



UNIVERSITAT DE
BARCELONA

Identification and characterisation of transcriptional regulators through proximity interaction mapping in two distinct systems

Michael Lewis

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.



UNIVERSITAT DE
BARCELONA

Universitat de Barcelona
Facultat de Farmàcia i Ciències de l'Alimentació
Programa de Doctorat en Biomedicina

Identification and characterisation of transcriptional
regulators through proximity interaction mapping in
two distinct systems

by

Michael Lewis

2022

Cover design and illustration: The research described in this thesis was performed from xyz at the IRB Barcelona, and was supported by a “la Caixa”-Severo Ochoa/IRB Barcelona International PhD fellowship. No parts of this thesis may be shared for any commercial purposes. Reproduction of this thesis may only be in the whole without any changes, and only if the original authorship

UNIVERSITAT DE BARCELONA
FACULTAT DE FARMÀCIA I CIÈNCIES DE L'ALIMENTACIÓ
PROGRAMA DE DOCTORAT

IRB Barcelona

Identification and characterisation of transcriptional
regulators through proximity interaction mapping in
two distinct systems

Memòria presentada per Michael Lewis
per optar al títol de Doctor per la Universitat de Barcelona

Director
Dr. Xavier Salvatella

Author
Mr. Michael Lewis

Tutor
Dr. Albert Tauler Girona

Institute for Research in
Biomedicine of Barcelona



INSTITUTE
FOR RESEARCH
IN BIOMEDICINE

Universitat De Barcelona
Facultat de Farmàcia i Ciències
de l'Alimentació, Programa de
Doctorat en Biomedicina



UNIVERSITAT DE
BARCELONA

Acknowledgments

I would like to start by thanking my thesis directors. First **Travis**, thank you for giving me the chance to join your lab. Your mentorship and knowledge have been invaluable to me over the course of this PhD. You have taught me a great deal about good, accurate and true science. I truly appreciate your constant, constructive feedback and open door/zoom chat policy that have enabled me to grow and work with confidence and independence to become a better scientist. It would be hard to find a better mentor, and I wish you all the best in your future endeavours.

Secondly, I would like to extend my gratitude to **Xavi** for taking me into his lab halfway through my PhD and in the middle of a pandemic! You have made me feel welcome from day one and been more than accommodating to my situation. I have learnt a great deal about a completely new and exciting area of biology thanks to your expertise and guidance. I am very grateful for you giving me the freedom and to find my ways in it, it has been a great journey together!

To my **thesis advisory committee**; Jens Luders, Roger Gomis and Sandra Peiró, thank you for following the developments of my research and the valuable input. I would also like to thank my **thesis defence committee**; Rual Mendez, Sandra Sdelci and Zoi Lygerou for their time and evaluation of my thesis.

This work would not have been possible without our collaborators. Thanks to **Moe Mahjoub**, **Sudipto Roy** and **Haotian Zhao** for their time, resources, and contribution to the formulation of manuscripts.

Thanks to our partner labs, especially the **Luders** and **Nebreda** lab. These groups when I enter, have always been full of helpful and happy people, ready to support you or lend anything, I am still very much indebted to you guys over the years! Our collaboration with the **Supek's** lab as well, has been great, especially at the start of my PhD. Thanks to all the facilities of IRB and PCB for doing a good job at running a professional service. Notably, thanks **Camille** of the IRB Biostatistics unit for her analysis and timely work. Thank you **Leyre** for all the help, explanations and emails, and for making IRB a more fun place.

I owe a special thanks to the great **Mass Spectrometry** department at IRB. To **Marta**, **Mar** and **Marina** for the organisation, understanding and running of so many critical parts of my project with diligence and precision. In addition, to **Gianluca** for the assiduous level of analysis provided, whilst always being engaged to provide improvements in next to no time. The facility is a great team and I hope it continues.

I am very grateful to **EBL** for accepting me into their space. Especially, **Isabelle Brun Heath** for always having time and interest in your project and providing genuine help in any area. As well as **Nuria Villegas** for being so attentive and running a super organised and always plenished lab! As well as the other EBL members for creating a friendly and enjoyable place to work.

I need to express my gratitude to the **LMB** lab, whom adopted me as one of their own. Firstly to, **Mireia**, **Jomi** and **Isabelle** for being so and helpful and capable. Least not forgetting; **Carla** and **Federica** I hope you find success in your post PhD careers, whatever that may be. **Borja**, you will make a great PI! **Mateusz**

as well, thanks for all the help I have asked of you, your knowledge was really valued in my last two years here. **Stase**, thanks for showing me the ropes and all the talks we had, you made the transition a lot easier, enjoy the mountains and good luck with everything! **Levante**, it was short but sweet, make the most of Barca and I hope your time at IRB will be great! **Paula** thanks for the all support in LMB, it was great working beside you and I hope you return to work soon with full health and vigour.

I spent some of the best years of my PhD in the Stracker lab thanks to an amazing set of people that were there. Firstly, I owe my respects to **Berta**, for pathing the way of the GEMC1 project and providing me with a kick-start to a great project. To **Sandra**, I learnt a great deal in the performance and execution of good science in my early years, thank you for being a great role model! **Isabel** also, thanks for your time and help with everything in my project. **Marina**, working besides you all those years has been a lot of fun, I miss listening and laughing at all your problems together, IRB was a lot more boring without you! Enjoy the rest of your time at NIH/Holiday-planning: ;) **Josep/Pepi**, me amigooo, err que tal? (con acento guri ;)). Thank you for all the laughs and immature humour that you gave, it made the days pass fast, even though they were less productive! **Maria**, never a dull moment in your company, our time was cut short, but I will remember it - full of jokes, good times and our silly chats taking the mick out of each other and life! **Lluiiiiiisss** best lab tech ever, thanks for also being so supportive, kind, and humorous it has always been fun chatting about Barca's/ManUs' woes. **I-laria**, my cilia buddy! Always full of energy, enthusiasm and pure positivity. I owe you so many beers for all the favours and advice you gave me! I wish you all the happiness in whatever field you end up in, you deserve it ☺. **Chithran**, not sure I could have completed this without all your help, one day you will be old and wise like me, I'm sure you will do a great job in the rest of your PhD, especially without me bothering you, thank you!

A special thanks goes to my fellow starters of this PhD, whom have gone on this long journey with me - the dream team - **Cris, Nevenka, Petra** and **David**. Without our regular beer escapades and shared venting of the PhD struggles I wouldn't have made it to the end. I'm sure in the next chapters of our lives we will be scattered all over the world, but I hope we can meetup in some bolera somewhere and share un cubo de cervezas and reminisce about good times.

Babette, I've been so lucky to have had you with me these last two years. To see your face, to hear your positivity at the end of each day has meant so much to me. To all those long days and weekends your spirit never wavered even when mine did, I really owe you for that Guapa ☺ Bedankt voor alle liefde die je me elke dag laat zien, Ik hoop dat we nog vele dagen samen zullen doorbrengen.

Mum & Dad. Thank you for getting me to this point, for encouraging my curiosity and pursuement of a career I really enjoy. To **Rob** for the encouragement of the funer things in life, providing much needed optimism and escapism from the toils and troubles of PhD life. Thank you all for the weekly support and reassurance in all areas of life. Spending my holidays to come see you guys and vice a versa gave me much needed grounding and respite to persevere. I hope that the most arduous days are past and we can enjoy more time in each other's presence.

Table of contents

| | |
|---|----|
| Acknowledgments..... | 2 |
| Summary..... | 8 |
| Abbreviations..... | 10 |
| Introduction..... | 13 |
| 1. Regulation of transcriptional specificity in development and disease | 14 |
| a. Transcription and Transcription Factors..... | 14 |
| b. Condensates in biology..... | 18 |
| c. TFs and Co-activators in disease..... | 22 |
| 2. Role of Geminin family transcriptional activators in the development of multiciliated tissues and ciliopathies | 24 |
| a. The transcriptional differentiation of multiciliated cells | 24 |
| b. Transcriptional activation of MCC differentiation by GEMC1 and MCIDAS..... | 30 |
| c. Consequences of MCC dysfunction..... | 36 |
| 3. Ligand activated transcription factors in prostate cancer..... | 39 |
| a. Androgen and Glucocorticoid receptors..... | 39 |
| b. AR Structure and signaling pathways..... | 39 |
| c. Prostate cancer (PC)..... | 41 |
| d. AR transcriptional cofactors and condensate formation | 44 |
| e. Castration-resistant prostate cancer (CRPC)..... | 45 |
| f. Small molecule inhibitors | 46 |
| Objectives..... | 49 |
| Materials & Methods | 52 |
| Results..... | 64 |
| Chapter I..... | 65 |
| Cellular Characterisation of GEMC1 and MCIDAS | 66 |
| GEMC1 and MCIDAS activate common and specific gene targets | 68 |
| Identification of the proximal interactomes of GEMC1 and MCIDAS | 70 |
| GEMC1 and MCIDAS proximally interact with Mediator and distinct SWI/SNF complexes | 73 |
| Sequence C-terminal to the CC domain regulates distinct SWI/SNF interactions | 75 |
| TIRT domain swapping reverses SWI/SNF specificity | 77 |
| Distinct SWI/SNF complexes influence transcriptional activation | 79 |
| Inhibition of BRD9 inhibits Multiciliogenesis in an MTEC model | 81 |
| BRD9 depletion inhibits multiciliogenesis in an inducible <i>in vitro</i> model..... | 83 |
| CCNO influences the expression of cell cycle components..... | 90 |

| | |
|--|-----|
| Potential roles of GEMC1 in Cancer | 92 |
| Chapter II | 95 |
| Identification of Androgen Receptor proximal interactions in AD293s | 96 |
| Proximal interactions of AR-VAR7 in AD293s | 98 |
| Proximal proteome of AR in stable cell lines | 100 |
| Mutating ARs Tyrosine residues depletes protein interactions | 102 |
| 22YtoS fails to translocate to the nucleus and interacts with Nucleoporins | 105 |
| AR Tyrosine mutants are transcriptionally impaired | 107 |
| Tyrosine mutants with AR Variant 7 depletes protein interactions | 109 |
| Small molecule inhibitors deplete the AR proteomic network | 111 |
| Small molecule inhibitors alter the homeostasis of specific transcriptional interactions | 113 |
| Discussion | 117 |
| 1. BioID proximity labelling identifies interaction networks in Multiciliogenesis | 118 |
| 2. BioID proximity labelling identifies interaction networks in PC | 122 |
| 3. Phase separation and Tyrosine Mutants | 124 |
| 4. Implications for disease | 127 |
| Conclusions | 131 |
| Disclosure | 132 |
| References | 134 |

Summary

Transcriptional regulation is a multifaceted and fundamental process, which underpins normal physiology, and its dysregulation is associated to numerous pathologies. Using two inherently different transcriptional programmes, we probed the complex nature of the components and systems that govern their organisation.

1) The differentiation of Multiciliated cells (MCCs) requires the sequential action of the atypical transcriptional activators GEMC1 and MCIDAS. Activation of the pathway results in the projection of dozens to hundreds of motile cilia from the apical surface of MCCs to promote the movement of fluids in the mammalian brain, airway or reproductive organs. How these factors operate and the extent to which they play redundant functions remains poorly understood. Here, we demonstrate that the transcriptional targets and proximal proteomes of GEMC1 and MCIDAS are highly similar. However, we identified distinct interactions with SWI/SNF subcomplexes; GEMC1 interacts primarily with the ARID1A containing BAF complex and MCIDAS with the BRD9 containing ncBAF complex. Inhibition of BRD9 impaired MCIDAS-mediated activation of several target genes and compromised MCC differentiation in two Multiciliogenesis models. We further began to decipher the regulatory role CCNO, a cyclin expressed during MCC differentiation, has on this complex process, highlighting a number of important interactions. Our data suggests that the differential engagement of distinct SWI/SNF subcomplexes by GEMC1 and MCIDAS is required for MCC-specific transcriptional regulation and that CCNO regulates the process.

2) Large transcriptional apparatus can form transcriptional condensates, in a process known as phase separation, aberrations of which can affect the activity of genes responsible for human disease. Castration resistant prostate cancer is a disease that is driven by the pathogenic Androgen Receptor (AR), a transcription factor which is currently thought to be “undruggable”. Here we demonstrate, that by altering the aromaticity of the AR through tyrosine mutations, we can impair the formation of condensates that results in defective translocation, protein interactions and transcriptional activity. Furthermore, using a panel of novel small molecule inhibitors, we could perturb AR phase separation and arrest its protein interactions and functions in tumour development.

Abbreviations

| | |
|---|----------|
| Activation function | AF |
| Air liquid interface: | ALI |
| Amino Acids | aa |
| Androgen deprivation therapy | ADT |
| Androgen Receptor | AR |
| Androgen response elements | ARE |
| Aryl hydrocarbon receptor | AHR |
| Basal body: | BB |
| Bayesian false discovery rate | BFDR |
| BioID affinity purifications | BioID-AP |
| BRG1/hBRM-associated factors | BAF |
| Castration-resistant prostate cancer | CRPC |
| Coiled-coil | CC |
| Cyclin O | CCNO |
| Deuterosome-dependent | DD |
| Dihydrotestosterone | DHT |
| DNA-binding domain | DBD |
| Doxycycline | DOX |
| Fold change | FC |
| Forkhead Box A | FOX |
| Geminin | GMNN |
| Geminin-like coiled-coil containing protein 1 | GEMC1 |
| Gene Ontology | GO |
| Glucocorticoid receptor | GR |
| Intrinsically disordered proteins | IDP |
| Intrinsically disordered region; | IDR |
| Ligand-binding domain | LBD |
| Mass Spectrometry | MS |
| Mediator Complex | MED |
| Medullary thymic epithelial cells | mTEC |
| MiniTurboID | MTID |
| Multiciliated cell | MCC: |
| Multicilin, Idas | MCIDAS |
| N-terminal domain | NTD |
| Nuclear localisation signal | NLS |
| Nucleotide enrichment score | NES |
| Pericentriolar material: | PCM |
| Polycomb group RING finger protein | PCGF1 |
| Polyglutamine | pQ |
| Prostate Cancer | PCM |
| Radial Glial Cells: | RGC |

Reduced generation of multiple motile cilia
RNA polymerase II
RNA Sequencing
Significance Analysis of INTeractome
Streptavidin
Super-enhancers
Transcription factors
Variant 7

RGMC
RNAP
RNA-Seq
SAINT
Strep
SE
TF
VAR7

Introduction

1. Regulation of transcriptional specification and disease

a. Transcription and Transcription Factors

Transcription, the process of copying the information in a strand of DNA into a molecule of messenger RNA (mRNA), is fundamental to all living organisms. Thus, understanding how transcription is regulated is of great importance in biology, as it orchestrates the vast majority of biological processes, from cell differentiation to apoptosis.

Transcription factors (TF) are DNA-interacting proteins that govern the expression of specific genes. They serve as key components of signalling pathways by binding to specific short DNA base pair patterns, termed motifs or cis-regulatory elements (CREs), in the promoters or enhancers of target genes (Rebeiz et al., 2017). Transcription is initiated by the promoter, that is key to controlling when and where in the organism a particular gene is expressed, and is further amplified by enhancers (Andersson et al., 2019).

Typically, a promoter sequence is a short region of DNA approximately 100–1,000 bp, located directly upstream or at the 5' end of a specific sequence in DNA termed a transcription start site (TSS) (Le et al., 2019). The core promoter sits nearest to the TSS and typically contains a well characterised DNA sequence, typically a TATA box or TATA recognition element (BRE), that indicates where a genetic sequence can be read and in which direction (Mishal et al., 2022). Further upstream, from a few hundred to kilo-base pairs are the proximal and distal promoters which contain many additional regulatory elements. Transcription is further regulated by position and orientation specific elements, including enhancers, insulators, and silencers. These contribute to the regulation of gene expression, as discussed later (Lin et al., 2018).

Transcription takes place in three key stages (Figure 1). The Initiation step requires the binding of several general TFs. This mediates the binding of RNA polymerase (RNAP), whilst ensuring the preceding DNA is unwound in a “transcription bubble,” at a selected TSS. An Elongation step then proceeds where the RNAP reads and synthesizes an mRNA product from complementary nucleotides, in typical short bursts of 11–15 nucleotides in length. Throughout elongation, there is a proofreading step that ensures the transcribed RNA molecule does not insert incorrect bases. Lastly, the termination step occurs where transcription ends, as the RNAP meets a termination sequence. The mRNA strand is then complete and detaches from DNA to be exported to the cytoplasm, where they are finally translated into proteins by the ribosome.

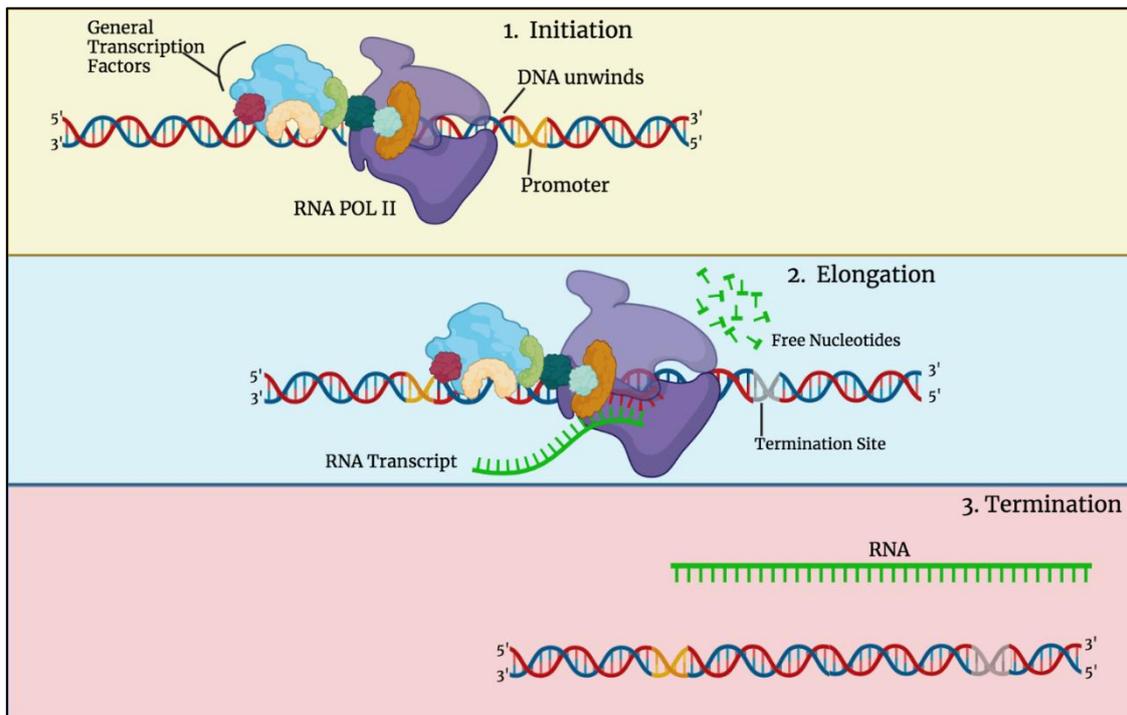


Figure 1 – Stages of Transcription. Three stages of transcription. 1) Initiation: RNA Polymerase binds to DNA at specific transcription start site, located in the core promoter. DNA is unwound prior to RNA Polymerase. 2) Elongation: nascent nucleotides are utilised to generate mRNA from complementary template strand of DNA as RNA polymerase moves along DNA. 3) Termination: transcription ends as the RNA Polymerase meets a termination sequence in the DNA. The mRNA strand is complete and it detaches from DNA. Created with BioRender.com.

Generally, TFs have at least one DNA binding domain, one or more transcription activation domains (TAD), which are responsible for recruiting transcriptional machinery to target promoters/enhancers, and often a dimerisation domain. In many cases, TF also have other protein–protein interaction domains (Brent and Ptashne 1985; Keegan et al., 1986). It is estimated in mammals that there are at least 1500 TFs present, conserved in at least 70 different sub-classes, reflecting homologies in their DNA-binding domains (Lambert 2018; Fulton et al., 2009; Vaquerizas et al., 2009). Three of the largest categories of site-specific DNA binding factors (which encompass over 50% of all TFs), interact with the genome through conserved DNA binding domains called zinc fingers, homeodomains, and helix–loop–helix domains (Vaquerizas et al., 2009).

Single TFs are capable of binding and altering the expression of thousands of genes, although often they function in unique multifaceted networks between TFs and TF binding proteins to govern specific expression patterns. These include interactions with cofactors, regulatory proteins, such as phosphatases and kinases, as well as interactions with dimerisation partners, subunits and inhibitory proteins. Moreover, as chromosomal DNA is organised into chromatin, partially to avoid uncontrolled transcription, TFs also

interact with several chromatin re-modelling proteins. These complexes are essential to ensure the accessibility of DNA and permit the opening of chromatin and gene transcription (Figure 2).

Some TFs are described as ‘pioneer factors’ due to their ability to initiate chromatin remodelling by binding to sites located in condensed, inactive chromatin. Additionally, they facilitate the binding of many other TFs to control tissue-specific gene activation.

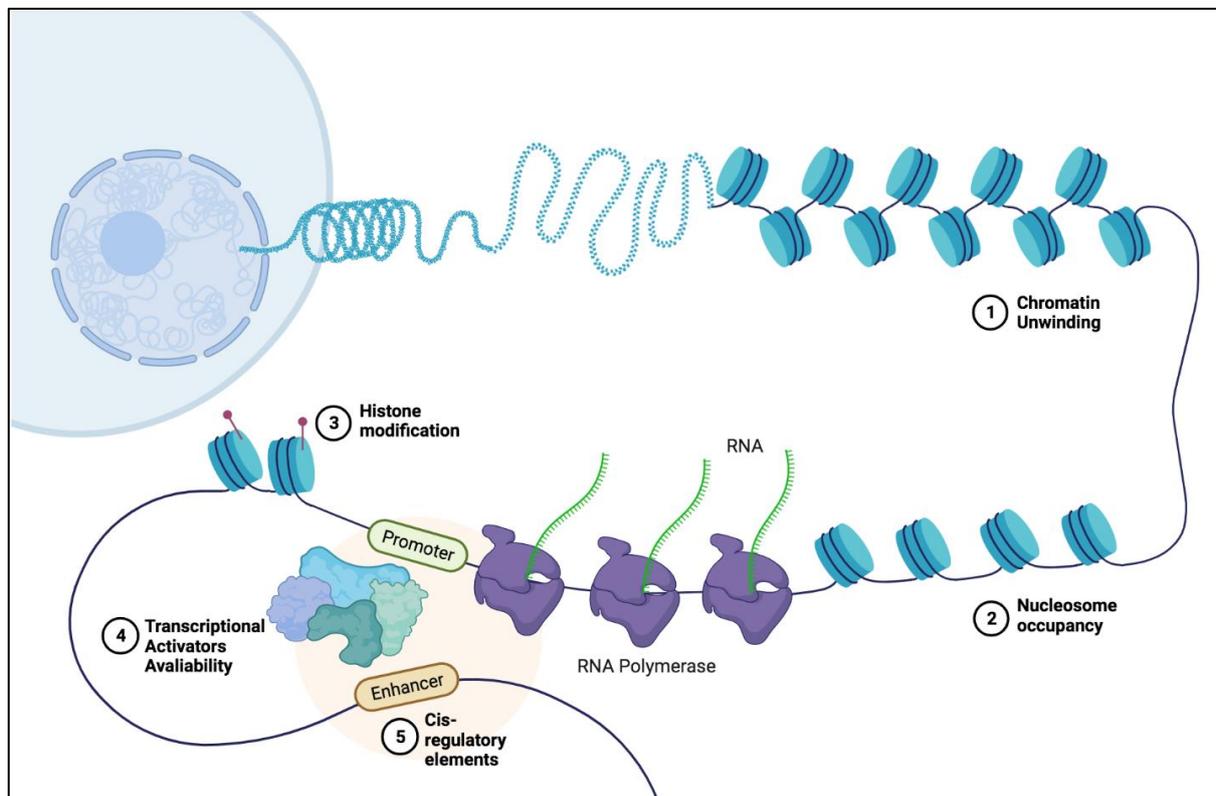


Figure 2 – Chromatin unwinding and transcription regulation. Factors that regulate transcription, not necessarily in order. DNA has to be unwound from a tightly compacted and condensed state around histones. A nucleosome is the basic structural unit of DNA packaging and their levels of occupancy are a common determinant of gene regulation. Histone modifications are marks that regulate chromatin compaction, nucleosome dynamics, and transcription. RNA Polymerase mediated transcription takes place around active promoters and cis-regulatory elements with the aid of transcriptional activators/co-factors. Cis regulatory elements are noncoding DNA sequences which play a critical role in gene expression regulation. Adapted from “Regulation of Transcription in Eukaryotic Cells”, by BioRender.com (2022). Retrieved from <https://app.biorender.com/biorender-templates>

Promoters represent critical elements that work in concert with other regulatory regions. Typically, transcription driven by only core promoters is of low output (Roeder et al., 1996). Enhancer elements collaborate with promoters from afar, creating long-range interactions or promoter-enhancer loops which fundamentally contribute to the control of transcription (Sanyal et al., 2021).

Enhancers are sequences of genomic DNA elements several hundred bps in length. They function to boost transcription of a target promoter independently of the enhancer's relative orientation and distance (Zabidi et al., 2016). Their sequence contains TF recognition sequences which enable a number of TFs to bind, which in turn can recruit an array of co-factors. Transcriptional co-factors (coactivators and corepressors) are typically unable to bind to DNA but mediate communication between regulatory TFs, the core promoter and enhancers (Zabidi et al., 2016).

TFs frequently bind to proximal gene promoter sequences, but they are more often bound to distal enhancers which can be located many MBs away from their target gene (Grosveld et al., 2021). Thus, enhancers enable promoters to achieve high rates of transcription in a tissue specific manner.

Emerging work suggests that enhancers function within larger domains, in clusters termed super-enhancers (SE). These SEs are clusters of smaller enhancers in close proximity, which are occupied by an unusually high density of transcriptional machinery (Figure 3). By compartmentalising RNAP and TFs, they have been shown to control mammalian gene expression and play essential roles in cell-type-specific processes (Hnisz et al., 2013; Whyte et al., 2013).

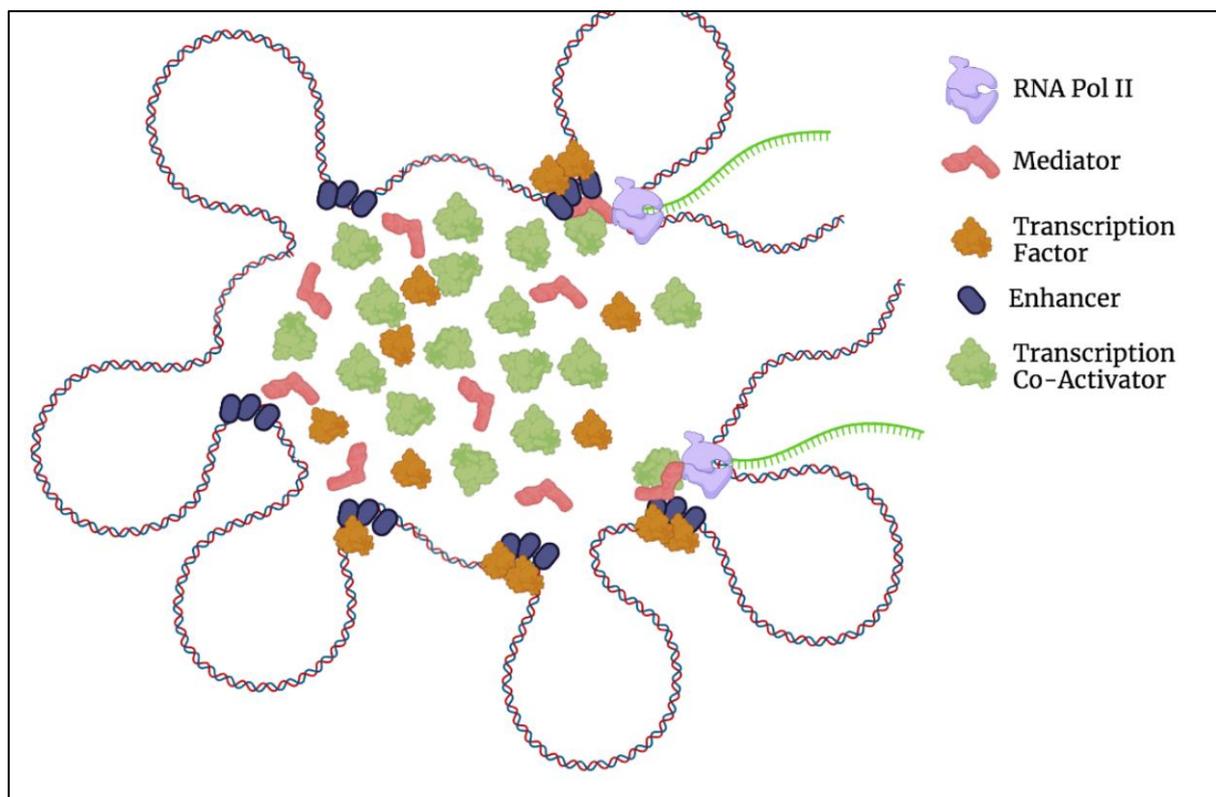


Figure 3 – Super Enhancers. A highly dense area of local enhancers that are occupied by large occupancy of transcriptional activator proteins are termed super enhancers. The Mediator complex is key to promote the recruitment of transcription factors and co-activators as well as linking distal enhancers to their promoters in regulating gene expression. Created with BioRender.com.

b. Condensates in biology

i. Intrinsically disordered proteins (IDRs)

Protein-Protein interactions normally depend on intricate geometric compatibility and noncovalent interactions between folded structures. However, the majority of TFs (>60%), contain intrinsically disordered regions (IDRs), that lack well-defined secondary and tertiary structures, which is thought to enhance its structural plasticity, enabling these domains to adapt to a multitude of conformations with distinct properties and thus provide multifunctionality (Liu et al., 2006; Fuxreiter et al., 2011; Guo et al., 2012; Minezaki et al., 2006; Radhakrishnan 1997). This phenomenon of so-called protein disorder has been known since the 1990s and questioned a key theme in molecular biology; that structure defines function, exemplified by the “lock-and-key” model.

Intrinsically disordered proteins (IDP's) are capable of an alternative method of interaction facilitated via long-range electrostatic attractions, which is often promiscuous (Borgia et al., 2018; Protter et al., 2018). This flexibility offers many advantages, primarily the ability to exhibit a wide range of both highly specific and low affinity (Kd) interactions (Arbesú et al., 2018; Gao et al., 2018).

Secondly, it is argued that IDRs increase the potential of proteins to possess differential post-translational modifications (typically phosphorylation, methylation and ubiquitination) which can drastically affect their affinity for their partners and their stability, thus vastly increasing the opportunities of refined regulation. (Tompa, 2002; Dunker et al., 2005; Haynes et al., 2006; Uversky et al., 2010; Pancsa and Tompa, 2012).

Lastly, IDPs are able to bind to multiple, distinct targets without sacrificing specificity (Dunker et al., 2005; Liu and Huang 2014). The TF's TADs are inherently disordered and repetitive in nature - thus, acting as flexible hubs for multiple interactions (Dunker et al., 2005; Ptashne and Gann, 1997). This engineers TFs to act as platforms to facilitate recruitment of a varied array of structural interaction partners and fulfil its function of intricately diverse, gene-specific transcription.

ii. Condensate formation in transcription

The cytoplasm and nucleus are very densely packed cellular environments and typically localise reactions and processes within membrane-bound organelles. However, there are many organelles that lack a physical barrier around them, such as the nucleoli, Cajal bodies or PML bodies (Kurihara et al., 2020; Gomes et al., 2018; Mao et al., 2011). Membraneless organelles vary in different states and function. It is thought that several membraneless organelles in cells arise by liquid-liquid phase separation (LLPS) or biomolecular condensation, comprised of their protein and RNA components (Bratek-Skicki et al., 2020). These condensates depend on weak multivalent interactions, such as those mediated by IDRs. This suggests that IDRs play a pivotal role in cellular function (Elbaum-Garfinkle et al., 2015; Wang et al., 2018).

Many of these organelles or compartments seem to form through intracellular phase transition such as LLPS (coined phase separation in this study) (Bergeron-Sandoval et al., 2016; Brangwynne et al., 2009, Feric et al., 2016, Hyman et al., 2014). Such condensates, produced by LLPS define environments that are ordered and thermodynamically driven with properties that are different from those of the nucleoplasm surrounding them (Figure 4).

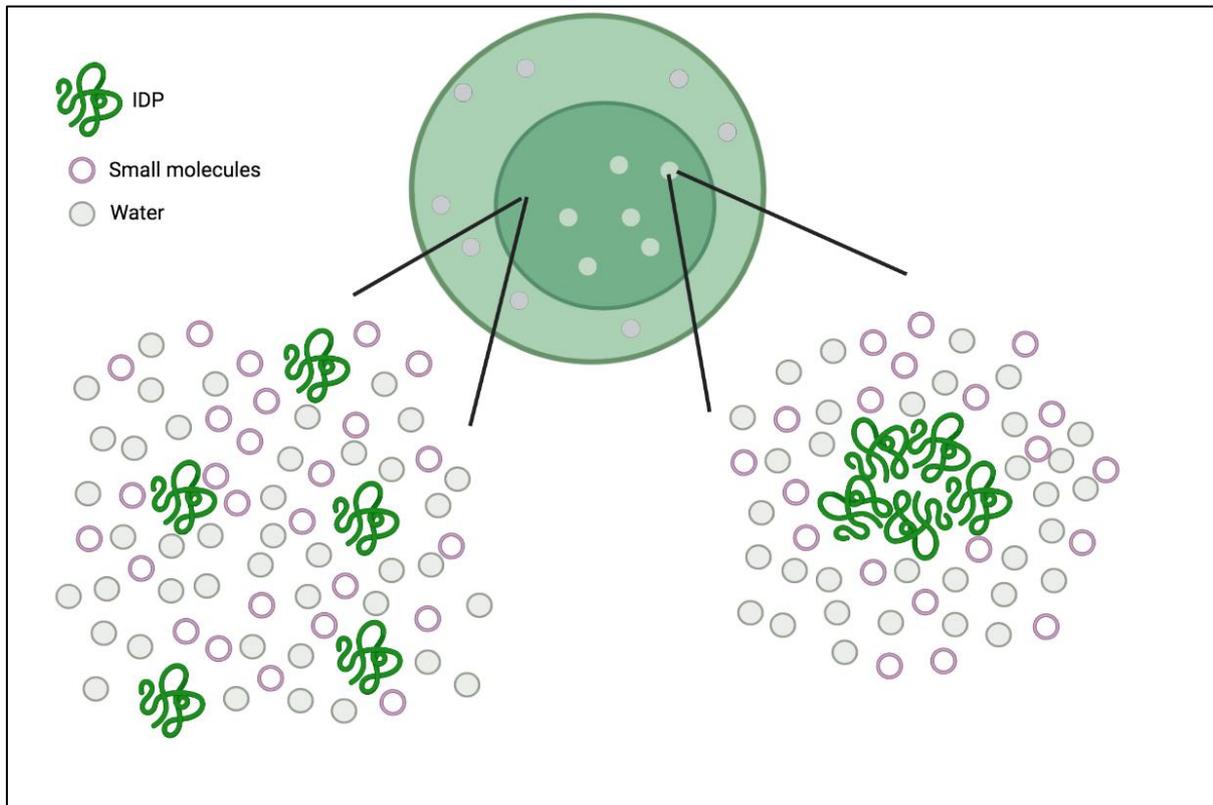


Figure 4 – Liquid-liquid phase separation (LLPS). Cell representing two forms of macromolecule organisation being disperse and random on the left or being part of phase separated membraneless organelles on the right. These condensates consist of many Intrinsically Disordered Proteins (IDP) and/or their protein or nucleic acid partners. Created with BioRender.com.

The macromolecular composition of condensates is specific and reversible. Our understanding of what governs the heterogeneity and propensity of condensates to form is in its infancy. However, the protein components of condensates have been classified as either scaffolds, which have been defined as the proteins that drive reversible condensate formation, or clients, which partition into pre-formed condensates (Banani et al., 2016, Banani et al., 2017; Feric et al., 2016, Zhang et al., 2015). Multivalent interactions (scaffold-scaffold and scaffold-client) allow the formation of networks, which is the basis for phase separation (Alberti et al., 2017).

DNA interaction data show SEs are in close spatial proximity with each other and the promoters of the genes they regulate (Downen et al., 2014; Ji et al., 2016). These have been shown to exhibit sharp transitions

of formation and dissolution, as the consequence of a single nucleation event (Mansour et al., 2014; Brown 2014) or collapsing when concentrated factors are depleted from chromatin (Lovén 2013; Chipumuro et al., 2014; Chapuy et al., 2013). Consequently, the latest research proposes that phase separation coordinates the spatiotemporal organisation of TFs and their co-factors to assemble at SEs. These form transcriptional condensates which activate gene expression (Hnisz et al., 2017; Boehning et al., 2018; Boija et al., 2018; Cho et al., 2018; Chong et al., 2018; Lu et al., 2018; Nair et al., 2019; Sabari et al., 2018; Shrinivas et al., 2019)

iii. SWI/SNF complex and Super Enhancers

SWI/SNF complexes consist of 12–15 subunits and are large multimeric molecular assemblies that regulate chromatin architecture (Kassabov et al., 2003) (Figure 5). SWI/SNF complexes contain an ATPase catalytic core and have been considered to have three sub-families: BAF (defined by ARID1A/B), PBAF (defined by ARID2, PBRM1, and BRD7) and ncBAF/GBAF (defined by BICRA/BICRAL, BRD9 (Mashtalir et al., 2018), which differ substantially via the incorporation of distinct subunits and in their targeting of different genes (Clapier et al., 2017).

SWI/SNF alters the assembly of nucleosomes and provides accessibility of TF and co-activators to their respective gene targets. Its importance is highlighted in that 20% of human cancers contain mutations in the genes encoding SWI/SNF subunit (Kadoch et al., 2013).

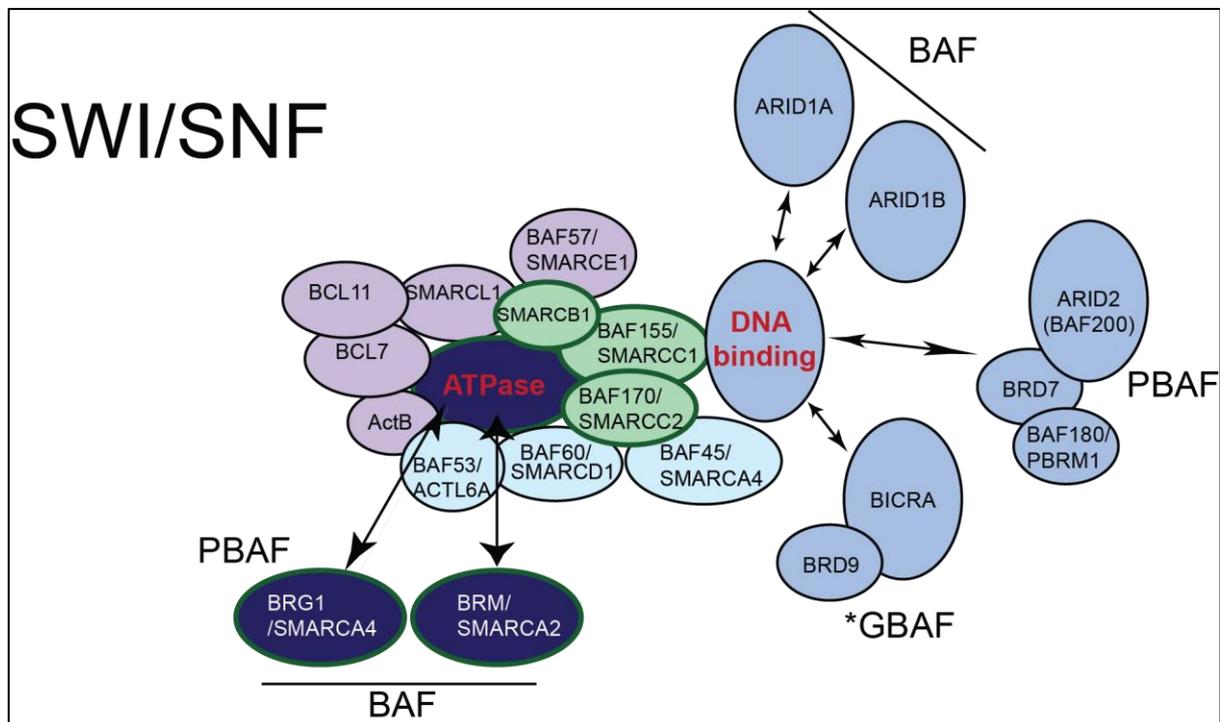


Figure 5 – SWI/SNF. Sub-unit map of the large chromatin modifier, and its interchangeable, ATPase and DNA binding domains.

The SWI/SNF complex is abundant in the nucleus and facilitates its role primarily at non-coding regulatory elements. These regions are often enhancers or SEs which have exceptionally high densities of transcriptional machinery to regulate genes, often involved in cell fate or identity (Whyte et al., 2013; Hnisz et al., 2013). These sites are distal to gene promoters and require looping interactions to potentiate RNAP-mediated transcription (Visel et al., 2009). Although enhancers play critical roles in the oncogenic process (Murakawa et al., 2016) little is known about the molecular machinery responsible for enhancer maintenance and/or activation.

SEs can also display transitions of formation and dissolution and can appear rapidly from single a nucleation event (Mansour et al., 2014; Brown et al., 2014). Based on the dynamics of these properties it has been put forward that the high concentration assembly of biomolecules at active SEs stems from phase separation of enriched factors at these genetic loci (Hnisz et al., 2017).

iv. Mediator

The Mediator complex, also known as Mediator, is a 33 multi-subunit protein complex that is key to formation of transcriptional condensates and RNAP activity (Figure 6) (Robinson et al., 2016; Allen et al., 2015).

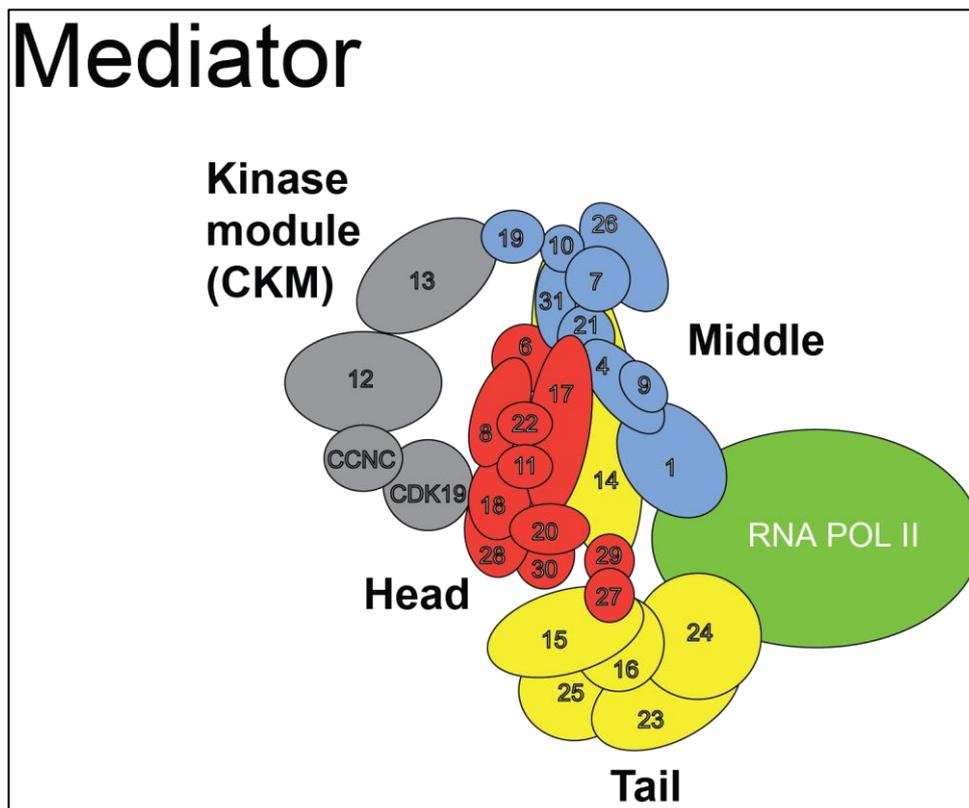


Figure 6 – Mediator. Schematic of the Mediator complex, comprised of three sub groups; tail, head and middle.

Mediator is aptly named, as it acts as a mediator or conduit between promoters and enhancers, linking DNA-bound transcription factors with RNAP to activate and repress gene expression (Weber et al., 2018).

Mediator can be divided into three main modules; head, middle, and a tail - along with an interchangeable kinase module (Robinson et al., 2016; Jeronimo et al., 2017). Approximately half of Mediator subunits contain IDRs, that confer its phase separation propensity (Nagulapalli et al., 2016). It has been put forward that this helps to provide an overall flexible structure of the complex so that it can recognise numerous interacting proteins and adopt diverse conformations to function in mechanistically distinct ways at different genes (Nagulapalli et al., 2016; Poss et al., 2013).

c. TFs and Co-activators in disease

i. Developmental disorders

Due to their vital biological roles in an array of processes, IDPs, including many TFs are frequently associated with the pathogenesis of a number of human diseases. The breadth of pathologies associated with TF malfunction is widespread. The stoichiometry of the transcriptional machinery is of importance, as nonspecific electrostatic interactions termed 'squenching' can cause inhibition of many genes, whereas overexpression may also have toxic effects on their expression (Kumar et al., 2021).

Pertinently, transcriptional co-activators are also intrinsically associated to disease. For example, CBP and P300 play a crucial role in development and are frequently mutated in Rubinstein-Taybi syndrome characterised by mental retardation, postnatal growth deficiency, microcephaly, and dysmorphic facial features (Roelfsema et al., 2005). Similar developmental disorders with comparable phenotypes have been associated with mutations in the BAF (BRG1/hBRM-associated factors) complex, a SWI/SNF family chromatin remodeler, that occur in Coffin-Siris syndrome (Santen et al., 2012; Tsurusaki et al., 2012). Likewise, mutations in the components of the Mediator co-activator are linked to a number of neurological and developmental disorders, including X linked mental retardation, Cerebellar Atrophy and axonal and Charcot-Marie-Tooth disease (Ding et al., 2008; Kaufmann et al., 2010; Leal et al., 2009).

From the initial stages of life, TFs carefully manage the sequence of events needed to achieve cell differentiation and identity. Starting from early embryonic development, the hallmark asymmetrical cell divisions along a basolateral cleavage are carried out by POU-TFs Oct 3 and 4. Nanog, Klf4, and Sox2 among others are essential to maintain the inner cell mass (ICM) and govern embryonic stem cell pluripotency (Chambers et al., 2003; Mitsui et al., 2003; Zhang et al., 2003), disturbances in the expression of these TFs is often lethal or leads to numerous developmental disorders.

ii. Cancer

Another hallmark of development is the Epithelial to mesenchymal transition (EMT). During this process, stationary apical-basal polarised epithelial cells transform into motile, front-back polarised mesenchymal

cells (Hay et al., 1995). This tightly controlled process administered by the delicate interplay of TFs is hijacked in cancer to enable tissue invasion and metastasis.

TFs influence every stage of cancer development and metastasis, the colonisation of the cancer to distant tumour organs. Typically, ‘gain of function’ oncogenes or ‘loss of function’ tumour-suppressors initiate the progression of the cancer that derive from mutation, gene amplification or epigenetic change. It is estimated that about 20% of oncogenes are TFs (Lambert et al., 2018).

One of the most well recognised oncogenic TFs is c-MYC. Its function involves the gene activation of 15% of the genome and its misregulation is associated with 30% of all cancers (Dang et al., 2012). Myc induces the expression of OPN and RhoA to help instigate EMT and therefore has multiple roles including; cell cycle processes, operating in primary tumour formation and metastasis (Bernards et al., 2004).

On the either side of the spectrum, p53 is the most well-established tumour suppressor TF. Upon activation, following cellular stresses such as DNA damage, p53 initiates a series of transcriptional programs that induces DNA repair, cell cycle arrest, metabolic reprogramming, apoptosis, or autophagy (Sionov et al., 1999; Menendez et al., 2009). *TP53* mutations that disrupt p53 function occur in nearly 50% of human cancers (Rayburn et al., 2005).

The Forkhead Box A (FOX) family includes more than 170 TFs that contain a common winged-helix DBD (Hannenhalli et al., 2009; Carlsson et al., 2002). They have been categorized into 19 subfamilies based on protein sequence homology (FOXA to FOXS) (Jackson et al 2010) and influence all aspects of cellular mechanics from proliferation to senescence. Unsurprisingly, the FOX family transcriptional imbalance is linked to a wide range of cancer types and potential usefulness as molecular biomarkers (see Bach et al., 2018). FOXA members are pioneer factors that have been shown to augment tissue-specific gene activation, such as estrogen and androgen modulation. They are therefore potential drug targets for both breast and prostate cancer (Jäggle et al., 2017; Iwafuchi et al., 2014; Cirillo et al., 2002)

Despite the majority of TFs being described as ‘undruggable’, a better understanding of TF regulation, protein/protein interaction and binding of their co-activators could provide new insight into their mechanism of gene transcription and open up new avenues to utilise TFs and their partners as drug targets for therapy.

2. Role of Geminin family transcriptional activators in the development of multiciliated tissues and ciliopathies

a. The transcriptional differentiation of multiciliated cells

i. Centrosomes and Centrioles

Canonical centrosomes are subcellular organelles approximately 1 μm in size that are comprised of two centrioles. They play an ever-increasing number of important roles including asymmetric division, maintenance of cell shape, cell orientation and migration, as well as meiotic recombination (Marjanović et al., 2015). Centrosomes are embedded within a matrix of proteins known as the pericentriolar material (PCM) that are responsible for microtubule nucleation and anchoring (Azimzadeh and Bornens, 2007).

Centrioles themselves are cylindrically shaped structures arranged perpendicular to each other, composed of nine sets of triplet microtubules symmetrically arranged around a central axis (Conduit et al., 2015). Being comprised of microtubules, the centriole is an inherently polar structure with microtubule minus ends positioned at the proximal end of the centriole.

However, the two centrioles of a given centrosome are not equivalent. Firstly, they differ in age, the older centriole, termed the Mother, arose at least one cell cycle earlier than the second, termed the Daughter centriole in a pathway termed the mother-daughter pathway (MD). Secondly, the Mother centrioles have distal and subdistal appendages, which are absent from the Daughter centriole. These sheet-like structures emanate from the centriole's distal end, and are required for supporting the recruitment of the PCM and MT anchoring in interphase cells (Delgehr et al. 2005) The daughter centriole is tethered to the mother centriole and matures to acquires such specializations following the next mitosis.

ii. Centrosome cycle

Centrosome replication is highly regulated and occurs in coordination with the cell cycle in proliferating cells. The pathway of the centrosome cycle is evolutionarily conserved among species and involves several core proteins, including Cep192, Cep152, Plk4, SCL-interrupting locus protein (STIL), and spindle assembly abnormal protein 6 (SAS-6) (Nigg & Holland 2018; Habedanck et al., 2005, Kim et al., 2013, Ohta et al., 2014). In the G1 phase of the cell cycle, each cell contains a centrosome with a pair of parental centrioles, each of the mother and daughter centrioles are able to nucleate one procentriole at its proximal end. These procentrioles elongate to full-length centrioles in S to late G2 phase, where the two centriole pairs separate and PCM recruitment promotes the formation of the mitotic spindle (Fujita et al., 2016). At

the end of mitosis, each daughter cell inherits a pair of centrioles resulting in two pairs of mother and daughter centrioles.

Centriole duplication initiates at the G1/S transition and is regulated by the activity of Plk4, a divergent member of the Polo-like kinase family. In mammalian cells, PLK4 activation is mediated by CEP152 and CEP192 that interact with PLK4 and promote its recruitment encircling the proximal end of the parent centriole (Hatch et al., 2010; Kim et al., 2013). This step defines the site of origin of centriole duplication along the mother centriole (Kim et al., 2013) as Plk4 overexpression causes multiple centrioles to form adjacent to the two existing centrioles (Habedanck et al., 2005).

PLK4 transforms from an initial ring-like localisation to a single dot at the G1–S phase transition on the wall of the parent centriole (Kim et al., 2013; Ohta et al., 2014). This transition requires binding of PLK4 to STIL, which activates PLK4 to autophosphorylate its kinase domain and promote its own kinase activity (Lopes et al., 2015; Moyer et al., 2015). Active PLK4 then phosphorylates STIL within the so-called STAN motif, triggering the centriolar recruitment of SAS-6. The SAS-6 proteins form rod-shaped homodimers with their coiled-coil domains to organise the central cartwheel, a stack of ring-like assemblies with nine-fold symmetry which initiate the structural foundation for the procentriole (Breslowe et al., 2019; Ohta et al., 2014; Moyer et al., 2015). The centriolar microtubules and other machinery are further primed for assembly by the Centrin SAS6-containing cartwheel (Zhao et al., 2019). After the cartwheel assembly, the centriole protein CPAP has a crucial role in the formation and stabilization of the triplet microtubule blades that make up the procentriole wall through interacting with multiple proteins including STIL (Cottee et al., 2013) CEP152 (Cizmecioglu et al., 2010), CEP120 (Lin et al., 2013) and CEP135 help to determine centriole length. The procentriole continues to elongate throughout G2, until the capping protein CP110 and the centrosomal protein 97 (CEP97) binds to the distal end of the procentriole to prevent further growth and maintain the proper centriole length.

iii. Primary and Motile Cilia

Centrioles are multifaceted and key components of the cell. They serve to enable the successful division of the cell through mitosis and are also integral in generating structures known as cilia, that play diverse functions in the cell. Cilia can be non-motile, where they are referred to as primary cilia, or one of many types of motile cilia.

Primary cilia can form during interphase or when the cell exits cell cycle into the quiescent G₀ phase (Bornens, 2002). The primary cilium (also referred to as a sensory cilium) is a microtubule based, hair-like organelle that projects from the surface of many cell types to control a number of important physiological processes.

In the formation of the primary cilium, or ciliogenesis, a fully mature mother centriole migrates to the apical surface of the cell and associates with the ciliary vesicle, it proceeds to dock with the plasma membrane to become a basal body (BB). This structure templates cilium formation, where ciliary trafficking machineries such as intraflagellar transport (IFT) complexes A and B deliver material to the growing cilium through the transition zone based on the kinesin and dynein motor proteins (Kim and Dynlacht, 2013). The cilium is comprised of axonemal microtubules that elongate from the distal end of the basal body and a ciliary membrane that surrounds the axoneme, this “antenna” like extension must first be disassembled prior to the start of mitosis so that the mature centriole can support the formation of the mitotic spindle (Rasi et al., 2009; Breslowe et al., 2019).

Classically, there are two classes of cilia distinguished by a central pair of microtubules, motile cilia which has typically a “9+2” structure, with dynein arms moving against the central pair to initiate ciliary movement. Whilst, Non-motile primary cilia or sensory cilia, lack this central pair and dynein arms, having a “9+0” structure. They are found on the apical surface of most cell types in the vertebrates, they sense extracellular chemicals osmotic, mechanical or optical signals (Berbari et al. 2009). Sensory cilia are also critical for the regulation of important developmental pathways, such as Wnt (Lancaster et al. 2011) or Sonic Hedgehog signalling (Breunig et al. 2008).

Motile cilia are characterised by their axonemal movement and co-ordinated beating dynamics. Generally, they are longer, and utilise energy from ATP hydrolysis to drive rhythmic motility of the axonemes. There are different subtypes, for example motile monocilia, such as the flagella on unicellular eukaryotes and sperm cells. In addition, solitary motile cilia that rotate are present on cells of the embryonic node to propel growth factors in a directional fashion for establishment of left-right body axis (Hirokawa et al., 2006). These cilia generally beat in a wavelike or corkscrew fashion in order to generate cellular locomotion or fluid movement (Kramer-Zucker et al., 2005).

iv. Multiple motile cilia

Multiciliated cells (MCCs) are specialised epithelial cells that project multiple motile cilia required for respiratory, reproductive, renal and brain functions in many vertebrates (Spassky et al., 2017). In humans, MCCs are present in the ependyma and choroid plexus of the brain to direct the flow of cerebrospinal fluid, the airways to clear mucus and pathogens, and in the efferent ducts and oviducts for spermatozoa and egg transport, respectively. Depending on the tissue, dozens to hundreds of motile cilia are generated per MCC that can beat in a coordinated, directional manner or generate turbulence through whip-like movement (Spassky et al., 2017; Yuan et al., 2019)

To direct fluid flow, cilia must beat in the right orientation. Thus, MCCs are polarized and aligned by the planar cell polarity (PCP) pathway with respect to the whole tissue (Butler et al., 2017). The ciliary beat is defined by the position of BBs within the plane of the apical membrane, BB rotational polarity, and subapical actin networks which are vital for BB spacing, as well as beating synchrony which is well reviewed (Meunier et al., 2019; Werner et al., 2011).

The advancement in knowledge surrounding the fundamental principles of MCC biology have mostly derived from four main models; the ciliated epidermis of the *Xenopus* embryo, the pronephros of zebrafish and the airway mucociliary or brain ependymal epithelia of the mouse. MCCs can also be found in a range of small-sized aquatic invertebrate organisms, where they are mainly involved in locomotion (Meunier et al., 2016, Tamm and Tamm 1988).

v. Centriole amplification

iv.

All MCCs are post-mitotic. However, although they have exited the cell cycle, they are remarkably able to rapidly generate hundreds of centrioles per cell. The MD pathway of centriole duplication is estimated to account for only around 10% of BB production in MCCs (Al Jord et al., 2014). The bulk of centriole production in MCCs is normally accomplished by the deuterosome-dependent (DD) pathway to facilitate centriole expansion which is discussed later. (Zhao et al., 2013; Klos Dehring et al., 2013 Spassky et al., 2017)

The ordered stages of the DD pathway are now well established (Figure 7). At the onset of centriole amplification, a Centrin-2 cloud or 'halo' accumulates around the pre-existing centrosome (Stage-I or Amplification (A)) and this signal intensifies to adopt a 'flower-like' shape (Stage-II or Growth(G)), corresponding to maturing centrioles that appear in a synchronized manner. And finally, the coordinated disengagement (Stage III or D) and release of centrioles and their migration to the apical surface as BBs takes place (Al Jord et al., 2014).

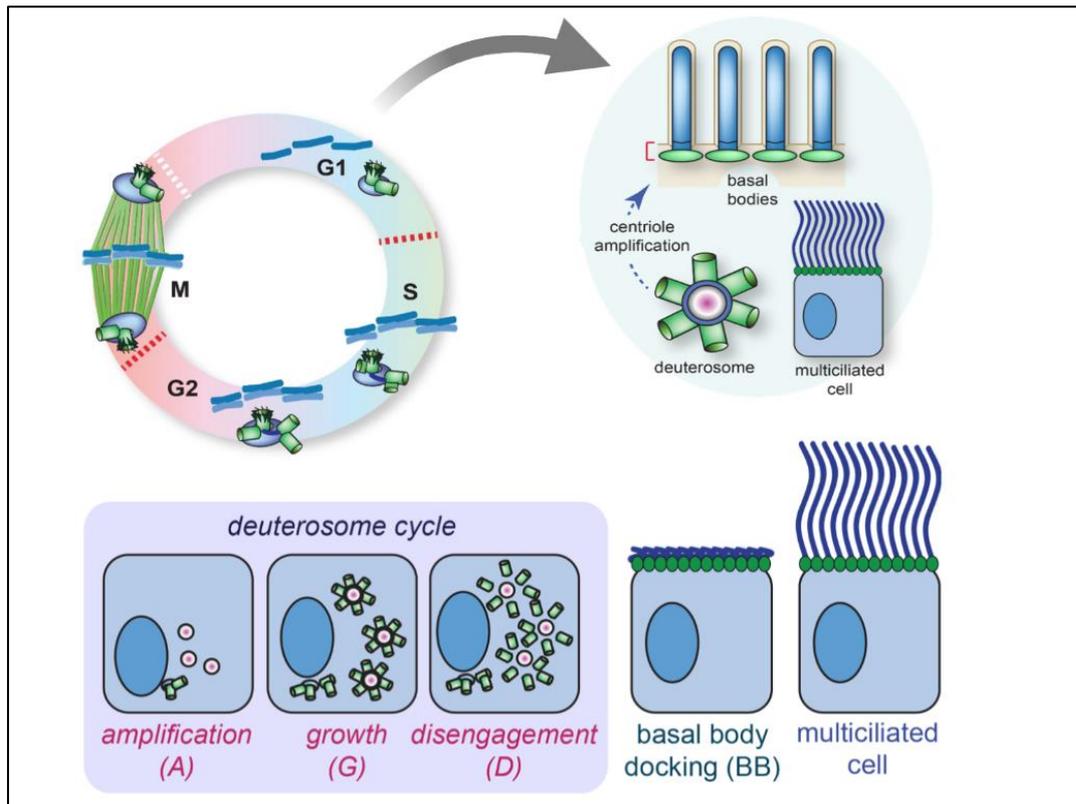


Figure 7 – Multiciliated cell production and the deuterosome dependent pathway. Centrosome amplification in Multiciliated cells occurs outside of the cell cycle. Centriole amplification creates produces hundreds of centrioles which can aggregate around deuterosomes in a ring like structure, in the growth stage. These centrioles disengage and migrate to the apical surface and dock as basal bodies, where a cilium can form.

vi. Deuterosomes and cell cycle regulation

The DD pathway, that uses a ring-shaped, electron dense structure, called the deuterosome, to facilitate centriole expansion of the bulk number of centrioles needed for MCC (Zhao et al., 2013; Klos Dehring et al., 2013; Spassky et al., 2017). Many deuterosomes can form during multiciliogenesis, each nucleating multiple procentrioles, thereby evading the restrictions that regulate the MD centriole duplication pathway to allow rapid BB production.

While the deuterosome was first described many decades ago, only recently has its major structural component, DEUP1, been identified (Zhao et al., 2013; Zhang et al., 2015). *DEUP1* evolved through a duplication of *CEP63*, that plays a key role in centriole duplication in the MD pathway. In addition to DEUP1, *CCDC78*, that localises to the acentriolar sites of centriole biogenesis and *CDC20B*, that localises to the perideuterosomal region, are the only other deuterosome specific proteins identified in both mouse

and frog MCCs (Klos Dehring et al., 2013; Revinski et al., 2018). Other components of the deuterosome are common to the MD pathway, including Pericentrin (PCNT), γ -tubulin, and CEP152, that are all located on the peripheral ring (Boutin et al., 2019). However, there are likely additional proteins that comprise the centre core or outer wall of the deuterosome that remain to be identified.

The prevailing view was that deuterosome synthesis occurred *de novo* and did not require centrosomal centrioles. Detailed live cell imaging of newly synthesized centrioles during cultured brain MCC differentiation showed that deuterosomes were seeded by the daughter centrosomal centriole (Al Jord et al., 2014). However, a handful of studies showed that deuterosomes with multiple centrioles could be produced when both parent centrioles were depleted (Nanjundappa et al. 2019; Zhao et al., 2019; Mercey et al., 2019). Deuterosomes were able to spontaneously synthesize from the pericentriolar material (PCM) in a manner that did not require PLK4, in contrast to earlier reports (Zhao et al., 2019). However recently, genetic ablation of PLK4 in mouse MCCs blocked procentriole assembly and multiciliogenesis without preventing MCC differentiation. (LoMastro et al., 2022). Discrepancies in results may be due to the level of PLK4 knockdown, which is upregulated ~20-fold in differentiating mTECs.

One alternative view that has been put forward is that deuterosomes are created in the PCM, and then only briefly associate with the daughter centrioles to facilitate the loading of procentrioles that occur there (Mercey et al., 2019). Interestingly, E2F4, a key TF in MCC differentiation, was shown to co-localise with various components of centriole amplification in the cytoplasm, including PCM1, and it was proposed that this forms the core of the fibrous granules that have been observed overlapping with or adjacent to deuterosomes. It will be important in future work to fully elucidate the composition of the PCM in MCCs and explore potential species-specific differences.

Strikingly, the deuterosome cycle appears to rely on much of the same cell cycle machinery that plays a role during stepwise MD centriole duplication in proliferating cells, although the precise regulatory details remain unclear. Following cell cycle exit, radial glial progenitors re-express cell cycle markers, including KI67, CDK1, CDK2 and phosphorylated histone H3-Serine10, without performing DNA synthesis or undergoing mitotic division (Al Jord et al., 2017). Through the use of a number of small molecule inhibitors or agonists, key roles in controlling the transitions between centriole templating, growth and disengagement phases were demonstrated for CDK1, CDK2, PLK1, APC/C and CDC20 (Al Jord et al., 2019, 2017).

However, recent work has questioned the necessity of the deuterosome for centriole amplification in MCCs. In mouse tracheal MCCs *in vitro*, depletion of DEUP1 caused a reduction in centriole amplification and CEP63 depletion did not influence centriole numbers in MCCs, due to apparent compensation from DEUP1 (Zhao et al., 2013). Depletion of CEP63 *in vivo*, that reduces MD pathway centriole duplication in neural progenitor cells, also did not have an impact on centriole numbers in MCCs (Marjanović et al., 2015; Mercey et al., 2019b). However, the deletion of DEUP1 *in vivo* revealed that mice were able to

generate MCCs normally in the absence of deuterosomes, indicating they are dispensable for procentriole amplification during multiciliogenesis in mice and frogs (Mercey et al., 2019b). Concomitant depletion of CEP63, to impair the MD pathway, led to a slight reduction in centriole amplification *in vitro*, indicating that there is likely minimal compensation from the MD pathway in the absence of DEUP1 (Mercey et al., 2019b). This finding remains to be extended to MCCs in all tissues, but it opens up the field to a number of exciting possibilities. They speculate that centriole synthesis occurs within a cloud of fibrogranular material and PCM, and that deuterosomes function to relieve the parental centrioles of the high numbers of procentrioles that form along their length. Supporting this possibility, flatworms and some species of ray-finned fish (such as zebrafish) are capable of producing MCCs and do not encode DEUP1 or generate deuterosomes (Kramer-Zucker et al., 2005; Kemp 1996; Azimzadeh 2012). This establishes that another route for *de novo* procentriole generation in MCCs can function independently of both the MD and DD pathways and provide compensation for their loss. It was proposed that this pathway may be equivalent to pericentriolar satellites that occur in non-MCCs given that they are scaffolded by PCM1 that is enriched in fibrogranular material (Mercey et al., 2019b; Ito et al., 2019).

b. Transcriptional activation of MCC differentiation by GEMC1 and MCIDAS

The process of MCC differentiation, or multiciliogenesis, requires the activation of a transcriptional program that specifies cell fate and allows the rapid generation of hundreds of centrioles which are amplified and become BBs to template motile cilia (Yan et al., 2016). This unique pathway has only just begun to be elucidated but has an important role in human development.

i. Geminin family proteins

Geminin is a protein with dual functions in cell cycle and differentiation. Geminin was first discovered as a replication control protein with a central coiled-coil (CC) domain that mediates its dimerization, reflecting its name (Balestrini et al., 2010). Its primary function is to bind to Cdt1, preventing origin firing and re-replication in the same cell cycle (McGarry and Kirschner, 1998; Wohlschlegel et al., 2000). Its expression oscillates during the cell cycle with high levels of the protein being detected in S, G2 and the first part of the M phase (Xouri et al., 2004).

Geminin (encoded by *GMNN*), GEMC1 (Geminin coiled-coil-domain containing protein 1, encoded by *GMNC*) and MCIDAS (Multicilin (MCI) and IDAS, encoded by *MCIDAS*) were identified due to the presence of a similar CC domain and all 3 proteins can homo or heterodimerize (Figure 8). This feature appears to be functionally relevant, as the CC domains are required for some of their key functions in all cases, although the physiological significance of heterodimerization remains unclear (Terré et al., 2016; Lalioti et al., 2019). Studies in zebrafish and mice identified GEMC1 as a critical upstream mediator of the MCC differentiation program that is activated upon Notch inhibition (Terré et al., 2016; Arbi et al., 2016; Kyrousi et al., 2015; Zhou et al., 2015; Lu et al., 2019). GEMC1 deficient mice lack MCCs in every tissue

where they are normally present, leading to infertility and penetrant hydrocephalus, a pathological condition characterized by the enlargement of the brain ventricles resulting from the accumulation of cerebrospinal fluid (CSF) (Terré et al., 2016). While MCCs are absent from the murine airway in GEMC1 deficient mice, no respiratory defects have been reported, potentially due to sterile housing conditions, although significant discrepancies remain regarding the lifespan of GEMC1 deficient mice derived in different colonies where they range from 9 days to over 2 years (Terré et al., 2016; Arbi et al., 2016).

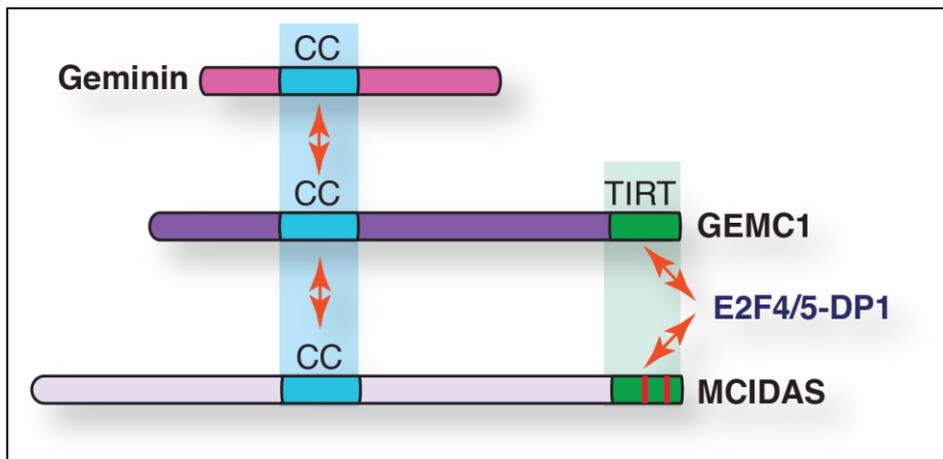


Figure 8 – Geminin Family Proteins. Three members of the Geminin family; Geminin, GEMC1 (Geminin coiled-coil-domain containing protein 1 and MCIDAS (Multicilin (MCI) and IDAS). Each are characterised by a coiled coil (CC) domain and a TIRT domain which is required for E2F4/5-DP1 binding and transcriptional activity

The induction of *Gemc1* and *Mcidas* in the developing murine brain is temporally distinct and MCIDAS deficient mice develop morphologically identifiable MCCs that express early transcription factors, including p73 and FOXJ1, but fail to amplify centrioles or generate cilia (Kyrousi et al., 2015; Lu et al., 2019). Air liquid interface (ALI) cultures from MCIDAS deficient mice showed no increase in the expression of genes implicated in centriole amplification, indicating that GEMC1 is not sufficient to activate multiciliogenesis when under its normal physiological regulation (Lu et al., 2019). Thus, a two-step process has been proposed; GEMC1 activates *MCIDAS* and other key transcription factors to promote MCC specification and MCIDAS subsequently activates the expression of genes required for Multiciliogenesis (Lu et al., 2019). The overexpression of *MCIDAS* can trigger MCC differentiation in frog skin and mouse cells, indicating that *MCIDAS* is one of the most crucial targets of GEMC1 (Kim et al., 2018; Stubbs et al., 2012). Consistent with this, overexpression of GEMC1 in frog embryos or murine ALI cultures also generates supernumerary MCCs in a manner that requires MCIDAS in ALI cultures (Zhou et al., 2015; Lu et al., 2019).

ii. MCC specification: Notch

The full effects of Notch inhibition on transcription are not clearly understood at the early steps of MCC differentiation, but fate decisions following Notch inhibition are controlled by the interplay between the Geminin family proteins (Figure 9)

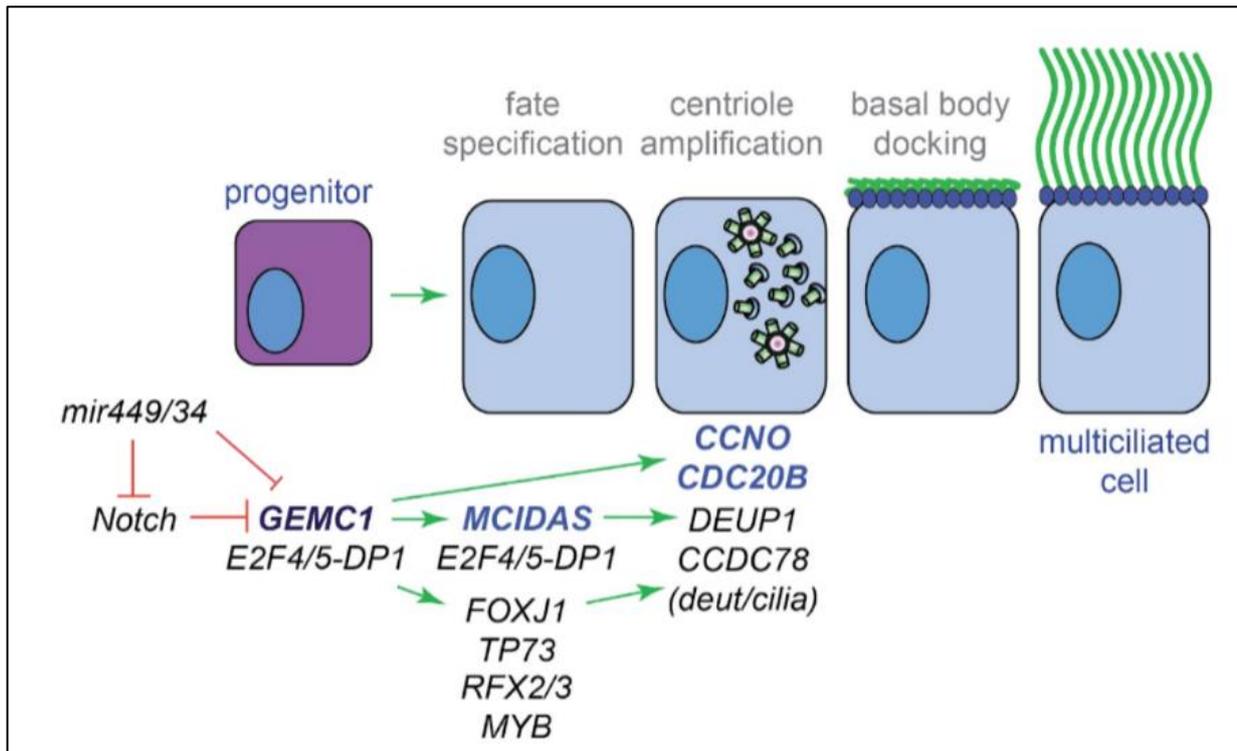


Figure 9 – Transcriptional pathway of GEMC1/MCIDAS mediated Multiciliogenesis. General depiction of the Multiciliogenesis pathway in the mammalian airway epithelium. NOTCH inhibition activates GEMC1 through an unknown mechanism. GEMC1 interacts with E2F5-DP1 through its C-terminal domain, to activate the MCC differentiation program. p73 act at an early step, prior to the induction of MCIDAS. MCIDAS induction required MYB, that also acts upstream of FOXJ1. Both GEMC1 and MCIDAS activate additional TFs, including FOXJ1, RFX2, RFX3. CCNO is induced by GEMC1 while MCIDAS induces DEUP1 amongst others. Centrosome amplification in Multiciliated cells occurs outside of the cell cycle and canonical and MCC specific cell cycle regulators may play roles in controlling the steps of centriole amplification and maturation to form MCCs. Diagram is not inclusive to all factors that are important in this pathway.

The inhibition of Notch signalling is established as a consistent early event in MCC differentiation from the study of frog skin, zebrafish pronephros and murine ependymal, fallopian tube (oviduct) and airway epithelia (Ma et al., 2017; Guseh et al., 2009; Pardo-Saganta et al., 2015; Tsao et al., 2009; Kessler et al.,

2015; Marcet et al., 2011). The precise details of Notch regulation remain unclear in all cases, but the *Mir-34/449* family of miRNAs has been implicated in Notch inhibition in frogs, zebrafish and mice and this miRNA family plays redundant roles in MCC formation in several tissues, including the brain, airway and male germline (Yuan et al., 2019; Marcet et al., 2011; Lize et al., 2010; Song et al., 2014; Wildung et al., 2019; Chevalier et al., 2015). The *Mir449* family has been linked more directly to the regulation of GEMC1 and MCIDAS as important contributors to MCC development (Arbi et al., 2016; Wildung et al., 2019; Loukas et al., 2021). Recently, In the murine airway, progenitor cells give rise to secretory (Clara) or MCC lineages in a Notch dependent manner (Kiyokawa et al., 2020). Genetic or pharmacological inhibition of Notch causes the trans-differentiation of secretory cells into MCCs, demonstrating a central role for Notch inhibition in initiating the MCC differentiation program in the airway (Pardo-Saganta et al., 2015).

iii. E2F4 and E2F5

As GEMC1 and MCIDAS lack clear DNA binding domains, it remained an open question as to how they could activate transcription. Work in frogs first showed that a C-terminal domain of MCIDAS, absent in Geminin, dubbed the TIRT domain due to a repetitive amino acid motif sequence, interacted with the E2F4 and E2F5 transcription factors heterodimerized with DP-1 (Ma et al., 2014). This connected the activity of MCIDAS to previous and subsequent work that established a key role for E2F4 and E2F5 in murine brain, airway and germline MCC development (Ma et al., 2014; Hsu et al., 2016; Lindeman et al., 1998; Danielian et al., 2007). While it lacks the specific TIRT amino-acid repetition, the C-terminus of GEMC1 contains a homologous “TIRT” domain that is also required for E2F4/5-DP1 binding and transcriptional activity (Terre et al., 2016; Arbi et al., 2016). The TIRT domains of GEMC1 and MCIDAS may not be functionally identical, immunoprecipitation experiments in mouse and human showed a higher affinity of GEMC1 for E2F5 than E2F4, and this specificity could be eliminated by replacing the TIRT domain of GEMC1 with that of MCIDAS. In contrast, MCIDAS was shown to have similar affinity to both E2F4 and E2F5, potentially enabling more efficient activation of the genes necessary for centriole amplification (Lu et al., 2019). One possibility that is consistent with current data is that GEMC1 displaces pocket protein inhibitors of E2F5, such as RB, p107 or p130, but does not strongly activate the MCC transcriptional program until it is reinforced by *MCIDAS* expression and its interactions with both E2F4 and E2F5. While more detailed biochemical and structural studies remain to be performed, the emerging picture is that GEMC1 and MCIDAS play sequential roles that utilize specific E2F4/5-DP1 interactions influenced by differences in their C-terminal domains.

While both E2F4 and E2F5 have been clearly linked to MCC generation, discerning their specific roles has been challenging in mice due to their essentiality and influence on the development of many tissues. E2F4 and E2F5 have been typically characterised as repressive E2Fs, in contrast to the activating E2Fs, E2F1-3, and bind to the regulatory regions of hundreds of genes to repress cell cycle genes during G0 and G1 in conjunction with the RB family proteins and interactions with the MuvB/DREAM complex (Fischer et al., 2017). Recent work in zebrafish showed that some tissue specificity exists with regards to their role in MCC

generation, but that this reflects the relative expression levels of *E2f4* or *E2f5* in those tissues, rather than specific functions of either factor, consistent with work in frogs that suggested redundancy between E2F4 and E2F5 (Ma et al., 2014; Chong et al., 2018).

In overexpression experiments in frogs or human cells, Geminin acted as a potent inhibitor of transcriptional activation by GEMC1 and MCIDAS through CC-mediated interactions and the formation of a ternary complex with the E2F4 or E2F5-DP1 transcription factors (Terre et al., 2016; Arbi et al., 2016; Ma et al., 2014). It was proposed that Geminin may prevent the activation of the multiciliogenesis program until dividing cells exit the cell cycle, as the consequent centriole amplification would cause mitotic spindle defects (Ma et al., 2014).

A crucial function of GEMC1 and MCIDAS is to activate numerous downstream TFs. This expanding list includes FOXJ1, FOXN4, RFX2, RFX3, MYB and P73, among others. Each of these TFs has been demonstrated to play critical, and in some cases cooperative, roles in enabling the gene expression of numerous proteins involved in MCC differentiation.

iv. Cyclin O (CCNO) and CDC20B

Two targets of the GEMC1-MCIDAS program, the poorly understood cell cycle proteins CDC20B and CCNO, appear to be required for the repurposing of the mitotic oscillator for centriole amplification, and have their highest expression levels during the earliest phase of deuterosome formation (Revinski et al., 2018; Funk et al., 2015; Núñez-Ollé et al., 2017).

CDC20B is a homologue of CDC20, which has a well-established role in the APC/C to drive progression of the cell through mitotic checkpoints. Furthermore, CDC20B was shown to interact with PLK1, and coordinate centriole disengagement of centrioles from the deuterosomes in mouse and *Xenopus* MCCs, although its functions have been proposed to be independent of APC/C interactions (Revinski et al., 2018).

CCNO is an atypical cyclin that is encoded from the same genomic locus as the MCC genes CDC20B, *Mir449a/b/c* and MCIDAS, and can interact with CDK1 and maybe CDK2 in some cell types (Roig et al., 2009). Its precise kinase partner(s) in MCCs or their targets have not been clearly elucidated.

CCNO mutant mice exhibit large, malformed deuterosomes and produce fewer BBs, suggesting that CCNO is required for deuterosome stability (Funk et al., 2015). MCCs generated in ALI cultures derived from CCNO knockout mice also display significant transcriptional dysregulation, suggesting that CCNO functions to restrain the activation of key MCC genes, including MCIDAS (Núñez-Ollé et al., 2017) but precisely how CCNO regulates deuterosome stability, transcription or how it influences CDK activity in MCCs remains unclear.

Reassigning the existing cell cycle regulation machinery appears to provide an elegant solution to regulating centriole amplification in the context of MCCs. Future work will be needed to understand its regulation and the specific roles of factors such as CDC20B and CCNO. It is certainly an interesting possibility that

CCNO and CDC20B could be specifically expressed in MCCs to help regulate CDK activity and promote differentiation.

v. p73

The TF p73 (encoded by *TP73* and *Trp73* in humans or mice) is a member of the p53 family that also includes p63, a marker of basal progenitor cells that give rise to MCCs in the airway (Pardo-Saganta et al., 2015). Given its relationship to p53, one of the most well studied tumor suppressors, p73 was extensively studied in mice and was shown to cause a number of developmental phenotypes, most of which have been recently linked to defects in MCC differentiation (Marques et al., 2019). The TP73 gene generates 2 major isoform groups via 2 promoters; the activating (Tap73) or an N-terminally truncated (DNp73) form lacking the transactivating domain. Reduced numbers of MCCs were observed in the airway, oviducts and efferent ducts of mice lacking both isoforms, or only Tap73, and CHIP experiments linked p73 directly to genes involved in multiciliogenesis, such as *FoxJ1*, *Rfx2*, or *Rfx3* (Nemajerova et al., 2016; Marshall et al., 2016). P73 forms a ternary complex with GEMC1 and E2F5 that is stabilised by both the CC and TIRT domains of GEMC1, and this complex can activate the *TP73* promoter (Lalioi et al., 2019; Nemajerova et al., 2016; Terré et al., 2019). P73 is induced by both GEMC1 and MCIDAS through their respective E2F4/5-DP1-containing complexes, although it remains unclear if MCIDAS also interacts with p73. P73 expression has been reported in p63+ basal cells and Radial Glial Cells (RGCs), that act as MCC progenitors in the airway and brain, respectively, suggesting an early role for p73 in MCC fate specification independent of GEMC1 or MCIDAS, although this remains controversial (Marques et al., 2019; Nemajerova et al., 2016; Marshall et al., 2016; Fujitani et al., 2017). The requirement for p73 appears to also vary depending on the isoform, tissue and cell type. In contrast to the airway, where Tap73 deletion in mice impaired MCC formation, these mice did not exhibit hydrocephalus or show impaired MCC formation in the brain due to compensation from the *miR-449a-c* cluster (Wildung et al., 2019). Combined deletion of *Tap73* and *miR-449a-c* impaired MCC generation in the choroid plexus (CP) and deletion of both TP73 isoforms impaired MCC generation in ependymal cells (Wildung et al., 2019; Gonzalez-Cano et al., 2016). Therefore, MCC fate may be subjected to tissue-specific feedback modulation, warranting further comparison of the transcriptional regulation and signalling pathways involved in different tissues.

vi. FOXJ1

FOXJ1 was the first transcription factor shown to be required for MCC differentiation and is frequently used as a marker of MCCs in all tissues (Blatt et al., 1999; Brody et al., 2000; You et al., 2003; Okada et al., 2004). FOXJ1 is now a well-established target of p73 and GEMC1, that likely work together to activate its expression at early stages of Multiciliogenesis (Terrer et al., 2016; Arbi et al., 2016; Lu et al., 2019; Nemajerova et al., 2016; Marshall et al., 2015). FOXJ1 is also regulated by MCIDAS, as the absence of MCIDAS activity strongly reduced FOXJ1 expression in frog skin and human airway cells, but airway MCCs

in MCIDAS deficient mice were FOXJ1 positive (Lu et al., 2019; Stubbs et al., 2012; Boon et al., 2014) FOXJ1 promotes MCC differentiation by regulating a cohort of genes involved in the production, assembly, transport, and docking of the inner and outer dynein arms, radial spokes and the central pair as well as genes that encode intraflagellar transport (IFT) proteins (You et al., 2004; Stubbs et al., 2008; Newton et al., 2012; Jacquet et al., 2009; Didon et al., 2013). Studies in the human airway demonstrated that the RFX3 TF functions as a transcriptional coactivator of FOXJ1, helping to induce the expression of cilia genes involved in differentiation towards the MCC lineage (Didon et al., 2013). Similarly, recent work demonstrated that FOXJ1 preferentially binds enhancers and is stabilized at promoters of cilia genes through cooperative interactions with the TF RFX2 (Quigley et al., 2017).

The stability of the MCC lineage is reportedly dependent on a constant FOXJ1 protein expression level in order to prevent cellular de-differentiation back into a glial-like morphology in mouse ependymal MCCs (Abdi et al., 2018). Recently, human patients with FOXJ1 mutations were shown to have defects in motile cilia, left right patterning defects and impaired basal body docking in MCCs (Wallmeier et al., 2019).

vii. SWI/SNF complex

The large chromatin remodeller - SWI/SNF complex, plays an integral part in the transcriptional response of many genes (See 1,b,iii and Figure 5).

Previously, a disordered, acidic region of Geminin, was shown to facilitate interactions with BRG1 (SMARCA4) an ATPase component of the SWI/SNF complex, that is known to interact with E2F4/5, and controls the expression of a number of TFs (Karamitros et al., 2015; Sanka et al., 2016; Seo et al., 2005; Nagl et al., 2007).

It remains to be determined if Geminin acts only by impairing GEMC1 or MCIDAS homodimerization and transcriptional activation through E2F4/5, or has a more complex role involving modulation of SWI/SNF. Likewise, given the similar homology between the geminin family proteins and their intricate functional differences, exploring the interaction of GEMC1 or MCIDAS with the SWI/SNF complexes in the regulation of multiciliogenesis remains an exciting prospect.

c. Consequences of MCC dysfunction

i. Ciliopathies

Defective cilia can be the cause of a diverse range of syndromes termed Ciliopathies. There are currently at least 190 known ciliopathy disease genes causing defects of the primary cilia that cause different phenotypes (Whewey et al., 2019). Some of the most severe-lethal ciliopathies disorders include Joubert syndrome, Bardet–Biedl syndrome, orofacial digital syndrome (OFD) and Meckel syndrome, but also more common

pathologies like polycystic kidney diseases (PKDs). Ciliopathies have been extensively reviewed elsewhere (Goetz et al., 2010; Hildebrandt et al., 2011)

While present only in a few tissues, defective MCCs contribute to congenital disorders, such as primary ciliary dyskinesia (PCD) and Reduced generation of multiple motile cilia (RGMC) (Boon et al., 2014; Wallmeier et al., 2014; Amiravet et al., 2016). Thus far, mutations in 2 genes have been implicated in RGMC: *MCIDAS* and *CCNO*. Both genes are located adjacent to each other on Chromosome 5 (5q11.2) that also contains *CDC20B* that encodes both a protein and the *Mir449a/b/c* genes. RGMC patients present clinically with hydrocephalus, recurrent airway infections and bronchiectasis (Figure 10).

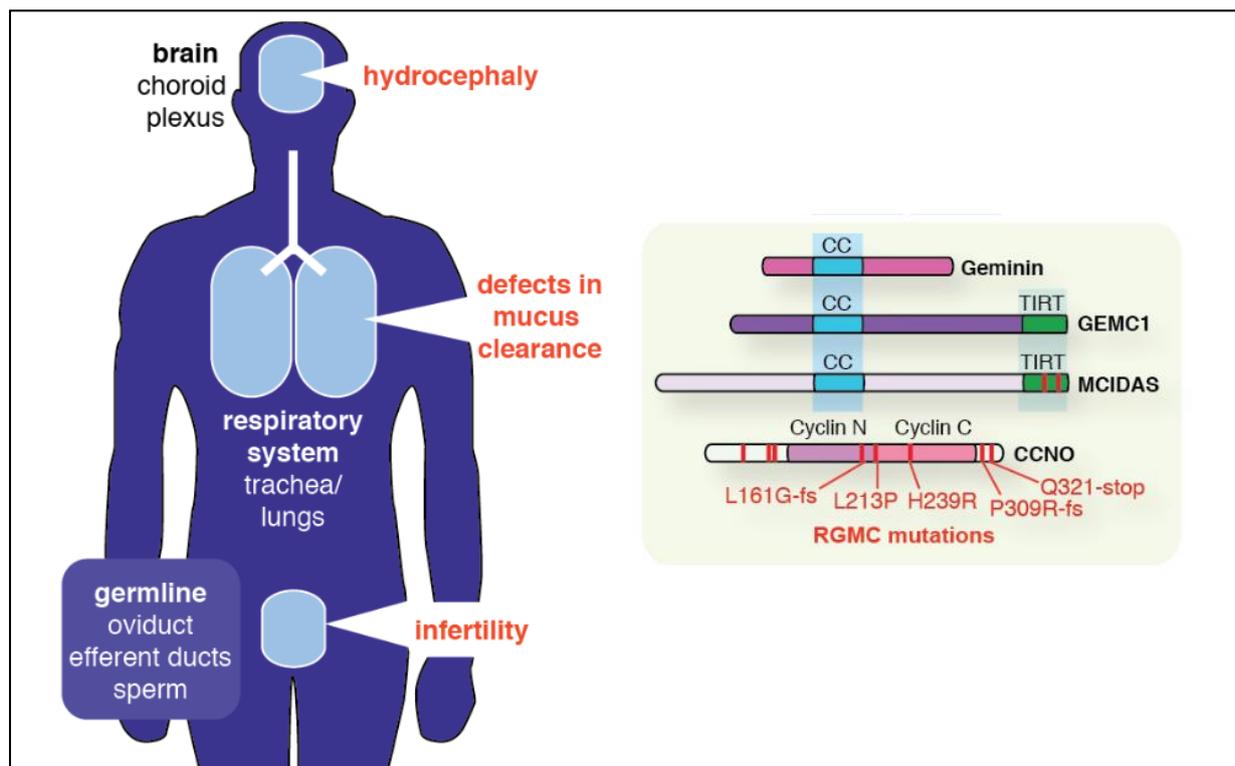


Figure 10 – RGMC clinical perspective. Reduced generation of multiple motile cilia (RGMC) patients present clinically with hydrocephalus, recurrent airway infections and bronchiectasis, as these are the areas where multiciliated cells are present in humans. Diagram on the right depicts current known mutations that are implicated in RGMC, two in *MCIDAS* and several in *CCNO*.

Due to the ages of the patients, little data regarding fertility has been reported, although some female patients have presented with infertility (Boon et al., 2014; Wallmeier et al., 2014; Amiravet et al., 2016). While RGMC represents an extreme case, it is possible that less severe mutations or tissue specific defects in the function or expression of key MCC factors may result in more subtle pathologies, such as subfertility, scoliosis or normal pressure hydrocephalus (Núñez-Ollé et al., 2018; Zhang et al., 2018).

Consistent with this possibility, a mutation in *GEMC1* was recently identified in a patient with congenital hydrocephalus, but respiratory defects or fertility issues were not reported (Laloti et al., 2019). *MCIDAS*

mutations that have been identified in RGMC are located in the TIRT domain of MCIDAS and interfere with binding of the E2F4/5 transcription factors, highlighting again the central importance of MCIDAS in the MCC transcriptional pathway (Boon et al., 2014).

A new report identified the NEK10 kinase in familial bronchiectasis and demonstrated its expression in airway MCCs (Chivukula et al., 2020). While these patients exhibited respiratory distress, similar to that observed in RGMC, they generated airway MCCs with subtle morphological defects and shorter cilia that were incapable of mucociliary clearance. Proteomic analysis of NEK10 deficient MCCs revealed widespread effects on the phosphorylation of many proteins required for motile cilia function, indicating that NEK10 may be a key regulator of many aspects of MCC function.

ii. Cancer

GEMC1 and MCIDAS appear to be expressed only in MCCs in normal tissues. Aberrant expression of key MCC specific genes can also be observed in many cancers and some evidence suggests they may play functional roles and become co-opted to support tumourigenesis (Gao et al., 2013; Cerami et al., 2012; Li et al., 2018).

The irregular expression or depletion of GEMC1 and MCIDAS in cancer cell lines has been shown to impact DNA replication and could potentially modulate transcription factors such as E2F4, p73 and FOXJ1 that have been implicated in cancer development and could provoke centriole amplification and mitotic defects (Caillat et al., 2013; Pefani et al., 2011; Balestrini et al., 2010; Li et al., 2022).

The potential of CCNO and CDC20B to impact cell cycle regulation or apoptosis is also potentially relevant to their expression in human cancers (Li et al., 2018; Roset et al. 2009), *CDC20* is an oncogene and has been proposed as a target in cancer. Overexpression of *CDC20* was observed in a variety of human tumours (Wang 2015). Moreover, higher expression of *CDC20* is associated with clinicopathological parameters in various types of human cancers such as pancreatic or breast cancer (Taniguchi et al., 2008; Karra et al., 2014).

There are some clear parallels and overlaps of the cell cycle machinery utilised in MCC generation and some tumour formation. Therefore, further understanding the details of the MCC transcriptional program and its regulation could provide potential insights into cancer pathogenesis.

3. Ligand activated transcription factors in prostate cancer

a. Androgen and Glucocorticoid receptors

The Androgen Receptor (AR) is a member of the class I nuclear receptor TF family, which includes the steroid receptors (SR) glucocorticoid receptor (GR), mineralocorticoid receptor (MR), estrogen receptor (ER), and progesterone receptor (PR).

The GR and the AR fall into the oxosteroid subfamily (Bledsoe et al., 2002) and are master regulators of several metabolic, homeostatic and differentiation processes. Both respond to binding of their respective hormone, causing activation and consequential gene specific transcription to fulfil their physiological roles. Glucocorticoids stimulate the GR to instigate an array of fundamental biological processes in the human body, including cell proliferation, metabolism, development, inflammation, and immune responses (Ramamoorthy et al., 2016; Vandevyver et al., 2014; Weikum et al., 2017)

Androgens and the AR are crucial for the development, differentiation and function of male reproductive organs (Bardin et al., 1998; Gao et al., 2005) AR is not only found in the prostate where it promotes the survival and growth of prostate cells (PC), but is also expressed in range of tissues and has unique functions in bone, muscle, cardiovascular, immune, neural and haemopoietic systems (Rana et al., 2014). AR is also expressed in cells of the efferent ducts that generate MCCs to facilitate movement of male gametes (Zhou et al., 2002).

b. AR Structure and signaling pathways

The AR is a 110 kDa phospho-protein encompassing approximately 919 amino acids (aa), encoded by the *AR* gene located on chromosome X at Xq11–12. Like many nuclear receptors, the *AR* gene is comprised of three distinct functional domains: an N-terminal domain (NTD), a 2-zinc finger DNA-binding domain (DBD) and a ligand-binding domain (LBD) (Figure 11) (Heinlein et al., 2004).

The NTD of the AR (spanning from amino acids 1–559) is the least conserved domain among the class I members, being less than 15–20% homologous (Simental et al., 1991; Jenster et al., 1995). The NTD is largely unstructured and intrinsically disordered, although it contains several motifs of helical propensity (Reid et al., 2002; Kumar et al., 2004; De Mol et al., 2016). It varies in the number of glutamine repeats (polyglutamine tract or polyQ) with shorter repeats equating to higher transcriptional activity (Beilin et al., 2000).

The NTD provides the majority of the transcriptional activity of the AR (Claessens et al., 2008), containing the activation function 1 (AF-1) domain. The AF-1 can be functionally subdivided into two transactivation units; Tau-1 (between amino acids 100–370) and Tau-5 (between amino acids 360–528).

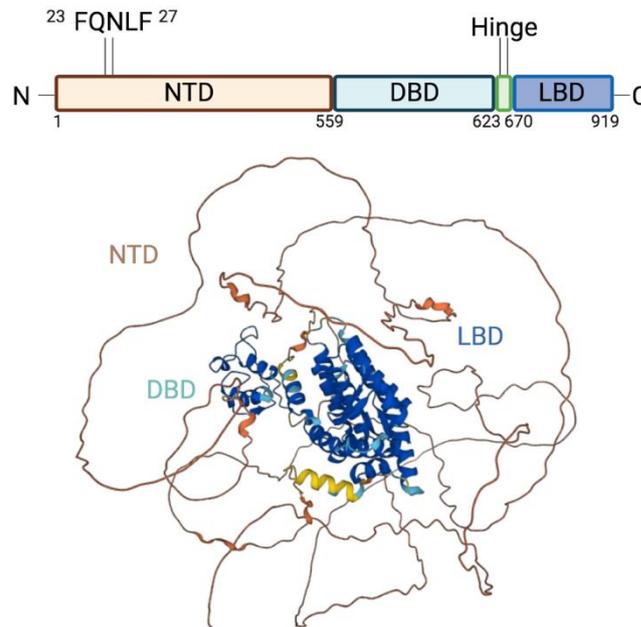


Figure 11 – Androgen Receptor (AR). The AR gene is comprised of three distinct functional domains: an N-terminal domain (NTD), a DNA-binding domain (DBD) and a ligand-binding domain (LBD). Predicted structure of the AR protein with AlphaFold (Deepmind), with high confidence structural prediction in blue to low confidence in yellow.

The DBD is the highly conserved domain among SRs, it has two zinc finger motifs that are responsible for DNA recognition and dimerization (Dalal et al., 2014). The DBD binds as a dimer to the androgen response elements (AREs) on the DNA. The hinge region serves as a short flexible linker between the DBD and the LBD and contains the nuclear localisation signal (NLS) for translocation (Claessens et al., 2008).

The LBD of the AR is responsible for ligand binding, it is only moderately conserved among the receptors, and contains AF-2 which is important for interaction with co-regulators and co-chaperones as well as a dimerisation and activation (Nadal et al., 2017; Estebanez-Perpina et al., 2007; Dubbink et al., 2004). Inactivate AR is normally located in the cytoplasm, bound to chaperones Hsp40 and Hsp70 (Eftekharzadeh et al., 2019). Upon binding to the steroid dihydrotestosterone (DHT), a conformational change occurs, dislodging the chaperones and exposing the AF-2 surface (Claessens et al., 2008).

The ${}_{23}\text{FQNLF}_{27}$ motif within the NTD interacts as an α -helix with its AR C-terminal LBD, creating the so-called N/C interaction (He 2004). AR forms a dimer and then translocates to the nucleus (Nadal et al., 2017). Finally, upon DNA binding, the AR dimer forms a complex with coactivator and coregulatory

proteins at the AF-1 and AF-2 regions and binds to AREs of promoter and the enhancer targeted genes for transcription (Takayama et al., 2008.) (Figure 12).

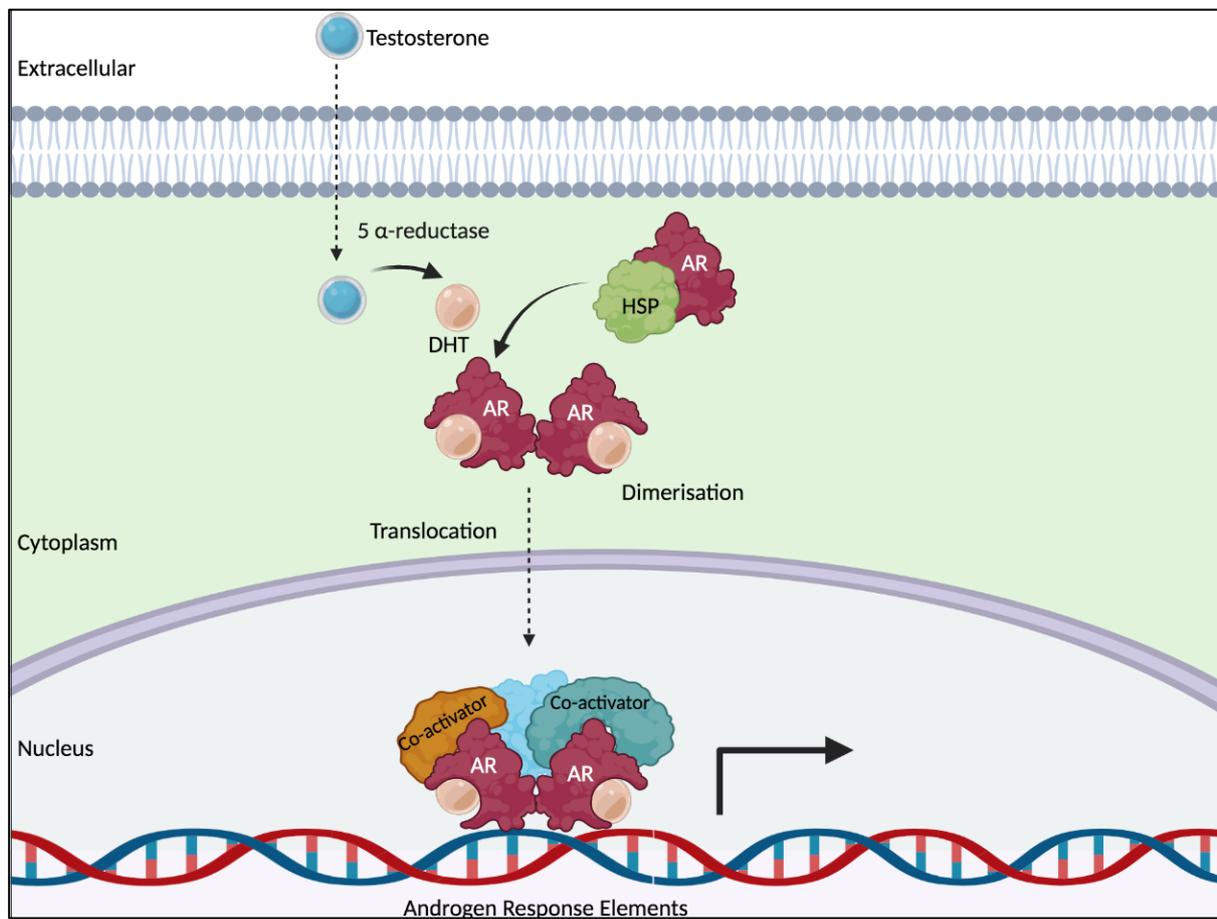


Figure 12 – Androgen Receptor activation. Androgen Receptor (AR) is normally located in the cytoplasm, bound to Heat Shock Proteins (HSP). Testosterone is shuttled into the cell where it is converted into dihydrotestosterone (DHT). This causes a conformation change in AR, which dimerises and translocates into the nucleus. The AR dimer forms complexes with coactivator and coregulatory proteins and binds to AREs where it can activate targeted gene expression.

c. Prostate cancer (PC)

This AR signalling pathway is exploited in the development and growth of Prostate Cancer (PC). With ~1.2 million new cases and nearly 360,000 deaths annually, PC is the second most common type of cancer and the fifth most common cause of cancer-related death in men worldwide (Bray et al., 2018). PC is unsurprisingly initially localised to the prostate. Prostatectomy and radiotherapy are often the first treatment options, but have limited long term success. From a clinical viewpoint, patients with metastasis-

free PC have a 5-year survival rate of 98.9%, whilst patients with metastatic PC on initial diagnoses have less than 30% survival rate (Siegel et al., 2018; Wu et al., 2014).

As such, PC remains a major public health concern across the globe with a poor prognosis and no cure. One factor contributing to its persistence in the male population is the heterogeneity of the disease (Haffner et al., 2021). The clinical presentation of PC can range from anything between a localised asymptomatic onset to a rapidly progressing lethal metastatic cancer. Furthermore, Primary PCs are often multifocal, having morphologically and genetically distinct tumour foci leading to therapeutic challenges (Haffner et al., 2021).

Upon diagnosis PC is classified in one of two groups. Firstly, Androgen-sensitive tumours (Around 80–90%), which are highly responsive and likely controlled with androgen deprivation therapy (ADT). The second being Androgen-insensitive tumours, which generally represent later-stage, more aggressive and have higher metastatic potential (Sharifi et al., 2005). Unfortunately, resistance to AR blockade and progression to metastatic castration-resistant prostate cancer (mCRPC) eventually emerges in almost all patients (Watson et al., 2015).

i. AR activating mutations

Answering the questions of how and why these tumours originate is key to targeted therapeutic strategies and has been the focus of studies for decades. It is now well understood that AR becomes overactivated in this pathway and is a key driver in PC formation (Huggins et al., 2002).

Within the AR gene about 1% of primary PC patients carry mutations or amplifications and approximately 60% in metastatic tumours (Abeshouse et al., 2015; Grasso et al., 2012). Approximately 160 AR mutations have been found in PC, almost all being single-nucleotide substitutions due to somatic rather than germline mutations. It is estimated that the majority of mutations occur in the LBD (45%), although one-third occur in the transactivating NTD (Gottlieb et al., 2012). The biological significance of these mutations results in a much broader ligand specificity, providing an invaluable advantage for cancer progression.

The most common example is the point mutation T877A, with a prevalence of about 30% in mCRPC patients (Gottlieb et al., 2012; Aurilio et al., 2020). The mutation enables the activation of AR by other adrenal androgens such as progesterone, DHEA and androstenediol (Veldscholte et al., 1990; Tan et al., 1997; Miyamoto et al., 1998). Similarly, AR mutations like L701H permit the AR to be activated by glucocorticoids or progesterone and as such also confer ADT resistance (Buttiglierio et al., 2015).

Alternatively, a distinct set of mutations, such as H874Y or W435L, cause enhanced binding of AR with its co-regulators and activators, increasing AR mediated transcription (Barbieri et al., 2012; Steinkamp et al., 2009). The AR pathway is a source of genomic alteration in PC, with translocations involving androgen-regulated promoter *TMPRSS2* to the TF ERG (*TMPRSS2:ERG*) gene fusion present in ~50% of localised PCs (Tomlins et al. 2009).

Other common genetic aberrations not only target the AR pathways but also PI3K–PTEN, WNT, DNA repair and components of the cell cycle in nearly all CRPCs and a high portion of primary PCs (Taylor et al. 2010; Abeshouse et al., 2015; Robinson et al. 2015).

ii. AR variants

In addition to *AR* amplification and point mutations, PC tissues exhibits a diverse array of structural *AR* alterations which facilitate resistance to ADT. It is well established that the *AR* undergoes many missplicing events to generate variants in PC (Bryce et al., 2016). To date, over 20 AR variants (AR-VARs) have been reported. The majority of them lack some of the C-terminal domain, including the LBD, but retain their ability to bind to DNA and localise to the nucleus, causing them to be constitutively active (Hu et al., 2017). AR-VARs can be detected in benign prostate tissue, hormone-naïve and CRPC, with the most frequent and highest expression detected in the CRPC sample (Guo et al., 2009; Hu 2009; Hörnberg et al., 2011, Qu et al., 2015).

AR-VAR7 (also known as AR-V7) is the most well examined variant, as it is endogenously detected at the protein level and the most consistently expressed variant in recurrent PC cell lines and CRPC specimens Luo, 2016; Haile 2016). Moreover, AR-VAR7 expression is specific to CRPC (Sharp et al., 2019) and may drive a progressive phenotype when canonical AR signalling is suppressed (Hu et al., 2012).

AR-VAR7 is truncated at the end of exon 3 and contains 16 amino acids from cryptic exon 3. AR-VAR7 stays active in the nucleus under androgen-depleted conditions, where it regulates both AR-regulated genes and a unique set of AR-independent genes including, gene sets related to cell cycle progression, such as the *UBE2C* gene (Hu et al., 2009). This suggests AR-VAR7 has an overlapping, but distinct role compared to full-length AR in prostate cancer cells (Dehm et al., 2011).

Another common constitutively active AR variant termed ARv567es, originates from exon skipping of exons 5–7, coding for a portion of the LBD (Aurilio et al., 2020). Its expression has been confirmed in multiple studies of PC and CRPC and is another proposed method of AR variant-dependent mechanism of resistance (Hörnberg et al., 2011; Liu et al., 2013, Watson et al., 2010). Expression of AR-V567es and/or AR-VAR7 was associated with poor survival (Hörnberg et al., 2011). Particularly, AR-VAR7 overexpression is linked with increased risk of remission after radical prostatectomy in hormone-naïve PC patients (Bono, et al., 2012; Smith et al., 2018).

Other variants including AR-V1 and AR-V9 are found to be conditionally active, depending on cellular context (Hu, 2011; Kohli 2017). In contrary to the hormone induced AR translocation pathway, the mechanisms controlling AR variant entry to the nucleus are less well known. For instance, some constitutively active AR variants, like AR-V567es, contain a NLS, whereas others like AR-V3, AR-VAR7, AR-VAR9 do not (Guo et al., 2009; Hu et al., 2009; Sun et al., 2010; Dehm et al., 2008; Hu et al., 2011). It has been hypothesised that within the unique COOH-terminal extension exists a NLS-like signal which enables the NLS-negative AR variants to enter the nucleus, or potentially dimerise with the full length AR-WT (Hu et al., 2011).

d. AR transcriptional cofactors and condensate formation

The coordination of TFs and its co-activators to assemble at precise elements of the genome to carry out activation of selected gene programmes is both complex and reversible. In conjunction with RNAP this transcriptional complex of machinery quickly assembles and disassembles at promoters and enhancers at specific loci at a timely and tightly controlled manner.

Therefore, unlike membrane-bound organelles, this highly dynamic system can rapidly respond to cellular environmental perturbations (Zhao et al., 2020). Additionally, the compartmentalisation of functionally related proteins enables enhanced capability of encountering its unique interacting partners in the cell. Moreover, it creates a reserve of proteins that can be resourced when required at separate condensates (Sabari et al., 2020). Emerging research suggests that it is the process of phase separation that synchronises these transcriptional condensates to facilitate its function, which are loosely held together by multivalence and weak low affinity-interactions (Hnisz et al., 2017).

Recently, biomolecular condensates were observed to form *in vitro* and *in vivo* by the C-terminal domain of RNAP (Lu et al., 2018; Guo et al., 2019) or by multiple sequence-specific TFs containing disordered amino acid regions at their TAD (Boijja et al., 2018; Chong et al., 2018).

Developing research demonstrates the SWI/SNF complex within SEs, which drive AR transcriptional activation and PC growth. Of the three SWI/SNF complexes (Discussed earlier, Figure 5), BAF has shown to be mostly required for enhancer activation (Alver et al., 2017; Wang et al., 2017). Inactivation of the ATPase components led to a targeted loss of chromatin accessibility at the core-enhancer circuitry, thereby attenuating cancer-promoting transcriptional programs and tempering the enhancer related expression of driver oncogenes (Xiao et al., 2022). Cancer cells can also develop genomic alterations that create SE at driver oncogenes (Chapuy et al., 2013; Hnisz et al., 2013; Zhang et al., 2016).

The Mediator complex (discussed earlier Figure 6) is fundamental to the recruitment and maintenance and execution of transcription mediated by phase separation. The Mediator complex have been well

documented with Biomolecular condensates. Both *in vitro* and *in vivo* (cells) studies have shown many of Mediator's subunits forming condensates in combination with multiple TF's, which provides significant support to the role of biocondensates as master regulators in specific transcriptional programmes (Sabari et al., 2018; Cho et al., 2018; Boija et al., 2018; Shi et al., 2002).

As a corollary, MED1 presence is considered a marker for SEs, along with H3K27Ac and BRD4 (Whyte et al., 2013). MED1 contains LXXLL motifs, typical of co-regulators for nuclear receptors, and has been found to interact with the AF-2 of AR (Chen et al., 2011) and stimulates its transcriptional activity in a ligand-dependent manner (Yuan et al., 1998; Wang et al., 2002; Wang et al., 2005). Although it is now known the Tau1 site in the AR NTD can facilitate AR binding to Mediator as well (Ren et al., 2000; Jin et al., 2012).

MED1 has been shown to drive the activity of AR in PC cells (Wang et al., 2002) and is overexpressed in primary PC patient samples (Vijayvargia et al., 2007; Jin et al., 2013). The Phosphorylation of MED1 enhances its association with the Mediator complex and its binding to AR (Belakavadi 2008). Whilst several other Mediator subunits have been found to act as co-regulators for AR including MED24, MED14 and MED19 (Yuan et al., 1998; Wang et al., 2002; Imberg-Kazdan et al., 2013; Cui et al., 2011).

e. Castration-resistant prostate cancer (CRPC)

Interestingly, MED1 phosphorylation at T1457 enhances its interaction with AR and it is this phosphorylated form of MED1 that interacts with AR-VAR7. This is important as it is the AR-VAR7 that lacks the ligand-binding domain and is constitutively active in CRPC (Chen et al., 2011; Watson et al., 2015). Moreover, CRPC models have increased MED1 phosphorylation, and AR-mediated MED1 recruitment to chromatin (Rasool et al., 2019).

Apart from abnormal AR splice variant expression, the shift of PC to CRPC can originate from many sources, such as aberrant AR activation, enhanced AR expression, hypersensitivity to androgens (Waltering et al., 2009), intra-tumoral steroidogenesis (Locke et al., 2008) alteration of AR co-regulators and more (see Huang et al., 2018). The spectrum of heterogeneity in CRPC formation has led to a consensus of re-defining what exactly constitutes CRPC (Heidenreich et al., 2014).

f. Small molecule inhibitors

The last decade has led to the development of different therapeutics for CRPC. As this cancer still relies on the AR signalling axis, CRPC patients continue to receive ADT (enzalutamide, abiraterone, apalutamide, darolutamide) as a first line treatment, often in combination with chemotherapy (docetaxel) or Cabazitaxel (de Bono et al., 2010). A second therapeutic option is immunotherapy with spirucel-T, and in those cases with bone metastasis, treatment with a radioligand (radium 223 dichloride) is a possibility (Parker et al., 2013; Kantoff et al., 2010).

However, even with these recent developments in the clinic, the median survival after reaching CRPC stage is in the range of 15–36 months, depending on disease burden (Higano et al., 2011; Kirby et al., 2011). Therefore, finding new strategies to fight CRPC is of paramount importance.

Small molecule inhibitors are attractive anti-tumour agents due to their small size (≤ 500 Da), highly penetrative features, and suitability for oral administration (Imai et al., 2006; Zhong et al., 2020). To date, there are a total of 89 anti-cancer small molecules approved in the United States and China. They cover a large scope including kinases, epigenetic regulatory proteins, DNA damage repair enzymes, and proteasomes (Zhong et al., 2021).

In PC there are two small molecule inhibitors (ARV-110 and ARV-471) designed by proteolysis targeting chimera (PROTAC) technology that have entered clinical trials (An et al., 2018). ARV-110, which specifically binds to AR and mediates AR degradation, showed promise in early phase I/II clinical trials. The latest observations demonstrated prostate-specific antigen (PSA) declines of over 50% in a subset of CRPC patients (Gao et al., 2022; Wang et al., 2020).

Targeting of ARs NTD would be advantageous over drugs specific to the LBD of AR due to the NTD being essential for the transcriptional activities of AR full length and AR-VARs that lack the LBD (Huang et al., 2018). However, the high level of IDRs in the AR-NTD has hindered the development of inhibitors for this domain (Sadar et al., 2020).

EPI compounds were the first drugs designed to inhibit the AR-NTD. Their exact mechanism of action was not fully understood, but has been suggested they prevent transcriptional coregulator interactions within the TAD (Urol et al., 2020). Studies showed EPI-001 was promising in reducing clinically relevant AR-dependent mechanisms of treatment resistance and restricted PC growth in preclinical studies (Andersen et al., 2010). The successor to EPI-001, EPI-506, was investigated in a Phase I study on CRPC, but was discontinued after showing only minor PSA decline; ultimately proving ineffective owing to its high metabolism rate and low potency (Obst et al., 2019). However, a second-generation of EPIs, including EPI-7386, entered phase I/II clinical trials in 2020 (Sadar et al., 2020).

Derivatives of EPI compounds have been recently developed by the Salvatella and Riera labs with improved pharmacodynamics, such as 1ae (Basu et al., 2022). Targeting biomolecular condensates with small molecule inhibitors is an exciting and largely unexplored avenue in PC. Elucidation of the mechanisms controlling the regulation and organisation of PS condensates is central to unlocking the potential in drugging the currently 'undruggable' TFs like AR.

Objectives

Objectives

1. Determine the overlapping and/or specific molecular functions of GEMC1/MCIDAS.
 - Understand how GEMC1 and MCIDAS find and activate genes.
2. Identify the targets of CCNO and determine how it influences the transcriptional targets of Geminin family proteins.
3. Understand the mechanisms governing the heterogeneous organisation, nuclear translocation and function of AR .
 - Investigate how Phase Separation effects the AR interactome and its transcriptional function.

Materials

&

Methods

Cloning

Geminin Family

Human GEMC1 or MCIDAS were amplified from previously described expression vectors (Terré et al., 2016) using forward primers containing *AscI* and reverse primers containing *NotI* restriction sites (GEMC1-*AscI*-F-AAGGCGCGCCATGAACACCATTTCTGCCTT; GEMC1-*NotI*-R-TTGC GGCCGCCTAACTGGGGACCCAGCGGAACT; MCIDAS-*AscI*-F-AAGGCGCGCCATGCAGGCGTGCGGGGGCGGC; MCIDAS-*NotI*-R-TTGC GGCCGCCTAACTGGGGACCCAGCGGAACT) using KOD Hot Start DNA Polymerase (Millipore) and cycling conditions recommended from the manufacturer (polymerase activation at 95 °C for 2 min, denaturation at 95 °C for 20 s, annealing at 55 °C for 10 s and extension at 70 °C for 50 s, repeated for 40 cycles). PCR products were purified using the PureLink Quick Gel Extraction Kit (Invitrogen) and cloned into pCR2.1-TOPO vector (Invitrogen). Top10 competent *E. coli* cells (Invitrogen) were transformed with pCR2.1-GEMC1/MCIDAS and colonies were selected in carbenicillin. Constructs were verified by restriction digestion and sequencing (Macrogen) with primers for the TOPO vector (T7 Promoter-F and M13-R). Afterwards, GEMC1 or MCIDAS were cut from the pCR2.1-TOPO vector by restriction digest with *AscI* (New England BioLabs (NEB), Ipswich, MA, USA) and *NotI*-HF (NEB), purified using the PureLink Quick Gel Extraction Kit (Invitrogen) and ligated into pcDNA5/FRT/TO-N-FLAG-BirA* using Quick Ligation Kit (NEB). Top10 competent *E. coli* cells (Invitrogen) were transformed with pcDNA5/FRT/TO-N-Flag-hBirA*-GEMC1/MCIDAS vector and carbenicillin selected. The constructs were confirmed by restriction digestion with *AscI* (NEB) and *NotI*-HF (NEB) and sequencing (Macrogen).

Geminin family Hybrids

Initial Geminin family hybrid plasmids were designed by swapping the C-termini of GEMC1 and MCIDAS directly after the CC domain to generate GEM-IDAS (C-terminus of MCIDAS) and MC-C1 (C-terminus of GEMC1). Later Hybrid domains were designed to exchange the TIRT domain. The GEM-IDAS-C1 = GEM-IDAS with the GEMC1 TIRT domain. GEM-IDA-C1 = GEM-IDAS with the GEMC1 TIRT and the internal deletion of a predicted ordered region upstream from the TIRT within the MCIDAS sequence. The MCI-GEM-S is MCI-C1 with the MCIDAS TIRT domain. All constructs were made by Genescript and incorporated into pcDNA3.1(+)-N-DYK: Cloning site: 5' KpnI, 3' NotI.

For Co-immunoprecipitation experiments the GEM-IDAS and MC-C1 hybrid cDNAs were PCR amplified using forward primer containing *HindIII* and reverse primer containing *NotI* restriction enzyme sites (GM-F-*HindIII*: GATCGATCAAGCTTACCATGAACACCATTTCTGCC; GM-R-*NotI*: GATCGATCGCGGCCGCCTACTGGGGACCCAGCGGAA; MG-F-*HindIII*: GATCGATCAAGCTTACCATGCAGGCGTGCGGGGGC; MG-R-*NotI*: GATCGATCGCGGCCGCCTAGACTGCTTAGGGACCCA) and cloned into the pXJ40 vector with one HA-tag at the N-terminus.

CCNO

CCNO was previously described (Terré et al., 2019). Briefly, Human CCNO cDNA was obtained as a template from Expressed Sequence Tag EST IMAGE: 6421733 (GenBank accession number BQ917277.1). This was amplified using 5'-GCGAATTCCATGGTGACCCCCCTGTCCCACCAGCC-3' and 5'-GCTCTAGATTATTTTCGAGCTCGGGGGCAGG-3' primers. hCCNO cDNA was cloned into the pBluescript (I) SK(+) vector (Stratagene) at the EcoRI and XbaI restriction sites and then into the mammalian expression vector pCDNA3.1 (Invitrogen) modified with an N-terminal myc-tag under the control of the constitutive CMV promoter (pcDNA3.1-myc-hCno).

PSLIK - Induction Multiciliogenesis Cell model

Plasmids FLAG-hBirA*-GEMC1/MCIDAS were amplified using forward primers and reverse primers containing SpeI-XhoI restriction sites (GEMC1 F – AAAAAGTATggactacaaagacgatgac, R-TTTTCTCGAGCTAAGACTGCCTAGGGACCCA), (Primers used for amplification- MCIDAS, F- AAAAAGTATggactacaaagacgatg, R-TTTTCTCGAGTCAACTGGGGACCCAGCGGAAC) using KOD Hot Start DNA Polymerase (Millipore) and cycling conditions recommended from the manufacturer (polymerase activation at 95 °C for 2 min, denaturation at 95 °C for 20 s, annealing at 55 °C for 10 s and extension at 70 °C for 50 s, repeated for 40 cycles). PCR products were purified using the PureLink Quick Gel Extraction Kit (Invitrogen) and ligated into a pENTR vector for final cloning into the PSLIK-Neo destination vector (Cat# 25736) using LR Clonase gateway reaction vector (Invitrogen Cat. No. 11791-020). Top10 competent *E. coli* cells (Invitrogen) were transformed and colonies were selected in neomycin. Constructs were verified by restriction digestion and sequencing (Macrogen).

AR plasmids

Constructs to express FLAG-MTID or its fusions to AR-WT; 8YtoS, 22YtoS as well as AR-WT-22YtoF, AR-VAR7, AR-VAR7-22YtoS and AR-VAR7-22YtoF fusion proteins were synthesized by Genscript and either cloned into pcDNA3.1(-) and subsequently cloned into pLenti-CMV-MCS-GFP-SV-puro using XbaI and BamHI to replace GFP or cloned directly into pLenti-CMV-MCS-GFP-SV-puro by Genscript using the same sites. Sequences were codon optimized for mammalian expression and verified by sequencing. pLenti-CMV-MCS-GFP-SV-puro was a gift from Paul Odgren (Addgene plasmid # 73582)

Cell culture and transfection

General

AD293 cells (Agilent) and HeLa cells were cultured in Dulbecco's Modified Eagle Medium (Gibco) with 10% FBS (Hyclone) were routinely tested for mycoplasma and found negative. For transient transfections, cells were seeded in either 10 cm or 6 well plates at 70% confluence and transfected with 10 µg or 3 µg of plasmid respectively, with polyethylenimine (Polysciences) for AD293 or Lipofectamine 2000 (ThermoFisher Scientific) for HeLa cells. The medium was changed 8 hrs post-transfection and cells were collected after 48 hrs. Full established cell lines used can be seen at Table 1.

Table 1. Established Cell Lines.

| Established Cell Lines | Company | Identifier |
|--|----------------------|-------------------|
| AD293 | Stratagene | Cat# 240085 |
| Hela | ATCC | Cat# CCL-2.2 |
| Mouse Tracheal Epithelial Cell (MTEC) cultures for (ALI) | Mahjoub et al., 2010 | N/A |
| PC3 | ATCC | CRL-1435 |
| LNCaP FGC | ATCC | CRL-1740 |
| DBTRG-05MG | ATCC | CRL-2020 |
| U2-OS | ATCC | HTB-96 |

Multiciliogenesis Induction Cell model

For induction experiments, expression was induced on D0 through addition of Doxycycline 1ug/ml with the addition of 5uM DAPT notch inhibitor (Sigma 208255-80-5). The next day media was changed to 1% FBS whilst maintaining Doxycycline and notch inhibition. On D3 the media was replenished with media containing the previously described with the addition of Noggin 20ng/ml (In house made from IRB Mass Spectrometry core facility). Measurements were taken D0, D3 and D7 where cells were plated on coverslips. Immunostainings were performed as previously described, antibodies utilised were Acetylated tubulin (Sigma T6793) and Centrin (#04-1624). With Secondary (Goat anti-Mouse IgG2b, Alexa Fluor™ 488 # A-21141) and (Goat anti-Mouse IgG2a, Alexa Fluor™ 568Catalog # A-21134). Utilisation of DBRD9 – a PROTAC degrader (Cat# 6943) against BRD9 was applied at D0 in conjunction with Doxycycline induction. PLA was first performed as previously described, slides were then counter stained with Acetylated tubulin before mounting. Fluorescence images were acquired and deconvoluted using a Leica TCS SP8 confocal microscope. Full list of inhibitors used in this thesis described in Table 2.

Table 2. Inhibitors used

| Inhibitor | Company | Identifier |
|------------------|----------------|-------------------|
| DBRD9 | TOCRIS | Cat #6943 |
| I-BRD9 | Sigma | Cat #SML1534 |
| DAPT | Sigma | Cat #208255-80-5 |

MTEC ALI cultures

All animal studies were performed following protocols that are compliant with guidelines of the Institutional Animal Care and Use Committee at Washington University and the National Institutes of Health. Mouse Tracheal Epithelial Cell (MTEC) cultures were established as previously described

(Nanjundappa et al., 2019; Silva et al., 2016). Briefly, C57BL/6J mice were euthanized at 2-4 months of age, trachea were excised, opened longitudinally to expose the lumen, placed in 1.5 mg/ml pronase E in DMEM/F-12 supplemented with antibiotics and incubated at 4°C overnight. Tracheal epithelial progenitor cells were dislodged by gentle agitation and collected in DMEM/F12 containing 10% FBS. After centrifugation at 4°C for 10 min at 800 g, cells were resuspended in DMEM/F12 with 10% FBS and plated in a Primaria Tissue Culture dish (Corning) for 3-4 h at 37°C with 5% CO₂ to adhere contaminating fibroblasts. Non-adhered cells were collected, concentrated by centrifugation, resuspended in an appropriate volume of MTEC-complete medium (described in (You and Brody, 2013; You et al., 2002)), and seeded at a density of 9×10^4 cells/cm² onto Transwell-Clear permeable filter supports (Corning) treated with 0.06mg/ml rat tail Collagen type I. Air liquid interface (ALI) was established after cells reached confluence by feeding cells with MTEC serum-free medium (You and Brody, 2013; You et al., 2002) only in the lower chamber. Cells were treated with either 5, 10 or 20uM I-BRD9 (#SML1534) in serum-free medium, which was exchanged every 2 days. Cells were cultured at 37°C with 5% CO₂, and media containing I-BRD9 replaced every 2 days for up to 12 days. Samples were harvested at ALI days 4, 8, and 12, fixed in either 100% ice-cold methanol or 4% paraformaldehyde in PBS at room temperature for 10 min.

AR cell culture

PC3 (ATCC; CRL-1435) and LNCaP clone FGC (ATCC; CRL-1740) cells were cultured in RPMI containing 4.5 g L⁻¹ glucose (Glutamax, Gibco) supplemented with 10% (v/v) charcoal stripped FBS (Gibco) and antibiotics, unless specified otherwise.

Lentiviral production and stable cell lines

Plasmids were co-transfected with lentiviral packaging plasmid vectors REV (Cat# 12253), RRE (Cat# 12251) and VSV-G (Cat# 8454) into AD293 cells with PEI (Sigma-Aldrich). Two days after transfection, virus-containing medium was collected and filtered through a 0.45- μ m low-protein-binding filtration cartridge. The virus containing media was directly used to infect primary cells in the presence of polybrene (8 μ g/mL) for 48 hours, before correct selection was applied corresponding to resistance gene, either 400 μ g/ml Neomycin was introduced for 72 hours or puromycin 2ug/ml. Recovered cells were then tested for successful introduction of protein expression via western blot.

Western blotting assays

For western blotting, cells were collected and lysed in RIPA buffer (1% NP-40, 0.1% sodium dodecyl sulfate (SDS), 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.5% sodium deoxycholate, 1 \times proteinase inhibitor cocktail (Roche)) on ice. Samples were sonicated using a Bioruptor XL sonicator (Diagenode) for 15 min with 15 sec intervals and centrifuged at 4°C for 15 min at 1300 rpm. The proteins were quantified and resolved on 8% SDS-polyacrylamide gels with 1% (v/v) β -mercaptoethanol, 0.01% (w/v) Bromophenol

blue. Transferred onto Nitrocellulose (Bio-Rad) 100V for 1 hour. The membrane was blocked with 5% milk powder (Bio-Rad) then incubated with specific antibodies at 4°C overnight. Following incubation with secondary antibodies, immunoblots were visualized LI-COR Odyssey CLx system (LI-COR Biotechnologies). Full antibody list can be seen at Table 3.

Table 3. Antibodies used

| Antibody | Company | Identifier |
|--|----------------|--------------------------------------|
| ARID1a/BAF250A | Sigma | Cat# 12354, RRID:AB_2637010 |
| BRD9 | Bethyl | Cat# A303-781A, RRID:AB_11218396 |
| Baf57/SMARCE1 | Bethyl | Cat# IHC-00129, RRID:AB_2192300) |
| Baf155/SMARCC1 | Abcam | Cat# ab26304, RRID:AB_470850 |
| Flag | Sigma | Cat# P2983, RRID:AB_439685 |
| Actin | Sigma | Cat# A4700, RRID:AB_476730 |
| LaminA | Sigma | Cat# SAB4501764, RRID:AB_10744653 |
| Goat anti-Mouse IgG (H+L) Alexa Fluor 680 | Invitrogen | Cat# A-21057, RRID:AB_2535723 |
| Goat anti-rabbit IgG (H+L) Alexa Fluor 680 | Invitrogen | Cat# A21088, RRID:AB_10373119 |
| Androgen Receptor [ER179(2)] | Abcam | Cat #ab108341 |
| Androgen Receptor (441) | SCBT | Cat #sc-7305 |
| GAPDH (ZG003) | Invitrogen | Cat #39-8600 |
| Goat anti-Rabbit IgG (H+L) Secondary Antibody, HRP | Invitrogen | Cat #65-6120 |
| Goat anti-Mouse IgG (H+L) Secondary Antibody, HRP | Invitrogen | Cat #G-21040 |
| Nup153 antibody (QE5) | Abcam | Cat #ab24700 |
| Med1 antibody | Abcam | Cat #ab64965 |

| | | |
|---|-------------------------------|---------------|
| Streptavidin antibody, Alexa Fluor™ 488 conjugate | ThermoScientific | Cat #S11223 |
| Centrin | Sigma | Cat #04-1624 |
| Actylated Tubulin | Sigma | Cat #T6793 |
| PML | Sigma | Cat #P6746 |
| FLAG | Sigma | Cat #F3165 |
| CDK1 | Millipore | Cat #MAB8878 |
| CDK2 | Santa Cruz | Cat #sc-163 |
| PLK1 | Calbiochem | Cat #DR1037 |
| CDC20 | Santa Cruz | Cat #sc-13162 |
| Anti-Phosphoserine/threonine | BD Transduction Laboratories™ | Cat #612548 |

BioID analysis of proximity interactions

AD293 cells were seeded and transfected the next day using PEI (Sigma-Aldrich) $\pm 50 \mu\text{M}$ biotin (IBA GmbH; 2-1016-002). For mass spectrometry, 5x15 cm² plates per condition were harvested 24 h after transfection by scraping cells into PBS, washing two times in PBS and snap freezing on dry ice. Cell pellets were lysed in 5 ml modified RIPA buffer (1% TX-100, 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 0.1% SDS, 0.5% sodium deoxycholate and protease inhibitors) on ice, treated with 250 U benzonase (Millipore) and biotinylated proteins were isolated using streptavidin-sepharose beads (GE Healthcare). Proteins were washed in ammonium bicarbonate and digested with trypsin. Mass spectrometry was performed as described previously (Silva et al., 2018).

Small scale BioID-AP followed by western blotting was carried out as described above, but using 3x15 cm² plates per condition and, following affinity purification with streptavidin-sepharose beads (GE Healthcare), samples were centrifuged briefly and the supernatant aspirated. Samples were resuspended in 100ul-150ul RIPA buffer and boiled at 95°C for 10 minutes. Lysates were separated on SDS-PAGE gels and transferred for western detection as described previously.

Co-immunoprecipitation

Co-immunoprecipitation assays followed by western blot were performed as previously described. Combinations of plasmids were co-transfected into AD293 cells (3μg plasmid, per 10 cm dish) using Lipofectamine 2000 (Thermo Fisher Scientific). 24 hrs post-incubation, cells were lysed in 800μl of RIPA

buffer (Thermo Fisher Scientific) supplemented with EDTA-free Protease Inhibitor Cocktail (Roche, 11836170001). Cell lysates were sonicated briefly, spun down and an aliquot taken for total cell lysate control. Remaining lysate was rotated O/N with 25µl of Protein-A-agarose beads (Roche) and 2µg of mouse anti-HA antibody (Santa Cruz, sc-7392). Beads were washed four times in RIPA buffer and boiled in 50µl of 1× SDS loading buffer. Total cell lysates (15µl, 1%) and IP (15µl, 30%) were resolved with SDS-PAGE gels, transferred to PVDF membranes, blocked in 2% BSA and probed with relevant primary antibodies for western detection.

siRNA Knockdown

For RNA interference experiments (RNAi) cells were transfected with siRNAs using Lipofectamine RNAiMAX (Invitrogen) according to manufacturer's protocol. RNAi oligos used were: siGEMC1.1 (5'-CCACAUCCUUUCUCAACUU-3') 40nM; siGEMC1.2 (5'-CCAGUUUCCGGAUGAUGCA-3') 80 nM; siCdt1 (5'- AACGUGGAUGAAGUACCCGAC-3') 50 nM; control luciferase siGL2 (5'-CGUACGCGGAAUACUUCGA-3') 40-80 nM. All siRNAs can be seen at Table 4.

Table 4 siRNAs

| Oligonucleotides | Sequence |
|------------------|---------------------------|
| siGEMC1.1 | 5'-CCACAUCCUUUCUCAACUU-3' |
| siCdt1 | 5'-UGCCAACUCUGGAAUCAA-3' |
| siGL2 | 5'-CGUACGCGGAAUACUUCGA-3' |

Quantitative real time-PCR (RT-qPCR)

Transfected cells were collected after two cold PBS washes by scraping into Tri-Reagent. RNA was isolated by chloroform extraction followed by centrifugation, isopropanol precipitation, 2 washes in 75% ethanol and resuspension in DEPC-treated water. Nucleic acid quantification was performed with a Nanodrop 8000 Instrument (ThermoFisher Scientific). cDNA was generated using 0.5-1µg of total RNA and a High Capacity RNA-to-cDNA Kit (Applied Biosystems). RT-qPCR was performed using the comparative CT method and a StepOne Real-Time PCR System (Applied Biosystems). Amplification was performed using Power SYBR Green PCR Master Mix (Applied Biosystems) or TaqMan Universal PCR Master Mix (Applied Biosystems). All assays were performed in duplicate. For TaqMan assays, *mActB* probe was used as an endogenous control for normalization and a specific Taqman probe was used for mouse or human *TP73* (Hs01056231_m1), *CCNO* (Hs004389588_91), *FOXJ1* (Hs00230964_m1). Full list of primers used can be found in Table 5.

Table 5. SYBR green and Taqman primers

| Oligonucleotides | Sequence | Identifier |
|------------------|-------------------------|------------|
| PKYMT1-F | CATGGCTCCTACGGAGAGGT | N/A |
| PKYMT1-R | ACATGGAACGCTTTACCGCAT | N/A |
| GAPDH-F | GCACAGTCAAGGCCGAGAAT | N/A |
| GAPDH-R | GCCTTCTCCATGGTGGTGAA | N/A |
| HSPA1L-F | TGATGCCAATGGTATTCTCAATG | N/A |
| HSPA1L-R | CAGGCGGCCCTTGTCAT | N/A |
| CCDC96-F | GACATCCTGACCAAGACGAAG | N/A |
| CCDC96-R | GAAGTGAGTCCTTGCTAGAAG | N/A |
| Taqman Probes | | |
| <i>TP73</i> | Hs01056231 | |
| <i>CCNO</i> | Hs004389588 | |
| <i>FOXJ1</i> | Hs00230964 | |

Immunofluorescence (IF)

For IF cells were grown on 8-well Lab Tek II chamber slides (Labclinics) to sub-confluence. Cells were grown on Poly-L-Lysine coated coverslips. Cells were fixed for 10 minutes in 4% paraformaldehyde on ice and permeabilised for 5 minutes in 0.2% Triton in 1X PBS at room temperature. After two washes in PBS, fixed cells were incubated for 1 h in blocking solution (5% BSA 0.1% Tween/PBS) and stained with primary antibodies in blocking solution for 1hr at room temperature in a humid chamber. The secondary antibodies used were Alexa Fluor 488 (goat anti-rabbit IgG), Alexa Fluor 488 (goat anti-mouse IgG), Alexa Fluor 568 (goat anti-rabbit IgG), and Alexa Fluor 568 (goat anti-mouse IgG) from Thermo Fisher Scientific and were used at a dilution of 1:500 in blocking solution. After primary and secondary antibody incubations, washes were performed with blocking solution. The last wash before mounting contained DAPI. Slides were air-dried, mounted in Vectashield (Vector Labs), sealed and stored in a cold dark chamber. Confocal fluorescence images were obtained on a Leica DM2500 SPE confocal system. Images were taken with 40x NA 1.15 oil or 63x NA 1.3 oil objectives and the standard LAS-AF software.

Proximity Ligation Assay (PLA)

Protein—protein interactions were studied using a Duolink In Situ Orange Starter Kit Mouse/Rabbit (Sigma DUO92102) following the manufacturer's protocol. Briefly, transfected HeLa cells were seeded in coverslips (company) and cultured overnight. Slides were washed with cold 1×PBS and fixed in 4% paraformaldehyde for 15 min, permeabilized using 0.1% Triton X-100 for 10 min and then blocked with blocking buffer for 1 hour at 37°C. The cover slips were blocked with Duolink Blocking Solution in a pre-heated humidified chamber for 30 min at 37°C. The primary antibodies were added to the coverslips and incubated overnight at 4°C. Then coverslips were washed with 1×Wash Buffer A and subsequently incubated with the two PLA probes (1:5 diluted in antibody diluents) for 1 h, then the Ligation-Ligase solution for 30 min, and the Amplification-Polymerase solution for 100 min in a pre-heated humidified chamber at 37°C. Before imaging, slides were washed with 1×Wash Buffer B and mounted with a cover slip using Duolink In Situ Mounting Medium with DAPI. Fluorescence images were acquired using a Leica TCS SP8 confocal microscope. The signal was detected as a distinct fluorescent dot and analysed by fluorescence microscopy. Negative controls consisted of samples treated as described above but with only primary, secondary or control IgG antibodies.

BrdU labelling for G1/S checkpoint assay by flow cytometry

For BrdU labelling, cells were plated 24 hours post transfection. After 24 hours, cells were pulsed with 10 µM BrdU for 20 min, and cells were trypsinized and fixed overnight in ice-cold 70% ethanol. DNA was denatured by 0.1M HCl and incubated at 100 °C. S-phase cells were stained using the BD Pharmigen FITC Mouse Anti- BrdU Set antibody (BD Biosciences). Cells were resuspended in 400 µl PBS containing 25 µg/ml propidium iodide (PI) and 0.1 mg/ml RNaseA and subjected to FACS analysis in a Gallios Flow Cytometer (Beckman Coulter). The percentage of BrdU positive cells and the cell cycle based on PI content was analysed with FlowJo software

CRISPR knockout of ARID1A

For the generation of stable ARID1A KO cells, AD293 cells were transfected with the NickaseNinja (ATUM) vector (pD1401-AD: CMV-Cas9N-2A-GFP, Cas9-ElecD) co-expressing two gRNAs targeting ARID1A. ARID1A gRNA sequences (GGGGAGCTCAGCGCGTAGGC and TGGCACTCCGGGCTCCGGCG) were designed using the ATUM gRNA Design Tool. 48 hrs post-transduction, positive GFP cells were sorted by FACS (BD FACSAria™ Fusion) and plated into 96-well plates. After 15 days, clones were collected and validated by western blot.

Phosphoproteomics

AD293 cells transfected with CCNO (3 x 15 cm plates). The next day the media was supplemented with 50µM of Biotin. 48 hours post transfection the cells were trypsinized and washed. Small scale BioID-AP was

performed as described above with sepharose beads. The proteins were digested of the beads and BioID-MS analysis was performed as described. For phosphoproteomic enrichment, the phosphopeptides were enriched using TiO₂ columns (Titansphere, GL Sciences, 5µm) in Eppendorf GELoader Tips. TiO₂ microcolumns were packed as described elsewhere (Larsen et al., 2005). Briefly, the protein digest was then diluted in a loading buffer containing 80% acetonitrile and 0.1% TFA and loaded on the EV1106 column. The column was washed with 30 µl of loading buffer and then with washing buffer (40 µl; 80% acetonitrile 2% TFA). The bound peptides were eluted with H₂O 0.1% formic acid and CH₃CN 0.1% formic acid.

MS analysis was carried out using the Orbitra p Eclipse™ Tribrid (Thermo Scientific) with detection system Orbitrap (MS1 120k), peptides were analyzed in positive ion mode. The phosphorylation peptides were analyzed using Proteome Discoverer (Thermo Scientific) / v2.5.0.000) with MaxQuant (Cox, 2008) v2.0.2.0.

RNA-Seq

Stable AR mutant PC3 cells were plated (1 million) in 6 -well plates. The next day the cells are washed twice with cold PBS and collected by scraping in Tri-Reagent (Sigma-Aldrich). RNA was isolated using the PureLink RNA Mini Kit (Thermo Fisher Scientific) following the manufacturer's instructions. Briefly, chloroform extraction followed by centrifugation resulted in a colourless upper aqueous phase that was mixed 1:1 with 100% ethanol. Sample containing RNA in 50% ethanol was bound to the spin cartridge, washed twice with Wash Buffer II and eluted in RNase-free water. RNA quantity was checked for quality via nanodrop. RNA sampling was conducted in National Cancer Institute (NCI). Sample mRNA was purified using the Nucleospin RNA Plus kit (Takara). Libraries were prepared using the Nextseq 500/550 High Output kit v2.5 (Illumina) following manufacturer's instructions and 75x75 PE sequencing was run on a NextSeq550 in the CCR Genomics core facility of the NCI. Gene differential expression was performed using DESeq2 (Love et al., 2014) with replicates as covariate. Pathway enrichment was assessed through the preranked version of Geneset Enrichment Analysis (GSEA) (Subramanian et al., 2005). GSEA was applied to the ranking defined by the log₂ Fold Change of the differential expression analysis using DESeq2. Genesets for analyses were from the Gene Ontology (GO) terms (Ashburner et al., 2000) as collected in the Hallmark collection (Liberzon et al., 2015) after retrieval from the MsigDB (Liberzon et al., 2011). Classification in coding and non-coding gene classes was performed according to the Gencode annotation version 19 (Harrow et al., 2012). For Repeat Masker analysis, sequences from repeat elements in the human Repbase database version 22.06 (Bao et al., 2015) were downloaded. Fold changes were computed between AR-WT or FLAG samples using normalised reads values of repeat sequences that had at least 100 raw reads.

Results

Chapter I

Cellular Characterisation of GEMC1 and MCIDAS

The Geminin family of proteins consists of three members, Geminin, GEMC1 (or GMNC) and MCIDAS. Geminin contains a central coiled-coil (CC) domain and its primary function is to bind to Cdt1 to prevent origin firing and re-replication in the same cell cycle (McGarry and Kirschner, 1998; Wohlschlegel et al., 2000). GEMC1 and MCIDAS were identified by their similar CC domain but have distinctive structures. Both proteins contain a Geminin-like CC with 47% sequence homology to each other in Humans (Figure 1A-C). The CC enables dimerisation and is somewhat conserved across species (Caillat et al., 2015). (Figure 1D). Both GEMC1 and MCIDAS contain a C-terminal TIRT domain, also with 47% homology (Figure 1A). This domain is not present in Geminin and is essential for transcriptional activation (Arbi et al., 2016; Terré et al., 2016; Lu et al., 2019).

Previously in the Stracker group, GEMC1 was knocked out in mice and did not cause a replication phenotype. Instead, the mice displayed growth impairment, developed hydrocephaly and were infertile, due to defects in the formation of MCCs (Terré et al., 2016). In contrast, mice lacking MCIDAS have columnar MCC like cells that are positive for MCC markers, including FOXJ1 and TP73, but they fail to amplify centrioles and generate motile cilia, consistent with the fact that GEMC1 expression preceded MCIDAS (Zhou et al 2015., Kim et al., 2018).

As both GEMC1 and MCIDAS do not contain any notable DNA binding domains, we sought to better understand their transcriptional roles in multiciliogenesis and find any functional differences that may exist and explain their specific functions. Analysis of their structural properties showed a high degree of intrinsic disorder in their structures, especially just after their CC domains (Figure 1E, F). Immunostaining in AD293 cells expressing either GEMC1 or MCIDAS following transient transfection showed a clear propensity to form nuclear foci that resemble condensates (Figure 1G). These foci did not overlap with PML bodies and may represent transcriptionally active sites (Figure 1H).

GEMC1 and MCIDAS activate common and specific gene targets

Transient overexpression of either GEMC1 or MCIDAS is sufficient to activate key genes implicated in MCC differentiation, including *TP73*, *MYB*, *CCNO* and *FOXJ1*, in a manner dependent on the CC and TIRT domain of either protein (Arbi et al., 2016; Lalioti et al., 2019; Lu et al., 2019; Terré et al., 2016, 2019). We took advantage of this facile system to try and understand their relative influence on gene expression. Microarray analysis of gene expression changes in AD293 cells expressing either GEMC1 or MCIDAS following transient transfection, revealed the activation of a similar core set of genes implicated in MCC differentiation (Figure 2A, B). This included several transcription factors (TF) important for MCC differentiation, including *FOXJ1*, *TP73*, *MYB* and *FOXN4*, as well as additional genes implicated in the MCC program, such as *CCNO* and *CDC20B* (Lu et al., 2019).

GEMC1 and MCIDAS expression led to the differential expression ($p < 0.05$, $fc \geq 0.25$ or ≤ -0.25) of a similar number of genes, and around 50% of those induced by GEMC1 were common to MCIDAS (Figure 2C). Functional enrichment analysis on gene ontologies (GO) revealed that GEMC1 and MCIDAS activated genes in overlapping GO categories, including Ribosome biogenesis, Developmental maturation, Transcription factor binding and Cilium, with the latter only significantly enriched by MCIDAS (Figure 2D). Despite the similarity in their transcriptional profiles, both proteins also activated some targets preferentially (Figure 2C). Several genes were highly enriched with *MCIDAS* in comparison to *GEMC1*, including *CCDC96* and *HSPA1L*, that have both been previously implicated in centrosome or cilium biology (Figure 2B) (van Dam et al., 2019; Firat-Karalar et al., 2014).

Figure 2

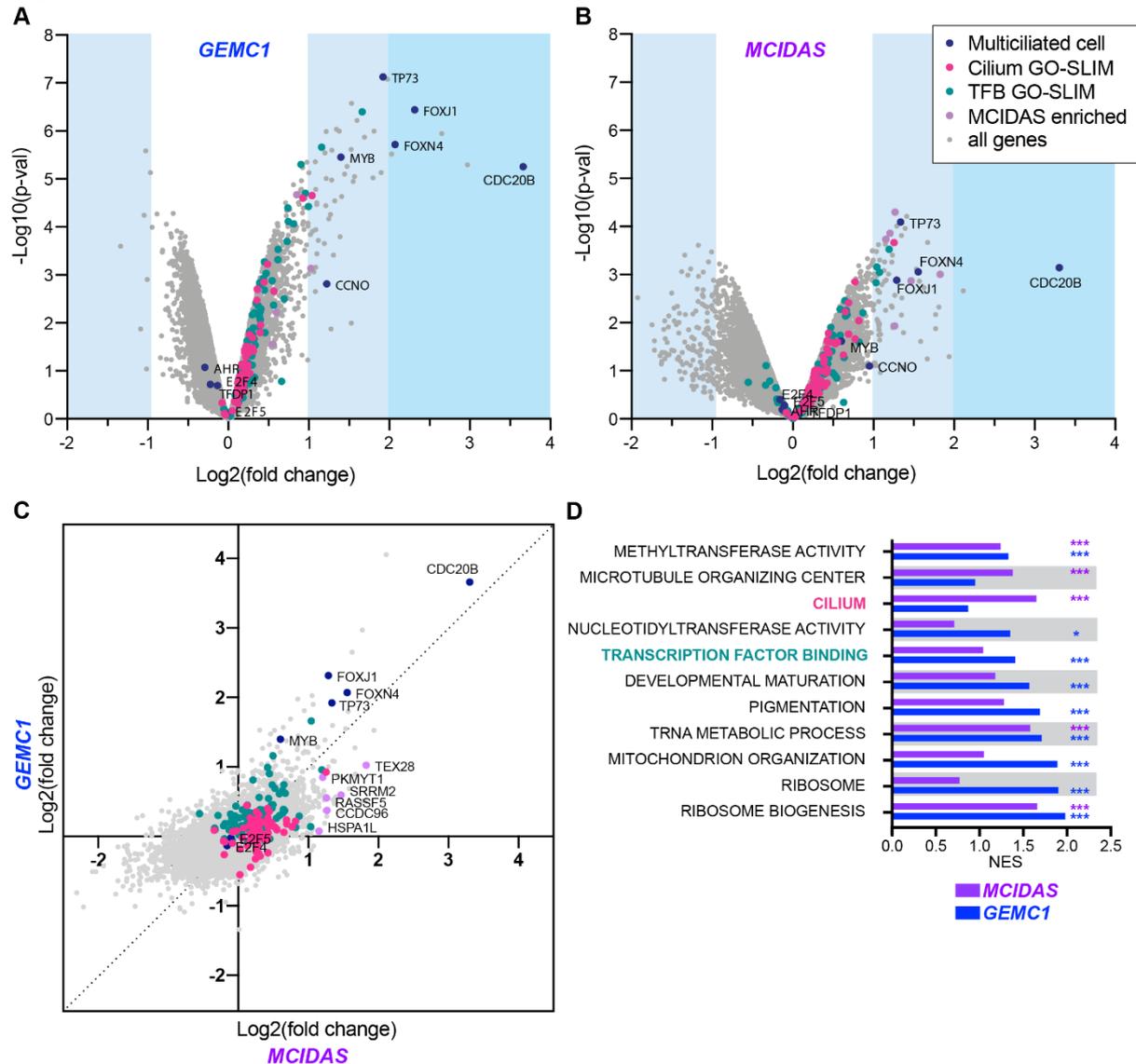


Figure 2: GEMC1 and MCIDAS activate overlapping and distinct target genes. **A.** Volcano plots ($-\log_{10}$ p-value vs \log_2 fold change) of microarray analysis of gene expression following transient transfection of *GEMC1* and **B.** *MCIDAS* in AD293 cells. Genes in the Cilium and Transcription Factor Binding Gene Ontologies (GO-SLIM categories) (TFB = Transcription factor binding), genes implicated specifically in MCCs or genes enriched with *MCIDAS* are indicated by colour (see legend, applies to A, B). **C.** Scatter plot comparing gene expression (\log_2 fold change) following GEMC1 or MCIDAS expression. Colour coding of gene categories is applied as in B. **D.** Gene set enrichment analysis of the gene expression data using GO-SLIM categories. The nominal enrichment score (NES) for each category with either GEMC1 or MCIDAS is shown.

Identification of the proximal interactomes of GEMC1 and MCIDAS

We wanted to better understand how GEMC1 and MCIDAS can regulate transcription and achieve distinct target specificity without any known DNA binding. To address this, we performed BioID-mass spectrometry (BioID-MS) to identify proximal interactors for each protein. Both genes were N-terminally tagged with a FLAG-tagged BirA* enzyme and expressed in AD293 cells supplemented with biotin for 20 hours (Figure 3A). Biotinylated proteins were affinity purified, trypsinised and identified by LC-MS/MS. The proximal interactomes of GEMC1 and MCIDAS were highly similar and the proteins identified were almost exclusively involved in transcription (Figure 3B, C). Validating the approach, known interactors of each protein were identified (Figure 3C, D). This included Geminin, that interacts with the CC domain of both proteins and its primary interactor CDT1, as well as E2F4/5-DP1 that interacts with the C-terminal TIRT domain of GEMC1 or MCIDAS (Arbi et al., 2016; Caillat et al., 2015; Ma et al., 2014; Terré et al., 2016).

Consistent with previous work, MCIDAS showed increased labeling of E2F4/5-DP1 compared to GEMC1 (Figure 3D, E) (Lu et al., 2019), and E2F3 and the E2F regulators RB (RB1) and p107 (RBL1) were also identified. Multiple components of the DREAM/MuvB complex, including MYB-B (MYBL2), LIN9, LIN52, and LIN54, that functions with E2F to regulate cell cycle related genes, were identified with both baits (Figure 3D, E). TRRAP, that was previously shown to be required for multiciliation, as well as many components of its associated chromatin regulatory complexes SAGA/STAGA, ATAC and NuA4, were identified (Figure 3D, E, Figure 4). In addition, many other transcriptional regulatory complexes and factors, including the p300 and CBP acetyltransferases, SWI/SNF, Mediator, LSD-CoREST, Polycomb (PRC1, PR-DUB), and COMPASS-like complexes were identified to different extents with both baits (Figure 3D, E and Figure 4). Similarly, TP73, that is transcriptionally activated by both GEMC1 and MCIDAS and shown to be a direct interactor of GEMC1, as well as the YAP1 and TEAD1 TF that interact physically and functionally with TP73, were identified with both baits (Figure 3D, E and Figure 4) (Lalioi et al., 2019; Papaspyropoulos et al., 2018; Strano et al., 2001; Wang et al., 2018).

In contrast, the Aryl hydrocarbon receptor (AHR), that is required for *CCNO* expression and MCC differentiation at some developmental stages, and its associated proteins ARNT and AIP, as well as ELMSAN1/MIDEAS and DNNTIP1, defining components of the MiDAC complex, were identified as proximal interactors highly enriched with GEMC1 compared to MCIDAS (Figure 3D,E) (Bantscheff et al., 2011; Villa et al., 2016). CDK2 and Cyclin A (CCNA2) were also highly enriched in GEMC1 transfected samples and have been shown to associate with the MiDAC complex, as well as regulate E2F4/5 and RB family members (Figure 3D, E) (Kent and Leone, 2019; Pagliuca et al., 2011).

To further validate some of these results, we performed small scale BioID affinity purifications (BioID-AP) followed by western blotting in transfected AD293 cells. We observed that expression of either GEMC1 or MCIDAS fused to BirA* led to similar, robust levels of YAP1 biotinylation that was not observed in controls (Figure 3G). Similarly, the SMARCE1(BAF57) component of the SWI/SNF

remodeler, that plays important roles in transcriptional regulation, was labeled to a similar extent by both GEMC1 and MCIDAS, but not BirA* alone, consistent with the proteomic data (Figure 3G).

Figure 3

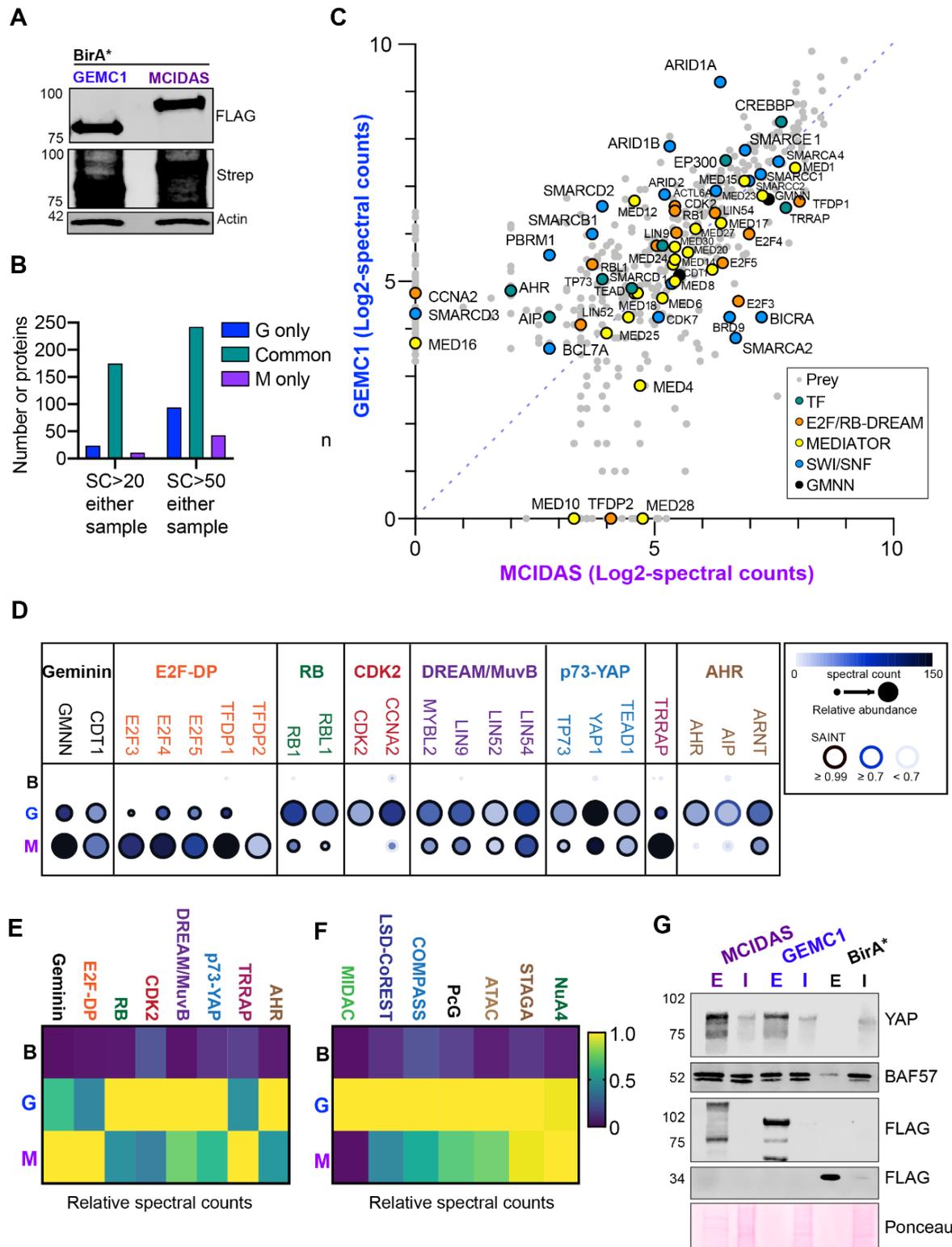


Figure 3: The proximal interactomes of GEMC1 and MCIDAS. **A.** Western blots from lysates of AD293 cells transfected with FLAG-BirA*-GEMC1 or FLAG-BirA*-MCIDAS. Expression of baits in cells with or without biotin supplementation is shown using the FLAG antibody (top panel) and labeling

in biotin supplemented samples with Strep-HRP (middle panel). **B.** Graph of protein number unique or common to each sample using a cutoff of >20 or >50 spectral counts (SC) with a SAINT score of >0.7 in one of the samples. **C.** Scatter plot of peptides identified in GEMC1 or MCIDAS samples (Log2- spectral counts of 2 technical replicates). All peptides with a SAINT score of >0.7 in one of the samples are shown. Specific subsets are highlighted including key TF, components of the E2F or DREAM complexes, SWI/SNF and Mediator related proteins. Specific subsets are shown in detail in Figures 2 and 3. **D.** Dot plot depicting relative abundance and spectral counts for selected proteins. B=BiA*, G=GEMC1 and M=MCIDAS. **E.** Heatmap of relative combined spectral counts for selected groups. Colour coding corresponds to protein groups defined in D. **F.** Heatmap of relative combined spectral counts for selected transcriptional complexes in each sample. Colour coding corresponds to protein complex groups defined in Figure 4. **G.** BioID-AP western blots for YAP1 and BAF57 (SMARCE1) a SWI/SNF component following expression GEMC1/MCIDAS baits (FLAG). Input (I) and streptavidin Eluate (E) (See A). Ponceau staining shown for loading and transfer control.

These results established that the proximal interactomes of GEMC1 and MCIDAS are highly similar and comprise primarily of proteins involved in transcription, including many factors required for multiciliation

Figure 4

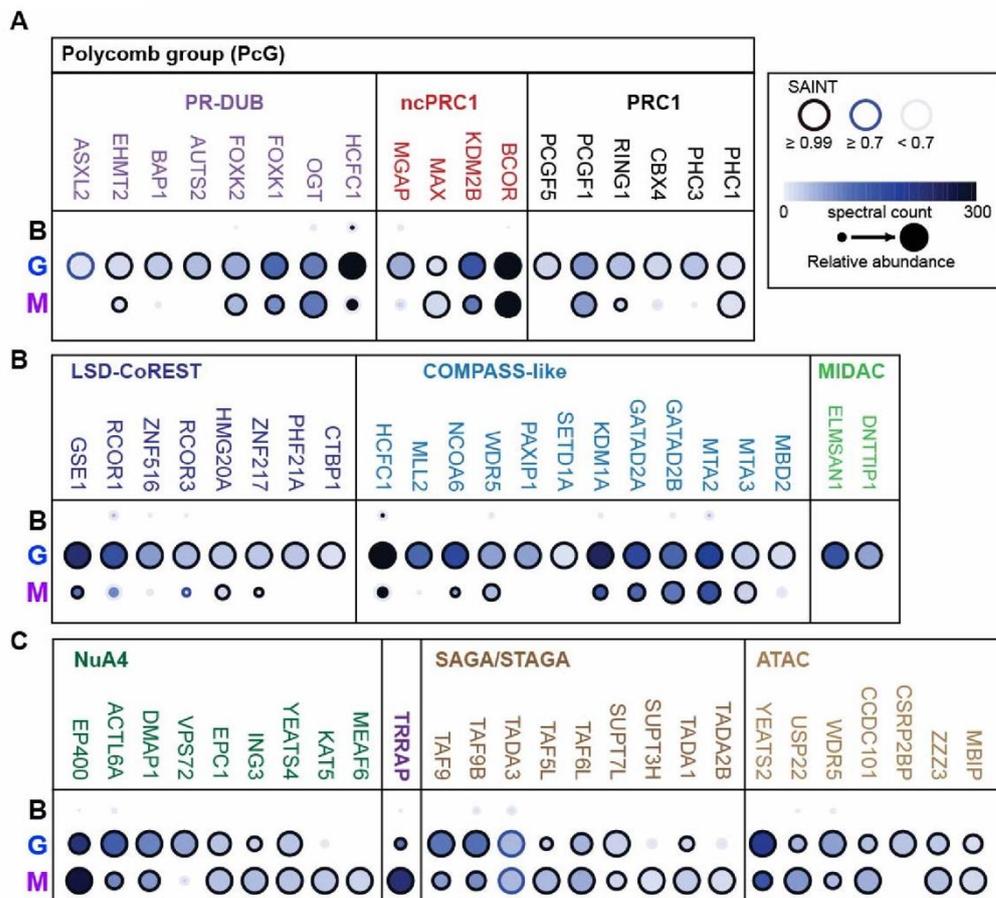


Figure 4. The proximal transcription factor complexes of GEMC1 and MCIDAS. A. Dot plot depicting relative abundance and spectral counts for selected Polycomb group (PcG) proteins grouped by

known complexes. Key applies to all panels. **B.** Dot plot depicting relative abundance and spectral counts for selected complexes enriched with GEMC1 compared to MCIDAS. **C.** Dot plots of transcriptional complexes associated with TRRAP. B=BirA*, G=GEMC1 and M=MCIDAS for all panels

GEMC1 and MCIDAS proximally interact with Mediator and distinct SWI/SNF complexes

While many other proteins involved in coactivator and corepressor complexes, as well as basal TFs, splicing regulators and RNA binding proteins, were identified, both samples were notably enriched for components of the SWI/SNF nucleosome remodeler and the Mediator complex, that is involved in the coordination of promoter and enhancer elements and recruitment of RNAP (Figure 3D and 5A) (Savas and Skardasi, 2018; Soutourina, 2018). Although there were differences in specific subunits, the overall abundance and identity of Mediator subunits identified was highly similar between GEMC1 and MCIDAS (Figure 5A-C).

While potentially hundreds of distinct SWI/SNF subcomplexes can be formed, 3 major subcomplexes, defined by distinct DNA and chromatin binding components, have been identified: cBAF, PBAF and ncBAF/GBAF, that each contain overlapping and unique subunits (Figure 5C) (Mashtalir et al., 2018). In addition, specificity of SWI/SNF complexes can be dictated by the specific ATPase subunit used, either SMARCA2 (BRM) or SMARCA4 (BRG1).

While GEMC1 and MCIDAS identified similar core subunits, we observed differential labeling of subunit-specific complexes and ATPase domains. In the case of GEMC1, we found primarily BAF, defined by the ARID1A/B subunits, and to a lesser extent PBAF subunits and near exclusive labeling of BRG1, a known interactor of Geminin (Figure 5A-C) (Seo et al., 2005). In contrast, MCIDAS labeled the ncBAF components BICRA (also referred to as GLSTCR1) and BRD9, that were not enriched with GEMC1, as well as both BRG1 and BRM ATPase subunits (Figure 5A-C).

MCIDAS also showed a notable stronger interaction to the Cell cycle components (Figure 5D). These data indicated that although GEMC1 and MCIDAS both labeled Mediator to a similar extent, GEMC1 and MCIDAS preferentially associated with distinct SWI/SNF subcomplexes; namely the ARID1A containing BAF complex in the case of GEMC1 and the BRD9 containing ncBAF in the case of MCIDAS.

Figure 5

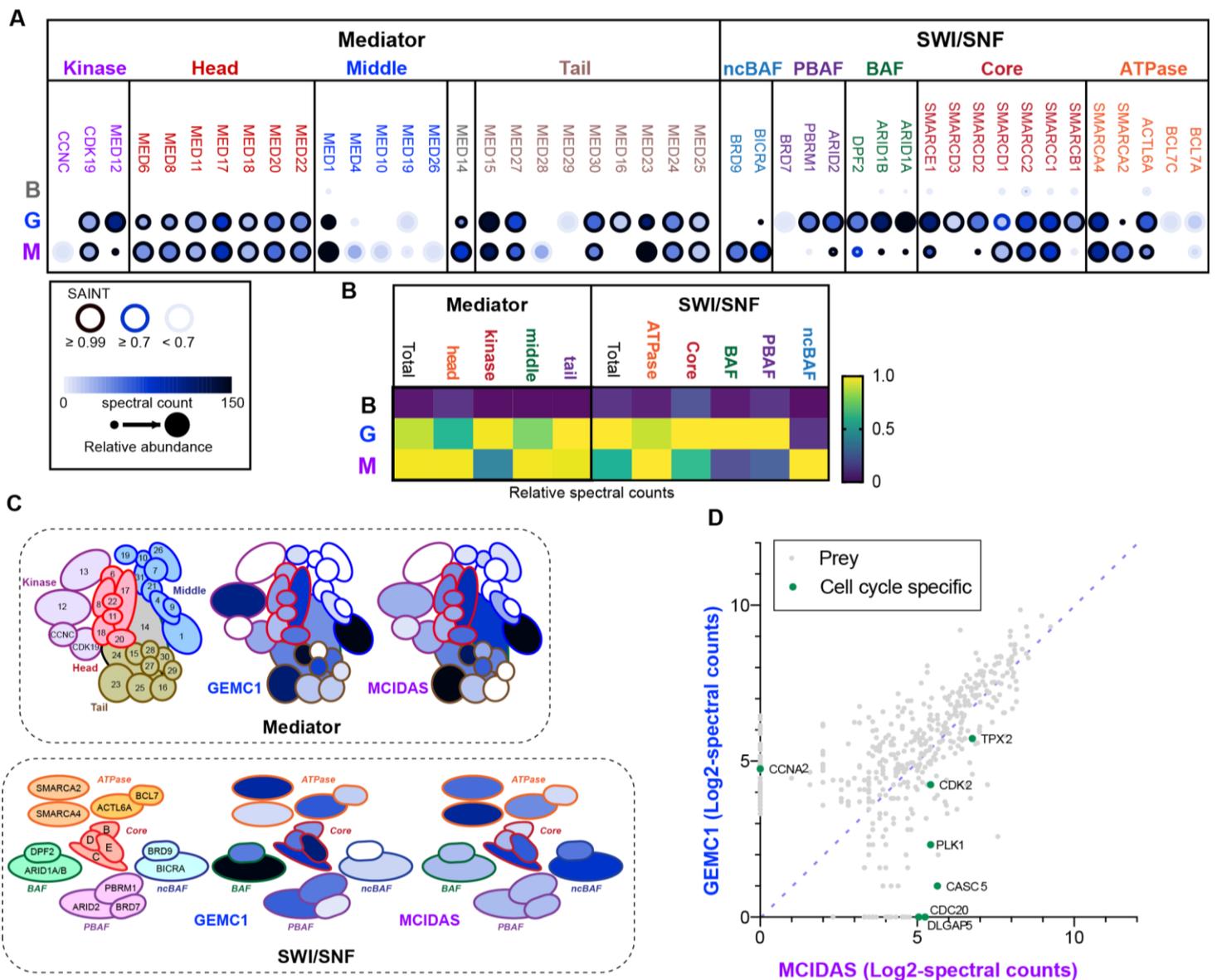


Figure 5. Interactions with Mediator and SWI/SNF. **A.** Dot plot depicting relative abundance, spectral counts and SAINT score for Mediator and SWI/SNF proteins. **B**=BirA*, **G**=GEMC1 and **M**=MCIDAS. **B.** Heatmap showing the relative abundance of the indicated subunits of either complex (groups are sum of components detailed in A). **C.** Schematic illustration of the Mediator and SWI/SNF complexes with spectral count data for each bait overlaid (colour key in A applied). **D.** Relative abundance of the Cell cycle proteins found between MCIDAS and GEMC1 BioID.

Sequence C-terminal to the CC domain regulates distinct SWI/SNF interactions

SWI/SNF nucleosome remodeling complexes occupy thousands of genomic loci and play important roles in transcriptional regulation during development, including the regulation of E2F target genes (Mittal and Roberts, 2020). As the enriched interactions between GEMC1 and BAF and MCIDAS and ncBAF emerged as one of the major differences, we sought to validate them and determine if they influenced the transcriptional activity of either protein. The most abundant interaction with a specific SWI/SNF component we identified was between GEMC1 and ARID1A. We therefore further examined this using BioID-AP followed by western. We observed ARID1A in Strep-purified lysates from both GEMC1 and MCIDAS proteins but enriched in GEMC1 samples, consistent with the BioID-MS data (Figure 6A). In contrast, we observed BRD9, a defining component of the ncBAF complex, highly enriched in MCIDAS samples, while the levels of core SWI/SNF components SMARCE1 (BAF57) and SMARCC1 (BAF155) were similar in each case and not identified in the BirA* controls (Figure 6A).

To try and understand what region of the protein conferred the BRD9 proximal interaction, we generated hybrid proteins by swapping the C-termini of GEMC1 and MCIDAS directly after the CC domain to generate GEM-IDAS (C-terminus of MCIDAS) and MC-C1 (C-terminus of GEMC1) (Figure 6B). In previous work, the C-terminal TIRT domain of MCIDAS was shown to have a higher affinity for E2F4/5-DP1 than GEMC1 when compared directly (Lu et al., 2019), consistent with our BioID results (Figure 3D). To determine if swapping the larger C-terminal domains also influenced this higher affinity interaction, we performed immunoprecipitations for GEMC1, MCIDAS, GEM-IDAS and MC-C1 using an HA-tag. As predicted, GEM-IDAS, containing the C-terminus of MCIDAS, immunoprecipitated higher levels of both E2F4-DP1 and E2F5-DP1 than GEMC1 or MC-C1 (Figure 6C).

We next examined the proximal interactions between the hybrid proteins and specific components of the BAF or ncBAF complex using BioID-AP westerns. GEM-IDAS again showed a similar profile as MCIDAS, interacting proximally with BRD9, and to a lesser extent with ARID1A. In contrast, MC-C1 did not exhibit detectable labeling of either specific component and showed reduced labeling of the core SMARCE1 component, despite showing similar levels of overall expression (Figure 6D).

To validate this further, we used the Proximity ligation assay (PLA) to examine the subcellular localisation of GEMC1, MCIDAS or the hybrid proteins, all tagged with FLAG, relative to BRD9. We observed that all of the proteins localised to the nucleus, in some cells forming discrete foci or localising diffusely throughout the nucleus (Figure 6E, lower panels). In PLA experiments, we observed signal generated by the proximity of the FLAG and BRD9 antibodies only in cells transfected with MCIDAS or GEM-IDAS (Figure 6E and F), consistent with the specific interaction of MCIDAS observed in BioID-MS (Figure 5A), as well as BioID-AP (Figure 6D).

Figure 6

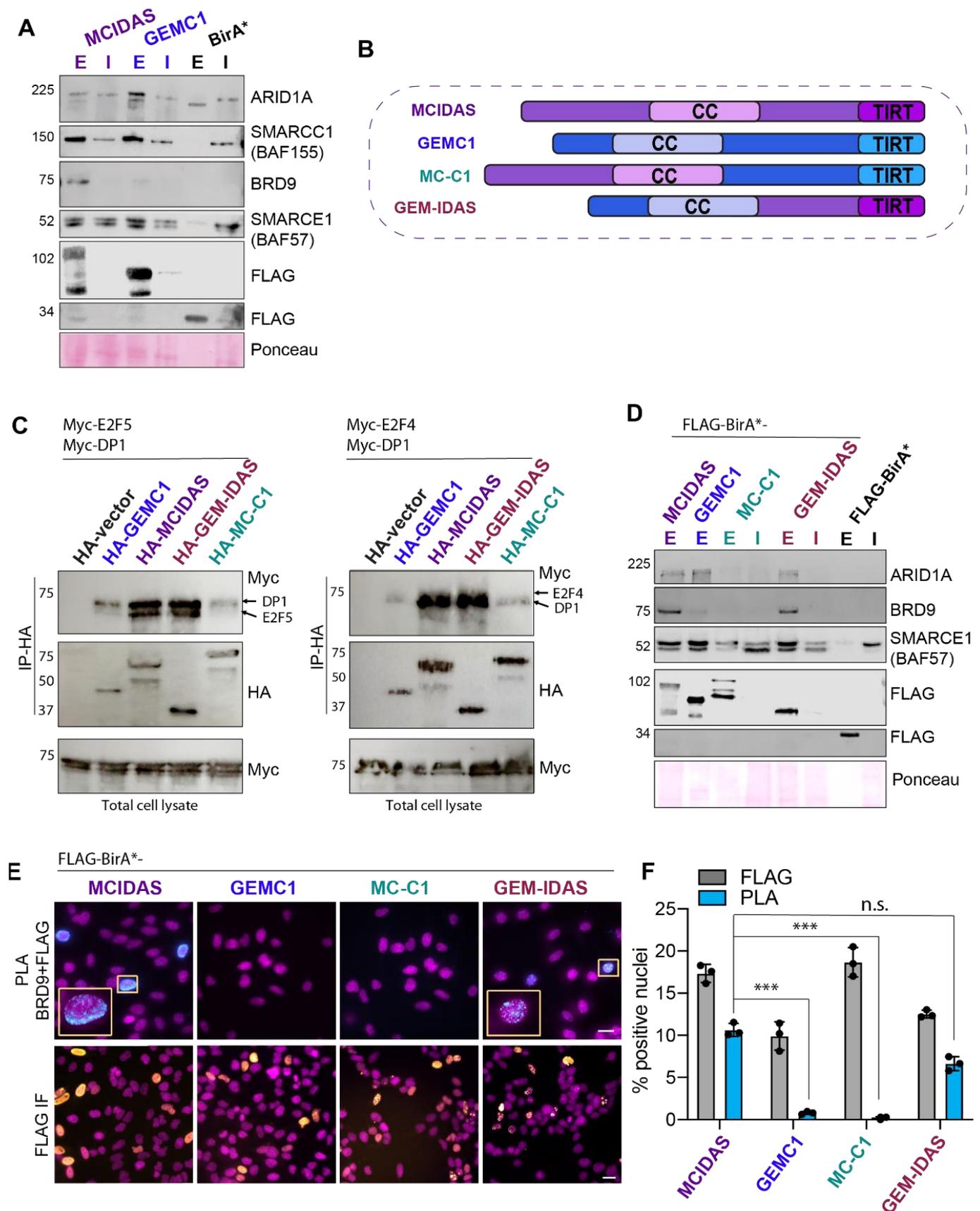


Figure 6: Specific interactions with SWI/SNF subcomplexes. **A**. BioID-AP western blots for ARID1A, BRD9 and core SWI/SNF subunits SMARCE1 (BAF155) and SMARCC1 (BAF57) demonstrate

the relative specificity for BAF and ncBAF complexes. **B.** Schematic illustration of the structure of GEMC1-MCIDAS hybrid proteins that exchange C-terminal domains after the CC domains. **C.** Co-immunoprecipitation experiments demonstrate enhanced E2F4/5-DP1 interactions with the MCIDAS C-terminus. AD293 cells were transfected with HA-tagged vector, GEMC1, MCIDAS, GEM-IDAS and MCI-C1, Myc-tagged DP1 and either Myc-E2F4 or Myc-E2F5. Lysates were immunoprecipitated with anti-HA antibodies and westerns carried out for Myc and HA following transfer to PVDF. Total lysates are shown blotted for Myc. Data presented is representative of 2 biological replicates. **D.** BioID-AP westerns blots for ARID1A, BRD9 and core SWI/SNF components following expression of hybrid proteins (see B for schematic). Ponceau staining shown for loading and transfer control. **E.** Representative immunofluorescence (IF) images of HeLa cells transfected with the indicated FLAG-BirA* tagged proteins (bottom panel, FLAG in yellow and DAPI (DNA) in magenta) and PLA-IF (top panels, PLA in cyan and DAPI (DNA) in magenta). Scale bar = 20 μ M. **F.** Quantification of the percentage of transfected HeLa cells and PLA positive cells. Mean (bar) with standard deviation and values of individual experiments (circles) are shown. For statistical analysis, a generalised linear model was performed with raw cell counts from 3 independent experiments. MCIDAS vs GEMC1, $p < 1e-06$ (***), MCIDAS vs MCI-C1, $p < 1e-06$ (***), MCIDAS vs GEM-IDAS, $p = 0.836$ (n.s).

TIRT domain swapping reverses SWI/SNF specificity

Given that the C-terminal of MCIDAS provides specificity towards SWI/SNF's ncBAF DNA binding domain - BRD9, we further probed what exactly what part of the C-terminal confers this result. To this end, we generated three more hybrid proteins (Figure 7A). Firstly, GEM-IDAS-C1, which is the GEM-IDAS hybrid, but with the TIRT domain of GEMC1. Secondly, GEM-IDA-C1, which is also like the GEM-IDAS hybrid, but with the TIRT domain of GEMC1 and the internal deletion of a predicted ordered region upstream from the TIRT within the MCIDAS sequence (leaving the disordered regions). Lastly, the MCI-C1 with the MCIDAS TIRT domain. We performed a Bio-ID-AP pull down on these plasmids after expression into AD293s. We found that having the MCIDAS TIRT domain put back into the C-terminal of MCI-C1 rescued BRD9 binding (Figure 7B). Contrastingly, the GEM-IDAS-C1 and GEM-IDA-C1 hybrid plasmids both registered no interaction with BRD9 or other components to the SWI/SNF complex (Figure 7C). These results were validated again through PLA experiments showing strong co-localisation between BRD9 and FLAG in MC-GEM-S but not GEM-IDAS-C1 (Figure 7D and E).

Together these results demonstrate that the C-terminal half of MCIDAS defines its SWI/SNF interactions and specifically its TIRT domain is responsible for its proximity to BRD9 containing ncBAF complexes. Furthermore, it indicates that both the N and C-terminal portions of GEMC1 are required for enhanced BAF-ARID1A interactions.

Figure 7

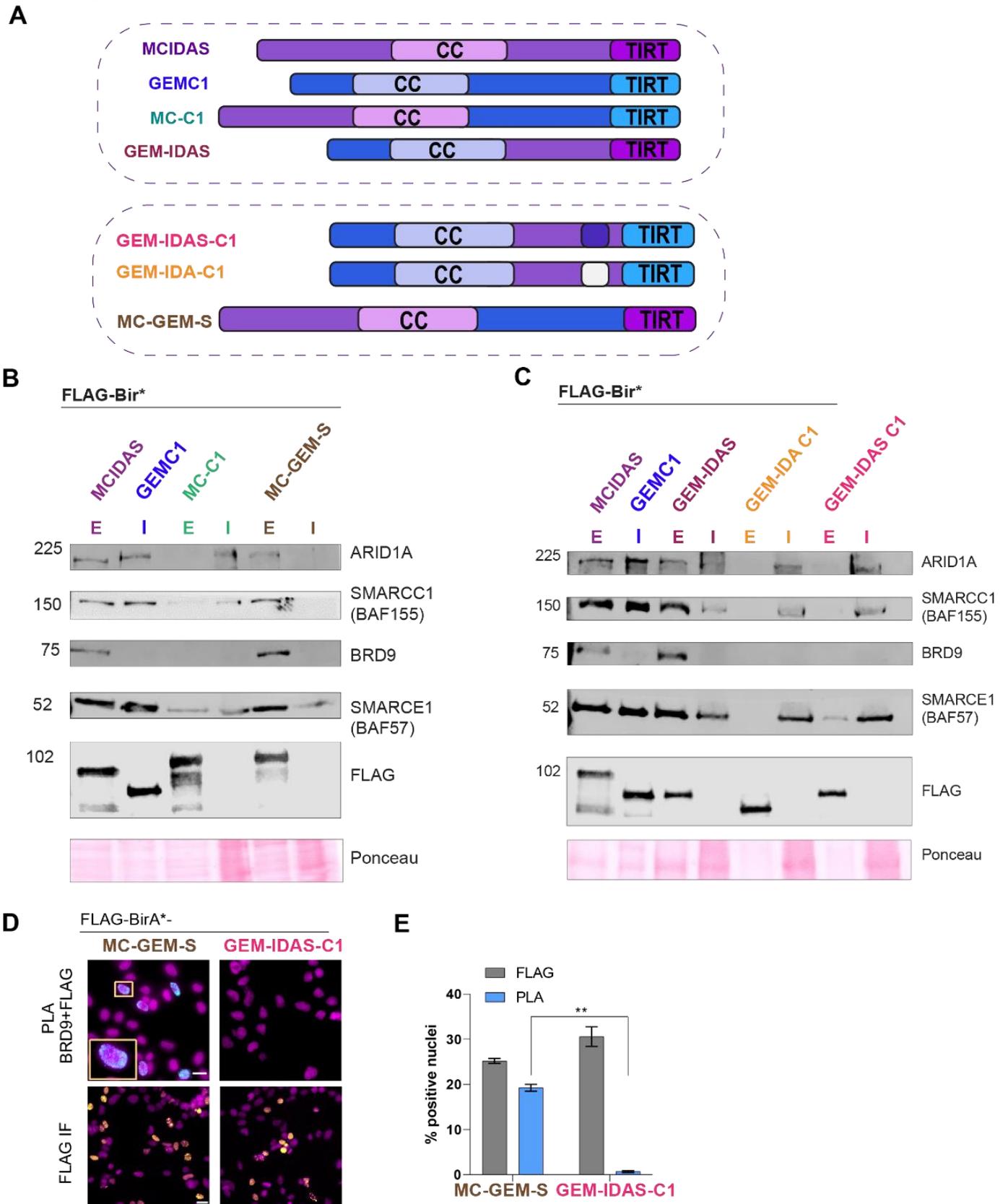


Figure 7: MCIDAS TIRT domain is required for the BRD9 interaction. A. Schematic illustration of the structure of GEMC1 and MCIDAS hybrid proteins that exchange C-terminal domains and TIRT

domains (Continuation of Figure 6). **B + C.** BioID-AP westerns blots for ARID1A, BRD9 and core SWI/SNF components following expression of hybrid proteins (see A for schematic) that demonstrate the relative specificity for BAF and ncBAF complexes. Ponceau staining shown for loading and transfer control. **D** Representative IF images of HeLa cells transfected with the indicated FLAG-BirA* tagged proteins (bottom panel, FLAG in yellow and DAPI (DNA) in magenta) and PLA-IF (top panels, PLA in cyan and DAPI (DNA) in magenta). Scale bar = 20 μ M. **E.** Quantification of the percentage of transfected HeLa cells and PLA positive cells. Mean (bar) with standard deviation and values of individual experiments (circles) are shown. For statistical analysis, a generalised linear model was performed with raw cell counts from 3 independent experiments, $p < 1e-05$ (**).

Distinct SWI/SNF complexes influence transcriptional activation

Given the clear interactions with SWI/SNF and their role in E2F-mediated transcription, we sought to determine if they were required for GEMC1 or MCIDAS-mediated transcriptional activity. We generated AD293 cells lacking ARID1A using CRISPR/CAS9 (Figure 8A) and analysed the ability of GEMC1 or MCIDAS to activate several common or enriched gene targets. In *ARID1A-KO* cells, neither protein could activate *FOXJ1* or *CCNO* following transfections (Figure 8B). In contrast, the activation of several genes preferentially activated by MCIDAS (Figure 2B), *CCDC96* and *HSPA1L*, were largely unaffected by ARID1A status (Figure 8C). Conversely, the activation of both of these genes was reduced following treatment with a small molecule inhibitor for BRD9 (I-BRD9, #SML1534) (Figure 7D) (Theodoulou et al., 2016). Treatment with I-BRD9 also strongly reduced the activation of *FOXJ1* by MCIDAS but had a minor effect on its activation by GEMC1 (Figure 8E).

As replacing the C-terminus of GEMC1 with that of MCIDAS resulted in a proximal interaction with BRD9, we analysed the ability of the hybrid proteins to activate *HSPA1L* and *CCDC96*, that were ARID1A independent but sensitive to I-BRD9. The GEM-IDAS fusion protein, that showed BRD9 interactions (Figure 6), activated both genes to a similar extent as MCIDAS, while the MC-C1 fusion protein showed only minor activation, similar to that of GEMC1 (Figure 8F). Together, these results demonstrated that ARID1A and BRD9 influence the transcriptional output of GEMC1 and MCIDAS. In addition, they predicted that ARID1A or BRD9 loss may affect subsets of GEMC1 or MCIDAS targets and impair multiciliogenesis.

Figure 8

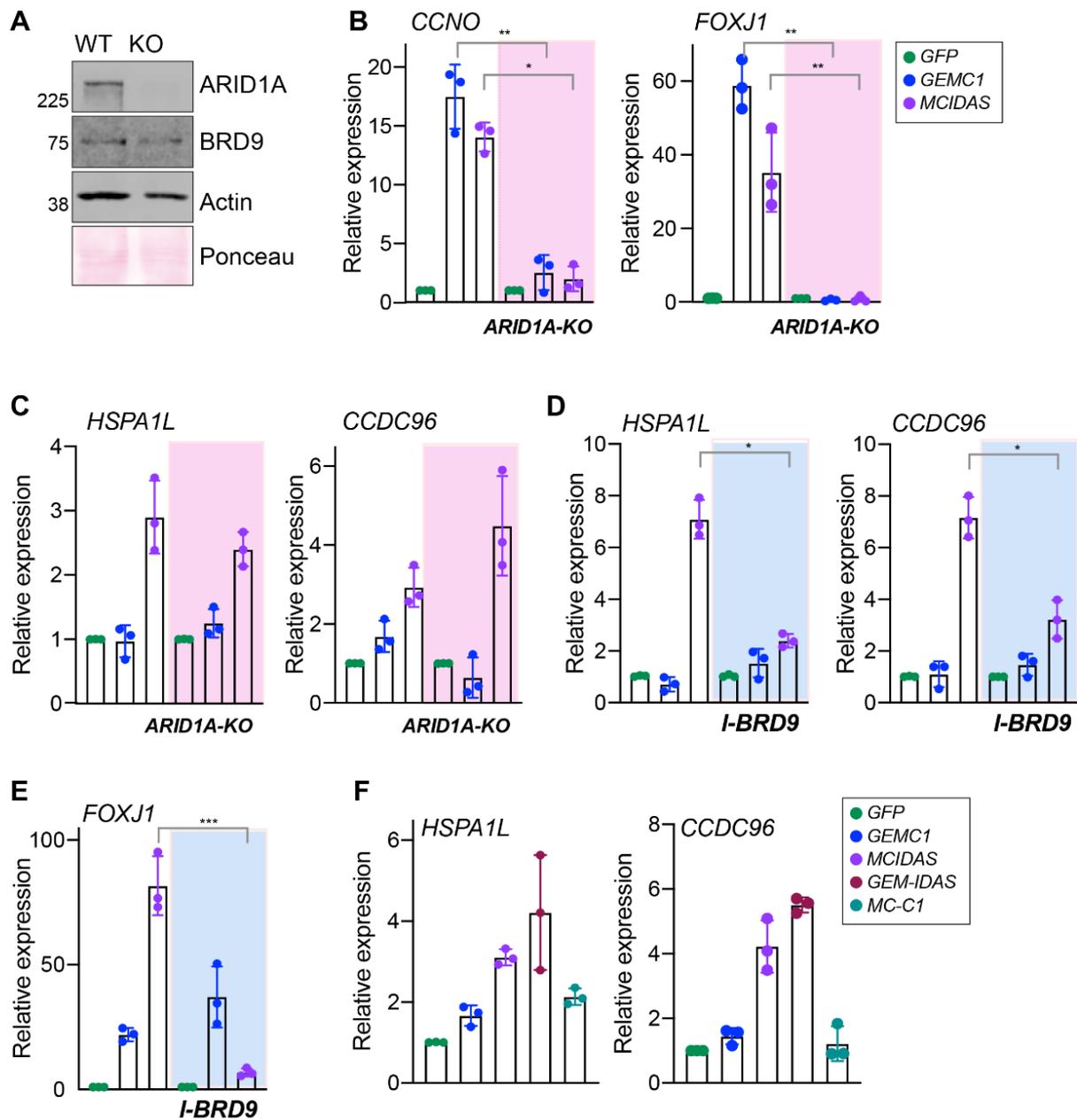


Figure 8. Transcriptional dependence on SWI/SNF subcomplexes. A. Validation of AD-293 ARID1A knock-out (KO) cells. Western blots for ARID1A and BRD9 are shown. Actin and Ponceau staining serve as loading and transfer controls. **B.** Quantitative real-time PCR (qRT-PCR) analysis of *CCNO* and *FOXJ1* expression in AD-293 cells following transfection with the indicated genes. WT vs ARID1A-KO p-values were 0.0027** and 0.0083** for *GEMC1* and 0.0142* and 0.0024** for *MCIDAS*. Transfected genes are colour coded in the key to the right. **C.** qRT-PCR analysis of *HSPA1L* and *CCDC96* in parental or ARID1A-KO AD293 cells. Statistical analysis as in B revealed no significant differences. **D.** qRT-PCR analysis of *HSPA1L* and *CCDC96* in parental or ARID1A-KO AD293 cells treated with 5 μ M of I-BRD9. WT vs I-BRD9 p-values were 0.0207* (*HSPA1L*), 0.0166*(*CCDC96*) and 0.0176*(*FOXJ1*) for *MCIDAS*. Colour coding as shown in panel B. **E.** qRT-PCR analysis of *FOXJ1* expression in AD293 cells following transfection with the indicated genes, $p=0.0003$ *** for *MCIDAS*. Colour coding as shown in panel B. **F.**

qRT-PCR analysis of *HSPAL1* and *CCDC96* expression in AD293 cells following transfection with the indicated genes. Transfected genes are colour coded in the key to the right. For B-F, results from 3 independent experiments are shown with the bars indicating the mean and the standard deviation shown. For statistical analysis a paired (B, C) or unpaired (D, E, F) t-test was used, $p \leq 0.001 = ***$, $p \leq 0.01 = **$ and $p \leq 0.05 = *$.

Inhibition of BRD9 inhibits Multiciliogenesis in an MTEC model

To test this possibility, we used Medullary thymic epithelial cells (mTECs) that can generate MCCs and other differentiated cell types characteristic of the airways when exposed to an Air Liquid Interphase (ALI) (Nanjundappa et al., 2019).

ALI cultures were treated with I-BRD9 during differentiation at 3 doses and examined for their ability to generate MCCs. Treatment with I-BRD9 reduced the expression of both *Foxj1* and *Trp73* at the mRNA and protein level in a dose-dependent manner (Figure 9A-C).

In addition, there was a dose-dependent reduction in cells exhibiting centriole amplification (inferred by staining for Centrin and DEUP1) (Figure 9A, D, E) and multiciliogenesis (inferred by staining with Acetylated tubulin antibodies (Ac-tubulin) (Figure 9A, F).

Together these data indicated that ncBAF function is essential for multiciliogenesis in murine ALI mTEC cultures.

Figure 9

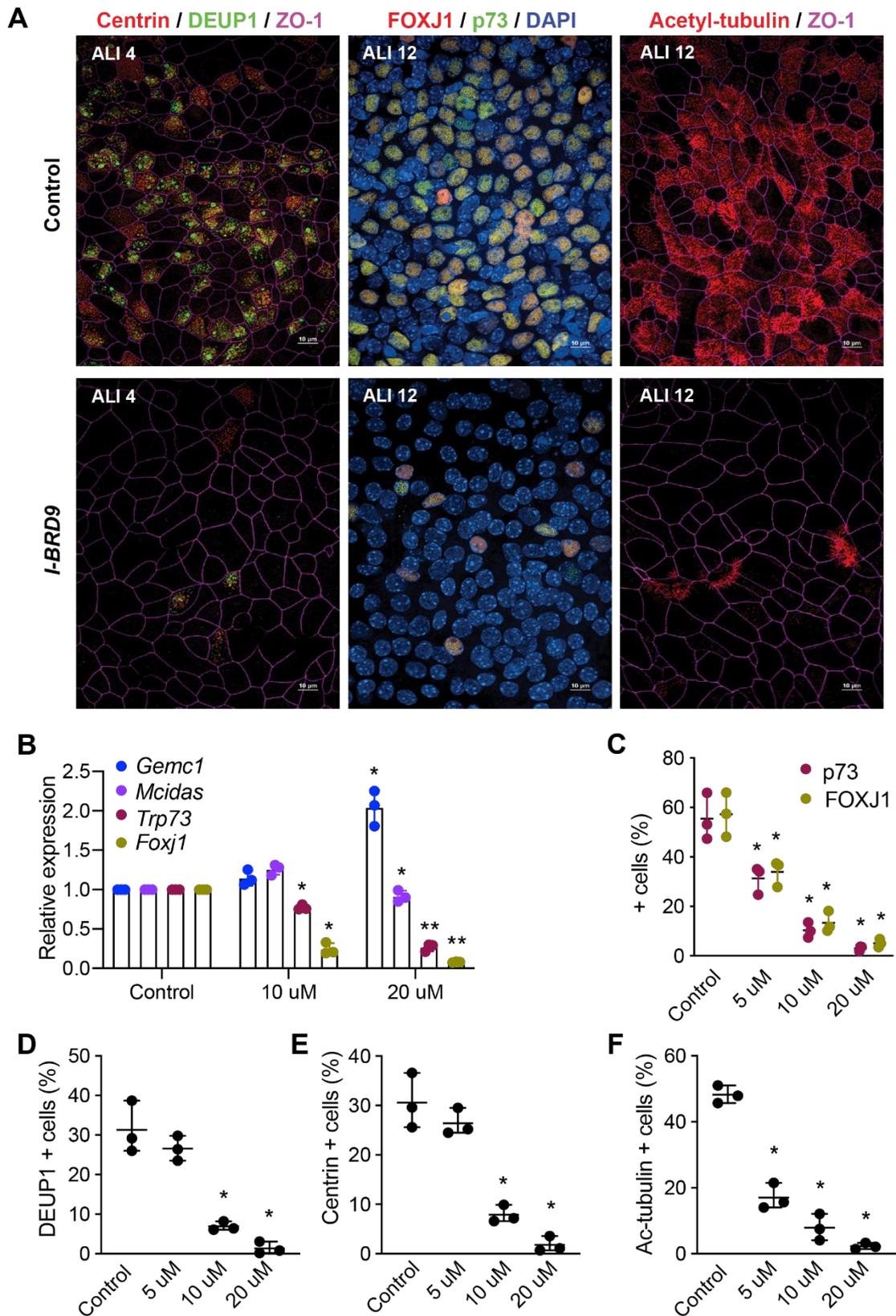


Figure 9: Requirement for BRD9 in multiciliation. **A.** Representative images of MTEC ALI cultures stained for Centrin and DEUP1 to analyse centriole amplification, FOXJ1 and p73 transcription factor targets of GEMC1/MCIDAS, or Acetylated tubulin (Ac-tubulin) to monitor ciliogenesis. Some images

stained with ZO-1 to define cell boundaries or DAPI to define the nuclear DNA. ALI culture day and treatment with 20 μ M I-BRD9 indicated. **B.** Quantitative real-time PCR analysis of *Gemc1*, *Mcidas*, *Trp73* and *Foxj1* in ALI cultures. Results from 3 independent experiments are shown with bar indicating mean and standard deviation. Paired t-tests were used for statistical analysis: Control vs. I-BRD9: *Foxj1* $p=0.0125^*$ (10 μ M) and 0.003^{**} (20 μ M), *Trp73* $p=0.0092^*$ (10 μ M) and 0.0087^{**} (20 μ M), *Gemc1* $p=0.0082^{**}$ (20 μ M). **C.** Fraction of cells committed to the MCC fate (p73+ or FoxJ1+) at ALI 12. Results are from 3 independent experiments. N (total cells) = control (2128), 5 μ M (1665), 10 μ M (1273), 20 μ M (1360). **D.** Fraction of cells undergoing centriole amplification (centrin-positive cells) or **E.** deuterosome formation (Deup1-positive cells) at ALI 4. Results are from 3 independent experiments. N (total cells) = control (1269), 5 μ M (1062), 10 μ M (1199), 20 μ M (1170). **F.** Fraction of ciliated cells at ALI 12. Results are from 3 independent experiments. N (total cells) = control (1814), 5 μ M (1295), 10 μ M (1113), 20 μ M (1099). Statistical analyses for C-F were done using one-way ANOVA. *denotes $p<0.05$.

BRD9 depletion inhibits multiciliogenesis in an inducible *in vitro* model

To further probe the specificity of interaction found between MCIDAS and BRD9 we created a cell model from Glioblastoma DBTRG-05MG cells that were able to generate multiple cilia through a Doxycycline-inducible system expressing either GEMC1 or MCIDAS after 7 days (Figure 10A and B). Using this time course we tested the system with addition of a proteolysis targeting chimera (PROTAC) degrader of BRD9 termed DBRD9 in MCIDAS expressing cells, to determine if it impacted Multiciliogenesis (Brien et al., 2018).

At D0, all cells showed one or no cilia, marked by acetylated tubulin (Figure 10B). At D3 there was a significant amplification in the production of centrioles marked by centrin staining, this is consistent with prominent acetylated tubulin presence at the surface of the cell, indicating budding multicilia. FLAG expression showed that GEMC1 and MCIDAS both primarily localised to the nucleus (Figure 10B).

Based on previous results that demonstrated specificity of interaction between MCIDAS and BRD9 and not GEMC1 we applied the DBRD9 degrader on the MCIDAS expressing cell line. Notably, DBRD9 inhibited the amplification of centrioles at this early stage in the MCIDAS inducible system. At D7 the cilia were fully extended with multiple extensions protruding per cell for both MCIDAS and GEMC1 (Figure 10B). Consistent with the impaired centriole amplification, treatment with DBRD9 almost completely prevented the production of multiple cilia in MCIDAS expressing cells.

Cells from each time point were counted and quantified and GEMC1 and MCIDAS showed the majority (60%) of FLAG expressing cells produced multiple centrioles (>20) and almost half of the cells in each

system resulted in MCCs (>10), whilst DBRD9 + MCIDAS significantly decreased both cilia and centriole amplification (Figure 9C).

Using the same system, we applied the PLA technique to visualize the interaction between BRD9 and either GEMC1 or MCIDAS (using FLAG) and the effects of this interaction with addition of DBRD9 (Figure 10D). After 3 days of induction MCIDAS can be seen with a number of foci representing the proximal interaction with BRD9. This was not observed in cells expressing MCIDAS or GEMC1 and treated with DBRD9 (Figure 10D, E). At D7, multiple cilia were fully produced and there remains a number of interactions between MCIDAS and BRD9 (average 6), while DBRD9 + MCIDAS and GEMC1 have significantly less interactions, averaging less than one per cell.

To further validate this interaction, we performed a BioID-AP western at the D3 time point (Figure 10F). We found that the DBRD9 successfully degrades BRD9 and this interaction was abolished in the Bio-ID pull down. In all, these results indicate BRD9 is a crucial regulator of MCIDAS mediated multiciliogenesis in this system.

Figure 10

Induction Model

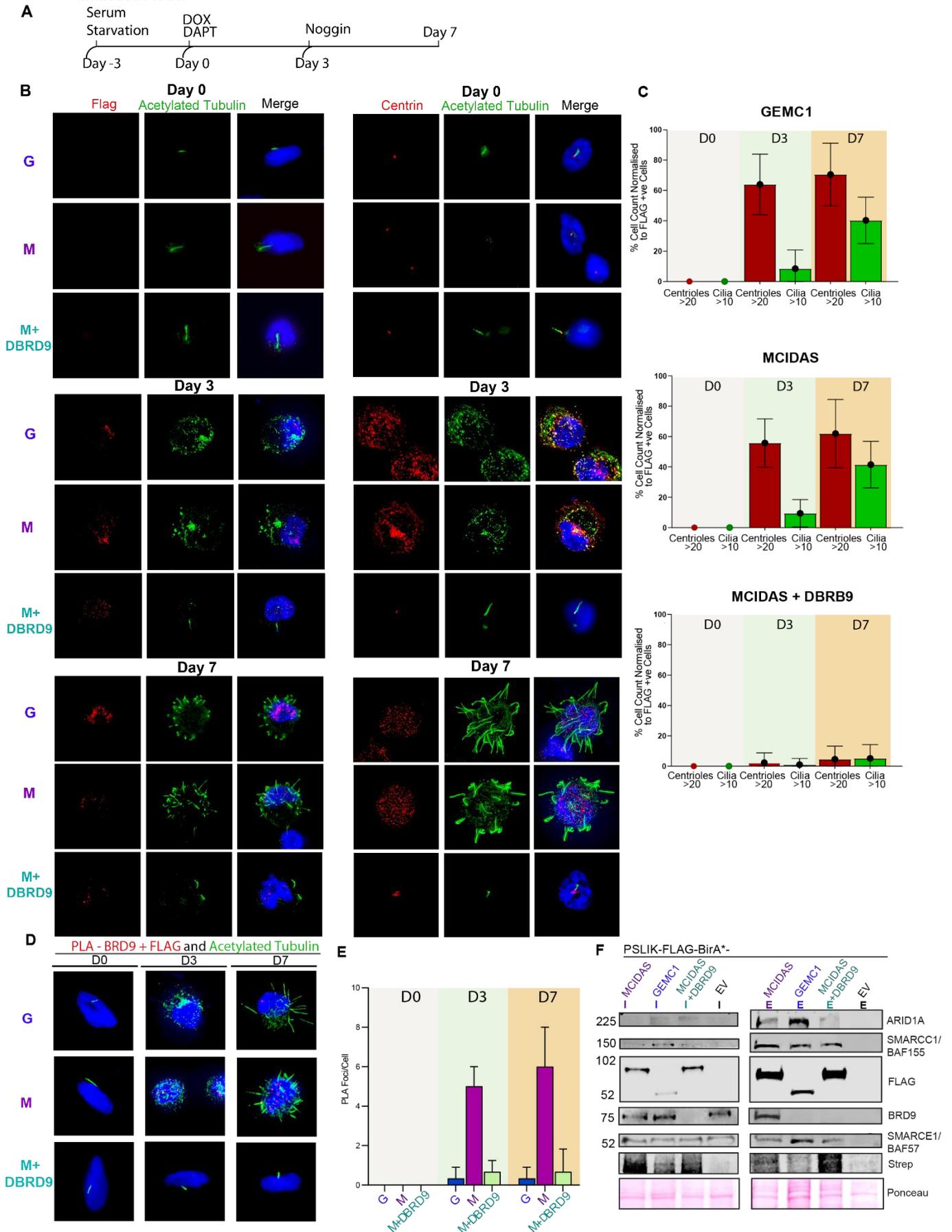


Figure 10: MCIDAS knockdown inhibits Multiciliogenesis in an *in vitro* cell model. **A.** Schematic demonstrating the time course of the inducible system used to generate MCCs from Glioblastoma DBTRG-05MG cells. For full details, see methods section. **B.** Immunofluorescence of the induction model with either G=PSLIK-FLAG-BirA-GEMC1, M=PSLIK-FLAG-BirA-MCIDAS or M+DBRD9=PSLIK-FLAG-BirA-MCIDAS + DBRD9 PROTAC (20uM). Images taken at three different time points D0, D3, D7 and stained with the indicated antibodies. **C.** Quantification of Immunofluorescence from induction models PSLIK-FLAG-BirA-GEMC1, PSLIK-FLAG-BirA-MCIDAS or PSLIK-FLAG-BirA-MCIDAS DBRD9 PROTAC (20uM). Centriole counts were derived from Centrin antibody and Cilia counts were derived from acetylated tubulin antibody, normalised as percentage from FLAG expressing cells. **D.** Proximity ligation assay (PLA) merged images taken at three different timepoints of the three inductions models described in B. Counterstained with Acetylated tubulin after PLA protocol was followed before mounting. **E.** Quantitation of D where Foci were counted per cell for each condition. N>100 cells counted per condition. **F.** Small scale BioID-AP followed by western blotting was carried out as previously described. Samples were taken from D3 of the induction models as described in B. I indicates Input and E indicates Eluate, western run separately but from same small scale BioID-AP.

Identification of CCNOs proximity interactome and influence on GEMC1 and MCIDAS

Previous work suggested that GEMC1 was negatively regulated by the poorly characterized Cyclin, CCNO (Cyclin O), potentially connecting CDK activity to its transcriptional functions (Funk et al., 2015). We sought to further understand CCNOs role in multiciliogenesis in context to its relationship with GEMC1 and MCIDAS. To explore this, we first performed microarray analysis on CCNO overexpression in AD293s, we found there to be a strong increase in expression of cell cycle specific genes and especially genes involved in the G2/M transition (Figure 11A and B)

To examine the interactors of CCNO we tagged the plasmid with FLAG-tagged BirA* and expressed it in AD293 cells supplemented with Biotin for 20 hours (Figure 11C). The proximal interactome of CCNO was identified and the majority of the top 25 protein interactors identified were involved in cell cycle regulation (Figure 11D). The GO molecular function search specifically highlighted regulators of the G2/M transition, such as CDC20, BUB1 and PLK1 which were also upregulated in the Microarray data (Figure 10E and 10A)

To determine any regulatory influence that CCNO may have on the interactome of GEMC1 and MCIDAS, we co-expressed CCNO in AD293 cells with either protein fused to BirA* and performed BioID-MS. We found that the overexpression of CCNO with GEMC1 greatly decreased the interaction with the majority of its transcriptional regulators in comparison to GEMC1 expression on its own (Figure 11F). This is in contrast to MCIDAS co-expression with CCNO, which showed an almost negligible change in its proximal interactome (Figure 11G).

Figure 11

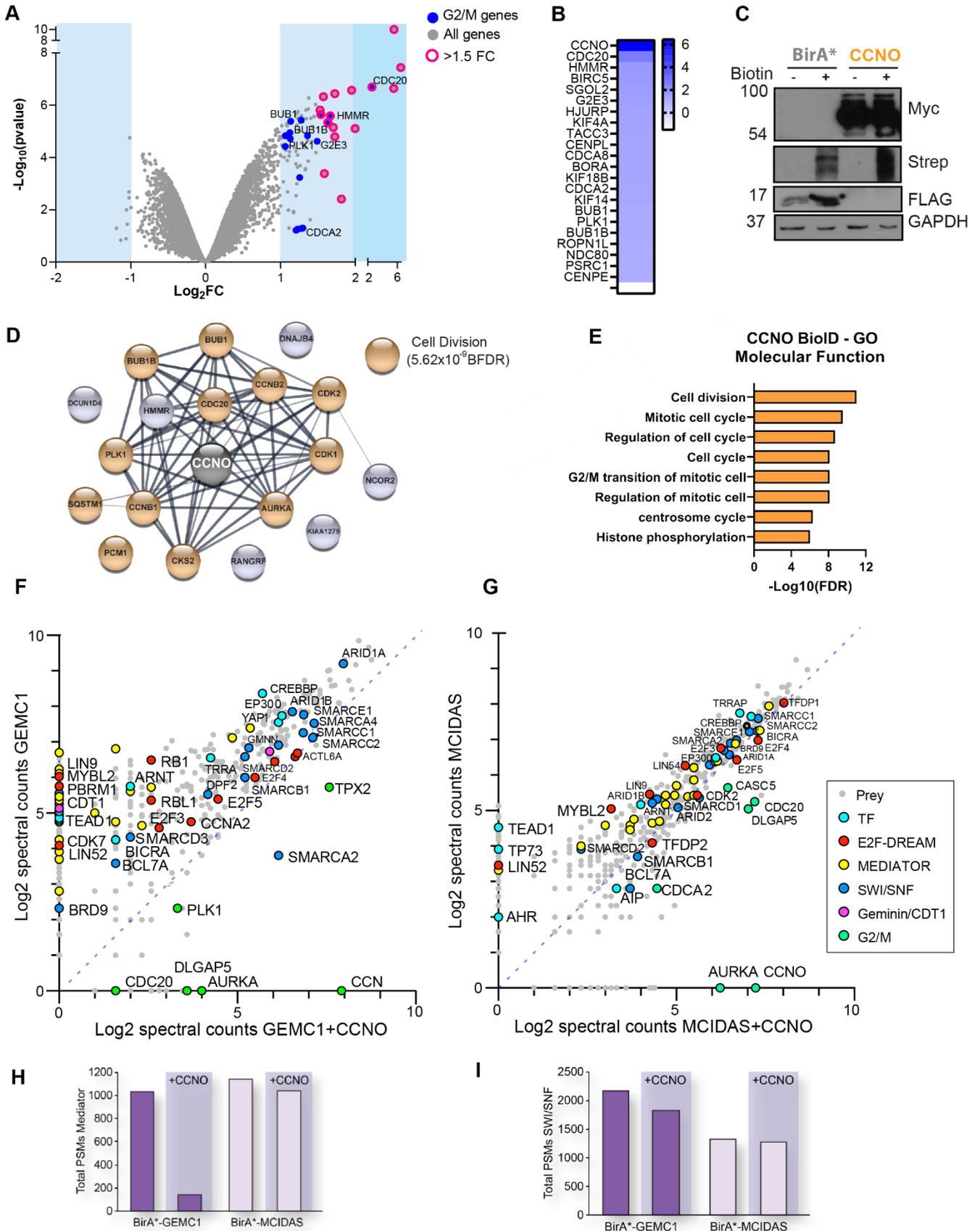


Figure 11: Transcriptional effects and proximal interactome of CCNO. A. Volcano plots ($-\text{Log}_{10}$ p-value vs Log_2 fold change) of microarray analysis of gene expression following transient transfection of

CCNO **B.** Gene set enrichment analysis of the gene expression data using GO-SLIM categories. The nominal enrichment score (NES) for each category in *CCNO* microarray **C.** Western blots from lysates of AD293 cells transfected with FLAG-BirA* or MYC-BirA*-*CCNO*. Expression of baits in cells with or without biotin supplementation is shown using the MYC antibody (top panel) and FLAG (second bottom) and labeling in biotin supplemented samples with Strep-HRP (middle panel). **D.** STRING network generated from top 25 proteins **E.** GO Molecular function of *CCNO* BioID top 25 proteins, Generated from STRING enrichment searches. BFDR categorization from STRING enrichment analysis. Network made in cytoscape. **F,G.** Scatter plot of peptides identified in GEMC1 or MCIDAS with *CCNO* co-expression (Log₂- spectral counts of 2 technical replicates). All peptides with a SAINT score of >0.7 in one of the samples are shown. Specific subsets are highlighted including key TF, components of the E2F or DREAM complexes, G2M, SWI/SNF and Mediator related proteins. **H,I.** As in D,E but Total peptide spectra matches (PSM) quantitated for each condition and complex.

Deeper analysis showed a substantial reduction in the interaction of GEMC1 with the Mediator complex but a less significant decrease in the interaction with E2F4/5 and SWI/SNF in the presence of *CCNO* that was not observed in MCIDAS + *CCNO* co-expression (Figure 11H, I and Figure 12A and B).

complex. Colour indicates the total number of spectral counts from 2 technical replicates. **B.** As in A but interaction mapping from the SWI/SNF chromatin remodeler with CCNO co-expression.

CCNO influences the expression of cell cycle components

As CCNO BioID and Microarray analysis highlighted specific G2M component interaction or upregulation respectfully, we sought to elucidate whether this occurrence was a consequence of a shift in the cell cycle or a direct effect of CCNO expression, or both. Using flow cytometry, we observed that CCNO overexpression in AD293 cells caused a reduction in cells entering mitosis in a time dependent manner (Figure 13A). Consistent with this, we observed reduced CDK1 levels in CCNO overexpressing cells (Figure 13B). We performed the BioID-AP followed by western blotting in CCNO transfected AD293 cells and validated a number of interactors including CDK2 and CDC20 (Figure 13C). We then applied a phosphorylated serine/threonine (p-S/T) antibody which identified a number of proteins that became phosphorylated after CCNO expression. The most abundant of these appeared around 62 kDa, which we hypothesised based on BioID data and size to be PLK1 (Figure 13C).

This was confirmed with a PLK1 specific antibody by a BioID-AP western blot which gave a band at the same size, as well as a slighter higher band, presumably resulting from Phosphorylation of PLK1. Reblotting with the p-S/T antibody confirmed this likelihood (Figure 13D). This indicated that CCNO overexpression in AD293s led to the phosphorylation of PLK1, potentially through the direct action of CCNO-CDK complexes. Based on these findings, we utilised phosphoproteomic techniques to examine the proteome wide effects of CCNO. We performed phosphoproteomic enrichment after BioID-AP of either CCNO or FLAG overexpression in AD293s (Figure 13E). Although PLK1 was not enriched in the CCNO sample above background controls, a number of PLK1 interactors were significantly phosphorylated as a consequence of CCNO overexpression. Whilst a number of significant G2/M proteins were enriched with CCNO overexpression that were not strongly present with FLAG like the CDK proteins (Figure 13E and F).

In all, these data indicate that GEMC1 and MCIDAS engage a similar core set of transcriptional machinery to activate transcription, but also differ in SWI/SNF subunit specificity and potentially other unique interactions to upregulate distinct gene sets. We hypothesize that CCNO provides negative feedback on GEMC1 through CDK2-mediated phosphorylation, and that the possible activation of PLK1 and depletion of CDK1 may allow cells to amplify centrioles without entering mitosis (Figure 13G).

Figure 13

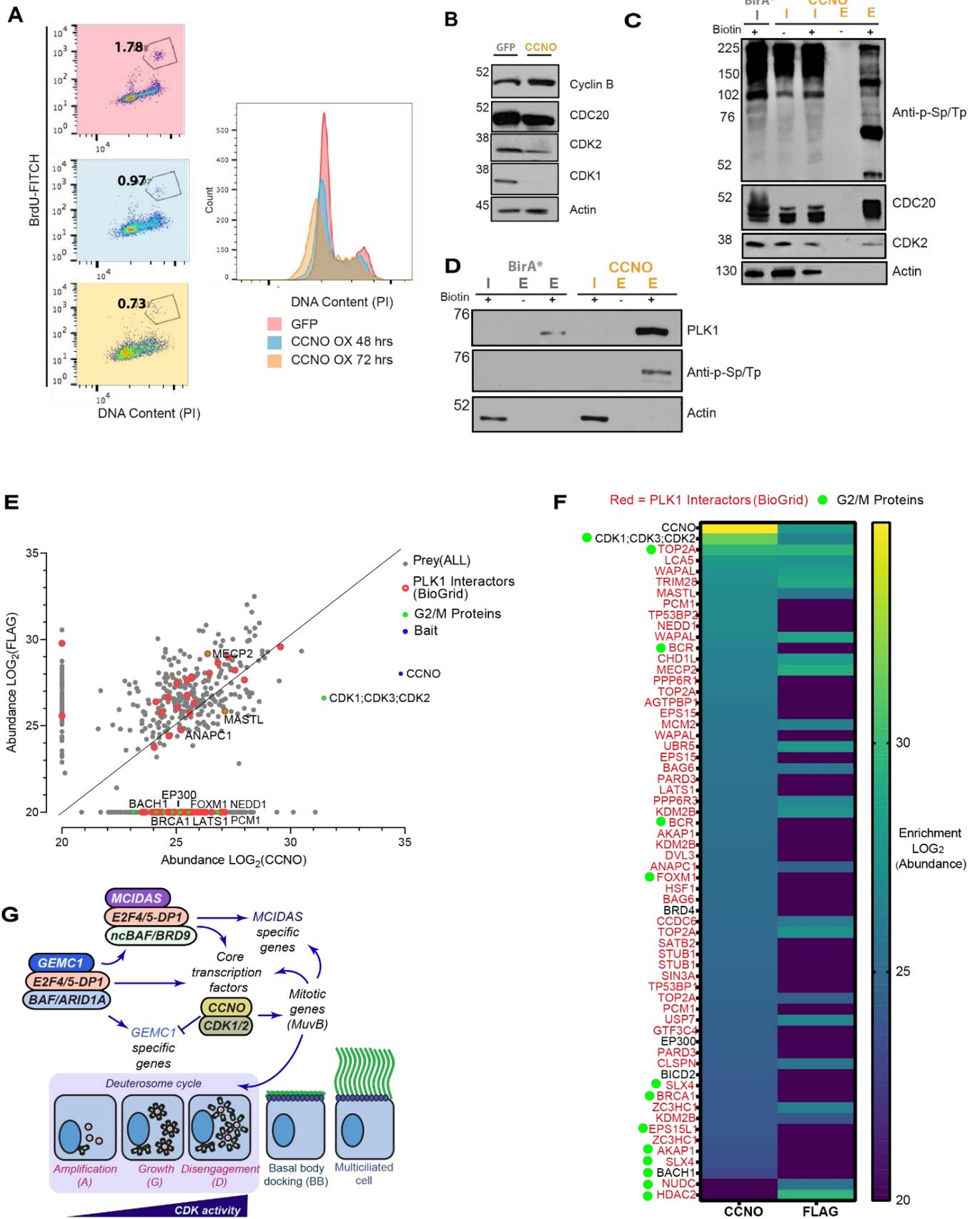


Figure 13: The influence of CCNO on the cell cycle. A. Flow cytometry of CCNO overexpression in AD293s after 48 and 72 hours stained with BrdU for proliferation and DNA (PI), gated area marks mitosis,

counts are quantified in right panel **B**. Western blot of overexpression of either CCNO or GFP in AD293 cells stained with indicated antibodies **C,D**. BioID-AP western blots for CCNO overexpression or Bir-A control in AD293s cells, probed with marked antibodies. Input (I) and Elute (E). **E**. Scatter graph of abundance of Phosphoproteomic enrichment of proteins from either BirA-CCNO or BirA-FLAG overexpression in AD293s after BioID-AP. Categorized by proteins inclusive of PLK1 interactors by Biogrid database or G2/M transition proteins according to GO search terms. **F**. Heatmap of most phospho-enriched proteins of CCNO from E. Repeat proteins occur from more than one phosphorylation site. Categorized by proteins inclusive of PLK1 interactors by Biogrid database or G2/M transition proteins according to GO search terms **G**. Schematic of hypothesized GEMC1 and MCIDAS activation of the transcriptional pathway in multiciliogenesis.

Potential roles of GEMC1 in Cancer

In vivo, GEMC1 and MCIDAS appear to be expressed only in MCCs. In addition to MCC differentiation, the Geminin family members and CCNO are expressed in different cancer cell types, suggesting their functions may be co-opted to support tumourgenesis, warranting further investigation (Balestrini et al., 2010; Pefani et al., 2011; Caillat et al., 2013; Li, 2018). We performed gene expression analysis using RT-PCR and observed that many cancer cell lines maintain a high level of GEMC1 and other members of the Geminin family, as well as their downstream target p73 (Figure 14A, B).

We wanted to determine the dependency of GEMC1 on cancer growth through siRNA mediated knockdown in different cancer cell lines. Two different siRNAs were utilised targeting GEMC1 and knockdown was confirmed by western blotting and RT-PCR analysis using the hTERT RPE-1 cell line, with siGEMC1.1 being a more potent oligonucleotide (Figure 14C, D). This knockdown strategy was applied to several cancer cell lines that showed the highest GEMC1 expression. Using the Alamar blue assay we observed a time dependent decrease in cell viability once GEMC1 was depleted across all cell lines compared to control (Figure 14E). Finally, the effect of the GEMC1 knockdown on the cell cycle was measured using flow cytometry in U2-OS cells. We observed a clear depletion in S phase and a shift to G1 phase with BrdU staining, indicating GEMC1s importance for DNA replication in these cells.

Figure 14

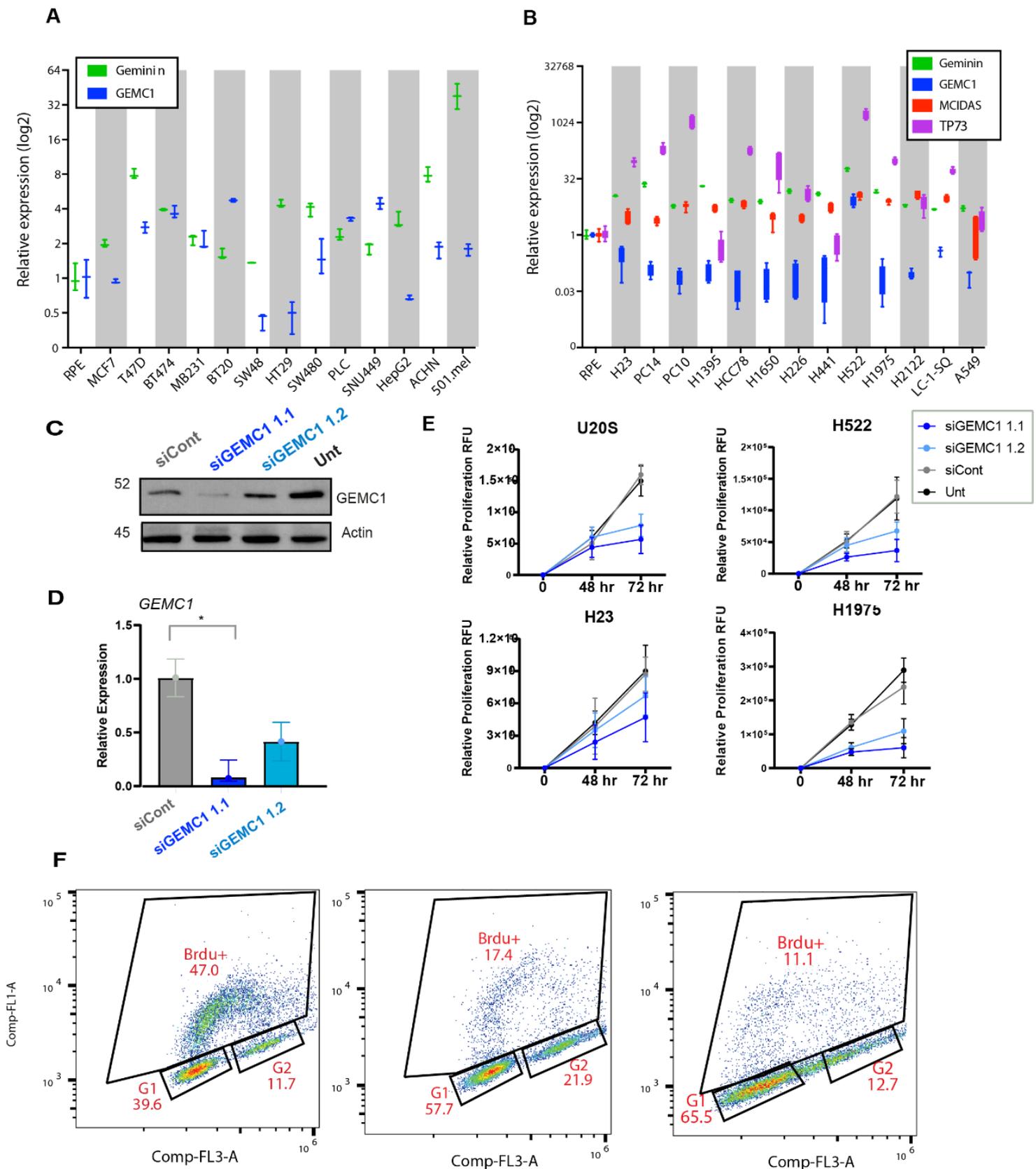


Figure 14: GEMC1 influences cancer cell growth. A.B RT-PCR of Geminin family genes quantified across a series of different cancer cell lines, Log2 relative expression normalised to hTERT RPE-1 cells.

(Taqman probe used) N=3 **C.** Western blot showing hTERT RPE-1 cell knockdown of GEMC1 using siRNA after 48 hours, transfected with two distinct oligonucleotides with lipofectamine. Antibodies used GEMC1 or actin. **D.** RT-PCR of GEMC1 (Taqman) of knockdown from C. **E.** Alamar blue assay of cell viability 48 hours and 72 hours after siRNA mediated GEMC1 knockdown. N=3. **F.** Brdu staining 48 hours and 72 hours after siRNA mediated GEMC1 knockdown, showing time dependent reduction in s phase replication.

Chapter II

Identification of Androgen Receptor proximal interactions in AD293s

As BioID proved to be a powerful method for identifying the transcriptional apparatus associated with GEMC1 and MCIDAS, we used it to examine proximal protein interactions with the AR, a transcription factor key to the pathogenesis of PC. The full length Wild-Type (WT) Androgen Receptor (AR) protein (22Q) was fused to a FLAG-MiniTurboID (MTID) tag by gene synthesis (Genscript) and a number of mutations were made in key domains (Figure 15A). The T1 mutant occurs in the DNA binding domain, the Tau1 and 5 mutations occur in the transactivation domain (TAD) and may influence recruitment of co-activators. The FQNLF motif is important for the interaction with the ligand-binding domain (LBD), and S767A has been shown to disrupt dimer formation (Nadal et al., 2017).

The expression and biotinylation of each mutant was confirmed through transient expression in AD293 cells (Figure 15A, B). Like many TFs, AR shows a high level of intrinsic disorder, especially in its NTD (Figure 15C). Correspondingly, the MTID-AR-WT is capable of forming nuclear foci or condensates that are morphologically similar to other AR-WT fusion proteins, with the addition of DHT in transfected AD293 cells, indicating the FLAG-MTID tag does not interfere with its ability to phase separate in this system.

The functionality of the MTID-AR-WT enzyme was tested in a pilot BioID-MS experiment using transiently transfected AD293 cells with 2 hours DHT and Biotin treatment. The proximal interactome of the top 75 AR hits consists largely of TFs, transcriptional coactivators/repressors, helicases and chromatin remodelers, including the SWI/SNF and Mediator complexes (Figure 15E).

Figure 15

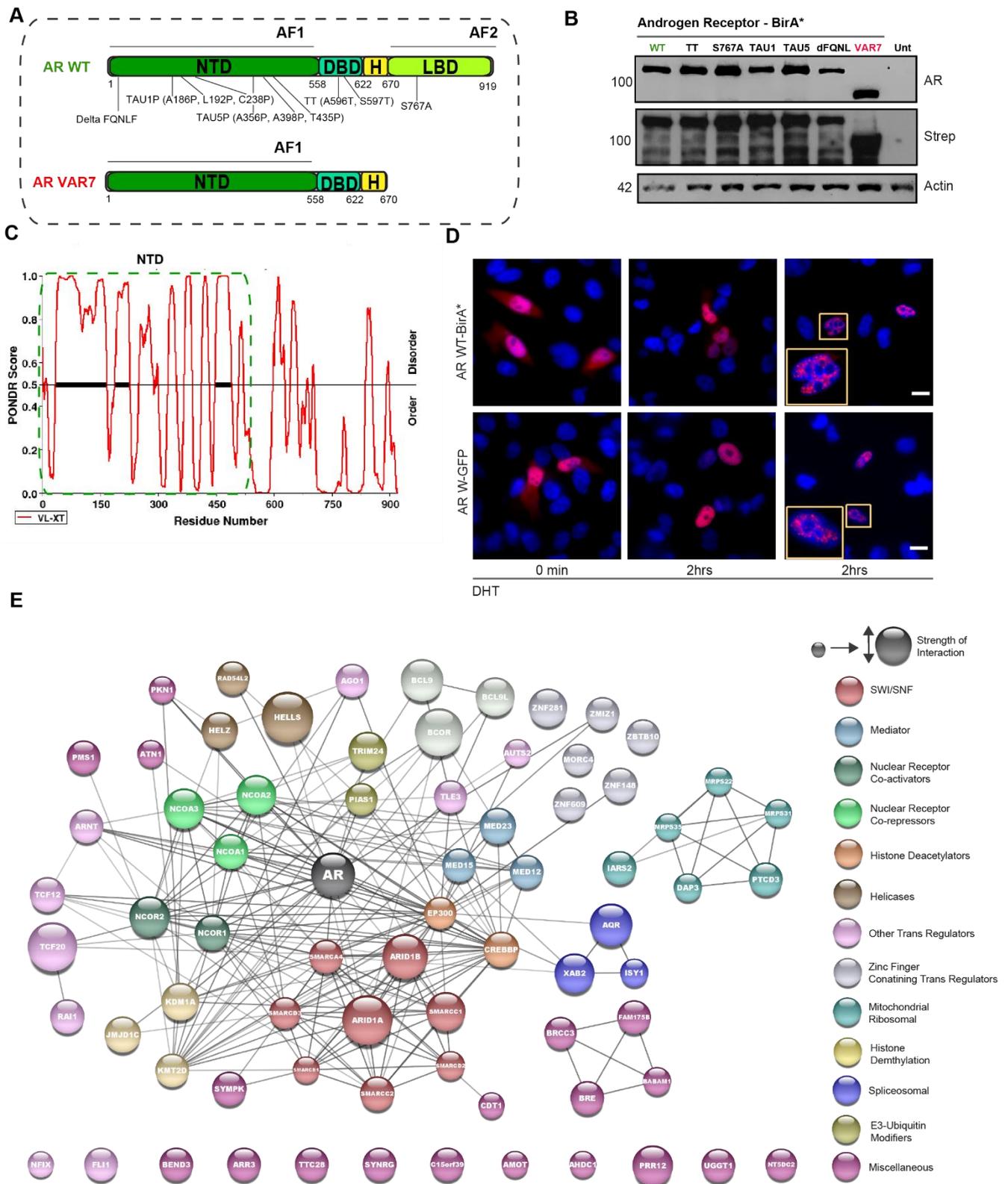


Figure 15: Androgen Receptor characterisation **A**. Schematic of Androgen Receptor (AR) domains and length, annotated with 6 specific mutations corresponding to the 6 AR-Mutants generated **B**. Western blot

showing expression of each AR mutant generated after over expression in AD293s, 48 hrs. Blotted with FLAG and Strep for biotinylation. **C.** Prediction of intrinsic disorder in the AR with NTD indicated. Sourced from <http://www.pondr.com/>. **D.** Representative immunofluorescence (IF) images of AD293 cells transfected with the indicated proteins with addition of DHT and stained with FLAG antibody and DAPI (DNA). Scale bar = 20 uM. **E.** STRING generated protein network of BioID-MS experiment of AR overexpression in AD293s, 48 hrs. Top 75 strongest interactors based on Tmean spectral counts and BFDR ≤ 0.02 , FC > 3. Size of node equates to strength of interaction by Tmean, colour coded for loose characterisation based on STRING GO search categorisations. DHT and Biotin 2 hrs. Made in Cytoscape. N=1.

Proximal interactions of AR-VAR7 in AD293s

Using the MTID-AR-VAR7 construct, we performed the BioID-MS experiment in AD293s with 2 hours of DHT and Biotin treatment. The top 50 interactors that were unique to VAR7 (FC > 15, BFDR < 0.2) were mapped using STRING (Figure 16A). Analysis with GO searches of the AR-VAR7 interactors highlighted clusters of proteins associated with RNA binding, mRNA processing and splicesomal type Biological Processes or Molecular Functions (Figure 16B and C).

The proximal interactomes of AR-WT and AR-VAR7 were compared in terms of the intensity of their interactions (Figure 16D). Generally, the AR-WT had a stonger intereaction in most of the prey proteins (BFDR < 0.05) in comparison to AR-VAR7. Of the known AR interactors (according to BioGrid database) and SWI/SNF subunits AR-WT also exhibited a stronger binding towards them compared to AR-VAR7. Whilst subunits of the Mediator complex showed no preference in their strength of interaction between the two proteins.

Figure 16

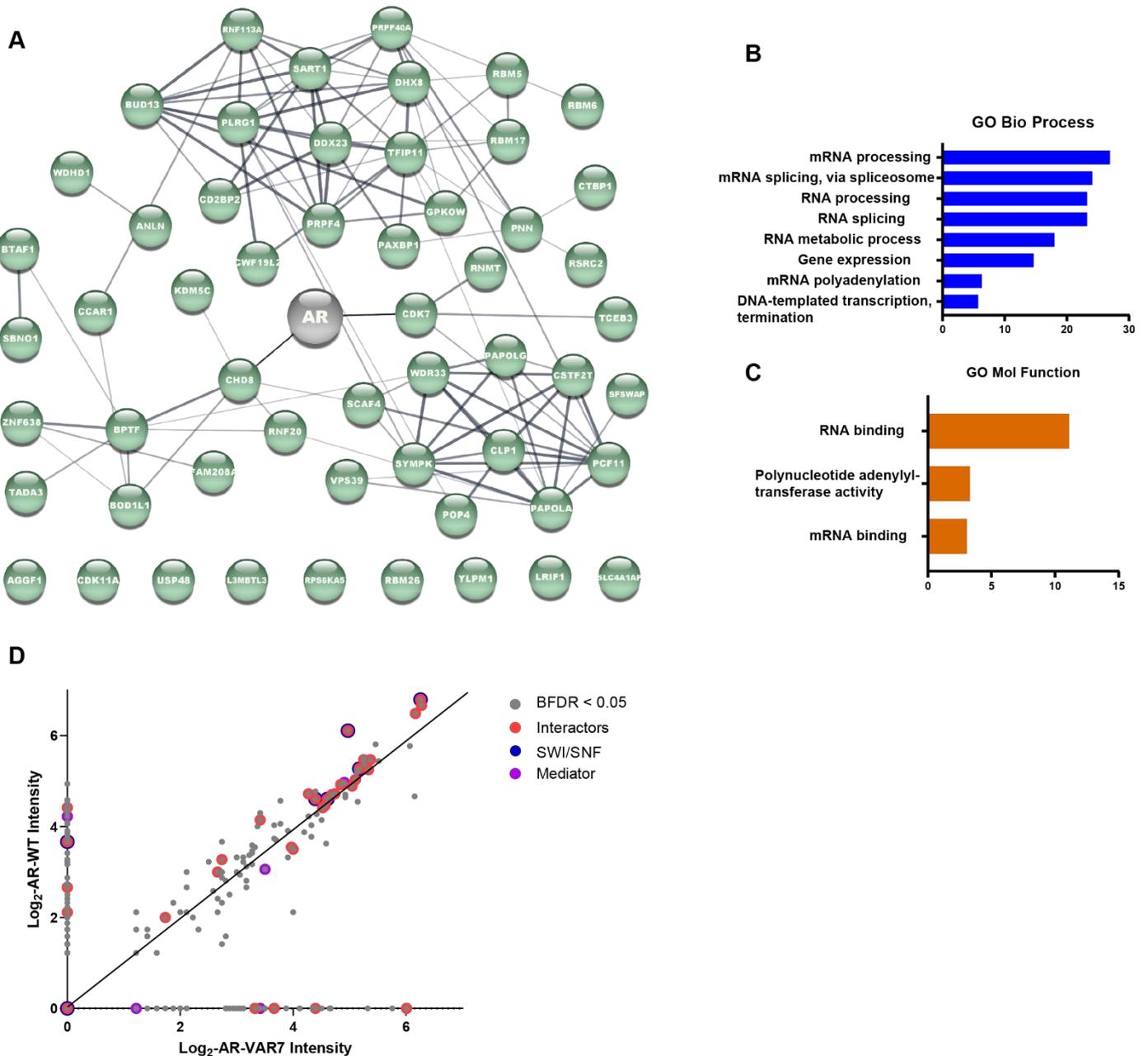


Figure 16: Androgen Receptor VAR7 Interactome in AD293. **A.** STRING generated protein network of BioID-MS experiment of AR-VAR7 overexpression in AD293s, 48 hrs. Top 50 strongest interactors of AR-VAR7, based on Tmean spectral counts and BFDR ≤ 0.2 , FC > 15 against AR-WT as control. DHT and Biotin 2 hrs. Made in Cytoscape. N=1. **B** and **C.** GO enrichment analysis searches of AR-VAR7 overexpression in AD293s network from A, generated from STRING database. **D.** Scatter plot of peptides identified in BioID-MS of AR-WT or AR-VAR7 samples (Log₂- SAINTe n=1). DHT and Biotin 2 hrs.

Proximal proteome of AR in stable cell lines

To ensure our results were not the result of overexpression, we successfully generated a number of stably expressing MTID-AR-WT cell lines using lentivirus transduction in several different prostate cancer backgrounds that are either AR positive, such as LNCaP (Figure 17A), or AR negative, such as PC3 and DU145 (Figure 17B, C). The proximal interactomes of the PC3 and LNCaP stable cell lines were determined from BioID-MS analysis and compared after the addition of DHT and Biotin for 2 hrs (Figure 17D). The strongest 100 interactors shown for LNCaP are strongly correlated with reported AR interactors in the BioGrid database, with the majority involved in transcription related activity, nuclear hormone receptor binding and chromosome organisation (Figure 17D, E).

Strikingly, there was a significant overlap of the total number and high confidence (BFDR ≤ 0.02) interactors shared between LNCaP and PC3 MTID-AR-WT cell lines, not just from such well-established partners such as FOXA1, EP300, CREBBP, but also in the intensity of interaction with the SWI/SNF and Mediator complexes (Figure 17F, G). In all, these data demonstrate the flexibility and effectiveness of proximal protein labelling with the MTID enzyme. Our results validated a large number of known AR interactors and will allow us to better understand AR and AR variant function across cell types.

Figure 17

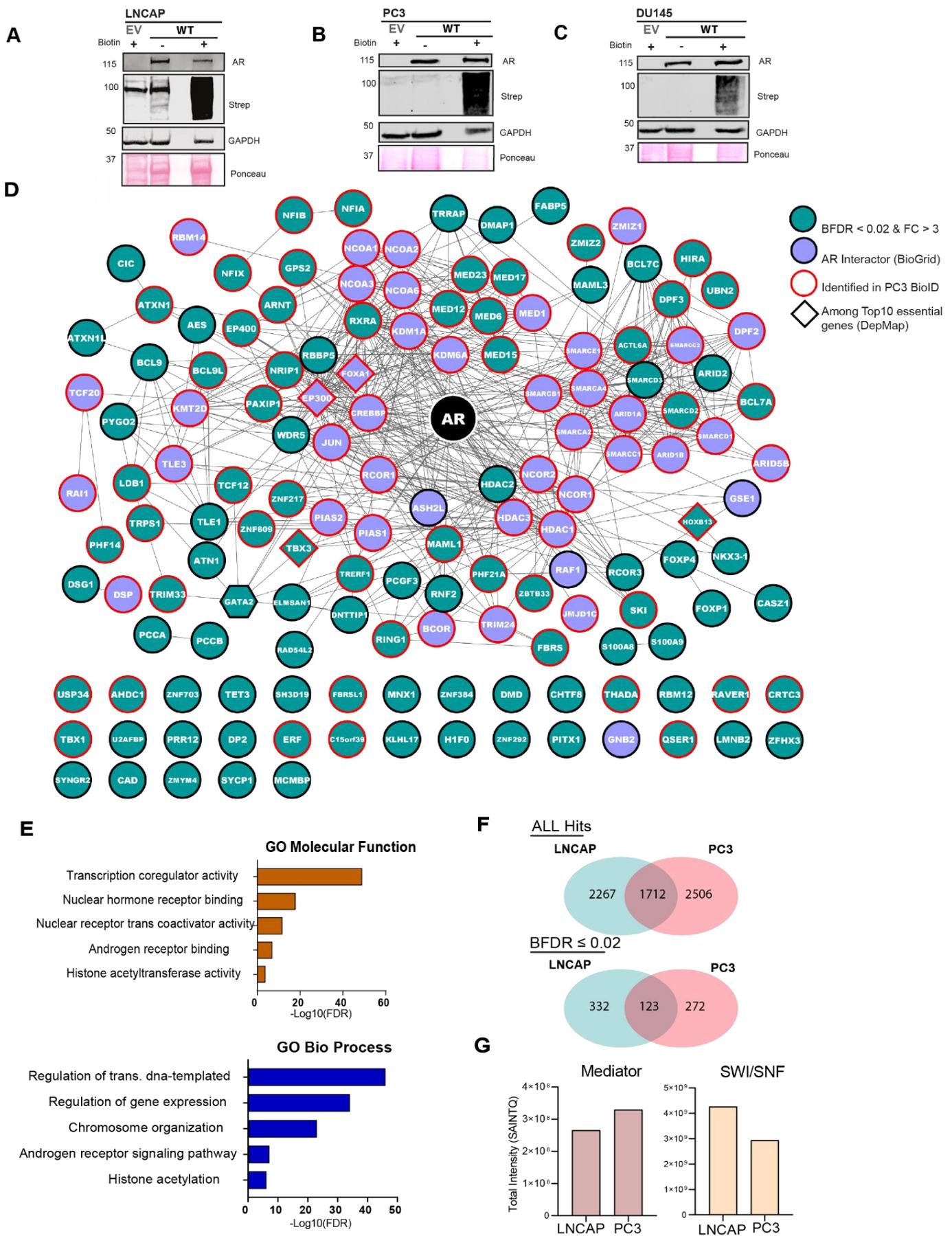


Figure 17: Analysis of proximal interaction in stable cell lines. A-C. Western blot showing expression of each MTID-AR-WT from generated Prostate cancer stable cell lines, LNCaP, PC3 and DU145 respectfully with or without biotin. Stained with AR or Strep antibodies. LNCaP panel shows background in biotin negative from biotin being present in growth medium of RPMI. EV=empty vector. **D.** STRING generated protein network of AR-FLAG-MTID stable cell line BioID-MS analysis of LNCA. Top 100 interactors and annotated with top PC3 interactors $\text{BFDR} \leq 0.02$, $\text{FC} > 3$. SAINTq of intensity from N=1. DHT, Biotin 2 hrs. BioGrid database used for known AR interactors and Depmap database of LNCaP for cancer dependency genes for Prostate Cancer annotated. Made in Cytoscape. **E.** GO enrichment analysis searches of LNCaP network from D, generated from STRING database. **F** Venn diagram highlighting overlap of proximal interactome from both AR-FLAG-MTID-WT in PC3 and LNCaP top hits as in D. **G.** SAINTq intensity values of total counts from FLAG-MTID-WT in LNCaP and PC3 cell lines of each complex.

Mutating ARs Tyrosine residues depletes protein interactions

AR condensates are stabilised by interactions between exposed aromatic residues, similar to condensates formed by various prion-like proteins (Li et al., 2018; Martin et al., 2020; Vernon et al., 2018). Previous research done by the Salvatella lab and others showed the importance of aromatic residues in the NTD, especially Tyrosines, of TFs in the propensity of the protein to phase separate *in vitro* and *In vivo* (Holehouse et al., 2021). We therefore generated two AR mutants that had either 8 or 22 Tyrosine (Y) to Serine (S) substitutions in the NTD to examine the functional and proteomic effects of depleting ARs aromatic character (Figure 18A). We fused both these mutants (8YtoS and 22YtoS) to a FLAG-tagged MTID enzyme and introduced them into PC3 cells using a lentiviral vector. Addition of biotin for 1 hour led to increased protein labeling, demonstrating that the MTID enzyme was functional (Figure 18B).

We carried out BioID-MS analysis of these Tyrosine mutants compared to MTID-AR-WT and collected samples after DHT treatment of 60 min. The SAINTq analysis revealed a significant reduction in the interactions of the majority of the strongest interactors of AR-WT in comparison to the 8YtoS mutant, although their remains a weak interaction with the transcriptional machinery (Figure 18C, and E). Enrichment analysis of the strongest interactors present in 8YtoS over AR-WT reveal according to STRING generated GO Bioprocesses - negative regulators of AR signaling, cell growth and post-translational modifications (Figure 18D, F and G). This suggests Tyrosine residues may play an important role in ARs interactome and warrants further investigation.

Figure 18

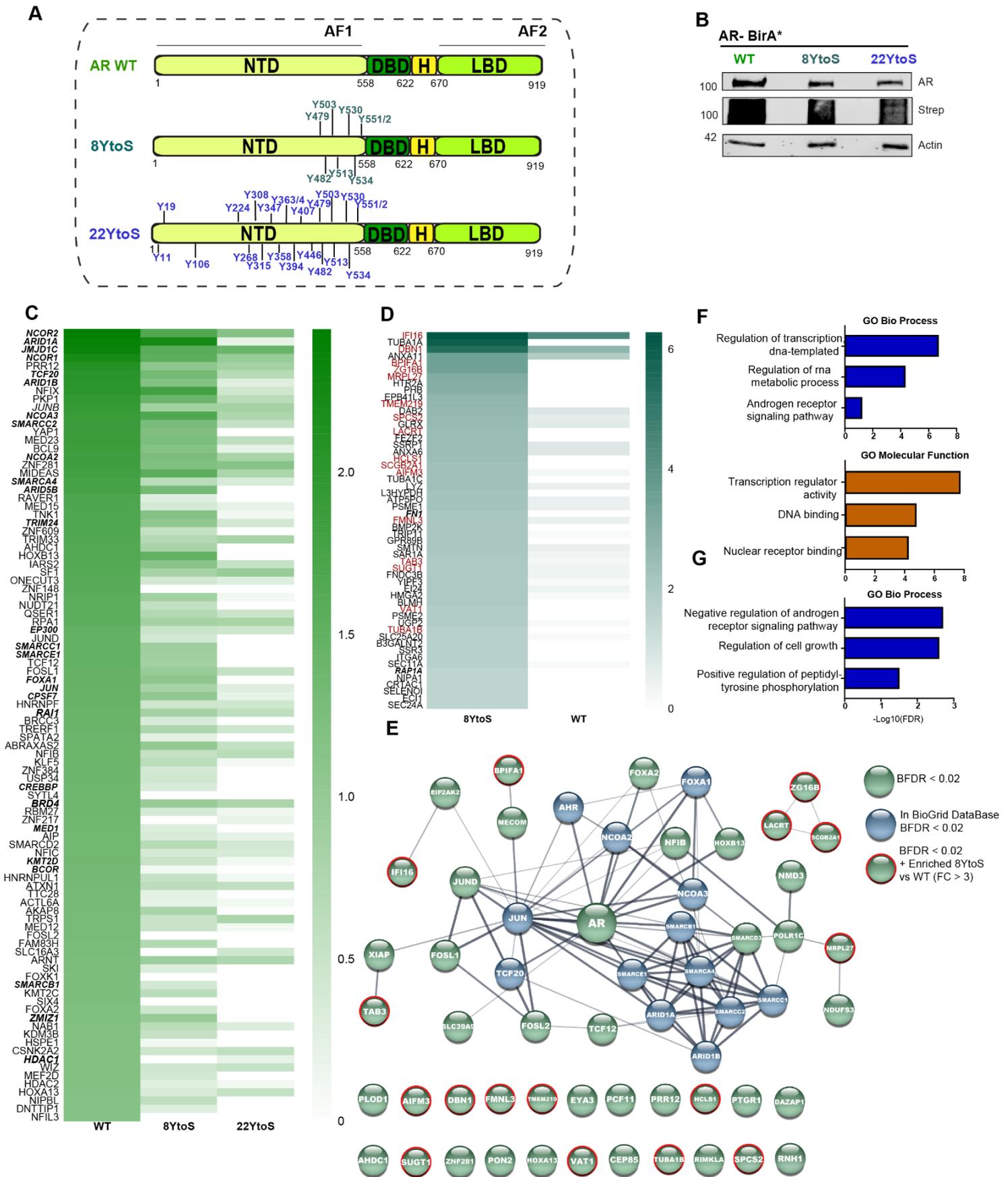


Figure 18: Mutating Tyrosine residues **A.** Schematic of AR domains and length (amino acid), MTID-AR-WT top and both Tyrosine MTID tagged mutants 8YtoS and 22YtoS below, annotated with their respective tyrosine to serine (YtoS) substitutions along the protein. **B.** Western blot expression of the generated PC3 stable cell lines MTID-AR-WT and both Tyrosine MTID AR mutants 8YtoS and 22YtoS below, stained with indicated antibodies. **C** Heat map generated from BioID-MS of top 100 AR-WT interactors $\text{BFDR} \leq 0.02$, $\text{FC} > 3$, compared to the Tyrosine mutants interaction of the same interactors. Scale represents strength of interaction of averaged SAINTq intensity values. DHT and Biotin 2 hrs. $N=2$. **D.** As in C but proteins listed as most enriched in 8YtoS compared to MTID-AR-WT, $\text{BFDR} \leq 0.02$, $\text{FC} > 3$. DHT and Biotin 2 hrs. **E.** BioID-MS of MTID 8YtoS AR mutant STRING network of SAINTq data, top 50 interactors based on intensity $\text{BFDR} < 0.02$, $\text{FC} \geq 1.5$. Made in Cytoscape. **F.** GO enrichment from STRING of 8YtoS network of E. **G.** GO Bio Process' of interactors enriched in 8YtoS compared to AR WT. $\text{BFDR} \leq 0.02$, $\text{FC} > 3$.

BioID-MS analysis of 22YtoS revealed an almost complete depletion of the AR-WT proteomic network (Figure 19A). Instead, 22YtoS identified several proteins related to nuclear transport with 5 nucleoporins identified amongst the top 50 most enriched proteins compared to AR-WT (Figure 19B).

Figure 19

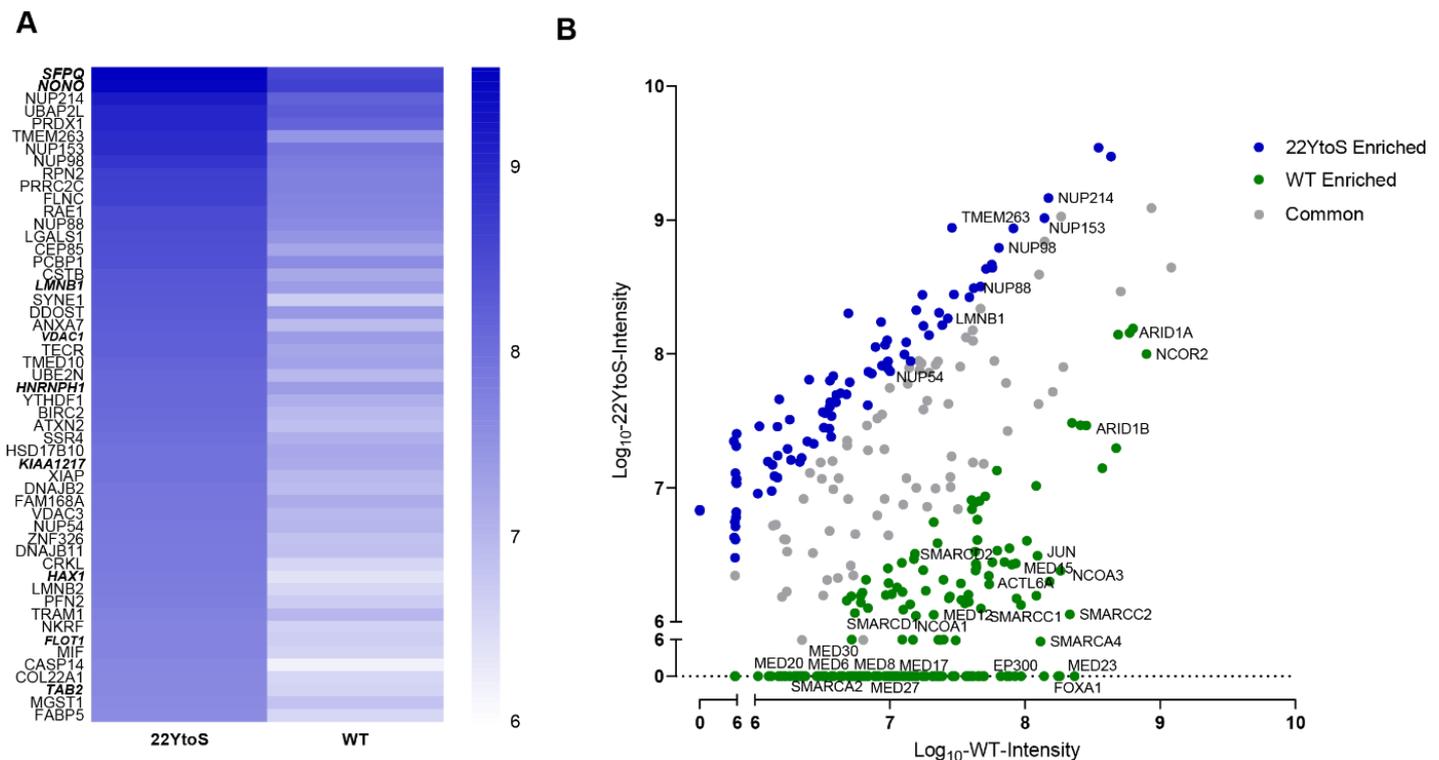


Figure 19: 22YtoS Tyrosine residues Interactome **A.** Heat map generated from BioID-MS of top 50 most enriched in 22YtoS compared to MTID-AR-WT, $\text{BFDR} \leq 0.02$, $\text{FC} > 3$. DHT and Biotin 2 hrs Scale represents strength of interaction of averaged SAINTq intensity values. DHT and Biotin 2 hrs. $N=2$. **B.** Scatter plot of peptides identified in BioID-MS of MTID-AR-WT or 22YtoS samples (Log_{10} - SAINTq $n=2$). DHT and Biotin 2 hrs.

22YtoS fails to translocate to the nucleus and interacts with Nucleoporins

To better interpret these results, we performed a series of immunofluorescence experiments with hormone treatment to visualize the relative localisation of AR-WT and 22YtoS in stably expressing PC3 cells. Without DHT treatment, we observed that AR-WT was present mainly in the cytoplasm with some nuclear localisation. After addition of DHT (60mins), almost all AR-WT had translocated to the nucleus, where anti-streptavidin antibody indicated most of the proximity interactions were taking place (Figure 20A). The 22YtoS did not follow this pattern of localisation, as addition of DHT caused the mutant AR to go from a mostly cytoplasmic localisation to a perinuclear staining, with some AR present in the nucleus but the majority of its interactions taking place on the nuclear membrane (Figure 20A).

The same hormone treatment in the PC3 AR-WT and 22YtoS cells was performed with samples taken for BioID-MS with and without DHT to determine how AR-WT responds to the ligand and to further examine the differences that may exist between their proteomic interaction. As expected, the addition of DHT (60mins) substantially increased the intensity of interactions between AR-WT and many of its core transcriptional partners such as EP300, SWI/SNF (Figure 20B). Whereas the overall interactome of 22YtoS did not change to any notable degree with DHT treatment, apart from further increasing the interaction strength of members of the Nucleoporin family (Figure 20C). Enrichment analysis of AR-WT protein interactors after DHT treatment revealed little overlap with 22YtoS, highlighting transcriptional or hormone related categories, whereas 22YtoS enrichment related to nuclear pore and cellular structure based ontologies (Figure 20D, E).

To validate these observations, we performed proximity ligation assays (PLA) for several of the top hits in AR-WT, including the SWI/SNF component ARID1A, the Mediator component MED1, and nucleoporin member NUP153. PLA signal was clearly evident for AR-WT with MED1 and ARID1A, whereas no interaction was observed for 22YtoS (Figure 20F). In contrast, 22YtoS showed clear interaction with NUP153 in the perinuclear space, but no interaction with MED1 or ARID1A (Figure 20F).

Figure 20

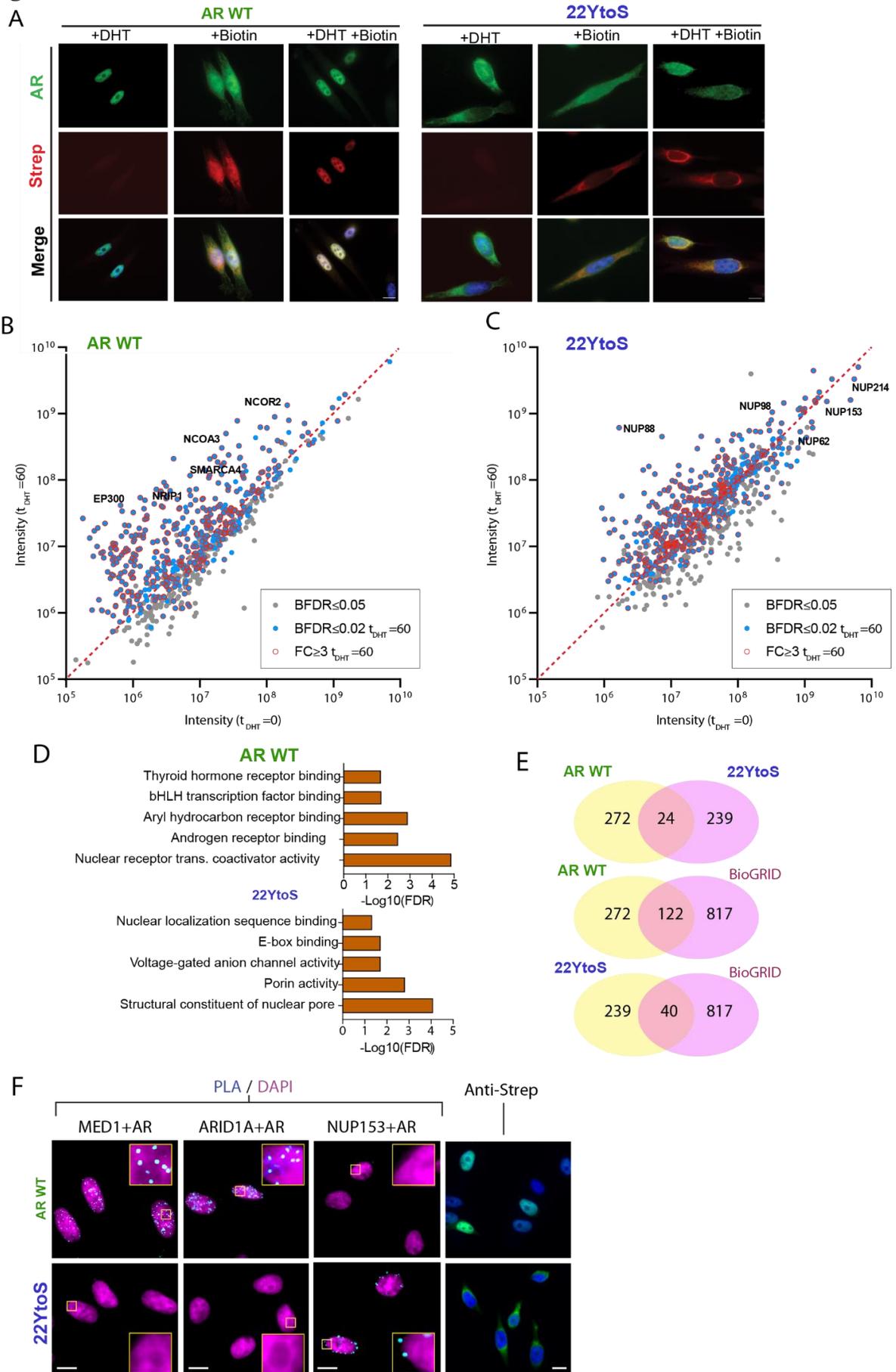


Figure 20: Translocation and the proximity interactome of MTID-AR-WT and 22YtoS. **A.** Representative immunofluorescence (IF) images of the MTID-AR-WT or 22YtoS stable cell lines with treatment of Biotin and DHT (2 hrs) or without. Stained with AR antibody and Streptavidin (Strep) for indicating potential interactors. **B,C.** Scatter plot of time course PPI network from BioID-MS in PC3 stable cell lines. Relative abundance of top proximity interacting proteins of AR-WT or 22YtoS ranked from SAINTq (BFDR \leq 0.02). T0 = Biotin \pm 50 μ M = 2 h, T60 = Biotin \pm 50 μ M = 1 h then tDHT = 1h). **D.** Enrichment analysis of B+C of top 70 PPI network ranked by abundance utilising SAINTq Data (BFDR \leq 0.02) of the AR-WT or 22YtoS respectfully. N=2. **E** Venn diagram of AR-WT or 22YtoS from B and SAINTq (BFDR \leq 0.02). BioGrid database was used for known AR interactors (<https://thebiogrid.org>). **F.** Proximity ligation assays (PLA) using the indicated antibodies are shown in cyan with DAPI staining in magenta in DHT treated PC3 cells. Streptavidin labeling is shown in green with DAPI in blue (far right panels) in DHT treated PC3 cells, scale bars 10 μ m.

AR Tyrosine mutants are transcriptionally impaired

The two Tyrosine mutants were investigated further with RNA-Seq to examine the potential transcriptional effects of their impaired protein-protein interactions. The PC3 stable cell lines were subjected to 2 hours of DHT treatment, whilst 22YtoS was treated with 2 hours DHT or 24 hours DHT to allow for a greater nuclear translocation time (Figure 21A, B). Globally, MTID-AR-WT and 8YtoS shared the most similar gene expression profile whilst 22YtoS and 22YtoS 24 hrs (DHT) were together the most distinct with a 10x greater number of downregulated differentially expressed genes (Figure 21B). Gene set enrichment analysis (GSEA) revealed that AR-WT was positively enriched for known AR targets (Figure 21C), which reached significance in weighted GSEA parameters ($p_{adj} < 0.01$) (Figure 21D). 8YtoS showed an overall slight negative correlation in GSEA analysis of the AR hallmark response, whereas 22YtoS and 22YtoS 24 hrs both exhibited a significant dysregulation in the known AR targets ($p_{adj} < 0.01$) (Figure 19D, E). Comparing AR-WT to the Tyrosine mutants directly highlights a number of pathways activated in AR-WT, including SWI/SNF subunit SMARCA4 targets and pathways hyperactive in CRPC, like E2F and MYC targets (Rasool et al., 2019; Takeda et al., 2018) (Figure 21F). In contrast, downregulated pathways of AR-WT against the mutants included stress response signaling and inflammatory responses, with 22YtoS showing the strongest dysregulation, indicating the longer DHT treatment in 22YtoS of 24 hrs rescued some typical AR transcriptional responses (Figure 21F). Taken together these results indicate that the Tyrosine AR mutants indicate a reduced propensity to phase separate, they are characterised by lower nuclear translocation rate, increased association with the nuclear pore, and reduced transcriptional activity in cells.

Figure 21

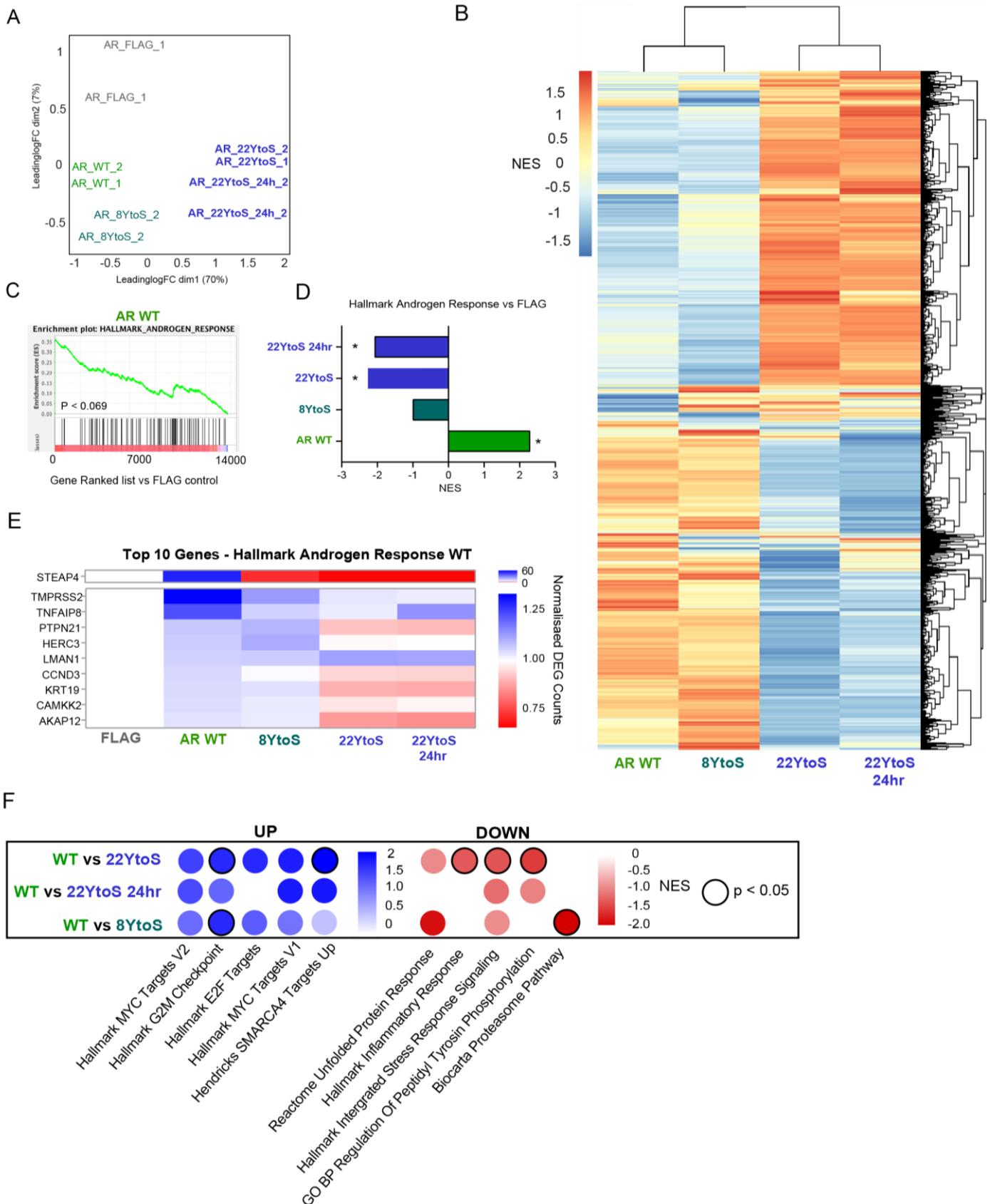


Figure 21: AR Tyrosine mutants Gene activation. A. Principle Component Plot of inputs from PC3 stably expressing MTID-AR-WT, 8YtoS, 22YtoS and 22YtoS 24 hrs (DHT) or FLAG control. **B.** Heatmap

of differentially expressed genes (DEG), indicated with NES=nucleotide enrichment score. Normalized to FLAG **C.** Random walk of the GSEA running enrichment score of Hallmark Androgen Response Pathway genes in AR-WT vs FLAG control, treated with DHT 2 hours. Weighted $p=1$. **D.** The GSEA running enrichment score of Hallmark Androgen Response Pathway in AR-WT, 8YtoS and 22YtoS, 22YtoS 24 hrs vs FLAG control, treated with DHT 2 hours. Weighted $p=1.5$. *= p value < 0.05 . **E.** Heatmap of normalised DEG Counts of Top 10 Genes in Hallmark Androgen Response in AR-WT vs FLAG compared to 8YtoS, 22YtoS or 22YtoS 24 hrs all vs FLAG. **F.** GSEA of enrichment score of a number of different pathways of AR-WT vs 8YtoS or 22YtoS or 22YtoS 24 hrs. Weighted $p=1$. Black ring indicates p value < 0.05 NES= Nucleotide enrichment score.

Tyrosine mutants with AR Variant 7 depletes protein interactions

To investigate whether the reduced engagement of the 22YtoS mutant with the transcriptional machinery and lower gene expression stemmed from its reduced propensity to phase separate, its aggregation in the nuclear membrane or its impaired nuclear localisation, we generated a new set of Tyrosine mutants in AR-VAR7 that is normally constitutively nuclear. Firstly, the AR-VAR7-22YtoS was generated, which lacks the LBD and may enable better translocation into the nucleus to examine the phase separation properties (Figure 22A). Secondly, we changed the 22 Tyrosines to Phenylalanines (F) to generate AR-VAR7-22YtoF, another aromatic residue that has a weaker tendency to reduce phase separation and can therefore operate as an intermediate or partial rescue of the S mutations (Figure 22A). Lastly, as a control for Phenylalanine, we generated the same plasmid in the WT background, AR-WT-22YtoF, to determine if there were effects of utilising Phenylalanine on nuclear translocation.

We again generated stable PC3 cell lines and tested the conjugated FLAG-MTID expression and enzyme biotinylation (Figure 22B). The translocation of the AR-VAR7 Tyrosine mutants was examined by immunofluorescence. MTID-AR-WT showed clear translocation to the nucleus upon DHT treatment, whilst deconvoluted images show clear condensate like foci in the nucleus (Figure 22C, D). The VAR7 remains largely nuclear without hormone treatment as expected, also with similar condensate formation (Figure 22C, D, E). AR-WT-22YtoF showed very similar translocation properties as the WT, while AR-VAR7-22YtoF maintained some cytoplasmic staining even with DHT treatment (Figure 22C, E). AR-VAR7-22YtoS demonstrated primarily nuclear localisation with some residual cytoplasmic staining and less clear or more aggregated like condensate formation (Figure 22D). The AR-VAR7-22YtoS much like AR-VAR7 localisation, did not significantly change with hormone treatment (Figure 22C, E).

As all stable cell lines show no major translocation defects, the next steps will be to take these forward for analysis with BioID- MS to determine whether altering the aromaticity translates into inhibited AR interaction with the VAR7 form of AR too.

Figure 22

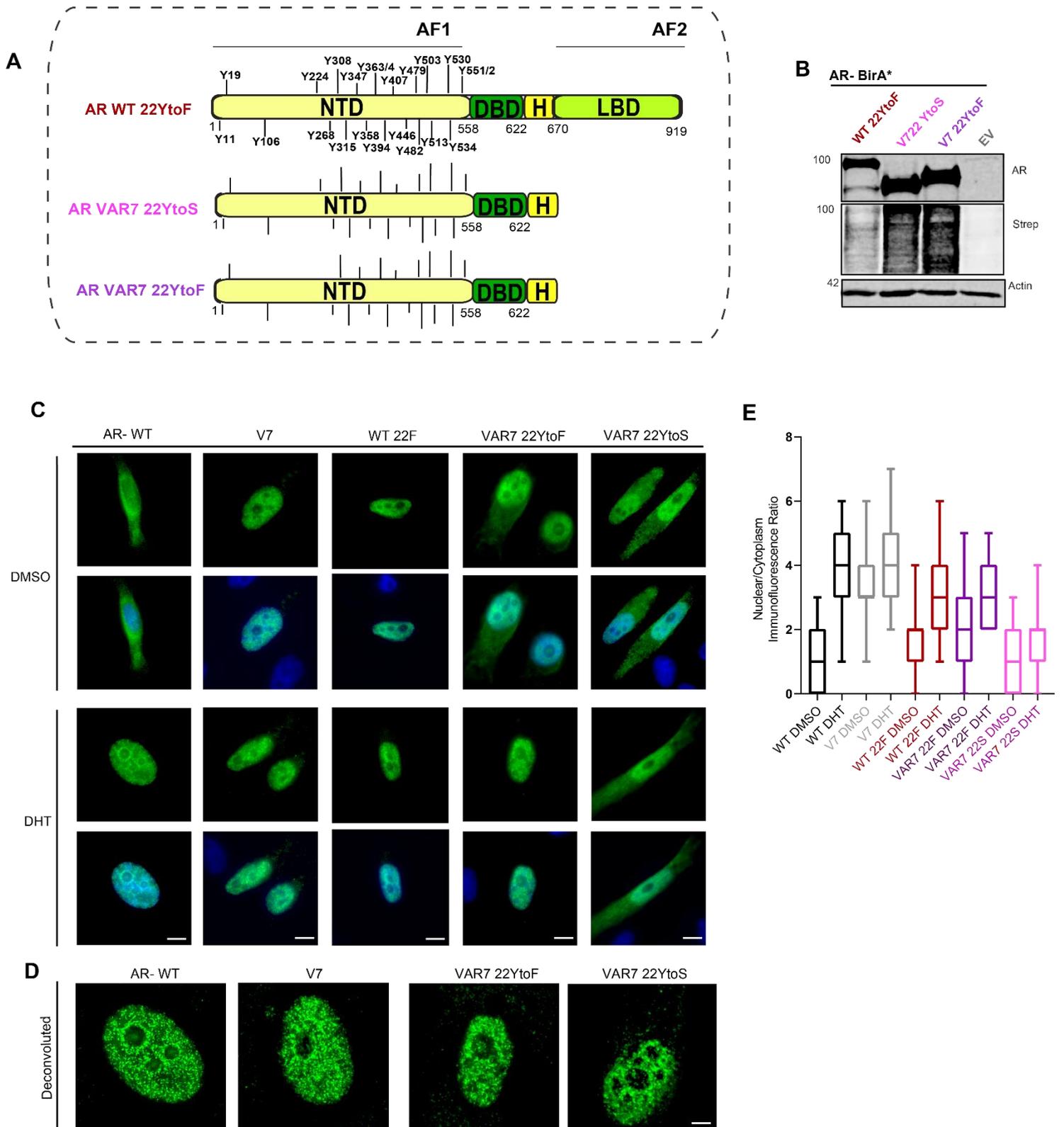


Figure 22: AR Variant 7 Tyrosine mutants rescue interaction. A. Schematic of Androgen Receptor (AR) domains and length (amino acid), AR-WT-22YtoF top and both Tyrosine AR-VAR7-22YtoS and AR-VAR7-22YtoF below, annotated with their respective tyrosine to serine (YtoS) or tyrosine to Phenylalanine (YtoF) substitutions along the protein. **B.** Western blot expression of the generated PC3 stable cell lines of

plasmids from A, stained with indicated antibodies. **C.** Representative immunofluorescence (IF) images of the PC3 stable cell lines from A+B with treatment of DHT (2 hrs) or without DMSO. Stained with FLAG. Scale bar= 10 μ m. **D.** Deconvoluted image of MTID-AR and mutants to show condensate formation. PC3 stable cell lines with DHT treatment. Stained with FLAG. Scale bar = 10 μ m **E.** Quantification from C of nucleic over cytoplasmic ratio of immunofluorescence intensity calculated in ImageJ.

Small molecule inhibitors deplete the AR proteomic network

As the NTD of AR is key to interactions with its co-activators, targeting it with small molecule inhibitors is a promising prospect with EPI-001 based compounds already in clinical trials (Sadar et al., 2020). Derivatives of EPI have been formulated, with potentially improved potency, including 1ae (Figure 23A, B) (Basu et al., 2022). Understanding the mechanistic effects of these small molecule inhibitors is therefore of clinical relevance. To determine if changes in the proximal interactome occur with the addition of these drugs, we performed BioID-MS with SAINTq analysis with a 1 hour pretreatment and then 2 hours with addition of DHT + Biotin in MTID-AR-WT expressing PC3 and LNCaP stable cell lines (Figure 23C).

1ae treatment caused a noticeable decrease in the intensity of interactions present in the PC3 cell line, whilst a small minority of proteins increased in the intensity of their interaction (Figure 23D). This effect was replicated in the PC3 cell lineage with EPI-001, with almost two-fold more proteins demonstrating reduced compared to an enhancement ($\text{Log}_2\text{FC} > 1.5$) (Figure 23E). The small molecule inhibitor treatment led to a more profound reduction in total AR interactors, including known transcriptionally related proteins (annotated) in the LNCaP cell line, with few interactors significantly increasing their intensity of interaction upon drug treatment (Figure 23F, G). GO searches showed the significantly decreased protein interactors are mainly categorised as transcriptional regulators, chromatin modifiers and specific androgen receptor related signaling partners (Figure 23H, I).

Figure 23

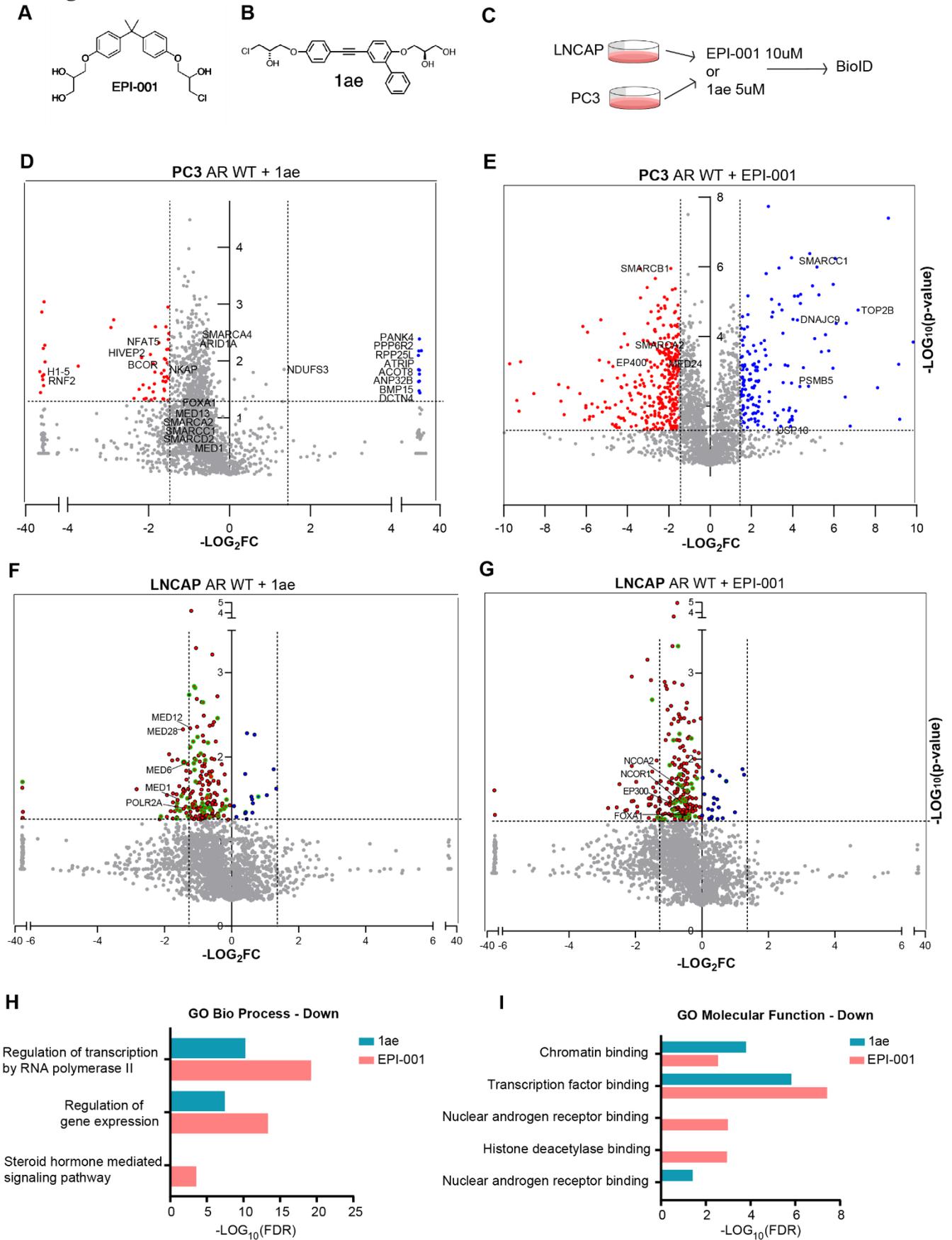


Figure 23: Small molecule inhibitors inhibit AR proteomic interactions. **A.** Chemical structure of compound EPI-001. **B.** Chemical structure of compound 1ae. **C.** Schematic of method for small molecule treatment of cells and BioID experiment. **D.** BioID-MS of PC3 MTID-AR-WT cells treated with 1 hour 1ae 5uM and then 2 hrs of DHT and Biotin. SAINTq intensity data of Log₂FC of intensity of interaction. Red highlights DOWN in reduced interaction pvalue < 0.05 and Log₂FC < -1.5 or UP in interaction. In blue pvalue < 0.05, Log₂FC > 1.5 or Log₂FC < -1.5. Annotated with proteins of interest. N=1. **E.** As in B BioID-MS of PC3 MTID-AR-WT cells but treated with 1 hour EPI-001 10uM and then 2 hrs of DHT and Biotin. LIMMA data of Log₂FC of intensity of interaction. Red highlights DOWN in reduced interaction pvalue < 0.05 and Log₂FC < -1.5 or UP in interaction in blue pvalue < 0.05 Log₂FC > 1.5 or Log₂FC < -1.5. Annotated with proteins of interest. N=1. **F.** as in D but utilising LNCaP MTID-AR-WT cells with SAINTq intensity data. Green circles represent proteins with BFDR < 0.02. N=3. **G.** As in E but utilising LNCaP MTID-AR-WT cells using SAINTq intensity data. Green circles represent proteins with BFDR < 0.02. N=3. **H,I.** Enrichment analysis of SAINTq intensity Data (BFDR < 0.02, DOWN=Log₂FC < -1.5) of the LNCaP MTID-AR-WT treated cells with small molecule inhibitors indicated above DOWN enriched protein interactions.

Small molecule inhibitors alter the homeostasis of specific transcriptional interactions

We further sought to determine the effects of small molecule inhibitors. Some of these affected key interactions, highlighted in Figure 23, were validated with PLA, using FLAG-AR + MED1 or ARID1A antibodies. Drug treatment led to a visible reduction in the total number of foci per cell, which was inversely correlated to increasing time of drug incubation in LNCaP cells (Figure 24A). This proved to be a significant reduction in the number of foci for MED1 with 1ae at all time points and ARID1A with both drugs at the longer time point of 8 hours (Figure 24B).

As the Mediator complex is a key interactor with AR, and notably important for phase separation, we highlighted the specific subdomain change in interaction we saw with these small molecule inhibitors in the BioID-MS of the LNCaP stable cell line. The MTID-AR-WT forms a strong interaction with the tail and middle parts of the Mediator complex (Figure 24C). EPI-001, and more strongly 1ae, caused the most significant loss of interaction across the Mediators middle and head subdomains, particularly MED1 and MED14 being the most significantly reduced (Figure 24C).

Lastly, using the BioID data of the LNCaP cell line with drug treatments 1ae and EPI-001, a network of the most enriched and significant interactors was formed (Figure 24D). These interactors were loosely characterised to functions related to transcriptional regulation, as well as other categories.

Interestingly both drugs led to enhanced AR interaction with E3 Ubiquitin protein ligase proteins, as well as the Polycomb group RING finger protein 1 (PCGF1). This suggests these small molecule inhibitors alter

the balance that exists between AR and its transcriptional co-activators and may lead to enhanced post-translational modification that could inhibit its function or target its degradation.

Figure 24

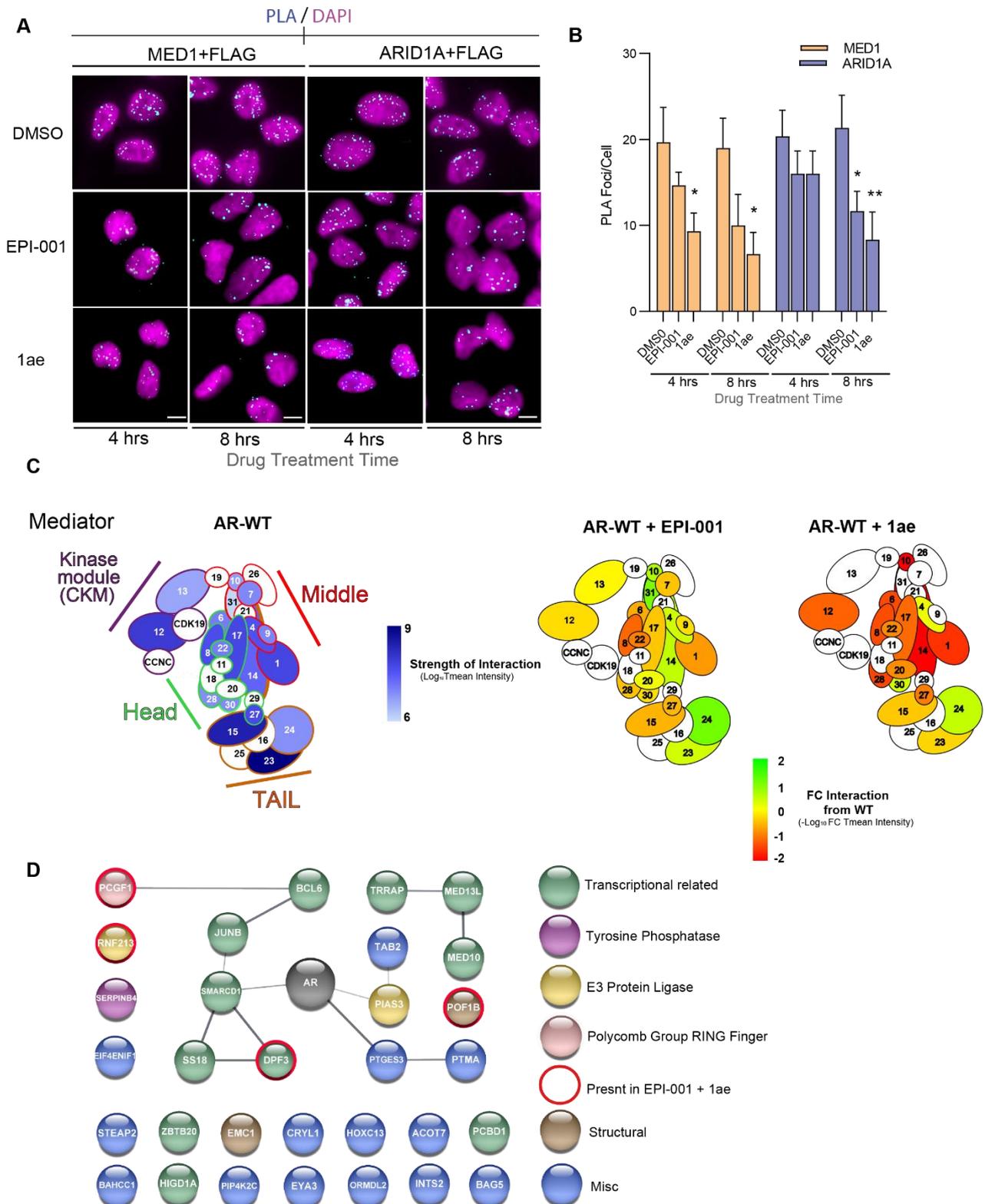


Figure 24: Small molecule inhibitors specific effects in AR's proteome. A Proximity ligation assay (PLA) using the indicated antibodies shown in cyan with DAPI staining in magenta in small molecule or

DMSO and DHT treated LNCaP MTID-AR-WT cells at times indicated. Scale bars=10 μ m. **B.** Quantitation of PLA foci shown in A per cell scored. **C.** BioID MS of MTID-AR-WT interaction with Mediator complex with and without small molecule inhibitors shown. SAINTq data, colour indicates strength of interaction from LogFC Tmean of intensity. **D.** Interactome of LNCaP MTID-AR-WT cell line BioID-MS treated with small molecule inhibitors. Most enriched and significant interactors FC > 1.5 compared to MTID-AR-WT and BFDR < 0.1. Generated from STRING. Made in Cytoscape.

Discussion

The TF AR and transcriptional regulators GEMC1 and MCIDAS are unified by their ability to drive specialised transcriptional programs to activate gene networks with important roles in cell fate decisions and oncology. In the first chapter, we exploited the proximity labelling technique of BioID to identify the complex and novel interaction networks that equip them to perform their specific transcriptional functions. This work is the first interactome study to compare the two atypical transcriptional activators GEMC1 and MCIDAS and implicate multiple SWI/SNF complexes in the transcriptional regulation of MCCs. Furthermore, we uncovered the molecular role of CCNO as an important regulator in Multiciliogenesis. In the second chapter we explored the role of phase separation on the heterogenic composition and organisation of the AR through multiple biochemical methods. We have uniquely demonstrated the importance of aromaticity in ARs translocation, interaction and function. Lastly, we attempted to alter the properties of these processes to gain a better understanding of how AR regulates transcription, which may help to generate beneficial therapeutic insights that will benefit PC treatment.

1. BioID proximity labelling identifies interaction networks in Multiciliogenesis

We and others have previously established that the two Geminin family members, GEMC1 and MCIDAS, are key factors that control the transcriptional regulation of MCC development (Terré et al. 2019; Lu et al., 2019). Yet, several lines of evidence point to distinct functions for GEMC1 and MCIDAS in MCC differentiation that is not adequately explained by our current understanding of their molecular interactions. In this work, we directly compared their transcriptional targets and proximal interactions and found that, while their targets and interactomes are highly similar, they also have distinct molecular interactions and transcriptional effects. Using BioID, we identified nearly every protein previously implicated in the transcriptional regulation of MCC differentiation to date, including E2F4/5-DP1, TRRAP, AHR, MYB, TP73, and CDK2, as well as many additional proteins implicated in their regulation, demonstrating the suitability of this strategy for interrogating transcriptional complexes (Lewis and Stracker, 2020). These results also indicated that the proposed role of GEMC1 and MCIDAS in DNA replication is likely limited to their interactions with Geminin/CDT1, as nearly all of the proteins identified were related to transcriptional regulation (Balestrini et al., 2010).

Previous work showed that the E2F4/E2F5-DP1 TFs were required for MCC differentiation in every tissue where MCCs are present (Chong et al., 2018b; Danielian et al., 2007, 2016; Ma et al.,

2014). Our data suggests that E2F3, as well as DP-2 (TFDP2), may also influence the transcriptional response, particularly with MCIDAS, that has higher affinity for E2F4 and E2F5 through its distinct C-terminal TIRT domain (Lu et al., 2019). Work by the Lygerou group demonstrated that GEMC1 was essential for the formation of MCCs in the mouse airway epithelium and responsible for the transcriptional activation of MCIDAS expression in radial glial cells (Arbi et al., 2016; Laloti et al., 2019). Research from the Roy group proposed a stepwise model of activation, as MCIDAS deficient mice developed morphologically recognisable MCCs that maintained expression of early TFs such as TP73 and FOXJ1, but failed to multiciliate, in contrast to GEMC1 KO mice that lack MCCs all together (Lu et al., 2019). These group's data and others indicate that GEMC1 activates MCIDAS and other key TFs to promote MCC specification and then MCIDAS consequently activates the expression of genes required for Multiciliogenesis. This is consistent with our array data that shows MCIDAS overexpression activating many more cilia related genes than GEMC1 (Figure 2).

How GEMC1 and MCIDAS achieve specificity remains unclear, but our work suggests that this may involve specific interactions with SWI/SNF complexes. We showed that GEMC1 almost exclusively interacts with BRG1 containing complexes, compared to MCIDAS that labeled both BRG1 and BRM (SMARCA2). MCIDAS showed almost exclusive binding to the BRD9 subunit of the ncBAF complex in comparison to GEMC1 in AD293s and our Inducible Multiciliogenesis model. However, moving the C-terminus of GEMC1 to MCIDAS (MC-C1) eliminated interactions with complex-specific subunits, which were then rescued by swapping the TIRT of MCIDAS back in the hybrid protein (MC-GEM-S) in AD293s (Figure 7). This indicates the TIRT domain is the key determinant of specificity that varies the interaction between the DNA-binding subdomains of the SWI/SNF. Observing the predicted Alphafold structures of the C-terminal domains of these proteins did not reveal any clear mechanistic insight, but did show that while they are somewhat similar, they are potentially varied enough to have distinct binding sites and affinities, consistent with biochemical data. Additionally, previous work showed the microRNA family Mir449a-c are key regulators of Multiciliogenesis and the initial stages of the transcriptional pathway (Wildung et al., 2019; Loukas et al., 2021). Understanding the complex relationship between these transcriptional regulators would benefit from a better understanding of their structures, further analysis of how the TIRT domains interact with TFs at the molecular level and a comprehensive view of their own transcriptional regulation by both TFs and miRNAs.

As both GEMC1 and MCIDAS identified multiple SWI/SNF complexes, one possibility is that they work together at some of the same genes through distinct enhancer/promoter interactions. It was shown that ARID1A containing BAF complexes bind more predominantly to H3K4

monomethylated enhancer regions, while ncBAF binds preferentially to H3K4-trimethylated promoter regions (Gatchalian et al., 2018; Michel et al., 2018). This theory is supported by the identification of common TFs that bind enhancer and promoter elements with both proteins in BioID experiments, as well as their similar labeling of the Mediator complex that bridges promoter and enhancer elements and recruits RNAPII (Jeronimo et al., 2016; Lee et al., 2011; Nemajero et al., 2016). Another possibility is that the weakened interactions of GEMC1 allow it to activate the MCC program in a manner that is reversible, if cell status is not appropriate. Work from the Kintner group proposed that the inhibition of MCIDAS by Geminin provided a mechanism for preventing centriole amplification in cycling cells (Ma et al., 2014). The relatively weak activation of these genes by GEMC1 compared to MCIDAS could make it more sensitive to Geminin inhibition, but further experiments are needed to test this idea.

It is known that MCC differentiation requires canonical cell cycle regulators, including CDK2, CDK1 and PLK1 to facilitate centriole amplification, although their precise roles remain largely unclear (Al Jord et al., 2017; Funk et al., 2015; Revinski et al., 2018; Wallmeier et al., 2014). GEMC1 was shown to be a target of CCNA2/CDK2 kinase activity (Balestrini et al., 2010) and previous work suggested that GEMC1 was negatively regulated by the poorly characterized Cyclin, CCNO (Cyclin O), potentially connecting CDK activity to its transcriptional functions (Funk et al., 2015). Consistent with this, we identified CCNA2 in our BioID experiments with GEMC1 and found that CCNO co-expression significantly reduced GEMC1's interactions. Notably, CCNO has little effect on the interactions of MCIDAS, indicating that its effects are specific for GEMC1. Our analysis of CCNO's interactome revealed a number of cell cycle components, many of which were also transcriptionally upregulated by CCNO.

We also noted that several G2/M specific factors, including PLK1 and CDC20 (a component of the APC/C), proximally associated with MCIDAS, but not GEMC1, and these associations were enhanced by CCNO co-expression. Co-expression of CCNO with MCIDAS led to the loss of interactions with TP73, SMARCD2 (a CDK target) and TEAD1 (a partner of YAP), indicating possible regulation of MCIDAS as well. YAP is a TF and component of the Hippo signaling pathway that responds to cell surface tension and mechanical forces (Lewis & Stracker, 2020). Several YAP interactors have been localised to BBs in mice, and in frogs, YAP nuclear translocation shows increased tissue stiffness in the multiciliated epithelium, suggesting YAP may influence MCC size or centriole numbers (Tu et al, 2018; Kim et al., 2020; Soldt et al., 2019). Thus, the emerging picture suggests that GEMC1 and MCIDAS are both likely targets of cell cycle-specific regulation, providing an elegant mechanism to facilitate their stepwise actions and fine tune transcriptional output with massive centriole amplification.

We validated the interaction between CCNO and CDC20B, which both have their highest expression levels during the earliest phases of MCC differentiation during the formation of deuterosomes. (Revinski et al., 2018; Funk et al., 2015; Núñez-Ollé et al., 2017) In CCNO deficient cells, deuterosomes are malformed and produce fewer BBs and defective MCCs (Funk et al., 2015; Núñez-Ollé et al., 2017). Transcriptional dysregulation is also apparent in Air liquid Interphase (ALI) cultures from CCNO KO mice indicating that CCNO is required for deuterosome stability and plays a role in the regulation of the Multiciliogenesis pathway (Funk et al., 2015). Recent studies showed that CDC20B localises to deuterostomes and regulates the release of centrioles during disengagement (Revinski 2018). Previous work showed that the phosphorylation of CDC20B by BUB1–PLK1 inhibits APC^{Cdc20} *in vitro* and is required for checkpoint signalling in human cells. (Luying Jia 2016). Therefore, it is likely that the components of the cell cycle machinery are being repurposed to regulate different stages of Multiciliogenesis, as proposed by Meunier and Spassky (Al Jord et al., 2017). Based on our work here (Figure 11 and 13), we speculate that CCNO negatively regulates GEMC1, presumably through CDK phosphorylation of GEMC1 itself, or potentially of its associated factors. CCNO may also modulate APC regulation by CDC20 phosphorylation through PLK1 and CDKs, keeping entry to mitosis inhibited and promoting the controlled formation and disengagement of centrioles.

Despite the limitation that our proteomic study was performed by overexpressing GEMC1 or MCIDAS in cycling AD293 cells, our approach is strongly validated by its ability to identify most of the factors known to be involved in MCC transcription. Due to technical challenges working with MCC models and the amount of material needed for BioID, performing such experiments in MCCs would be currently unfeasible. We and others have shown that the initial transcriptional functions of GEMC1 and MCIDAS are recapitulated in AD293 cells and the proteomic data picks up nearly all proteins previously implicated in MCC transcription. Notably absent from our results were the RFX2/3 transcription factors, that are expressed at low levels in AD293 (<http://www.proteinatlas.org>). This could be one reason why overexpression of GEMC1 and MCIDAS is incapable of driving efficient multiciliation in this cell type (Bisgrove et al., 2012; Chung et al., 2014; Didon et al., 2013). This notion is supported by the inducible Multiciliogenesis model created in this study using a Glioma cell line, DBTRG-05MG, which has the highest expression levels of RFX2/3 according to the TCGA dataset.

This model is similar to that reported by the Kintner group using mouse cells, but does not require the use of E2F4 overexpression (Ma et al., 2014). In our cell lines, GEMC1 or MCIDAS induction

resulted in markers of multiciliated cells being apparent from D3 (Figure 10). Results obtained from PLA and BioID-AP pull downs in this system solidified observations we made in AD293s, demonstrating functional significance of the SWI/SNF interactions in Multiciliogenesis.

2. BioID proximity labelling identifies interaction networks in PC

Due to difficulties in targeting AR therapeutically, widespread efforts are being made to understand the proteomic landscape and functional significance of its interaction partners. Some interactors, for example Cyclic adenosine monophosphate (cAMP) and p300, have already been linked with distinct sets of target genes (Ianculescu et al., 2012). These and other co-activators such as BRD4, represent potentially more druggable targets than AR that are being investigated as alternative ways to perturb AR gene activation and signalling, as discussed later in more detail (Asangani et al., 2014; Waddell et al., 2021).

Previous BioID studies have been performed on AR in AD293 cells and LAPC4 cells (Lempiäinen et al., 2017; Vélot et al., 2021.). The AR interaction networks generated from our PC stable cell lines have been partially validated from these previous studies, as we identified many members of the same transcriptional machinery and a large majority of AR interactors reported in BioGrid. The only caveat from previous studies is the reliance on transient transfection, whereas our stable cell line system is more manipulatable, maintains expression to near physiological levels, and is less variable. The results presented here push the field further, providing insights into the interaction changes during the transition of AR following ligand binding in both hormone dependent and independent PC cell lines. Generally, AR is bound HSP in the cytosol prior to ligand activation. This family of proteins often appear in FLAG-BirA* controls and are often not returned as significant hits, something noted in Contaminant Repository for Affinity Purification (the CRAPome), which is carefully considered when analysing all significant interactions (Mellacheruvu et al., 2013).

Utilisation of the Mini Turbo ID (MTID) with AR, enabled more flexible and rapid labelling times, allowing us to look temporally at interactions after hormone treatment. Looking at the condition of 0 min DHT in our time course (Figure 20), AR maintains cytoplasmic localisation and its interaction partners are associated to biological process relating to regulation of cellular component organisation, actin or vesicle mediated transport and K63-linked deubiquitination, which has been linked with AR stability (Lim et al., 2017). We identified novel interactions with

several proteins at this time point. For instance, we noted several interactions with breast cancer anti-estrogen resistance (BCAR) proteins, like BCAR1 which serves as a structural hub in cellular signaling and may correlate to PC progression (Heumann et al., 2018).

We saw a rapid change in the proteome of AR after only a 15 min interval of DHT + biotin, with the AR becoming almost entirely nuclear and hits being largely transcriptionally related. We see therefore, a large amount of the transcriptional machinery already in proximity to the AR within 15 minutes, with a stronger enrichment of chromatin remodellers than Mediator, possibly reflecting limited active transcription at this point. Typically identified interactors of AR at the 15 minute time point included translocation proteins that can shuffle between the cytoplasm and the nucleus, as well as structural proteins that serve as an interface between TFs. Such proteins may offer potentially novel targets to disrupt pathogenic AR translocation or at least reduce its transcriptional output.

After 60 mins of ligand treatment, fully translocated AR consistently identified members of large chromatin modifiers, TFs and Mediator subunits, indicative of proteins found in active, phase separated condensates. Unexpectedly, we discovered a number of mitochondrial proteins as significant hits including Mitochondrial Ribosomal Proteins (MRPS), Isoleucyl-TRNA Synthetase 2, Mitochondrial (IARS2) and Pentatricopeptide Repeat Domain 3 (PTCD3). Mitochondrial type proteins were found across all cell lines tested with the highest number occurring in AD293 and PC3 cells. AR was recently shown to also localise to mitochondria, potentially regulating mitochondrial function and retrograde signaling, although this relationship warrants further investigation with respect to PC (Bajpai et al., 2019).

Using the same system, we examined the proximal interactome of the AR-VAR7, a ligand independent variant that has been proposed to be clinically relevant. To our knowledge this is the first report to examine the proximal interactome of AR-VAR7 in any cell line. Although, generated from AD293s, clear differences in comparison to the AR-WT can be seen, including splicesomal and RNA processing type interactors that are unique for AR-VAR7. The Ring finger protein 113A (RNF113A) was identified, which is an E3 ubiquitin-protein ligase that functions in pre-mRNA splicing and promotes the proliferation, migration and invasion of cancer, its expression often correlating with poor prognosis in certain cancer types (Wang et al., 2018; Lear et al., 2017). Further work is being carried out to determine the AR-VAR7 interactome in PC lines to understand the functionality of the variant which is found more prominently in CRPC.

There are a number of possible avenues to pursue in relation to determining functional relevancy for any number of AR co-activators identified in this study. For example, a similar BioID study

in PC identified the Kruppel-like Factor 4 in the cell line LAPC4, and through biochemical methods they determined its significance to AR function (Vélot et al., 2021). Due to time constraints and other particular aims of the AR project, these leads were not fully explored. However, it would be important to show dependency of AR function on a number of the interactions in relevant PC models in future work, as it was seen that a large number of proteins are expressed at diverse levels throughout the progression of PC, from benign prostate cells to CRPC (Kumari et al., 2017; Chmelar et al., 2006). Importantly, the specificity of interaction between AR and its potential pivotal interactor should be assessed to establish any potential off target toxicity that may come as a result of therapeutic intervention (Dahiya and Heemers 2022). Potentially, it would be interesting to apply the BioID technique to xenograft tumours, as was performed for the TF c-MYC (Dingar et al., 2015). This could be done specifically, in CRPC in conjunction with different drug treatments or conditions, like enzalutamide or hypoxia, in order to gain more clinically relevant insights to AR.

Additional new technologies based on BioID have been developed, such as Split-TurboID, which consists of two inactive fragments of TurboID that biotinylate proximal proteins when the two constitutive protein-protein interaction occurs. This allows greater targeting specificity, which could be used with AR to decipher interactomes at transcriptionally active regions (Cho et al., 2020). Crosslinking (XL)-MS based proteomic studies have been used with AR at chromatin regions, which confirmed many known AR coregulators reported here such as FOXA1, P300 and the newly identified interactor GRHL2 (Paltoglou et al., 2017). Interesting recent work utilised a chromatin immunoprecipitation coupled with selective isolation of chromatin-associated protein method (ChIP-SICAP) and captured the chromatin protein network of AR in VCaP cells that represent CRPC, to highlight a number of clinically relevant candidates (Launonen et al., 2021). Moreover, a vast number of quantitative proteomics studies have been performed over the years ranging from PC cell lines to patient samples (reviewed Sadeesh et al., 2021). The depth of these studies is beyond the scope of this thesis, but as protein expression profiles can change from a non-malignant sample to metastatic tissue in the same patient, further work should adopt a more integrative approach. These could combine genomic, transcriptomic, and proteomic alterations to tackle the ever-changing landscape of this disease (Kwon et al., 2020; Sadeesh et al., 2021).

3. Phase separation and Tyrosine Mutants

The weight of evidence behind the process of phase separation orchestrating transcriptional machinery to coherently form at SEs and mediate RNAP transcription is growing. As GEMC1

and MCIDAS do not have identifiable DNA binding domains, their target gene specificity is likely dictated by multivalent protein interactions rather than specific contact with DNA elements, a property indicative of phase separation. GEMC1 and MCIDAS form clear nuclear condensates in multiple cell lines tested in this thesis, which did not correlate to PML type foci in the nucleus. Further experiments such as Fluorescence recovery after photobleaching (FRAP), super-resolution Stimulated emission depletion (STED) microscopy, or *in vitro* based droplet assays, are needed to clarify its propensity to phase separate. However, foci type condensates can be seen in immunofluorescence based images in other overexpression studies or in MCC-bearing tissues tissue (Kyrousi et al., 2015; Lu et al., 2019). Additionally, as our results show a vast number of transcriptional interactors, many of which are already implicated in phase separated condensates, like SWI/SNF components and Mediator, it is easy to imagine the relatively small proteins GEMC1 and MCIDAS are integrated in this process to recruit the large transcriptional complexes needed to active their gene program. For this reason, we speculate that our extensive attempts to ChIP either GEMC1 or MCIDAS to identify direct gene targets were unsuccessful because GEMC1 and MCIDAS are not located in close enough proximity to DNA.

However, our data provides a number of leads to pursue this possibility in future work, in order to better understand the mechanism that confer transcriptional specificity. In addition, approaches such as ATAC-seq, that measures changes in chromatin density, could be applied to infer target genes of GEMC1 and MCIDAS in more detail.

AR on the other hand, is a DNA binding-domain containing TF that has been previously shown to form condensates *in vitro* and *in vivo* (Black et al., 2004, 2004b; Zhang et al., 2021; Basu et al., 2022). The data that was generated from the AR project illustrated how the phase separation capacity of AR can be altered by changing the aromaticity of the receptor's NTD. As the 8YtoS mutant showed decreased propensity to form condensates *in vitro* (Basu et al., 2022). It is becoming evident that the molecular grammar of phase separation is stipulated by pi-pi interactions between two aromatic residues, or the pi-cation interactions of aromatic residues and positively charged amino acids (Wang et al., 2018). Previous studies substituting tyrosine residues with serine on prion like domains changed the propensity of phase transition and inhibited their capacity to form hydrogels (Kato et al., 2012; Kwon et al., 2013). With the 8YtoS substitutions in AR's NTD we saw a decreased interaction of its known co-activators and decreased transcriptional activity, consistent with previous mutagenesis based studies on an array of different TFs (Basu et al., 2020; Boija et al., 2018; Wang et al., 2021; Zhang et al., 2022). Whilst positively enriched protein interactions indicated perturbed AR signaling and cell growth with the 8YtoS mutant. This data indicates that the structural properties of AR's NTD domain facilitate the phase separation

process. This is likely indicative for other nuclear hormone receptors and transcriptional regulators, similar to condensates seen formed by various prion-like proteins (Li et al., 2018; Martin et al., 2020; Vernon et al., 2018).

We found that replacing all 22YtoS significantly inhibited transcription based protein interactions and gene activation typical of AR. Unexpectedly, this Tyrosine mutant which still had an intact NLS, showed impaired nuclear translocation. 22YtoS preferentially interacted with nucleoporins on the nuclear membrane and PLA revealed a perinuclear accumulation compared to the AR-WT (Figure 20). Nucleoporins contain FG repeats that we speculate directly interact with the aromatic residues in the AR NTD and hinder the membrane crossing mediated by nuclear import receptors and adaptor proteins (Milles et al., 2015). Nuclear pore complexes undergo phase separation themselves, highlighting a complex relationship between permeabilisation and condensate formation (Nag et al., 2022). Studies in large globular proteins, which substituted surface residues with aromatic ones, resulted in an increase in the rate of nuclear translocation, further supporting our results (Frey et al., 2018). This suggests a potentially universal nuclear import mechanism different from the linear sequence information encoded in nuclear localisation signals (Basu et al., 2022).

In light of these findings, AR-VAR7 was used in conjunction with the 22YtoS mutations. This was found to be more localised to the nucleus with ligand stimulation than the same mutant in the AR-WT background, but visually it did not form typical transcriptional condensates. A 22YtoF plasmid was created to act as a rescue experiment, as phenylalanine is an aromatic amino acid with a lesser propensity to phase separate. This plasmid showed more residual nuclear staining and better response to ligand stimulation indicating less perceived translocation interference. The analysis of their respective proximal interactions was performed but not included in this thesis due to time restraints.

One alternative explanation regarding the effects of the tyrosine mutants seen in this study that cannot be fully ruled out is associated with post-translational modifications (Hofweber & Dormann 2019). Substituting tyrosine residues to serine inherently alters some known kinase targets that may alter the function of AR (Liu et al., 2010; Mahajan et al., 2007). Although, in house phosphoproteomic experiments only revealed one phosphorylation site in AR-Y224 in our system, to fully exclude this uncertainty one could add inhibitors to the known tyrosine kinases or perform parallel experiments with phosphomimic mutants. Additionally, potential interference with AR's capacity to dimerise should also be assessed, as both inter- and intramolecular NTD-LBD

interactions facilitate ARs dimerisation which may be affected in these mutants and could contribute to decreased affinity with its interaction partners (Yu et al., 2020).

Equally, it should not be ignored the alternate mechanisms that exist to describe the biogenesis, maintenance, and disassembly of membraneless compartments of macromolecules. Musacchio argues that many properties held by the PS viewpoint can rather be explained by Site Specific Interactions (SSIs), whereby highly specific interactions, rather than unspecific ones, appear to be the main driver of biogenesis of subcellular compartments (Musacchio, 2022). It is true that to define PS is an arduous task, as there are countless ways to modulate the solubility of a protein to form droplets *in vitro* and it is difficult to correlate with overexpression experiments in cells that do not necessarily represent the physiological environment (Musacchio, 2022). Definition experiments come under their own scrutiny like FRAP that different studies have kinetics of recovery ranging between sub-seconds to hours (McSwiggen et al, 2019). Thus, this study like many others in the field falls short of excluding other potential drivers of compartment biogenesis under limited resources and time. However, PS is still an emerging field with many unknowns and the tools needed to improve its defining parameters and translation into the *in vivo* context are still developing. Only with advancements in super-resolution fluorescence methods and effective interdisciplinary minded approaches can these shortcomings be fulfilled.

4. Implications for disease

We identified distinct SWI/SNF subcomplexes used by GEMC1 and MCIDAS that influence their transcriptional outputs and inhibition of BRD9 impaired Multiciliogenesis in mTECs and our Induction Multiciliogenesis model. These observations may have significance to ependymal MCC differentiation in mice or human patients with SWI/SNF mutations and hydrocephaly (Diets et al. 2019; Cao and Wu 2015).

The discovery of the transcriptional specificity being reliant on the TIRT domain may also be of clinical significance. Mutations in the TIRT domain of MCIDAS have been implicated in RGMC, a condition in which patients have malformed MCCs, and may be linked to the pathogenesis of other ciliopathies (Boon et al., 2014; Wallmeier et al., 2014 N nez-Oll  et al., 2018; Zhang et al., 2018). Consequently, restoring the aberrant binding of the TIRT domain with specific SWI/SNF components or E2F4/5 TFs, could potentially re-establish defective Multiciliogenesis. Approaches

such as CRISPR/Cas9 gene editing could be used in future work to correct point mutations in RGMC patients as restore MCC function.

We measured the expression levels of key MCC specific genes in several cancer cell backgrounds and found that they showed significant upregulation. GEMC1 was overexpressed in certain cell lines consistent with previous reports which suggests it may play functional roles in DNA replication (Li et al., 2018; Cerami et al., 2012; Gao et al., 2013). This warrants additional studies of the transcriptional roles of GEMC1 and MCIDAS expression in cancer. Whether or not the expression of these genes are pivotal driving forces in cancer progression or by-products of pathway dysregulation in these backgrounds should be determined. It is not hard to envisage these proteins contributing to the pathology through the over activation of TFs such as E2F4, TP73, FOXJ1 and SWI/SNF components, that have already been implicated in cancer development (Zheng et al., 2021; Bisso et al., 2011; Liu et al., 2019). Contrastingly, it was proposed that re-activation of the GMNC-MCIDAS Multiciliogenesis pathway can be tumour suppressive. Through Notch interference one study was able to generate tumours that are similar to Choroid plexus papilloma in mice. Disruption of Notch restored multiciliation and decreased tumour growth, while GEMC1-MCIDAS overexpression rescued multiciliation defects and suppresses tumour cell proliferation (Li et al., 2022). These studies indicate an intricate and perhaps evolving relationship between cancer causation, progression and the Multiciliogenesis pathway.

Likewise, the identified potential regulators of MCCs we highlighted, CCNO and CDC20, have a multifaceted association with oncology. There is a lot of interest of CDC20 as an oncogene and its pathogenesis to cancer. Overexpression of CDC20 was observed in a variety of human tumours (Cheng et al., 2019; Wang et al., 2015). Moreover, higher expression of CDC20 is associated with clinicopathological parameters in various types of human cancers, such as pancreatic or breast cancer (Taniguchi et al., 2008; Karra et al., 2014). CCNO is relatively understudied in oncology but its depletion was seen to significantly induce cancer cell apoptosis, both *in vitro* and *in vivo* (Li et al., 2018). Moreover, PLK1 is well known to be highly expressed in many cancers, where its expression is associated with poor prognosis (Ramani et al., 2015; Tut et al., 2015). Thus, one hypothesis arising from our work is that CCNO overexpression in cancer could be activating PLK1 and CDC20, thus having oncogenic effects. Analysis of CCNO overexpression showed correlation with a subset of cilia genes in some cancers, but it has yet to be determined the significance of this. As the G2M programme is clearly associated with oncogenesis, this may have some clinical importance or prognostic value and warrants further investigation in the future.

The significance of AR in prostate cancer is undisputed. However, current treatments for CRPC are limited and generally ineffective as long term options. There are many emerging studies relating phase separation to cancer development (Wang et al., 2020; Spann et al., 2019). The Salvatella group and collaborators designed next generation compounds based on knowledge of the properties of phase separation and the pharmacokinetics of small molecule inhibitors that function by covalently attaching to multiple aromatic groups to the AR-NTD. In this study we utilised the small molecule EPI-001, which is known to inhibit AR-NTD, derivatives of which are currently under clinical investigation (Le Moigne et al., 2020). These reduce AR's propensity to phase separate and we showed significant and specific decreases in interactions with AR's transcriptional machinery, many of which are known to be highly expressed at SEs, following EPI-001 treatment (Hnisz et al., 2013). Work with our collaborators using a newly generated compound, 1ae, demonstrated enhanced potency *in vitro* and inhibited tumourigenesis in a murine CRPC model resistant to enzalutamide (Basu et al., 2022). Understanding why enhanced AR phase separation from addition of small molecule inhibitors perturbs AR's interactions and facilitates AR dysregulation remains elusive. It may be that exogenous enhancement of phase separation can alter the normal physiological balance that exists between AR and its interactors with respect to heterogeneity and assembly kinetics. This could lead to enhanced condensate formation, but condensates are now non-functional or even detrimental to the cell. We noted ubiquitin protein ligase proteins and PCGF proteins as interactors in the small molecule treated samples that may alter enhanced post-translational modifications and target ARs degradation. This is supported by previous studies showing AR partition into condensates with SPOP, an adaptor of the cullin3-RING ubiquitin ligase (Bouchard et al., 2018). Whilst most recently, a dose-dependent decrease in the level of soluble AR was found with 1ae-treated LNCaP cells (Basu et al., 2022).

Interestingly, another recent study showed that increasing phase separation does not necessarily correlate with TF transcriptional activation (Trojanowski et al., 2022). Although this was performed with many artificial components, it highlights the complex nature of this process.

One group of proteins enriched from BioID experiments with small molecule treated samples of AR were RNA type binding, splicing and general processing proteins. Enhancers produce short and long RNAs radiating from the center that vary in stability, but these RNAs are known to play key structural roles in transcriptionally active condensates (Suzuki et al., 2017; Sharp et al., 2022). Altering the aromaticity of AR may lead to increased attraction of RNAs and their binding proteins, which results in an excess of negatively charged RNA molecules causing repulsion, which has been shown to cause condensate dissolution (Banerjee et al. 2017; Milin and Deniz 2018). Alternatively, the enriched RNA molecules found with small molecule treatment may be a direct

consequence of malfunctioning transcription increasing their local concentration around AR target sites. It would therefore be interesting to see Hi-C coupled with ChIP-seq (HiChIP-seq) experiments performed in this context to analyse the three dimensional changes in enhancer based genomic localisations of AR with these drugs.

Treating SE as potential drug targets is an exciting future prospect in drug development. Our working models both implicate GEMC1, MCIDAS and AR in their respective systems, at the heart of SEs, that may form through phase separation. SE defects have been linked to multiple genetic diseases, including cancer (Thandapani 2019; Sabari et al., 2019; Loven et al., 2015) metabolic (Ounzain et al., 2015) and immune diseases (Herranz et al., 2014; Wang et al., 2019). Equally, SEs are responsible for activating cell identity genes. Consequently, they hold therapeutic potential, as SEs are more prone to external stimuli than any other genomic regions which makes them good targets, as inhibitory responses should be obtained with low drug concentrations (Bruter et al., 2021). Therefore, they have the potential to enhance the potency and specificity of drugs, which persists as a major therapeutic challenge in disease. Inhibitors against the bromodomain protein BRD4, which is an important marker in SE activity, has shown reduced tumour growth and suppression of oncogenic Myc, and its downstream transcriptional program in preclinical models of acute myeloid leukemia and multiple myeloma (Delmore et al., 2011; Zuber et al., 2011). Consequently, there are number of clinical trials targeting other bromodomains BRD2/3/4/T (Alqahtani et al., 2019). The knockdown of the BRD4, and two Mediator subunits reduced enhancer dependent gene activation and repressed proliferation of colorectal cancer cells (Kuuluvainen et al., 2018). More pertinently, elimination of the SMARCA2 and SMARCA4 SWI/SNF's ATPases with a proteolysis-targeting chimera (PROTAC) inhibited tumour growth in xenograft and drug resistant models of 'enhancer-addicted' PC (Xiao et al., 2022).

Going forward, it will be important to fully define the pathological SEs seen in cancers from the physiological in the transcriptome (Shin et al., 2018). Armed with this better mechanistic knowledge, disentangling their formation, or rather aggregation, will be site specific and detrimental to their transcriptional output. Although, more *in vivo* evidence is still required to further support SEs as pathological intervention targets, further characterisation of their makeup, the co-activators, the assembly and disassembly in physiology and disease could lead to widespread and potent therapeutic strategies.

Conclusions

- Proximity labelling with BioID identified the proximal protein-protein interaction network of GEMC1 and MCIDAS that are key activators of the Multiciliogenesis transcriptional program.
 - GEMC1 and MCIDAS share a similar proximal interaction network but activated a distinct set of genes.
 - MCIDAS has preferential binding specificity to the BRD9 subunit of the ncBAF SWI/SNF complex that is dependent on its unique TIRT domain.
 - Inhibition of BRD9 prevented MCCs formation in a Multiciliogenesis mouse model.
 - Depletion of BRD9 using PROTAC inhibited Multiciliogenesis in an inducible *in vitro* model.
 - CCNO regulates the interaction of GEMC1 to transcriptional regulators.
 - CCNO activates the G2/M program and interacts with a number of G2/M specific proteins that have been implicated in multiciliogenesis.
-
- The AR proximity interaction network was mapped across several PC lines and in a hormone dependent manner.
 - Tyrosine residues are important for efficient translocation of the AR across the nuclear membrane.
 - Reducing phase separation of the AR inhibited interactions with the transcriptional machinery and gene activation of PC specific genes.
 - Small molecule inhibitors decreased protein-protein interactions and may be therapeutically useful for targeting the IDRs of oncogenic transcription factors and other disease-associated proteins.

Disclosure

To abide by any copyright infringements, it is declared that parts of the introductory text and figures are taken from the authors review in *Seminars in Cell & Developmental Biology*.

Full Reference;

Lewis, M., & Stracker, T. H. (2021). Transcriptional regulation of multiciliated cell differentiation. *Seminars in Cell & Developmental Biology*, 110, 51–60.
<https://doi.org/10.1016/j.semcdb.2020.04.007>.

References

- Abdi, K., Lai, C.-H., Paez-Gonzalez, P., Lay, M., Pyun, J., & Kuo, C. T. (2018). Uncovering inherent cellular plasticity of multiciliated ependyma leading to ventricular wall transformation and hydrocephalus. *Nature Communications*, *9*(1), 1655. <https://doi.org/10.1038/s41467-018-03812-w>
- Al Jord, A., Lemaître, A.-I., Delgehyr, N., Faucourt, M., Spassky, N., & Meunier, A. (2014). Centriole amplification by mother and daughter centrioles differs in multiciliated cells. *Nature*, *516*(7529), 104–107. <https://doi.org/10.1038/nature13770>
- Al Jord, A., Shihavuddin, A., Servignat d'Aout, R., Faucourt, M., Genovesio, A., Karaiskou, A., ... Meunier, A. (2017). Calibrated mitotic oscillator drives motile ciliogenesis. *Science (New York, N.Y.)*, *358*(6364), 803–806. <https://doi.org/10.1126/science.aan8311>
- Alberti, S. (2017). Phase separation in biology. *Current Biology : CB*, *27*(20), R1097–R1102. <https://doi.org/10.1016/j.cub.2017.08.069>
- Allen, B. L., & Taatjes, D. J. (2015). The Mediator complex: a central integrator of transcription. *Nature Reviews. Molecular Cell Biology*, *16*(3), 155–166. <https://doi.org/10.1038/nrm3951>
- Alqahtani, A., Choucair, K., Ashraf, M., Hammouda, D. M., Alloghbi, A., Khan, T., ... Nemunaitis, J. (2019). Bromodomain and extra-terminal motif inhibitors: a review of preclinical and clinical advances in cancer therapy. *Future Science OA*, *5*(3), FSO372. <https://doi.org/10.4155/fsoa-2018-0115>
- Amirav, I., Wallmeier, J., Loges, N. T., Menchen, T., Pennekamp, P., Mussaffi, H., ... Omran, H. (2016). Systematic Analysis of CCNO Variants in a Defined Population: Implications for Clinical Phenotype and Differential Diagnosis. *Human Mutation*, *37*(4), 396–405. <https://doi.org/10.1002/humu.22957>
- An, S., & Fu, L. (2018). Small-molecule PROTACs: An emerging and promising approach for the development of targeted therapy drugs. *EBioMedicine*, *36*, 553–562. <https://doi.org/10.1016/j.ebiom.2018.09.005>
- Andersson, R., & Sandelin, A. (2020). Determinants of enhancer and promoter activities of regulatory elements. *Nature Reviews Genetics*, *21*(2), 71–87. <https://doi.org/10.1038/s41576-019-0173-8>
- Arbesú, M., Iruela, G., Fuentes, H., Teixeira, J. M. C., & Pons, M. (2018). Intramolecular Fuzzy Interactions Involving Intrinsically Disordered Domains. *Frontiers in Molecular Biosciences*, *5*, 39. <https://doi.org/10.3389/fmolb.2018.00039>
- Asangani, I. A., Dommeti, V. L., Wang, X., Malik, R., Cieslik, M., Yang, R., ... Chinnaiyan, A. M. (2014). Therapeutic targeting of BET bromodomain proteins in castration-resistant prostate cancer. *Nature*, *510*(7504), 278–282. <https://doi.org/10.1038/nature13229>
- Aurilio, G., Cimadamore, A., Mazzucchelli, R., Lopez-Beltran, A., Verri, E., Scarpelli, M., ... Montironi, R. (2020). Androgen Receptor Signaling Pathway in Prostate Cancer: From Genetics to Clinical Applications. *Cells*, *9*(12). <https://doi.org/10.3390/cells9122653>
- Azimzadeh, J., & Bornens, M. (2007). Structure and duplication of the centrosome. *Journal of Cell Science*, *120*(Pt 13), 2139–2142. <https://doi.org/10.1242/jcs.005231>

- Azimzadeh, J., Wong, M. L., Downhour, D. M., Sánchez Alvarado, A., & Marshall, W. F. (2012). Centrosome loss in the evolution of planarians. *Science (New York, N.Y.)*, *335*(6067), 461–463. <https://doi.org/10.1126/science.1214457>
- Bajpai, P., Koc, E., Sonpavde, G., Singh, R., & Singh, K. K. (2019). Mitochondrial localization, import, and mitochondrial function of the androgen receptor. *The Journal of Biological Chemistry*, *294*(16), 6621–6634. <https://doi.org/10.1074/jbc.RA118.006727>
- Balestrini, A., Cosentino, C., Errico, A., Garner, E., & Costanzo, V. (2010). GEMC1 is a TopBP1-interacting protein required for chromosomal DNA replication. *Nature Cell Biology*, *12*(5), 484–491. <https://doi.org/10.1038/ncb2050>
- Banani, S. F., Lee, H. O., Hyman, A. A., & Rosen, M. K. (2017). Biomolecular condensates: organizers of cellular biochemistry. *Nature Reviews. Molecular Cell Biology*, *18*(5), 285–298. <https://doi.org/10.1038/nrm.2017.7>
- Banani, S. F., Rice, A. M., Peeples, W. B., Lin, Y., Jain, S., Parker, R., & Rosen, M. K. (2016). Compositional Control of Phase-Separated Cellular Bodies. *Cell*, *166*(3), 651–663. <https://doi.org/10.1016/j.cell.2016.06.010>
- Banerjee, P. R., Milin, A. N., Moosa, M. M., Onuchic, P. L., & Deniz, A. A. (2017). Reentrant Phase Transition Drives Dynamic Substructure Formation in Ribonucleoprotein Droplets. *Angewandte Chemie (International Ed. in English)*, *56*(38), 11354–11359. <https://doi.org/10.1002/anie.201703191>
- Bantscheff, M., Hopf, C., Savitski, M. M., Dittmann, A., Grandi, P., Michon, A.-M., ... Drewes, G. (2011). Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. *Nature Biotechnology*, *29*(3), 255–265. <https://doi.org/10.1038/nbt.1759>
- Bardin, C. W., & Catterall, J. F. (1981). Testosterone: a major determinant of extragenital sexual dimorphism. *Science (New York, N.Y.)*, *211*(4488), 1285–1294. <https://doi.org/10.1126/science.7010603>
- Basu, S., Martínez-Cristóbal, P., Pesarrodona, M., Frigolé-Vivas, M., Szulc, E., Lewis, M., ... Salvatella, X. (2022). Androgen receptor condensates as drug targets. *BioRxiv*, 2022.08.18.504385. <https://doi.org/10.1101/2022.08.18.504385>
- Beilin, J., Ball, E. M., Favaloro, J. M., & Zajac, J. D. (2000). Effect of the androgen receptor CAG repeat polymorphism on transcriptional activity: specificity in prostate and non-prostate cell lines. *Journal of Molecular Endocrinology*, *25*(1), 85–96. <https://doi.org/10.1677/jme.0.0250085>
- Belakavadi, M., Pandey, P. K., Vijayvargia, R., & Fondell, J. D. (2008). MED1 phosphorylation promotes its association with mediator: implications for nuclear receptor signaling. *Molecular and Cellular Biology*, *28*(12), 3932–3942. <https://doi.org/10.1128/MCB.02191-07>
- Berberi, N. F., O'Connor, A. K., Haycraft, C. J., & Yoder, B. K. (2009). The primary cilium as a complex signaling center. *Current Biology : CB*, *19*(13), R526-35. <https://doi.org/10.1016/j.cub.2009.05.025>
- Bergeron-Sandoval, L.-P., Safaee, N., & Michnick, S. W. (2016). Mechanisms and Consequences of Macromolecular Phase Separation. *Cell*, *165*(5), 1067–1079. <https://doi.org/https://doi.org/10.1016/j.cell.2016.05.026>

- Bisso, A., Collavin, L., & Del Sal, G. (2011). p73 as a pharmaceutical target for cancer therapy. *Current Pharmaceutical Design*, *17*(6), 578–590. <https://doi.org/10.2174/138161211795222667>
- Blatt, E. N., Yan, X. H., Wuerffel, M. K., Hamilos, D. L., & Brody, S. L. (1999). Forkhead transcription factor HFH-4 expression is temporally related to ciliogenesis. *American Journal of Respiratory Cell and Molecular Biology*, *21*(2), 168–176. <https://doi.org/10.1165/ajrcmb.21.2.3691>
- Bledsoe, R. K., Montana, V. G., Stanley, T. B., Delves, C. J., Apolito, C. J., McKee, D. D., ... Xu, H. E. (2002). Crystal structure of the glucocorticoid receptor ligand binding domain reveals a novel mode of receptor dimerization and coactivator recognition. *Cell*, *110*(1), 93–105. [https://doi.org/10.1016/s0092-8674\(02\)00817-6](https://doi.org/10.1016/s0092-8674(02)00817-6)
- Boehning, M., Dugast-Darzacq, C., Rankovic, M., Hansen, A. S., Yu, T., Marie-Nelly, H., ... Zweckstetter, M. (2018). RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nature Structural & Molecular Biology*, *25*(9), 833–840. <https://doi.org/10.1038/s41594-018-0112-y>
- Boija, A., Klein, I. A., Sabari, B. R., Dall'Agnesse, A., Coffey, E. L., Zamudio, A. V., ... Young, R. A. (2018). Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell*, *175*(7), 1842–1855.e16. <https://doi.org/10.1016/j.cell.2018.10.042>
- Boon, M., Wallmeier, J., Ma, L., Loges, N. T., Jaspers, M., Olbrich, H., ... Omran, H. (2014). MCIDAS mutations result in a mucociliary clearance disorder with reduced generation of multiple motile cilia. *Nature Communications*, *5*, 4418. <https://doi.org/10.1038/ncomms5418>
- Borgia, A., Borgia, M. B., Bugge, K., Kissling, V. M., Heidarsson, P. O., Fernandes, C. B., ... Schuler, B. (2018). Extreme disorder in an ultrahigh-affinity protein complex. *Nature*, *555*(7694), 61–66. <https://doi.org/10.1038/nature25762>
- Bornens, M. (2002). Centrosome composition and microtubule anchoring mechanisms. *Current Opinion in Cell Biology*, *14*(1), 25–34. [https://doi.org/https://doi.org/10.1016/S0955-0674\(01\)00290-3](https://doi.org/https://doi.org/10.1016/S0955-0674(01)00290-3)
- Bornens, M. (2002). Centrosome composition and microtubule anchoring mechanisms. *Current Opinion in Cell Biology*, *14*(1), 25–34. [https://doi.org/https://doi.org/10.1016/S0955-0674\(01\)00290-3](https://doi.org/https://doi.org/10.1016/S0955-0674(01)00290-3)
- Bouchard, J. J., Otero, J. H., Scott, D. C., Szulc, E., Martin, E. W., Sabri, N., ... Mittag, T. (2018). Cancer Mutations of the Tumor Suppressor SPOP Disrupt the Formation of Active, Phase-Separated Compartments. *Molecular Cell*, *72*(1), 19–36.e8. <https://doi.org/10.1016/j.molcel.2018.08.027>
- Brangwynne, C. P., Eckmann, C. R., Courson, D. S., Rybarska, A., Hoege, C., Gharakhani, J., ... Hyman, A. A. (2009). Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science (New York, N.Y.)*, *324*(5935), 1729–1732. <https://doi.org/10.1126/science.1172046>
- Bratek-Skicki, A., Pancsa, R., Meszaros, B., Van Lindt, J., & Tompa, P. (2020). A guide to regulation of the formation of biomolecular condensates. *The FEBS Journal*, *287*(10), 1924–1935. <https://doi.org/10.1111/febs.15254>
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, *68*(6), 394–424. <https://doi.org/10.3322/caac.21492>

- Brent, R., & Ptashne, M. (1985). A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell*, *43*(3 Pt 2), 729–736. [https://doi.org/10.1016/0092-8674\(85\)90246-6](https://doi.org/10.1016/0092-8674(85)90246-6)
- Breslow, D. K., & Holland, A. J. (2019). Mechanism and Regulation of Centriole and Cilium Biogenesis. *Annual Review of Biochemistry*, *88*, 691–724. <https://doi.org/10.1146/annurev-biochem-013118-111153>
- Brody, S. L., Yan, X. H., Wuerffel, M. K., Song, S. K., & Shapiro, S. D. (2000). Ciliogenesis and left-right axis defects in forkhead factor HFH-4-null mice. *American Journal of Respiratory Cell and Molecular Biology*, *23*(1), 45–51. <https://doi.org/10.1165/ajrcmb.23.1.4070>
- Brown, J. D., Lin, C. Y., Duan, Q., Griffin, G., Federation, A., Paranal, R. M., ... Plutzky, J. (2014). NF- κ B directs dynamic super enhancer formation in inflammation and atherogenesis. *Molecular Cell*, *56*(2), 219–231. <https://doi.org/10.1016/j.molcel.2014.08.024>
- Bryce, A. H., & Antonarakis, E. S. (2016). Androgen receptor splice variant 7 in castration-resistant prostate cancer: Clinical considerations. *International Journal of Urology : Official Journal of the Japanese Urological Association*, *23*(8), 646–653. <https://doi.org/10.1111/iju.13134>
- Butler, M. T., & Wallingford, J. B. (2017). Planar cell polarity in development and disease. *Nature Reviews. Molecular Cell Biology*, *18*(6), 375–388. <https://doi.org/10.1038/nrm.2017.11>
- Buttiglieri, C., Tucci, M., Bertaglia, V., Vignani, F., Bironzo, P., Di Maio, M., & Scagliotti, G. V. (2015). Understanding and overcoming the mechanisms of primary and acquired resistance to abiraterone and enzalutamide in castration resistant prostate cancer. *Cancer Treatment Reviews*, *41*(10), 884–892. <https://doi.org/10.1016/j.ctrv.2015.08.002>
- Caillat, C., Fish, A., Pefani, D. E., Taraviras, S., Lygerou, Z., & Perrakis, A. (2015). The structure of the GemC1 coiled coil and its interaction with the Geminin family of coiled-coil proteins. *Acta Crystallographica. Section D, Biological Crystallography*, *71*(Pt 11), 2278–2286. <https://doi.org/10.1107/S1399004715016892>
- Caillat, C., Pefani, D.-E., Gillespie, P. J., Taraviras, S., Blow, J. J., Lygerou, Z., & Perrakis, A. (2013). The Geminin and Idas coiled coils preferentially form a heterodimer that inhibits Geminin function in DNA replication licensing. *The Journal of Biological Chemistry*, *288*(44), 31624–31634. <https://doi.org/10.1074/jbc.M113.491928>
- Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., ... Schultz, N. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discovery*, *2*(5), 401–404. <https://doi.org/10.1158/2159-8290.CD-12-0095>
- Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., & Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell*, *113*(5), 643–655. [https://doi.org/10.1016/s0092-8674\(03\)00392-1](https://doi.org/10.1016/s0092-8674(03)00392-1)
- Chapuy, B., McKeown, M. R., Lin, C. Y., Monti, S., Roemer, M. G. M., Qi, J., ... Bradner, J. E. (2013). Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. *Cancer Cell*, *24*(6), 777–790. <https://doi.org/10.1016/j.ccr.2013.11.003>
- Chen, W., & Roeder, R. G. (2011). Mediator-dependent nuclear receptor function. *Seminars in Cell & Developmental Biology*, *22*(7), 749–758. <https://doi.org/10.1016/j.semcdb.2011.07.026>

- Cheng, S., Castillo, V., & Sliva, D. (2019). CDC20 associated with cancer metastasis and novel mushroom-derived CDC20 inhibitors with antimetastatic activity. *International Journal of Oncology*, *54*(6), 2250–2256. <https://doi.org/10.3892/ijo.2019.4791>
- Chevalier, B., Adamiok, A., Mercey, O., Revinski, D. R., Zaragosi, L.-E., Pasini, A., ... Marcet, B. (2015). miR-34/449 control apical actin network formation during multiciliogenesis through small GTPase pathways. *Nature Communications*, *6*(1), 8386. <https://doi.org/10.1038/ncomms9386>
- Chipumuro, E., Marco, E., Christensen, C. L., Kwiatkowski, N., Zhang, T., Hatheway, C. M., ... George, R. E. (2014). CDK7 inhibition suppresses super-enhancer-linked oncogenic transcription in MYCN-driven cancer. *Cell*, *159*(5), 1126–1139. <https://doi.org/10.1016/j.cell.2014.10.024>
- Chivukula, R. R., Montoro, D. T., Leung, H. M., Yang, J., Shamseldin, H. E., Taylor, M. S., ... Sabatini, D. M. (2020). A human ciliopathy reveals essential functions for NEK10 in airway mucociliary clearance. *Nature Medicine*, *26*(2), 244–251. <https://doi.org/10.1038/s41591-019-0730-x>
- Chmelar, R., Buchanan, G., Need, E. F., Tilley, W., & Greenberg, N. M. (2007). Androgen receptor coregulators and their involvement in the development and progression of prostate cancer. *International Journal of Cancer*, *120*(4), 719–733. <https://doi.org/10.1002/ijc.22365>
- Cho, K. F., Branon, T. C., Udeshi, N. D., Myers, S. A., Carr, S. A., & Ting, A. Y. (2020). Proximity labeling in mammalian cells with TurboID and split-TurboID. *Nature Protocols*, *15*(12), 3971–3999. <https://doi.org/10.1038/s41596-020-0399-0>
- Cho, W.-K., Spille, J.-H., Hecht, M., Lee, C., Li, C., Grube, V., & Cisse, I. I. (2018). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science (New York, N.Y.)*, *361*(6400), 412–415. <https://doi.org/10.1126/science.aar4199>
- Chong, S., Dugast-Darzacq, C., Liu, Z., Dong, P., Dailey, G. M., Cattoglio, C., ... Tjian, R. (2018). Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science (New York, N.Y.)*, *361*(6400). <https://doi.org/10.1126/science.aar2555>
- Cizmecioglu, O., Arnold, M., Bahtz, R., Settele, F., Ehret, L., Haselmann-Weiss, U., ... Hoffmann, I. (2010). Cep152 acts as a scaffold for recruitment of Plk4 and CPAP to the centrosome. *The Journal of Cell Biology*, *191*(4), 731–739. <https://doi.org/10.1083/jcb.201007107>
- Claessens, F., Denayer, S., Van Tilborgh, N., Kerkhofs, S., Helsen, C., & Haelens, A. (2008). Diverse roles of androgen receptor (AR) domains in AR-mediated signaling. *Nuclear Receptor Signaling*, *6*, e008. <https://doi.org/10.1621/nrs.06008>
- Coffey, K., & Robson, C. N. (2012). Regulation of the androgen receptor by post-translational modifications. *The Journal of Endocrinology*, *215*(2), 221–237. <https://doi.org/10.1530/JOE-12-0238>
- Conduit, P. T., Wainman, A., & Raff, J. W. (2015). Centrosome function and assembly in animal cells. *Nature Reviews Molecular Cell Biology*, *16*(10), 611–624. <https://doi.org/10.1038/nrm4062>
- Cottee, M. A., Muschalik, N., Wong, Y. L., Johnson, C. M., Johnson, S., Andreeva, A., ... van Breugel, M. (2013). Crystal structures of the CPAP/STIL complex reveal its role in centriole assembly and human microcephaly. *ELife*, *2*, e01071. <https://doi.org/10.7554/eLife.01071>

- Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, *26*(12), 1367–1372. <https://doi.org/10.1038/nbt.1511>
- Cui, X., Xu, D., Lv, C., Qu, F., He, J., Chen, M., ... Yu, H. (2011). Suppression of MED19 expression by shRNA induces inhibition of cell proliferation and tumorigenesis in human prostate cancer cells. *BMB Reports*, *44*(8), 547–552. <https://doi.org/10.5483/bmbrep.2011.44.8.547>
- Dahiya, U. R., & Heemers, H. V. (2022). Analyzing the Androgen Receptor Interactome in Prostate Cancer: Implications for Therapeutic Intervention. *Cells*, *11*(6). <https://doi.org/10.3390/cells11060936>
- Dalal, K., Roshan-Moniri, M., Sharma, A., Li, H., Ban, F., Hessein, M., ... Rennie, P. S. (2014). Selectively targeting the DNA-binding domain of the androgen receptor as a prospective therapy for prostate cancer. *The Journal of Biological Chemistry*, *289*(38), 26417–26429. <https://doi.org/10.1074/jbc.M114.553818>
- Danielian, P. S., Bender Kim, C. F., Caron, A. M., Vasile, E., Bronson, R. T., & Lees, J. A. (2007). E2f4 is required for normal development of the airway epithelium. *Developmental Biology*, *305*(2), 564–576. <https://doi.org/10.1016/j.ydbio.2007.02.037>
- de Bono, J. S., Logothetis, C. J., Molina, A., Fizazi, K., North, S., Chu, L., ... Scher, H. I. (2011). Abiraterone and increased survival in metastatic prostate cancer. *The New England Journal of Medicine*, *364*(21), 1995–2005. <https://doi.org/10.1056/NEJMoa1014618>
- de Bono, J. S., Oudard, S., Ozguroglu, M., Hansen, S., Machiels, J.-P., Kocak, I., ... Sartor, A. O. (2010). Prednisone plus cabazitaxel or mitoxantrone for metastatic castration-resistant prostate cancer progressing after docetaxel treatment: a randomised open-label trial. *Lancet (London, England)*, *376*(9747), 1147–1154. [https://doi.org/10.1016/S0140-6736\(10\)61389-X](https://doi.org/10.1016/S0140-6736(10)61389-X)
- De Mol, E., Fenwick, R. B., Phang, C. T. W., Buzón, V., Szulc, E., de la Fuente, A., ... Salvatella, X. (2016). EPI-001, A Compound Active against Castration-Resistant Prostate Cancer, Targets Transactivation Unit 5 of the Androgen Receptor. *ACS Chemical Biology*, *11*(9), 2499–2505. <https://doi.org/10.1021/acscchembio.6b00182>
- Dehm, S. M., Schmidt, L. J., Heemers, H. V., Vessella, R. L., & Tindall, D. J. (2008). Splicing of a novel androgen receptor exon generates a constitutively active androgen receptor that mediates prostate cancer therapy resistance. *Cancer Research*, *68*(13), 5469–5477. <https://doi.org/10.1158/0008-5472.CAN-08-0594>
- Dehm, S. M., & Tindall, D. J. (2011). Alternatively spliced androgen receptor variants. *Endocrine-Related Cancer*, *18*(5), R183-96. <https://doi.org/10.1530/ERC-11-0141>
- Delgehyr, N., Sillibourne, J., & Bornens, M. (2005). Microtubule nucleation and anchoring at the centrosome are independent processes linked by ninein function. *Journal of Cell Science*, *118*(Pt 8), 1565–1575. <https://doi.org/10.1242/jcs.02302>
- Delmore, J. E., Issa, G. C., Lemieux, M. E., Rahl, P. B., Shi, J., Jacobs, H. M., ... Mitsiades, C. S. (2011). BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell*, *146*(6), 904–917. <https://doi.org/10.1016/j.cell.2011.08.017>

- Didon, L., Zwick, R. K., Chao, I. W., Walters, M. S., Wang, R., Hackett, N. R., & Crystal, R. G. (2013). RFX3 Modulation of FOXJ1 regulation of cilia genes in the human airway epithelium. *Respiratory Research*, *14*(1), 70. <https://doi.org/10.1186/1465-9921-14-70>
- Ding, N., Zhou, H., Esteve, P.-O., Chin, H. G., Kim, S., Xu, X., ... Boyer, T. G. (2008). Mediator links epigenetic silencing of neuronal gene expression with x-linked mental retardation. *Molecular Cell*, *31*(3), 347–359. <https://doi.org/10.1016/j.molcel.2008.05.023>
- Dingar, D., Kalkat, M., Chan, P.-K., Srikumar, T., Bailey, S. D., Tu, W. B., ... Raught, B. (2015). BioID identifies novel c-MYC interacting partners in cultured cells and xenograft tumors. *Journal of Proteomics*, *118*, 95–111. <https://doi.org/10.1016/j.jpro.2014.09.029>
- Downen, J. M., Fan, Z. P., Hnisz, D., Ren, G., Abraham, B. J., Zhang, L. N., ... Young, R. A. (2014). Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell*, *159*(2), 374–387. <https://doi.org/10.1016/j.cell.2014.09.030>
- Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., ... Obradovic, Z. (2001). Intrinsically disordered protein. *Journal of Molecular Graphics & Modelling*, *19*(1), 26–59. [https://doi.org/10.1016/s1093-3263\(00\)00138-8](https://doi.org/10.1016/s1093-3263(00)00138-8)
- Dunker, A. K., Cortese, M. S., Romero, P., Iakoucheva, L. M., & Uversky, V. N. (2005). Flexible nets. The roles of intrinsic disorder in protein interaction networks. *The FEBS Journal*, *272*(20), 5129–5148. <https://doi.org/10.1111/j.1742-4658.2005.04948.x>
- Eftekharzadeh, B., Banduseela, V. C., Chiesa, G., Martínez-Cristóbal, P., Rauch, J. N., Nath, S. R., ... Salvatella, X. (2019). Hsp70 and Hsp40 inhibit an inter-domain interaction necessary for transcriptional activity in the androgen receptor. *Nature Communications*, *10*(1), 3562. <https://doi.org/10.1038/s41467-019-11594-y>
- Elbaum-Garfinkle, S., Kim, Y., Szczepaniak, K., Chen, C. C.-H., Eckmann, C. R., Myong, S., & Brangwynne, C. P. (2015). The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(23), 7189–7194. <https://doi.org/10.1073/pnas.1504822112>
- Feric, M., Vaidya, N., Harmon, T. S., Mitrea, D. M., Zhu, L., Richardson, T. M., ... Brangwynne, C. P. (2016). Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell*, *165*(7), 1686–1697. <https://doi.org/https://doi.org/10.1016/j.cell.2016.04.047>
- Firat-Karalar, E. N., Sante, J., Elliott, S., & Stearns, T. (2014). Proteomic analysis of mammalian sperm cells identifies new components of the centrosome. *Journal of Cell Science*, *127*(Pt 19), 4128–4133. <https://doi.org/10.1242/jcs.157008>
- Fischer, M., & Müller, G. A. (2017). Cell cycle transcription control: DREAM/MuvB and RB-E2F complexes. *Critical Reviews in Biochemistry and Molecular Biology*, *52*(6), 638–662. <https://doi.org/10.1080/10409238.2017.1360836>
- Fujita, H., Yoshino, Y., & Chiba, N. (2016). Regulation of the centrosome cycle. *Molecular & Cellular Oncology*, *3*(2), e1075643. <https://doi.org/10.1080/23723556.2015.1075643>

- Fujitani, M., Sato, R., & Yamashita, T. (2017). Loss of p73 in ependymal cells during the perinatal period leads to aqueductal stenosis. *Scientific Reports*, 7(1), 12007. <https://doi.org/10.1038/s41598-017-12105-z>
- Fulton, D. L., Sundararajan, S., Badis, G., Hughes, T. R., Wasserman, W. W., Roach, J. C., & Sladek, R. (2009). TFCat: the curated catalog of mouse and human transcription factors. *Genome Biology*, 10(3), R29. <https://doi.org/10.1186/gb-2009-10-3-r29>
- Funk, M. C., Bera, A. N., Menchen, T., Kualess, G., Thriene, K., Lienkamp, S. S., ... Arnold, S. J. (2015). Cyclin O (Ccno) functions during deuterosome-mediated centriole amplification of multiciliated cells. *The EMBO Journal*, 34(8), 1078–1089. <https://doi.org/10.15252/emboj.201490805>
- Furlong, E. E. M., & Levine, M. (2018). Developmental enhancers and chromosome topology. *Science (New York, N.Y.)*, 361(6409), 1341–1345. <https://doi.org/10.1126/science.aau0320>
- Fuxreiter, M., Simon, I., & Bondos, S. (2011). Dynamic protein-DNA recognition: beyond what can be seen. *Trends in Biochemical Sciences*, 36(8), 415–423. <https://doi.org/10.1016/j.tibs.2011.04.006>
- Gao, A., Shrinivas, K., Lepeudry, P., Suzuki, H. I., Sharp, P. A., & Chakraborty, A. K. (2018). Evolution of weak cooperative interactions for biological specificity. *Proceedings of the National Academy of Sciences of the United States of America*, 115(47), E11053–E11060. <https://doi.org/10.1073/pnas.1815912115>
- Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., ... Schultz, N. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science Signaling*, 6(269), p11. <https://doi.org/10.1126/scisignal.2004088>
- Gao, W., Bohl, C. E., & Dalton, J. T. (2005). Chemistry and structural biology of androgen receptor. *Chemical Reviews*, 105(9), 3352–3370. <https://doi.org/10.1021/cr020456u>
- Gao, X., Burris III, H. A., Vuky, J., Dreicer, R., Sartor, A. O., Sternberg, C. N., ... Petrylak, D. P. (2022). Phase 1/2 study of ARV-110, an androgen receptor (AR) PROTAC degrader, in metastatic castration-resistant prostate cancer (mCRPC). *Journal of Clinical Oncology*, 40(6_suppl), 17. https://doi.org/10.1200/JCO.2022.40.6_suppl.017
- Goetz, S. C., & Anderson, K. V. (2010). The primary cilium: a signalling centre during vertebrate development. *Nature Reviews Genetics*, 11(5), 331–344. <https://doi.org/10.1038/nrg2774>
- Gomes, E., & Shorter, J. (2019). The molecular language of membraneless organelles. *The Journal of Biological Chemistry*, 294(18), 7115–7127. <https://doi.org/10.1074/jbc.TM118.001192>
- Gonzalez-Cano, L., Fuertes-Alvarez, S., Robledinos-Anton, N., Bizy, A., Villena-Cortes, A., Fariñas, I., ... Marin, M. C. (2016). p73 is required for ependymal cell maturation and neurogenic SVZ cytoarchitecture. *Developmental Neurobiology*, 76(7), 730–747. <https://doi.org/10.1002/dneu.22356>
- Gottlieb, B., Beitel, L. K., Nadarajah, A., Paliouras, M., & Trifiro, M. (2012). The androgen receptor gene mutations database: 2012 update. *Human Mutation*, 33(5), 887–894. <https://doi.org/10.1002/humu.22046>

- Grasso, C. S., Wu, Y.-M., Robinson, D. R., Cao, X., Dhanasekaran, S. M., Khan, A. P., ... Tomlins, S. A. (2012). The mutational landscape of lethal castration-resistant prostate cancer. *Nature*, *487*(7406), 239–243. <https://doi.org/10.1038/nature11125>
- Grosveld, F., van Staalduinen, J., & Stadhouders, R. (2021). Transcriptional Regulation by (Super)Enhancers: From Discovery to Mechanisms. *Annual Review of Genomics and Human Genetics*, *22*, 127–146. <https://doi.org/10.1146/annurev-genom-122220-093818>
- Guo, X., Bulyk, M. L., & Hartemink, A. J. (2012). Intrinsic disorder within and flanking the DNA-binding domains of human transcription factors. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, 104–115.
- Guo, Y. E., Manteiga, J. C., Henninger, J. E., Sabari, B. R., Dall'Agnesse, A., Hannett, N. M., ... Young, R. A. (2019). Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature*, *572*(7770), 543–548. <https://doi.org/10.1038/s41586-019-1464-0>
- Guo, Z., Yang, X., Sun, F., Jiang, R., Linn, D. E., Chen, H., ... Qiu, Y. (2009). A novel androgen receptor splice variant is up-regulated during prostate cancer progression and promotes androgen depletion-resistant growth. *Cancer Research*, *69*(6), 2305–2313. <https://doi.org/10.1158/0008-5472.CAN-08-3795>
- Guseh, J. S., Bores, S. A., Stanger, B. Z., Zhou, Q., Anderson, W. J., Melton, D. A., & Rajagopal, J. (2009). Notch signaling promotes airway mucous metaplasia and inhibits alveolar development. *Development (Cambridge, England)*, *136*(10), 1751–1759. <https://doi.org/10.1242/dev.029249>
- Habchi, J., Tompa, P., Longhi, S., & Uversky, V. N. (2014). Introducing protein intrinsic disorder. *Chemical Reviews*, *114*(13), 6561–6588. <https://doi.org/10.1021/cr400514h>
- Haffner, M. C., Zwart, W., Roudier, M. P., True, L. D., Nelson, W. G., Epstein, J. I., ... Yegnasubramanian, S. (2021). Genomic and phenotypic heterogeneity in prostate cancer. *Nature Reviews Urology*, *18*(2), 79–92. <https://doi.org/10.1038/s41585-020-00400-w>
- Haile, S., & Sadar, M. D. (2011). Androgen receptor and its splice variants in prostate cancer. *Cellular and Molecular Life Sciences : CMLS*, *68*(24), 3971–3981. <https://doi.org/10.1007/s00018-011-0766-7>
- Hay, E. D. (1995). An overview of epithelio-mesenchymal transformation. *Acta Anatomica*, *154*(1), 8–20. <https://doi.org/10.1159/000147748>
- Haynes, C., Oldfield, C. J., Ji, F., Klitgord, N., Cusick, M. E., Radivojac, P., ... Iakoucheva, L. M. (2006). Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. *PLoS Computational Biology*, *2*(8), e100–e100. <https://doi.org/10.1371/journal.pcbi.0020100>
- He, B., Gampe, R. T. J., Kole, A. J., Hnat, A. T., Stanley, T. B., An, G., ... Wilson, E. M. (2004). Structural basis for androgen receptor interdomain and coactivator interactions suggests a transition in nuclear receptor activation function dominance. *Molecular Cell*, *16*(3), 425–438. <https://doi.org/10.1016/j.molcel.2004.09.036>
- Heidenreich, A., Bastian, P. J., Bellmunt, J., Bolla, M., Joniau, S., van der Kwast, T., ... Mottet, N. (2014). EAU guidelines on prostate cancer. Part II: Treatment of advanced, relapsing, and castration-resistant prostate cancer. *European Urology*, *65*(2), 467–479. <https://doi.org/10.1016/j.eururo.2013.11.002>

- Heinlein, C. A., & Chang, C. (2004). Androgen receptor in prostate cancer. *Endocrine Reviews*, *25*(2), 276–308. <https://doi.org/10.1210/er.2002-0032>
- Heumann, A., Heinemann, N., Hube-Magg, C., Lang, D. S., Grupp, K., Kluth, M., ... Hinsch, A. (2018). High BCAR1 expression is associated with early PSA recurrence in ERG negative prostate cancer. *BMC Cancer*, *18*(1), 37. <https://doi.org/10.1186/s12885-017-3956-3>
- Higano, C. S., & Crawford, E. D. (2011). New and emerging agents for the treatment of castration-resistant prostate cancer. *Urologic Oncology*, *29*(6 Suppl), S1-8. <https://doi.org/10.1016/j.urolonc.2011.08.013>
- Hirokawa, N., Tanaka, Y., & Okada, Y. (2012). Cilia, KIF3 molecular motor and nodal flow. *Current Opinion in Cell Biology*, *24*(1), 31–39. <https://doi.org/https://doi.org/10.1016/j.ceb.2012.01.002>
- Hirokawa, N., Tanaka, Y., & Okada, Y. (2012). Cilia, KIF3 molecular motor and nodal flow. *Current Opinion in Cell Biology*, *24*(1), 31–39. <https://doi.org/https://doi.org/10.1016/j.ceb.2012.01.002>
- Hnisz, D., Abraham, B. J., Lee, T. I., Lau, A., Saint-André, V., Sigova, A. A., ... Young, R. A. (2013). Super-enhancers in the control of cell identity and disease. *Cell*, *155*(4), 934–947. <https://doi.org/10.1016/j.cell.2013.09.053>
- Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K., & Sharp, P. A. (2017). A Phase Separation Model for Transcriptional Control. *Cell*, *169*(1), 13–23. <https://doi.org/10.1016/j.cell.2017.02.007>
- Hofweber, M., & Dormann, D. (2019). Friend or foe-Post-translational modifications as regulators of phase separation and RNP granule dynamics. *The Journal of Biological Chemistry*, *294*(18), 7137–7150. <https://doi.org/10.1074/jbc.TM118.001189>
- Holehouse, A. S., Ginell, G. M., Griffith, D., & Böke, E. (2021). Clustering of Aromatic Residues in Prion-like Domains Can Tune the Formation, State, and Organization of Biomolecular Condensates. *Biochemistry*, *60*(47), 3566–3581. <https://doi.org/10.1021/acs.biochem.1c00465>
- Hörnberg, E., Ylitalo, E. B., Crnalic, S., Antti, H., Stattin, P., Widmark, A., ... Wikström, P. (2011). Expression of androgen receptor splice variants in prostate cancer bone metastases is associated with castration-resistance and short survival. *PLoS One*, *6*(4), e19059. <https://doi.org/10.1371/journal.pone.0019059>
- Hsu, J., & Sage, J. (2016). Novel functions for the transcription factor E2F4 in development and disease. *Cell Cycle (Georgetown, Tex.)*, *15*(23), 3183–3190. <https://doi.org/10.1080/15384101.2016.1234551>
- Hu, J., Wang, G., & Sun, T. (2017). Dissecting the roles of the androgen receptor in prostate cancer from molecular perspectives. *Tumour Biology: The Journal of the International Society for Oncodevelopmental Biology and Medicine*, *39*(5), 1010428317692259. <https://doi.org/10.1177/1010428317692259>
- Hu, R., Dunn, T. A., Wei, S., Isharwal, S., Veltri, R. W., Humphreys, E., ... Luo, J. (2009). Ligand-independent androgen receptor variants derived from splicing of cryptic exons signify hormone-refractory prostate cancer. *Cancer Research*, *69*(1), 16–22. <https://doi.org/10.1158/0008-5472.CAN-08-2764>

- Hu, R., Isaacs, W. B., & Luo, J. (2011). A snapshot of the expression signature of androgen receptor splicing variants and their distinctive transcriptional activities. *The Prostate*, *71*(15), 1656–1667. <https://doi.org/10.1002/pros.21382>
- Hu, R., Lu, C., Mostaghel, E. A., Yegnasubramanian, S., Gurel, M., Tannahill, C., ... Luo, J. (2012). Distinct transcriptional programs mediated by the ligand-dependent full-length androgen receptor and its splice variants in castration-resistant prostate cancer. *Cancer Research*, *72*(14), 3457–3462. <https://doi.org/10.1158/0008-5472.CAN-11-3892>
- Huang, Y., Jiang, X., Liang, X., & Jiang, G. (2018). Molecular and cellular mechanisms of castration resistant prostate cancer. *Oncology Letters*, *15*(5), 6063–6076. <https://doi.org/10.3892/ol.2018.8123>
- Huggins, C., & Hodges, C. V. (2002). Studies on prostatic cancer: I. The effect of castration, of estrogen and of androgen injection on serum phosphatases in metastatic carcinoma of the prostate. 1941. *The Journal of Urology*, *168*(1), 9–12. [https://doi.org/10.1016/s0022-5347\(05\)64820-3](https://doi.org/10.1016/s0022-5347(05)64820-3)
- Hyman, A. A., Weber, C. A., & Jülicher, F. (2014). Liquid-liquid phase separation in biology. *Annual Review of Cell and Developmental Biology*, *30*, 39–58. <https://doi.org/10.1146/annurev-cellbio-100913-013325>
- Ianculescu, I., Wu, D.-Y., Siegmund, K. D., & Stallcup, M. R. (2012). Selective roles for cAMP response element-binding protein binding protein and p300 protein as coregulators for androgen-regulated gene expression in advanced prostate cancer cells. *The Journal of Biological Chemistry*, *287*(6), 4000–4013. <https://doi.org/10.1074/jbc.M111.300194>
- Imai, K., & Takaoka, A. (2006). Comparing antibody and small-molecule therapies for cancer. *Nature Reviews. Cancer*, *6*(9), 714–727. <https://doi.org/10.1038/nrc1913>
- Imberg-Kazdan, K., Ha, S., Greenfield, A., Poultney, C. S., Bonneau, R., Logan, S. K., & Garabedian, M. J. (2013). A genome-wide RNA interference screen identifies new regulators of androgen receptor function in prostate cancer cells. *Genome Research*, *23*(4), 581–591. <https://doi.org/10.1101/gr.144774.112>
- Ito, D., Zitouni, S., Jana, S. C., Duarte, P., Surkont, J., Carvalho-Santos, Z., ... Bettencourt-Dias, M. (2019). Pericentrin-mediated SAS-6 recruitment promotes centriole assembly. *eLife*, *8*. <https://doi.org/10.7554/eLife.41418>
- Jacquet, B. V., Salinas-Mondragon, R., Liang, H., Therit, B., Buie, J. D., Dykstra, M., ... Ghashghaei, H. T. (2009). FoxJ1-dependent gene expression is required for differentiation of radial glia into ependymal cells and a subset of astrocytes in the postnatal brain. *Development (Cambridge, England)*, *136*(23), 4021–4031. <https://doi.org/10.1242/dev.041129>
- Jäggle, S., Busch, H., Freißen, V., Beyes, S., Schrempp, M., Boerries, M., & Hecht, A. (2017). SNAIL1-mediated downregulation of FOXA proteins facilitates the inactivation of transcriptional enhancer elements at key epithelial genes in colorectal cancer cells. *PLoS Genetics*, *13*(11), e1007109. <https://doi.org/10.1371/journal.pgen.1007109>
- Jenster, G., van der Korput, H. A., Trapman, J., & Brinkmann, A. O. (1995). Identification of two transcription activation units in the N-terminal domain of the human androgen receptor. *The Journal of Biological Chemistry*, *270*(13), 7341–7346. <https://doi.org/10.1074/jbc.270.13.7341>

- Jeronimo, C., & Robert, F. (2017). The Mediator Complex: At the Nexus of RNA Polymerase II Transcription. *Trends in Cell Biology*, 27(10), 765–783. <https://doi.org/10.1016/j.tcb.2017.07.001>
- Ji, X., Dadon, D. B., Powell, B. E., Fan, Z. P., Borges-Rivera, D., Shachar, S., ... Young, R. A. (2016). 3D Chromosome Regulatory Landscape of Human Pluripotent Cells. *Cell Stem Cell*, 18(2), 262–275. <https://doi.org/10.1016/j.stem.2015.11.007>
- Jin, F., Claessens, F., & Fondell, J. D. (2012). Regulation of androgen receptor-dependent transcription by coactivator MED1 is mediated through a newly discovered noncanonical binding motif. *The Journal of Biological Chemistry*, 287(2), 858–870. <https://doi.org/10.1074/jbc.M111.304519>
- Kadoch, C., Hargreaves, D. C., Hodges, C., Elias, L., Ho, L., Ranish, J., & Crabtree, G. R. (2013). Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nature Genetics*, 45(6), 592–601. <https://doi.org/10.1038/ng.2628>
- Kang, Y., & Massagué, J. (2004). Epithelial-mesenchymal transitions: twist in development and metastasis. *Cell*, 118(3), 277–279. <https://doi.org/10.1016/j.cell.2004.07.011>
- Kantoff, P. W., Higano, C. S., Shore, N. D., Berger, E. R., Small, E. J., Penson, D. F., ... Schellhammer, P. F. (2010). Sipuleucel-T immunotherapy for castration-resistant prostate cancer. *The New England Journal of Medicine*, 363(5), 411–422. <https://doi.org/10.1056/NEJMoa1001294>
- Karamitros, D., Patmanidi, A. L., Kotantaki, P., Potocnik, A. J., Bähr-Ivacevic, T., Benes, V., ... Taraviras, S. (2015). Geminin deletion increases the number of fetal hematopoietic stem cells by affecting the expression of key transcription factors. *Development (Cambridge, England)*, 142(1), 70–81. <https://doi.org/10.1242/dev.109454>
- Karra, H., Repo, H., Ahonen, I., Löyttyniemi, E., Pitkänen, R., Lintunen, M., ... Kronqvist, P. (2014). Cdc20 and securin overexpression predict short-term breast cancer survival. *British Journal of Cancer*, 110(12), 2905–2913. <https://doi.org/10.1038/bjc.2014.252>
- Kassabov, S. R., Zhang, B., Persinger, J., & Bartholomew, B. (2003). SWI/SNF unwraps, slides, and rewraps the nucleosome. *Molecular Cell*, 11(2), 391–403. [https://doi.org/10.1016/s1097-2765\(03\)00039-x](https://doi.org/10.1016/s1097-2765(03)00039-x)
- Kato, M., Han, T. W., Xie, S., Shi, K., Du, X., Wu, L. C., ... McKnight, S. L. (2012). Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell*, 149(4), 753–767. <https://doi.org/10.1016/j.cell.2012.04.017>
- Kaufmann, R., Straussberg, R., Mandel, H., Fattal-Valevski, A., Ben-Zeev, B., Naamati, A., ... Elpeleg, O. (2010). Infantile cerebral and cerebellar atrophy is associated with a mutation in the MED17 subunit of the transcription preinitiation mediator complex. *American Journal of Human Genetics*, 87(5), 667–670. <https://doi.org/10.1016/j.ajhg.2010.09.016>
- Kawaguchi, T., Takeda, A., Ogiyama, Y., Yamauchi, Y., Murata, M., Suzuki, T., ... Morozumi, K. (2012). Peculiar renal endarteritis in a patient with acute endocapillary proliferative glomerulonephritis. *Clinical Nephrology*, 77(2), 151–155. <https://doi.org/10.5414/CN106909>
- Keegan, L., Gill, G., & Ptashne, M. (1986). Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. *Science (New York, N.Y.)*, 231(4739), 699–704. <https://doi.org/10.1126/science.3080805>

- Kemp, A. (1996). Role of epidermal cilia in development of the Australian lungfish, *Neoceratodus forsteri* (Osteichthyes: Dipnoi). *Journal of Morphology*, 228(2), 203–221. [https://doi.org/10.1002/\(SICI\)1097-4687\(199605\)228:2<203::AID-JMOR9>3.0.CO;2-5](https://doi.org/10.1002/(SICI)1097-4687(199605)228:2<203::AID-JMOR9>3.0.CO;2-5)
- Kessler, M., Hoffmann, K., Brinkmann, V., Thieck, O., Jackisch, S., Toelle, B., ... Meyer, T. F. (2015). The Notch and Wnt pathways regulate stemness and differentiation in human fallopian tube organoids. *Nature Communications*, 6, 8989. <https://doi.org/10.1038/ncomms9989>
- Kim, H. Y., Jackson, T. R., Stuckenholtz, C., & Davidson, L. A. (2020). Tissue mechanics drives regeneration of a mucociliated epidermis on the surface of *Xenopus* embryonic aggregates. *Nature Communications*, 11(1), 665. <https://doi.org/10.1038/s41467-020-14385-y>
- Kim, S., & Dynlacht, B. D. (2013). Assembling a primary cilium. *Current Opinion in Cell Biology*, 25(4), 506–511. <https://doi.org/https://doi.org/10.1016/j.ceb.2013.04.011>
- Kim, S., Ma, L., Shokhirev, M. N., Quigley, I., & Kintner, C. (2018). Multicilin and activated E2f4 induce multiciliated cell differentiation in primary fibroblasts. *Scientific Reports*, 8(1), 12369. <https://doi.org/10.1038/s41598-018-30791-1>
- Kirby, M., Hirst, C., & Crawford, E. D. (2011). Characterising the castration-resistant prostate cancer population: a systematic review. *International Journal of Clinical Practice*, 65(11), 1180–1192. <https://doi.org/10.1111/j.1742-1241.2011.02799.x>
- Kiyokawa, H., & Morimoto, M. (2020). Notch signaling in the mammalian respiratory system, specifically the trachea and lungs, in development, homeostasis, regeneration, and disease. *Development, Growth & Differentiation*, 62(1), 67–79. <https://doi.org/10.1111/dgd.12628>
- Klos Dehring, D. A., Vladar, E. K., Werner, M. E., Mitchell, J. W., Hwang, P., & Mitchell, B. J. (2013). Deuterosome-mediated centriole biogenesis. *Developmental Cell*, 27(1), 103–112. <https://doi.org/10.1016/j.devcel.2013.08.021>
- Kohli, M., Ho, Y., Hillman, D. W., Van Etten, J. L., Henzler, C., Yang, R., ... Dehm, S. M. (2017). Androgen Receptor Variant AR-V9 Is Coexpressed with AR-V7 in Prostate Cancer Metastases and Predicts Abiraterone Resistance. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, 23(16), 4704–4715. <https://doi.org/10.1158/1078-0432.CCR-17-0017>
- Kramer-Zucker, A. G., Olale, F., Haycraft, C. J., Yoder, B. K., Schier, A. F., & Drummond, I. A. (2005). Cilia-driven fluid flow in the zebrafish pronephros, brain and Kupffer's vesicle is required for normal organogenesis. *Development (Cambridge, England)*, 132(8), 1907–1921. <https://doi.org/10.1242/dev.01772>
- Kumar, R., Betney, R., Li, J., Thompson, E. B., & McEwan, I. J. (2004). Induced alpha-helix structure in AF1 of the androgen receptor upon binding transcription factor TFIIF. *Biochemistry*, 43(11), 3008–3013. <https://doi.org/10.1021/bi035934p>
- Kumari, S., Senapati, D., & Heemers, H. V. (2017). Rationale for the development of alternative forms of androgen deprivation therapy. *Endocrine-Related Cancer*, 24(8), R275–R295. <https://doi.org/10.1530/ERC-17-0121>
- Kurihara, M., Kato, K., Sanbo, C., Shigenobu, S., Ohkawa, Y., Fuchigami, T., & Miyanari, Y. (2020). Genomic Profiling by ALaP-Seq Reveals Transcriptional Regulation by PML Bodies through

- DNMT3A Exclusion. *Molecular Cell*, 78(3), 493-505.e8.
<https://doi.org/10.1016/j.molcel.2020.04.004>
- Kuuluvainen, E., Domènech-Moreno, E., Niemelä, E. H., & Mäkelä, T. P. (2018). Depletion of Mediator Kinase Module Subunits Represses Superenhancer-Associated Genes in Colon Cancer Cells. *Molecular and Cellular Biology*, 38(11). <https://doi.org/10.1128/MCB.00573-17>
- Kwon, I., Kato, M., Xiang, S., Wu, L., Theodoropoulos, P., Mirzaei, H., ... McKnight, S. L. (2013). Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. *Cell*, 155(5), 1049–1060. <https://doi.org/10.1016/j.cell.2013.10.033>
- Kwon, O. K., Ha, Y.-S., Na, A.-Y., Chun, S. Y., Kwon, T. G., Lee, J. N., & Lee, S. (2020). Identification of Novel Prognosis and Prediction Markers in Advanced Prostate Cancer Tissues Based on Quantitative Proteomics. *Cancer Genomics & Proteomics*, 17(2), 195–208.
<https://doi.org/10.21873/cgp.20180>
- Kyrousi, C., Arbi, M., Pilz, G.-A. A., Pefani, D.-E. E., Lalioti, M.-E. E., Ninkovic, J., ... Taraviras, S. (2015). Mcdas and gemc1 are key regulators for the generation of multiciliated ependymal cells in the adult neurogenic niche. *Development (Cambridge)*, 142(21), 3661–3674.
<https://doi.org/10.1242/dev.126342>
- Lalioti, M.-E. E., Arbi, M., Loukas, I., Kaplani, K., Kalogeropoulou, A., Lokka, G., ... Taraviras, S. (2019). GemC1 governs multiciliogenesis through direct interaction with and transcriptional regulation of p73. *Journal of Cell Science*, 132(11). <https://doi.org/10.1242/jcs.228684>
- Lambert, S. A., Jolma, A., Campitelli, L. F., Das, P. K., Yin, Y., Albu, M., ... Weirauch, M. T. (2018). The Human Transcription Factors. *Cell*, 172(4), 650–665. <https://doi.org/10.1016/j.cell.2018.01.029>
- Lancaster, M. A., Schroth, J., & Gleeson, J. G. (2011). Subcellular spatial regulation of canonical Wnt signalling at the primary cilium. *Nature Cell Biology*, 13(6), 700–707.
<https://doi.org/10.1038/ncb2259>
- Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P., & Jørgensen, T. J. D. (2005). Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Molecular & Cellular Proteomics : MCP*, 4(7), 873–886.
<https://doi.org/10.1074/mcp.T500007-MCP200>
- Launonen, K.-M., Paakinaho, V., Sigismondo, G., Malinen, M., Sironen, R., Hartikainen, J. M., ... Palvimo, J. J. (2021). Chromatin-directed proteomics-identified network of endogenous androgen receptor in prostate cancer cells. *Oncogene*, 40(27), 4567–4579. <https://doi.org/10.1038/s41388-021-01887-2>
- Le, N. Q. K., Yapp, E. K. Y., Nagasundaram, N., & Yeh, H.-Y. (2019). Classifying Promoters by Interpreting the Hidden Information of DNA Sequences via Deep Learning and Combination of Continuous FastText N-Grams. *Frontiers in Bioengineering and Biotechnology*, 7, 305.
<https://doi.org/10.3389/fbioe.2019.00305>
- Leal, A., Huehne, K., Bauer, F., Sticht, H., Berger, P., Suter, U., ... Rautenstrauss, B. (2009). Identification of the variant Ala335Val of MED25 as responsible for CMT2B2: molecular data, functional studies of the SH3 recognition motif and correlation between wild-type MED25 and

- PMP22 RNA levels in CMT1A animal models. *Neurogenetics*, *10*(4), 275–287.
<https://doi.org/10.1007/s10048-009-0183-3>
- Lear, T., Dunn, S. R., McKelvey, A. C., Mir, A., Evankovich, J., Chen, B. B., & Liu, Y. (2017). RING finger protein 113A regulates C-X-C chemokine receptor type 4 stability and signaling. *American Journal of Physiology. Cell Physiology*, *313*(5), C584–C592. <https://doi.org/10.1152/ajpcell.00193.2017>
- Lewis, M., & Stracker, T. H. (2021). Transcriptional regulation of multiciliated cell differentiation. *Seminars in Cell & Developmental Biology*, *110*, 51–60. <https://doi.org/10.1016/j.semcdb.2020.04.007>
- Li, L., Cao, Y., Zhou, H., Li, Y., He, B., Zhou, X., ... Ye, L. (2018). Knockdown of CCNO decreases the tumorigenicity of gastric cancer by inducing apoptosis. *Oncotargets and Therapy*, *11*, 7471–7481. <https://doi.org/10.2147/OTT.S176252>
- Lim, J. J., Lima, P. D. A., Salehi, R., Lee, D. R., & Tsang, B. K. (2017). Regulation of androgen receptor signaling by ubiquitination during folliculogenesis and its possible dysregulation in polycystic ovarian syndrome. *Scientific Reports*, *7*(1), 10272. <https://doi.org/10.1038/s41598-017-09880-0>
- Lin, Y.-N., Wu, C.-T., Lin, Y.-C., Hsu, W.-B., Tang, C.-J. C., Chang, C.-W., & Tang, T. K. (2013). CEP120 interacts with CPAP and positively regulates centriole elongation. *The Journal of Cell Biology*, *202*(2), 211–219. <https://doi.org/10.1083/jcb.201212060>
- Lindeman, G. J., Dagnino, L., Gaubatz, S., Xu, Y., Bronson, R. T., Warren, H. B., & Livingston, D. M. (1998). A specific, nonproliferative role for E2F-5 in choroid plexus function revealed by gene targeting. *Genes & Development*, *12*(8), 1092–1098. <https://doi.org/10.1101/gad.12.8.1092>
- Liu, G., Sprenger, C., Sun, S., Epilepsia, K. S., Haugk, K., Zhang, X., ... Plymate, S. (2013). AR variant ARv567es induces carcinogenesis in a novel transgenic mouse model of prostate cancer. *Neoplasia (New York, N.Y.)*, *15*(9), 1009–1017. <https://doi.org/10.1593/neo.13784>
- Liu, J., Perumal, N. B., Oldfield, C. J., Su, E. W., Uversky, V. N., & Dunker, A. K. (2006). Intrinsic disorder in transcription factors. *Biochemistry*, *45*(22), 6873–6888. <https://doi.org/10.1021/bi0602718>
- Liu, L., Zhang, P., Shao, Y., Quan, F., & Li, H. (2019). Knockdown of FOXJ1 inhibits the proliferation, migration, invasion, and glycolysis in laryngeal squamous cell carcinoma cells. *Journal of Cellular Biochemistry*, *120*(9), 15874–15882. <https://doi.org/10.1002/jcb.28858>
- Liu, Z., & Huang, Y. (2014). Advantages of proteins being disordered. *Protein Science: A Publication of the Protein Society*, *23*(5), 539–550. <https://doi.org/10.1002/pro.2443>
- Lizé, M., Klimke, A., & Dobbelstein, M. (2011). MicroRNA-449 in cell fate determination. *Cell Cycle*, *10*(17), 2874–2882. <https://doi.org/10.4161/cc.10.17.17181>
- Locke, J. A., Guns, E. S., Lubik, A. A., Adomat, H. H., Hendy, S. C., Wood, C. A., ... Nelson, C. C. (2008). Androgen levels increase by intratumoral de novo steroidogenesis during progression of castration-resistant prostate cancer. *Cancer Research*, *68*(15), 6407–6415. <https://doi.org/10.1158/0008-5472.CAN-07-5997>

- LoMastro, G. M., Drown, C. G., Maryniak, A. L., Jewett, C. E., Strong, M. A., & Holland, A. J. (2022). PLK4 drives centriole amplification and apical surface area expansion in multiciliated cells. *ELife*, *11*. <https://doi.org/10.7554/eLife.80643>
- Loo, C.-S., Gatchalian, J., Liang, Y., Leblanc, M., Xie, M., Ho, J., ... Zheng, Y. (2020). A Genome-wide CRISPR Screen Reveals a Role for the Non-canonical Nucleosome-Remodeling BAF Complex in Foxp3 Expression and Regulatory T Cell Function. *Immunity*, *53*(1), 143-157.e8. <https://doi.org/10.1016/j.immuni.2020.06.011>
- Lopes, C. A. M., Jana, S. C., Cunha-Ferreira, I., Zitouni, S., Bento, I., Duarte, P., ... Bettencourt-Dias, M. (2015). PLK4 trans-Autoactivation Controls Centriole Biogenesis in Space. *Developmental Cell*, *35*(2), 222–235. <https://doi.org/10.1016/j.devcel.2015.09.020>
- Loukas, I., Skamnelou, M., Tsaridou, S., Bournaka, S., Grigoriadis, S., Taraviras, S., ... Arbi, M. (2021). Fine-tuning multiciliated cell differentiation at the post-transcriptional level: contribution of miR-34/449 family members. *Biological Reviews of the Cambridge Philosophical Society*, *96*(5), 2321–2332. <https://doi.org/10.1111/brv.12755>
- Lovén, J., Hoke, H. A., Lin, C. Y., Lau, A., Orlando, D. A., Vakoc, C. R., ... Young, R. A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell*, *153*(2), 320–334. <https://doi.org/10.1016/j.cell.2013.03.036>
- Lu, H., Yu, D., Hansen, A. S., Ganguly, S., Liu, R., Heckert, A., ... Zhou, Q. (2018). Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nature*, *558*(7709), 318–323. <https://doi.org/10.1038/s41586-018-0174-3>
- Luo, J. (2016). Development of AR-V7 as a putative treatment selection marker for metastatic castration-resistant prostate cancer. *Asian Journal of Andrology*, *18*(4), 580–585. <https://doi.org/10.4103/1008-682X.178490>
- Ma, L., Quigley, I., Omran, H., & Kintner, C. (2014). Multicilin drives centriole biogenesis via E2f proteins. *Genes & Development*, *28*(13), 1461–1471. <https://doi.org/10.1101/gad.243832.114>
- Ma, M., & Jiang, Y.-J. (2007). Jagged2a-notch signaling mediates cell fate choice in the zebrafish pronephric duct. *PLoS Genetics*, *3*(1), e18. <https://doi.org/10.1371/journal.pgen.0030018>
- Mansour, M. R., Abraham, B. J., Anders, L., Berezovskaya, A., Gutierrez, A., Durbin, A. D., ... Look, A. T. (2014). Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science (New York, N.Y.)*, *346*(6215), 1373–1377. <https://doi.org/10.1126/science.1259037>
- Mao, Y. S., Zhang, B., & Spector, D. L. (2011). Biogenesis and function of nuclear bodies. *Trends in Genetics : TIG*, *27*(8), 295–306. <https://doi.org/10.1016/j.tig.2011.05.006>
- Marcet, B., Chevalier, B., Coraux, C., Kodjabachian, L., & Barbry, P. (2011). MicroRNA-based silencing of Delta/Notch signaling promotes multiple cilia formation. *Cell Cycle*, *10*(17), 2858–2864. <https://doi.org/10.4161/cc.10.17.17011>
- Marjanović, M., Sánchez-Huertas, C., Terré, B., Gómez, R., Scheel, J. F., Pacheco, S., ... Stracker, T. H. (2015). CEP63 deficiency promotes p53-dependent microcephaly and reveals a role for

- the centrosome in meiotic recombination. *Nature Communications*, 6, 7676. <https://doi.org/10.1038/ncomms8676>
- Marques, M. M., Villoch-Fernandez, J., Maeso-Alonso, L., Fuertes-Alvarez, S., & Marin, M. C. (2019). The Trp73 mutant mice: A ciliopathy model that uncouples ciliogenesis from planar cell polarity. *Frontiers in Genetics*, 10(MAR), 1–9. <https://doi.org/10.3389/fgene.2019.00154>
- Marshall, C. B., Mays, D. J., Beeler, J. S., Rosenbluth, J. M., Boyd, K. L., Santos Guasch, G. L., ... Pietenpol, J. A. (2016). p73 Is Required for Multiciliogenesis and Regulates the Foxj1-Associated Gene Network. *Cell Reports*, 14(10), 2289–2300. <https://doi.org/10.1016/j.celrep.2016.02.035>
- Martino, F., Pal, M., Muñoz-Hernández, H., Rodríguez, C. F., Núñez-Ramírez, R., Gil-Carton, D., ... Llorca, O. (2018). RPAP3 provides a flexible scaffold for coupling HSP90 to the human R2TP co-chaperone complex. *Nature Communications*, 9(1), 1501. <https://doi.org/10.1038/s41467-018-03942-1>
- Mashtalir, N., D'Avino, A. R., Michel, B. C., Luo, J., Pan, J., Otto, J. E., ... Kadoch, C. (2018). Modular Organization and Assembly of SWI/SNF Family Chromatin Remodeling Complexes. *Cell*, 175(5), 1272–1288.e20. <https://doi.org/10.1016/j.cell.2018.09.032>
- McGarry, T. J., & Kirschner, M. W. (1998). Geminin, an inhibitor of DNA replication, is degraded during mitosis. *Cell*, 93(6), 1043–1053. [https://doi.org/10.1016/s0092-8674\(00\)81209-x](https://doi.org/10.1016/s0092-8674(00)81209-x)
- McSwiggen, D. T., Mir, M., Darzacq, X., & Tjian, R. (2019). Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. *Genes & Development*, 33(23–24), 1619–1634. <https://doi.org/10.1101/gad.331520.119>
- Mellacheruvu, D., Wright, Z., Couzens, A. L., Lambert, J.-P., St-Denis, N. A., Li, T., ... Nesvizhskii, A. I. (2013). The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nature Methods*, 10(8), 730–736. <https://doi.org/10.1038/nmeth.2557>
- Mercey, O., Al Jord, A., Rostaing, P., Mahuzier, A., Fortoul, A., Boudjema, A.-R., ... Meunier, A. (2019). Dynamics of centriole amplification in centrosome-depleted brain multiciliated progenitors. *Scientific Reports*, 9(1), 13060. <https://doi.org/10.1038/s41598-019-49416-2>
- Mercey, O., Levine, M. S., LoMastro, G. M., Rostaing, P., Brotslaw, E., Gomez, V., ... Holland, A. J. (2019). Massive centriole production can occur in the absence of deuterosomes in multiciliated cells. *Nature Cell Biology*, 21(12), 1544–1552. <https://doi.org/10.1038/s41556-019-0427-x>
- Milin, A. N., & Deniz, A. A. (2018). Reentrant Phase Transitions and Non-Equilibrium Dynamics in Membraneless Organelles. *Biochemistry*, 57(17), 2470–2477. <https://doi.org/10.1021/acs.biochem.8b00001>
- Minezaki, Y., Homma, K., Kinjo, A. R., & Nishikawa, K. (2006). Human transcription factors contain a high fraction of intrinsically disordered regions essential for transcriptional regulation. *Journal of Molecular Biology*, 359(4), 1137–1149. <https://doi.org/10.1016/j.jmb.2006.04.016>
- Mishal, R., & Luna-Arias, J. P. (2022). Role of the TATA-box binding protein (TBP) and associated family members in transcription regulation. *Gene*, 833, 146581. <https://doi.org/10.1016/j.gene.2022.146581>

- Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., ... Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell*, *113*(5), 631–642. [https://doi.org/10.1016/s0092-8674\(03\)00393-3](https://doi.org/10.1016/s0092-8674(03)00393-3)
- Mittal, P., & Roberts, C. W. M. (2020). The SWI/SNF complex in cancer - biology, biomarkers and therapy. *Nature Reviews. Clinical Oncology*, *17*(7), 435–448. <https://doi.org/10.1038/s41571-020-0357-3>
- Miyamoto, H., Yeh, S., Lardy, H., Messing, E., & Chang, C. (1998). Delta5-androstenediol is a natural hormone with androgenic activity in human prostate cancer cells. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(19), 11083–11088. <https://doi.org/10.1073/pnas.95.19.11083>
- Moyer, T. C., Clutario, K. M., Lambrus, B. G., Daggubati, V., & Holland, A. J. (2015). Binding of STIL to Plk4 activates kinase activity to promote centriole assembly. *The Journal of Cell Biology*, *209*(6), 863–878. <https://doi.org/10.1083/jcb.201502088>
- Murakawa, Y., Yoshihara, M., Kawaji, H., Nishikawa, M., Zayed, H., Suzuki, H., ... Hayashizaki, Y. (2016). Enhanced Identification of Transcriptional Enhancers Provides Mechanistic Insights into Diseases. *Trends in Genetics : TIG*, *32*(2), 76–88. <https://doi.org/10.1016/j.tig.2015.11.004>
- Musacchio, A. (2022). On the role of phase separation in the biogenesis of membraneless compartments. *The EMBO Journal*, *41*(5), e109952. <https://doi.org/10.15252/embj.2021109952>
- Nadal, M., Prekovic, S., Gallastegui, N., Helsen, C., Abella, M., Zielinska, K., ... Estébanez-Perpiñá, E. (2017). Structure of the homodimeric androgen receptor ligand-binding domain. *Nature Communications*, *8*, 14388. <https://doi.org/10.1038/ncomms14388>
- Nag, N., Sasidharan, S., Uversky, V. N., Saudagar, P., & Tripathi, T. (2022). Phase separation of FG-nucleoporins in nuclear pore complexes. *Biochimica et Biophysica Acta. Molecular Cell Research*, *1869*(4), 119205. <https://doi.org/10.1016/j.bbamcr.2021.119205>
- Nagl, N. G. J., Wang, X., Patsialou, A., Van Scoy, M., & Moran, E. (2007). Distinct mammalian SWI/SNF chromatin remodeling complexes with opposing roles in cell-cycle control. *The EMBO Journal*, *26*(3), 752–763. <https://doi.org/10.1038/sj.emboj.7601541>
- Nagulapalli, M., Maji, S., Dwivedi, N., Dahiya, P., & Thakur, J. K. (2016). Evolution of disorder in Mediator complex and its functional relevance. *Nucleic Acids Research*, *44*(4), 1591–1612. <https://doi.org/10.1093/nar/gkv1135>
- Nair, S. J., Yang, L., Meluzzi, D., Oh, S., Yang, F., Friedman, M. J., ... Rosenfeld, M. G. (2019). Phase separation of ligand-activated enhancers licenses cooperative chromosomal enhancer assembly. *Nature Structural & Molecular Biology*, *26*(3), 193–203. <https://doi.org/10.1038/s41594-019-0190-5>
- Nanjundappa, R., Kong, D., Shim, K., Stearns, T., Brody, S. L., Loncarek, J., & Mahjoub, M. R. (2019). Regulation of cilia abundance in multiciliated cells. *ELife*, *8*. <https://doi.org/10.7554/eLife.44039>
- Nemajerova, A., Kramer, D., Siller, S. S., Herr, C., Shomroni, O., Pena, T., ... Lizé, M. (2016). TAp73 is a central transcriptional regulator of airway multiciliogenesis. *Genes & Development*, *30*(11), 1300–1312. <https://doi.org/10.1101/gad.279836.116>

- Newton, F. G., zur Lage, P. I., Karak, S., Moore, D. J., Göpfert, M. C., & Jarman, A. P. (2012). Forkhead transcription factor Fd3F cooperates with Rfx to regulate a gene expression program for mechanosensory cilia specialization. *Developmental Cell*, *22*(6), 1221–1233. <https://doi.org/10.1016/j.devcel.2012.05.010>
- Nigg, E. A., & Holland, A. J. (2018). Once and only once: mechanisms of centriole duplication and their deregulation in disease. *Nature Reviews Molecular Cell Biology*, *19*(5), 297–312. <https://doi.org/10.1038/nrm.2017.127>
- Núñez-Ollé, M., Jung, C., Terré, B., Balsiger, N. A., Plata, C., Roset, R., ... Gil-Gómez, G. (2017). Constitutive Cyclin O deficiency results in penetrant hydrocephalus, impaired growth and infertility. *Oncotarget*, *8*(59), 99261–99273. <https://doi.org/10.18632/oncotarget.21818>
- Núñez-Ollé, M., Stracker, T. H., & Gil-Gómez, G. (2018, February). CCNO mutations in NPH? *Aging*, Vol. 10, pp. 158–159. <https://doi.org/10.18632/aging.101379>
- Obst, J. K., Wang, J., Jian, K., Williams, D. E., Tien, A. H., Mawji, N., ... Sadar, M. D. (2019). Revealing Metabolic Liabilities of Ralaniten To Enhance Novel Androgen Receptor Targeted Therapies. *ACS Pharmacology & Translational Science*, *2*(6), 453–467. <https://doi.org/10.1021/acspsci.9b00065>
- Ohta, M., Ashikawa, T., Nozaki, Y., Kozuka-Hata, H., Goto, H., Inagaki, M., ... Kitagawa, D. (2014). Direct interaction of Plk4 with STIL ensures formation of a single procentriole per parental centriole. *Nature Communications*, *5*, 5267. <https://doi.org/10.1038/ncomms6267>
- Okada, A., Ohta, Y., Brody, S. L., Watanabe, H., Krust, A., Chambon, P., & Iguchi, T. (2004). Role of foxj1 and estrogen receptor alpha in ciliated epithelial cell differentiation of the neonatal oviduct. *Journal of Molecular Endocrinology*, *32*(3), 615–625. <https://doi.org/10.1677/jme.0.0320615>
- Pagliuca, F. W., Collins, M. O., Lichawska, A., Zegerman, P., Choudhary, J. S., & Pines, J. (2011). Quantitative proteomics reveals the basis for the biochemical specificity of the cell-cycle machinery. *Molecular Cell*, *43*(3), 406–417. <https://doi.org/10.1016/j.molcel.2011.05.031>
- Panca, R., & Tompa, P. (2012). Structural disorder in eukaryotes. *PLoS One*, *7*(4), e34687. <https://doi.org/10.1371/journal.pone.0034687>
- Papaspyropoulos, A., Bradley, L., Thapa, A., Leung, C. Y., Toskas, K., Koennig, D., ... O'Neill, E. (2018). RASSF1A uncouples Wnt from Hippo signalling and promotes YAP mediated differentiation via p73. *Nature Communications*, *9*(1), 424. <https://doi.org/10.1038/s41467-017-02786-5>
- Pardo-Saganta, A., Tata, P. R., Law, B. M., Saez, B., Chow, R. D.-W., Prabhu, M., ... Rajagopal, J. (2015). Parent stem cells can serve as niches for their daughter cells. *Nature*, *523*(7562), 597–601. <https://doi.org/10.1038/nature14553>
- Park, J. J., Irvine, R. A., Buchanan, G., Koh, S. S., Park, J. M., Tilley, W. D., ... Coetzee, G. A. (2000). Breast cancer susceptibility gene 1 (BRCA1) is a coactivator of the androgen receptor. *Cancer Research*, *60*(21), 5946–5949.
- Parker, C., Nilsson, S., Heinrich, D., Helle, S. I., O'Sullivan, J. M., Fosså, S. D., ... Sartor, O. (2013). Alpha emitter radium-223 and survival in metastatic prostate cancer. *The New England Journal of Medicine*, *369*(3), 213–223. <https://doi.org/10.1056/NEJMoa1213755>

- Pefani, D.-E., Dimaki, M., Spella, M., Karantzelis, N., Mitsiki, E., Kyrrousi, C., ... Lygerou, Z. (2011). Idas, a novel phylogenetically conserved geminin-related protein, binds to geminin and is required for cell cycle progression. *The Journal of Biological Chemistry*, *286*(26), 23234–23246. <https://doi.org/10.1074/jbc.M110.207688>
- Poss, Z. C., Ebmeier, C. C., & Taatjes, D. J. (2013). The Mediator complex and transcription regulation. *Critical Reviews in Biochemistry and Molecular Biology*, *48*(6), 575–608. <https://doi.org/10.3109/10409238.2013.840259>
- Protter, D. S. W., Rao, B. S., Van Treeck, B., Lin, Y., Mizoue, L., Rosen, M. K., & Parker, R. (2018). Intrinsically Disordered Regions Can Contribute Promiscuous Interactions to RNP Granule Assembly. *Cell Reports*, *22*(6), 1401–1412. <https://doi.org/10.1016/j.celrep.2018.01.036>
- Ptashne, M., & Gann, A. (1997). Transcriptional activation by recruitment. *Nature*, *386*(6625), 569–577. <https://doi.org/10.1038/386569a0>
- Qu, Y., Dai, B., Ye, D., Kong, Y., Chang, K., Jia, Z., ... Shi, G. (2015). Constitutively active AR-V7 plays an essential role in the development and progression of castration-resistant prostate cancer. *Scientific Reports*, *5*, 7654. <https://doi.org/10.1038/srep07654>
- Quigley, I. K., & Kintner, C. (2017). Rfx2 Stabilizes Foxj1 Binding at Chromatin Loops to Enable Multiciliated Cell Gene Expression. *PLoS Genetics*, *13*(1), e1006538. <https://doi.org/10.1371/journal.pgen.1006538>
- Radhakrishnan, I., Pérez-Alvarado, G. C., Parker, D., Dyson, H. J., Montminy, M. R., & Wright, P. E. (1997). Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. *Cell*, *91*(6), 741–752. [https://doi.org/10.1016/s0092-8674\(00\)80463-8](https://doi.org/10.1016/s0092-8674(00)80463-8)
- Ramamoorthy, S., & Cidlowski, J. A. (2016). Corticosteroids: Mechanisms of Action in Health and Disease. *Rheumatic Diseases Clinics of North America*, *42*(1), 15–31, vii. <https://doi.org/10.1016/j.rdc.2015.08.002>
- Ramani, P., Nash, R., Sowa-Avugrah, E., & Rogers, C. (2015). High levels of polo-like kinase 1 and phosphorylated translationally controlled tumor protein indicate poor prognosis in neuroblastomas. *Journal of Neuro-Oncology*, *125*(1), 103–111. <https://doi.org/10.1007/s11060-015-1900-4>
- Rana, K., Davey, R. A., & Zajac, J. D. (2014). Human androgen deficiency: insights gained from androgen receptor knockout mouse models. *Asian Journal of Andrology*, *16*(2), 169–177. <https://doi.org/10.4103/1008-682X.122590>
- Rasi, M. Q., Parker, J. D. K., Feldman, J. L., Marshall, W. F., & Quarumby, L. M. (2009). Katanin knockdown supports a role for microtubule severing in release of basal bodies before mitosis in *Chlamydomonas*. *Molecular Biology of the Cell*, *20*(1), 379–388. <https://doi.org/10.1091/mbc.e07-10-1007>
- Rasool, R. U., Natesan, R., Deng, Q., Aras, S., Lal, P., Sander Effron, S., ... Asangani, I. A. (2019). CDK7 Inhibition Suppresses Castration-Resistant Prostate Cancer through MED1 Inactivation. *Cancer Discovery*, *9*(11), 1538–1555. <https://doi.org/10.1158/2159-8290.CD-19-0189>

- Rayburn, E., Zhang, R., He, J., & Wang, H. (2005). MDM2 and human malignancies: expression, clinical pathology, prognostic markers, and implications for chemotherapy. *Current Cancer Drug Targets*, 5(1), 27–41. <https://doi.org/10.2174/1568009053332636>
- Rebeiz, M., & Tsiantis, M. (2017). Enhancer evolution and the origins of morphological novelty. *Current Opinion in Genetics & Development*, 45, 115–123. <https://doi.org/10.1016/j.gde.2017.04.006>
- Reid, J., Kelly, S. M., Watt, K., Price, N. C., & McEwan, I. J. (2002). Conformational analysis of the androgen receptor amino-terminal domain involved in transactivation. Influence of structure-stabilizing solutes and protein-protein interactions. *The Journal of Biological Chemistry*, 277(22), 20079–20086. <https://doi.org/10.1074/jbc.M201003200>
- Ren, Y., Behre, E., Ren, Z., Zhang, J., Wang, Q., & Fondell, J. D. (2000). Specific structural motifs determine TRAP220 interactions with nuclear hormone receptors. *Molecular and Cellular Biology*, 20(15), 5433–5446. <https://doi.org/10.1128/MCB.20.15.5433-5446.2000>
- Revinski, D. R., Zaragosi, L.-E. E., Boutin, C., Ruiz-Garcia, S., Deprez, M., Thomé, V., ... Barbry, P. (2018). CDC20B is required for deuterosome-mediated centriole production in multiciliated cells. *Nature Communications*, 9(1), 4668. <https://doi.org/10.1038/s41467-018-06768-z>
- Robinson, D., Van Allen, E. M., Wu, Y.-M., Schultz, N., Lonigro, R. J., Mosquera, J.-M., ... Chinnaiyan, A. M. (2015, July). Integrative Clinical Genomics of Advanced Prostate Cancer. *Cell*, Vol. 162, p. 454. <https://doi.org/10.1016/j.cell.2015.06.053>
- Robinson, P. J., Trnka, M. J., Bushnell, D. A., Davis, R. E., Mattei, P.-J., Burlingame, A. L., & Kornberg, R. D. (2016). Structure of a Complete Mediator-RNA Polymerase II Pre-Initiation Complex. *Cell*, 166(6), 1411-1422.e16. <https://doi.org/10.1016/j.cell.2016.08.050>
- Roeder, R. G. (1996). The role of general initiation factors in transcription by RNA polymerase II. *Trends in Biochemical Sciences*, 21(9), 327–335.
- Roelfsema, M. R. G., & Hedrich, R. (2005). In the light of stomatal opening: new insights into “the Watergate”. *The New Phytologist*, 167(3), 665–691. <https://doi.org/10.1111/j.1469-8137.2005.01460.x>
- Roig, M. B., Roset, R., Ortet, L., Balsiger, N. A., Anfosso, A., Cabellos, L., ... Gil-Gómez, G. (2009). Identification of a novel cyclin required for the intrinsic apoptosis pathway in lymphoid cells. *Cell Death and Differentiation*, 16(2), 230–243. <https://doi.org/10.1038/cdd.2008.145>
- Roset, R., & Gil-Gómez, G. (2009). Measurement of changes in Cdk2 and cyclin o-associated kinase activity in apoptosis. *Methods in Molecular Biology (Clifton, N.J.)*, 559, 161–172. https://doi.org/10.1007/978-1-60327-017-5_12
- Sabari, B. R., Dall’Agnese, A., Boija, A., Klein, I. A., Coffey, E. L., Shrinivas, K., ... Young, R. A. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. *Science (New York, N.Y.)*, 361(6400). <https://doi.org/10.1126/science.aar3958>
- Sabari, B. R., Dall’Agnese, A., & Young, R. A. (2020). Biomolecular Condensates in the Nucleus. *Trends in Biochemical Sciences*, 45(11), 961–977. <https://doi.org/10.1016/j.tibs.2020.06.007>

- Sadar, M. D. (2020). Discovery of drugs that directly target the intrinsically disordered region of the androgen receptor. *Expert Opinion on Drug Discovery*, 15(5), 551–560. <https://doi.org/10.1080/17460441.2020.1732920>
- Sadeesh, N., Scaravilli, M., & Latonen, L. (2021). Proteomic Landscape of Prostate Cancer: The View Provided by Quantitative Proteomics, Integrative Analyses, and Protein Interactomes. *Cancers*, 13(19). <https://doi.org/10.3390/cancers13194829>
- Sankar, S., Yellajoshiyula, D., Zhang, B., Teets, B., Rockweiler, N., & Kroll, K. L. (2016). Gene regulatory networks in neural cell fate acquisition from genome-wide chromatin association of Geminin and Zic1. *Scientific Reports*, 6(July), 1–16. <https://doi.org/10.1038/srep37412>
- Santen, G. W. E., Kriek, M., & van Attikum, H. (2012). SWI/SNF complex in disorder. *Epigenetics*, 7(11), 1219–1224. <https://doi.org/10.4161/epi.22299>
- Savas, S., & Skardasi, G. (2018). The SWI/SNF complex subunit genes: Their functions, variations, and links to risk and survival outcomes in human cancers. *Critical Reviews in Oncology/Hematology*, 123, 114–131. <https://doi.org/10.1016/j.critrevonc.2018.01.009>
- Senapati, D., Kumari, S., & Heemers, H. V. (2020). Androgen receptor co-regulation in prostate cancer. *Asian Journal of Urology*, 7(3), 219–232. <https://doi.org/10.1016/j.ajur.2019.09.005>
- Seo, S., Herr, A., Lim, J.-W., Richardson, G. A., Richardson, H., & Kroll, K. L. (2005). Geminin regulates neuronal differentiation by antagonizing Brg1 activity. *Genes & Development*, 19(14), 1723–1734. <https://doi.org/10.1101/gad.1319105>
- Sharifi, N., Gulley, J. L., & Dahut, W. L. (2005). Androgen deprivation therapy for prostate cancer. *JAMA*, 294(2), 238–244. <https://doi.org/10.1001/jama.294.2.238>
- Sharp, A., Coleman, I., Yuan, W., Sprenger, C., Dolling, D., Rodrigues, D. N., ... Plymate, S. R. (2019). Androgen receptor splice variant-7 expression emerges with castration resistance in prostate cancer. *The Journal of Clinical Investigation*, 129(1), 192–208. <https://doi.org/10.1172/JCI122819>
- Sharp, P. A., Chakraborty, A. K., Henninger, J. E., & Young, R. A. (2022). RNA in formation and regulation of transcriptional condensates. *RNA (New York, N.Y.)*, 28(1), 52–57. <https://doi.org/10.1261/rna.078997.121>
- Shi, Y., Chen, J., Zeng, W.-J., Li, M., Zhao, W., Zhang, X.-D., & Yao, J. (2021). Formation of nuclear condensates by the Mediator complex subunit Med15 in mammalian cells. *BMC Biology*, 19(1), 245. <https://doi.org/10.1186/s12915-021-01178-y>
- Shiratori, H., & Hamada, H. (2006). The left-right axis in the mouse: from origin to morphology. *Development (Cambridge, England)*, 133(11), 2095–2104. <https://doi.org/10.1242/dev.02384>
- Shrinivas, K., Sabari, B. R., Coffey, E. L., Klein, I. A., Boija, A., Zamudio, A. V., ... Chakraborty, A. K. (2019). Enhancer Features that Drive Formation of Transcriptional Condensates. *Molecular Cell*, 75(3), 549–561.e7. <https://doi.org/10.1016/j.molcel.2019.07.009>
- Siegel, R. L., Miller, K. D., & Jemal, A. (2018). Cancer statistics, 2018. *CA: A Cancer Journal for Clinicians*, 68(1), 7–30. <https://doi.org/10.3322/caac.21442>

- Simental, J. A., Sar, M., Lane, M. V, French, F. S., & Wilson, E. M. (1991). Transcriptional activation and nuclear targeting signals of the human androgen receptor. *The Journal of Biological Chemistry*, 266(1), 510–518.
- Smith, M. R., Saad, F., Chowdhury, S., Oudard, S., Hadaschik, B. A., Graff, J. N., ... Small, E. J. (2018). Apalutamide Treatment and Metastasis-free Survival in Prostate Cancer. *The New England Journal of Medicine*, 378(15), 1408–1418. <https://doi.org/10.1056/NEJMoa1715546>
- Song, R., Walentek, P., Sponer, N., Klimke, A., Lee, J. S., Dixon, G., ... He, L. (2014). miR-34/449 miRNAs are required for motile ciliogenesis by repressing cp110. *Nature*, 510(7503), 115–120. <https://doi.org/10.1038/nature13413>
- Soutourina, J. (2018). Transcription regulation by the Mediator complex. *Nature Reviews. Molecular Cell Biology*, 19(4), 262–274. <https://doi.org/10.1038/nrm.2017.115>
- Spannl, S., Tereshchenko, M., Mastromarco, G. J., Ihn, S. J., & Lee, H. O. (2019). Biomolecular condensates in neurodegeneration and cancer. *Traffic (Copenhagen, Denmark)*, 20(12), 890–911. <https://doi.org/10.1111/tra.12704>
- Spassky, N., & Meunier, A. (2017). The development and functions of multiciliated epithelia. *Nature Reviews. Molecular Cell Biology*, 18(7), 423–436. <https://doi.org/10.1038/nrm.2017.21>
- Stubbs, J. L., Oishi, I., Izpisua Belmonte, J. C., & Kintner, C. (2008). The forkhead protein Foxj1 specifies node-like cilia in *Xenopus* and zebrafish embryos. *Nature Genetics*, 40(12), 1454–1460. <https://doi.org/10.1038/ng.267>
- Sun, S., Sprenger, C. C. T., Vessella, R. L., Haugk, K., Soriano, K., Mostaghel, E. A., ... Plymate, S. R. (2010). Castration resistance in human prostate cancer is conferred by a frequently occurring androgen receptor splice variant. *The Journal of Clinical Investigation*, 120(8), 2715–2730. <https://doi.org/10.1172/JCI41824>
- Suzuki, H. I., Young, R. A., & Sharp, P. A. (2017). Super-Enhancer-Mediated RNA Processing Revealed by Integrative MicroRNA Network Analysis. *Cell*, 168(6), 1000-1014.e15. <https://doi.org/10.1016/j.cell.2017.02.015>
- Takayama, K., & Inoue, S. (2013). Transcriptional network of androgen receptor in prostate cancer progression. *International Journal of Urology: Official Journal of the Japanese Urological Association*, 20(8), 756–768. <https://doi.org/10.1111/iju.12146>
- Tamm, S., & Tamm, S. L. (1988). Development of macrociliary cells in *Beroë*. I. Actin bundles and centriole migration. *Journal of Cell Science*, 89 (Pt 1), 67–80. <https://doi.org/10.1242/jcs.89.1.67>
- Tan, J., Sharief, Y., Hamil, K. G., Gregory, C. W., Zang, D. Y., Sar, M., ... French, F. S. (1997). Dehydroepiandrosterone activates mutant androgen receptors expressed in the androgen-dependent human prostate cancer xenograft CWR22 and LNCaP cells. *Molecular Endocrinology (Baltimore, Md.)*, 11(4), 450–459. <https://doi.org/10.1210/mend.11.4.9906>
- Taniguchi, K., Momiyama, N., Ueda, M., Matsuyama, R., Mori, R., Fujii, Y., ... Shimada, H. (2008). Targeting of CDC20 via small interfering RNA causes enhancement of the cytotoxicity of chemoradiation. *Anticancer Research*, 28(3A), 1559–1563.

- Taylor, B. S., Schultz, N., Hieronymus, H., Gopalan, A., Xiao, Y., Carver, B. S., ... Gerald, W. L. (2010). Integrative genomic profiling of human prostate cancer. *Cancer Cell*, 18(1), 11–22. <https://doi.org/10.1016/j.ccr.2010.05.026>
- Thandapani, P. (2019). Super-enhancers in cancer. *Pharmacology & Therapeutics*, 199, 129–138. <https://doi.org/10.1016/j.pharmthera.2019.02.014>
- Theodoulou, N. H., Bamborough, P., Bannister, A. J., Becher, I., Bit, R. A., Che, K. H., ... Humphreys, P. G. (2016). Discovery of I-BRD9, a Selective Cell Active Chemical Probe for Bromodomain Containing Protein 9 Inhibition. *Journal of Medicinal Chemistry*, 59(4), 1425–1439. <https://doi.org/10.1021/acs.jmedchem.5b00256>
- Tomlins, S. A., Bjartell, A., Chinnaiyan, A. M., Jenster, G., Nam, R. K., Rubin, M. A., & Schalken, J. A. (2009). ETS gene fusions in prostate cancer: from discovery to daily clinical practice. *European Urology*, 56(2), 275–286. <https://doi.org/10.1016/j.eururo.2009.04.036>
- Tompa, P. (2002). Intrinsically unstructured proteins. *Trends in Biochemical Sciences*, 27(10), 527–533. [https://doi.org/10.1016/s0968-0004\(02\)02169-2](https://doi.org/10.1016/s0968-0004(02)02169-2)
- Tsao, P.-N., Vasconcelos, M., Izvolsky, K. I., Qian, J., Lu, J., & Cardoso, W. V. (2009). Notch signaling controls the balance of ciliated and secretory cell fates in developing airways. *Development (Cambridge, England)*, 136(13), 2297–2307. <https://doi.org/10.1242/dev.034884>
- Tsurusaki, Y., Okamoto, N., Ohashi, H., Kosho, T., Imai, Y., Hibi-Ko, Y., ... Matsumoto, N. (2012). Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nature Genetics*, 44(4), 376–378. <https://doi.org/10.1038/ng.2219>
- Tu, F., Sedzinski, J., Ma, Y., Marcotte, E. M., & Wallingford, J. B. (2018). Protein localization screening in vivo reveals novel regulators of multiciliated cell development and function. *Journal of Cell Science*, 131(3). <https://doi.org/10.1242/jcs.206565>
- Tut, T. G., Lim, S. H. S., Dissanayake, I. U., Descallar, J., Chua, W., Ng, W., ... Lee, C. S. (2015). Upregulated Polo-Like Kinase 1 Expression Correlates with Inferior Survival Outcomes in Rectal Cancer. *PLoS One*, 10(6), e0129313. <https://doi.org/10.1371/journal.pone.0129313>
- Uversky, V. N., & Dunker, A. K. (2010). Understanding protein non-folding. *Biochimica et Biophysica Acta*, 1804(6), 1231–1264. <https://doi.org/10.1016/j.bbapap.2010.01.017>
- van Soldt, B. J., Qian, J., Li, J., Tang, N., Lu, J., & Cardoso, W. V. (2019). Yap and its subcellular localization have distinct compartment-specific roles in the developing lung. *Development (Cambridge, England)*, 146(9). <https://doi.org/10.1242/dev.175810>
- Vandevyver, S., Dejager, L., & Libert, C. (2014). Comprehensive overview of the structure and regulation of the glucocorticoid receptor. *Endocrine Reviews*, 35(4), 671–693. <https://doi.org/10.1210/er.2014-1010>
- Vaquerezas, J. M., Kummerfeld, S. K., Teichmann, S. A., & Luscombe, N. M. (2009, April). A census of human transcription factors: function, expression and evolution. *Nature Reviews. Genetics*, Vol. 10, pp. 252–263. <https://doi.org/10.1038/nrg2538>

- Veldscholte, J., Ris-Stalpers, C., Kuiper, G. G., Jenster, G., Berrevoets, C., Claassen, E., ... Mulder, E. (1990). A mutation in the ligand binding domain of the androgen receptor of human LNCaP cells affects steroid binding characteristics and response to anti-androgens. *Biochemical and Biophysical Research Communications*, 173(2), 534–540. [https://doi.org/10.1016/s0006-291x\(05\)80067-1](https://doi.org/10.1016/s0006-291x(05)80067-1)
- Villa, M., Crotta, S., Dingwell, K. S., Hirst, E. M. A., Gialitakis, M., Ahlfors, H., ... Wack, A. (2016). The aryl hydrocarbon receptor controls cyclin O to promote epithelial multiciliogenesis. *Nature Communications*, 7, 12652. <https://doi.org/10.1038/ncomms12652>
- Visel, A., Rubin, E. M., & Pennacchio, L. A. (2009). Genomic views of distant-acting enhancers. *Nature*, 461(7261), 199–205. <https://doi.org/10.1038/nature08451>
- Waddell, A. R., Huang, H., & Liao, D. (2021). CBP/p300: Critical Co-Activators for Nuclear Steroid Hormone Receptors and Emerging Therapeutic Targets in Prostate and Breast Cancers. *Cancers*, 13(12). <https://doi.org/10.3390/cancers13122872>
- Wallmeier, J., Al-Mutairi, D. A., Chen, C.-T. T., Loges, N. T., Pennekamp, P., Menchen, T., ... Omran, H. (2014). Mutations in CCNO result in congenital mucociliary clearance disorder with reduced generation of multiple motile cilia. *Nature Genetics*, 46(6), 646–651. <https://doi.org/10.1038/ng.2961>
- Wallmeier, J., Frank, D., Shoemark, A., Nöthe-Menchen, T., Cindric, S., Olbrich, H., ... Omran, H. (2019). De Novo Mutations in FOXJ1 Result in a Motile Ciliopathy with Hydrocephalus and Randomization of Left/Right Body Asymmetry. *American Journal of Human Genetics*, 105(5), 1030–1039. <https://doi.org/10.1016/j.ajhg.2019.09.022>
- Waltering, K. K., Helenius, M. A., Sahu, B., Manni, V., Linja, M. J., Jänne, O. A., & Visakorpi, T. (2009). Increased expression of androgen receptor sensitizes prostate cancer cells to low levels of androgens. *Cancer Research*, 69(20), 8141–8149. <https://doi.org/10.1158/0008-5472.CAN-09-0919>
- Wang, J., Choi, J.-M., Holehouse, A. S., Lee, H. O., Zhang, X., Jahnel, M., ... Hyman, A. A. (2018). A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell*, 174(3), 688-699.e16. <https://doi.org/10.1016/j.cell.2018.06.006>
- Wang, L., Hou, Z., Hasim, A., Abuduerheman, A., Zhang, H., Niyaz, M., ... Sheyhidin, I. (2018). RNF113A promotes the proliferation, migration and invasion, and is associated with a poor prognosis of esophageal squamous cell carcinoma. *International Journal of Oncology*, 52(3), 861–871. <https://doi.org/10.3892/ijo.2018.4253>
- Wang, L., Zhang, J., Wan, L., Zhou, X., Wang, Z., & Wei, W. (2015). Targeting Cdc20 as a novel cancer therapeutic strategy. *Pharmacology & Therapeutics*, 151, 141–151. <https://doi.org/10.1016/j.pharmthera.2015.04.002>
- Wang, Q., Carroll, J. S., & Brown, M. (2005). Spatial and temporal recruitment of androgen receptor and its coactivators involves chromosomal looping and polymerase tracking. *Molecular Cell*, 19(5), 631–642. <https://doi.org/10.1016/j.molcel.2005.07.018>
- Wang, Q., Sharma, D., Ren, Y., & Fondell, J. D. (2002). A coregulatory role for the TRAP-mediator complex in androgen receptor-mediated gene expression. *The Journal of Biological Chemistry*, 277(45), 42852–42858. <https://doi.org/10.1074/jbc.M206061200>

- Wang, W., Chen, Y., Xu, A., Cai, M., Cao, J., Zhu, H., ... He, Q. (2020). Protein phase separation: A novel therapy for cancer? *British Journal of Pharmacology*, *177*(22), 5008–5030. <https://doi.org/10.1111/bph.15242>
- Wang, Y., Jiang, X., Feng, F., Liu, W., & Sun, H. (2020). Degradation of proteins by PROTACs and other strategies. *Acta Pharmaceutica Sinica. B*, *10*(2), 207–238. <https://doi.org/10.1016/j.apsb.2019.08.001>
- Watson, P. A., Arora, V. K., & Sawyers, C. L. (2015). Emerging mechanisms of resistance to androgen receptor inhibitors in prostate cancer. *Nature Reviews. Cancer*, *15*(12), 701–711. <https://doi.org/10.1038/nrc4016>
- Watson, P. A., Chen, Y. F., Balbas, M. D., Wongvipat, J., Socci, N. D., Viale, A., ... Sawyers, C. L. (2010). Constitutively active androgen receptor splice variants expressed in castration-resistant prostate cancer require full-length androgen receptor. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(39), 16759–16765. <https://doi.org/10.1073/pnas.1012443107>
- Weber, H., & Garabedian, M. J. (2018). The mediator complex in genomic and non-genomic signaling in cancer. *Steroids*, *133*, 8–14. <https://doi.org/10.1016/j.steroids.2017.11.007>
- Weikum, E. R., Knuesel, M. T., Ortlund, E. A., & Yamamoto, K. R. (2017). Glucocorticoid receptor control of transcription: precision and plasticity via allostery. *Nature Reviews. Molecular Cell Biology*, *18*(3), 159–174. <https://doi.org/10.1038/nrm.2016.152>
- Werner, M. E., Hwang, P., Huisman, F., Taborek, P., Yu, C. C., & Mitchell, B. J. (2011). Actin and microtubules drive differential aspects of planar cell polarity in multiciliated cells. *The Journal of Cell Biology*, *195*(1), 19–26. <https://doi.org/10.1083/jcb.201106110>
- Wheway, G., & Mitchison, H. M. (2019). Opportunities and Challenges for Molecular Understanding of Ciliopathies-The 100,000 Genomes Project. *Frontiers in Genetics*, *10*, 127. <https://doi.org/10.3389/fgene.2019.00127>
- Whyte, W. A., Orlando, D. A., Hnisz, D., Abraham, B. J., Lin, C. Y., Kagey, M. H., ... Young, R. A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*, *153*(2), 307–319. <https://doi.org/10.1016/j.cell.2013.03.035>
- Wildung, M., Esser, T. U., Grausam, K. B., Wiedwald, C., Volceanov-Hahn, L., Riedel, D., ... Lizé, M. (2019). Transcription factor TAp73 and microRNA-449 complement each other to support multiciliogenesis. *Cell Death and Differentiation*, *26*(12), 2740–2757. <https://doi.org/10.1038/s41418-019-0332-7>
- Wohlschlegel, J. A., Dwyer, B. T., Dhar, S. K., Cvetic, C., Walter, J. C., & Dutta, A. (2000). Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. *Science (New York, N.Y.)*, *290*(5500), 2309–2312. <https://doi.org/10.1126/science.290.5500.2309>
- Wright, P. E., & Dyson, H. J. (2009). Linking folding and binding. *Current Opinion in Structural Biology*, *19*(1), 31–38. <https://doi.org/10.1016/j.sbi.2008.12.003>
- Wu, J. N., Fish, K. M., Evans, C. P., Devere White, R. W., & Dall'Era, M. A. (2014). No improvement noted in overall or cause-specific survival for men presenting with metastatic prostate cancer over a 20-year period. *Cancer*, *120*(6), 818–823. <https://doi.org/10.1002/cncr.28485>

- Xiao, L., Parolia, A., Qiao, Y., Bawa, P., Eyunni, S., Mannan, R., ... Chinnaiyan, A. M. (2022). Targeting SWI/SNF ATPases in enhancer-addicted prostate cancer. *Nature*, *601*(7893), 434–439. <https://doi.org/10.1038/s41586-021-04246-z>
- Xouri, G., Squire, A., Dimaki, M., Geverts, B., Verveer, P. J., Taraviras, S., ... Lygerou, Z. (2007). Cdt1 associates dynamically with chromatin throughout G1 and recruits Geminin onto chromatin. *The EMBO Journal*, *26*(5), 1303–1314. <https://doi.org/10.1038/sj.emboj.7601597>
- Yan, X., Zhao, H., & Zhu, X. (2016). Production of Basal Bodies in bulk for dense multicilia formation. *F1000Research*, *5*. <https://doi.org/10.12688/f1000research.8469.1>
- You, Y., Huang, T., Richer, E. J., Schmidt, J.-E. H., Zabner, J., Borok, Z., & Brody, S. L. (2004). Role of f-box factor foxj1 in differentiation of ciliated airway epithelial cells. *American Journal of Physiology. Lung Cellular and Molecular Physiology*, *286*(4), L650-7. <https://doi.org/10.1152/ajplung.00170.2003>
- Yu, X., Yi, P., Hamilton, R. A., Shen, H., Chen, M., Foulds, C. E., ... O'Malley, B. W. (2020). Structural Insights of Transcriptionally Active, Full-Length Androgen Receptor Coactivator Complexes. *Molecular Cell*, *79*(5), 812-823.e4. <https://doi.org/10.1016/j.molcel.2020.06.031>
- Yuan, C. X., Ito, M., Fondell, J. D., Fu, Z. Y., & Roeder, R. G. (1998). The TRAP220 component of a thyroid hormone receptor-associated protein (TRAP) coactivator complex interacts directly with nuclear receptors in a ligand-dependent fashion. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(14), 7939–7944. <https://doi.org/10.1073/pnas.95.14.7939>
- Yuan, S., Liu, Y., Peng, H., Tang, C., Hennig, G. W., Wang, Z., ... Yan, W. (2019). Motile cilia of the male reproductive system require miR-34/miR-449 for development and function to generate luminal turbulence. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(9), 3584–3593. <https://doi.org/10.1073/pnas.1817018116>
- Zabidi, M. A., & Stark, A. (2016). Regulatory Enhancer-Core-Promoter Communication via Transcription Factors and Cofactors. *Trends in Genetics: TIG*, *32*(12), 801–814. <https://doi.org/10.1016/j.tig.2016.10.003>
- Zhang, F., Wong, S., Lee, J., Lingadahalli, S., Wells, C., Saxena, N., ... Lallous, N. (2021). Dynamic phase separation of the androgen receptor and its coactivators to regulate gene expression. *BioRxiv*, 2021.03.27.437301. <https://doi.org/10.1101/2021.03.27.437301>
- Zhang, H., Elbaum-Garfinkle, S., Langdon, E. M., Taylor, N., Occhipinti, P., Bridges, A. A., ... Gladfelter, A. S. (2015). RNA Controls PolyQ Protein Phase Transitions. *Molecular Cell*, *60*(2), 220–230. <https://doi.org/10.1016/j.molcel.2015.09.017>
- Zhang, S., & Mitchell, B. J. (2015). Centriole biogenesis and function in multiciliated cells. *Methods in Cell Biology*, *129*, 103–127. <https://doi.org/10.1016/bs.mcb.2015.03.015>
- Zhang, X., Jia, S., Chen, Z., Chong, Y. L., Xie, H., Feng, D., ... Zhao, C. (2018, December). Cilia-driven cerebrospinal fluid flow directs expression of urotensin neuropeptides to straighten the vertebrate body axis. *Nature Genetics*, Vol. 50, pp. 1666–1673. <https://doi.org/10.1038/s41588-018-0260-3>
- Zhao, H., Chen, Q., Fang, C., Huang, Q., Zhou, J., Yan, X., & Zhu, X. (2019). Parental centrioles are dispensable for deuterosome formation and function during basal body amplification. *EMBO Reports*, *20*(4). <https://doi.org/10.15252/embr.201846735>

- Zhao, H., Zhu, L., Zhu, Y., Cao, J., Li, S., Huang, Q., ... Zhu, X. (2013). The Cep63 paralogue Deup1 enables massive de novo centriole biogenesis for vertebrate multiciliogenesis. *Nature Cell Biology*, 15(12), 1434–1444. <https://doi.org/10.1038/ncb2880>
- Zhao, Y. G., & Zhang, H. (2020). Phase Separation in Membrane Biology: The Interplay between Membrane-Bound Organelles and Membraneless Condensates. *Developmental Cell*, 55(1), 30–44. <https://doi.org/10.1016/j.devcel.2020.06.033>
- Zheng, Q., Fu, Q., Xu, J., Gu, X., Zhou, H., & Zhi, C. (2021). Transcription factor E2F4 is an indicator of poor prognosis and is related to immune infiltration in hepatocellular carcinoma. *Journal of Cancer*, 12(6), 1792–1803. <https://doi.org/10.7150/jca.51616>
- Zhong, L., Li, Y., Xiong, L., Wang, W., Wu, M., Yuan, T., ... Yang, S. (2021). Small molecules in targeted cancer therapy: advances, challenges, and future perspectives. *Signal Transduction and Targeted Therapy*, 6(1), 201. <https://doi.org/10.1038/s41392-021-00572-w>
- Zhong, S., Jeong, J.-H., Chen, Z., Chen, Z., & Luo, J.-L. (2020). Targeting Tumor Microenvironment by Small-Molecule Inhibitors. *Translational Oncology*, 13(1), 57–69. <https://doi.org/10.1016/j.tranon.2019.10.001>
- Zhou, F., Narasimhan, V., Shboul, M., Chong, Y. L., Reversade, B., & Roy, S. (2015). Gmnc Is a Master Regulator of the Multiciliated Cell Differentiation Program. *Current Biology: CB*, 25(24), 3267–3273. <https://doi.org/10.1016/j.cub.2015.10.062>
- Zhou, Q., Nie, R., Prins, G. S., Saunders, P. T. K., Katzenellenbogen, B. S., & Hess, R. A. (2002). Localization of androgen and estrogen receptors in adult male mouse reproductive tract. *Journal of Andrology*, 23(6), 870–881.
- Zuber, J., Shi, J., Wang, E., Rappaport, A. R., Herrmann, H., Sison, E. A., ... Vakoc, C. R. (2011). RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature*, 478(7370), 524–528. <https://doi.org/10.1038/nature10334>