TAZ-CAMTA1 and YAP-TFE3 modulate the basal TAZ/YAP transcriptional program by recruiting the ATAC histone acetyltransferase complex

Nicole Merritt^{1†}, Keith Garcia^{1,2†}, Dushyandi Rajendran³, Zhen-Yuan Lin³, Xiaomeng Zhang⁴, Katrina M. Mitchell^{4,5}, Nicholas Borchering⁶, Colleen Fullenkamp¹, Michael Chimenti⁷, Anne-Claude Gingras³, Kieran F. Harvey^{4,5,8}, and Munir R. Tanas^{1,2,9*}.

1. Department of Pathology, University of Iowa, Iowa City, IA.
2. Cancer Biology Graduate Program, University of Iowa, Iowa City, IA.
3. Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, ON.
4. Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia.
5. Sir Peter MacCallum Department of Oncology, The University of Melbourne, Parkville, Victoria, Australia.
6. Department of Pathology and Immunology, Washington University, St. Louis, MO.
7. Iowa Institute of Human Genetics, Carver College of Medicine, University of Iowa, Iowa City, IA.
8. Department of Anatomy and Developmental Biology, and Biomedicine Discovery Institute, Monash University, Clayton, Australia.
9. Holden Comprehensive Cancer Center, University of Iowa, Iowa City, IA. †Indicates equal contribution.

*Address correspondence to:

Munir R. Tanas
Assistant Professor
Department of Pathology
University of Iowa
25 S. Grand Ave., 1023 ML
Iowa City, IA 52242

Phone: 319-335-6637

FAX: 319-335-8453

E-mail: munir-tanas@uiowa.edu

Summary: TAZ-CAMTA1 and YAP-TFE3 alter the TAZ/YAP transcriptional program by recruiting the ATAC complex and modifying the chromatin landscape.

Abstract: Epithelioid hemangioendothelioma (EHE) is a vascular sarcoma that metastasizes early and lacks an effective medical therapy. The TAZ-CAMTA1 and YAP-TFE3 fusion proteins are chimeric transcription factors and initiating oncogenic drivers of EHE. A combined proteomic/genetic screen identified YEATS2 and ZZZ3, components of the Ada2a-containing histone acetyltransferase (ATAC) complex, as key interactors of both TAZ-CAMTA1 and YAP-TFE3 despite the dissimilarity of the C terminal fusion partners CAMTA1 and TFE3. An integrative next generation sequencing approach showed the fusion proteins drive expression of a unique transcriptome distinct from TAZ and YAP by simultaneously hyperactivating a TEAD-based transcriptional program and modulating the chromatin environment via interaction with the ATAC complex. Interaction of the ATAC complex with both TAZ-CAMTA1 and YAP-TFE3 indicates the histone acetyltransferase complex is an oncogenic driver in EHE and potentially other sarcomas. Furthermore, the ATAC complex is an enzymatic transcriptional cofactor required for both fusion proteins in EHE, representing a unifying therapeutic target for this sarcoma. Gene fusions are the most common genetic alterations activating TAZ and YAP in cancer, and this study serves as a template for identifying epigenetic modifiers recruited by the C terminal fusion partners of other TAZ/YAP gene fusions occurring in gliomas, carcinomas, and other sarcomas.

Introduction: Epithelioid hemangioendothelioma (EHE) is a vascular sarcoma that arises over a wide age range and can occur essentially anywhere anatomically. When it occurs in the lung or the liver, it frequently presents with metastasis early in its clinical course (1). Two mutually exclusive gene fusions define this sarcoma which lacks an effective medical therapy. The *WWTR1-CAMTA1* (2) (3) (4) gene fusion is encoded by a t(1;3) (p36;q25) chromosomal translocation, resulting in a TAZ-CAMTA1 fusion

protein present in 85-90% of all EHE tumors (3). The other ~10-15% of EHE contain a t(X;11) (p11;q22) chromosomal translocation encoding a *YAP1-TFE3* gene fusion (5). Homologous to TAZ-CAMTA1, YAP-TFE3 fuses the N-terminus of YAP in frame to the C-terminus of TFE3 (transcription factor E3) (5).

The Hippo pathway is a serine/threonine kinase cascade involved in regulating tissue growth and organ size and is composed of the STE20-like protein kinases 1 and 2 (MST1/2; homologs of *Drosophila* Hpo) (6-11) and the large tumor suppressor 1 and 2 (LATS1/2; homologs of *Drosophila* Wts) (12-14). In mammalian cells the nuclear effectors of the Hippo pathway are TAZ (gene is *WWTR1*) and YAP, two paralogous transcriptional coactivators (10, 15) (11). Hippo signaling limits TAZ and YAP activity by phosphorylating various serine residues, promoting their accumulation in the cytoplasm and ubiquitin-mediated degradation (10) (11). In the nucleus, TAZ and YAP bind to the TEA domain (TEAD) family of transcription factors to drive their transcriptional program (16) (17). TAZ and YAP have emerged as central oncoproteins (18) in many different cancers including breast (19) (20), colorectal (21), liver (22), lung (23), pancreas (24), thyroid (25), and sarcomas (26-30). Despite this observation, point mutations within the Hippo signaling pathway are relatively rare in most cancers (18). This is especially true of YAP and TAZ, which instead appear to be more commonly mutated by virtue of gene fusions including *WWTR1(TAZ)-CAMTA1* and *YAP1-TFE3*.

TAZ-CAMTA1 transforms cells by driving a predominantly TAZ-related transcriptional program that is mediated by binding to TEAD4 (4). Although homologous regions of YAP and TAZ are found in the YAP-TFE3 and TAZ-CAMTA1 fusion proteins, functional studies of YAP-TFE3 have not been performed. Furthermore, the mechanism by which these hybrid transcription factors regulate transcription and subsequent cell transformation has not been defined. Herein we identify the Ada2a-containing histone acetyltransferase (ATAC) complex as a key epigenetic modifier of the TAZ-CAMTA1 and YAP-TFE3 oncogenic transcriptional programs. The ATAC complex is one of two metazoan complexes that

incorporate the GCN5 or closely related PCAF histone acetyltransferase subunit, and primarily acetylate lysine 9 of histone 3 (H3K9) and to a lesser degree lysine 14 (H3K14) (31).

There is emerging evidence that the ATAC complex is an oncogenic driver in cancer. *YEATS2*, one of the subunits of the ATAC complex, has been shown to be functionally important to the pathogenesis of non-small cell lung cancer (32). Amplification of *YEATS2* has been identified in ovarian cancer, head and neck cancer, esophageal cancer, and uterine cancer (32), as well as both well-differentiated and dedifferentiated liposarcoma (33). Herein we show that the ATAC complex is a key epigenetic mediator of the oncogenic properties of both the TAZ-CAMTA1 and YAP-TFE3 fusion proteins and thereby a common therapeutic target for essentially all EHE.

Results:

TAZ-CAMTA1 and YAP-TFE3 are chimeric transcription factors

The *WWTR1* (TAZ)-*CAMTA1* gene fusion fuses the end of exon 2 or exon 3 of *WWTR1* to a breakpoint with exon 9 of *CAMTA1* (2, 3). This results in the N terminus of TAZ being fused in frame to the C terminus of CAMTA1. The N terminus of TAZ contains the TEAD binding domain and 14-3-3 protein binding site (34). The WW domain may be absent or present, depending on the type of gene fusion. The C terminus of CAMTA1 contributes the transactivating domain, the TIG domain, ankyrin domain, IQ domain, and nuclear localization signal (**Figure 1A**) (35) (4). Similarly, *YAP1*-*TFE3* fuses exon 1 from *YAP1* in frame to exon 4 of *TFE3* resulting in the N terminus of YAP fused in frame to the C terminus of TFE3. The N terminus of YAP contains the TEAD binding domain (36) (**Figure 1A**). The C terminus of TFE3 contains the transactivating domain, basic helix loop helix domain, and leucine zipper (5, 37). TFE3 is a member of the MiTF family of basic helix-loop-helix containing transcription factors, consisting of 4 structurally similar genes (*TFE3*, *TFE3B*, *TFEC*, and *MiTF*), that are involved in regulating tissue-specific functions of cell differentiation (37, 38). TFE3 is ubiquitously expressed and activates transcription by binding to the μ E3 motif within the immunoglobulin heavy-chain enhancer (39, 40). TFE3 has recently been shown to promote cell survival in starvation conditions by regulating

lysosomal biogenesis and homeostasis while also having a role in clearance of cellular debris (41, 42).

TFE3 is also involved in gene fusions in renal cell carcinoma and alveolar soft part sarcoma (37, 43, 44).

YAP-TFE3 requires components of both YAP and TFE3 to transform cells

In order to dissect the function of the YAP-TFE3 fusion protein, we expressed it in NIH 3T3 cells, which have been utilized to study various components of the Hippo pathway (11),(16),(17). We also expressed the fusion protein in SW872 cells, a human liposarcoma cell line, chosen because it has an attenuated ability to grow in an anchorage independent manner in either soft agar or poly-HEMA (poly 2-hydroxyethyl methacrylate; data not shown) and because there are no EHE cell lines available. Both TAZ-CAMTA1 and YAP-TFE3 were able to promote colony formation in soft agar in SW872 cells (**Figure S2E**). To investigate YAP-TFE3's ability to transform cells, full-length YAP, full-length TFE3, the truncated portion of YAP or TFE3 involved in the fusion protein, as well as YAP-TFE3 were expressed in NIH 3T3 cells (**Figure S1**) and evaluated in soft agar (**Figure 1B**) and poly-HEMA (**Figure 1C**). YAP-TFE3 promoted anchorage independent growth while full length YAP, full length TFE3, or the truncated portions of YAP or TFE3 involved in the fusion protein did not. Expressing the same constructs in SW872 cells (**Figure S1**) and evaluating proliferation in poly-HEMA yielded similar results (**Figure 1D**). The above findings indicate that YAP-TFE3 is a neomorphic protein combining properties of both YAP and TFE3 to transform cells.

YAP and TAZ dynamically shuttle between the cytoplasm and nucleus at a rate which is regulated by both the Hippo pathway and Hippo pathway-independent mechanisms (45). Furthermore, the abundance of both YAP and TAZ is limited by ubiquitin-mediated degradation which is controlled by the Hippo pathway (46, 47). YAP and TAZ transit between the nucleus and cytoplasm and their relative abundance in these subcellular compartments is modulated by multiple features including cell confluence and mechanotransduction cues (48) (45). Like YAP, we found YAP-TFE3 to be localized within the nucleus when cells were plated under sparse conditions. Under confluent conditions, YAP was enriched in the cytoplasm and its abundance was reduced, while YAP-TFE3 remained predominantly nuclear (**Figure**

1E-F; Figure S2A-B). Nuclear and cytoplasmic fractionation confirmed that YAP is no longer localized in the nucleus and instead accumulates in the cytoplasm where it is subsequently degraded. In contrast, YAP-TFE3 remains in the nucleus during cell confluence (**Figure S2C**). Similarly, YAP localization shifts from the nucleus into the cytoplasm upon cell detachment, while YAP-TFE3 remains constitutively within the nucleus (**Figure S2D**). Taken together, the above data suggest that YAP-TFE3 is no longer subject to Hippo pathway-mediated suppression, likely because the key residues targeted by LATS1/2 (S127 and S381) (11) are not present in YAP-TFE3.

YAP-TFE3 and TAZ-CAMTA1 drive tumor initiation and proliferation *in vivo*

To study the function of YAP/TAZ fusion proteins *in vivo*, NIH 3T3 cells or SW872 cells transduced with empty vector, TAZ-CAMTA1, or YAP-TFE3 were injected into subcutaneous tissue of *Prkdc^{scid}Il2rg^{tm1Wjl}/SzJ* (NSG) mice. Mice with NIH 3T3-derived tumors expressing YAP-TFE3 or TAZ-CAMTA1 demonstrated a shorter overall survival compared to empty vector ($p < 0.0001$ for both YAP-TFE3 and TAZ-CAMTA1) (**Figure 1G**). Tumors expressing YAP-TFE3 and TAZ-CAMTA1 formed palpable masses at ~7 days post-injection (**Figure 1H**) indicating time to tumor initiation was similar for NIH 3T3 cells expressing either fusion protein. NIH 3T3 cells containing empty vector formed tumors, as has been described (49), but tumor initiation was delayed by approximately 10 days. The slopes of the growth curves for the two fusion proteins were similar to empty vector, and this was confirmed by Ki-67 labeling (**Figure S3A and B**). Mice were sacrificed when tumors reached 2 cm in the greatest dimension (according to animal use protocol). NIH 3T3 cell-derived tumors expressing YAP-TFE3 or TAZ-CAMTA1 exhibited a more rounded (epithelioid) cytomorphology and a greater degree of cellular pleomorphism as compared to control tumors (**Figure S3A**). Necropsy and pathological evaluation of the lungs demonstrated the presence of microscopic metastases in the YAP-TFE3 ($p = 0.0299$) but not the TAZ-CAMTA1 (not significant) cohorts (**Figure 1I**). Expression of YAP-TFE3 and TAZ-CAMTA1 in SW872 cells promoted tumor growth *in vivo* (**Figure 1J**) similar to that seen in NIH 3T3 cells. Cytological changes in SW872 cell-derived tumors were less pronounced than in NIH 3T3 cells (**Figure**

S3B). The above data show that both fusion proteins play a role in tumor initiation and can drive tumorigenesis *in vivo*.

TEAD transcription factors mediate the oncogenic activity of both YAP-TFE3 and TAZ-CAMTA1

TAZ and YAP are transcriptional co-activators that lack DNA binding domains and require binding with transcription factors containing DNA binding domains to drive transcription. The best characterized of these transcription factors are TEAD1-4. TAZ and YAP interact with TEAD1-4 via the TEAD binding domain in the YAP and TAZ amino termini (16, 17), while the WW domains of YAP and TAZ mediate interactions with other transcription factors including TBX5 (50), Runx2 (51), PPAR- γ (51) and SMADs (52)). We previously showed that TEAD4 is required for the oncogenic function of TAZ-CAMTA1 (4). To determine if the WW domain is essential for cell transformation, alanine was substituted for the two essential tryptophans in the WW domain of TAZ-CAMTA1 (TAZ-CAMTA1 Δ WW). While the S51A mutant (disrupts TEAD binding) of TAZ-CAMTA1 significantly reduced colony formation in soft agar/growth on poly-HEMA (4), the TAZ-CAMTA1 Δ WW mutant did not (**Figure S2F and G**). This confirms that TAZ-CAMTA1 function is dependent on the TEAD binding domain but not the WW domain.

YAP-TFE3 (**Figure 1A**) contains the TEAD binding domain, but not the WW domain, suggesting that the TEAD binding domain mediates the YAP-TFE3 transcriptional program. To test this hypothesis, Tead1 was knocked down in NIH-3T3 cells expressing YAP-TFE3 (**Figure 2A; Table S1**). Knock-down of Tead1 reduced anchorage independent growth in soft agar (**Figure 2B**) and poly-HEMA (**Figure 2C**) with two different short hairpin RNAs (shRNAs). A YAP-TFE3 S94A mutant was created to disrupt its interaction with TEAD (homologous to S94 in full-length YAP and required for TEAD binding (16)), which was confirmed by co-immunoprecipitation (**Figure 2D**). The S94A mutation completely abrogated YAP-TFE3 driven anchorage independent growth in soft agar (**Figure 2E**) and poly-HEMA (**Figure 2F**). By quantitative RT-PCR, the YAP-TFE3 S94A mutant showed a 7-fold decrease in expression of *Ctgf*, a canonical YAP transcriptional target, compared to the YAP-TFE3 control (**Figure 2G**). Using the

8XGTTC luciferase reporter containing 8 TEAD consensus binding sequences (53), YAP-TFE3 S94A showed decreased luciferase activity compared to the YAP-TFE3 control (**Figure 2H**). Collectively, these data suggest that YAP-TFE3, like TAZ-CAMTA1, requires TEAD transcription factors to drive its oncogenic transcriptional program (**Figure 2I**).

TAZ-CAMTA1 and YAP-TFE3 induce overlapping transcriptomes but differ compared to TAZ and YAP

From this working model, we hypothesized that since TEAD binding is essential for YAP-TFE3 and TAZ-CAMTA1 driven transformation, the YAP-TFE3 and TAZ-CAMTA1 transcriptomes would essentially recapitulate the YAP and TAZ transcriptional programs, respectively. To test this hypothesis, RNA-Seq was performed on NIH 3T3 cells and SW872 cells stably expressing TAZ4SA (hyperactivated form of TAZ due to alanine being substituted for 4 serines that are phosphorylated by the LATS1/2 kinases (54)), TAZ-CAMTA1, CAMTA1, YAP5SA (hyperactivated form of YAP due to alanine being substituted for 5 serines that are phosphorylated by the LATS1/2 kinases (11)), YAP-TFE3, and TFE3, as well as cells transduced with empty vector control (**Figure S1; Table S2 and S3**). Differentially expressed genes (FDR < 0.05) showed considerable overlap between the TAZ4SA and TAZ-CAMTA1 transcriptomes; 70% of genes differentially expressed in the TAZ-CAMTA1 transcriptome were also present in the TAZ4SA transcriptome with a $-\log_{10}[\text{hypergeometric density}]$ (HGD) of 105.6 ($p = 1.92 \times 10^{-106}$). By comparison, a much smaller percentage of overlap (9%) was identified between the TAZ-CAMTA1 and CAMTA1 gene sets with an HGD of 34.1 ($p = 4.71 \times 10^{-35}$). A significant percentage (29%) of differentially expressed genes in the TAZ-CAMTA1 transcriptional program were not induced by either TAZ4SA or CAMTA1 (**Figure 3A; Table S2**) in NIH 3T3 cells. Similarly, 31% of the TAZ-CAMTA1 gene set in SW872 cells were unique and not found in the TAZ or CAMTA1 transcriptional programs (**Figure 3E; Table S3**). Like NIH 3T3 cells, the amount of overlap between TAZ-CAMTA1 and TAZ4SA was greater (67% of TAZ-CAMTA1 genes; HGD of 72.6) than that between TAZ-CAMTA1 and CAMTA1 (9% of TAZ-CAMTA1 genes; HGD of 27.6).

The YAP-TFE3 transcriptional program demonstrated similar findings. In NIH 3T3 cells, 78% of genes differentially expressed in the YAP-TFE3 transcriptome were present in the YAP5SA transcriptional program. Comparatively, only 10% of differentially expressed genes in the YAP-TFE3 gene set were also present in the TFE3 transcriptome. In NIH 3T3 cells 20% of the differentially expressed genes were unique to the YAP-TFE3 transcriptome and not found in the YAP5SA or TFE3 transcriptional programs (**Figure 3B; Table S2**). Similarly, in SW872 cells 47% of the YAP-TFE3 transcriptional program was found only in the YAP-TFE3 transcriptome and not in YAP5SA or TFE3 controls (**Figure 3F; Table S3**). Mirroring findings in NIH 3T3 cells, the amount of overlap between YAP-TFE3 and YAP5SA (53% of YAP-TFE3 genes; HGD of 220.1) was greater than that between YAP-TFE3 and TFE3 (0.2% of YAP-TFE3 genes; HGD of 1.48).

Principal component analysis (PCA) was then performed in NIH 3T3 cells to determine whether variance seen in differential expression could be used to group similar transcriptomes. This analysis showed that the TAZ-CAMTA1 and YAP-TFE3 transcriptomes grouped more closely together than the transcriptomes of their cognate full length proteins (**Figure 3C**). In SW872 cells, the TAZ-CAMTA1 and YAP-TFE3 transcriptomes diverges from one another, but were still distinct from the TAZ4SA and YAP5SA transcriptomes (**Figure S4A**). This is further support that the fusion proteins drive a transcriptional program that incorporates some elements of the TAZ/YAP transcriptomes but is also significantly altered from the baseline TAZ and YAP transcriptional program. This was corroborated by qRT-PCR experiments, which showed that TAZ-CAMTA1 and YAP-TFE3 promote upregulation of *bona fide* TAZ/YAP transcriptional targets *CTGF* and *CYR61*, but also genes unique to the TAZ-CAMTA1 and YAP-TFE3 transcriptomes such as *Fras1* (an extracellular matrix protein), *ErbB3* (Erb-B2 Receptor Tyrosine Kinase 3) (**Figure 3D**), and *FBLN5* (an extracellular matrix protein) (**Figure 3I**).

iPathwayGuide analysis performed on SW872 cells showed that the top 6 cancer-related signatures activated by TAZ-CAMTA1 in SW872 cells (**Figure 3G**) are the PI3K-Akt signaling pathway, Hippo signaling pathway (**Figure S5A**), focal adhesion, proteoglycans, and extracellular matrix (ECM)-receptor interaction (mirroring qRT-PCR results in **Figure 3D and I**). The top 6 cancer-related signatures

activated by YAP-TFE3 in SW872 cells (Figure 3H) also included the PI3K-Akt signaling pathway, as well as the MAPK and JAK-STAT signaling pathways. A microRNA and apoptosis signature were also identified.

The TAZ-CAMTA1 and YAP-TFE3 chromatin binding profiles are enriched for both TEAD and non-TEAD transcription factor motifs

To determine if the modified transcriptional programs of the fusion proteins were due to altered DNA binding, we performed chromatin immunoprecipitation (ChIP) sequencing studies on SW872 cells using Flag-tagged TAZ4SA, TAZ-CAMTA1, CAMTA1, YAP5SA, YAP-TFE3, and TFE3 constructs. Chromatin and cross-linked proteins were immunoprecipitated with an anti-flag antibody. Annotation of ChIP peaks with respect to their nearest gene transcriptional start site (TSS) demonstrated that CAMTA1 and TFE3 ChIP peaks were more localized in close proximity to the TSS (0-1 kb). In contrast, TAZ4SA, TAZ-CAMTA1, YAP5SA, and YAP-TFE3 predominantly occupied distal intergenic sequences (10 kb or greater) consistent with enhancer sequences (**Figure 4A-D**). Thus, TAZ-CAMTA1 and YAP-TFE3 maintain the genomic occupancy characteristics of increased preference for distal intergenic and intronic sites demonstrated by TAZ and YAP (55-57).

ChIP peaks were annotated to their nearest gene features (as described in Methods). Mirroring the RNA-Seq data, 56% of TAZ-CAMTA1 genes containing ChIP peaks were present in the TAZ-CAMTA1 data set, but not the TAZ4SA or CAMTA1 data sets (**Figure 4E**). Similarly, 46% of YAP-TFE3 genes containing ChIP peaks were unique to the YAP-TFE3 data set and not present in the YAP5SA or TFE3 controls (**Figure 4F**). Overall, the chromatin binding profile of TAZ-CAMTA1 more closely resembled TAZ4SA (HGD of 52.9) than CAMTA1 (HGD of 43.8). Likewise, YAP-TFE3 bound chromatin was more like that of YAP5SA (HGD of 166.8) than TFE3 (HGD of 11.8). Unbiased motif enrichment analysis (**Table S4**) showed that approximately the same proportion of TAZ4SA and TAZ-CAMTA1 peaks contained the TEAD binding sequence (58% vs. 61%, respectively) (**Figure 4E**). Interestingly, there was a decrease in the frequency of FOS/JUN binding sites in TAZ-CAMTA1 (31%) as compared to TAZ4SA (50%), which appeared to be offset by an increase in Early growth response gene-2 (EGR2)

binding motifs (23% vs. 13% for TAZ4SA). EGR2 is a transcription factor important to neural development (58) as is CAMTA1 (59). The majority of CAMTA1 ChIP peaks (69%) contained the EGR2 motif, suggesting that the C terminus of CAMTA1 interacts with EGR2 and contributes additional DNA binding activity to TAZ-CAMTA1.

Motif enrichment analysis showed a decrease in the fraction of YAP-TFE3 bound peaks containing the TEAD binding sequence (38%) as compared to YAP5SA bound peaks (51%). Microphthalmia-associated transcription factor (MiTF) is a member of the TFE/MiTF transcription factor family known to complex with TFE3 (38). Interestingly, the MiTF binding motif was enriched in YAP-TFE3 (58%) compared to YAP5SA (16%), suggesting that, like TAZ-CAMTA1, TFE3 confers DNA binding properties to the YAP-TFE3 protein in addition to events mediated by YAP's interaction with TEAD transcription factors. This analysis explains, at least in part, our observation that the chromatin binding profiles of the fusion proteins overlap significantly, but also differ from, full length TAZ and YAP.

To validate the ChIP-Seq experiment, a plot was generated that arrayed differentially expressed genes in terms of increasing statistical significance along the Y axis, and genes in terms of increasing ChIP peak “score” (i.e. statistical significance) along the X axis for TAZ-CAMTA1 (**Figure 4G**), YAP-TFE3, and controls (**Figure S4 B-F; Table S5**). This analysis revealed well known YAP/TAZ target genes such as *CCN1* (*CYR61*), *CCN2* (*CTGF*), *AMOTL2*, *AJUBA*, *LATS2*, and *TEAD1*, and genes that have not been identified as YAP/TAZ targets, such as *MAFK*. 13% of differentially expressed genes for TAZ-CAMTA1 were directly bound by the transcription factor (**Figure S5B**). Similarly, 15% of differentially expressed genes for YAP-TFE3 were directly bound by the fusion protein (**Figure S5C**). Kyoto encyclopedia of genes and genomes (KEGG) analysis of TAZ-CAMTA1 bound genes that are also differentially expressed in SW872 TAZ-CAMTA1 cells showed enrichment for RHO GTPase dependent pathways, EPHA dependent pathways, RUNX2 dependent pathways, as well as other pathways known to be regulated by TAZ/YAP (51, 53, 60) (**Figure S5D**). Genes bound by the fusion proteins or controls were additionally validated by ChIP-qPCR (**Figure 4H**).

TAZ-CAMTA1 and YAP-TFE3 augment the euchromatin landscape

Previous studies indicated that TAZ and YAP form physical complexes with multiple different with chromatin modifying proteins (61). To test the hypothesis that the fusion proteins modify chromatin structure, Assay for Transposase-Accessible Chromatin (ATAC)-sequencing studies were performed on SW872 cells expressing the above constructs. Heat map and histogram analysis of ATAC peaks showed that TAZ-CAMTA1 and CAMTA1 preferentially opened up the chromatin structure in areas of the genome more distal to the TSS than TAZ4SA (**Figure 5A, Figure S6A, B**). Although less pronounced, YAP-TFE3 preferentially opened chromatin architecture more distal to the TSS than YAP5SA (**Figure 5B and Figure S6A, B**). These results are in keeping with the genomic occupancy profiles of the fusion proteins defined by ChIP-Seq (**Figure 4C, D**) which were predominantly bound to enhancer sequences.

Analysis of differentially accessible chromatin annotated to the nearest gene feature (as described in Methods; note that multiple ATAC peaks could be assigned to the same gene) (**Figure 5C, D, and Figure S6C, D**) showed that 15,779 genes with annotated ATAC peaks were shared between TAZ-CAMTA1 and all of the controls while 16,809 genes were shared between YAP-TFE3 and all of the controls, representing the baseline levels of transposase accessible chromatin. We initially hypothesized that the accessible chromatin landscape for the fusion proteins would overwhelmingly resemble that of full length TAZ and YAP. However hypergeometric analysis showed that the accessible chromatin profile of TAZ-CAMTA1 most closely resembled that of CAMTA1 (**Figure 5C**), while the $-\log_{10}(\text{hypergeometric density})$ HGD between YAP-TFE3 and YAP5SA (106.0) was roughly equivalent to that between YAP-TFE3 and TFE3 (98.0) (**Figure 5D**). This analysis (**Figure 5C, D**) also showed that both TAZ-CAMTA1 and YAP-TFE3 promoted chromatin accessibility for unique genes not found in the control data sets, making them available for either transcriptional activation or repression. More transposase accessible, non-baseline genes were present with expression of TAZ-CAMTA1 (3147 genes) and YAP-TFE3 (2499 genes) than TAZ4SA (656 genes) and YAP5SA (1629 genes), respectively. 'Diffbind' differential binding affinity analysis of consensus-derived peak sets was carried out with respect to empty vector

controls. Dimensionality reduction with PCA showed that TAZ-CAMTA1 and YAP-TFE3 samples clustered together with CAMTA1 and TFE3, suggesting a greater degree of similarity in ATAC peak affinity patterns, while TAZ4SA or YAP5SA were distinct and separate (**Figure 5E**). The results are largely consistent with the hypergeometric analysis mentioned above. The homology of the resultant chromatin landscape for the two fusion proteins, despite the structural dissimilarity of the C termini of CAMTA1 and TFE3, is another piece of evidence supporting a convergent function for the two fusion proteins. Integrating annotation results from all three NGS modalities (RNA-, ChIP-, and ATAC-seq) showed that the majority of genes bound by the fusion proteins, as well as those which were differentially expressed, were within euchromatin regions of the genome. (**Figure 5F, G**). To further evaluate whether the fusion protein DNA binding sites were nested within transposase accessible chromatin, we evaluated the combined ChIP-Seq and ATAC-Seq data for *MAFK* and *HOXA1*, two genes that were bound by at least one or both of the fusion proteins and simultaneously differentially expressed (**Figure 4G; Figure S4F**). Evaluating the combined ChIP/ATAC signal tracks for *MAFK* and *HOXA1* in TAZ-CAMTA1 and YAP-TFE3 expressing cells (**Figure S6E, F**) suggest that the fusion proteins combine the DNA binding activity of the TEAD domains of TAZ/YAP with the chromatin remodeling properties of CAMTA1 and TFE3.

The TAZ-CAMTA1 and YAP-TFE3 interactomes are enriched for transcriptional regulatory proteins

To address the hypothesis that CAMTA1 and TFE3 confer altered chromatin remodeling properties to TAZ-CAMTA1 and YAP-TFE3, we utilized proximity-dependent biotinylation (BioID) and tandem mass spectrometry. TAZ-CAMTA1, YAP-TFE3, and full length controls were fused on their N terminus to BirA*, an abortive biotin ligase, in a tetracycline-inducible vector, and stably transfected in Flp-In T-Rex HEK 293 cells. Proteins in proximity to the BirA*-tagged bait proteins were biotinylated, purified, and analyzed by tandem mass spectrometry and statistical analysis (**Figure 6A**). Prey proteins detected with at least 10 spectral counts in at least one biological replicate were represented in dot plot format as a

composite of the relative abundance and spectral count, where the size of the dots represents the relative abundance and darkness represents the spectral count (the false discovery rate is represented as the color coding at the periphery of the dot). 68 prey proteins were identified that established high-confidence proximity interactions with the fusion proteins or the full-length controls (**Figure S7A-C**). The YAP-TFE3 proximity interactome was composed predominantly of histone modifying enzymes (35%), transcriptional regulators (26%), proteins involved in chromatin remodeling (18%), and transcription factor/co-activators (15%) (**Figure 6B**). Similarly, the TAZ-CAMTA1 proximity interactome was composed predominantly of histone modifying enzymes (27%), transcriptional regulators (18%), transcription factor/co-activators (15%), and chromatin remodeling (13%). A smaller percentage of the TC interactome was involved in mediating cell-cell junctions (5%) and SUMO processing (4%). To identify proteins to prioritize for further study, we subtracted proteins that showed a decreased proximity interaction with the fusion proteins relative to full-length TAZ and YAP; many of which (e.g. MPDZ, PATJ) are part of the Crumbs complex, and the decreased interaction can be explained in part by loss/mitigation of the WW domain. This left 55 and 34 prey proteins in the TAZ-CAMTA1 and YAP-TFE3 interactomes, respectively. Known components of the Hippo pathway were subtracted from further analysis, leaving 49 proteins in the TAZ-CAMTA1 data set, and 31 proteins in the YAP-TFE3 data set. We then retained only prey proteins that demonstrated an increased interaction with the fusion proteins relative to full length TAZ and YAP, leaving 47 proteins in the TAZ-CAMTA1 data set and 28 remaining proteins in the YAP-TFE3 data set. The two data sets were intersected to demonstrate 27 shared proteins (**Figure 6C-E**). As shown in the table listing the function of the shared proteins (**Figure 6D**) and a chart giving an overview of their composition (**Figure 6E**), the shared prey proteins were enriched for epigenetic/transcriptional regulators. Proteins with weak interactions with both fusion proteins (generally these were in the lowest category of spectral count/relative abundance and FDR >5%) were eliminated, leaving 18 proteins. The observation that so many proximity interactions were shared between TAZ-CAMTA1 and YAP-TFE3, despite having very different C-termini, indicated these shared prey proteins

may be important to the oncogenic mechanism of TAZ-CAMTA1 and YAP-TFE3. For this reason, these proteins were prioritized for the subsequent RNAi screen (overall algorithm is summarized in **Figure 6F**).

The ATAC complex mediates the transforming potential of TAZ-CAMTA1

To further prioritize these 18 prey proteins, an RNAi screen was performed on SW872 cells expressing TAZ-CAMTA1 using 5 short hairpin constructs (shRNA) for each gene/protein (**Table S6**). These lines were subsequently evaluated by poly-HEMA assay (**Figure 7A**) utilizing the observation that TAZ-CAMTA1 promoted anchorage independent growth in SW872 cells on poly-HEMA (**Figure 7B**) and in soft agar (**Figure S2E**). The number of shRNAs that demonstrated a deficit in anchorage independent growth as compared to shEV and shNT were recorded to rank order the genes/proteins in the interactome for further study (**Figure 7C**). Genes that demonstrated 5 shRNAs promoting a deficit in anchorage independent growth included *YEATS2* (**Figure 7C, D**), *ZZZ3* (**Figure 7C, E**), and *EP400* (**Figure 7C, F**). Other genes/prey proteins with more than 2 shRNAs showing a phenotype included *NCOA2* (**Figure 7G**) and *TCF20* (**Figure 7H**). Remaining prey proteins/genes with at least one shRNA showing a phenotype are also shown (**Figure 7C; Figure S8A-F**). *YEATS2* knock-down and a reduction in anchorage independent growth was validated in SW872 cells expressing YAP-TFE3 using the two shRNAs with the best knock-down (**Figure S8G-I**).

Significantly, *YEATS2* and *ZZZ3*, the two top ranking hits in the RNAi screen, are both members of the Ada2a-containing histone acetyltransferase complex (ATAC). *YEATS2* is the scaffolding protein for other subunits of the ATAC complex and binds to H3K27ac via its YEATS domain (62). The ZZ domain of *ZZZ3* binds the histone H3 tail and in combination with *YEATS2* anchors the ATAC complex to acetylated H3 (63). These proteins bind to other proteins in the complex including Ada2a as well the GCN5 catalytic subunit (62-64). GCN5 or the closely related PCAF subunit (65) are responsible for acetylating H3K9, which subsequently opens up the surrounding chromatin structure promoting transcription (62). In addition, KAT14, a third subunit of the ATAC complex, was identified by BioID mass spectrometry to interact with TAZ-CAMTA1 and YAP-TFE3 (**Figure 6C**), further emphasizing the

functional significance of this complex and an additional impetus to prioritize *YEATS2* and *ZZZ3* for further study.

Knock-down of *YEATS2* in SW872 TAZ-CAMTA1 cells was confirmed by quantitative RT-PCR (**Figure 8A**) and resulted in loss of anchorage independent growth by soft agar assay (**Figure 8B**). Similarly, knock-down of *ZZZ3* in SW872 TAZ-CAMTA1 cells was confirmed by western blot (**Figure 8C**) and the loss of anchorage independent phenotype was replicated by soft agar assay (**Figure 8D**). We then performed co-immunoprecipitation studies to validate the BioID experiments. Of the three subunits of the ATAC complex identified by BioID mass spectrometry, KAT14 showed the strongest interaction with TAZ-CAMTA1 and YAP-TFE3 by co-immunoprecipitation (data not shown). Pull-down of Flag-TAZ-CAMTA1 in HEK293 cells transfected with Myc-KAT14 demonstrated that TAZ-CAMTA1 binds the KAT14 subunit of the ATAC complex (**Figure 8E**). Similarly, Flag-YAP-TFE3 cells co-immunoprecipitated with Myc-KAT14 in HEK 293 cells (**Figure 8F**). Co-immunoprecipitation studies similarly demonstrated an interaction between Flag-TAZ-CAMTA1 and Myc-tagged KAT14 in SW872 cells (**Figure 8G**). Investigation into the role of the ATAC complex in sarcomas is limited, although amplification of *YEATS2* has also been identified in sarcomas including well differentiated liposarcoma and dedifferentiated liposarcoma (33). To determine if expression of *YEATS2* and *ZZZ3* could be prognostically important in sarcoma clinical samples, we utilized RNA-Seq expression data from The Cancer Genome Atlas (TCGA) database. Via Kaplan-Meier analysis, we showed that increased expression of *YEATS2* (**Figure 8H**) or *ZZZ3* (**Figure 8I**) predicted a poorer prognosis across different histological types of sarcoma, suggesting the ATAC complex may be a key oncogenic driver in multiple sarcomas in addition to EHE.

Discussion:

TAZ-CAMTA1 and YAP-TFE3 drive a predominantly TEAD-based transcriptional program

Our current studies show that rather than simply recapitulating the TAZ and YAP transcriptional programs, TAZ-CAMTA1 and YAP-TFE3 drive a transcriptome that overlaps with, but is significantly different from that of each full-length protein. Our studies show that the degree to which the fusion protein transcriptomes do overlap with the TAZ/YAP transcriptional programs is due to shared TEAD transcription factor binding. This is borne out via evaluation of the chromatin binding profiles of the two fusion proteins which showed similar levels of enrichment for TEAD binding sites compared to full length TAZ and YAP. Our findings are in keeping with previous functional studies with TAZ-CAMTA1 showing that it requires binding to TEAD4 in order to drive its oncogenic transcriptional program and transform cells (4). In this way, the fusion proteins reproduce a key characteristic of full length TAZ and YAP in that they require binding to the TEAD family of transcription factors in order to drive their oncogenic transcriptional programs (18, 48). Among genes that were specifically induced by TAZ-CAMTA1 and YAP-TFE3 were select members of the PI3 kinase-Akt and MAP kinase pathways, and genes encoding extracellular matrix proteins. As such, additional studies are warranted to further dissect the downstream functional consequences of the TAZ-CAMTA1 and YAP-TFE3 transcriptional programs in EHE in order to fully understand how these proteins drive this disease.

The fusion protein transcriptomes differ from the full length TAZ/YAP transcriptomes due to altered DNA binding

Although the transcriptional programs of TAZ-CAMTA1 and YAP-TFE3 significantly overlap with the TAZ/YAP transcriptomes, they also differ in important ways. In both cell lines examined here, approximately 20-40% of differentially expressed genes were unique to the TAZ-CAMTA1- and YAP-TFE3-induced transcriptomes, indicating that the fusion proteins do not only regulate transcription via TEADs. Accordingly, unbiased motif enrichment analysis of TAZ-CAMTA1 revealed an enrichment for the EGR2 binding motif while YAP-TFE3 showed an enrichment for the MiTF binding motif. This indicates that the genomic occupancy of TAZ-CAMTA1 and YAP-TFE3 is mediated by the C termini of CAMTA1 and TFE3 in addition to interactions with the TEAD transcription factors. Algorithms

predicting the binding sites of transcription factors that contain multiple DNA binding sites, or are influenced by the DNA binding sites of interacting transcription factors have been generated (66). These bioinformatic approaches combined with structure-function studies should yield additional insight into how these chimeric transcription factors interface with the genome.

The fusion protein transcriptomes differ from the TAZ/YAP transcriptomes due to shifts in the chromatin landscape

We found that TAZ-CAMTA1 and YAP-TFE3 promoted marked changes in the chromatin landscape. In general, the fusion proteins increased the amount of transcriptionally active chromatin when compared to TAZ and YAP overexpression. Although the genes embedded in transcriptionally accessible portions of the genome showed a degree of overlap between the fusion proteins and TAZ/YAP, hypergeometric analysis showed that the euchromatin landscape for TAZ-CAMTA1 more closely resembled that of CAMTA1 while that of YAP-TFE3 showed an almost equivalent level of similarity to TFE3 as it did to YAP. This observation is reflected in the PCA analysis, which showed that the distribution of transposase accessible chromatin for the fusion proteins was more closely related to each other, CAMTA1, and TFE3 than to TAZ/YAP. This indicates that CAMTA1 and TFE3 confer chromatin remodeling properties to the fusion proteins, while the N termini of TAZ and YAP essentially contribute the TEAD binding domain. This is reminiscent of previous studies that have demonstrated that TAZ/YAP's interactions with epigenetic modifiers are key to their function. In *Drosophila*, Yorkie (YAP orthologue) has been shown to be a key interactor of NcoA6, a subunit of the Trithorax-related histone H3 lysine 4 methyltransferase complex. Importantly, conjugation of NcoA6 to Scalloped (TEAD orthologue) alone was sufficient to functionally activate the Yorkie transcriptional program (67, 68). In mammalian/cancer cells, YAP/TAZ driven transcription is facilitated by binding to BRD4, a member of the bromodomain and extraterminal (BET) protein family of epigenetic modifiers that mediates interaction between enhancer sequences and associated promoters (69).

Although the overlap in euchromatin profile between the fusion proteins and CAMTA1/TFE3 can be partially explained by DNA binding elements conferred by the C termini of CAMTA1 and TFE3, it also raises questions about how the TAZ-CAMTA1:ATAC and YAP-TFE3:ATAC complexes form. What remains to be determined is whether the fusion proteins dynamically recruit epigenetic modifiers to contact points in the genome where TEAD and non-TEAD binding sites (e.g. MiTF or EGR2) are in close proximity? Or, are the fusion proteins recruited to stable pools of epigenetic complexes anchored in certain regions of the genome? Additional studies are warranted to distinguish between these two possibilities.

The enhanced euchromatin landscape in cells expressing TAZ-CAMTA1 and YAP-TFE3 is mediated by the ATAC complex

Although we found that TAZ-CAMTA1 and YAP-TFE3 interact with numerous epigenetic complexes, the combined proteomic/RNAi screen demonstrated that the ATAC complex is the most functionally relevant. Ada2a-containing acetyltransferase (ATAC) and Spt-Ada-Gcn5-acetyltransferase (SAGA) are two closely related metazoan histone acetyltransferase complexes that both contain GCN5 or the closely related PCAF histone acetyltransferase catalytic subunits. While the SAGA complex is predominantly localized to promoter sequences, the ATAC complex is found at both promoter and enhancer sequences. This is consistent with our observation that TAZ-CAMTA1 and YAP-TFE3, which are predominantly localized at enhancer sequences, interact primarily with the ATAC complex rather than the SAGA complex. Although the ATAC complex can regulate both housekeeping and tissue specific functions, its binding to enhancer sequences is cell-type specific. Furthermore, knock-down of subunits of the ATAC complex affected only a limited set of active genes rather than globally altering gene expression, another key piece of data demonstrating that the ATAC complex modulates a specific transcriptional program. The implication of these data is that the ATAC complex is recruited to a particular set of target genes and interacts with a specific set of transcription factors rather than ubiquitously interacting with many different transcription factors across the genome. This is corroborated

by our BioID mass spectrometry data which showed that the ATAC complex preferentially binds to both fusion proteins as compared to full-length TAZ and YAP (70).

Importantly, our study also indicates that the ATAC complex represents a potential therapeutic target for EHE, a sarcoma currently lacking an effective medical therapy, regardless of which fusion protein it harbors. In addition, the ATAC complex may also represent a key oncogenic driver in other sarcomas. RNA expression profiling data show that expression of *YEATS2* and *ZZZ3* have prognostic significance in sarcoma clinical samples across various histological types. Future studies elucidating therapeutic strategies targeting the ATAC complex are warranted.

TAZ-CAMTA1 and YAP-TFE3 have convergent oncogenic properties

Despite the observation that CAMTA1 and TFE3 are from entirely different families of proteins, several lines of evidence mentioned above reveal that the function of TAZ-CAMTA1 and YAP-TFE3 are remarkably similar. As mentioned above, the two fusion proteins are similar in terms of the genes they induce, their chromatin-binding profiles, and how they alter the chromatin landscape. This raises the question of whether other TAZ/YAP gene fusions show similar convergence of function. Instead of point mutations, gene fusions are the most common genetic alterations of TAZ(*WWTR1*) and YAP in cancer. Structurally similar *WWTR1*(TAZ)/*YAP1* gene fusions have also been identified in nasopharyngeal carcinoma (*YAP1-MAMLD2*) (Valouev et al., 2014), lung cancer (*YAP1-BIRC2*; *WWTR1-SLC9A9*) (Dhanasekaran et al., 2014), ependymomas (*YAP1-MAMLD1*) (Pajtlar et al., 2015), ossifying fibromyxoid tumor (*KDM2A-WWTR1*) (71), poromas/porocarcinomas (*YAP1-MAML2* and *YAP1-NUTM1*) (72), *NF2*-wildtype meningiomas (*YAP1-MAML2*, *YAP1-PYGO1*, *YAP1-LMO1*) (73), and sclerosing epithelioid fibrosarcoma (*YAP1-KMT2A*) (74, 75). Here, we have outlined an approach integrating NGS modalities with combined proteomic/genetic screens to gain insight into the transcriptional regulatory mechanisms of the TAZ-CAMTA1 and YAP-TFE3 fusion proteins. At its most essential, the mechanism of the two fusion proteins are the same in that a TEAD-based transcriptional program is modulated by interaction

with an epigenetic modifier, the ATAC complex. Additional studies are needed to determine if YAP/TAZ/TEAD fusion proteins also drive tumorigenesis via the ATAC complex or do so by recruiting different chromatin-modifying complexes.

Materials and Methods:

Expression constructs

Double Flag TAZ-CAMTA1 and Double Flag CAMTA1 were cloned into pBabeNeo as previously described. Double Flag TAZ4SA was derived utilizing previously constructed TAZ S89A (4). Additional S66A, S117A, and S311A mutations were introduced by site directed mutagenesis using the QuikChange

II Site-directed mutagenesis kit (Agilent #200521) and the following primers:

TAZ S66A Primer – Forward

5' – CTCGCGCCAGTCCGCCACCGACTCGTCG – 3'

TAZ S66A Primer – Reverse

5' – CGACGAGTCGGTGGCGGACTGGCGCGAG – 3'

TAZ S117A Primer – Forward

5' – CCTCCGCCAGCAGGCCCTACGACGTGACCGAC – 3'

TAZ S117A Primer – Reverse

5' – GTCGGTCACGTCTAGGCCTGCTGGCGGAGG – 3'

TAZ S311A Primer – Forward

5' – CATTCGAGGGAGCAGGCCACTGACAGTGGCCT – 3'

TAZ S311A Primer – Reverse

5' – AGGCCACTGTCAAGTGGCCTGCTCCCTCGAATG – 3'

Double Flag YAP, YAP-TFE3 and TFE3 constructs were cloned into pBabeNeo as previously described (4).

YAP S61A Primer – Forward

5' – TCCGCGGGGACGCGGAGACCGAC – 3'

YAP S61A Primer – Reverse

5' – GTCGGTCTCCGCGTCCCCGCGGA – 3'

YAP S109A Primer – Forward

5' – CACTCCCGACAGGCCGCTACTGATGCAGGCAC – 3'

YAP S109A Primer – Reverse

5' – GTGCCTGCATCAGTAGCGGCCTGTCGGGAGTG – 3'

YAP S127A Primer – Forward

5' – AGCATGTTTCGAGCTCATGCCTCTCCAGCTTCT – 3'

YAP S127A Primer – Reverse

5' – AGAAGCTGGAGAGGCATGAGCTCGAACATGCT – 3'

YAP S164A Primer – Forward

5' – CAGCATCTTCGACAGTCTGCTTTTGAGATACCTGATG – 3'

YAP S164A Primer – Reverse

5' – CATCAGGTATCTCAAAGCAGACTGTCGAAGATGCTG – 3'

YAP S397A Primer – Forward

5' – CTATCACTCTCGAGATGAGGCTACAGACAGTGGACTAAGC – 3'

YAP S397A Primer – Reverse

5' – GCTTAGTCCACTGTCTGTAGCCTCATCTCGAGAGTGATAG – 3'

The pPGS-3HA-TEAD1 construct was obtained courtesy of Dr. Kunliang Guan.

Triple Myc-tagged KAT14 plasmid (Vector ID: VB18 p18-119pmf; Vector name: pRP[Exp]-Puro-CMV>BamHI/3xMyc-KAT14/EcorRI) was synthesized by Vector builder, Shenandoah, TX.

Cell culture

NIH 3T3 mouse fibroblasts and SW872 human liposarcoma cells were obtained from American Type Culture Collection (ATCC, Manassas, VA, USA). NIH 3T3 cells were cultured in DMEM media containing 10% bovine calf serum (Invitrogen-Life Technologies), 1 mM sodium pyruvate, and 50 µg/mL penicillin/streptomycin. SW872 cells were cultured in DMEM media (according to ATCC recommendations) containing 10% fetal bovine serum (Invitrogen-Life Technologies), 1 mM sodium pyruvate, and 50 µg/mL penicillin/streptomycin. All cells were cultured at 37 °C and 5% CO₂.

Transfection and retroviral transduction

Retroviral transfection with pBabeNeo constructs was performed by transfecting PhoenixA retroviral packaging cells with 10 µg of plasmid DNA. Transfection was done using Lipofectamine Reagent and Plus Reagent (Invitrogen-Life Technologies) according to manufacturer's instructions. Supernatant was collected at 48 and 72 hours after transfection, filtered with 0.45 µm filters and supplemented with 8 µg/mL polybrene (EMD Millipore, Burlington, MA, USA). Serial transductions (48 and 72-hour

supernatants) were applied to either SW872 or NIH 3T3 cells for 8 hours each. Pooled stable lines were generated by selecting with G418 for two weeks.

RNA interference-mediated silencing

PLKO.1-puro constructs for knock-down of Tead1 were obtained from Sigma-Aldrich (MISSION shRNAs). Constructs were transfected into LentiX (HEK293T) cells, along with pcMVΔ8.12 and pVSVG packaging plasmids, using Lipofectamine Reagent and Plus Reagent (Invitrogen) according to manufacturer's instructions. Supernatants were collected at 48 hours, filtered using 0.45 μm filters, and polyethylene glycol (PEG) 8000 was added to a final concentration of 12%, and stored at 4 °C overnight. The following day, 48 hour supernatants were centrifuged and viral pellets were resuspended in 0.45 μm filtered 72 hour media and supplemented with 8 μg/mL (EMD Millipore, Burlington, MA, USA) added to either SW872 or NIH 3T3 cells overnight. Pooled stable lines were generated by selection in puromycin.

Antibodies

Anti-FLAG antibody (mouse monoclonal clone M2 (catalog # F3165) utilized for immunofluorescence and western blot (1:1000) was obtained from Sigma-Aldrich (St. Louis, MO, USA). Anti-β-actin antibody (AC-15; catalog #A5441) utilized for western blot (1:10,000) was obtained from Sigma-Aldrich (St. Louis, MO, USA). Alexa 568-conjugated secondary antibody (catalog# A11031) was obtained from Invitrogen-Life Technologies (Grand Island, NY, USA). Anti-alpha tubulin antibody (clone DM1A, catalogue #T9026) was obtained from Sigma-Aldrich (St. Louis, MO, USA). Anti-H3 antibody (clone 1G1, catalogue# sc-517576) was obtained from Santa Cruz Biotechnology (Dallas, TX, USA). Anti-TEAD1 antibody (clone EPR3967(2), catalogue #AB133533) was obtained from Abcam (Cambridge, MA, USA). Anti-HA antibody (sc-7392 clone F-7) was obtained from Santa Cruz Biotechnology (Dallas TX, USA). Anti-ZZZ3 antibody (catalogue# ab118800) was obtained from Abcam (Cambridge, MA, USA). Anti-Myc antibody (catalogue # 2272S) was obtained from Cell Signaling Technologies (Danvers, MA, USA). Horseradish peroxidase-conjugated secondary antibodies used for western blots were

obtained from Santa Cruz Biotechnology (catalog# sc-2055, sc-2054, or sc-2033) and used at 1:5,000 or 1:10,000).

Western blot

Cell pellets were lysed in radioimmunoprecipitation assay (RIPA) buffer, containing cOmplete™ Protease Inhibitor Cocktail (EDTA-free) (Roche) and PhosSTOP™ Phosphatase Inhibitor Cocktail (Roche) according to the manufacturer's instructions. Total protein concentration was measured using Pierce BCA™ Protein Assay Kit (ThermoFisher Scientific, Waltham, MA, USA). Between 50 and 100 µg of total protein was loaded onto a gradient (4-15%) polyacrylamide gel (BioRad, Hercules, CA, USA). Proteins were then transferred to a polyvinylidene difluoride (PVDF) membrane and probed with antibodies described above. Each experiment was repeated at least twice.

Co-immunoprecipitation

1 mg of total cellular protein (whole cell lysates) were cleared by incubating with 2.0 µg of the control IgG and 20 µL vortexed Protein G PLUS-Agarose beads (Santa Cruz, sc-2002) for 30 min with rocking at 4 °C. Supernatant transferred and incubated with 10 µg primary antibody or IgG control and incubated at 4 °C overnight with rocking. 40 µL of vortexed Protein G PLUS agarose beads added and incubated at 4 °C on a rocker platform for 4 hrs. After centrifugation and removal of the supernatant, the cell pellet was re-suspended and washed 3 times with 1 mL cold 1X PBS followed by an additional wash with 1 mL 1X RIPA buffer. The pellet/beads were then resuspended in 2X Laemmli buffer, boiled, then loaded onto polyacrylamide gel.

Nuclear and cytoplasmic fractionation

Nuclear and cytoplasmic fractionation was performed with the Nuclear Extract Kit (Active Motif, 40010). Cells collected from 15 cm cell culture plates were centrifuged and resuspended in 500 µL 1X Hypotonic

buffer and incubated for 5 min on ice. 25 μ L of detergent was added to resuspended cells which were subsequently vortexed for 10s. After centrifugation, the supernatant cytoplasmic fraction was removed and stored. Remaining nuclear pellet was resuspended in 50 μ L complete lysis buffer followed by vortexing. After centrifugation, the supernatant nuclear fraction was removed and stored.

Immunofluorescence staining

Cells were plated in chambered slides (2.5×10^4 for sparse conditions or 5×10^4 for confluent conditions) and allowed to grow until desired timepoint. Cells were fixed with 4% paraformaldehyde (in 1X PBS) for 15 minutes and then washed with 1X PBS. Cells were permeabilized and blocked with 0.3% Triton X-100 and 3% fetal bovine serum for 30 minutes. Primary antibodies, as described above, were diluted in 3% fetal bovine serum and incubated on the cells overnight at 4 °C in a humidity chamber. The following morning, the primary antibody was removed, cells were washed with 1X PBS, and then incubated with Alexa Fluor 568-conjugated secondary antibody (described above) for 45 minutes to 1 hour at room temperature. Immunofluorescence was visualized using a Leica DFC3000G microscope/camera and Leica Application Suite Advanced Fluorescence imaging software or an Olympus BX-61 microscope (Tokyo, Japan) with cellSens imaging software. Each experiment was repeated at least twice.

Soft agar assay

The base layer of 0.5% agarose was plated into 6-well plates (2 mLs/well) and allowed to solidify. For SW872 cells, 5×10^3 cells/2 mL in 0.35% agarose was added to form the top layer. For NIH 3T3 cells, 1×10^4 cells/2 mL in 0.35% agarose was added to form the top layer. Plates were left in the hood and allowed to solidify at room temperature for 3 hours. Colonies were allowed to grow for 2-3 weeks at 37 °C and 5% CO₂ before imaging. Each experiment was repeated at least twice.

Poly-HEMA assay

Poly(2-hydroxyethyl methacrylate), or poly-HEMA, solution was made at 20 mg/mL in 95% ethanol. 96-well tissue culture plates were coated with 130 μ L of poly-HEMA solution and UV sterilized overnight. Cells were plated (2×10^3 cells/100 μ L for SW872 or 3×10^3 cells/100 μ L for NIH 3T3) into multiple wells and incubated at 37 °C and 5% CO₂. Proliferation was assessed every other day by adding 10 μ L of Cell Counting Kit-8 reagent (Dojindo Molecular Technologies, Inc., Rockville, MD, USA). The plates were then incubated for 1 hour at 37 °C and 5% CO₂ and then absorbance at 450 nm was measured using a Synergy H1 Hybrid Multi-Mode Microplate Reader (Biotek, Winooski, VT, USA).

Xenografts

NOD scid gamma (NSG) mice were obtained from The Jackson Laboratory (Bar Harbor, ME, USA) and carry the strain NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ. All animal work was approved by the University of Iowa Institutional Animal Use and Care Committee. SW872 cells containing either empty vector, YAP-TFE3, or TAZ-CAMTA1 and NIH 3T3 cells containing empty vector, YAP-TFE3, or TAZ-CAMTA1 (1×10^7 cells/500 μ L PBS) were injected into the flank of NSG mice. Mice were 6 months of age at the time of NIH 3T3 injection, and 2-3 months of age at the time of SW872 injection. Xenografts were repeated twice, using 5-10 mice per group.

Quantitative reverse transcription polymerase chain reaction

SW872 and NIH 3T3 cells were scraped and collected using TRIzol reagent (Ambion-Life Technologies). Total RNA was isolated using the PureLink™ RNA mini kit (Invitrogen-ThermoFisher Scientific). On-column DNase (Invitrogen) treatment was performed according to manufacturer's instructions. Purified RNA was quantified using a NanoDrop (ThermoFisher) and 1 μ g was converted to cDNA using SuperScript™ III First-Strand Synthesis System (Invitrogen) according to manufacturer's instructions. PCR amplification was performed in technical triplicates on the Applied Biosystems QuantStudio 3 Real-Time PCR System (Applied Biosystems-ThermoFisher). TaqMan Universal PCR Master Mix (Applied Biosystems-ThermoFisher) and PrimeTime standard qPCR primer/probe sets from Integrated DNA

Technologies (Iowa City, IA, USA) were utilized. The qPCR cycling conditions were as follows:

95°C^{10:00}(95°C^{0:15}, 60°C^{1:00})₄₀. Relative quantitation was performed utilizing the delta-delta C_T method and *POLR2A/Polr2a* (RNA Polymerase II) C_T values as the reference control. Each experiment was repeated at least twice. The following primers and probes were used:

Human

CTGF-F 5'-ACCAATGACAACGCCTCC-3'

CTGF-R 5'-TTGGAGATTTTGGGAGTACGG-3'

CTGF-Probe 5'-TGCGAAGCTGACCTGGAAGAGAAC-3'

CYR61-F 5'-GGGATTTCTTGGTCTTGCTG-3'

CYR61-R 5'-CCAATGACAACCCTGAGTGC-3'

CYR61-Probe 5'-TTCTTGCCCTTTTTCAGGCTGCTG-3'

AJUBA-F 5'-GCCTCTACCACACCCAGT-3'

AJUBA-R 5'-ACAGTACACAGAGCCATTGAC -3'

AJUBA-Probe 5'-/56-FAM/TGCAACGCA/ZEN/AAGTTCGCCCA/3IABkFQ/-3'

FBLN5-F 5'-CCTCTTATATGCCGCTTTGGA-3'

FBLN5-R 5'-CCCGCCTTCAGTATTGATGC-3'

FBLN5-Probe 5'-/56-FAM/TCTGTTGCA /ZEN/ CACTCGTCCACATCC/3IABkFQ/-3'

MAP1A-F 5'-CCTAAGTGCTGCAGACTCTT-3'

MAP1A-R 5'-AGGATGATGGTCTCTAGGATCTC-3'

MAP1A-Probe 5'-/56-FAM/TCTGGCCT /ZEN/GGGGTGTAGGA /3IABkFQ/-3'

YEATS2-F 5'-GAACTGACCGTAGATGAAGACA-3'

YEATS2-R 5'-GGAACCACAATGCTGAGAGTAG-3'

YEATS2-Probe 5'-TATCACAACCTGCCACTTCCCCTGC-3'

ZZZ3-F 5'-TCTTGGACTGCTTGAGAACG-3'

ZZZ3-R 5'-TGAATCAGATCATGTGGCACT-3'

ZZZ3-Probe 5'-CAGACGATTGCTGTACTCGAGGCT-3'

POLR2A-F 5'-TCAGCATGTTGGACTCGATG-3'

POLR2A-R 5'-CGTATTCGCATCATGAACAGC-3'

POLR2A-Probe 5'-ACCACCTCTTCCTCCTCTTGCATCT-3'

Mouse

Ctgf-F 5'-TTGACAGGCTTGGCGATT-3'

Ctgf-R 5'-GTTACCAATGACAATACCTTCTGC-3'

Ctgf-Probe 5'-CCGGATGCACTTTTTGCCCTTCT-3'

Cyr61-F 5'-GTGCAGAGGGTTGAAAAGAAC-3'

Cyr61-R 5'-GGAGGTGGAGTTAACGAGAAAC-3'

Cyr61-Probe 5'-TCGGTGCCAAAGACAGGAAGCC-3'

Fras1-F 5'-CACTGTCACCATATCCAACGA-3'

Fras1-R 5'-GATGACCTTGATGCTCAGAACT-3'

Fras1-Probe 5'-/56-FAM/TGAAGAAGC/ZEN/CGCATACCAAGTCCG/3IABkFQ/-3'

Errb3-F 5'-CGTGAAGTGTACCAAGATCCT-3'

Errb3-R 5'-GACTGTCCTGAAAACATTGAGC-3'

Errb3-Probe 5'-/56-FAM/ACAAGATCC/ZEN/CTGCACTGGACCC/3IABkFQ/-3'

Scn1b-F 5'-CATCTCCTGTAAGCGTCGTAG-3'

Scn1b-R 5'-GCACCTCATTCTCATAGCGTAG-3'

Scn1b-Probe 5'-/56-FAM/AAATTCCTC/ZEN/TGTCCCCTTCTGGCG/3IABkFQ/-3'

Polr2a-F 5'-GTCAGCATGTTGGACTCAATG-3'

Polr2a-R 5'-TCGAATCCGCATCATGAACAG-3'

Polr2a-Probe 5'-ACCACCTCTTCCTCCTCTTGCATCT-3'

Luciferase Reporter Assay

HEK293 cells (4×10^5) stably containing empty vector, *YT*, *YAP*, *YT S94A*, *YAP 5SA*, *TC*, *TAZ*, *TC S51A*, or *TAZ 4SA* were transfected with the pGL3b-8xGTIC firefly luciferase reporter plasmid (containing 8 TEAD1-4 binding sites; GTIC also known as MCAT and Hippo response element) (53) and the renilla luciferase reporter plasmid (pRL-TK *Renilla*, (Promega, Madison, WI, USA) in a well from a 6-well plate (biological triplicates). Two days after transfection, the cells were collected and lysed, and extracts were assayed in technical triplicate for firefly and renilla luciferase activity using the Dual Luciferase Reporter Assay (Promega) and a Synergy H1 Hybrid Multi-Mode Microplate Reader (Biotek, Winooski, VT, USA). Each experiment was repeated at least twice.

RNA sequencing

NIH 3T3 and SW872 cells were stably transfected, in triplicate, expressing the following proteins empty vector (EV), *YAP 5SA*, *TFE3*, *YAP-TFE3* (YT), *TAZ 4SA*, *CAMTA1*, and *TAZ-CAMTA1* (TC). All cell lines, except those containing *CAMTA1*, were grown to confluence for 48 hours before collection. Cell lines containing *CAMTA1* were collected at sub-confluence. Total RNA was extracted using Trizol Reagent (Ambion-Life Technologies). Total RNA was isolated using the PureLink™ RNA mini kit (Invitrogen-ThermoFisher Scientific). On-column DNase (Invitrogen) treatment was performed according to manufacturer's instructions. Transcription profiling using RNA-Seq was performed by the University of Iowa Genomics Division using manufacturer recommended protocols. Initially, 500 ng of DNase I-treated total RNA was used to enrich for polyA containing transcripts using oligo(dT) primers bound to beads. The enriched RNA pool was then fragmented, converted to cDNA and ligated to sequencing adaptors containing indexes using the Illumina TruSeq stranded mRNA sample preparation kit (Cat. #RS-122-2101, Illumina, Inc., San Diego, CA). The molar concentrations of the indexed libraries were measured using the 2100 Agilent Bioanalyzer (Agilent Technologies, Santa Clara, CA) and combined equally into pools for sequencing. The concentration of the pools was measured using the Illumina Library Quantification Kit (KAPA Biosystems, Wilmington, MA) and sequenced on the Illumina HiSeq 4000 genome sequencer using 75 bp paired-end SBS chemistry.

Barcoded samples were pooled and sequenced using an Illumina HiSeq 4000 in the Iowa Institute of Human Genetics (IIHG) Genomics Core Facility. Average reads per sample for SW872 samples was 93 M (+/- 16 M). Reads were demultiplexed and converted to FASTQ format, then processed with ‘bcbio-nextgen’ (<https://github.com/chapmanb/bcbio-nextgen>; version 1.0.8). SW872 reads were aligned against the ‘hg38’ reference genome (genome FASTA and annotations derived from Ensembl release 78 (ftp://ftp.ensembl.org/pub/release78/fasta/homo_sapiens/dna/Homo_sapiens.GRCh38.dna.toplevel.fa.gz)). Reads from NIH 3T3 cells were aligned against ‘mm10’ (genome FASTA and annotations derived from ftp://ftp.ensembl.org/pub/release-97/fasta/mus_musculus/) reference genome using ‘Hisat2’ (76). Reads were also pseudo-aligned to the transcriptome using Salmon (77). Transcript-level abundances were converted to gene-level counts with ‘tximport’ (78). QC was performed with ‘qualimap’ and ‘samtools’ (79). Differential gene expression analysis was performed with DESeq2 (80). Gene lists were analyzed using Advaita’s ‘iPathwayGuide’ (<https://www.advaitabio.com/ipathwayguide>; Advaita Bioinformatics, Ann Arbor, MI).

ChIP sequencing

SW872 cells stably expressing Flag-tagged EV, YAP 5SA, TFE3, YT, TAZ 4SA, CAMTA1, and TC were grown to 90% confluence in 15 cm tissue culture plates (in biological triplicate). Cells were cross-linked and prepared using the SimpleChIP® Enzymatic Chromatin IP Kit (Cell Signaling, #9003, Danvers, MA). Protocol was performed according to manufacturer’s instructions. Lysates were sonicated for 3 sets of 20-sec pulses in ice bath, using a VirTis Virsonic 60 Sonicator at setting 6, with a 30 second ice bath incubation between sonication pulses. Chromatin digestion and concentration was verified to be within the desired range using the 2100 Agilent Bioanalyzer (Agilent Technologies, Santa Clara, CA). The SimpleChIP® ChIP-seq DNA Library Prep Kit for Illumina (Cell Signaling, #56795) was used to prepare the DNA library according of manufacturer’s instructions. Immunoprecipitation was performed with the anti-DYKDDDDK Tag Rabbit monoclonal antibody (Cell Signaling #14793) or provided isotype

rabbit IgG control (Cell Signaling). The molar concentrations of the indexed libraries were measured using the 2100 Agilent Bioanalyzer (Agilent Technologies, Santa Clara, CA) and combined equally into pools for sequencing using an Illumina HiSeq 4000 in the Iowa Institute of Human Genetics (IIHG) Genomics Core Facility.

ChIP-seq analysis was carried out using ‘bcbio-nextgen’ in “analysis: chip-seq” mode. Reads were pooled across replicates to obtain improved read depth, then trimmed for low quality and adapter read-through with ‘Atropos’ (version 1.1.16) (81) with adapter sequences (‘AGATCGGAAGAG’, ‘CTCTTCCGATCT’) and options set as ‘--quality-cutoff=5 --minimum-length=25’. Trimmed reads were aligned to the ‘GRCh37/hg19’ reference assembly with ‘BWA-MEM’ (82) using default parameters. Multi-mapping reads were removed before peak calling, and ‘greylist’ regions were detected and removed (<https://github.com/roryk/chipseq-greylist>). Peaks were called using ‘macs2’ (83) with command-line options ‘-f BAMPE’, ‘-g 2.7e9’, ‘-B’, and ‘-q 0.10.’ (FDR=0.1) Narrow peak files were imported into the R and annotated with ‘ChIPseeker’ (84) and ‘clusterProfiler’ packages (85). Peak ranges for motif analysis were written out as FASTA using ‘Biostrings’ (<https://bioconductor.org/packages/release/bioc/html/Biostrings.html>) and ‘BSgenome’ packages (<https://bioconductor.org/packages/release/bioc/html/BSgenome.html>). Motif enrichment analysis was carried out with MEME suite (<http://meme-suite.org>) using the ‘AME’ and ‘FIMO’ modules and HOCOMOCO (v11) database.

ChIP-qPCR

PCR amplification was performed in technical triplicates on the Applied Biosystems QuantStudio 3 Real-Time PCR System (Applied Biosystems-ThermoFisher). The PCR reaction contained SimpleChip Universal master mix (catalogue # 88989) Cell Signaling Technologies (Danvers MA, USA), 5uM of primers and 2uL of ChIP DNA or 2% input DNA. Cycling conditions as follows, 3 minutes at 95C for

initial denaturation followed by 40 cycles of 15 seconds at 95 C then 1 minute at 60C. Percent input was calculated.

The following primers were used:

CTGF-F 5'-GCCAATGAGCTGAATGGAGT-3'

CTGF-R 5'-CAATCCGGTGTGAGTTGATG-3'

MAFK-F 5'-GCTCTGTAAACATCTGGTGACT-3'

MAFK-R 5'-GGGAGCACCATGAGAACTT -3'

NEAT1-F 5'-CTCAACAACATCCGGGAAGA -3'

NEAT1-R 5'-CGCTTCTCTTTAACAGTGCTTT -3'

MAP1A-F 5'-TGCAGAAGCAGCCAAGAA-3'

MAP1A-R 5'-TTTGCCGCAGTGTGTGA3'

ATAC sequencing

50,000 cells were trypsinized and pelleted. Cells were washed once with 50 µL of cold 1x PBS buffer.

After pelleting, cell pellet was resuspended in 50 µL cold lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl₂, 0.1% IGEPAL CA-630). The suspension was centrifuged again and supernatant discarded leaving a nuclear pellet. A transposition reaction was prepared containing 25 µL Tagment DNA buffer (#15027866), 2.5 µL TDE1 Nextera Tn5 Transposase (#15027865), and 22.5 µL Nuclease Free H₂O. The nuclear pellet was resuspended in the transposition reaction mixture. The transposition reaction was incubated at 37°C for 30 min. After transposition, fragmented DNA purified using a Qiagen miniElute PCR purification kit (catalogue #28004) per the manufacturer's instructions. Transposed DNA fragments were PCR amplified using the Nextera XT Index Kit (#15055293) per the manufacturer's instruction to generate indexed libraries.

Analysis of ATAC-seq data was carried out as follows. Reads were trimmed using 'NGmerge' in 'adapter removal' mode (86). Trimmed reads were then aligned as described above [command line

settings: “-c 250 -t 1 -v 3’] and sorted with ‘samtools’ (79). Peaks were called with ‘genrich’ (<https://github.com/jsh58/Genrich>) in ‘atacseq’ mode [with flags set as ‘-E blacklists/wgEncodeHg19ConsensusSignalArtifactRegions.bed -v -j -r -e MT, GL000191.1, GL000192.1, GL000193.1, GL000195.1, GL000199.1, GL000205.1, GL000206.1, GL000208.1, GL000212.1, GL000214.1, GL000216.1, GL000217.1, GL000219.1, GL000220.1, GL000222.1, GL000223.1, GL000224.1, GL000225.1, GL000226.1, GL000228.1, GL000235.1, GL000243.1’].] Genrich removes PCR duplicates, excludes named regions, and blacklist regions, and properly accounts for multi-mapping reads. Peaks were called individually on each replicate. Peak files were imported along with BAM files into DiffBind (<https://bioconductor.riken.jp/packages/3.0/bioc/html/DiffBind.html>) (87) and a consensus peakset was determined. A binding affinity matrix containing (normalized) read counts for each sample at each consensus site was computed. ‘DESeq2’ (80) was used to test contrasts for differential chromatin accessibility. Finally, PCA analysis was performed to show how samples cluster at differentially accessible chromatin.

BioID Mass spectrometry and SAINT analysis

Full-length YAP, TFE3, TAZ, and CAMTA1, along with YAP-TFE3 and TAZ-CAMTA1 were cloned into pcDNA5 FRT/TO FLAG-BirAR118G vectors, containing an N-terminal *E.coli* biotin ligase with an R118G mutation, as described in Roux et al. (88) and Lambert et al. (89). These bait proteins of interest were then stably expressed in Flp-In™ T-REx™ HEK293 cells as previously described (90). Expression was then induced using 1 µg/mL tetracycline for 24 hours. The cells were also treated with 50 µM biotin at the time of induction. After 24 hours, cells harvested and processed (89). Two biological replicates were made for each cell line. Affinity purification and proximity biotinylation coupled to mass spectrometry were performed as described in Lambert et al. (89). SAINT (significance analysis of interactome) analysis (91) was performed on the mass spectrometry data, using 10 controls compressed to 5. Only proteins with iProphet protein probability ≥ 0.95 were used. Results are expressed in dotplot

format. Columns show bait proteins, while the rows list the names of the identified prey proteins. Each prey protein is represented as a dot, with color signifying average spectral count, the darkness indicating average spectral count between the two biological replicates (the darker the dot, the higher the average spectral count), and the size represents the relative abundance. The ring around the dot indicates the false discovery rate (the darker the ring, the higher the confidence). The data was filtered so that each prey had a minimum of 10 spectral counts in at least one of the biological replicates. Results were also filtered to exclude components of trypsin, biotin, and streptavidin.

Statistics

For soft agar assays, statistical significance was evaluated using an unpaired two-tailed *t*-test. For poly-HEMA proliferation assays, statistical significance was evaluated using fold change increase in proliferation at day 10 with an unpaired two-tailed *t*-test. For quantitative RT-PCR, standard deviation was calculated from fold change values for each triplicate. For Kaplan-Meier curves significances was determined by Log-rank (Mantel-Cox) test. Graphs were made using the average of the technical replicates. Each experiment was repeated at least twice. Hypergeometric testing was performed using the `phyper()` function in the stats R package (v3.6.3) set to assess enrichment and the lower tail set to false. Hypergeometric density was calculated using the related `dhyper` function and converted using the negative log10 of the output. For gene expression data, the population was set to the total number of recovered genes with the mean of 5 counts across all samples. Both the ChIP-seq and ATAC-seq, the population was set as the total number of genes annotated across all conditions. Error bars were used to define one standard deviation. For all panels, **** $p < 0.0001$, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Supplementary Materials

Fig. S1. YAP-TFE3, TAZ-CAMTA1, and controls are expressed in NIH 3T3 cells, SW 872 cells, and HEK293 cells.

Fig. S2. YAP-TFE3 is not regulated by the Hippo pathway and required binding to TEAD

transcription factors to drive its oncogenic transcriptional program.

Fig. S3. YAP-TFE3 and TAZ-CAMTA1 driven xenograft in NOD.Cg-PrkdcscidIl2rgtm1Wjl/SzJ (NSG) mice form tumors in vivo.

Fig. S4. Differential expression of YAP-TFE3 and TAZ-CAMTA1 in SW872 cells.

Fig. S5. Pathway analysis of TAZ-CAMTA1 differential expression and DNA binding shows enrichment of Hippo pathway genes as well as other Hippo related pathways.

Fig. S6. TAZ-CAMTA1 and YAP-TFE3 alter the chromatin landscape.

Fig. S7. BioID Mass Spectrometry shows that YAP-TFE3 and TAZ-CAMTA1 have an altered interactome compared to YAP and TAZ

Fig. S8. RNAi screen identifies epigenetic modifiers important for the activity of TAZ-CAMTA1 and YAP-TFE3.

Table S1. Tead1 shRNA sequences

Table S2. NIH 3T3 differentially expressed genes

Table S3. SW872 differentially expressed genes

Table S4. ChIP AME Motif Enrichment Analysis

Table S5. Differentially expressed genes also bound by ChIP-Seq

Table S6. Sequences from shRNA screen

References:

1. F. A. Goldblum JR, and Weiss SW in *Enzinger and Weiss's Soft Tissue Tumors, 6th Ed.* (Elsevier Saunders, Philadelphia, 2014), chap. 22, pp. 681-688.
2. C. Errani *et al.*, A novel WWTR1-CAMTA1 gene fusion is a consistent abnormality in epithelioid hemangioendothelioma of different anatomic sites. *Genes, chromosomes & cancer* **50**, 644-653 (2011).
3. M. R. Tanas *et al.*, Identification of a disease-defining gene fusion in epithelioid hemangioendothelioma. *Sci Transl Med* **3**, 98ra82 (2011).
4. M. R. Tanas *et al.*, Mechanism of action of a WWTR1(TAZ)-CAMTA1 fusion oncoprotein. *Oncogene* **35**, 929-938 (2016).
5. C. R. Antonescu *et al.*, Novel YAP1-TFE3 fusion defines a distinct subset of epithelioid hemangioendothelioma. *Genes, chromosomes & cancer* **52**, 775-784 (2013).

6. K. F. Harvey, C. M. Pflieger, I. K. Hariharan, The *Drosophila* Mst ortholog, hippo, restricts growth and cell proliferation and promotes apoptosis. *Cell* **114**, 457-467 (2003).
7. S. Wu, J. Huang, J. Dong, D. Pan, hippo encodes a Ste-20 family protein kinase that restricts cell proliferation and promotes apoptosis in conjunction with salvador and warts. *Cell* **114**, 445-456 (2003).
8. R. S. Udán, M. Kango-Singh, R. Nolo, C. Tao, G. Halder, Hippo promotes proliferation arrest and apoptosis in the Salvador/Warts pathway. *Nat Cell Biol* **5**, 914-920 (2003).
9. S. Pantalacci, N. Tapon, P. Leopold, The Salvador partner Hippo promotes apoptosis and cell-cycle exit in *Drosophila*. *Nat Cell Biol* **5**, 921-927 (2003).
10. J. Dong *et al.*, Elucidation of a universal size-control mechanism in *Drosophila* and mammals. *Cell* **130**, 1120-1133 (2007).
11. B. Zhao *et al.*, Inactivation of YAP oncoprotein by the Hippo pathway is involved in cell contact inhibition and tissue growth control. *Genes & development* **21**, 2747-2761 (2007).
12. R. W. Justice, O. Zilian, D. F. Woods, M. Noll, P. J. Bryant, The *Drosophila* tumor suppressor gene warts encodes a homolog of human myotonic dystrophy kinase and is required for the control of cell shape and proliferation. *Genes & development* **9**, 534-546 (1995).
13. T. Xu, W. Wang, S. Zhang, R. A. Stewart, W. Yu, Identifying tumor suppressors in genetic mosaics: the *Drosophila* lats gene encodes a putative protein kinase. *Development (Cambridge, England)* **121**, 1053-1063 (1995).
14. N. Tapon *et al.*, salvador Promotes both cell cycle exit and apoptosis in *Drosophila* and is mutated in human cancer cell lines. *Cell* **110**, 467-478 (2002).
15. F. D. Camargo *et al.*, YAP1 increases organ size and expands undifferentiated progenitor cells. *Current biology : CB* **17**, 2054-2060 (2007).
16. B. Zhao *et al.*, TEAD mediates YAP-dependent gene induction and growth control. *Genes & development* **22**, 1962-1971 (2008).
17. H. Zhang *et al.*, TEAD transcription factors mediate the function of TAZ in cell growth and epithelial-mesenchymal transition. *The Journal of biological chemistry* **284**, 13355-13362 (2009).
18. K. F. Harvey, X. Zhang, D. M. Thomas, The Hippo pathway and human cancer. *Nature reviews. Cancer* **13**, 246-257 (2013).
19. S. W. Chan *et al.*, A role for TAZ in migration, invasion, and tumorigenesis of breast cancer cells. *Cancer research* **68**, 2592-2598 (2008).
20. D. Zhao, X. Zhi, Z. Zhou, C. Chen, TAZ antagonizes the WWP1-mediated KLF5 degradation and promotes breast cell proliferation and tumorigenesis. *Carcinogenesis* **33**, 59-67 (2012).
21. L. Wang *et al.*, Overexpression of YAP and TAZ is an independent predictor of prognosis in colorectal cancer and related to the proliferation and metastasis of colon cancer cells. *PloS one* **8**, e65539 (2013).
22. L. Zender *et al.*, Identification and validation of oncogenes in liver cancer using an integrative oncogenomic approach. *Cell* **125**, 1253-1267 (2006).
23. M. Xie *et al.*, Prognostic significance of TAZ expression in resected non-small cell lung cancer. *J Thorac Oncol* **7**, 799-807 (2012).
24. W. Zhang *et al.*, Downstream of mutant KRAS, the transcription regulator YAP is essential for neoplastic progression to pancreatic ductal adenocarcinoma. *Science signaling* **7**, ra42 (2014).
25. T. de Cristofaro *et al.*, TAZ/WWTR1 is overexpressed in papillary thyroid carcinoma. *Eur J Cancer* **47**, 926-933 (2011).
26. C. A. Fullenkamp *et al.*, TAZ and YAP are frequently activated oncoproteins in sarcomas. *Oncotarget* **7**, 30094-30108 (2016).
27. N. M. Merritt *et al.*, A comprehensive evaluation of Hippo pathway silencing in sarcomas. *Oncotarget* **9**, 31620-31636 (2018).
28. I. Isfort, S. Elges, M. Cyra, Prevalence of the Hippo Effectors YAP1/TAZ in Tumors of Soft Tissue and Bone. **9**, 19704 (2019).

29. T. S. Eisinger-Mathason *et al.*, Deregulation of the Hippo pathway in soft-tissue sarcoma promotes FOXM1 expression and tumorigenesis. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E3402-3411 (2015).
30. L. E. Crose *et al.*, Alveolar rhabdomyosarcoma-associated PAX3-FOXO1 promotes tumorigenesis via Hippo pathway suppression. *The Journal of clinical investigation* **124**, 285-296 (2014).
31. Z. Nagy *et al.*, The metazoan ATAC and SAGA coactivator HAT complexes regulate different sets of inducible target genes. *Cellular and molecular life sciences : CMLS* **67**, 611-628 (2010).
32. W. Mi *et al.*, YEATS2 links histone acetylation to tumorigenesis of non-small cell lung cancer. *Nature communications* **8**, 1088 (2017).
33. H. C. Beird *et al.*, Genomic profiling of dedifferentiated liposarcoma compared to matched well-differentiated liposarcoma reveals higher genomic complexity and a common origin. *Cold Spring Harbor molecular case studies* **4**, (2018).
34. F. Kanai *et al.*, TAZ: a novel transcriptional co-activator regulated by interactions with 14-3-3 and PDZ domain proteins. *The EMBO journal* **19**, 6778-6791 (2000).
35. A. Finkler, R. Ashery-Padan, H. Fromm, CAMTAs: calmodulin-binding transcription activators from plants to human. *FEBS Lett* **581**, 3893-3898 (2007).
36. C. J. Gaffney *et al.*, Identification, basic characterization and evolutionary analysis of differentially spliced mRNA isoforms of human YAP1 gene. *Gene* **509**, 215-222 (2012).
37. E. C. Kauffman *et al.*, Molecular genetics and cellular features of TFE3 and TFEB fusion kidney cancers. *Nature reviews. Urology* **11**, 465-475 (2014).
38. E. Steingrimsson *et al.*, Mitf and Tfe3, two members of the Mitf-Tfe family of bHLH-Zip transcription factors, have important but functionally redundant roles in osteoclast development. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 4477-4482 (2002).
39. H. Beckmann, L. K. Su, T. Kadesch, TFE3: a helix-loop-helix protein that activates transcription through the immunoglobulin enhancer muE3 motif. *Genes & development* **4**, 167-179 (1990).
40. P. S. Henthorn, C. C. Stewart, T. Kadesch, J. M. Puck, The gene encoding human TFE3, a transcription factor that binds the immunoglobulin heavy-chain enhancer, maps to Xp11.22. *Genomics* **11**, 374-378 (1991).
41. J. A. Martina *et al.*, The nutrient-responsive transcription factor TFE3 promotes autophagy, lysosomal biogenesis, and clearance of cellular debris. *Science signaling* **7**, ra9 (2014).
42. L. Slade, T. Pulinilkunnil, The MiTF/TFE Family of Transcription Factors: Master Regulators of Organelle Signaling, Metabolism, and Stress Adaptation. *Molecular cancer research : MCR* **15**, 1637-1643 (2017).
43. P. Argani *et al.*, A novel CLTC-TFE3 gene fusion in pediatric renal adenocarcinoma with t(X;17)(p11.2;q23). *Oncogene* **22**, 5374-5378 (2003).
44. T. Klatte *et al.*, Renal cell carcinoma associated with transcription factor E3 expression and Xp11.2 translocation: incidence, characteristics, and prognosis. *American journal of clinical pathology* **137**, 761-768 (2012).
45. S. A. Manning, B. Kroege, K. F. Harvey, The regulation of Yorkie, YAP and TAZ: new insights into the Hippo pathway. **147**, (2020).
46. C. Y. Liu *et al.*, The hippo tumor pathway promotes TAZ degradation by phosphorylating a phosphodegron and recruiting the SCF{beta}-TrCP E3 ligase. *The Journal of biological chemistry* **285**, 37159-37169 (2010).
47. W. Huang *et al.*, The N-terminal phosphodegron targets TAZ/WWTR1 protein for SCFbeta-TrCP-dependent degradation in response to phosphatidylinositol 3-kinase inhibition. *The Journal of biological chemistry* **287**, 26245-26253 (2012).
48. F. Zanconato, M. Cordenonsi, S. Piccolo, YAP/TAZ at the Roots of Cancer. *Cancer cell* **29**, 783-803 (2016).

49. R. G. Greig *et al.*, Tumorigenic and metastatic properties of "normal" and ras-transfected NIH/3T3 cells. *Proceedings of the National Academy of Sciences of the United States of America* **82**, 3698-3701 (1985).
50. M. Murakami, M. Nakagawa, E. N. Olson, O. Nakagawa, A WW domain protein TAZ is a critical coactivator for TBX5, a transcription factor implicated in Holt-Oram syndrome. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 18034-18039 (2005).
51. J. H. Hong *et al.*, TAZ, a transcriptional modulator of mesenchymal stem cell differentiation. *Science (New York, N.Y.)* **309**, 1074-1078 (2005).
52. X. Varelas *et al.*, TAZ controls Smad nucleocytoplasmic shuttling and regulates human embryonic stem-cell self-renewal. *Nat Cell Biol* **10**, 837-848 (2008).
53. S. Dupont *et al.*, Role of YAP/TAZ in mechanotransduction. *Nature* **474**, 179-183 (2011).
54. Q. Y. Lei *et al.*, TAZ promotes cell proliferation and epithelial-mesenchymal transition and is inhibited by the hippo pathway. *Molecular and cellular biology* **28**, 2426-2436 (2008).
55. F. Zanconato *et al.*, Genome-wide association between YAP/TAZ/TEAD and AP-1 at enhancers drives oncogenic growth. *Nat Cell Biol* **17**, 1218-1227 (2015).
56. G. G. Galli *et al.*, YAP Drives Growth by Controlling Transcriptional Pause Release from Dynamic Enhancers. *Molecular cell* **60**, 328-337 (2015).
57. C. Stein *et al.*, YAP1 Exerts Its Transcriptional Control via TEAD-Mediated Activation of Enhancers. *PLoS genetics* **11**, e1005465 (2015).
58. L. E. Warner *et al.*, Mutations in the early growth response 2 (EGR2) gene are associated with hereditary myelinopathies. *Nature genetics* **18**, 382-384 (1998).
59. C. Long *et al.*, Ataxia and Purkinje cell degeneration in mice lacking the CAMTA1 transcription factor. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 11521-11526 (2014).
60. D. N. Edwards, V. M. Ngwa, The receptor tyrosine kinase EphA2 promotes glutamine metabolism in tumors by activating the transcriptional coactivators YAP and TAZ. **10**, (2017).
61. I. M. Moya, G. Halder, Discovering the Hippo pathway protein-protein interactome. *Cell research* **24**, 137-138 (2014).
62. W. Mi, H. Guan, J. Lyu, YEATS2 links histone acetylation to tumorigenesis of non-small cell lung cancer. **8**, 1088 (2017).
63. W. Mi, Y. Zhang, The ZZ-type zinc finger of ZZZ3 modulates the ATAC complex-mediated histone acetylation and gene activation. **9**, 3759 (2018).
64. P. A. Grant *et al.*, Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes & development* **11**, 1640-1650 (1997).
65. Z. Nagy, L. Tora, Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene* **26**, 5341-5357 (2007).
66. J. Vandel, O. Cassan, S. Lèbre, C. H. Lecellier, L. Bréhélin, Probing transcription factor combinatorics in different promoter classes and in enhancers. **20**, 103 (2019).
67. H. Oh *et al.*, Yorkie promotes transcription by recruiting a histone methyltransferase complex. *Cell reports* **8**, 449-459 (2014).
68. Y. Qing *et al.*, The Hippo effector Yorkie activates transcription by interacting with a histone methyltransferase complex through Ncoa6. *eLife* **3**, (2014).
69. F. Zanconato, G. Battilana, M. Forcato, Transcriptional addiction in cancer cells is mediated by YAP/TAZ through BRD4. **24**, 1599-1610 (2018).
70. A. R. Krebs, K. Karmodiya, M. Lindahl-Allen, K. Struhl, L. Tora, SAGA and ATAC histone acetyl transferase complexes regulate distinct sets of genes and ATAC defines a class of p300-independent enhancers. *Molecular cell* **44**, 410-423 (2011).

71. Y. C. Kao *et al.*, Expanding the molecular signature of ossifying fibromyxoid tumors with two novel gene fusions: CREBBP-BCORL1 and KDM2A-WWTR1. *Genes, chromosomes & cancer* **56**, 42-50 (2017).
72. S. Sekine *et al.*, Recurrent YAP1-MAML2 and YAP1-NUTM1 fusions in poroma and porocarcinoma. *The Journal of clinical investigation* **130**, 3827-3832 (2019).
73. P. Sievers *et al.*, YAP1-fusions in pediatric NF2-wildtype meningioma. **139**, 215-218 (2020).
74. F. Puls *et al.*, Recurrent Fusions Between YAP1 and KMT2A in Morphologically Distinct Neoplasms Within the Spectrum of Low-grade Fibromyxoid Sarcoma and Sclerosing Epithelioid Fibrosarcoma. *The American journal of surgical pathology*, (2020).
75. Y. C. Kao *et al.*, Recurrent YAP1 and KMT2A Gene Rearrangements in a Subset of MUC4-negative Sclerosing Epithelioid Fibrosarcoma. *The American journal of surgical pathology* **44**, 368-377 (2020).
76. D. Kim, B. Langmead, S. L. Salzberg, HISAT: a fast spliced aligner with low memory requirements. **12**, 357-360 (2015).
77. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression. *Nature methods* **14**, 417-419 (2017).
78. C. Sonesson, M. I. Love, M. D. Robinson, Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research* **4**, 1521 (2015).
79. H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)* **25**, 2078-2079 (2009).
80. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**, 550 (2014).
81. J. P. Didion, M. Martin, Atropos: specific, sensitive, and speedy trimming of sequencing reads. **5**, e3720 (2017).
82. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)* **25**, 1754-1760 (2009).
83. Y. Zhang *et al.*, Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9**, R137 (2008).
84. G. Yu, L. G. Wang, Q. Y. He, ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics (Oxford, England)* **31**, 2382-2383 (2015).
85. G. Yu, L. G. Wang, Y. Han, Q. Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics : a journal of integrative biology* **16**, 284-287 (2012).
86. J. M. Gaspar, NGMerge: merging paired-end reads via novel empirically-derived models of sequencing errors. *BMC bioinformatics* **19**, 536 (2018).
87. C. S. Ross-Innes *et al.*, Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389-393 (2012).
88. K. J. Roux, D. I. Kim, M. Raida, B. Burke, A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *The Journal of cell biology* **196**, 801-810 (2012).
89. J. P. Lambert, M. Tucholska, C. Go, J. D. Knight, A. C. Gingras, Proximity biotinylation and affinity purification are complementary approaches for the interactome mapping of chromatin-associated protein complexes. *Journal of proteomics* **118**, 81-94 (2015).
90. A. L. Couzens *et al.*, Protein interaction network of the mammalian Hippo pathway reveals mechanisms of kinase-phosphatase interactions. *Science signaling* **6**, rs15 (2013).
91. H. Choi *et al.*, SAINT: probabilistic scoring of affinity purification-mass spectrometry data. *Nature methods* **8**, 70-73 (2011).

Acknowledgements: We thank Nicholas Scalora for critical reading of the manuscript. RNA-seq, ChIP-Seq, and ATAC-Seq data presented herein were obtained at the Genomics Division of the Iowa Institute

of Human Genetics which is supported, in part, by the University of Iowa Carver College of Medicine. This work was supported by a University of Iowa Sarcoma Multidisciplinary Oncology Group pilot award (M.R.T), by a grant from the Veterans Health Administration Merit Review Program 1 I01 BX003644-01 (M.R.T), by a grant from the National Institutes of Health, National Cancer Institute 1 R01 CA237031-01A1 (M.R.T), and by an NCI Core Grant P30 CA086862 (University of Iowa Holden Comprehensive Cancer Center). K.F.H is a National Health and Medical Research Council Senior Research Fellow (APP1078220). A.C.G was supported by the Canadian Institutes of Health Research (FDN 144301) and the Terry Fox Research Institute and is the Tier1 Canada Research Chair in Functional Proteomics. Proteomics analysis was performed at the Network Biology Collaborative Centre at the Lunenfeld-Tanenbaum Research Institute, a facility supported by Canada Foundation for Innovation funding, by the Ontarian Government and by Genome Canada and Ontario Genomics (OGI-139).

Author contributions: Conceptualization, K.F.H. and M.R.T.; Methodology, M.C., A.C.G., and M.R.T.; Software, M.C and N.B.; Formal analysis, M.C., N.B., D.R., Z.Y.L., A.C.G. and M.R.T.; Investigation, N.M., K.G., D.R., Z.Y.L., X.Z., K.M., C.F., M.C., and M.R.T.; Resources, A.C.G., K.F.H., and M.R.T.; Writing-Original Draft, N.M. and M.R.T.; Writing-Review & Editing, A.C.G., K.F.H., and M.R.T.; Supervision, A.C.G., K.F.H., and M.R.T.; Funding Acquisition, A.C.G., K.F.H., and M.R.T.

Data and Code Availability:

The accession number for the RNA-Seq data reported in this paper for NIH 3T3 cells is GEO: GSE152736.

The accession number for the RNA-Seq data reported in this paper for SW872 cells is GEO: GSE152737

The accession number for the ChIP-Seq data reported in this paper is GEO: GSE152778

The accession number for the ATAC-Seq data reported in this paper is GEO: GSE152733

Figures

Figure 1

Fig. 1. YAP-TFE3 transforms cells *in vitro* and *in vivo* and combines properties of both YAP and TFE3. (A) Structure of full-length TAZ, CAMTA1, TAZ-CAMTA1 (TC), YAP, TFE3, YAP-TFE3 (YT), truncated portion of YAP (YAP trunc) and the truncated portion of TFE3 (TFE3 trunc). (B) Soft agar assay in NIH 3T3 cells transduced with empty vector (EV), full-length YAP, YAP trunc, TFE3, TFE3 trunc, and YAP-TFE3. (C) Poly-HEMA proliferation assay in NIH 3T3 cells (same constructs as part B). (D) Poly-HEMA proliferation assay in SW872 cells (same constructs as part B). (E) Immunofluorescence images in NIH 3T3 cells expressing YAP, YAP-TFE3, and TFE3 during sparse and confluent conditions, represented graphically (% nuclear positivity during cell confluence is normalized to % positivity in sparse conditions) in (F). (G) Overall survival curve for NSG mice containing xenografted NIH 3T3 cells transduced with empty vector (EV), YAP-TFE3 (YT), and TAZ-CAMTA1

(TC). **(H)** Tumor growth curve for NSG mice containing xenografted NIH 3T3 cells transduced with empty vector (EV), YAP-TFE3 (YT), and TAZ-CAMTA1 (TC). **(I)** Gross and microscopic pathology of pulmonary metastases derived from NIH 3T3 cells derived tumors expressing YAP-TFE3. Graphically represented on the right. **(J)** Tumor growth curve for NSG mice containing xenografted SW872 cells transduced with empty vector (EV), YAP-TFE3 (YT), and TAZ-CAMTA1 (TC).

Figure 2

Fig. 2. Tead1 mediates the oncogenic transcriptional program of YAP-TFE3. (A) Knockdown of Tead1 in NIH 3T3 YAP-TFE3 cells. (B) Soft agar assay in NIH 3T3 YAP-TFE3 cells with two Tead1 shRNA (shTead#1 and shTead#2) with colonies shown below. (C) Poly-HEMA analysis in NIH 3T3 YAP-TFE3 cells with Tead1 knock-down. Spheroid formation on poly-HEMA is shown below. (D) Co-immunoprecipitation experiment in NIH 3T3 cells stably expressing Flag-YAP-TFE3 or Flag-YAP-TFE3 S94A. 3HA-TEAD1 is transiently transfected. * indicates MW of Flag-YAP-TFE3 (E) Soft agar assay in NIH 3T3 cells expressing YAP-TFE3 (YT), YT S94A, YAP, or YAPS94A. (F) Poly-HEMA assay in NIH 3T3 cells expressing same constructs present in E. Spheroid formation on poly-HEMA is shown below. (G) Quantitative RT-PCR showing reduced *Ctgf* expression in NIH 3T3 YAP-TFE3 (YT) S94A cells. (H) Luciferase reporter assay (8XGTIIIC luciferase reporter construct) in NIH 3T3 cells expressing YT or YT S94A. (I) Model diagram of YAP-TFE3 driving transcription by binding to TEAD1.

Figure 3

Fig. 3. TAZ-CAMTA1 and YAP-TFE3 induce different but overlapping transcriptomes compared to TAZ and YAP. (A) Number of differentially expressed genes (FDR=5%) in NIH 3T3 cells expressing TAZ-CAMTA1 and full-length controls. HG density represents the $-\log_{10}(\text{hypergeometric density})$. p value for hypergeometric analysis included. (B) Number of differentially expressed genes (FDR=5%) in NIH 3T3 cells expressing YAP-TFE3 and full-length controls. (C) Principal component analysis of RNA expression after variance-stabilizing transformation in NIH 3T3 cells. (D) Validation of RNA-Seq in NIH 3T3 cells by qRT-PCR for key genes. (E) Number of differentially expressed genes (FDR=5%) in SW872 cells expressing TAZ-CAMTA1 and full-length controls and (F) YAP-TFE3 and full-length controls. (G) Scatter plot of two types of pathway enrichment evidence: probability of over-representation (pORA) and probability of accumulation (pAcc) as calculated by iPathwayGuide for SW872 TAZ-CAMTA1 cells and (H) YAP-TFE3 cells. (I) Validation of RNA-Seq differential expression in SW872 cells by qRT-PCR for key genes.

Figure 4

Fig. 4. TAZ-CAMTA1 and YAP-TFE3 occupy TEAD and non-TEAD transcription factor motifs.

(A) Heat map of ChIP binding to transcriptional start site (TSS) regions and average TSS profile histogram for TAZ-CAMTA1 and controls. (B) Heat map and histogram for YAP-TFE3 and controls. (C) Distribution of transcription factor binding loci relative to TSS for the fusion proteins and controls. (D) Distribution of annotated functional DNA binding sites for the fusion proteins and controls. (E) Intersection of gene annotations for TAZ-CAMTA1 ChIP peaks and controls. Proportion of peaks containing selected transcription factor (TF) motifs (consensus sequences included) shown for each construct below. HG density represents the $-\log_{10}(\text{hypergeometric density})$. (F) Intersection of gene annotations for YAP-TFE3 ChIP peaks and controls. Proportion of peaks containing selected TF motifs shown for each construct below. (G) Scatterplot showing that many TAZ-CAMTA1 bound genes are also differentially expressed. (H) Validation of ChIP-Seq in SW872 cells in selected genes by ChIP-qPCR.

Figure 5

Fig. 5. TAZ-CAMTA1 and YAP-TFE3 promote the formation of chromatin accessible to transcriptional regulation. (A) Heat map of ATAC peak distribution around the transcriptional start site (TSS) region and average ATAC TSS profile for TAZ-CAMTA1 and controls and (B) for YAP-TFE3 and controls. (C) Intersection of gene annotations for transposase accessible chromatin peaks in TAZ-

CAMTA1 and controls and **(D)** YAP-TFE3 and controls. HG density represents the –
 $\log_{10}$ (hypergeometric density). p value for hypergeometric analysis included. **(E)** DiffBind principal
 component analysis for fusion proteins and controls. **(F)** Overlay of differentially expressed genes, bound
 genes, with transposase accessible chromatin for TAZ-CAMTA1. **(G)** Overlay of differentially expressed
 genes, bound genes, with transposase accessible chromatin for YAP-TFE3.

Figure 6

Fig. 6. The TAZ-CAMTA1 and YAP-TFE3 interactomes are enriched for transcriptional regulators and chromatin modifiers. (A) Overview of BioID method. (B) Summary of YAP-TFE3 and TAZ-CAMTA1 interactomes. (C) Dot plot representation of chromatin modifiers shared by YAP-TFE3 and TAZ-CAMTA1 interactomes. Larger dots indicate higher relative abundance. Darker color of dot indicates higher average spectral count. Color of border demonstrates the Bayesian false discovery rate. (D) Table of chromatin modifiers shared by YAP-TFE3 and TAZ-CAMTA1 interactomes. (E) Classification of prey proteins shared between TAZ-CAMTA1 and YAP-TFE3 interactomes. (F) Algorithm to prioritize the TAZ-CAMTA1 and YAP-TFE3 interactomes for the subsequent RNAi screen.

Figure 7

Fig. 7. The ATAC complex mediates the transforming capacity of TAZ-CAMTA1 and YAP-TFE3.

(A) Overview of RNAi screen. SW872 cells expressing TAZ-CAMTA1 are transduced with 5 shRNAs for each gene in the shared TAZ-CAMTA1/YAP-TFE3 interactome and proliferation assayed on poly-HEMA coated plates. (B) Poly-HEMA assay with SW872 cells expressing TAZ-CAMTA1. (C) Table with rank ordering of shared prey proteins after RNAi screen according to shRNAs causing a decrease in anchorage independent growth. (D-H) Proliferation curves for prey proteins/genes with the highest number of shRNAs demonstrating a reduction in anchorage independent growth on poly-HEMA coated plates.

Figure 8

Fig. 8. The ATAC complex is critical for the oncogenic phenotype of the fusion proteins. (A) Knock-down of *YEATS2* in SW872 TAZ-CAMTA1 cells by qRT-PCR. (B) Soft agar assay for SW872 TAZ-CAMTA1 cells after *YEATS2* knock-down. (C) Knock-down of *ZZZ3* evaluated by western blot. (D) Soft agar assay after *ZZZ3* knock-down. (E) Co-immunoprecipitation experiment (Co-IP) in HEK293 cells expressing Flag-TAZ-CAMTA1. Myc-KAT14 is transiently transfected. (F) Co-IP in HEK293 cells expressing Flag-YAP-TFE3. Myc-KAT14 is transiently transfected. (G) Co-IP in SW872 cells expressing TAZ-CAMTA1. Myc-KAT14 is transiently transfected. Overall survival curves as a function of *YEATS2* (H) and *ZZZ3* RNA expression (I) across different histological types of sarcoma utilizing RNA-Seq expression data for The Cancer Genome Atlas (TCGA). (J) Working model of TAZ-CAMTA1:ATAC complex.