyte differentiation. Further analysis will be required to understand whether this cell cycle deregulation is a consequence of the overall differentiation impairment discussed above or if this could contribute to the failure of chondrocyte commitment.

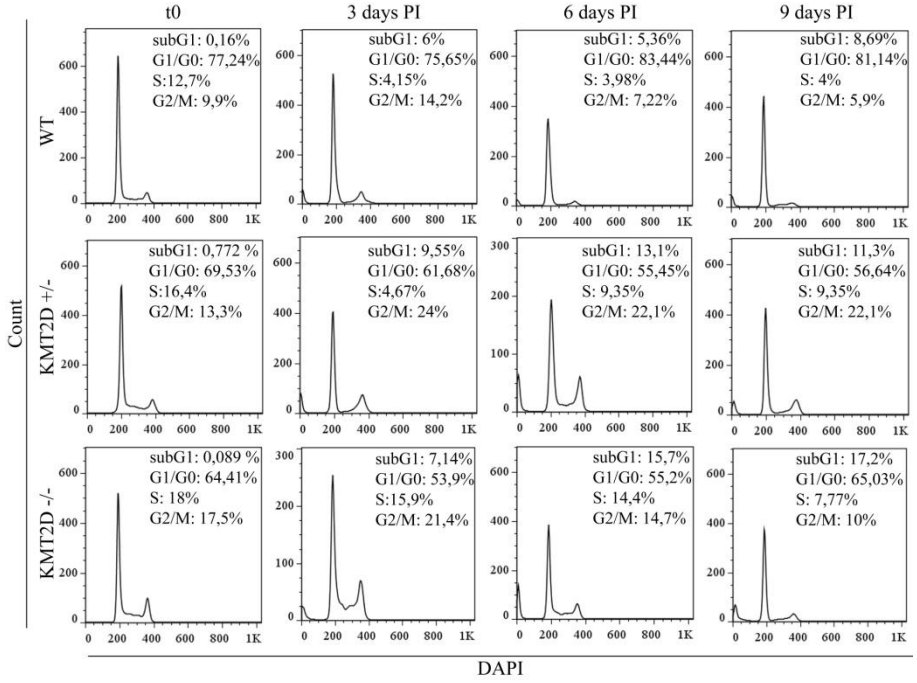


Figure 10. Cell cycle profiles of mutated clones during chondrogenic differentiation Flow cytometry analysis of cell cycle of WT, KMT2D$^{+/-}$ and KMT2D$^{-/-}$ undifferentiated clones (t0) and at indicated time points during chondrogenesis. On each graph percentages of cells in distinct cell cycle phases are reported.

2.1.6 Chondrocyte lineage-specific transcription factors and markers show altered expression patterns in *KMT2D* mutant iMSCs

Chondrogenesis is a multi-step process tightly regulated by extracellular cues that determine the expression and nuclear shuttling of time-specific transcription factors that, in turn, induce the expression of specific proteins required for proper formation of cartilage and bones. We decided to analyze by qRT-PCR, expression of some key transcription factors (TFs) and lineage-specific markers

in order to understand if the altered phenotype that we observed upon the total or partial loss of MLL4 is also accompanied by an altered transcriptional program.

SOX9 is the key TF required for the induction of chondrogenesis. It is already expressed in MSCs and it is down regulated during the chondrogenic process[11]. We observed that SOX9 transcript level was slightly decreased in undifferentiated iMSCs carrying *KMT2D* mutations, with respect to WT cells (Fig. 8a, left panel). RUNX2 is the TF responsible for the induction of hypertrophic phase in proliferating chondrocyte but it is also expressed in MSCs[11]. In accordance, although RUNX2 resulted being expressed in undifferentiated cells, its transcript level was reduced in mutated cells, compared to WT iMSCs (Fig. 8a, central panel). The ATF4 transcription factor is also responsible for the chondrocytes hypertrophy but it is induced only in late phases upon terminal differentiation[12]. We found that in $KMT2D^{-/-}$ cells there is an increase of the expression of ATF4 in undifferentiated MSCs respect to both WT and $KMT2D^{+/-}$ cells (Fig. 8a, right panel). These first results suggest that partial or total MLL4 loss alter the expression of some lineage-specifying TFs, already in the undifferentiated state of MSCs.

We next analyze the expression of the same TFs during differentiation to assess whether their deregulation was maintained or increased during chondrogenesis. Physiologically, SOX9 should be down regulated after the proliferating phase of chondrogenesis.[13] Indeed, in WT cells we observed a decrease of *SOX9* RNA level at 3 days PI. We did not observe differences in *SOX9* dynamics between WT and

mutated cells. Indeed, also in mutated cells we observed a decrease of SOX9 expression at 3 days PI (Fig. 8b, left panel).

RUNX2 expression has a more complex dynamic: it is expressed in MSCs but it is down-regulated during the first phase of chondrogenesis. Thereafter, it is up regulated again in the early phase of hypertrophy[11]. Observing the expression pattern of RUNX2 during the time-course analysis, WT cells showed that the *in vitro* model recapitulated its dynamic regulation during chondrogenesis. However, the same analyses showed that the expression pattern of RUNX2 was altered in $KMT2D^{+/-}$ and $KMT2D^{-/-}$ cells, compared to WT cells (Fig. 8b, central panel).

ATF4 is a TF controlling the cell cycle and the late phase of chondrogenesis[12] and it should be induced only in the late phase of hypertrophy. By analyzing the early time points of the differentiation process, we measured a mild increment of its expression in both WT and in the heterozygous cells, which did not occur in homozygous clone (Fig. 8b, right panel). Considering the altered expression pattern of the analyzed TFs, we asked whether these alterations could affect the expression of their targets, such as extracellular matrix proteins. We analyzed the expression of COLLAGEN type II and type X, AGGRECAN and the metallopeptidase MMP13. The obtained data showed that while these genes were strongly induced upon differentiation of wild-type cells, this increment was strongly reduced in *KMT2D* mutant cells (Fig. 8c). This finding suggested a failure in the production of extracellular matrix that is a fundamental requirement for chondrocyte differentiation. Considering these results, we suggest that one of the *KMT2D* function during chondrogenesis is

the regulation of the correct transcriptional program. Indeed, MLL4 impairment causes an overturning of the TFs dynamics and consequently, mutated cells are no more able to produce the extracellular matrix.

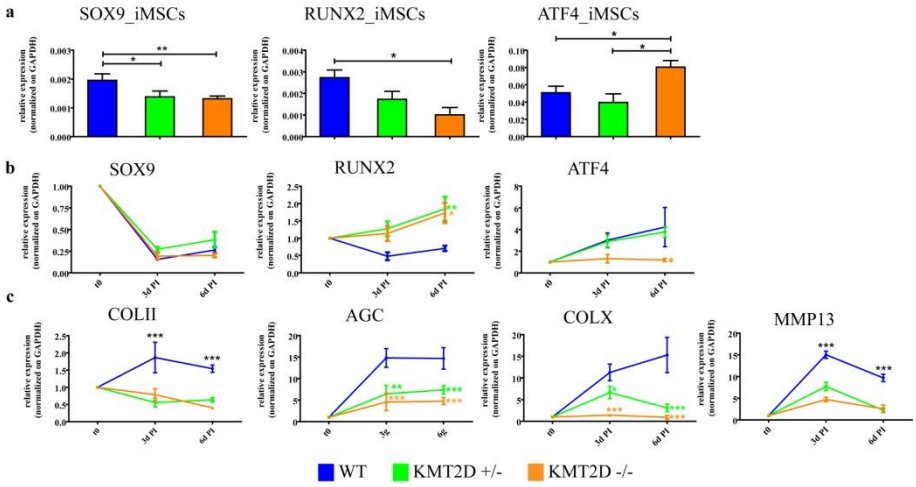


Figure 8. Gene expression analysis of lineage-specific TFs and markers in mutated clones

a, qRT-PCR analysis of key TFs of the chondrocyte lineage (SOX9, RUNX2 and ATF4) in undifferentiated iMSCs. Data are normalized on GAPDH and represented as mean of 3 independent experiments +SEM. One way Anova analysis was performed. *: p-value< 0,05; **: p-value < 0,001; ***: p-value < 0,0001. **b**, **c**, qPCR analysis of (**b**) chondrocyte lineage-specific TF and (**c**) extracellular matrix proteins produced by chondrocytes (COLX, collagen type X; COLII, collagen type II; AGC, aggrecan; MMP13 matrix metallopeptidase) and of transcription factors during the differentiation process . Data are normalized on GAPDH and expressed as fold change on undifferentiated cells. Values represent mean of 3 independent experiments ±SEM. One way Anova analysis was performed. *: p-value< 0,05; **: p-value < 0,001; ***: p-value < 0,0001.

2.1.7 LSD1 inhibitor rescue the MLL4 loss phenotype

As mentioned above, MLL4 is a monomethyl transferase responsible for the deposition of H3K4me1 on enhancer regions. We showed that haploinsufficiency or total loss of MLL4 cause a decrease in the global level of H3K4me1 (Fig 3b). Moreover, our data also indicate that partial or total loss of MLL4 causes different alterations in MSCs and chondrocytes thus resulting in deregulation of chondrogenesis (Fig. 4-8). In light of these findings, we asked whether the observed phenotype is due, at least in part, to the global decrease of H3K4me1 level. To address this issue, we take advantage from the possibility to counterbalance the reduced level ofH3K4me1 by targeting H3K4-specific histone demethylases. In particular, we focus our attention on the inhibition of the enzymatic activity of the H3K9 and H3K4 demethylase, LSD1. This demethylase acts as a co-repressor mediating the demethylation of H3K4me1/2 and favors enhancers decommissioning upon differentiation signals[14]. We supposed that if the observed defects upon MLL4 loss are due to the decreased H3K4me1 levels, the inhibition of an enzyme that is responsible for the demethylation could restore the monomethyl level thus rescuing, at least in part, chromatin state in the targeted cells. To this end, we administered different doses of LSD1 inhibitor (GSK2875552)[15] during the chondrocyte differentiation of WT and homozygous clone. We measured the effects of the drug treatment both in terms of morphology of the micromass and in terms of extracellular matrix deposition. Strikingly, we observed an improvement of the differentiation capacity of homozygous clone (Fig. 9a). Specifically our data showed that in response to LSD1 inhibition, the micromass

did not disaggregate, remaining compact, similarly to the WT cells. Moreover, quantification of Alcian blue staining showed that LSD1 inhibitor treatment restores the production of extracellular matrix (Fig. 9b). Together, these data indicate that the inhibition of the enzymatic activity of LSD1 rescues, at least in part, the failure of chondrocyte differentiation that occurs in homozygous clone.

Notably, the drug treatment causes some defects also in micromass of WT cells, which do not maintain its full and compact structure. This may suggest that not only a decreased level of H3K4me1 causes alteration in chondrogenesis, but also that the maintenance of correct balance between methylation and demethylation is important for proper chondrocyte differentiation (Fig. 9a, upper panel)

Further analyses are required to confirm that the phenotypic rescue is the consequence of the restoration of the methylation level. However, these preliminary results indicate that modifying the dynamic of histone methylation deposition affects chondrocytes differentiation. Of importance with this experiment, we showed that we have a robust tool with which we can measure the effects of drug treatment on chondrogenic potential. This is a fundamental aspect for the development of high-throughput drug screening approach to identify and characterize compounds that could ameliorate some of the clinical features affecting KS patients.

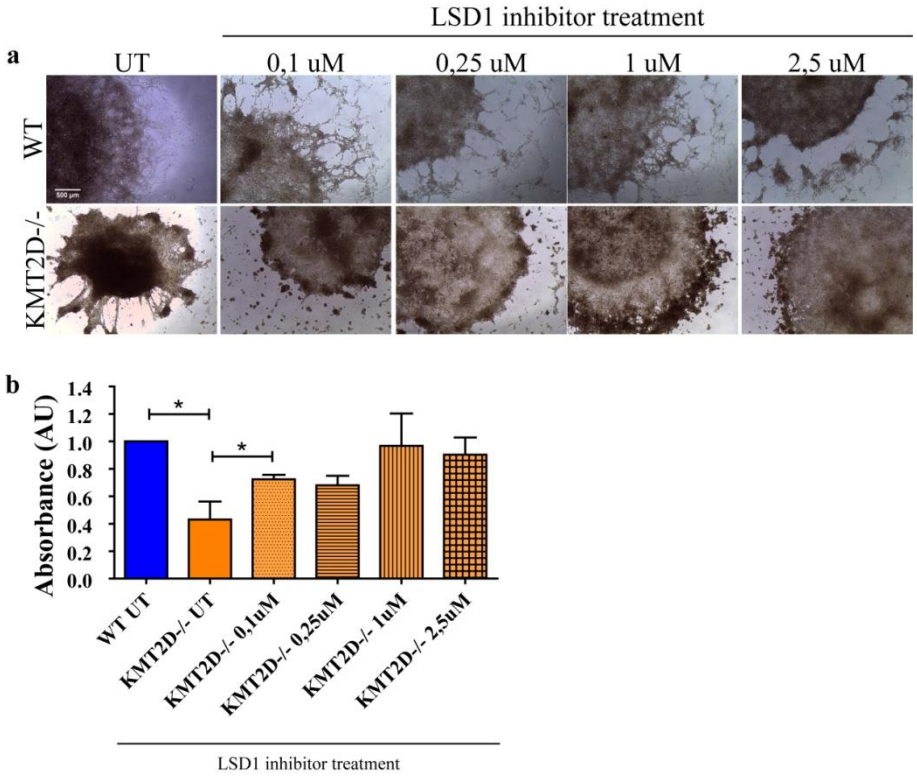


Figure 9. Inhibition of H3K4 demethylation restores the chondrocyte differentiation
a, bright field images of WT and KMT2D$^{-/-}$ differentiated chondrocytes treated with different doses of KDM1A inhibitor (GSK2875552). Scale bar: 500µm. **b**, quantification of alcian blue staining of WT and KMT2D$^{-/-}$ differentiated chondrocytes. The absorbance was normalized on the absorbance of WT untreated cells. Values are represented as mean of 3 independent experiments + SEM. T-test analysis was performed. *: p-value<0,05.

2.1.8 ATR inhibition partially rescue the chondrogenic potential

As described above, iMSCs carrying inactive MLL4 show an altered mechanotransduction, as demonstrated by the delocalization of YAP/TAZ (Fig. 4d-e). Classically, mechanotransduction means the transduction of extracellular mechanical stimuli inside of the cell. However, recent evidences pointed out the possibility of mechanical stimuli coming from inside of the cell. In particular, the nucleus can change its structure and, thanks to protein transducers, this changing can influence the structure and the behavior of a cell[16]. How exactly this changes is transduced to the cell is yet unknown. Recently, Kumar et al. showed that ATR is able to sense alteration of the nucleus structure. ATR relocate on the nuclear envelope and phosphorilate CHK1[17]. In this view, ATR can be considered a transducer of nuclear structure alteration.

It has been demonstrated that modification of the chromatin condensation influences the shape and the stiffness of the nucleus [16]. In our mutated cells we observed a global decrease of H3K4me1 and H3K27ac. These histone modifications typically characterize an open accessible chromatin. Giving these premises, we postulate that decrease of histone modifications, that serve to maintain an open chromatin state, could increase the compaction of chromatin thus increasing the nuclear stiffness. This structural change may be sensed by ATR that, with an unknown mechanism, could block the progression of chondrogenesis. If this suggestion is true, inhibition of ATR should release this block and, thereby, restore the chondrogenic capacity of mutated cells. To test this hypothesis, we treated WT and

mutated cells with an ATR inhibitor during the differentiation process. Interestingly, mutated cells seem to rescue the chondrogenic capacity, upon ATR inhibition (Fig 10a). At the end of differentiation process, micromass structure of mutated cells appeared more solid indicating an increased capacity in maintaining the organization of the micromass (Fig 10b). Mutated cells also improved the deposition of extracellular matrix. Indeed, quantification of Alcian blue staining revealed the restoration of the deposition of extracellular matrix at the same level of WT cells in both clones, but especially the heterozygous clone (Fig 10c). Strikingly, the phenotypic rescue observed is dose-responsive, indicating that it depends on the level of ATR activation. To rule out the possibility that the observed effect is due to an altered level of DNA damage of MSCs, we measured the level of phosphorilated H2AX (pH2AX) on WT and mutated MSCs. ATR dependent phosohorilation of H2AX marks DNA regions with double strand break. Usually, it is used to quantify the level of DNA damage and, consequently, the level of ATR activation[18]. Performing immunofluorescence on pH2AX we did not detect differences between the WT and mutated iMSCs. The foci number is equal among WT, KMT2D$^{+/-}$ and KMT2D$^{-/-}$ cells, indicating that mutated cells do not show a higher level of DNA damage level (Fig. 10 d-e). This indicates that the phenotypic rescue that we observed upon ATR inhibition is not due to its activity in DNA damage response. Even if these are preliminary data, we postulate that in mutated cells the decrease of histone marks typical of open chromatin causes the condensation of chromatin thus inducing a structural alteration of

nucleus. This is sensed by ATR that activate an unknown response thus affecting the chondrogenic capacity of mutated cells.

Figure 10. ATR inhibition partially restores the chondrogenic potential of mutated cells. **a**, overview of Alcian blue staining of WT, KMT2D$^{+/-}$ and KMT2D$^{-/-}$ cells differentiated into chondrogenic lineage and treated with different doses of ATR inhibitor, as indicated. UT: untreated cells. **b**, images of Alcian blue staining of WT, KMT2D$^{+/-}$ and KMT2D$^{-/-}$ cells differentiate into chondrogenic lineage and treated with different doses of ATR inhibitor, as indicated. Scale bar: 500µm. **c**, quantification of Alcian blue staining of WT, KMT2D$^{+/-}$ and KMT2D$^{-/-}$ cells differentiated into chondrogenic lineage and treated with ATR inhibitor, as indicated. Data are represented as fold increase respect to untreated WT cells. Data represent mean +SEM of 3 independent experiments. **d**, confocal images of WT, KMT2D$^{+/-}$ and KMT2D$^{-/-}$ MSCs stained with pH2AX (green) and DAPI (blue). Scale bar: 20µm. **e**, quantification of the number of pH2AX foci from immunofluorescence in **d**.

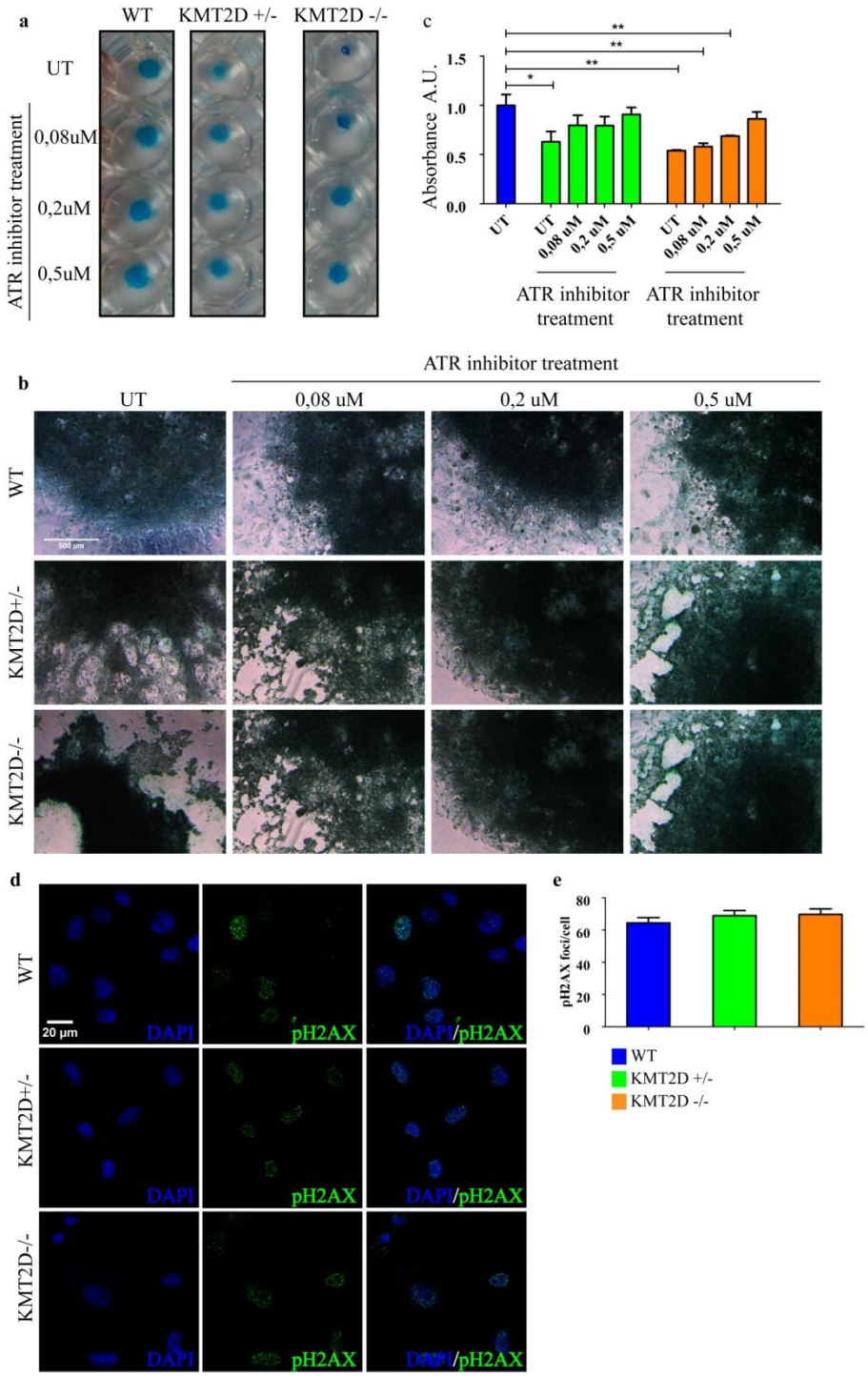


Figure 10. ATR inhibition partially restores the chondrogenic potential of mutated cells

2.1.9 Mll4 loss causes craniofacial defects in medaka animal model

Our *in vitro* findings suggest that *KMT2D* has a role in orchestrating the chondrogenic process. In KS patients tissues such as bones and cartilage are affected. Thus, in a wider view, the deregulation of chondrocytes development, caused by MLL4 impairments, may explain the skeletal symptoms of KS. Nevertheless, considering only these *in vitro* results, this assumption results fetched. To give robustness to our hypothesis, we asked whether MLL4 loss causes defects in chondrogenesis also in developing organisms. To this end, we aimed to generate an *in vivo* model for MLL4 deficiency, which could confirm our *in vitro* molecular findings and could be used to test potential therapeutic approaches

In collaboration with the laboratory of Ivan Conte (TIGEM), we developed a strategy in order to understand the role of KMT2D *in vivo* using the medaka animal model. This animal model is advantageous for the study of developmental diseases since it has an external fertilization and development, permitting more tightly controlled exposure regimens. Moreover, the presence of a transparent egg chorion is helpful for noninvasive direct observation of the early developmental stages (gastrulation and somitogenesis)[19]. By using morpholino to reduce the expression of the target gene, we attempted to recapitulate the haploinsufficiency of *KMT2D*. Morpholino (MO) is a modified DNA sequence that, through the binding to the target sequence, blocks access to other molecules[20]. We used two MO sequences directed against KMT2D. One of the two morpholino was directed to 5-UTR region, leading to the knock-down of the gene,

caused by the interferences with the progression of the ribosomal initiation complex. This MO should mimic the effects of those mutations that decrease the MLL4 stability and activity. The other MO sequence was designed to alter the normal splicing of *KMT2D,* by preventing small nuclear ribonucleoproteins complexes from binding it. In particular, the second MO reproduced a splicing-mutation that causes the exclusion of exon 11 from the final mRNA of KMT2D. Interestingly, with both morpholino medaka embryo showed a clear phenotype (Fig 11). The double staining of Alcian blue (for cartilage tissues) and Alizarin red (for bone tissues) show obvious craniofacial defects. In both knock-down and splice-site MO, embryos show shorter and wider kinked ceratohyal cartilage arch compared to WT embryo (Fig 11, top panels). Moreover, in the ventral view we observed that the embryos injected with the splicing morpholino have a longer distance between kinked ceratohyal cartilage and meckel cartilage when compared with control embryos (Fig 11, bottom panels). These results confirm our hypothesis that MLL4 alterations (in terms of expression but also in terms of splicing pattern) cause defects in the development of the osteo-chondro tissues. Overall, these findings validate, at least in part, our *in vitro* results as the lineage that we found altered upon MLL4 impairment is the same affected by morpholino treatment in medaka embryos, suggesting a role for *KMT2D* in skeletal development. Likewise, it may be possible that defects observed in KS patients result from abnormalities during development of the same skeletal tissues due to impairment of MSC differentiation towards these lineages.

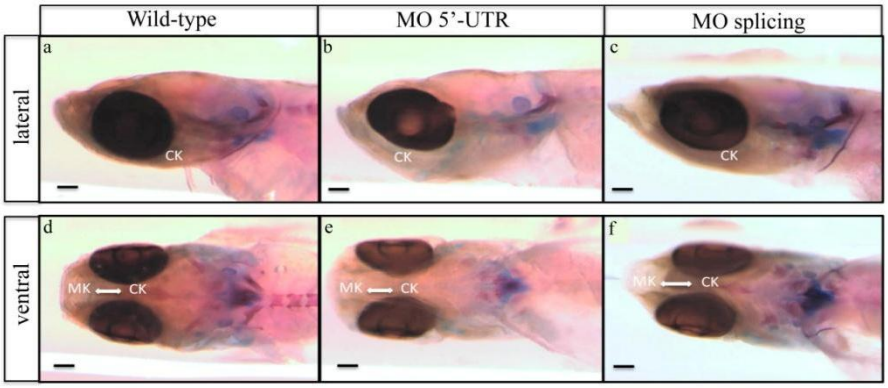


Figure 11. KMT2D loss impairs craniofacial development of medaka
Phenotypic analysis of medaka embryos with double Alcian blue and Alizarin red staining. **a, d,** wild-type embryos; **b, e,** embryos injected with 5'-UTR morpholino; **c, f,** embryos injected with splicing morpholino. **a-c**, lateral view. **d-f**, ventral view. CK, kinked ceratohyal cartilage. MK, meckel cartilage.

2.2 Materials and methods

2.2.1 Cell Culture

hTERT-immortalized human mesenchymal stem cells (iMSCs) were cultured at 37°C and 5% CO_2 in 1:1 DMEM/F-12 medium (Gibco #11320-074) supplemented with 10% Fetal Bovine Serum (Euroclone #ECS0180L). iMSCs expressing Cas9WT were generated by transducing iMSCs with lentiviral vector generated using the plasmid pCW-Cas9WT (Addgene # 50661). Transduced cells were selected with 1µg/ml of puromycin. The expression of sgRNAs were obtained by transducing iMSCs-Cas9WT with lentiviral vectors generated using the plasmid pLX-sgRNA (Addgene #50662). Transduced cells were selected with 2µg/ml of blasticidin.

For adipocyte differentiation, cells were seeded with a density of 1×10^4 cells/cm^2 in iMSCs medium. The day after, medium was changed with adipogenesis medium (Gibco #A10410-01) supplemented with Stem-Pro Adipogenesis supplement (Gibco #10065-01). Medium was changed every 4 days. For a complete differentiation, cells were cultivated for 13 days and then processed for further analysis.

For osteoblasts differentiation, cells were seeded with a density of 5×10^3 cells/cm^2 in iMSCs medium. The day after, medium was changed with osteogenesis medium (Gibco #A10069-01) supplemented with Stem-Pro Osteogenesis supplement (Gibco #10066-01). Medium was changed every 4 days. For complete differentiation, cells were cultivated for 13 days and then processed for further analysis.

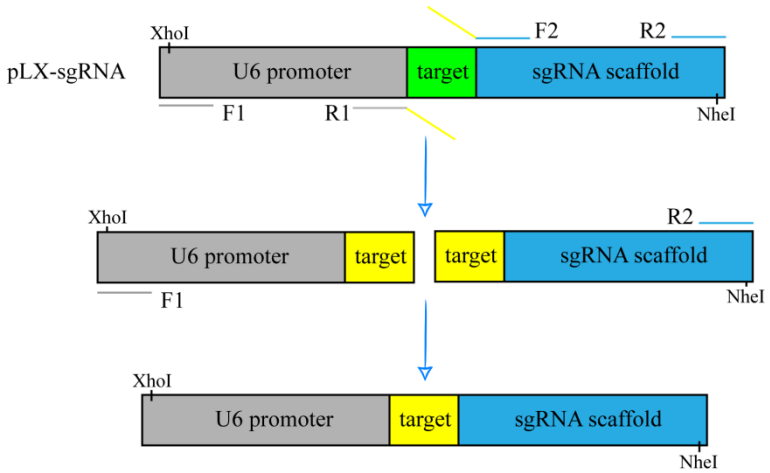
For chondrocytes differentiation, cells were seeded with a density of 5×10^3 cells/cm^2 in iMSCs medium. The day after, medium was changed with chondrogenesis medium (Gibco #A10069-01)

supplemented with Stem-Pro Chondrogenesis supplement (Gibco #10064-01). Medium was changed every 4 days. For complete differentiation, cells were cultivated for 13 days and then processed for further analysis.

For chondrocyte micromass culture, a cell solution of 1.6×10^7 viable cells/ml was generated. Micromass cultures were generated by seeding 5-µL droplets of cell solution. After cultivating micromass cultures for 2 hours under high humidity conditions, iMSCs medium was added. The day after, the medium was changed with chondrogenesis medium. Medium was changed every 4 days. For complete differentiation, cells were cultivated for 13 days and then processed for further analysis.

2.2.2 sgRNAs design and cloning

sgRNAs were designed using the online tool e-Crisp (http://www.e-crisp.org/E-CRISP/). Once designed, they were cloned in the pLX-sgRNA vector following the instruction from Wang et al. 2014[21]. Briefly, new target sequences was cloned into pLX-sgRNA between the XhoI and NheI sites using overlap-extension PCR (see figure) followed by restriction/ligation into pLX-sgRNA vector.



For the cloning PCRs, Phusion® High-Fidelity DNA Polymerase (NEB #M0530) was used. For the enzymatic digestion, the XhoI enzyme (NEB #R0146) and NheI enzymes (NEB #R3131) were used. The ligation was performed using T7 DNA ligase (NEB # M0318). The oligonucleotides used for sgRNA cloning are listed in the table. sgRNA target sequence is underlined

NAME	CLONING OLIGOs
sgRNA1-R1	gagagtgagggctttacctcggtgtttcgtcctttcc
sgRNA1-F2	gaggtaaagccctcactctcgttttagagctagaaatagcaa
sgRNA2-R1	gctctctgcagctgcagccacggtgtttcgtcctttcc
sgRNA2-F2	gtggctgcagctgcagagagcgttttagagctagaaatagcaa
sgRNA3-R1	tgatgtgtagagggcccctcggtgtttcgtcctttcc
sgRNA3-F2	gaggggccctctacacatcagttttagagctagaaatagcaa
sgRNA4-R1	cacagcaggtggaggaagccggtgtttcgtcctttcc
sgRNA4-F2	ggcttcctccacctgctgtggttttagagctagaaatagcaa
F1	aaactcgagtgtacaaaaaagcaggctttaaag
R2	aaagctagctaatgccaactttgtacaagaaagctg

2.2.3 Surveyor assay

Genomic DNA was extracted from iMSCs transduced with lentiviral vectors expressing the Cas9 and one of the 4 sgRNAs designed. With specific oligonucleotides the genomic region containing the sgRNA target region was PCR amplified using Phusion® High-Fidelity DNA Polymerase (NEB #M0530). The PCR products were purified using the Wizard® SV Gel and PCR Clean-Up System (Promega # A9282).

The PCR products were subjected to denaturation and slow renaturation following this program:

- 95°C 10m
- 95°C to 85° -2°C/s, 85°C 1m
- 85°C to 75° -0,3°C/s, 75°C 1m
- 75°C to 65° -0,3°C/s, 65°C 1m
- 65°C to 55° -0,3°C/s, 55°C 1m
- 55°C to 45° -0,3°C/s, 45°C 1m
- 45°C to 35° -0,3°C/s, 35°C 1m
- 35°C to 25° -0,3°C/s, 15°C 1m
- 4°C

Then, the PCR products were digested for 45m at 37°C with the T7 endonucleases (NEB #M0302). The digestions were then verified on 2% agarose gel.

2.2.4 Lentiviral vectors production

All the lentiviral vectors used in this work were third generation lentiviral vectors. HEK 293T packaging cells were seeded at 5×10^3 cells/cm^2 cell density in DMEM medium (Euroclone #ECB7501L) supplemented with 10% Fetal Bovine Serum, 100 U/ml Pen/Strep (Life Technologies #15140-122) 1X Non Essential Amino Acids (Euroclone #ECB3054D), 1mM Sodium Pyruvate (Euroclone #ECM0542D), 2mM L-Glutamine (Life Technologies #25030-081). The day after, cells were transfected with the calcium phosphate method with the packaging vectors pREV, pVSVG, pMDL and with the transfer vector using the 1: 1,44: 2: 5,12 ratio. 16 hours after medium was changed with ISCOVE medium (Euroclone,

#ECB2072L) supplemented with 10% Fetal Bovine Serum , 100 U/ml Pen/Strep, 1X Non Essential Amino Acids, 1mM Sodium Pyruvate. After 30 hours, the medium was harvested, filtered with 0,22 µm filter, and centrifugated for 2h at 20000rpm. Supernatant was discarded and viral vectors pellet was resuspended in PBS. Viral vectors were stored at -80°C.

2.2.5 Animal studies

All the experiments on medaka animal model were performed by the laboratory of Ivan Conte in TIGEM.

Morpholino sequences were injected into fertilized one-cell embryos as previously described[22].

The two morpholino sequences (Gene Tool, LLC) are:

- KMT2D down regulation
  CCCTGCTGCTGCTTTGATCTCTTTG
- KMT2D splicing alteration
  CCAAAGAAAGCTTCCTGTAAAAGA

2.2.6 Immunofluorescence

For immunofluorescence of undifferentiated iMSCs, cells were seeded on coverslips coated with 0,1% gelatin (Sigma Aldrich # G1393). For immunofluorescence of micromass, cells were seeded on 24 well plate with imaging plastic bottom. When needed, cells where fixed with 4% paraformaldehyde for 15 minutes at 4°C. Coverslips were processed as described: permeabilization and blocking with PBS/1% BSA/0.3% Triton X-100 (blocking solution) for 1 hour at room temperature, followed by incubation with primary antibody (diluted in the blocking

solution) for 2 hours at RT (or overnight at 4°C for micromass culture), 3 washes in the blocking solution and incubation with secondary antibodies (diluted in the blocking solution), DAPI for nuclear staining and phalloidin-TRITC for 30 minutes at room temperature. Images were acquired using a Leica TCS SP5 confocal microscope with HCX PL APO 63x/1.40 objective. Confocal z stacks were acquired with sections of 0.35 μm. In cases where image analysis was performed, image acquisition settings were kept constant.

TARGET PROTEIN	PRODUCER	CODE
YAP/TAZ	Cell Signaling	8418
pH2AX	Cell Signaling	9718

2.2.7 Flow Cytometry Analysis

For cell cycle analysis cells were fixed with 70% cold ethanol and then washed 2 times with PBS. Cells were then incubated in the staining solution (PBS with 1μg/ml DAPI and with 0,1% TRITON) in order to have 10^6 cells/ml. Cells were incubated for 30 minutes in the dark at 4°C and then analyzed with FACS CANTO.

2.2.8 Histochemical analysis

To detect adipogenesis, cells were washed with PBS and fixed in 4% formaldehyde for 1 hour at RT, washed with PBS and then stained for 30 min – 1 hr with fresh and filtered Oil-Red O solution (Sigma-Aldrich # O0625) composed of 3 parts of a 0.5% stock solution in isopropanol and 2 parts of distilled water. Then cells were washed three times with distilled water.

To detect osteogenesis, cells were washed with PBS, fixed with ice-cold 70% ethanol and incubated with filtered 2% (p/v in distilled

water) Alizarin red solution (Sigma # A5533) for 15 min. Then cells were washed three times with distilled water.

To detect chondrogenesis, cells were washed with PBS and fixed in 4% formaldehyde for 1 hr at RT and then washed with PBS. Cells were incubated with Alcian blue solution (1 g/L in 0.1 M HCl, Sigma-Aldrich # B8438) for 6 hrs at room temperature and then extensively washed with PBS. To measure Alcian blue deposition, dry wells were incubated with 1 ml of 6 N Guanidine HCl for 1 hour and then the absorbance was measured with a spectrophotometer measured between 600 and 650 nm.

2.2.9 Protein Extraction and Western Blot Analysis

For histone modifications and KMT2D detection, total protein extracts were obtained as follows. Cells were washed twice with cold PBS, harvested by scrapping in 1 ml cold PBS and centrifuged for 5 minutes at 1500 rpm. Pellet was resuspended in acid buffer (10mM Hepes pH 8, 10mM KCl, 0,1mM MgCl2, 0,1mM EDTA pH 8, 2mM PMSF, 0,1mM DTT) in order to have 10^7 cells/ml. Cells were left at 4°C 10 minutes and then were centrifugated for 10 minutes at 5000 rpm at 4°C. The supernatant (the cytosolic extract) was discarded and the pellet was resuspended in 0,2N HCl in order to have $4x10^7$ cells/ml and left O/N at 4°C on rotating wheel. The day after, proteins were recovered by centrifugation for 10 minutes at 4000rpm at 4°C. Supernatant was recovered and protein concentration was measured with Bradford assay (Biorad #5000006) according to manufacturer's instructions. The absorbance was measured at λ=595 using SAFAS spectrophotometer (SAFAS, Monaco).

For western blots analysis, 20 μg (histone modifications detection) or 50μg (KMT2D detection) of protein samples were boiled and loaded onto a pre-cast Bolt 4-12% Bis-Tris Plus gels (Novex #NW04122BOX) and ran in Bolt MES running buffer (Novex #B0002) (histone modifications detection) or onto a 6% SDS acrylamide gel and run in Tris/Glicine buffer. After electrophoresis, proteins were transferred to a nitrocellulose membrane. Membranes were blocked in PBS-T containing 5% Blotting-Grade Blocker (BIO-RAD #170-6404) (blocking buffer), for 1 h at RT with constant agitation and incubated with indicated primary antibody O/N at 4°C with agitation. The membrane was then washed three times with PBS-T, each time for 5 min, followed by incubation with secondary antibody HRP-conjugated for 1 h at RT. ECL reagents (GE Healthcare #RPN2232) was used to initiate the chemiluminescence of HRP. The chemiluminescent signal was captured using LAS3000 system (GE Healthcare).

TARGET PROTEIN	PRODUCER	CODE
H3	Cell Signaling	9715
H3K4me1	Abcam	ab8895
H3K4me3	Millipore	07-473
H3K27ac	Abcam	ab4729
KMT2D	Bethyl	A300-BL1185
LAMIN B	SantaCruz	sc-6217

2.2.10 RNA Extraction and Expression Level Quantification

Total RNAs were extracted from cells with TRIzol (Ambion #15596018), according to the manufacturer's instructions. Quantitative real-time PCR analysis was performed with SuperScript III One-Step SYBR Green kit (Invitrogen #11746). Relative gene

expression levels were determined using comparative Ct method, normalizing data on endogenous GAPDH.

NAME	SEQUENCE
SOX9-FW	CAGTACCCGCACTTGCAC
SOX9-RE	GCTTCTCGCTCTCGTTCAG
RUNX2-FW	GTGGCCTTCAAGGTGGTAG
RUNX2-RE	TAACAGCAGAGGCATTCCG
ATF4-FW	CTATACCCAACAGGGCATCC
ATF4-RE	GTCCCTCCAACAACAGCAAG
COL II-FW	CAGGATGTCCAGGAGGCT
COL II-RE	GCTTCCACACATCCTTATCATT
AGC-FW	GGAGTCCAACTCTTCAAGGTG
AGC-RE	ATGGTCTGAAGTTTCTACAGTGACA
COL X-FW	CCCAACACCAAGACACAGTT
COL X-RE	GTGGACCAGGAGTACCTTGC
MMP13-FW	GCAGCTGTTCACTTTGAGGA
MMP13-RE	CATCATATCTCCAGACCTGGTT
GAPDH-FW	AGGTGAAGGTCGGAGTCAAC
GAPDH-RE	CCATGTAGTTGAGGTCAATGAAG

2.2.11 Statistical analyses

Statistical parameters are reported in figure legends and include: number of replicates analyzed (n), dispersion and precision measures (mean ± SEM) and statistical significance (p-value). Data have been statistically assessed by one-tailed Student's *t*-test or 2-way ANOVA, when appropriate and as indicated in figure legends. In figures, asterisks means * = $P<0.05$, ** = $P<0.01$, *** = $P<0.001$, ns = not

significant. P<0.05 and lower were considered significant. All experiments were performed in at least triplicate biological replicates. ImageJ software was used for imaging analysis.

Bibliography

1. Tátrai, P. *et al.* Combined introduction of Bmi-1 and hTERT immortalizes human adipose tissue-derived stromal cells with low risk of transformation. *Biochem. Biophys. Res. Commun.* **422,** 28–35 (2012).

2. Micale, L. *et al.* Molecular Analysis, Pathogenic Mechanisms, and Readthrough Therapy on a Large Cohort of Kabuki Syndrome Patients. *Hum. Mutat.* **35,** 841–850 (2014).

3. Calo, E. & Wysocka, J. Modification of Enhancer Chromatin: What,How,andWhy. *Mol. Cell* **49,** 825–837 (2013).

4. Lee, J. K., Hu, J. C. Y., Yamada, S. & Athanasiou, K. A. Initiation of Chondrocyte Self-Assembly Requires an Intact Cytoskeletal Network. *Tissue Eng. Part A* **22,** 318–325 (2016).

5. Panciera, T., Azzolin, L., Cordenonsi, M. & Piccolo, S. Mechanobiology of YAP and TAZ in physiology and disease. *Nat. Rev. Mol. Cell Biol.* (2017). doi:10.1038/nrm.2017.87

6. Barry, F., Boynton, R. E., Liu, B. & Murphy, J. M. Chondrogenic Differentiation of Mesenchymal Stem Cells from Bone Marrow: Differentiation-Dependent Gene Expression of Matrix Components. *Exp. Cell Res.* **268,** 189–200 (2001).

7. Langelier, E., Suetterlin, R., Hoemann, C. D., Aebi, U. & Buschmann, M. D. The Chondrocyte Cytoskeleton in Mature Articular Cartilage: Structure and Distribution of Actin, Tubulin, and Vimentin Filaments. *J. Histochem. Cytochem.* **48,** 1307–1320 (2000).

8. Cooper, K. L. *et al.* Multiple phases of chondrocyte enlargement underlie differences in skeletal proportions. *Nature* **495,** 375–378 (2013).

9. Kozhemyakina, E., Lassar, A. B. & Zelzer, E. A pathway to bone : signaling molecules and transcription factors involved in chondrocyte development and maturation. **142,** 817–831 (2016).

10. Gomez, V. & Hergovich, A. *Cell-Cycle Control and DNA-Damage Signaling in Mammals*. *Genome Stability: From Virus to Human Application* **1,** (Elsevier Inc., 2016).

11. Long, F. Building strong bones: molecular regulation of the osteoblast

lineage. *Nat. Rev. Mol. Cell Biol.* **13,** 27–38 (2011).

12. Wang, W. *et al.* Atf4 regulates chondrocyte proliferation and differentiation during endochondral ossification by activating Ihh transcription. *Development* **136,** 4143–4153 (2009).

13. U, B. D. C. *et al.* Transcriptional mechanisms of chondrocyte differentiation ಜ. 389–394 (2000).

14. Maiques-Diaz, A. & Somervaille, T. C. LSD1: biologic roles and therapeutic targeting. *Epigenomics* **8,** 1103–1116 (2016).

15. Mohammad, H. P. *et al.* A DNA Hypomethylation Signature Predicts Antitumor Activity of LSD1 Inhibitors in SCLC. *Cancer Cell* **28,** 57–69 (2015).

16. Bustin, M. & Misteli, T. Nongenetic functions of the genome. *Science (80-.).* **352,** aad6933-aad6933 (2016).

17. Kumar, A. *et al.* ATR mediates a checkpoint at the nuclear envelope in response to mechanical stress. *Cell* **158,** 633–646 (2014).

18. Kidiyoor, G. R., Kumar, A. & Foiani, M. ATR-mediated regulation of nuclear and cellular plasticity. *DNA Repair (Amst).* **44,** 143–150 (2016).

19. Lin, C.-Y., Chiang, C.-Y. & Tsai, H.-J. Zebrafish and Medaka: new model organisms for modern biomedical research. *J. Biomed. Sci.* **23,** 19 (2016).

20. Corey, D. R. & Abrams, J. M. Morpholino antisense oligonucleotides: tools for investigating vertebrate development. *Genome Biol.* **2,** REVIEWS1015 (2001).

21. Wang, T. Genetic Screens in Human Cells Using. **80,** (2014).

22. Conte, I. *et al.* miR-204 is required for lens and retinal development via Meis2 targeting. *Proc. Natl. Acad. Sci.* **107,** 15491–15496 (2010).

3. Summary, conclusions and future perspectives

3.1 Histone modifying enzymes are important during the development process

Kabuki syndrome (KS; OMIM 147920) is a multiple congenital anomaly disorder with wide and variable clinical spectrum. KS is a rare autosomal disease identified in 1981 whose causative genes have been discovered only in 2011. However, the diagnosis of this disease still remains complicated. The described symptoms are mighty variable both in terms of type of symptoms and in terms of severity. Furthermore, although the causative genes have been identified, in a large cohort of patients no mutation has been identified.

Considering the relatively recent discovery of the causative genes, we still lack molecular insights explaining the etiology and pathogenesis of this disease. The two identified genes, *KMT2D* and *KDM6A*, are both histone modifying enzymes, participating in the formation and activity of the MLL4 complex. Together with DNA methylation, chromatin remodeling, nuclear organization, histone exchange and ncRNA molecules, histone modifying enzymes determine the chromatin state of a cell and, consequently, its identity. Indeed, the establishment of a specific epigenetic state is required for maintaining the undifferentiated state of stem cells. During development, specific extracellular signals are able to induce changes in the epigenetic state of stem cells, inducing differentiation[1]. Considering the importance of maintenance of cell identity, it is not surprising if deregulation of histone modifying enzymes is involved with human diseases. For example, the MLL family members are known to be tumor suppressors in the hematopoietic lineage because inactivating point

mutations (as for *KMT2B* in non- Hodgkin's lymphoma) or chromosomal translocations (as for *KMT2A* in acute myeloid and acute lymphoid leukemias) alters the function of MLL proteins, promoting stem cell propagation rather than differentiation. Also activating and inactivating mutations within EZH2 (a member of the Polycomb-group family responsible for the methylation of H3K9 and H3K27) were recently reported in hematological cancers. Considering KS causative genes, inactivating mutations in *KDM6A* gene were the first cancer-associated mutations identified in a histone demethylase. *KDM6A* mutations were found in a number of cancers, including multiple myeloma and medulloblastoma. Also *KMT2D* has been identified as tumor suppressor in medulloblastoma. In both cases, majority of mutations identified were inactivating mutations, similarly to those mutations found in KS patients. These finding support the hypothesis that deregulation of histone modifying enzymes activity destabilizes the cell capacity to control its fate. In general, the timing of mutation occurrence and the targeted cells could explain the different contribution of the altered function of these chromatin players to the mentioned disease. It has been proposed that if these genetic insults occur in somatic cells during lifespan of individuals, mutations would give rise to cancer. On the other hand, if mutations occur in germ line or during embryogenesis, developmental disorder may arise[2]. In addition to KS, other developmental disorders with mutations in histone modifying enzymes exist. The Rubinstein–Taybi syndrome is caused by a haploinsufficiency or dominant negative functions of CBP (CREB Binding Protein), which is an acetyltransferases targeting both histones and non-histone proteins.

CBP is involved in the transcriptional co-activation of many different transcription factors, thereby influencing gene expression. Interestingly, the symptoms of this disease are similar to KS manifestations: moderate to severe mental retardation, dysmorphic facial features and skeletal abnormalities[3]. Similarly, Kleefstra syndrome also shows severe mental retardation, craniofacial abnormalities and hypotonia. In this case, the molecular cause is a haploinsufficiency of EHMT1 (*KMT1D*), an H3K9 methyltransferase that mediate transcription repression[4]. These findings may indicate that these developmental disorders are not solely due to a reduction in the catalytic activity of a certain enzyme, but it may be that the alteration of the entire chromatin structure disturbs the normal developmental process resulting in multi-organs disease, such as KS.

3.2 MLL4 impairment alter chondrocyte differentiation *in vitro* and *in vivo*

In this thesis we proposed the development of an *in vitro* disease model to understand the molecular causes of KS. Craniofacial abnormalities, short stature and joint laxity are some of the key features of KS patients. The tissues affected by these symptoms are bones, cartilages and tendons. These three tissues have common precursor stem cells, the mesenchymal stem cells (MSCs). For this reason, we decided to develop our *in vitro* model using MSCs. In these cells we aimed to reproduce, through the CRISPR/Cas9 technology, a specific non sense mutation occurring on *KMT2D* gene. We were able to generate two different clones, one with the insertion of a premature STOP codon in both alleles and the other only in one

allele. In both cases, the resulting mutations affected the methyl transferase activity. The presence of a heterozygous ($KMT2D^{+/-}$) and a homozygous clone ($KMT2D^{-/-}$) allows us to study the effects of the complete inactivation of the protein in the homozygous clone but also the effects of the haploinsufficiency in the heterozygous clone, recapitulating the genetic background of KS patients. This is important if we consider that all the known mutations are heterozygous. In this way we can study the effect of total inactivation of MLL4 but, in parallel, we can also correlate the effects of haploinsufficiency with the observed phenotype and, thereby, correlate it with the patient's symptoms.

Apparently, mutated MSCs did not show severe defects when compared to WT cells, in terms of cell growth and stem cell maintenance. However, we detected relevant differences in cell size and shape. Mutated cells resulted smaller, with reduced cell volume and showed less spreading of actin cytoskeleton, indicative of disoriented actin fibers. Nevertheless, we found relevant differences when we induced MSCs to differentiate into chondrocytes. The finding that the MLL4 impairment did not alter MSCs self-renewing potential but rather impinged on differentiation capacity is in line with the notion that MLL4 is not necessary for the maintenance of stem cells state but rather for the acquisition of a specific cell fate[5].

Our data showed that mutated cells were not able to fully differentiate into chondrocyte. When differentiated as micromass, both clones failed to maintain the structure of the micromass and in the production of extracellular matrix. Noteworthy, the heterozygous clone shows an intermediate phenotype between the WT and the homozygous clone.

This finding is important and may give robustness to our model in recapitulating KS disease. It is known that homozygous deletion of *KMT2D* is not compatible with life[6]. Indeed, all known mutations are heterozygous mutation, suggesting that also in human, homozygous mutation of *KMT2D* is lethal. The obtained results suggest that the haploinsufficiency is sufficient to cause alterations during chondrogenesis and that totally loss of the protein activity worsens the observed phenotype.

Another indicator of altered cell lineage commitment is the alteration in timing of cell cycle exit. During differentiation, mutated cells were delayed in exiting from cell cycle, accumulating in G2/M phase and increasing the apoptotic cells. The correct regulation of cell cycle is an important aspect for chondrogenesis. Indeed, different signal pathways (like IHH or FGF signal) control the initiation of the proliferation stage of chondrocytes, the proliferation rate and also the exit from cell cycle that is important for the hypertrophic phase. Deregulation of this control can cause a spectrum of developmental cartilage-related disorders underlying the importance of the control of cell cycle[7].

To corroborate the *in vitro* results we tried to understand the effect of MLL4 impairment *in vivo*. Encouragingly, in medaka animal model morpholino knock-down induced relevant defects in craniofacial development confirming the role of MLL4 in osteo-chondrogenesis.

The defects in cranio-facial development that we observed upon KMT2D down-regulation are in line with the results of a recent work in which mutations of *KDM6A*, the other gene found mutated in KS patients, causes developments defects in mice, including cranio-facial

anomalies. In particular, in the paper the authors define KS as a neurocristopathies, meaning a disease caused by defects in neural crest cells (NCCs) development or migration[8]. As discussed above, KS patients, together with craniofacial defects, show cardiac malformations and mild to moderate brain deficits. From a developmental point of view, all these different apparatus have in common the contribution of NCCs to tissue formations. Indeed, upon specification, NCCs migrate ventrally giving their contribute to the development of heart, sensory ganglia, preosteoblasts and chondrocytes of anterior facial skeleton, melanocytes, dorsal root ganglia and enteric nervous system. It may be that, the alterations that we observed in undifferentiated MSCs are already present in the NCCs precursors. These defects may be responsible for the problems that we observe during the mesenchymal to chondrocytes transition but could also be responsible of additional changing in NCCs behavior. For example, a specific characteristic of NCCs is their migratory capacity. Considering the cytoskeleton alteration that we observe in MSCs, it is considerable the idea that MLL4 defects can also cause defects in the cytoskeleton structure of NCCs thus causing migration problems. It would be interesting to test this hypothesis in NCCs, inducing the same genetic mutation that we introduce in MSCs by adopting the genome editing approach. We could, indeed, understand at which developmental stage, NCCs, MSCs or chondrocytes, the *KMT2D* mutations have the major contribution to the etiology of the KS disease.

3.3 MLL4 impairment alters MSCs morphology and mechanotransduction

Looking at the cellular morphology of chondrocytes during differentiation, we observed a different cell shape between WT and mutated cells. KMT2D$^{+/-}$ and KMT2D $^{-/-}$ cells appear as rounded cells without cytoplasmic protrusions and with less cell-cell and cell-matrix contacts. The chondrogenic differentiation should let MSCs to acquire a cuboidal shape with cortical actin fibers. This is not achieved by mutated cells as they did not show any actin organization at three days of differentiation, thereby indicating early defects in chondrogenesis.

Another indication of a failed differentiation was the reduced production of extracellular matrix, as demonstrated by Alcian blue staining and also by gene expression analyses that show how mutated cells were not able to upregulate genes required for the production of extracellular matrix proteins. These data not only were in line with defects in differentiation, but may also explain the altered cell shape observed in differentiating cells. These cells do not have the "scaffold" needed for the matrix-cellular interaction because they are not able to produce extracellular matrix. Nevertheless, we observed alteration in cell shape and in cytoskeleton organization already in the undifferentiated state. As discussed above, mutated MSCs showed a less organized actin cytoskeleton and a reduced cell spreading. It is more likely that these alterations observed in MSCs are the driver for the impairments of chondrogenesis. Indeed, Lee et al. demonstrated that correct initiation of chondrogenesis requires a functional actin networks in MSCs in order to be able to respond to the cell-cell and cell-matrix contacts[9].

Intracellular signals that induce lineage commitment of MSCs are not only directed by extracellular molecules but also by mechanical stimuli that cells receive. Mechanotransduction permits cell sensing mechanical forces such as the extracellular elasticity and cell geometry, which influence cell lineage specification. Different works show how the substrate stiffness and the cell shape can influence the cell fate choice, overlapping, and in many cases exceeding, the soluble factors.

For example, MSCs differentiation toward osteoblast lineage implies acquisition of a spread cell shape while adipocytes become rounded. It has been demonstrated that, forcing the cell shape, fate determination can be changed. Indeed, MSCs instructed to become adipocytes by chemical signals, differentiate toward osteoblasts lineage if forced to have a spreading shape[10]. In addition, MSCs cultured on substrates with different elasticity show unusual passive differentiation. Mimicking brain elasticity, MSCs are induced to assume the morphology and to increase cell specific markers of neurons. On the other hand, mimicking stiff substrates, typical of osteoblasts environment, forces MSCs to assume the morphology of osteoblasts and to express typical osteoblasts markers[11]. For chondrogenesis, less data are currently available about the influence of cell shape and cell microenvironment on differentiation. Primary chondrocytes forced in monolayer culture show dedifferentiation to a more fibroblastic cell type with reduced expression of collagen type II[12]. Drug impairment of actin cytoskeleton showed that it plays a central role in maintaining the correct structure of chondrocyte morphology. Moreover, drug induction of the classical cortical actin cytoskeleton, with the

acquisition of the typical chondrocyte morphology, favors chondrocytes differentiation[13]. These data support the notion that cell shape is not merely a consequence of differentiation, but it can be the driver for lineage commitment. In this context, it is not surprising that our mutant MSCs, with altered cell shape and mechanotransduction signaling to the nucleus, failed in chondrocyte differentiation. The different cell shape and a different actin cytoskeleton organization of mutated cells could impinge on the capacity of mutated cells to respond to mechanical stimuli. To investigate this aspect, we analyzed the most relevant nuclear sensors of altered mechanotransduction, YAP and TAZ. They are transcriptional cofactors that shuttle between the cytoplasm and the nucleus where they associate with several promoter-specific transcription factors. It has been demonstrated that different extracellular elasticity but also different cell shape imply different localization and activity of YAP/TAZ. With stiff matrix, such as that one produced by osteoblasts, YAP/TAZ are nuclear and transcriptionally active. In adipocytes, with more elastic matrix, YAP/TAZ are no more nuclear and degraded in the cytoplasm. The same behavior is present when cells assume a spread shape (nuclear localization, active) or a rounded shape (cytoplasmatic localization, inactive)[14]. In WT MSCs, YAP/TAZ are nuclear, as expected by their spreading shape. Interesting, both heterozygous and homozygous clones show decrease in nuclear YAP/TAZ, resembling the pattern of YAP/TAZ on soft matrix. This may indicate that mutated cells were unable to sense the extracellular environment and, consequently, to sense the mechanical cues, mandatory for chondrogenesis.

3.4 Total or partial loss of MLL4 alters the expression of chondrogenic TFs

A differentiation process is orchestrated by different transcription factors (TFs), including signal-dependent TFs, which are responsive to extracellular cues, and lineage-specifying TFs, that drive cell specification and identity. Analyzing some TFs relevant for chondrogenesis, we noticed alterations of their expression in mutated MSCs. Indeed, $KMT2D^{+/-}$ and $KMT2D^{-/-}$ MSCs down regulated SOX9 and RUNX2, two TFs that should be expressed in MSCs. On the other hand, they expressed higher levels of ATF4, a late chondrocyte TF. These results may indicate a deregulation of the transcriptional program. As consequence, mutated cells were unable to induce the expression of the extracellular matrix proteins, which are direct targets of these TFs. These findings indicate that, also at the molecular level, mutated MSCs are unable to activate the proper transcriptional program that should let MSCs to differentiate correctly towards chondrocytes.

The changing in the expression pattern of these TFs could rely on the reduced enzymatic activity of MLL4. As expected, in mutated clones we observed a decrease in the H3K4me1 levels together with a decrease of H3K27ac. These two histone modifications are considered markers of active enhancers. We postulated that MLL4 impairment causes the inactivation of those enhancers required for the establishment of the correct transcriptional program. This hypothesis is further supported by the finding that the other causative gene of KS is *KDM6A*, a H3K27 demethylases that favors the onset of H3K27ac on cis-regulatory elements. MLL4 and KDM6A work together in a

multiprotein complex to induce the activation of poised enhancers: MLL4 deposits the monomethylation while KDM6A induce the demethylation of H3K27, thus facilitating the acetylation by CBP/p300. It has been demonstrated that MLL4 has not only the methyltransferase activity, but it is also needed for the stability of some component of the complex, such as KDM6A[15]. In this regard, it is convincible that MLL4 haploinsufficiency not only impairs H3K4me1 levels but also alters the activity of the ASCOM complex preventing the activation of those enhancers important for chondrogenesis. This hypothesis is supported, at least in part, by our data showing the possibility of rescuing chondrogenesis by inhibiting the H3K4-specific histone demethylase LSD1. Even if more analyses are required, the fact that, inhibiting the removal of methyl group from H3K4 partially rescues the chondrocyte differentiation indicates that the observed phenotype is due, at least in part, to the decreased H3K4me1 level.

3.5 Conclusions and future perspectives

Although the hypothesis that MLL4 impairment causes alteration in the activation of specific chondrogenic enhancers is reasonable, it could simplify a more complex situation. Indeed, different works demonstrate how the activity of MLL4 is not only due to its methyltransferase activity because there is only a minor effect on transcription upon its depletion[16]. Nevertheless, MLL4 has been described as modulator of enhancers and promoter interaction[5], beside having a critical function for the maintenance of genomic stability[17]. Considering these premises, we advanced the hypothesis that

deregulation of histone modification deposition, induced by MLL4 loss, has minor effect on activation of enhancers but, instead, affects the chromatin structure. In addition to being the site for storage of genetic material and gene transcription, the nucleus, which is generally the stiffest element of all eukaryotic cells, plays a central role in mechanical strain transduction[18]. The cytoskeleton forms a mechanically continuous network within the cell and transmits extracellular mechanical signals from sites of matrix adhesion to the nucleus that is able to adapt its conformation to the different stimuli. At the same time, changes in nuclear architecture might also impact on how forces are transmitted through the cell and how cell responds to extracellular signals[19]. Indeed, Heo et al. show how increasing of nuclear stiffness implies changes in mechanical forces inside a cell, making cells more sensitive to mechanical stimuli[18]. Among other elements, chromatin state determines the structure, the strength and the size of the nucleus. For example, during mitosis, the nuclear membrane disassembles and reassembles in telophase. During this stage, chromatin serves as a physical scaffold on which the nuclear membrane fragments are immobilized, facilitating their fusion into larger membrane sheets[20]. Another example is the targeting of compact heterochromatin to the nuclear periphery. The peripheral sequestration of heterochromatin may enhance the structural robustness of the nucleus and strengthen its ability to withstand physical challenges, such as mechanical forces exerted during cell migration or in mechanically active tissues[21,22]. Also the condensation of chromatin can influence nuclear characteristics. Indeed, chromatin decondensation weakens the ability of the nucleus to resist mechanical

force, beyond the structure of nuclear envelope. Indeed down-regulation of the Prdm3 and Prdm16 methyltransferases (which facilitate heterochromatic histone H3 Lys9 methylation and promote heterochromatinization) or reducing the ability of the linker histone H1 to bind to nucleosomes and thus reducing the level of heterochromatin, cause changes in nuclear shape and strength. These findings suggest that chromatin by itself contributes significantly, and directly, to determining the mechanical properties of nuclei[19], thereby influencing the mechanotransduction signaling. In this scenario, we may think that the global decrease of H3K4me1 and of H3K27ac causes a more heterochromatic state, favoring chromatin condensation and nuclear shrinkage. This can cause changing in nuclear architecture and consequently its stiffness. The change in nuclear stiffness may imply a global change in the cell structure (as we observed), in mechanical forces of cells and, consequently, a change in the mechanotransduction (as we observed). This function of the genome may be especially relevant for those cells exposed to mechanical stress, such as cardiomyocytes and migrating cells. Transgenic mice that over express HMGN5 (an architectural chromatin protein known to promote chromatin decompaction) only in the heart, are born with cardiomyocytes lacking heterochromatin, yet their heart and cardiomyocytes appear normal. However, most animals die within 3 months of birth, showing marked cardiac hypertrophy with extremely large cells and nuclei. Loss of heterochromatin gradually diminishes the ability of the nucleus to withstand the mechanical forces, such as of the continuously beating heart[21]. This aspect may be important in the picture of KS. Indeed, affected tissues seem to be those tissues

subjected to mechanical stress (bone and cartilage, hearth, muscle) or to migration (tissues derived from the neural crest cells).

A first but not conclusive indication of an altered nuclear mechanics induced by MLL4 impairment is the partial rescue that we observed with treatment with ATR inhibitor. Although ATR is considered to be a DNA damage response kinase, it is clear that it plays a relevant role in a variety of cellular aspects independent of DNA damage. Indeed, recently ATR has been shown to respond to a variety of stress stimuli, like mechanical stress, osmotic stress and thermal shock. Mechanical and chemical alterations of nuclear envelope (NE) are able to activate ATR[23]. During mitosis, condensing chromatin impose topological stress to the NE. Through an unknown mechanism, ATR can sense these forces, relocates on the nuclear envelope and locally phosphorilate CHK1[24] and presumably other yet undefined targets. In this view, ATR can be considered as a transducer of nuclear mechanical forces. In our mutated cells, the altered chromatin structure could cause alteration on the stiffness and elasticity of the NE. These modifications could be sensed by ATR that, through unknown mechanisms and among other factors, could alter the chondrogenic potential. In this view, the inhibition of ATR rescues the differentiation capacity, relaxing a stalled situation that prevented the chondrogenic differentiation.

In conclusion, here we developed a cellular model for KS. In our model, impairment of MLL4 causes differentiation defects for the chondrocyte lineages. Even if further studies are required to confirm the validity of this model, we could consider it a relevant disease model system that permit to gain insights the molecular mechanisms

involved in the etiology of this rare genetic disease. We focused our attention on a specific lineage which resulted altered upon *KMT2D* mutation in the model system and its perturbation could explain some of the clinical features affecting KS patients. Furthermore, cranial chondrocytes belong to NCCs whose involvement in the pathology begins to be elucidated. Moreover, the medaka-based *in vivo* model confirmed the relevance of these results as the knock-down of *KMT2D* caused defects in the development of cranio-facial structure. The consistency between the *in vitro* and the *in vivo* model give us the possibility to bring forward our hypothesis about the molecular mechanisms that causes KS, test them in the cellular model and then confirm it in the adopted animal model. Understanding the molecular mechanisms is fundamental for the development of therapeutic approaches that could be tested *in vitro* and confirmed *in vivo*.

Studying the pathogenesis of KS is not trivial because of the variable spectrum of symptoms that the patients show. There is not a clear cellular type or a specific pathway that, if altered, may explain all the symptoms. This is not surprising considering the causative genes of KS, which codify for histone modifying enzymes that act globally on the chromatin. Considering their role in activation of lineage specific enhancers, one possible mechanism of action could involve the impairment of gene activation thus leading to alteration of commitment of specific lineages, inducing developmental defects. However, this is not an exhaustive explanation. It has been demonstrated that depletion of MLL4 enzymatic activity has minor effect on global transcription. For these reasons, we hypothesize a more global effect of MLL4 impairment. MLL4 mutant cells showed

a decreased level of H3K4me1 and H3K27ac. Considering that these histone modifications are typical of an open chromatin state, we postulate that MLL4 alteration could cause a more condensed, heterochromatic state that alters the nuclear structure and stiffness. These nuclear alterations could be transduced to all the cellular structure including the actin cytoskeleton, causing defective mechanotransduction signals. If this holds true, cells that are more sensitive to mechanical stimuli or that are forced to mechanical stress may be more damaged by this type of alteration, including chondrocytes. This assumption is in line with the relatively new view of the chromatin that, not only controls the gene expression, but also participates to control the nuclear, and consequently, cell architecture.

Bibliography

1. Lee, T. F. *et al.* Chromatin Regulatory Mechanisms in Pluripotency. *Annu Rev Cell Dev Biol.* **223,** 648–657 (2011).

2. Distefano, M. D. Histone-modifying enzymes: regulators of developmental decisions and drivers of human disease. *Epigenomics* **4,** 213–223 (2012).

3. Petrij, F. *et al.* Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. *Nature* **376,** 348–351 (1995).

4. Kleefstra, T. Disruption of the gene Euchromatin Histone Methyl Transferase1 (Eu-HMTase1) is associated with the 9q34 subtelomeric deletion syndrome. *J. Med. Genet.* **42,** 299–306 (2005).

5. Wang, C. *et al.* Enhancer priming by H3K4 methyltransferase MLL4 controls cell fate transition. *Proc. Natl. Acad. Sci.* **113,** 11871–11876 (2016).

6. Lee, J. *et al.* H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. *Elife* **2,** 1–25 (2013).

7. Kozhemyakina, E., Lassar, A. B. & Zelzer, E. A pathway to bone : signaling molecules and transcription factors involved in chondrocyte development and maturation. **142,** 817–831 (2016).

8. Shpargel, K. B., Starmer, J., Wang, C., Ge, K. & Magnuson, T. UTX-guided

neural crest function underlies craniofacial features of Kabuki syndrome. *Proc. Natl. Acad. Sci. U. S. A.* **114,** E9046–E9055 (2017).

9. Lee, J. K., Hu, J. C. Y., Yamada, S. & Athanasiou, K. A. Initiation of Chondrocyte Self-Assembly Requires an Intact Cytoskeletal Network. *Tissue Eng. Part A* **22,** 318–325 (2016).

10. McBeath, R., Pirone, D. M., Nelson, C. M., Bhadriraju, K. & Chen, C. S. Cell shape, cytoskeletal tension, and RhoA regulate stemm cell lineage commitment. *Dev. Cell* **6,** 483–495 (2004).

11. Engler, A. J., Sen, S., Sweeney, H. L. & Discher, D. E. Matrix Elasticity Directs Stem Cell Lineage Specification. *Cell* **126,** 677–689 (2006).

12. Mark, K. V. O. N. D. E. R. & Mark, H. Immunological and biochemical studies of collagen type transition during in vitro chondrogenesis of chick limb mesodermal cells. *J. Cell Biol.* **73,** 736–747 (1977).

13. Woods, A., Wang, G. & Beier, F. RhoA/ROCK signaling regulates Sox9 expression and actin organization during chondrogenesis. *J. Biol. Chem.* **280,** 11626–11634 (2005).

14. Dupont, S. *et al.* Role of YAP/TAZ in mechanotransduction. *Nature* **474,** 179–184 (2011).

15. Herz, H. *et al.* Enhancer-associated H3K4 monomethylation by Trithorax-related , the Drosophila homolog of mammalian Mll3 / Mll4. **2,** 2604–2620 (2012).

16. Dorighi, K. M. *et al.* Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from Promoters Independently of H3K4 Monomethylation. *Mol. Cell* **66,** 568–576.e4 (2017).

17. Kantidakis, T. *et al.* Mutation of cancer driver MLL2 results in transcription stress and genome instability. **4,** 408–420 (2016).

18. Heo, S. J. *et al.* Differentiation alters stem cell nuclear architecture, mechanics, and mechano-sensitivity. *Elife* **5,** 1–21 (2016).

19. Bustin, M. & Misteli, T. Nongenetic functions of the genome. *Science (80-.).* **352,** aad6933-aad6933 (2016).

20. Wandke, C. & Kutay, U. Enclosing chromatin: Reassembly of the nucleus after open mitosis. *Cell* **152,** 1222–1225 (2013).

21. Furusawa, T. *et al.* Chromatin decompaction by the nucleosomal binding protein HMGN5 impairs nuclear sturdiness. *Nat. Commun.* **6,** 1–10 (2015).

22. Gerlitz, G. & Bustin, M. The role of chromatin structure in cell migration. *Trends Cell Biol.* **21,** 6–11 (2012).

23. Kidiyoor, G. R., Kumar, A. & Foiani, M. ATR-mediated regulation of nuclear and cellular plasticity. *DNA Repair (Amst).* **44,** 143–150 (2016).

24. Kumar, A. *et al.* ATR mediates a checkpoint at the nuclear envelope in response to mechanical stress. *Cell* **158,** 633–646 (2014).