SAPIENZA
UNIVERSITÀ DI ROMA

ARID1A supports the expression of YAP target genes by facilitating the binding of TEAD on enhancers in epithelial breast cells

PhD program in "Human Biology and Medical Genetics" – XXXVII Cycle

Candidate:
Francesco Greco

Tutor:
Prof. Valerio Fulci

Coordinator:
Prof.ssa Laura Stronati

A.A.
2023-24

Contents

1. Introduction

1.1 Co-transcriptional factor YAP

The emergence of multicellular organisms on our planet was the fundamental starting point for the development of today's biological diversity. The organisation of multicellular organisms into tissues and organs required the evolution of regulatory mechanisms to control the proliferation and differentiation of their component cells[1]. Despite the physiological and evolutionary importance of this mechanism, the superior structuring into organs and tissues has been a real puzzle for scientists for several years. The discovery of the Hippo pathway has represented one of the first steps towards understanding of these mechanisms[2,3].

Controlling the growth of internal organs is the best-known function, but not the only one that can be traced back to the Hippo pathway. In fact, the latest research has established its involvement in many fundamental biological processes, from the control of individual cell fate[4] to the regulation of tissue architecture[5]. It is therefore not surprising that the Hippo pathway is conserved in many species, from the arthropods in which it was discovered to mammals[6].

In contrast to more conventional signalling pathways linked to a specific ligand-receptor pair, the Hippo pathway is capable of responding to a wide range of biochemical and mechanical signals[5]. In response to these stimuli, Hippo activation negatively regulates the activity of its downstream target YAP and its paralog TAZ[2,7]. YAP and TAZ are transcriptional coactivators therefore their actions are localised in the nucleus. To inhibit these activities, Hippo components alter the phosphorylation state of YAP and TAZ proteins, inducing their cytoplasmic retention and degradation[8].

In order to drive the expression of their target genes, YAP and TAZ must enter the nucleus in a dephosphorylated form; however, translocation to the nucleus is not sufficient to activate their transcriptional programme. In fact, both YAP and TAZ lack a DNA-binding domain and therefore their recruitment on chromatin depends on their interactions with DNA-binding factors[9]. Although several transcription factors have been identified that can interact with YAP[10], members of the TEA family (TEADs 1-4) represent the interactors involved in the majority of YAP-mediated responses[10–13]. Structural studies of these proteins shed light on the mutualistic relationship between the co-activator YAP and these transcription factors. In fact, unlike other transcription factors, TEADs 1-4 do not have a trans-activating domain[14]. YAP, on the other hand, does[9]. To carry out their transcriptional role, TEADs therefore need to bind YAP and thus they share a common transcriptional programme[15].

A Gene Ontology (GO) analysis of the genes to whose promoter the YAP/TAZ-TEAD complex binds reveals an enrichment of terms such as regulation of cell migration, organisation of the extracellular matrix and regulation of cell proliferation[16].

The study of YAP and TAZ has revealed a predominant role in regulating growth in the early stages of embryonic development[3,8,11,12] as well as in later stages in different tissues[3,8,12]. YAP/TAZ activity is typically regulated by the Hippo pathway in healthy adult tissues. Any tissue damage induces the activation of a repair and regeneration response in which YAP/TAZ play a key role[3,11,12].

1.2 YAP is regulated by multiple *stimuli*

A wide variety of external and internal cues can modulate YAP activity in a way that may involve or not the activation of the Hippo pathway[17] . Hippo-activating stimuli include cell-cell adhesion, basal-apical polarity, mechanotransduction

and cytoskeletal activity, cell density, EMT and G-protein-coupled receptor signaling[8,17–19]. In addition to the above, YAP/TAZ activity can be regulated by more classical extracellular signalling pathways, including WNTs and ligands that activate the interleukin-6 receptor subunit beta[17].

As mentioned above, the activation of the Hippo pathway involves the phosphorylation cascade -mediated by MST- and LATS kinases, resulting in the functional inhibition of YAP[20,21] and the PDZ-domain transcriptional coactivator (TAZ)[22](Fig. 1).

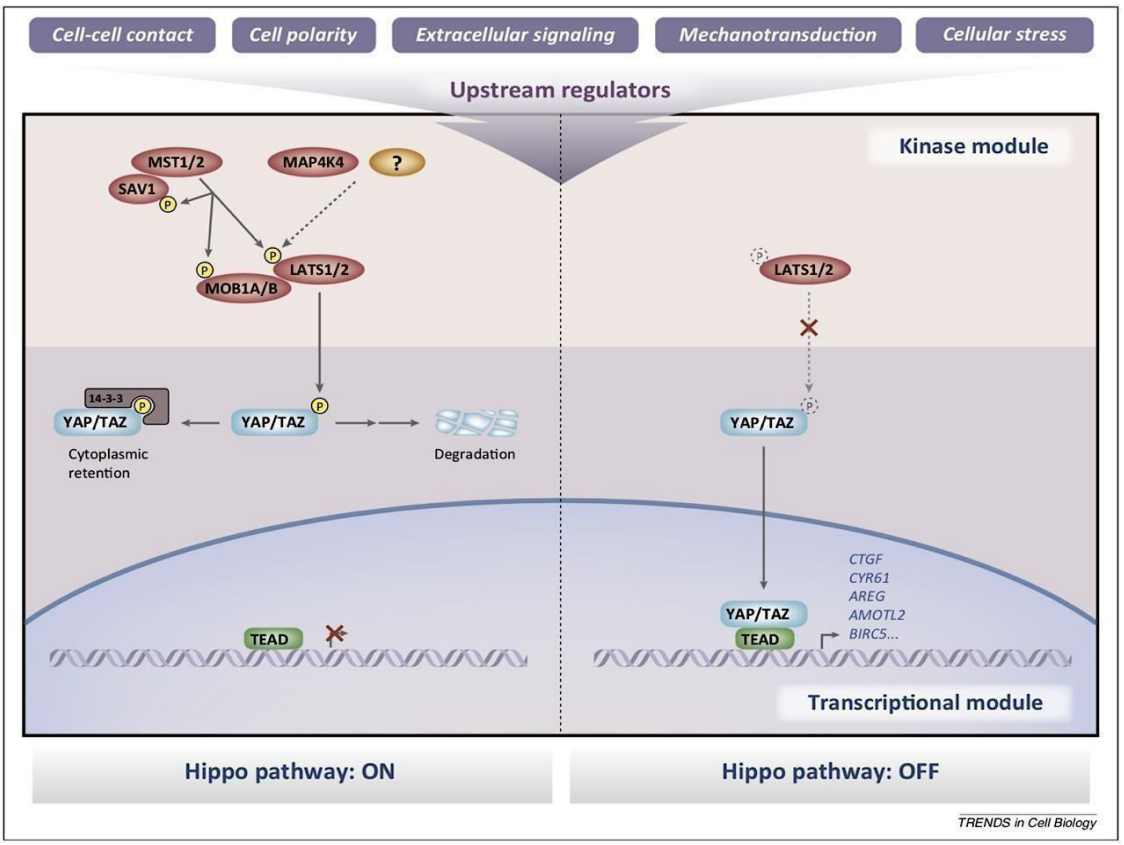

Fig. 1 Hippo Pathway (Hansen et al 2015[23])

More specifically, YAP is directly phosphorylated by LATS1/2 on five serine residues lying within the consensus sequence HxRxxS[24]. In support of this evidence, the introduction of alanine substitutions at all five serine residues (YAP-5SA mutant) results in the complete abolition of LATS2 and MST2 activity on YAP[25]. The phosphorylation of S112, in particular, allows the exposure of a

binding consensus for the 14-3-3 protein, thereby inducing the relocalisation of YAP to the cytoplasm or proteasome-mediated degradation[9,26]. This phenomenon has also been observed in TAZ[27].

The model whereby the downstream binding with 14-3-3 protein would be sufficient to induce relocalisation has several critical points, especially if certain experimental evidence[8] is taken into account. In particular, immunofluorescence studies utilising antibodies specific for the phosphorylated YAP demonstrated nuclear retention in mouse embryonic cells (NIH3T3) grown at low confluence. This is despite the fact that this condition should maintain predominantly Yap non-phosphorylated and localised in the nucleus[28].

Further insights were provided by transgenic mouse models capable of inducibly expressing the mutant form of Yap (S127A). It has been demonstrated that this mutant form displays an increased propensity to localise at the nuclear level, rendering it an optimal candidate for investigations into organ growth processes involving Yap[7]. In contrast, it was observed that the localisation of Yap (S127A) in the intestinal epithelium of transgenic mice is not restricted to the nucleus. This suggests that phosphorylation of Yap at this site is not the primary determinant of its subcellular localisation, at least in the intestine[29]. In their commentary on this topic, Piccolo and colleagues hypothesise the existence of alternative phosphorylation sites involved in binding to other complexes. Moreover, alternative mechanisms by which phosphorylation by LATS1/2 affects Yap1 are also conceivable[30].

In support of Piccolo's comment, AURORA kinase has been found to phosphorylate YAP at serine 397 (firstly annotated as S381 by Zhao and colleagues[31]) in the nucleus, thereby promoting its tumorigenic transcriptional programme in a triple-negative breast cancer model (MDA-MB-231 cell line)[32].

Consistent with the opposite effects of these chemical modifications on YAP, a novel mechanism of phosphorylation has recently been discovered that confers on YAP the ability to support stemness properties in cancer stem cells. Indeed, CDK7 has been shown to interact with and phosphorylate YAP at S127 and S397 in the nucleus, both of which are required to enhance its transcriptional activity[33]. It should be noted that both of these serine residues have been reported as phosphorylation sites for the LATS kinase activity, which determines the Hippo negative control on YAP through cytoplasmic retention and subsequent degradation[31].

A substantial body of evidence from numerous studies has demonstrated the involvement of other kinases in regulating the same residues targeted by LATS in the control of YAP1[8]. In fact, the first kinase that was reported to phosphorylate YAP was Akt. Co-purification studies in human cell lines demonstrated that AKT can phosphorylate the S127 residue of YAP, thereby inducing binding to 14-3-3. Therefore, AKT induces the cytoplasmic localisation of YAP, thereby preventing the latter from co-activating the pro-apoptotic factor p73[26].

In light of the aforementioned data, it can be posited that phosphorylation represents an ever-present event in the cytoplasmic translocation of YAP. Nevertheless, it may represent a non-primary event in the inactivation of YAP/TAZ[8]. It can be reasonably assumed that any upstream mechanism would not impede phosphorylation by LATS1/2, given that this event can occur in both the cytoplasm and the nucleus. Piccolo et al. postulate that the inhibitory effects exerted by LATS1/2 may serve as a mechanism to reinforce an upstream regulatory process that is not necessarily associated with the Hippo pathway[30]. As hypothesised by several authors, it is possible that phosphorylation at the same sites but in different compartments by different kinases could provide a

complex mechanism to regulate YAP activity based on cell context and in response to multiple cues[32,33].

It is important to note that serines are not the only target residues on YAP. Threonines, which are phosphorylated by kinases such as c-Abl and Yes/Src, also play a role in the regulation of YAP[34]. In the former case, phosphorylation on Y357 by c-Abl induces a stabilisation of the YAP protein which in turns appears to determine the binding affinity for the p73 factor in the first place. Furthermore, the authors observed a phosphorylation-dependent alteration in the transcriptional activity of YAP. Specifically, under normal conditions, YAP binds p73 on the p21 gene promoter rather than the Bax promoter. In response to DNA damage, YAP-p73 has been observed to support Bax expression. The key difference between the two conditions is the level of phosphorylation of YAP. In the absence of genotoxic stress, YAP is predominantly unphosphorylated, whereas following DNA damage, phosphorylation occurs. The resulting mechanism thus involves a change in the binding site of the YAP-p73 complex, such that the expression of pro-apoptotic factors is only observed in response to DNA damage[35].

Indeed, the threonine kinase activity of Yes/Scr has been demonstrated to trigger a transcriptional programme with an opposite effect to that verified with c-Abl. In their study, Rosenbluh et al.[34] initially demonstrated that YAP can form a complex with the transcription factor TBX5 and β-catenin. The formation of this complex is independent of the Yes1 kinase activity. Conversely, it would appear that phosphorylation is responsible for the activation of the YAP-TBX5-□-catenin complex, thereby enabling it to bind to the promoters of specific genes that are involved in the process of survival[34].

It should be noted that post-translational modifications are not the only mechanism capable of regulating the localisation and thus the activity of YAP. In

between these are the classical ligand-receptor signalling pathways that are independent of Hippo, including WNTs, ligands that activate G protein-coupled receptors (GPCRs) and the beta subunit of the interleukin-6 receptor.[17]. With regard to WNT, a complex crosstalk with YAP and TAZ was observed[17]. In fact, in cells in a "WNT OFF" state, they participate in the destruction complex by interacting with Axin1 and thus remain anchored in the cytoplasm. In this state, the destruction complex is fully able to carry out its main function of degrading beta-catenin. The nuclear relocalisation of YAP/TAZ can be achieved by treatment with Wnt ligands or by removing components of the complex. In turn, YAP is able to regulate WNT signalling since its nuclear relocalization makes the destruction complex unable to exert its negative activity on beta-catenin[36].

1.3 The role of YAP in cancer

To date, it is evident that the physiological activity of YAP is functionally reprogrammed to aberrantly support cancer cell proliferation, progression, migration and metastasis[8].

Numerous studies of YAP/TAZ expression and nuclear localisation have shown that the expression or activity of these co-factors are enhanced in the vast majority of tumour samples compared to normal tissues[3,37,38]. These studies demonstrated a correlation between altered expression or activation status of YAP and poor prognosis. In this regard, a meta-analysis of more than 21 studies involving thousands of patients demonstrated that over-expressed and activated YAP is associated with reduced survival[39]. In certain instances, the overexpression or increased activity of YAP has been demonstrated to be contingent upon genetic alterations involving components of the Hippo pathway or amplifications of YAP[40]. Indeed, the chromosomal region (11q22) encompassing the YAP locus is subject to amplification in a multitude of human

cancers, as evidenced by several studies[40–51]. Moreover, mutations in the core components of the Hippo pathway have been frequently observed in cancerous tissues[52–56]. Nevertheless, the occurrence of genetic alterations is not sufficiently prevalent to account for the phenomenon of overexpression and increased activity of YAP/TAZ in tumours[57].

Altered YAP expression exerts its influence at various levels in the promotion of tumour growth and disease progression. YAP is involved in a number of cancer-related biological processes, including cell proliferation, apoptosis, stem cell characteristics and the tumour microenvironment[38].

1.4 SWI/SNF chromatin remodelling complex

In eukaryotic cells, the nuclear DNA is wrapped around the histone octamer in order to assemble into nucleosomes, which represent the fundamental unit of chromatin[58]. The formation of higher-ordered structures in the nucleus, known as topologically associated domains (TADs), is determined by hierarchical folding and interactions[59,60]. The high constriction observed in the nucleus is subject to dynamic remodelling, which ensures accessibility to a vast number of factors involved in gene expression throughout the lifespan of cells[61,62]. Furthermore, chromatin remodelling is crucial for its assembly, DNA replications and DNA repair[63–65].

The actions of histone modifying enzymes[66] and ATP-dependent chromatin remodelling complexes[67] are synergistic, with both regulating these processes in conjunction with one another. The first type of enzyme acts by affixing various post-translational modifications on specific residues of the histone tail, thereby generating an extensive histone code comprising acetylation, methylation, phosphorylation, SUMOylation and ubiquitination[68,69].

ATP-dependent chromatin remodelling complexes are capable of reading histone modifications and utilising this information to facilitate site-specific unwrapping, mobilisation, exchange or ejection of nucleosomes through the process of ATP hydrolysis[70–72].

Actually, remodelers are classified in 4 distinct families (SWI/SNF, ISWI, NURD/Mi-2/CHD and INO80), based on the presence of additional domains on their ATPase domain[73].

The SWI/SNF genes were initially identified in yeast through genetic and biochemical screening as these mutations affect the mating-type switching (SWI)[74] and sucrose fermentation (Sucrose Non-Fermenting-SNF) pathways[75]. Subsequent biochemical analysis led to the characterization of the proteins encoded by these genes, revealing their capacity to assemble in large complexes[76]. The significance of these complexes is evidenced by their evolutionary conservation across diverse organisms, from *Drosophila*[77,78] to mammals[79].

In mammals, 29 genes encode more than 15 different subunits that can assemble in three distinct SWI/SNF complexes: canonical BRG1/BRM-associated factor (BAF), polybromo-associated BAF (PBAF) and ncBAF complexes. The purification of these subunits has revealed a cell line-specific combination that gives rise to a heterogeneous range of complexes[80].

The significance of the diverse array of complexes became evident in genetic studies of pivotal processes such as embryogenesis and cell fate determination[81–85]. Consequently, in vertebrates, the non-redundancy and combinatorial assembly of these subunits may have contributed to the diversification of gene expression patterns by a restricted set of genes.[86].

Despite the observed differences in composition, the mammalian SWI/SNF complexes exhibit a distinctive functional structure comprising three interacting

modules: BAF core, ATPase module and an accessory module which confers target specificity[87,88].

The accessory module may include three related but distinct proteins which provide an AT-rich interaction domain to the entire complex: protein 1A (ARID1A), protein 1B (ARID1B), protein 2 (ARID2)[87]. The incorporation of these products into the remodelling complex is mutually exclusive, thus defining three biochemical distinct forms of SWI/SNF[89]. Specifically, ARID1A and ARID1B define BAF-A and BAF-B complexes respectively. Conversely, ARID2 is found in the PBAF (Polybromo BAF) complex[87].

The ARID1A and ARID1B proteins are essential for the complete assembly of the BAF core module and for establishing a connection with the subunit with ATPase activity module[87]. Both proteins are expressed in cells, although their presence in the complex is mutually exclusive[89]. The resulting complexes exhibit minimal overlap in their genomic localization and exert distinct effects on gene transcription through specific interactions with different transcription factors[89].

1.5 ARID1A

ARID1A (also known as BAF250a, p270 or SMARCF1) is a key component of the mammalian SWI/SNF complex. In contrast to its yeast counterpart (SWI1), ARID1A exhibits high affinity for DNA, albeit without sequence specificity[90]. Nevertheless, ChIP-seq analysis in endometrial cancer cells revealed a correlation between the enriched motifs in ARID1A target peaks and the DNA binding motifs for JUNB, TEAD4, RUNX1, EWS:ERG, and NF1[91].

Further studies lend support to the notion that ARID1A plays a pivotal role in facilitating the co-localisation of SWI/SNF complexes with a number of transcription factors. For example, in breast cancer cells ARID1A is essential for

the co-localisation of SWI/SNF with FOXA1 on ER *cis*-regulatory elements[92]. Furthermore, in a colon cancer cell line, accessible regions dependent on ARID1A/B are bound by AP-1 subunits[93]. In the same study Kelso and colleagues demonstrated that ARID1A plays a crucial role in regulating global chromatin accessibility. Furthermore, the impact of ARID1A on chromatin accessibility is especially notable in regions designated as enhancers.[93].

Higher metazoans have evolved a multitude of sophisticated mechanisms to precisely regulate gene expression in essential biological processes. One such sophisticated method is the pausing of RNA Pol II[94,95]. The primary factor responsible for RNA pol II pausing is negative elongation factor (NELF), which impedes the effective elongation of RNA Pol II in the early stage of transcription[96]. Subsequently, other factors were identified as regulators of RNA Pol II pausing[97,98], including ARID1A[99]. It is noteworthy that ARID1A directly promotes the accumulation of paused RNA Pol II rather than the establishment of an active state within the chromatin[99].

As has been previously reported, several components of the SWI/SNF core complex are essential for early embryogenesis[82,100,101]. Gao and colleagues demonstrated that ARID1A is required for mouse embryogenesis and for the function of embryonic stem (ES) cells. Indeed, heterozygous mutant embryos are arrested at a late stage of development whereas complete loss of ARID1A determines developmental arrest at an early stage[102].

Moreover, ARID1A plays a direct role in maintaining DNA stability and facilitating repair processes[103,104].

In particular, ARID1A plays a crucial role in the recruitment of Topoisomerase IIa (TOP2A). Indeed, even though ARID1B globally replaces ARID1A function, still ARID1A depleted cells continue to exhibit genomic instability due to the failure to resolve R-loops by TOP2A on prone sites[104].

Shen and colleagues delineated a functional interplay between ARID1A and Ataxia telangiectasia and RAD3-related protein (ATR)[103]. ATR along with ataxia telangiectasia-mutated (ATM) are two kinases involved in the control of cell responses to DNA damage. The downstream effects following activation of these kinases are determined by a decision three in which the engagement of different factors determine multiple outcomes; although all share the purpose of preserving genomic integrity through cell-cycle delay and recruitment of DNA repair machinery [105–107].In order to gain insight into the mechanisms that regulate ATR activity, Shen et al. conducted a proteomic assay which identified ARID1A as a direct interactor of ATR. This interaction results in the recruitment of SWI/SNF complex which in turn induces the DNA damage checkpoint and promotes DSB end resection[103].

1.6 ARID1A in cancer

Over the course of several decades, the role of epigenetic abnormalities in tumorigenesis was fully accepted by the scientific community. [108]. The expeditious implementation of omics methodologies enabled researchers to delineate a conceptualised "cancer epigenome". This modified epigenetic landscape is characterised by alterations in DNA methylation, histone code and chromatin accessibility, which have a profound impact on gene expression in cancerous disease[109]. The overwhelming majority of epigenetic alterations have their origin in genetic mutations affecting key factors that regulate chromatin state[110]. As the number of tumour genomes sequenced increased, the number of identified "driver" mutations also rose. The term "driver mutations" is used to describe genetic alterations that are causative of oncogenesis. A salient feature of driver mutations is their high prevalence across a multitude of tumour types, with a particular propensity for certain cancers [111].

In the pre-omics era, initial studies of malignant cells demonstrated a genetic loss of subunits belonging to the SWI/SNF complex. It is the case of BAF47, which is lost in a rare childhood cancer known as malignant rhabdoid tumour (MRT)[112]. Further observation in adrenocortical carcinoma and several other cancer cells evidenced the inactivation of both BRG1 and BRM, which suggests that the SWI/SNF complex may have a tumour suppressor function [113,114].

Of the proteins that comprise SWI/SNF complex, ARID1A is the most frequently mutated in cancers[115]. To date, the AACR GENIE project has highlighted that the frequency of mutations affecting ARID1A is approximately 10% across all tumour types analysed[116,117]. It is anticipated that this percentage will increase significantly when data from exclusively ovarian and endometrial cancers are included. Indeed, 57% of ovarian clear cell carcinoma (OCCC) cases analysed exhibited inactivating mutations in ARID1A[118]. Additionally, ARID1A mutations were observed in 30% of endometrioid carcinomas. Considering the findings by Wiegand and colleagues regarding the presence of mutations and transcript absence in preneoplastic lesions, the hypothesis has been put forth that ARID1A loss represents an early event in the transformation of endometriosis into cancer[119]. The hypothesis is further supported by the more recent characterisation of cancer driver genes and mutations, which also extend this trigger function to breast cancers[120].

ARID1A mutations are distributed throughout the entire gene body without any evidence of clustering. The genetic alterations predominantly comprise frameshift mutations which are predicted to result in a loss of function or alternative splicing[115].

The multiple studies presented thus far concur in identifying ARID1A as a tumour suppressor gene, although some evidence appears to indicate the contrary in specific circumstances. Indeed, the survival of patients with ARID1A

mutations exhibits divergent trends contingent on the cohort under examination[121]. In support of this, Bai and colleagues demonstrated that BRG1 is highly expressed in biopsies from breast cancer patients which correlates with lower survival rates. Moreover, the knockdown of BRG1 has been observed to result in a reduction in proliferation, migration, and invasiveness when compared to control breast cancer cells (MDA-MB-231).[122]

In light of this evidence, Sun and colleagues verified that Arid1a loss protected mice livers from tumour initiation[123]. Indeed, the use of chemically induced models revealed a reduction in the number and size of liver tumours in the double KO, with no discernible differences in histology and proliferation compared to the control. Moreover, in mice with MYC overexpression, Arid1a loss did not accelerate cancer initiation as expected by a tumour suppressor gene. Conversely, mice model exhibiting MYC and ARID1A overexpression burden had shorter survival and hepatomegaly. Therefore, the data on loss and gain of function agree with one another in supporting the conclusion that ARID1A acts as a tumour promoter in the initiation process[123].

To elucidate the mechanism by which ARID1A exerts its function, they conducted RNA-seq analysis in the generated mouse models. Therefore, it was confirmed that ARID1A upregulates the expression of enzymes that are responsible for the production of ROS, which are well-established mediators of liver injury and hepatocarcinogenesis[123]. To further evaluate a cancer promoting role for ARID1A, they analysed the expression pattern in hepatocarcinoma (HCC) samples from patients. In accordance with the findings of a previous study[124], ARID1A was observed to be highly expressed in 86% of the samples, while 15% exhibited no expression. The adjacent non-cancerous tissue in all samples displayed weak or moderate staining. Conversely, there is a substantial loss of ARID1A expression in metastatic tissue in comparison to the paired

primary HCC. This phenomenon has already been observed in endometrial cancer, where only metastatic populations exhibit deleterious mutations in the ARID1A gene. Among the cancer driver genes analysed, ARID1A was the only gene that exhibited mutations in subclonal branches [125].

Furthermore, genomic analysis in HCC samples highlighted a prevalence of mono-allelic deleterious mutations, suggesting a phenotypical role for ARID1A even in haploinsufficient state[123].

These data prompted Sun and colleagues to investigate whether ARID1A's activity in tumorigenesis could be time- and dose-dependent. To achieve this, the researchers ablated Arid1a in mice with already initiated tumours. The resulting tumours were harvested and transferred to immunodeficient mice wherein tumours lacking Arid1a exhibited accelerated growth compared to WT[123].

Moreover, a comparative analysis of tumours with homozygous and heterozygous Arid1a loss demonstrated overlapping chromatin accessibility and gene expression profiles. Furthermore, mice with heterozygous Arid1a expression exhibited reduced overall survival, accelerated tumour growth and a metastasis burden comparable to that observed in homozygous Arid1a-deficient mice.

In conclusion, these data demonstrate a time-dependent and dose-dependent activity, thereby indicating that Arid1a promotes tumour initiation but is detrimental for cancer progression in late stages[123].

1.7 The functional interplay between YAP and ARID1A

While a substantial body of literature describes multiple mechanisms by which YAP activity is regulated through cytoplasm retention, comparatively less is known about its regulation in the nucleus[126,127].

In a recent study, Piccolo and colleagues identified several components of the SWI/SNF complex as nuclear interactors of YAP/TAZ. Indeed, co-immunoprecipitation analysis demonstrated that ARID1A (and not ARID1B) directly mediates the interaction between YAP and the remodeller complex[128]. Moreover, the researchers evaluated YAP activity in response to ARID1A depletion, revealing an increase in the expression of genes downstream of YAP/TEAD, without significant effects on YAP localization and phosphorylation status[128].

As has been widely reported, the shuttling of YAP from the cytoplasm to the nucleus is regulated by mechanical cues. Indeed, it is sequestered in the cytoplasm compartment in cells experiencing low mechanical *stimuli*, whereas it is free to translocate to the nucleus under high mechanical stress[126]. These two conditions can be replicated *in vitro* respectively by cells experiencing high or low cell density[127].

By applying different mechanical cues, Piccolo and colleagues demonstrated that ARID1A regulates YAP activity in response to mechanotransduction. Indeed, high mechanical cues induce F-ACTIN polymerization which sequesters ARID1A/SWI-SNF pool, thereby impairing its negative effects on YAP. Conversely, ARID1A interacts freely with YAP, thereby preventing the formation of the YAP/TEAD complex on chromatin[128].

It should be noted that other researchers have previously identified a physical interaction between the SWI/SNF complex and TAZ. Indeed, Skibisnki and colleagues identified the WW-domain in YAP/TAZ proteins as the portion responsible for binding with several subunits of SWI/SNF complex containing PPxY motifs such as ARID1A[129]. Nevertheless, in this instance, the SWI/SNF complex is responsible for sustaining TAZ transcriptional programmes, thereby ensuring the maintenance of basal cell homeostasis in mammary gland tissue[130].

A substantial body of research has demonstrated a physical interaction between these two factors, which results in the modulation of the YAP/TEAD transcriptional program. For instance, ARID1A functions as a suppressor of the YAP proliferative programme, thereby inducing cardiomyocyte maturation in the postnatal heart[131].

Another example of this pivotal interaction can be observed in the liver's response to injury. In this regard, previous studies have demonstrated that adult hepatocytes can undergo reprogramming into liver progenitor-like cells (LPLC) in response to injuries[132,133]. In this cellular context, Arid1a is essential for hepatocytes to gain the competence to undergo reprogramming in response to liver injury. Indeed, Arid1a enables a permissive chromatin state which facilitates the Yap-dependent LPLC programme[134].

In cancer models, it was previously demonstrated that overexpression of Yap in a constitutively active state (S127A[135]) in mouse liver did not result in any anomalies. Nevertheless, the oncogenic potential of Yap S127A is sustained by the co-expression of constitutively active beta-catenin, which results in a lethal burden of liver tumours [136]. In this tumour model, concomitant overexpression of ARID1A results in accelerated oncogenesis[123]. Consequently, further research is required to assess the consequences of this highly context-dependent interaction.

2. Aim of the study

YAP plays a significant role as an oncogenic driver in a multitude of tumour types[38]. In the majority of cases, its hyperactivity is not associated with any genomic abnormalities[137]. Consequently, it can be reasonably suggested that the aberrant activity of YAP in cancer should be sought in epigenetic dysregulation[138].

Conversely, ARID1A is primarily regarded as a tumour suppressor gene, whose activity is inactivated by recurrent mutations in numerous cancer types[139].

To date, numerous studies have demonstrated a functional interplay between these two factors in both physiological processes and oncogenesis[128,129,131]. Nevertheless, no unified mechanistic model has been established, and instead, a context-dependent interplay has been observed.

Indeed, Piccolo and colleagues reported that ARID1A acts as a tumour suppressor by inhibiting YAP activity in the nucleus, sequestering it and thus preventing it from binding TEAD on chromatin in cells experiencing low mechanical cues[128].

In contrast to this observation, in the liver compartment, the reprogramming of cells to repair injury as well as oncogenic processes are both driven by the activity of YAP through the support of ARID1A[123,134].

The objective of this project is to investigate the interactions between YAP and ARID1A in a cellular model of breast cancer and its healthy counterpart under high mechanics conditions.

The use of these two models will allow us to clarify the scope of these interactions and to validate the aforementioned cell context activity of ARID1A in a YAP-dependent manner, using a cell model competent to regulate YAP activity and its transformed counterpart in which YAP activity is unleashed.

Given the primary function of ARID1A as part of the chromatin remodelling complex and the dependence of YAP on the transcription factor TEAD, we hypothesise that ARID1A functions with YAP by regulating chromatin accessibility at TEAD-bound regions.

To this end, we will analyse available omics data by cross-referencing ATAC-seq results in ARID1A KO cells with ChIP-seq data for TEAD.

We will also analyse gene expression in YAP or ARID1A depleted cells by whole transcriptome analysis to investigate any overlapping effects and thus a common set of genes regulated by these two factors. Finally, we will take advantage of Cas9 genome editing to generate ARID1A KO clones to validate the dependence of the YAP-TEAD transcriptional programme upon ARID1A activity.

3. Materials and methods

3.1 Cell culture and transfection

MCF10A and MCF10AT cells obtained from Luca Azzolin were grown in DMEM F12 (Euroclone #ECM0090L) supplemented with 5% horse serum (Euroclone #ECS0091L), 2 mM L-glutamine (Euroclone #ECB3000D), 1% penicillin/streptomycin solution (Euroclone ECB3001D) and freshly embedded with insulin 0.01 mg/ml (Sigma-Aldrich #91077C-250MG), EGF 20 ng/ml (Peprotech #AF-100-15), hydrocortisone 500 ng/ml (Sigma-Aldrich #A16292) and cholera toxin 100 ng/ml (Sigma-Aldrich #C8052).

All cells were maintained in a humidified incubator at 37 °C, 5% CO2 and periodically tested for mycoplasma contamination.

YAP and ARID1A silencing were obtained by transfecting 30 nM of siRNA (listed in table 1) with Interferin as transfection reagent (Polyplus #409). siRNA targeting luciferase reversed sequence was used as negative control (sequence listed in table 1).

Cas9 expressing vector was obtained from Addgene (#62988). Single Guide RNA sequence (listed in table 1) to target ARID1A gene were obtained from CHOP CHOP webtool[140] and cloned as duplex in the Cas9 vector as described in Zang protocol[141]. The resulting vector was then Sanger sequenced to verify the insertion of the guiding sequence. It was then transfected into MCF10A cells using Jetprime transfection reagent (Polyplus #115) according to the manufacturer's instructions. Thirty-six hours after transfection, cells expressing Cas9 were selected by puromycin selection (1 ug/ml).

After 48 hours, surviving cells were diluted and seeded into 96-well plates for monoclonal cell isolation. MCF10A clones were first tested for ARID1A KO by

Western blot and finally genetically characterised by Sanger sequencing in the Cas9 target sequence.

Table 1:

Object	Sequence 5'-3' sense	Sequence 5'-3' antisense
siRNA Negative control	AGCUUCAUAAGGCGCA UGC[dT][dT]	GCAUGCGCCUUAUGA AGCU[dT][dT]
siRNA YAP	GACAUCUUCUGGUCA GAGA[dT][dT]	UCUCUGACCAGAAGA UGUC[dT][dT]
siRNA ARID1A	CCUUGGUUACACUCG CCAA[dT][dT]	UUGGCGAGUGUAACC AAGG[dT][dT]
gRNA ARID1A	AAACCGCAGCGGTACC CGATGACCC	CACCGGGTCATCGGG TACCGCTGCG

3.2 RNA isolation and Real Time quantitative PCR (RT-qPCR)

Total RNA was isolated from cells using Direct-zol RNA Miniprep kit (Zymo research #R2053) and quantified with Qubit fluorometric assay (Thermo Fisher #Q10211).

Thus, 1 μg of total RNA was reverse transcribed 42 °C using Protoscript II Reverse transcriptase (New England Biolabs GmbH #0368) and random hexamers (Takara Bio Europe #3801) according to the manufacturer's protocol. qPCR amplification of cDNAs was performed using specific primers (listed in table 2) with GoTaq® qPCR Master Mix (#A6002 Promega) in a CFX Connect instrument (BioRad).

The expression level of each gene was assessed using the 2^(-ΔΔCt) method by normalisation to GAPDH levels.

Table 2:

Gene	Sequence 5'-3' Forward	Sequence 5'-3' Reverse
YAP	GCACCTCTGTGTTTTAAGGGTCT	CAACTTTTGCCCTCCTCCAA
ARID1A	TGAACAACAAGGCAGATGGG	CACAGCAGGCAGATTTGTCA
CCN1	CCTTGTGGACAGCCAGTGTA	ACTTGGGCCGGTATTTCTTC
CCN2	AGGAGTGGGTGTGTGACGA	CCAGGCAGTTGGCTCTAATC
LPCAT3	GTTCTGCAGTCGGGTTTCCA	CCAGGCAGTTGGCTCTAATC
AXL	GTGGGCAACCCAGGGAATAT	GTACTGTCCCGTGTCGGAAAG
AGFG2	CGTGGAAATGAGGTTTGCCG	CTTCCCTTCTGGGATGGAGC
CLU	CAGGCCATGGACATCCACTT	GTCATCGTCGCCTTCTCGTA
TFP1	GTGCTGCCTCTCTCTATGCC	AGGAAATCCACCGACCTGTCAA
NRP1	GCCAGAGGAGTACGATCAGC	CAGGTCTGCTGGTTTTGCAC

3.3 Protein isolation and Western blot analysis

Cells were harvested by centrifugation and resuspended in lysis buffer (NaCl 150 mM, Tris HCl pH 8 50 mM, EDTA 2 mM, NP-40 0.5%, glycerol 10%, protease inhibitor cocktail (Sigma Aldrich#P8340). A total of 30 ug of total protein extract were loaded on a 8% SDS PAGE gel and gels were blotted by wet protein transfer to 0,45 μm nitrocellulose membrane (GE10600002 Amersham Protran). After membrane blocking with 5% (*w*/*v*) non-fat milk (A0830 PanReac Applichem) in

TBS-tween 0.1%, membranes were incubated with specific antibodies (anti-YAP Cell Signaling #14074, anti-ARID1A Proteintech #303004-1-AP, anti-GAPDH Cell Signaling #2118, β-tubulin Sigma Aldrich #T7816, anti-TEAD4 Abcam #58310) and developed with chemiluminescence reagents (Supersignal West Pico PLUS Chemiluminescent Substrate Thermo Scientific #34577). For control WB of ChIP against TEAD-4, anti-light chain mouse was used as secondary antibody (Cell signaling #91196S). Images were acquired with Chemidoc MP Imaging System Bio-rad. Densitometric analysis was conducted on Image Lab software from Biorad.

3.4 Chromatin immunoprecipitation (ChIP) qPCR

Approximately 1x10^7 cells were crosslinked in 150mm dishes adding 1/10 of volume with formaldehyde solution (HEPES pH 8 50mM, NaCl 100mM, EDTA pH 8 1mM, EGTA 0,5 mM, formaldehyde 11%) for 10 minutes at room temperature on shacking, followed by glycine quenching for 5 minutes.

After three washes with ice-cold PBS, fixed cells were scraped and collected by centrifugation.

A preliminary lysis process occurred in "buffer 1" (Tris HCl pH 8 50mM, EDTA pH 8 2mM, NP40 0,1%, glycerol 10%, protease inhibitor cocktail Sigma #P8340) on wheel to disrupt cell outer membrane, followed by douncing with potter. Subsequent centrifugation permits to collect nuclei and to discard cytoplasmic fraction. Collected nuclei were resuspended in "lysis buffer 2" (Tris HCl pH 8 50mM, EDTA pH8 5mM, SDS 1%, protease inhibitor cocktail) for 20' on wheel and transferred to sonication tubes. Cross-linked chromatin was disrupted with Cup Horn sonicator (Fisher scientific #FB705) until 150-200 bp fragments were obtained.

Each chromatin immunoprecipitation was set up with approximately 150μg of chromatin and 3μg of specific antibody for targeting TEAD1 or iso-specific antibody as negative control (anti-TEF3 NG2 clone Santa Cruz #sc-101184, purified mouse IgG Bio-rad #PMP01).

1/10 of volume calculated for IP reactions was taken from chromatin solution and conserved at this stage as a fraction of input. IP reactions occurred overnight on wheel at 4°C diluted 1:10 in volume with dilution buffer (Tris HCl pH 8 50mM, EDTA pH8 5mM, NaCl 200mM, NP40 0,5%, protease inhibitor cocktail).

The subsequent day, magnetic beads (Dynabeads protein G Invitrogen Thermo scientific #10003D) were injected into the IP tubes in order to capture antibody-chromatin complexes as reported by manufacturers. Immuno-complexes collected by magnetic beads were subjected to subsequent washes and finally eluted by two consecutive extractions.

Chromatin was reverse-cross linked overnight on shacking at 67°C and treated with Rnase A and proteinase K. The input and ChIP DNA were finally purified with phenol/chloroform solution, precipitated and resuspended in molecular grade H2O. Finally, ChIP-qPCR was performed with primers amplifying CCN2 promoter region and HBB promoter region as negative control, since it is not expressed in breast epithelium (primers listed in table 3).

Table 3:

Gene promoter	Sequence 5'-3' Forward	Sequence 5'-3' Reverse
CCN2	TGTGCCAGCTTTTTCAGACG	TGAGCTGAATGGAGTCCTACACA
HBB	GCTTCTGACACAACTGTGTTCACTAGC	CACCAACTTCATCCACGTTCACC

3.5 RNA sequencing (RNA-seq)

Total RNA extracted from MCF10A and MCF10AT cell lines (treated with siRNAs against either ARID1A, YAP or a control sequence) was sent for transcriptomic analysis through a service provider. Paired end (2x150) sequencing of 60 M reads per sample was performed.

3.6 Bioinformatic analysis

3.6.1 RNA-seq

Fastq files were checked using FastQC (available online: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapters (AGATCGGAAGAGCACACGTCTGAACTCCAGTCA, AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT) were trimmed using cutadapt v 3.5[142] with the following command line options:
-m 30:30 --trim-n --nextseq-trim=20
Reads were aligned to the hg38 human genome using hisat v 2.2.1[143] . Duplicate reads were removed using Sambamba v0.8.2[144] .
Read counts at gene level (gencode v43 basic gene annotation) were obtained using htseq-count v 0.12.4 with the following command line options:

```
-r pos \
 -m intersection-nonempty \
-f bam \
-s reverse \
-i gene_name \
--nonunique none \
--secondary-alignments ignore
```

3.6.2 ATAC-seq

Raw sequences were obtained from SRA (NCBI). Reads were aligned to the hg38 human genome sequence using STAR[145] with the following command line option:

--alignIntronMax 1
to suppress spliced alignments.
Quality controls were performed using custom perl scripts and R scripts to compute Transcription Start Site Enrichment (TSSE) score following ENCODE definition[146] after filtering reads overlapping ENCODE blacklisted regions. The ENCODE reference TSS for hg38 was used. All samples displayed a TSSE > 6.
ATAC-peaks, differentially accessible regions and motifs enriched at regions whose accessibility was impaired in ARID1A KO were identified using HOMER[147] . Enhancer regions were identified based on HMM Chromatin states annotation and peaks lying on enhancers were selected using bedtools[148] intersect. Heatmaps and peak coverage images were obtained using deeptools2[149].

3.7 Statistical analysis

PCA, sample clustering and statistical testing for Differential gene expression was carried out using DEseq2[150]. Batch effect removal from rlog transformed data was performed using limma package[151]. GO term enrichment was computed using the gprofiler2 R package[152].
R Session Info is reported in supplementary data.
RT-qPCR data and WB densitometric analysis were assessed in significance using the unpaired Student T test for comparison between groups using GraphPad software.

4. Results

4.1 ARID1A loss impairs chromatin accessibility on enhancers regions

Given ARID1A's primary function as a component of the SWI/SNF complex, we conducted an analysis to ascertain the impact of ARID1A deficiency on chromatin accessibility in various cancer types.

In order to achieve this, an analysis was conducted on data deposited on the Sequence Read Archive (SRA) from Assay for Transposase-Accessible Chromatin sequencing (ATAC-Seq) experiments in MCF7, HCT116 and endometrial cells knockout for ARID1A[92,93,153].

First of all, we inquired as to whether it would be feasible to delineate a set of affected regions that exhibit common characteristics. In order to answer this question, we initially undertook a comparison of the aforementioned regions with the histone modification data obtained from ChIP-seq in the same cell lines (ENCODE data). The comparative analysis revealed a significant decline in chromatin accessibility in the regions designated as enhancers based on histone marks. Indeed, out of 50908 peaks exhibiting diminished accessibility upon ARID1A loss in MCF7 cells, 13,3% and 17,7% could be respectively classified as weak and active enhancers. Furthermore, 57% of the less accessible chromatin displays characteristics of quiescent chromatin. In HCT116 and endometrial cells, the proportion of enhancers with reduced accessibility in ARID1A KO is 60.9% and 72.8%, respectively.

In light of this observation, we sought to ascertain the impact of ARID1A depletion on chromatin accessibility, focusing our investigation on regulatory regions defined as enhancers. As illustrated in Figure 2, the depletion of ARID1A

in cells resulted in a discernible yet modest decline in chromatin accessibility when compared to the parental cell line.

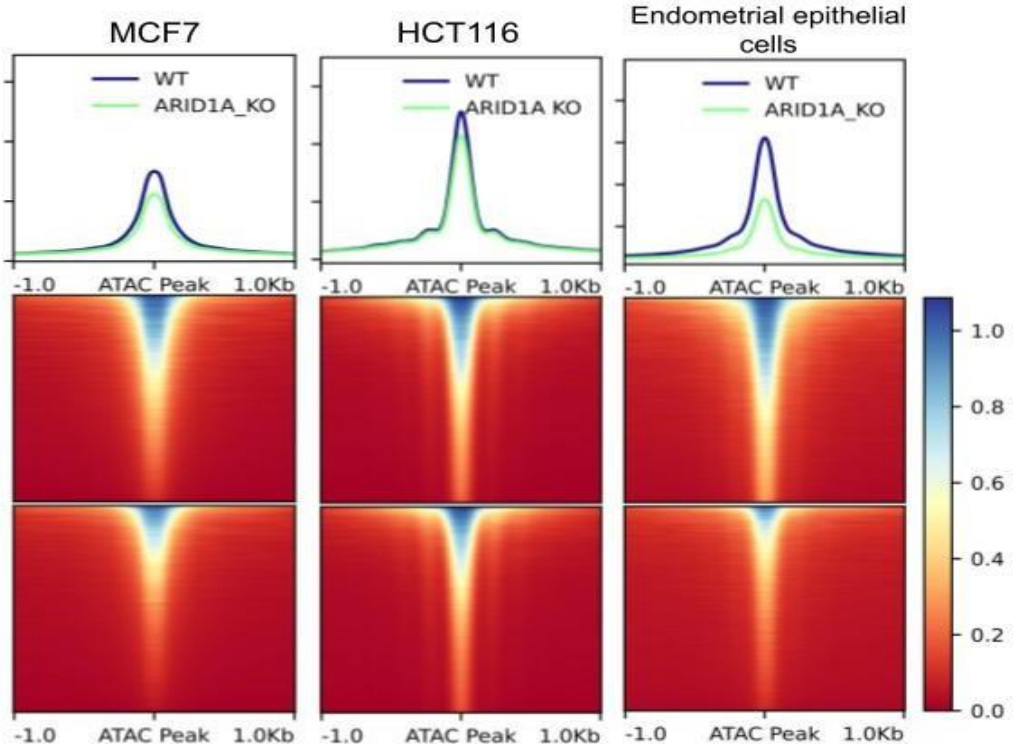

Fig. 2 ARID1A loss slightly affects chromatin accessibility on all enhancer regions: Heatmaps of ATAC-seq peaks in ARID1A KO and parental cell line. Each peak is stacked in rows ordered in descending order of signal strength

4.2 ARID1A regulates chromatin accessibility at TEAD binding sites

As previously stated, ARID1A has been observed to co-bind chromatin with several transcription factors.

In order to identify regulatory elements that are dependent on ARID1A activity, we conducted an enrichment motif analysis on enhancer regions that were made less accessible in KO of all cell lines, using the HOMER[154] software.

The top-most enriched motif matches the consensus sequence bound by AP-1 transcription factors. This finding is not unexpected, given that AP-1 co-binding with the SWI/SNF complex has already been demonstrated to play a critical role in establishing a permissive state for AP-1 binding regions, with ARID1A being

a key factor in this process[93,155,156]. In the second position we have found FOXA1, whose functional interplay with ARID1A has been documented[157].

The third most represented motif is the TEAD family binding consensus, with approximately 15,8% of the total peaks affected by loss of accessibility (Table 4).

Rank	Motif	P-value	% of Targets	Best Match
1		1e-9359	46.40%	Fra1(Bzip)
2		1e-824	12.64%	FOXA1(Forkhead)
3		1e-567	15.84%	TEAD(TEA)

Table 4 Consensus sequence recognized by BZip (AP-1 family), FOXA1 and TEAD are the most frequent in the regulatory regions negatively affected by ARID1A loss.

We then re-analysed the ATAC-seq data, this time focusing on peaks with binding regions for TEAD within them, as reported in the ChIP-seq data for this factor. In comparison to the observation of total enhancer peaks affected by ARID1A loss, TEAD binding regions on enhancers demonstrated a markedly reduced level of accessibility when compared to the parental cell lines. This effect was observed consistently across all cancer cell lines analysed. (fig. 3).

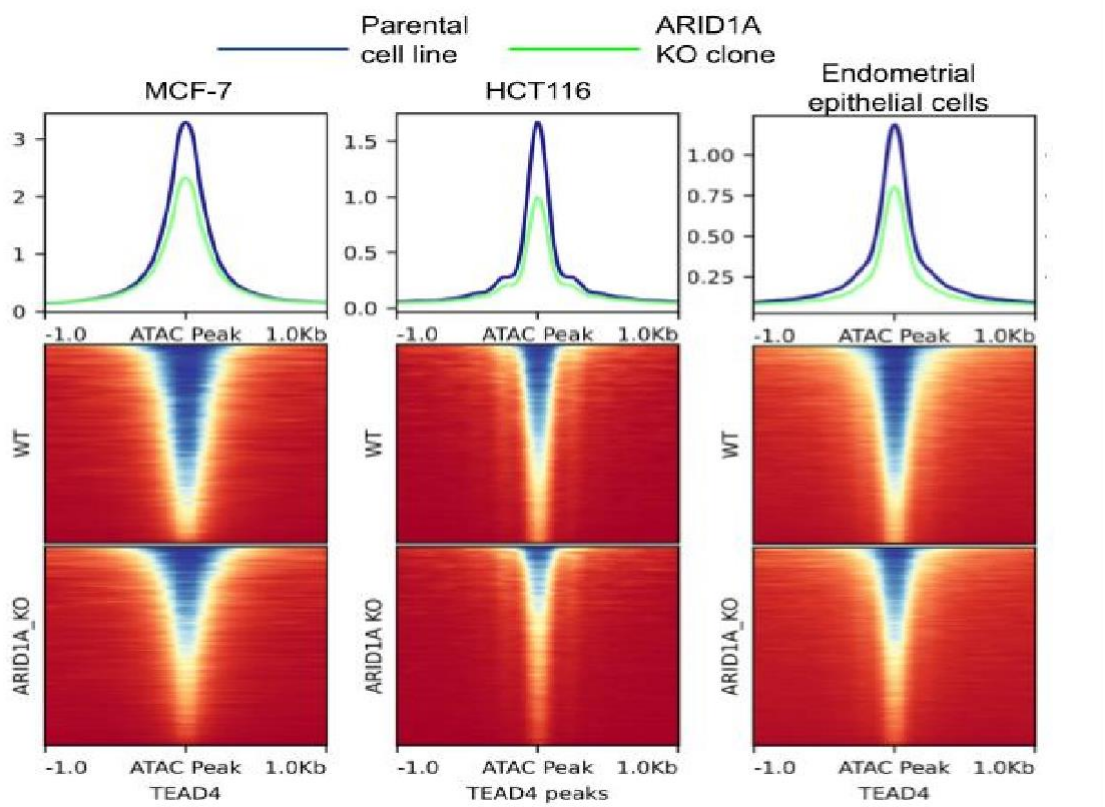

Fig.3 ATAC-seq unveils a significant reduction of chromatin accessibility at TEAD4-bound enhancers regions in three different cell lines: Heatmaps of ATAC-seq peaks in ARID1A KO and parental cell line. Each peak is stacked in rows ordered in descending order of signal strength.

4.3 YAP and ARID1A regulate a common set of genes in breast epithelial cell lines

To understand how changes in chromatin accessibility affect TEAD transcriptional programmes, we analysed gene expression in ARID1A or YAP-depleted cells. In fact, as has been thoroughly established, TEAD activity is primarily dependent on the formation of a complex with YAP to initiate transcription. To this end, we performed RNA-seq analysis in MCF10A cells and its transformed counterpart MCF10AT cell line. MCF10A cells are an immortalised but non-tumorigenic human mammary epithelial cell line that provides a good model to study the oncogenic effects of YAP. Indeed, these

epithelial cells physiologically respond to mechanical *stimuli* that inhibit proliferation through the "Hippo ON status" but over-activated YAP drives the acquisition of oncogenic features[158]. Conversely, MCF10AT represents a cellular context in which H-RAS transformations remove the negative control of the Hippo pathway on YAP, allowing it to enter the nucleus and drive transcription even at low mechanics[159]. The efficacy of silencing was assessed both at the protein level by Western blot analysis (Figure 4A-B-C) and at the mRNA expression level by real-time qPCR (Figure 5). Densitometric analysis showed that depletion of ARID1A did not result in a significant increase in YAP protein expression, and vice versa. (Fig. 4B-C). This finding excluded the possibility that the reduced expression of YAP targets upon ARID1A KD was dependent on the reduced protein level of YAP.

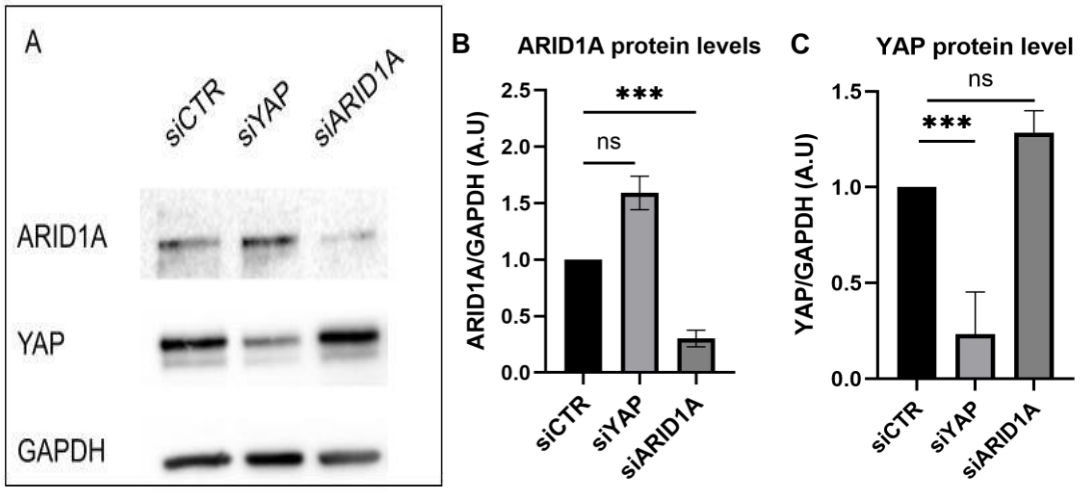

Fig. 4 Expression of YAP and ARID1A proteins in MCF10A KD cells:

(A) Western blot (WB) for ARID1A/YAP in MCF10A treated with siRNA. GAPDH is used as internal control for normalisation. (B-C) Densitometric analysis of ECL signals from WB in MCF10A treated with siRNA against ARID1A or YAP or control. YAP and ARID1A signals are normalised on GAPDH. $n = 4$. For all samples (compared to siCTR), unpaired one-tailed t-test p-value < 0,05 (*), < 0,01 (**), < 0,001(***).

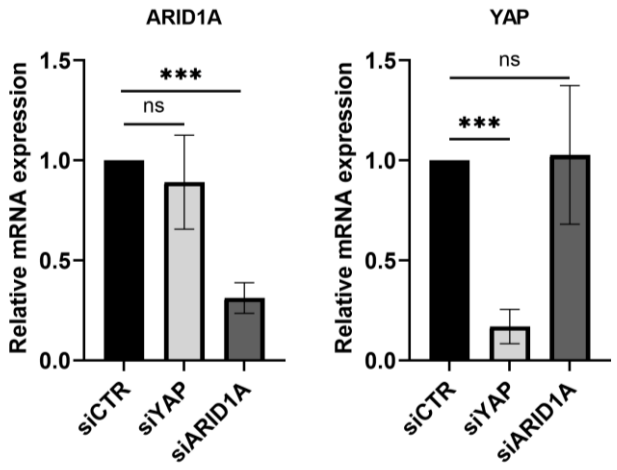

Fig. 5 Expression of YAP and ARID1A mRNAs in MCF10A KD cells:
Real Time qPCR in YAP and ARID1A depleted MCF10A cells in comparison to control. $n = 3$. For all samples (compared to siCTR), unpaired one-tailed t-test p-value < 0,05 (*), < 0,01 (**), < 0,001(***).

In accordance with protein expression, no influence was observed between the mRNA levels of YAP and ARID1A, thus precluding the possibility that increased protein levels could depend upon a boost in transcription of these two factors.
To assess the similarity between samples and the weight of variance introduced by our experimental conditions, we performed Principal Component Analysis (PCA). In our case, PCA identified differences between ARID1A KD and YAP KD as the main source of variance in our samples. The second component is represented by differences between control and silenced samples for YAP and ARID1A (Figure 6). Removal of the batch effect was effective and resulted in a well-established clustering of replicates between samples (Fig. 6).

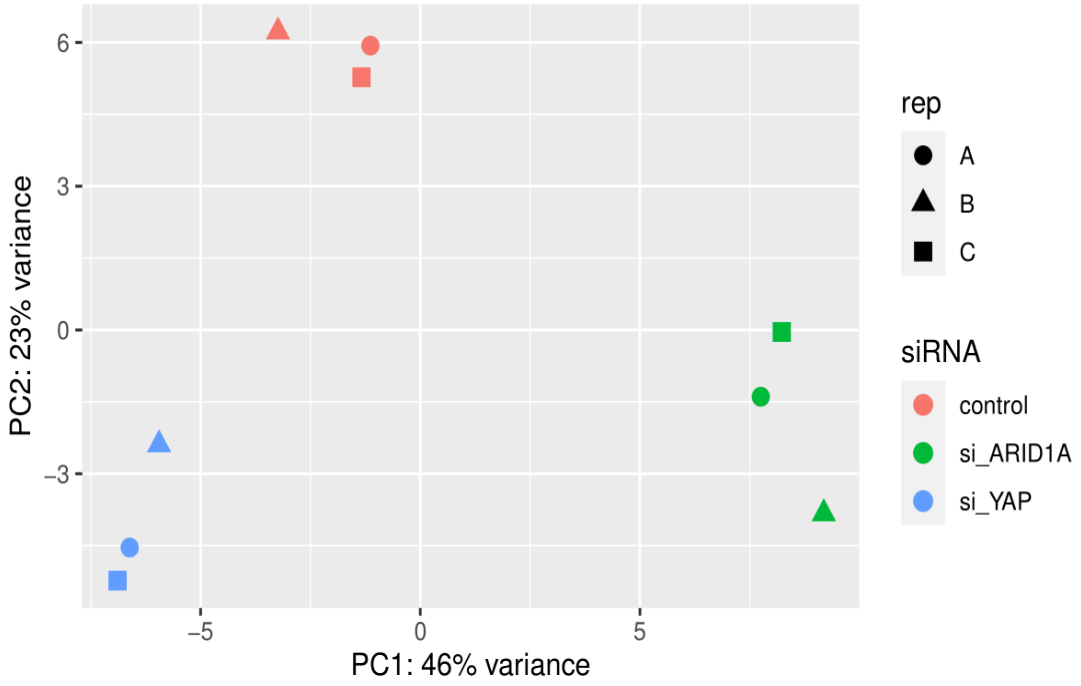

Fig.6 Principal component analysis (PCA) of RNA-Seq experimental data from MCF10A cells: The variance between ARID1A KD cells and YAP KD cells is greater than the differences between these two conditions and control cells. Batch effect was successfully removed using the limma package on R (PCA analysis without batch effect removal in Supplementary).

Next, we examined the effect of silencing on the top 50 genes for variance. Consistent with the PCA analysis, the main source of variability comes from the two knock-downs. Specifically, the ARID1A KD mainly induces an opposite effect on the expression of this restricted subset of genes compared to the YAP KD (Figure 7).

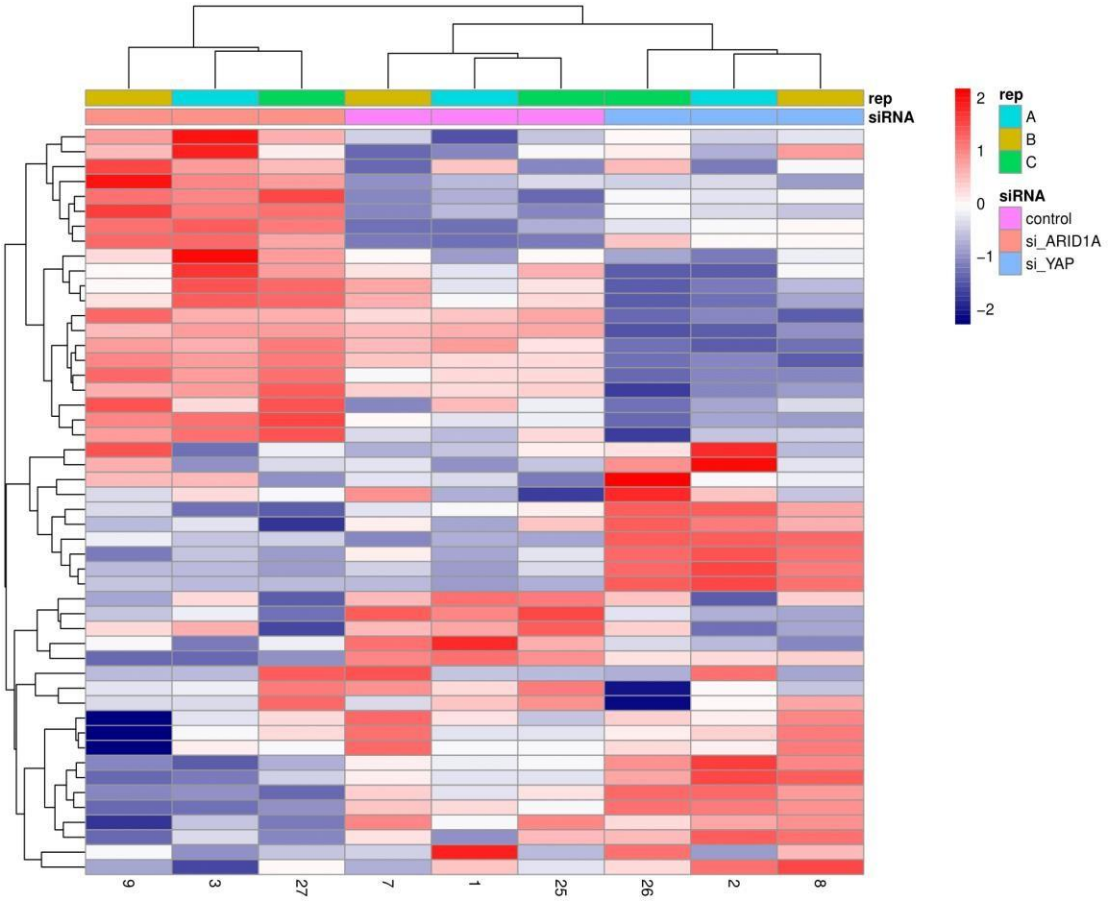

Fig.7 Heatmap of top 50 genes for variance in MCF10A siRNA treated cells and control:
The top 50 genes were selected for variance in their expression, independent of experimental design. The batch effect was successfully removed using the limma package in R (heatmap without batch effect removal in the Supplementary file).

Despite this specular effect on this subset of genes, whole transcriptome analysis of ARID1A or YAP-depleted MCF10A cells reveals a modest overlap between these two conditions. In fact, of the 312 genes significantly affected by both YAP and ARID1A KD, 271 genes are concordantly modulated by both YAP and ARID1A (Figure 8).

In MCF10AT cells, the same analysis revealed an extensive number of genes concordantly affected by these two conditions. To be specific, 350 genes are concordantly modulated by both YAP and ARID1A out of 373 genes affected (figure 4 in Supplementary).

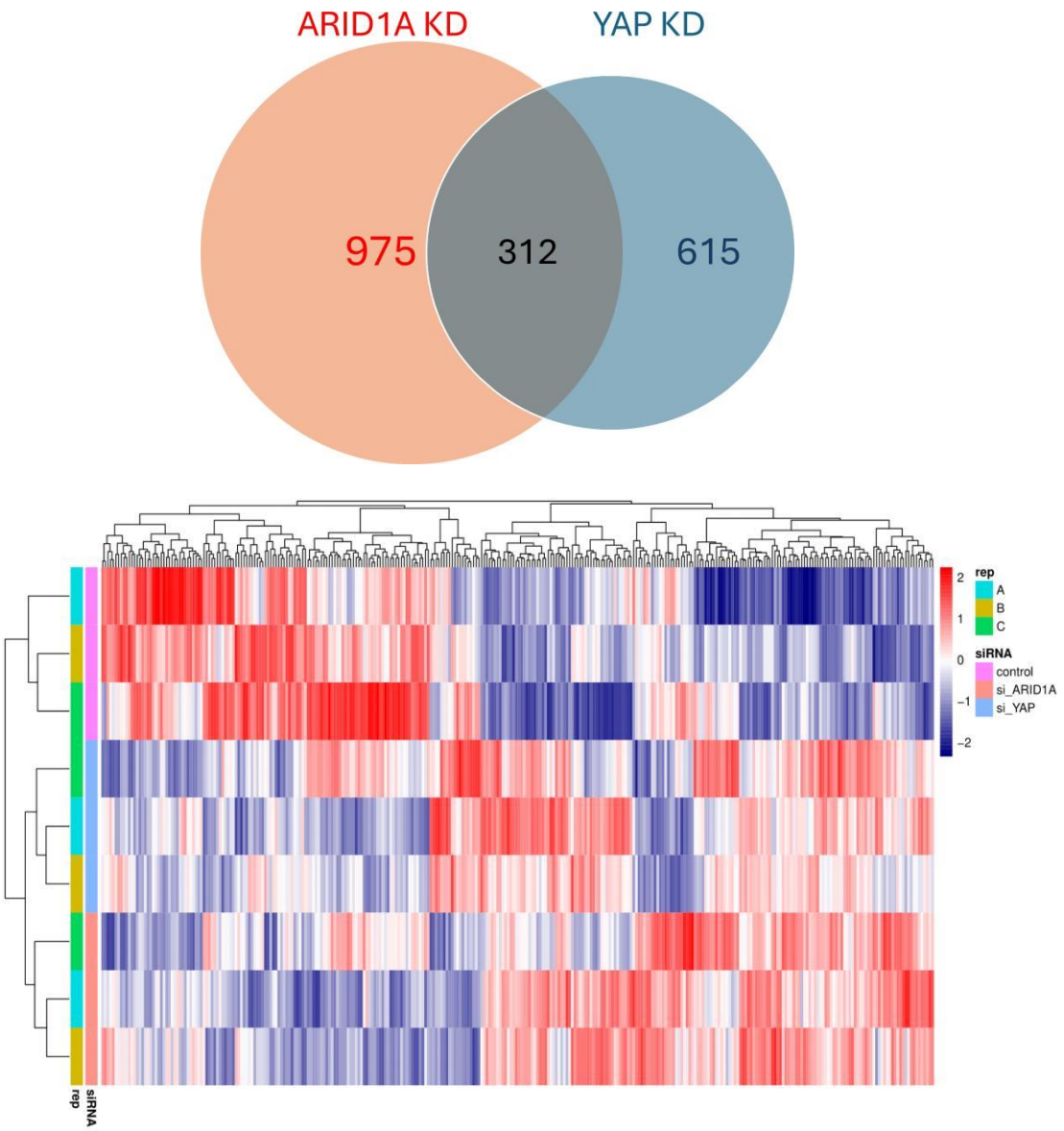

Fig. 8 Transcriptome analysis of ARID1A or YAP depleted MCF10A cells:
Venn diagram shows differentially expressed genes in the two analysed conditions. Intersections between ARID1A and YAP KD count on 312 genes, 271 of which are concordantly affected. The analysed gene expression is collected in a heatmap, which hierarchically clusters conditions. Each individual gene of the human transcriptome occupies one column, intersecting one row for each condition analysed. The Z-score on the right relates differential gene expression to a colour scale. $n = 3$

We then asked whether the genes differentially expressed upon ARID1A or YAP KD share common features in terms of biological function. To do this, we performed Gene Ontology (GO) enrichment analysis on the downregulated and

upregulated gene sets in these two conditions compared to control (GO term of upregulated genes in supplementary fig.6-7). Considering only downregulated genes, 507 GO terms are enriched in ARID1A KD cells while only 150 are enriched in the YAP KD condition. Out of these 150, 112 GO terms are commonly enriched with genes negatively affected by ARID1A KD (fig. 9)

A comprehensive examination of the Gene Ontology (GO) terms indicated that both conditions are characterised by an enrichment of genes involved in cell migration and proliferation. In particular, the depletion of YAP or ARID1A results in the downregulation of genes linked to the aforementioned GO terms (Fig. 10-11).

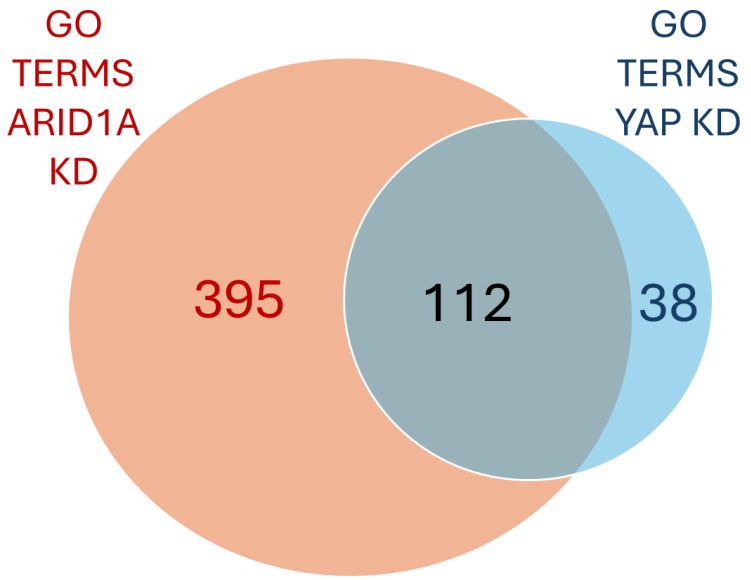

Fig.9 Venn diagram for GO term enrichment analysis: Out of 150 GO terms enriched in downregulated genes in YAP KD cells, 112 are shared with ARID1A KD.

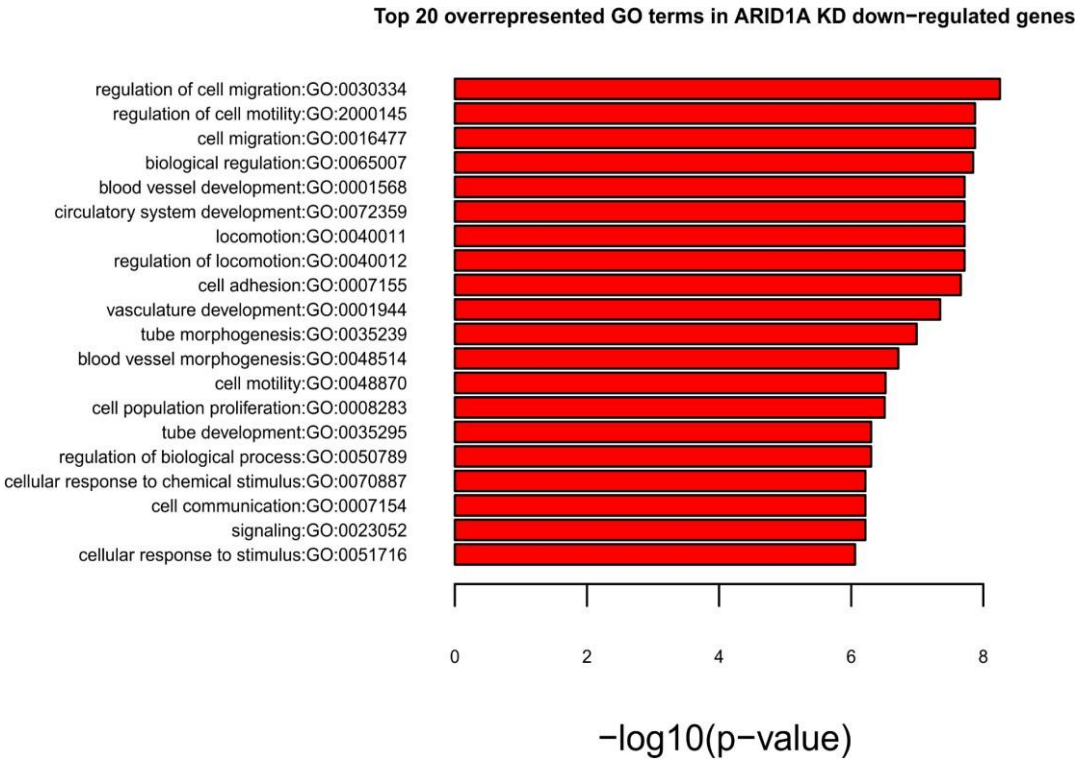

Fig. 10 Gene ontology enrichment analysis for downregulated genes in MCF10A ARID1A KD cells compared to control. Top 20 significantly enriched GO terms in ARID1A depleted cells.

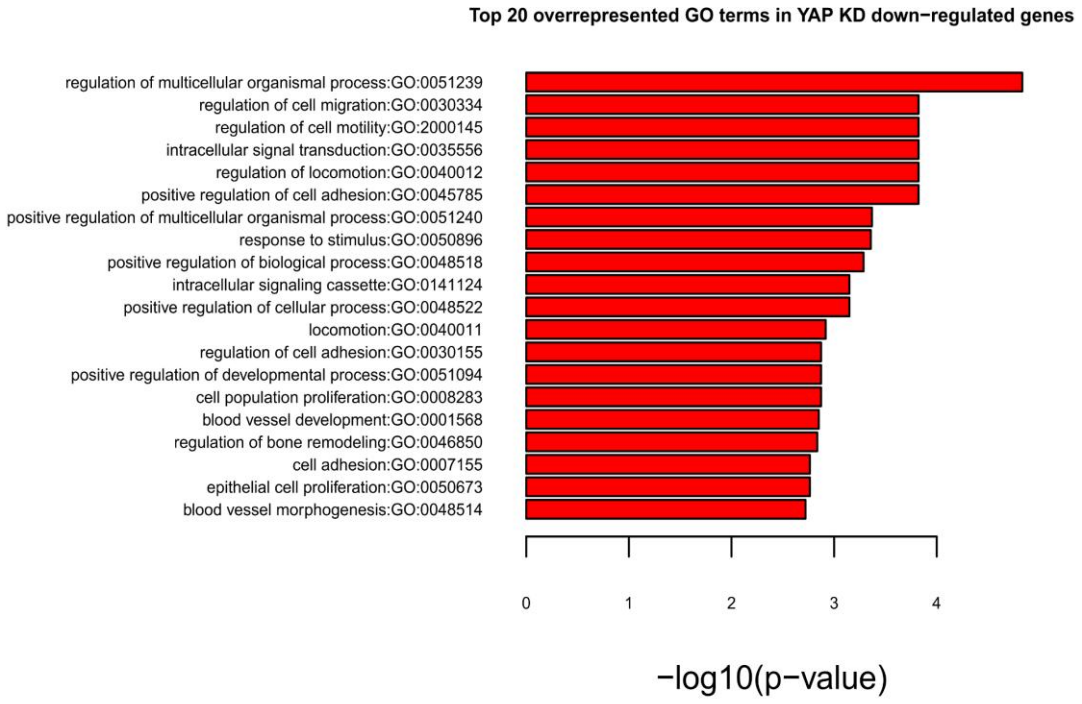

Fig. 11 Gene ontology enrichment analysis for downregulated genes in MCF10A YAP KD cells compared to control. Top 20 significantly enriched GO terms in YAP depleted cells.

To assess the significance of this overlap in terms of differential expression and biological functions, we compared the list of genes downregulated in ARID1A KD with gene sets collected in the GSEA database[160].

Consistent with the chromatin accessibility data, genes with a consensus sequence for TEAD-1 in their promoter are present in the list of downregulated genes more than by chance (fig. 12A).

Using the same approach, we have compared the above list with the set of 57 genes associated as a conserved signature of YAP[161]. Also in this case, our ARID1A-dependent gene list is significantly enriched for canonical YAP targets (fig. 12B). No significant enrichment was reported for genes involved in the Hippo pathway, suggesting that ARID1A does not affect YAP target genes by upregulating its inhibitory factors. (Supplementary Figure 8).

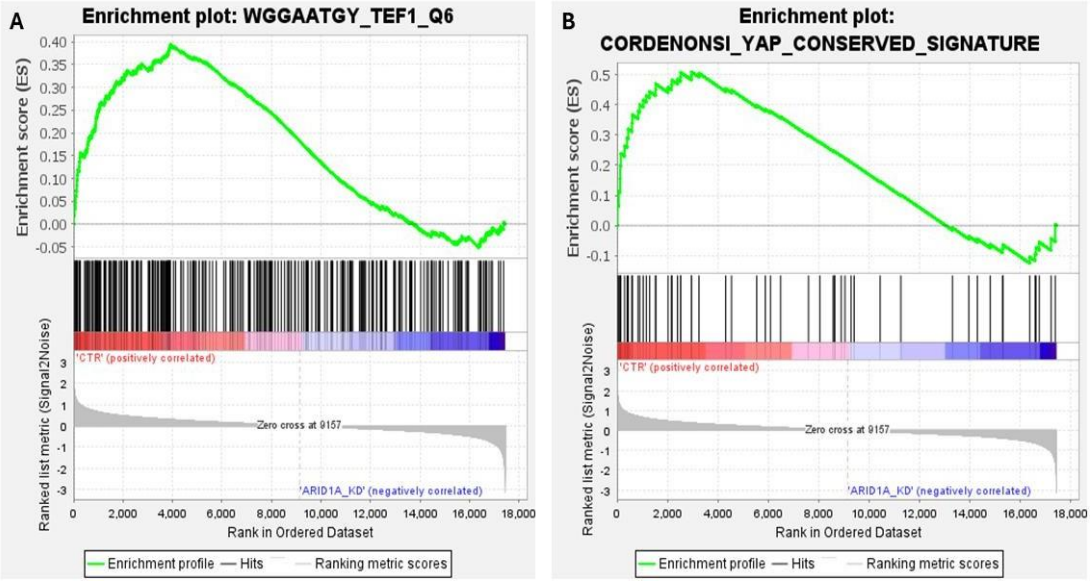

Fig. 12A Genes with TEAD-1 motif in their promoter are downregulated in ARID1A KD MCF10A cells: Enrichment plot for genes with at least one conserved motif in their promoter recognized by TEAD-1. Out of 278 genes with TEAD-1 binding site, 101 are significantly downregulated in ARID1A KD. FDR q-value < 0,05 Fig. 12B «YAP signature genes» represent a significant portion of downregulated genes in ARID1A KD MCF10A cells: Enrichment plot for the gene-set «YAP signature» in control cells compared to ARID1A depleted cells. Out of 55 genes belonging to this list, 27 are downregulated in ARID1A KD. FDR q-value < 0,01

Having established the overlap between the YAP-TEAD transcriptional programme and ARID1A-dependent genes, we focused our attention on a restricted set of genes whose expression is significantly affected by the depletion of both factors (fig. 13).

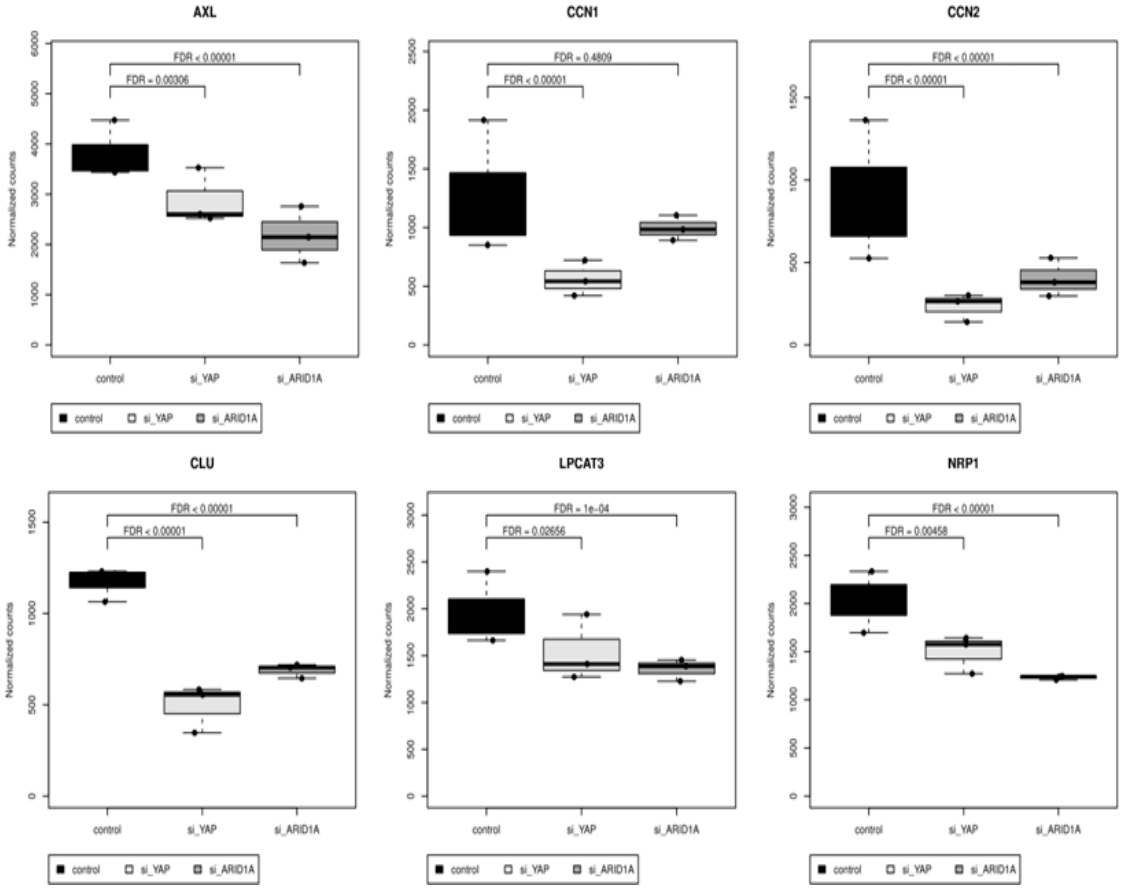

Fig. 13 Set of genes concordantly affected by YAP or ARID1A depletion in MCF10A cells:
Box plot shows the normalised number of reads for a set of genes concordantly downregulated upon YAP or ARID1A knockdown, compared to control. Statistical significance was assessed using adjusted statistical tests for multiple comparisons.

As a final check on the transcriptomic data, we evaluated the expression of this set of genes by real-time qPCR. As shown in Figure 14, AGFG2, AXL, CCN2 (best known as CTGF and a major target of YAP) and LPCAT3 were significantly downregulated in both conditions. Contrary to *in silico* data, CCN1 (also known as CYR61) and CLU were not significantly downregulated in ARID1A-depleted cells (data not shown) while NRP1 was not significantly downregulated in YAP

KD cells.

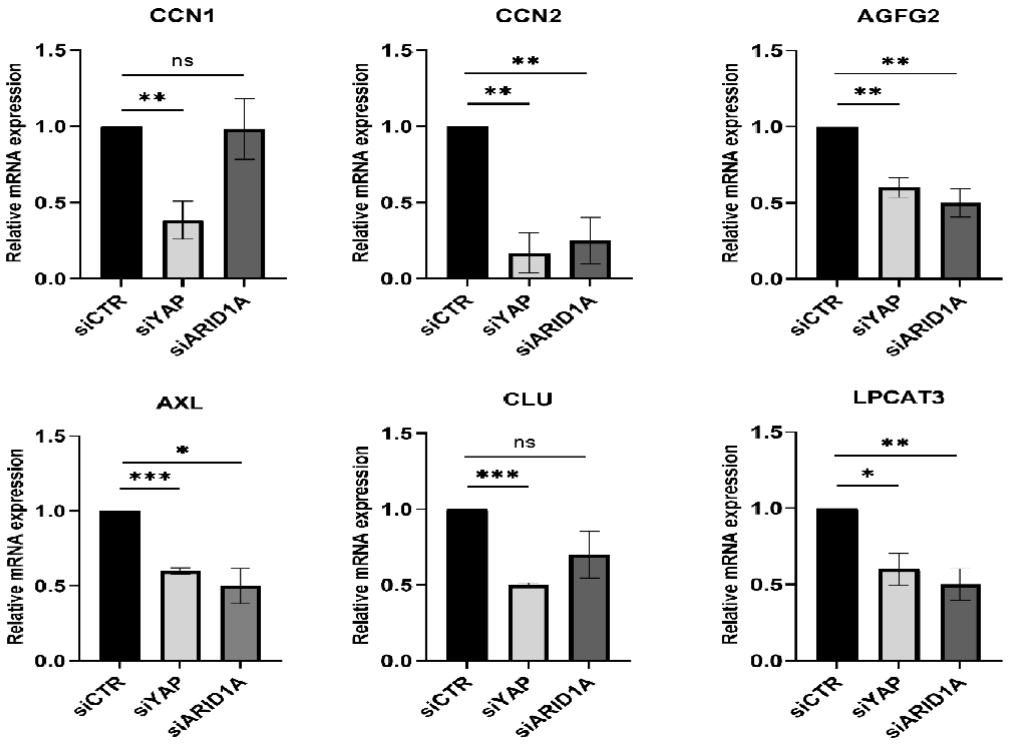

Fig. 14 Relative expression of genes concordantly affected by ARID1A KD or YAP KD in MCF10A cells:

Real Time qPCR in YAP and ARID1A depleted MCF10A cells in comparison to control. $n = 3$. For all samples (compared to siCTR), unpaired t-test p-value < 0,05 (*), < 0,01 (**), < 0,001(***).

4.4 TEAD occupancy on promoters of its targets is affected by ARID1A loss

In order to evaluate the extent of the involvement of ARID1A in the YAP-TEAD transcriptional programme, we took advantage of the Cas9 system to abolish ARID1A expression in MCF10A cells. The target sequence recognised by CAS9 on the ARID1A coding sequence was selected using CHOP-CHOP software, filtering to recognise exons closest to the 5' start site and with less off-target[140]. Western blot analysis for ARID1A protein confirmed the loss of ARID1A in the selected clone (fig. 15A), as did Sanger sequencing, which revealed a 1-base

insertion at the Cas9 cut point (Supplementary fig. 9). The ARID1A KO clone partially resembles the molecular phenotype observed in KD cells, expressing significantly lower levels of CCN2, AXL and NRP1 mRNAs compared to WT (fig. 15B).

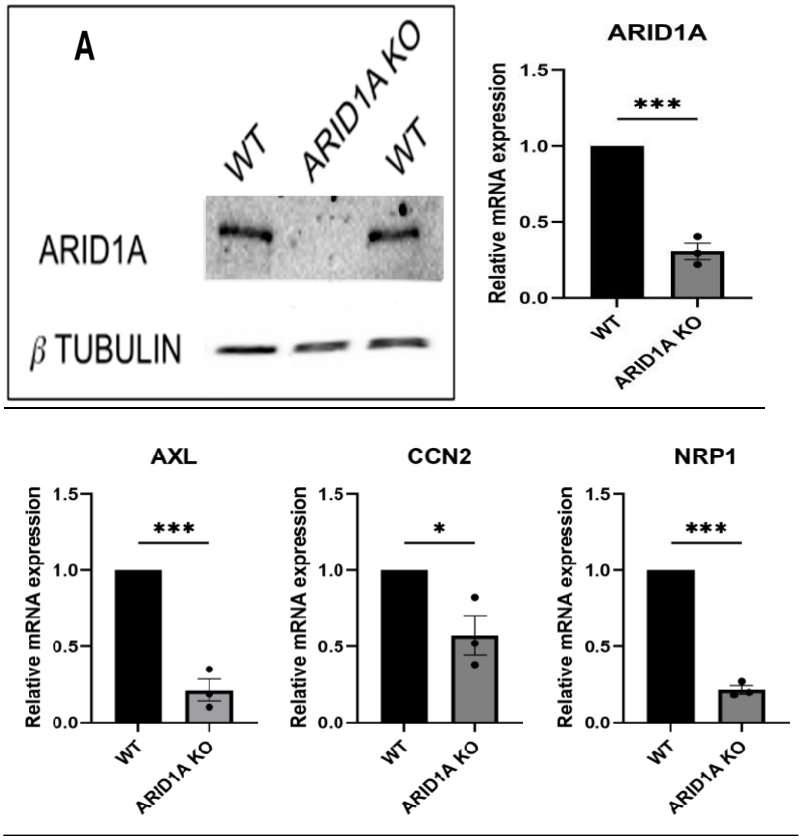

Fig. 15 Validation of ARID1A KO in MCF10A clone:
(A) Western blot for ARID1A in MCF10A clone. Beta tubulin as internal control.
(B) Real Time qPCR in ARID1A KO MCF10A cells in comparison to WT. *n = 3*.
For all samples (compared to WT), unpaired t-test p-value < 0,05 (*), < 0,01 (**), < 0,001(***).

In light microscopy, MCF10A ARID1A KO appear phenotypically different and behave quite differently compared to WT counterparts. In fact, as shown in Figure 16, KO clones appear more elongated with a patchy cytoplasm and tend to aggregate in *insulae*. In contrast, WT cells appear rounded with dense cytoplasm and are uniformly distributed across the dish.

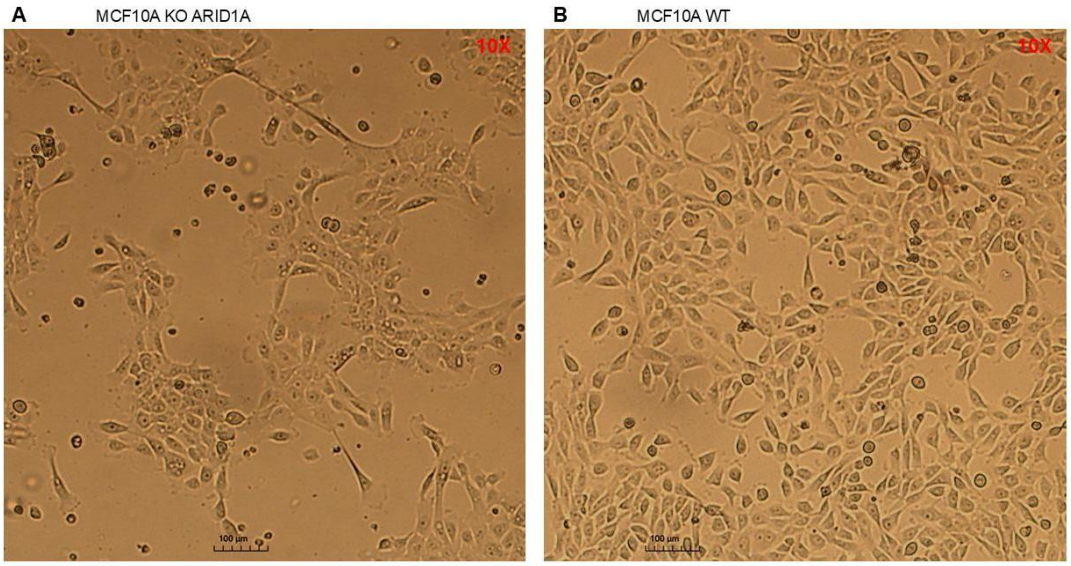

Fig.16 Morphological changes of ARID1A KO MCF10A cells:
Photo acquired in light microscopy with 10x optical zoom. Scale bar 100 um

The clone generated in this way was used to validate the activity of ARID1A in promoting TEAD binding to induce YAP target gene expression.

In order to do so, we performed ChIP-qPCR analysis for TEAD-4 in MCF10A KO compared to the WT clone. In agreement with data collected in other cell lines, ARID1A KO shows a significant reduction in TEAD binding in the promoter region of the CCN2 (CTGF) gene (Fig. 17).

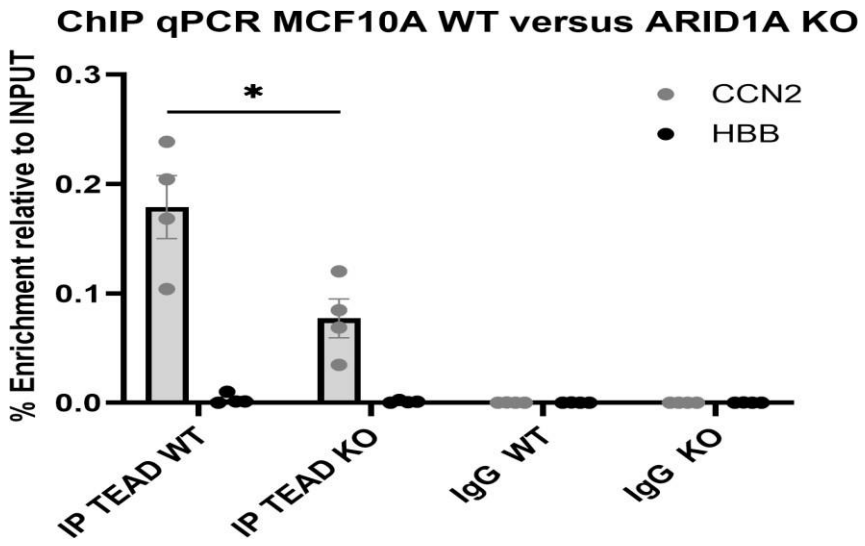

Fig. 17 ChIP-qPCR for TEAD4 in MCF10A ARID1A KO versus WT:

CCN2 enrichment is expressed as a fraction of the input for IP against TEAD4 compared to iso-specific control IgG. HBB is used as a negative control for TEAD IP. $n = 4$ Statistical analysis was performed with unpaired t-test, p-value < 0,05 (*)

5. Discussion

YAP is a co-transcriptional factor that works in conjunction with the TEAD 1-4 transcription factors to control crucial processes such as cell fate, organogenesis and response to injury[9,162]. In all these processes, YAP is able to respond to a variety of *stimuli*, but it is equally important that its activity can be switched off when not required[8]. The first reported mechanism of control of YAP is the activation of the Hippo pathway, which results in the cytoplasmic retention and degradation of YAP[163]. Loss of function of Hippo components or the convergence of other mechanisms can lead to YAP hyperactivity, which in turn is reported to drive tumorigenesis in many cancers[38].

ARID1A is an accessory protein whose presence in the SWI/SNF remodelling complex enables it to target specific chromatin regions by binding to transcription factors[153].

The activity of ARID1A through the SWI/SNF complex could result in both increased accessibility of regulatory regions on chromatin or a loss of accessibility, depending on the factors that are recruited[93]. Despite this bivalent effect on transcription, ARID1A is generally reported as a tumour suppressor gene[116]. In fact, it mainly contributes to the recruitment of factors involved in the DNA damage response[103], cell cycle progression and apoptosis[164], and its activity is lost in many tumours due to recurrent inactivating mutations[165].

To date, numerous studies have demonstrated a functional interplay between these two factors in both physiological processes and oncogenesis[128,129,131]. Nevertheless, no unified mechanistic model has been established, and instead, a context-dependent interplay has been observed.

The present PhD project offers a novel account of the complex interplay between ARID1A and YAP activity in breast epithelial cells

under high mechanical cues.

To clarify the scope of these interactions and to validate the aforementioned cell context activity we used a breast epithelial cell model competent to regulate YAP activity (MCF10A) and its transformed counterpart in which YAP activity is unleashed by H-RAS oncogene (MCF10AT)[166].

Given the primary function of ARID1A as part of the chromatin remodelling complex and the dependence of YAP on the transcription factor TEAD, we hypothesise that ARID1A functions with YAP by regulating chromatin accessibility at TEAD-bound regions.

To this end, we analysed available ATAC-seq data from HCT116, MCF7 and endometrioid cells in which ARID1A expression was abolished. In all cell lines analysed, loss of ARID1A slightly reduces chromatin accessibility, mainly at enhancer regions. Enrichment analysis for transcription factor motifs in the less accessible enhancers revealed AP-1, FOXA1 and TEAD as the most enriched. While the mechanisms underlying ARID1A mediated activation of AP-1 and FOXA1 have been thoroughly investigated,[93,155,156], the interplay of ARID1A and TEAD transcription factors is still poorly understood.

Given the fact that enhancers constitute the majority of YAP-TEAD cis-bound regions and are essential for their transcriptional activity[167], we crossed ATAC-seq results in ARID1A KO MCF7 cells with ChIP-seq data for TEAD obtained from the same cell line. In fact, loss of ARID1A in these TEAD-bound enhancers determined a strong reduction in accessibility.

Even though this effect was observed consistently across all cancer cell lines analysed, a major limitation of our omics approach is the lack of confirmatory data in the cell line used for subsequent analysis. Although MCF7 are already transformed cells, they still retain epithelial characteristics and Hippo responsiveness[168], so we assumed that these data could be a reasonable starting

point for our question. At this point we moved to our cell models and performed transcriptomic analysis in YAP or ARID1A depleted cells using siRNA. In MCF10A cells, although the overlap between these two conditions was modest, an interestingly large proportion of the genes differentially expressed in each condition were concordantly affected by depletion of both factors. Indeed, Gene Set Enrichment Analysis (GSEA) revealed that known targets of YAP, as well as genes with a TEAD1 motif in their promoter, were collectively significantly downregulated in ARID1A KD cells. Although ARID1A loss has been demonstrated to exert a dual impact on gene expression, with both positive and negative effects, in our cell context the majority of YAP target genes are observed to be repressed in response to ARID1A knockdown. Collectively, in non-tumorigenic cells such as MCF10A, our transcriptomic analysis demonstrated that ARID1A plays a role in promoting the YAP oncogenic transcriptional programme[169].

Although we observed a significant downregulation of YAP signature genes in MCF10AT in ARID1A KD (as shown in the Supplementary Table), we report a poorer overlap in the affected genes between the two cell lines analysed. In this context, we hypothesise that H-RAS may alter YAP dependencies on ARID1A for some genes, thus bypassing the limitations imposed by ARID1A loss.

We then queried differentially expressed genes for functional analysis using GO term enrichment analysis. As expected, depletion of YAP leads to a downregulation of genes associated with proliferation and migration. Surprisingly, these GO terms in most cases match the GO terms significantly enriched in the list of ARID1A-dependent genes in MCF10A cells.

Finally, we took advantage of MCF10A ARID1A KO clone to establish a mechanistic role for ARID1A in enabling TEAD-4 binding to the promoters of its main target gene CTGF[170]. We also report, as qualitative data, the acquisition of a

different cellular phenotype and behaviour for the ARID1A KO. We could hypothesise that the transcriptional changes resulted in substantial alterations of the biological processes, as reported in our functional GO term analysis. Lack of data on this aspect is a gap in our study that we aim to fill with proliferation and migration assays.

Taken together all this evidence supports a model in which ARID1A promotes TEAD binding and therefore represents a prerequisite to sustain YAP transcriptional programme in cells experiencing high mechanical stress.

Complementary to our works, Piccolo and colleagues have recently elucidated the mechanism by which the SWI/SNF complex sequesters YAP in the nucleus, thereby preventing its binding with TEAD and consequently controlling their activity[128]. In particular, ARID1A inhibits the formation of YAP-TEAD complexes exclusively in response to low mechanical cues, which correlates with a condition in MCF10A where YAP localisation is predominantly cytoplasmic. Consequently, ARID1A may regulate those YAP molecules that evade Hippo pathway negative control. This would explain the tumour-suppressor activity of ARID1A in the liver, whereby its loss renders YAP hyperactivation capable of driving oncogenesis[128].

These observations are consistent with the model presented here. Indeed, mechanosensitive cells like MCF10A, which were used in this study, were subjected to conditions that resulted in the exposure to high mechanical cues, which determine the nuclear shuttling of YAP and consequently, its binding with TEAD.

A physiological condition in which mechanotransduction releases YAP activity is reprogramming in response to injury[17]. A well-studied example of this process is the response of the liver to injury. In this instance, Li and colleagues have reported that Arid1a plays a primary role in maintaining a permissive chromatin

state, which is essential for reprogramming. To be specific, following the occurrence of injury, Arid1a directly binds to Yap on accessible chromatin sites, thereby supporting its activity in the reprogramming of hepatocytes[134]. Our transcriptomic data in cells responsive to mechanical cues strongly supports these observations.

In a pathological context, tumours were described as wounds that do not heal for the changes in microenvironments that recall first steps of injury[171]. Nowadays, we know that this is still a simplification because cancer cells are able to profoundly shape tissue to obtain a supportive microenvironment to their growth[172]. In this dynamic context, tumour cells experience increasing compressive stress due to gain of volume and interactions between cells and Extracellular matrix (ECM)[173]. All these alterations synergistically promote YAP shuttling to the nucleus, allowing its oncogenic function to be fully expressed without genetic alterations in Hippo pathway[174]. Even though YAP activity is essential for tumour progression, its hyperactivity is not able to initiate oncogenesis as a stand-alone factor. Indeed, *in vivo* injections in liver mice of hepatoblastoma cell lines harbouring a constitutively active form of YAP (S127A) was not sufficient to induce liver anomaly[175]. In this carcinoma context, ARID1A activity presented in this PhD project fits well. In fact, as already mentioned, Arid1a is able to promote cancer initiation in liver mice through promotion of driver factors like Myc or Yap[123]. Collectively, our data on gene expression and GO term analysis direct us on a model in which ARID1A loss is detrimental for YAP-dependent biological processes like proliferation and migration in cells immortalised but non tumorigenic.

It can thus be postulated that while ARID1A is necessary in the early stages of cancer, its activity is detrimental to tumour progression and metastasis. Indeed, Arid1a loss alters chromatin accessibility and thus gene expression, enabling cell-

motility and metastasis-related genes to be upregulated in the last stage of tumourigenesis[123]. In this scenario of a well-established tumour, cancer cells harbouring ARID1A null mutations and YAP hyperactivation could reshape their transcriptional programme to boost cell cycle progression and metastatic feature[176].

Further experiments will be required to validate this model and characterise the effects of simultaneous hyperactivation of YAP and loss of ARID1A in *in vitro* and *in vivo* models of cancer progression.

6. References

1. Kango-Singh, M. & Singh, A. Regulation of organ size: Insights from the Drosophila Hippo signaling pathway. *Dev. Dyn.* **238**, 1627–1637 (2009).
2. Pan, D. The Hippo Signaling Pathway in Development and Cancer. *Dev. Cell* **19**, 491–505 (2010).
3. Yu, F.-X., Zhao, B. & Guan, K.-L. Hippo Pathway in Organ Size Control, Tissue Homeostasis, and Cancer. *Cell* **163**, 811–828 (2015).
4. Driskill, J. H. & Pan, D. Control of stem cell renewal and fate by YAP and TAZ. *Nat. Rev. Mol. Cell Biol.* **24**, 895–911 (2023).
5. Ma, S., Meng, Z., Chen, R. & Guan, K.-L. The Hippo Pathway: Biology and Pathophysiology. *Annu. Rev. Biochem.* **88**, 577–604 (2019).
6. Mo, J., Park, H. W. & Guan, K. The Hippo signaling pathway in stem cell biology and cancer. *EMBO Rep.* **15**, 642–656 (2014).
7. Camargo, F. D. *et al.* YAP1 Increases Organ Size and Expands Undifferentiated Progenitor Cells. *Curr. Biol.* **17**, 2054–2060 (2007).
8. Piccolo, S., Dupont, S. & Cordenonsi, M. The biology of YAP/TAZ: hippo signaling and beyond. *Physiol. Rev.* **94**, 1287–1312 (2014).
9. Varelas, X. The Hippo pathway effectors TAZ and YAP in development, homeostasis and disease. *Development* **141**, 1614–1626 (2014).
10. Kim, M.-K., Jang, J.-W. & Bae, and S.-C. DNA binding partners of

YAP/TAZ. *BMB Rep.* **51**, 126–133 (2018).

11. Felley-Bosco, E. & Stahel, R. Hippo/YAP pathway for targeted therapy. *Transl. Lung Cancer Res.* **3**, 75–83 (2014).
12. Gong, R. & Yu, F.-X. Targeting the Hippo Pathway for Anti-cancer Therapies. *Curr. Med. Chem.* **22**, 4104–4117 (2015).
13. Lin, K. C., Park, H. W. & Guan, K.-L. Regulation of the Hippo Pathway Transcription Factor TEAD. *Trends Biochem. Sci.* **42**, 862–872 (2017).
14. Zhou, Y. *et al.* The TEAD Family and Its Oncogenic Role in Promoting Tumorigenesis. *Int. J. Mol. Sci.* **17**, 138 (2016).
15. Zhao, B. *et al.* TEAD mediates YAP-dependent gene induction and growth control. *Genes Dev.* **22**, 1962–1971 (2008).
16. Stein, C. *et al.* YAP1 Exerts Its Transcriptional Control via TEAD-Mediated Activation of Enhancers. *PLOS Genet.* **11**, e1005465 (2015).
17. Pocaterra, A., Romani, P. & Dupont, S. YAP/TAZ functions and their regulation at a glance. *J. Cell Sci.* **133**, jcs230425 (2020).
18. Totaro, A., Panciera, T. & Piccolo, S. YAP/TAZ upstream signals and downstream responses. *Nat. Cell Biol.* **20**, 888–899 (2018).
19. Yu, F.-X. *et al.* Regulation of the Hippo-YAP pathway by G-protein-coupled receptor signaling. *Cell* **150**, 780–791 (2012).
20. Dong, J. *et al.* Elucidation of a Universal Size-Control Mechanism in Drosophila and Mammals. *Cell* **130**, 1120–1133 (2007).

21. Zhang, J., Smolen, G. A. & Haber, D. A. Negative Regulation of YAP by LATS1 Underscores Evolutionary Conservation of the Drosophila Hippo Pathway. *Cancer Res.* **68**, 2789–2794 (2008).

22. Lei, Q.-Y. *et al.* TAZ Promotes Cell Proliferation and Epithelial-Mesenchymal Transition and Is Inhibited by the Hippo Pathway. *Mol. Cell. Biol.* **28**, 2426–2436 (2008).

23. Hansen, C. G., Moroishi, T. & Guan, K.-L. YAP and TAZ: a nexus for Hippo signaling and beyond. *Trends Cell Biol.* **25**, 499–513 (2015).

24. Zhao, B., Li, L., Tumaneng, K., Wang, C.-Y. & Guan, K.-L. A coordinated phosphorylation by Lats and CK1 regulates YAP stability through SCFβ-TRCP. *Genes Dev.* **24**, 72–85 (2010).

25. Dan, I., Watanabe, N. M. & Kusumi, A. The Ste20 group kinases as regulators of MAP kinase cascades. *Trends Cell Biol.* **11**, 220–230 (2001).

26. Basu, S., Totty, N. F., Irwin, M. S., Sudol, M. & Downward, J. Akt Phosphorylates the Yes-Associated Protein, YAP, to Induce Interaction with 14-3-3 and Attenuation of p73-Mediated Apoptosis. *Mol. Cell* **11**, 11–23 (2003).

27. TAZ: a novel transcriptional co-activator regulated by interactions with 14-3-3 and PDZ domain proteins. *EMBO J.* **19**, 6778–6791 (2000).

28. Wada, K.-I., Itoga, K., Okano, T., Yonemura, S. & Sasaki, H. Hippo pathway regulation by cell morphology and stress fibers. *Development* **138**,

3907–3914 (2011).

29. Barry, E. R. *et al.* Restriction of intestinal stem cell expansion and the regenerative response by YAP. *Nature* **493**, 106–110 (2013).

30. Piccolo, S., Panciera, T., Contessotto, P. & Cordenonsi, M. YAP/TAZ as master regulators in cancer: modulation, function and therapeutic approaches. *Nat. Cancer* **4**, 9–26 (2023).

31. Zhao, B., Li, L., Tumaneng, K., Wang, C.-Y. & Guan, K.-L. A coordinated phosphorylation by Lats and CK1 regulates YAP stability through SCFβ-TRCP. *Genes Dev.* **24**, 72–85 (2010).

32. Chang, S.-S. *et al.* Aurora A kinase activates YAP signaling in triple-negative breast cancer. *Oncogene* **36**, 1265–1275 (2017).

33. Lv, M. *et al.* CDK7-YAP-LDHD axis promotes D-lactate elimination and ferroptosis defense to support cancer stem cell-like properties. *Signal Transduct. Target. Ther.* **8**, 1–19 (2023).

34. Rosenbluh, J. *et al.* β-Catenin-driven cancers require a YAP1 transcriptional complex for survival and tumorigenesis. *Cell* **151**, 1457–1473 (2012).

35. Levy, D., Adamovich, Y., Reuven, N. & Shaul, Y. Yap1 phosphorylation by c-Abl is a critical step in selective activation of proapoptotic genes in response to DNA damage. *Mol. Cell* **29**, 350–361 (2008).

36. Azzolin, L. *et al.* YAP/TAZ Incorporation in the β-Catenin Destruction Complex Orchestrates the Wnt Response. *Cell* **158**, 157–170 (2014).

37. Thompson, B. J. YAP/TAZ: Drivers of Tumor Growth, Metastasis, and Resistance to Therapy. *BioEssays* **42**, 1900162 (2020).

38. Zanconato, F., Cordenonsi, M. & Piccolo, S. YAP/TAZ at the Roots of Cancer. *Cancer Cell* **29**, 783–803 (2016).

39. Sun, Z. *et al.* Prognostic Value of Yes-Associated Protein 1 (YAP1) in Various Cancers: A Meta-Analysis. *PloS One* **10**, e0135119 (2015).

40. Warren, J. S. A., Xiao, Y. & Lamar, J. M. YAP/TAZ Activation as a Target for Treating Metastatic Cancer. *Cancers* **10**, 115 (2018).

41. Weber, R. G., Sommer, C., Albert, F. K., Kiessling, M. & Cremer, T. Clinically distinct subgroups of glioblastoma multiforme studied by comparative genomic hybridization. *Lab. Investig. J. Tech. Methods Pathol.* **74**, 108–119 (1996).

42. Imoto, I. *et al.* Expression of cIAP1, a target for 11q22 amplification, correlates with resistance of cervical cancers to radiotherapy. *Cancer Res.* **62**, 4860–4866 (2002).

43. Baldwin, C., Garnis, C., Zhang, L., Rosin, M. P. & Lam, W. L. Multiple microalterations detected at high frequency in oral cancer. *Cancer Res.* **65**, 7561–7567 (2005).

44. Snijders, A. M. *et al.* Rare amplicons implicate frequent deregulation of cell fate specification pathways in oral squamous cell carcinoma. *Oncogene* **24**, 4232–4242 (2005).

45. Hermsen, M. *et al.* Chromosomal changes in relation to clinical outcome in larynx and pharynx squamous cell carcinoma. *Cell. Oncol. Off. J. Int. Soc. Cell. Oncol.* **27**, 191–198 (2005).

46. Bashyam, M. D. *et al.* Array-based comparative genomic hybridization identifies localized DNA amplifications and homozygous deletions in pancreatic cancer. *Neoplasia N. Y. N* **7**, 556–562 (2005).

47. Dai, Z. *et al.* A comprehensive search for DNA amplification in lung cancer identifies inhibitors of apoptosis cIAP1 and cIAP2 as candidate oncogenes. *Hum. Mol. Genet.* **12**, 791–801 (2003).

48. Lambros, M. B. K. *et al.* Analysis of ovarian cancer cell lines using array-based comparative genomic hybridization. *J. Pathol.* **205**, 29–40 (2005).

49. Imoto, I. *et al.* Identification of cIAP1 as a candidate target gene within an amplicon at 11q22 in esophageal squamous cell carcinomas. *Cancer Res.* **61**, 6629–6634 (2001).

50. Hélias-Rodzewicz, Z. *et al.* YAP1 and VGLL3, encoding two cofactors of TEAD transcription factors, are amplified and overexpressed in a subset of soft tissue sarcomas. *Genes. Chromosomes Cancer* **49**, 1161–1171 (2010).

51. In melanoma, Hippo signaling is affected by copy number alterations and YAP1 overexpression impairs patient survival - Menzel - 2014 - Pigment Cell & Melanoma Research - Wiley Online Library. https://onlinelibrary.wiley.com/doi/10.1111/pcmr.12249.

52. Lai, Z.-C. *et al.* Control of Cell Proliferation and Apoptosis by Mob as Tumor Suppressor, Mats. *Cell* **120**, 675–685 (2005).

53. Oh, J.-E. *et al.* Alterations in the NF2/LATS1/LATS2/YAP Pathway in Schwannomas. *J. Neuropathol. Exp. Neurol.* **74**, 952–959 (2015).

54. Miyanaga, A. *et al.* Hippo pathway gene mutations in malignant mesothelioma: revealed by RNA and targeted exon sequencing. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer* **10**, 844–851 (2015).

55. Strazisar, M., Mlakar, V. & Glavac, D. The expression of COX-2, hTERT, MDM2, LATS2 and S100A2 in different types of non-small cell lung cancer (NSCLC). *Cell. Mol. Biol. Lett.* **14**, 442–456 (2009).

56. Crequer, A. *et al.* Inherited MST1 deficiency underlies susceptibility to EV-HPV infections. *PloS One* **7**, e44010 (2012).

57. Jia, J., Zhang, W., Wang, B., Trinko, R. & Jiang, J. The Drosophila Ste20 family kinase dMST functions as a tumor suppressor by restricting cell proliferation and promoting apoptosis. *Genes Dev.* **17**, 2514–2519 (2003).

58. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997).

59. Bonev, B. & Cavalli, G. Organization and function of the 3D genome. *Nat. Rev. Genet.* **17**, 661–678 (2016).

60. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified

by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).

61. Padeken, J., Methot, S. P. & Gasser, S. M. Establishment of H3K9-methylated heterochromatin and its functions in tissue differentiation and maintenance. *Nat. Rev. Mol. Cell Biol.* **23**, 623–640 (2022).
62. Cenik, B. K. & Shilatifard, A. COMPASS and SWI/SNF complexes in development and disease. *Nat. Rev. Genet.* **22**, 38–58 (2021).
63. Clapier, C. R. & Cairns, B. R. The Biology of Chromatin Remodeling Complexes. *Annu. Rev. Biochem.* **78**, 273–304 (2009).
64. Smits, V. A. J., Alonso-de Vega, I. & Warmerdam, D. O. Chromatin regulators and their impact on DNA repair and G2 checkpoint recovery. *Cell Cycle* **19**, 2083–2093 (2020).
65. Willhoft, O. & Costa, A. A structural framework for DNA replication and transcription through chromatin. *Curr. Opin. Struct. Biol.* **71**, 51–58 (2021).
66. Beyond the Double Helix: Writing and Reading the Histone Code - Wang - 2004 - Novartis Foundation Symposia - Wiley Online Library. https://onlinelibrary.wiley.com/doi/abs/10.1002/0470862637.ch2.
67. Clapier, C. R. & Cairns, B. R. The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* **78**, 273–304 (2009).
68. Wang, Y. *et al.* Beyond the Double Helix: Writing and Reading the Histone Code. in *Reversible Protein Acetylation* 3–21 (John Wiley & Sons, Ltd, 2004). doi:10.1002/0470862637.ch2.

69. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403**, 41–45 (2000).

70. Owen-Hughes, T. Pathways for remodelling chromatin. *Biochem. Soc. Trans.* **31**, 893–905 (2003).

71. Levine, M. & Tjian, R. Transcription regulation and animal diversity. *Nature* **424**, 147–151 (2003).

72. Cosma, M. P. Ordered Recruitment: Gene-Specific Mechanism of Transcription Activation. *Mol. Cell* **10**, 227–236 (2002).

73. Tang, L., Nogales, E. & Ciferri, C. Structure and function of SWI/SNF chromatin remodeling complexes and mechanistic implications for transcription. *Prog. Biophys. Mol. Biol.* **102**, 122–128 (2010).

74. Stern, M., Jensen, R. & Herskowitz, I. Five SWI genes are required for expression of the HO gene in yeast. *J. Mol. Biol.* **178**, 853–868 (1984).

75. Neigeborn, L. & Carlson, M. GENES AFFECTING THE REGULATION OF SUC2 GENE EXPRESSION BY GLUCOSE REPRESSION IN SACCHAROMYCES CEREVISIAE. *Genetics* **108**, 845–858 (1984).

76. Peterson, C. L. & Herskowitz, I. Characterization of the yeast *SWI1, SWI2*, and *SWI3* genes, which encode a global activator of transcription. *Cell* **68**, 573–583 (1992).

77. Tamkun, J. W. *et al.* brahma: A regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF 2 SWI 2. *Cell*

68, 561–572 (1992).

78. The HMG-domain protein BAP111 is important for the function of the BRM chromatin-remodeling complex *in vivo* | PNAS. https://www.pnas.org/doi/abs/10.1073/pnas.091533398.

79. Wang, W. *et al.* Diversity and specialization of mammalian SWI/SNF complexes. *Genes Dev.* **10**, 2117–2130 (1996).

80. Wang, W. *et al.* Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. *EMBO J.* **15**, 5370–5382 (1996).

81. Dunaief, J. L. *et al.* The retinoblastoma protein and BRG1 form a complex and cooperate to induce cell cycle arrest. *Cell* **79**, 119–130 (1994).

82. Bultman, S. *et al.* A Brg1 Null Mutation in the Mouse Reveals Functional Differences among Mammalian SWI/SNF Complexes. *Mol. Cell* **6**, 1287–1295 (2000).

83. Lessard, J. *et al.* An Essential Switch in Subunit Composition of a Chromatin Remodeling Complex during Neural Development. *Neuron* **55**, 201–215 (2007).

84. Chi, T. H. *et al.* Reciprocal regulation of CD4/CD8 expression by SWI/SNF-like BAF complexes. *Nature* **418**, 195–199 (2002).

85. Lemon, B., Inouye, C., King, D. S. & Tjian, R. Selectivity of chromatin-remodelling cofactors for ligand-activated transcription. *Nature* **414**, 924–928 (2001).

86. Wu, J. I., Lessard, J. & Crabtree, G. R. Understanding the Words of Chromatin Regulation. *Cell* **136**, 200–206 (2009).

87. Mashtalir, N. *et al.* Modular Organization and Assembly of SWI/SNF Family Chromatin Remodeling Complexes. *Cell* **175**, 1272-1288.e20 (2018).

88. Odnokoz, O., Wavelet-Vermuse, C., Hophan, S. L., Bulun, S. & Wan, Y. ARID1 proteins: from transcriptional and post-translational regulation to carcinogenesis and potential therapeutics. *Epigenomics* **13**, 809 (2021).

89. Raab, J. R., Resnick, S. & Magnuson, T. Genome-Wide Transcriptional Regulation Mediated by Biochemically Distinct SWI/SNF Complexes. *PLOS Genet.* **11**, e1005748 (2015).

90. Wilsker, D. *et al.* The DNA-binding properties of the ARID-containing subunits of yeast and mammalian SWI/SNF complexes. *Nucleic Acids Res.* **32**, 1345–1353 (2004).

91. Suryo Rahmanto, Y. *et al.* Inactivation of Arid1a in the endometrium is associated with endometrioid tumorigenesis through transcriptional reprogramming. *Nat. Commun.* **11**, 2717 (2020).

92. Nagarajan, S. *et al.* ARID1A influences HDAC1/BRD4 activity, intrinsic proliferative capacity and breast cancer treatment response. *Nat. Genet.* **52**, 187–197 (2020).

93. Kelso, T. W. R. *et al.* Chromatin accessibility underlies synthetic lethality of SWI/SNF subunits in ARID1A-mutant cancers. *eLife* **6**, e30506 (2017).

94. Adelman, K. & Lis, J. T. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat. Rev. Genet.* **13**, 720–731 (2012).

95. Levine, M. Paused RNA Polymerase II as a Developmental Checkpoint. *Cell* **145**, 502–511 (2011).

96. Gilchrist, D. A. *et al.* Pausing of RNA Polymerase II Disrupts DNA-Specified Nucleosome Organization to Enable Precise Gene Regulation. *Cell* **143**, 540–551 (2010).

97. Chen, F. X. *et al.* PAF1, a Molecular Regulator of Promoter-Proximal Pausing by RNA Polymerase II. *Cell* **162**, 1003–1015 (2015).

98. Rahl, P. B. *et al.* c-Myc Regulates Transcriptional Pause Release. *Cell* **141**, 432–445 (2010).

99. Trizzino, M. *et al.* The Tumor Suppressor ARID1A Controls Global Transcription via Pausing of RNA Polymerase II. *Cell Rep.* **23**, 3933–3945 (2018).

100. Kim, J. K. *et al.* Srg3, a mouse homolog of yeast SWI3, is essential for early embryogenesis and involved in brain development. *Mol. Cell. Biol.* **21**, 7787–7795 (2001).

101. Klochendler-Yeivin, A. *et al.* The murine SNF5/INI1 chromatin remodeling factor is essential for embryonic development and tumor suppression. *EMBO Rep.* **1**, 500–506 (2000).

102. Gao, X. *et al.* ES cell pluripotency and germ-layer formation require the

SWI/SNF chromatin remodeling component BAF250a. *Proc. Natl. Acad. Sci.* **105**, 6656–6661 (2008).

103. Shen, J. *et al.* ARID1A Deficiency Impairs the DNA Damage Checkpoint and Sensitizes Cells to PARP Inhibitors. *Cancer Discov.* **5**, 752–767 (2015).
104. Tsai, S. *et al.* ARID1A regulates R-loop associated DNA replication stress. *PLOS Genet.* **17**, e1009238 (2021).
105. Harper, J. W. & Elledge, S. J. The DNA Damage Response: Ten Years After. *Mol. Cell* **28**, 739–745 (2007).
106. Ciccia, A. & Elledge, S. J. The DNA Damage Response: Making It Safe to Play with Knives. *Mol. Cell* **40**, 179–204 (2010).
107. Jackson, S. P. & Bartek, J. The DNA-damage response in human biology and disease. *Nature* **461**, 1071–1078 (2009).
108. Jones, P. A. & Baylin, S. B. The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.* **3**, 415–428 (2002).
109. Baylin, S. B. & Jones, P. A. A decade of exploring the cancer epigenome — biological and translational implications. *Nat. Rev. Cancer* **11**, 726–734 (2011).
110. Morgan, M. A. & Shilatifard, A. Chromatin signatures of cancer. *Genes Dev.* **29**, 238–249 (2015).
111. Dawson, M. A. & Kouzarides, T. Cancer Epigenetics: From Mechanism to Therapy. *Cell* **150**, 12–27 (2012).

112. Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer | Nature. https://www.nature.com/articles/BF28212.

113. Reisman, D. N. *et al.* Concomitant down-regulation of BRM and BRG1 in human tumor cell lines: differential effects on RB-mediated growth arrest vs CD44 expression. *Oncogene* **21**, 1196–1207 (2002).

114. Wong, A. K. C. *et al.* BRG1, a Component of the SWI-SNF Complex, Is Mutated in Multiple Human Tumor Cell Lines. *Cancer Res.* **60**, 6171–6177 (2000).

115. Kadoch, C. & Crabtree, G. R. Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci. Adv.* **1**, e1500447 (2015).

116. Fontana, B. *et al.* ARID1A in cancer: Friend or foe? *Front. Oncol.* **13**, (2023).

117. The AACR Project GENIE Consortium *et al.* AACR Project GENIE: Powering Precision Medicine through an International Consortium. *Cancer Discov.* **7**, 818–831 (2017).

118. Jones, S. *et al.* Frequent Mutations of Chromatin Remodeling Gene ARID1A in Ovarian Clear Cell Carcinoma. *Science* **330**, 228–231 (2010).

119. Wiegand, K. C. *et al.* ARID1A Mutations in Endometriosis-Associated Ovarian Carcinomas. *N. Engl. J. Med.* **363**, 1532–1543 (2010).

120. Bailey, M. H. *et al.* Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* **173**, 371-385.e18 (2018).

121. Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **6**, pl1 (2013).

122. Bai, J. *et al.* BRG1 Is a Prognostic Marker and Potential Therapeutic Target in Human Breast Cancer. *PLOS ONE* **8**, e59772 (2013).

123. Sun, X. *et al.* Arid1a Has Context-Dependent Oncogenic and Tumor Suppressor Functions in Liver Cancer. *Cancer Cell* **32**, 574-589.e6 (2017).

124. Zhao, J. *et al.* The Clinicopathologic Significance of BAF250a (ARID1A) Expression in Hepatocellular Carcinoma. *Pathol. Oncol. Res. POR* **22**, 453–459 (2016).

125. Gibson, W. J. *et al.* The genomic landscape and evolution of endometrial carcinoma progression and abdominopelvic metastasis. *Nat. Genet.* **48**, 848–855 (2016).

126. Totaro, A., Panciera, T. & Piccolo, S. YAP/TAZ upstream signals and downstream responses. *Nat. Cell Biol.* **20**, 888–899 (2018).

127. Panciera, T., Azzolin, L., Cordenonsi, M. & Piccolo, S. Mechanobiology of YAP and TAZ in physiology and disease. *Nat. Rev. Mol. Cell Biol.* **18**, 758–770 (2017).

128. Chang, L. *et al.* The SWI/SNF complex is a mechanoregulated inhibitor of YAP and TAZ. *Nature* **563**, 265–269 (2018).

129. Li, W. *et al.* A Homeostatic Arid1a-Dependent Permissive Chromatin State Licenses Hepatocyte Responsiveness to Liver-Injury-Associated YAP

Signaling. *Cell Stem Cell* **25**, 54-68.e5 (2019).

130. Skibinski, A. *et al.* The Hippo Transducer TAZ Interacts with the SWI/SNF Complex to Regulate Breast Epithelial Lineage Commitment. *Cell Rep.* **6**, 1059–1072 (2014).

131. Boogerd, C. J. *et al.* Cardiomyocyte proliferation is suppressed by ARID1A-mediated YAP inhibition during cardiac maturation. *Nat. Commun.* **14**, 4716 (2023).

132. Yanger, K. *et al.* Robust cellular reprogramming occurs spontaneously during liver regeneration. *Genes Dev.* **27**, 719–724 (2013).

133. Tarlow, B. D. *et al.* Bipotential Adult Liver Progenitors Are Derived from Chronically Injured Mature Hepatocytes. *Cell Stem Cell* **15**, 605–618 (2014).

134. Li, W. *et al.* A Homeostatic Arid1a-Dependent Permissive Chromatin State Licenses Hepatocyte Responsiveness to Liver-Injury-Associated YAP Signaling. *Cell Stem Cell* **25**, 54-68.e5 (2019).

135. Zhou, D. *et al.* Mst1 and Mst2 Maintain Hepatocyte Quiescence and Suppress Hepatocellular Carcinoma Development through Inactivation of the Yap1 Oncogene. *Cancer Cell* **16**, 425–438 (2009).

136. Tao, J. *et al.* Activation of β-Catenin and Yap1 in Human Hepatoblastoma and Induction of Hepatocarcinogenesis in Mice. *Gastroenterology* **147**, 690–701 (2014).

137. He, Z., Li, R. & Jiang, H. Mutations and Copy Number Abnormalities of

Hippo Pathway Components in Human Cancers. *Front. Cell Dev. Biol.* **9**, (2021).

138. Franklin, J. M., Wu, Z. & Guan, K.-L. Insights into recent findings and clinical application of YAP and TAZ in cancer. *Nat. Rev. Cancer* **23**, 512–525 (2023).
139. Jiang, T., Chen, X., Su, C., Ren, S. & Zhou, C. Pan-cancer analysis of ARID1A Alterations as Biomarkers for Immunotherapy Outcomes. *J. Cancer* **11**, 776–780 (2020).
140. Labun, K. *et al.* CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.* **47**, W171–W174 (2019).
141. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
142. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
143. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype | Nature Biotechnology. https://www.nature.com/articles/s41587-019-0201-4.
144. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032–2034 (2015).
145. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*

29, 15 (2012).

146. Terms and Definitions – ENCODE. https://www.encodeproject.org/data-standards/terms/#enrichment.

147. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).

148. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinforma. Oxf. Engl.* **26**, 841–842 (2010).

149. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).

150. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

151. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

152. Kolberg, L., Raudvere, U., Kuzmin, I., Vilo, J. & Peterson, H. gprofiler2 -- an R package for gene list functional enrichment analysis and namespace conversion toolset g:Profiler. *F1000Research* **9**, ELIXIR-709 (2020).

153. Suryo Rahmanto, Y. *et al.* Inactivation of Arid1a in the endometrium is associated with endometrioid tumorigenesis through transcriptional reprogramming. *Nat. Commun.* **11**, 2717 (2020).

154. Heinz, S. *et al.* Simple Combinations of Lineage-Determining Transcription

Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **38**, 576–589 (2010).

155. Vierbuchen, T. *et al.* AP-1 Transcription Factors and the BAF Complex Mediate Signal-Dependent Enhancer Selection. *Mol. Cell* **68**, 1067-1082.e12 (2017).

156. Wilson, M. R., Reske, J. J. & Chandler, R. L. AP-1 Subunit JUNB Promotes Invasive Phenotypes in Endometriosis. *Reprod. Sci.* **29**, 3266–3277 (2022).

157. Xu, G. *et al.* ARID1A determines luminal identity and therapeutic response in estrogen-receptor-positive breast cancer. *Nat. Genet.* **52**, 198–207 (2020).

158. Transforming properties of YAP, a candidate oncogene on the chromosome 11q22 amplicon. https://www.pnas.org/doi/epdf/10.1073/pnas.0605579103 doi:10.1073/pnas.0605579103.

159. Plunkett, C. *et al.* H-Ras Transformation of Mammary Epithelial Cells Induces ERK-Mediated Spreading on Low Stiffness Matrix. *Adv. Healthc. Mater.* **9**, 1901366 (2020).

160. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).

161. M, C. *et al.* The Hippo transducer TAZ confers cancer stem cell-related traits on breast cancer cells. *Cell* **147**, (2011).

162. Wu, Z. & Guan, K.-L. Hippo Signaling in Embryogenesis and Development. *Trends Biochem. Sci.* **46**, 51–63 (2021).

163. Ma, S., Meng, Z., Chen, R. & Guan, K.-L. The Hippo Pathway: Biology and Pathophysiology. *Annu. Rev. Biochem.* **88**, 577–604 (2019).

164. Mandal, J., Mandal, P., Wang, T.-L. & Shih, I.-M. Treating ARID1A mutated cancers by harnessing synthetic lethality and DNA damage response. *J. Biomed. Sci.* **29**, 71 (2022).

165. Kadoch, C. *et al.* Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat. Genet.* **45**, 592–601 (2013).

166. Franklin, J. M., Ghosh, R. P., Shi, Q., Reddick, M. P. & Liphardt, J. T. Concerted localization-resets precede YAP-dependent transcription. *Nat. Commun.* **11**, 4581 (2020).

167. Zanconato, F. *et al.* Genome-wide association between YAP/TAZ/TEAD and AP-1 at enhancers drives oncogenic growth. *Nat. Cell Biol.* **17**, 1218–1227 (2015).

168. Ma, S. *et al.* Hippo signalling maintains ER expression and ER+ breast cancer growth. *Nature* **591**, E1–E10 (2021).

169. Zanconato, F. *et al.* Genome-wide association between YAP/TAZ/TEAD and AP-1 at enhancers drives oncogenic growth. *Nat. Cell Biol.* **17**, 1218–1227 (2015).

170. Holbourn, K. P., Acharya, K. R. & Perbal, B. The CCN family of proteins: structure–function relationships. *Trends Biochem. Sci.* **33**, 461–473 (2008).

171. Dvorak, H. F. Tumors: Wounds that do not heal--Redux. *Cancer Immunol. Res.* **3**, 1 (2015).

172. Arneth, B. Tumor Microenvironment. *Medicina (Mex.)* **56**, 15 (2020).

173. Blanco, B. *et al.* Mechanotransduction in tumor dynamics modeling. *Phys. Life Rev.* **44**, 279–301 (2023).

174. Piccolo, S., Panciera, T., Contessotto, P. & Cordenonsi, M. YAP/TAZ as master regulators in cancer – modulation, function and therapeutic approaches. *Nat. Cancer* **4**, 9 (2022).

175. Tao, J. *et al.* Activation of β-Catenin and Yap1 in Human Hepatoblastoma and Induction of Hepatocarcinogenesis in Mice. *Gastroenterology* **147**, 690–701 (2014).

176. Fetiva, M. C. *et al.* Oncogenic YAP mediates changes in chromatin accessibility and activity that drive cell cycle gene expression and cell migration. *Nucleic Acids Res.* **51**, 4266–4283 (2023).

7. Supplementary

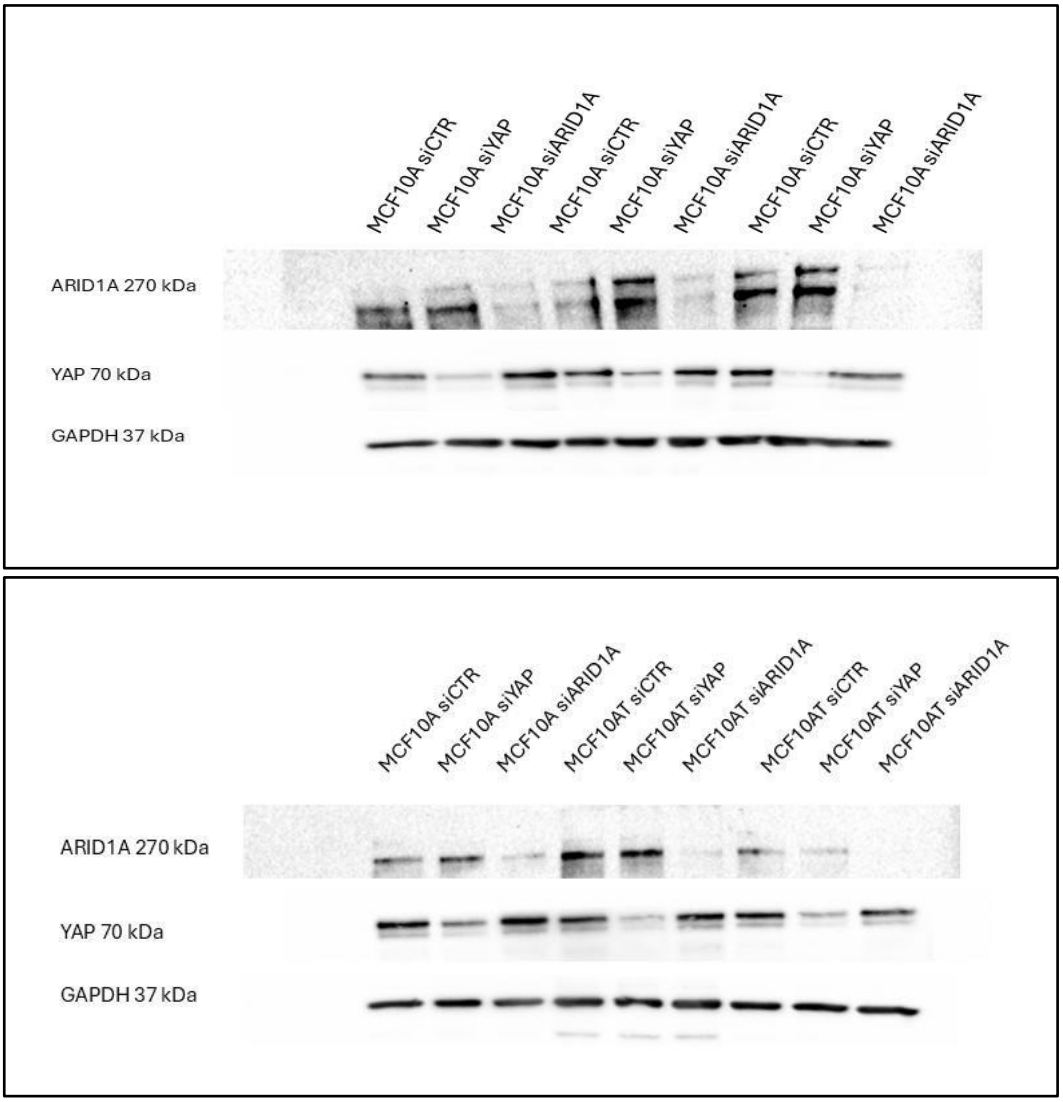

Fig.1 Total protein extracts were subjected to WB analysis to ascertain silencing efficacy for YAP or ARID1A in MCF10A and MCF10AT cells.

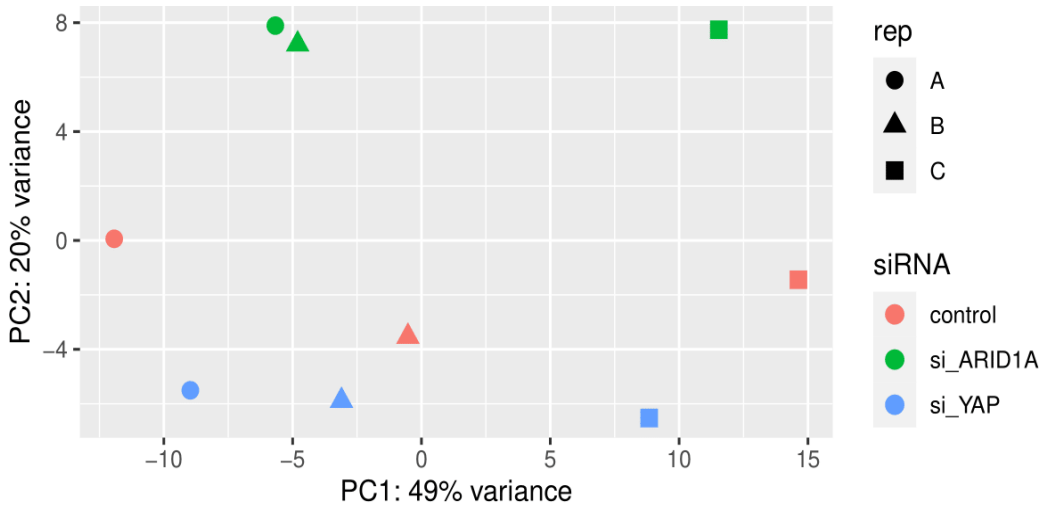

Fig.2 Principal component analysis (PCA) on RNA-seq data derived from YAP KD or ARID1A KD in MCF10A cells, compared to control. No procedure of batch effect removal was performed.

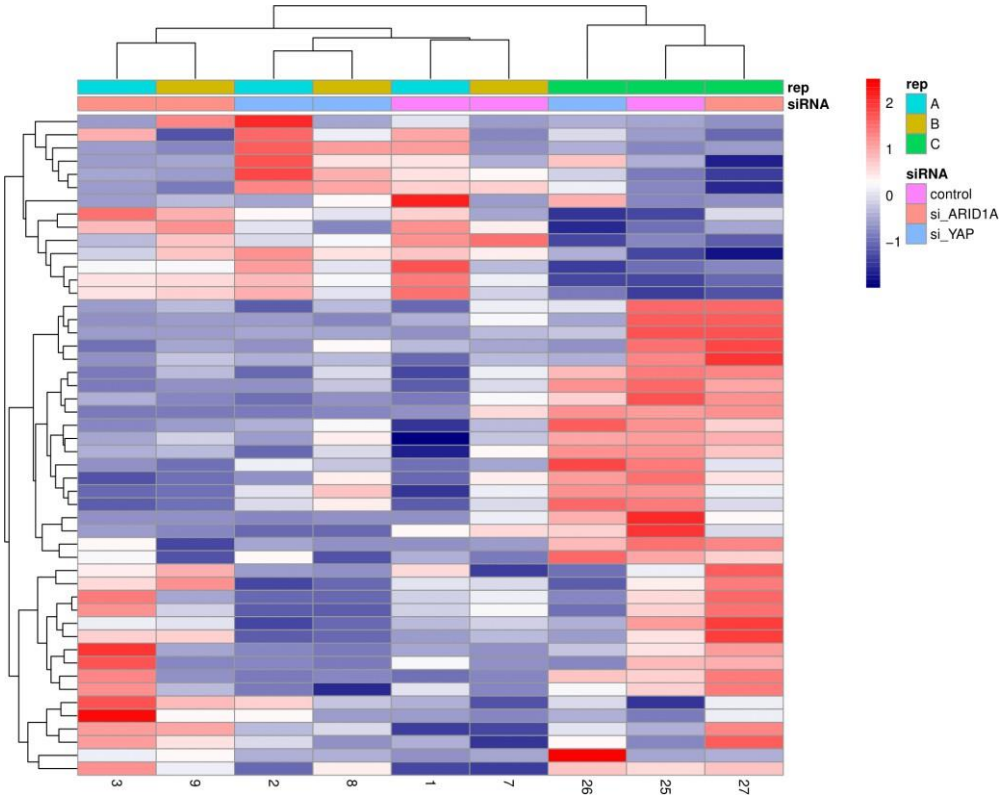

Fig.3 Heatmap of top 50 genes for variance in MCF10A siRNA treated cells and control:
The top 50 genes were selected for variance in their expression, independent of experimental design. No procedure of batch effect removal was performed. The Z-score on the right relates differential gene expression to a colour scale. $n = 3$

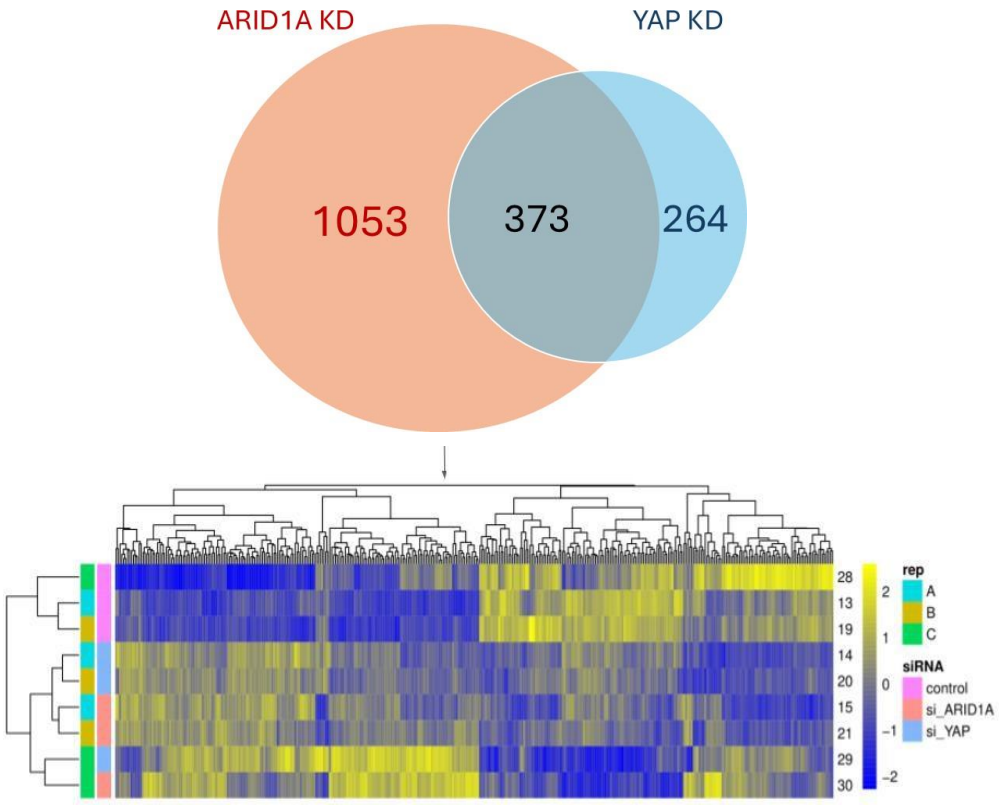

Fig. 4 Transcriptome analysis of ARID1A or YAP depleted MCF10AT cells:
Venn diagram shows differentially expressed genes in the two analysed conditions. Intersections between ARID1A and YAP KD count on 373 genes, 350 of which are concordantly affected. The analysed gene expression is collected in a heatmap which hierarchically clusters conditions. Each individual gene of the human transcriptome occupies one column, intersecting one row for each condition analysed. The Z-score on the right relates differential gene expression to a colour scale. $n = 3$

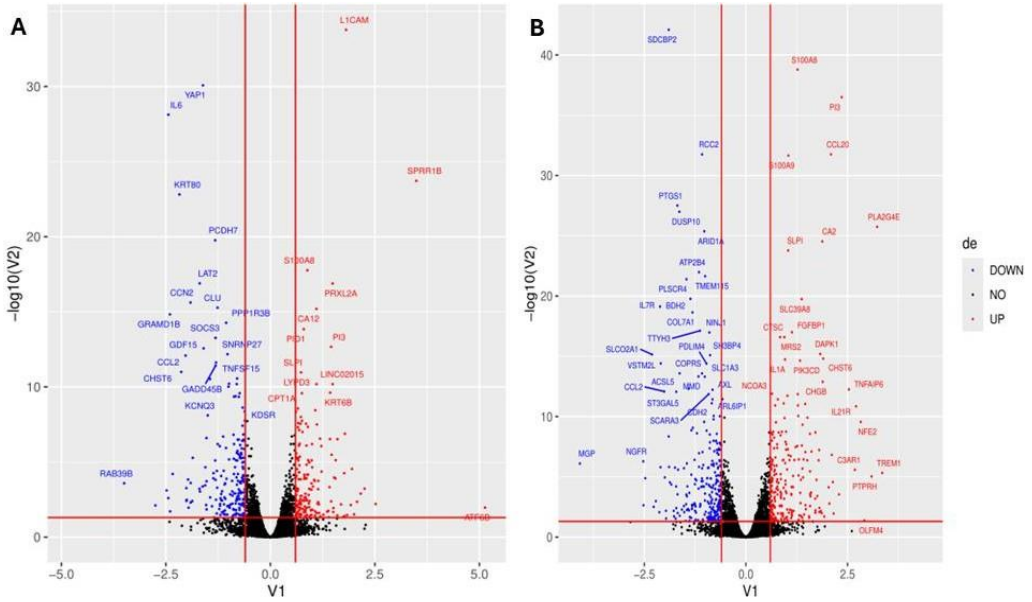

Fig. 5A Volcano plot for transcriptomic data in YAP KD MCF10A cells compared to control. The *x*-axis shows the amount of variance in expression, while the *y*-axis shows the statistical significance (-log10 P value). Fig. 5B Volcano plot for transcriptomic data in ARID1A KD MCF10A cells compared to control.

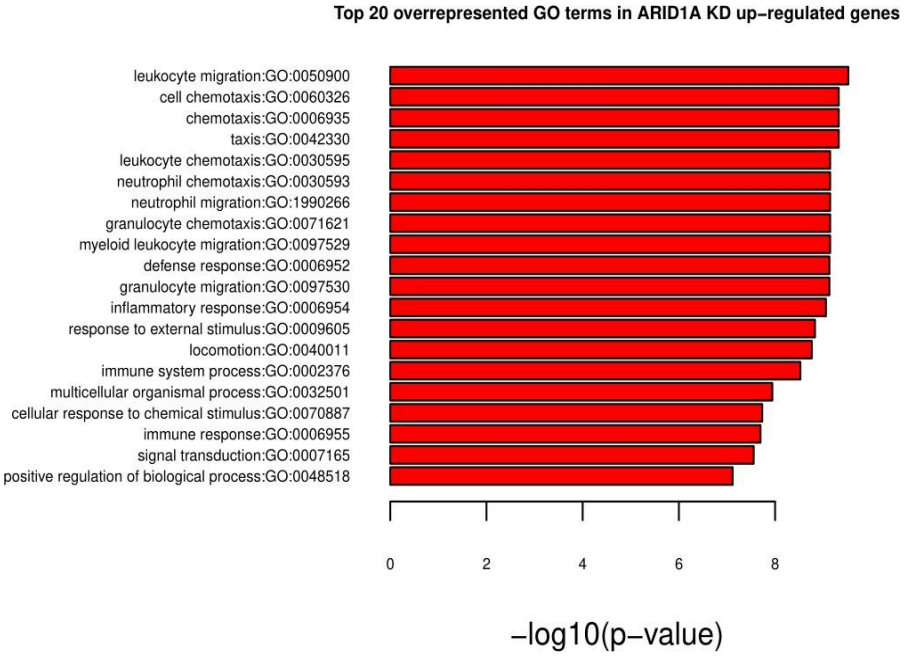

Fig.6 GO terms enrichment analysis in ARID1A KD top 20 up-regulated genes in MCF10A

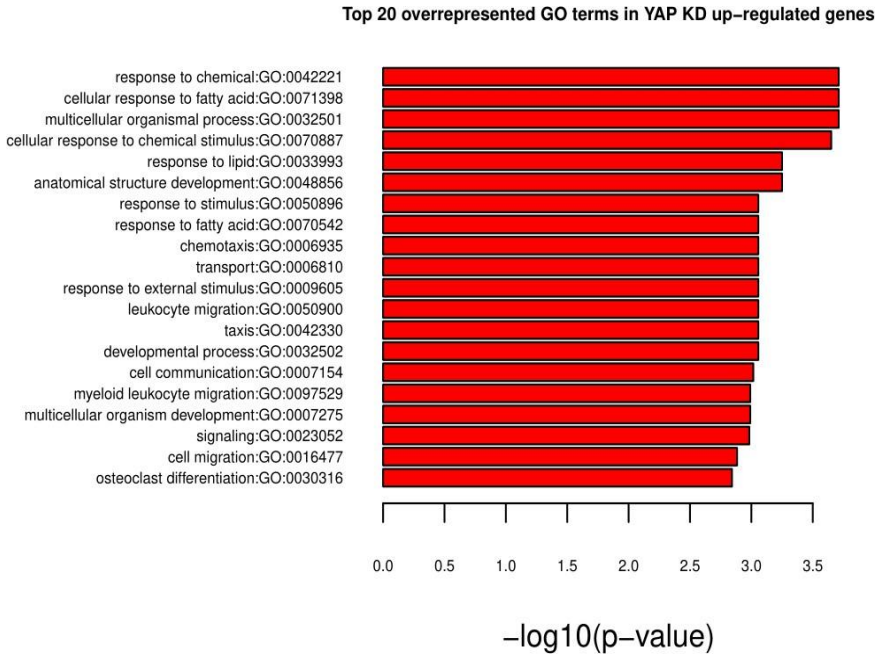

Fig.7 GO terms enrichment analysis in ARID1A KD top 20 up-regulated genes in MCF10A

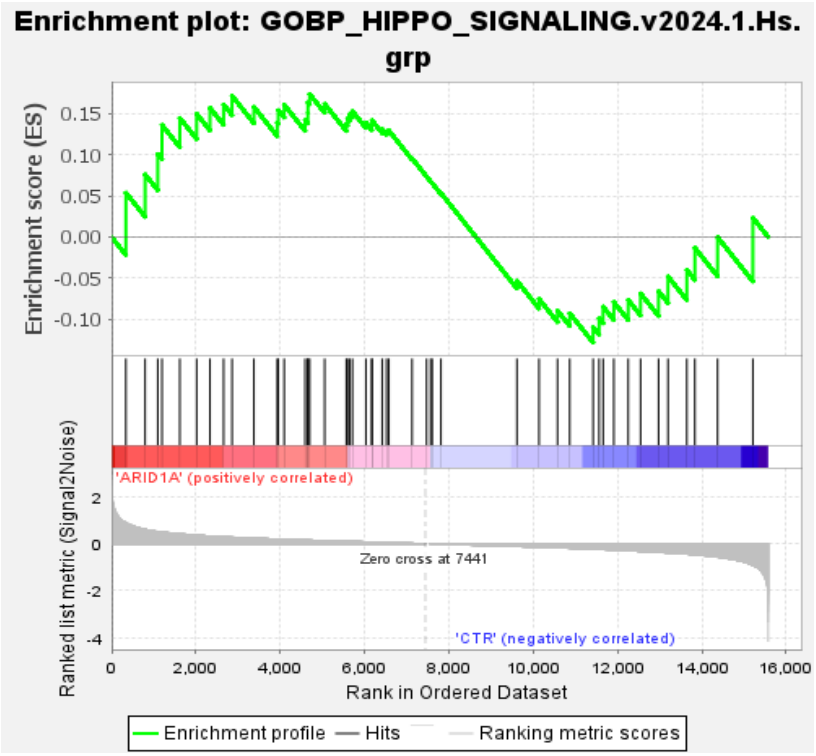

Fig.8 GSEA enrichment plot for the gene set "Hippo signalling" in ARID1A-depleted cells compared to control. There are no significant differences in the enrichment score between the two conditions, indicating that ARID1A does not upregulate the transcription of factors involved in the negative control of YAP activity.

```
ARID1A          CAGCCATACGGGTCCCAGACCCCGCA-GCGGTACCCGATGACCATGCAGGGCCGGGCG
10_T7           CAGCCATACGGGTCCCAGACCCCGCAAGCGGTACCCGATGACCATGCAGGGCCGGGCG
11_T7           CAGCCATACGGGCCCCAGACCCCGCAAGCGGTACCCGATGACCATGCAGGGCCGGGCG
1_T7            CAGCCATACGGGTCCCAGACCCCGCAAGCGGTACCCGATGACCATGCAGGGCCGGGCG
9_T7            CAGCCATACGGGTCCCAGACCCCGCAAGCGGTACCCGATGACCATGCAGGGCCGGGCG
12_T7           CAGCCATACGGGTCCCAGACCCCGCAAGCGGTACCCGATGACCATGCAGGGCCGGGCG
                ************ ************* ********************************
```

Fig.9 Fasta files of the genomic region (exon 2) spanning the target site of Crispr-Cas9 in ARID1A coding sequence. In red is the insertion of 1 base in the KO clone, compared to the WT sequence.

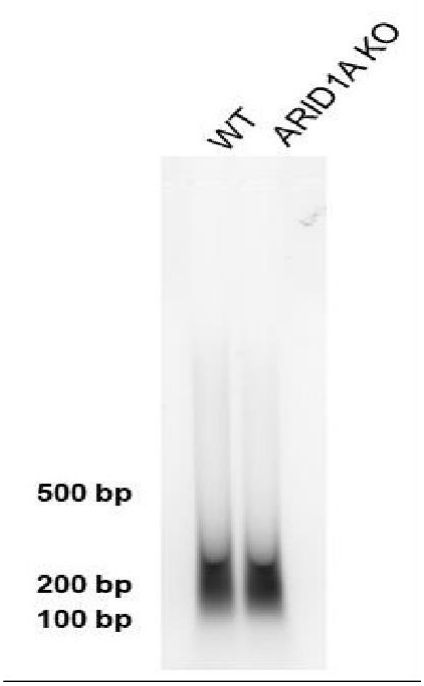

Fig. 10 Size control of sheared chromatin used for ChIP analysis
After decrosslinking and purification protocol, sheared chromatin was run on 1.2% agarose gel to verify the optimal range of fragments between 150-500 bp.

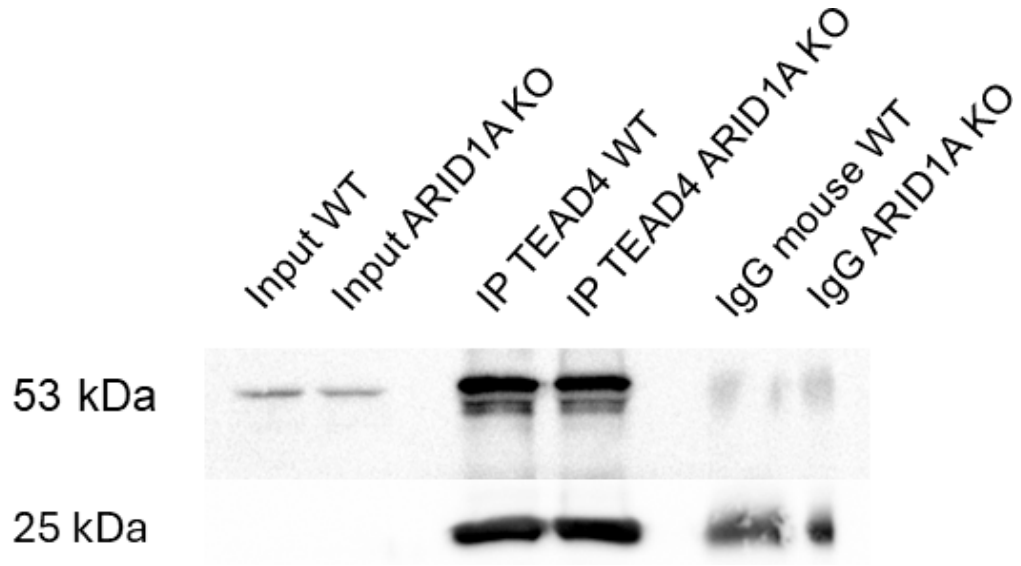

Fig.11 WB for IP against TEAD-4 in ChIP experiment
The same quantity of isospecific IgG was employed as a negative control for IP. A secondary antibody anti-light chain mouse (HRP) conjugated was employed.
10% *v/v* of IP was carried out from sheared chromatin and loaded as Input.

8. Data Availability

ATAC-seq MCF-7:
ATAC-seq data: SRA: PRJNA554641; GEO: GSE134270
ChIP-seq and HMM Chromatin states (ENCODE):
TEAD4 ChIP: ENCODE ENCFF751VAZ
HMM_Chrom_states: ENCODE ENCFF985EWD

ATAC-seq HCT116:
ATAC-seq data: PRJNA396039; GEO: GSE101966
TEAD4 ChIP: ENCODE ENCFF412ZJN
HMM_Chrom_states: ENCODE ENCFF513PJK

ATAC-seq Endometrial epithelial cells
ATAC-seq data SRA: PRJNA417567; GEO: GSE106658
TEAD4 ENCODE ENCFF974NUK
HMM_Chrom_states for Z12 cell line were generously shared by Prof Chandler (https://www.sciencedirect.com/science/article/pii/S2211124720313553?via%3Dihub)

Reference data:
TSS hg38 reference: ENCODE ENCFF493CCB
ENCODE Unified GRCh38 Blacklist: ENCODE ENCFF356LFX
Gene annotation (RNA-seq): Gencode basic gene annotation v43(GRCh38.p13)

R sessionInfo()

```
R version 4.4.1 (2024-06-14)
Platform: x86_64-pc-linux-gnu
Running under: Ubuntu 22.04.3 LTS

Matrix products: default
BLAS:   /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.10.0
LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.10.0

locale:
 [1] LC_CTYPE=en_US.UTF-8          LC_NUMERIC=C
 [3] LC_TIME=en_GB.UTF-8           LC_COLLATE=en_US.UTF-8
 [5] LC_MONETARY=en_GB.UTF-8       LC_MESSAGES=en_US.UTF-8
 [7] LC_PAPER=en_GB.UTF-8          LC_NAME=C
 [9] LC_ADDRESS=C                  LC_TELEPHONE=C
[11] LC_MEASUREMENT=en_GB.UTF-8    LC_IDENTIFICATION=C

time zone: Europe/Rome
tzcode source: system (glibc)

attached base packages:
[1] stats4stats     graphics  grDevices utils     datasets
[7] methods   base
```

other attached packages:
 [1] limma_3.60.6 ggplot2_3.5.1
 [3] DESeq2_1.44.0 SummarizedExperiment_1.34.0
 [5] Biobase_2.64.0 MatrixGenerics_1.16.0
 [7] matrixStats_1.4.1 GenomicRanges_1.56.2
 [9] GenomeInfoDb_1.40.1 IRanges_2.38.1
[11] S4Vectors_0.42.1 BiocGenerics_0.50.0
[13] RColorBrewer_1.1-3 pheatmap_1.0.12

loaded via a namespace (and not attached):
 [1] generics_0.1.3 utf8_1.2.4
 [3] SparseArray_1.4.8 lattice_0.20-45
 [5] magrittr_2.0.3 grid_4.4.1
 [7] jsonlite_1.8.9 Matrix_1.4-0
 [9] httr_1.4.7 fansi_1.0.6
[11] UCSC.utils_1.0.0 scales_1.3.0
[13] codetools_0.2-18 abind_1.4-8
[15] cli_3.6.3 rlang_1.1.4
[17] crayon_1.5.3 XVector_0.44.0
[19] munsell_0.5.1 withr_3.0.1
[21] DelayedArray_0.30.1 S4Arrays_1.4.1
[23] tools_4.4.1 parallel_4.4.1
[25] BiocParallel_1.38.0 dplyr_1.1.4
[27] colorspace_2.1-1 locfit_1.5-9.10
[29] GenomeInfoDbData_1.2.12 vctrs_0.6.5
[31] R6_2.5.1 lifecycle_1.0.4
[33] zlibbioc_1.50.0 pkgconfig_2.0.3
[35] pillar_1.9.0 gtable_0.3.5
[37] glue_1.8.0 Rcpp_1.0.13
[39] statmod_1.5.0 tidyselect_1.2.1
[41] tibble_3.2.1 compiler_4.4.1

9. Acknowledgements

I would like to express my deepest gratitude to my supervisor, Professor Valerio Fulci, for his unwavering support and insightful critiques throughout my research journey. His deep commitment to academic excellence and meticulous attention to detail have significantly shaped this dissertation.

I am also grateful for the ministerial funding of the PRIN 2017 project, through which Professor Fulci has made this Ph.D. project possible.