ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

BIOLOGIA CELLULARE E MOLECOLARE

Ciclo 37

Settore Concorsuale: 05/I1 - GENETICA

Settore Scientifico Disciplinare: BIO/18 - GENETICA

THE INTEGRATOR-PP2A COMPLEX AT THE CROSSROADS OF DEVELOPMENT AND CANCER

Presentata da: Luca Grillini

Coordinatore Dottorato

Anna Maria Porcelli

Supervisore

Giovanni Perini

Co-supervisore

Alessandro Gardini

Esame finale anno 2025

ABSTRACT

Transcriptional regulation is critically mediated by the Integrator complex and the Protein Phosphatase 2A (PP2A), which together form the Integrator–PP2A complex (INTAC). This complex plays a critical role in the pause-release mechanism of RNA Polymerase II (RNAPII), which spatiotemporally regulates gene expression; therefore, perturbations in this system can have significant biological consequences.

This thesis investigates the role of the PP2A-Integrator axis in two contexts: neurodevelopmental disorders caused by the R182W mutation in PPP2R1A, an essential subunit of PP2A, and prostate cancer progression involving the loss of Integrator complex Subunit 6 (INTS6), another fundamental component of the INTAC complex.

In the first study, Induced Pluripotent Stem Cells (iPSCs) were chosen as a model to study the R182W mutation. We first observed morphological abnormalities and defective neural network formation. Then, transcriptomic and genomic analysis revealed extensive dysregulation of genes critical for neurogenesis and altered RNAPII binding patterns. Finally, mass spectrometry data showed disrupted interactions between PP2A and the Integrator complex, suggesting that defective PP2A-Integrator interactions contribute to the pathogenesis of neurodevelopmental disorders associated with PPP2R1A mutations.

In the second study, we examined INTS6 in the context of prostate cancer. Contrary to its known tumor-suppressive function, CRISPR/Cas9-mediated knockout of INTS6 in DU145 prostate cancer cells resulted in reduced proliferation and tumor growth both *in vitro* and *in vivo*. Gene expression profiling revealed dysregulation of pathways involved in cell cycle progression and differentiation, while the investigation of DDX26B's (a homolog of INTS6) potential roles showed that alternative splicing of exon 11 affects its interaction with the Integrator complex and its ability to compensate for INTS6 loss.

The connection between these studies lies in the crucial role of PP2A-Integrator interactions in maintaining proper transcriptional regulation. Disruption of this axis - whether through the impairment of PPP2R1A functions or the loss of INTS6 - leads to significant changes in gene expression, thus affecting differentiation and proliferation. This highlights the PP2A-Integrator complex as a key factor in both neurodevelopmental disorders and cancer.

INDEX

3 - INTS6 role in prostate cancer progression

INTRODUCTION

1.1 Transcriptional regulation

Transcriptional regulation is a fundamental aspect of cellular biology, serving as the cornerstone for gene expression and cellular function. In eukaryotic organisms, despite each cell containing an identical genome, the selective expression of genes allows for cell specialization and the execution of complex biological processes such as development, differentiation, and proliferation. The precise control of transcription - the synthesis of RNA from a DNA template by RNA polymerase II (RNAPII) - is essential for maintaining cellular homeostasis and enabling dynamic responses to developmental cues and environmental stimuli.

The transcription cycle

The transcription cycle of RNAPII is a highly regulated process comprising several stages: pre-initiation complex assembly, initiation, promoter-proximal pausing, elongation, and termination (Core & Adelman, 2019) (**Fig.1**). Each stage is modulated by a network of transcription factors, coactivators, and chromatin remodelers that collectively ensure accurate and efficient gene expression.

During the initiation, the assembly of the Pre-Initiation Complex (PIC) at the core promoter involves general transcription factors along with RNAPII (Hahn, 2004). The mediator complex acts as a bridge between transcriptional activators and the pre-initiation complex, facilitating transcriptional regulation in response to signaling pathways (Soutourina, 2018). Following initiation, RNAPII often enters a promoter-proximal paused state approximately 30–50 nucleotides downstream of the Transcription Start Site (TSS) (Adelman & Lis, 2012). This pausing is mediated by the Negative Elongation Factor (NELF) and DRB Sensitivity-Inducing Factor (DSIF) (Vos et al., 2018), serving as a regulatory checkpoint that allows for rapid transcriptional activation in response to stimuli.

The transition from pausing to productive elongation is regulated by the Positive Transcription Elongation Factor b (P-TEFb), a kinase complex composed of Cyclin-Dependent Kinase 9 (CDK9) and Cyclin T1 or T2 (Peterlin & Price, 2006). P-TEFb phosphorylates the C-terminal domain (CTD) of RNAPII at serine 2 residues within its heptapeptide repeats (YSPTSPS), as well as NELF and DSIF. Phosphorylation leads to the dissociation of NELF and conversion of DSIF into a positive elongation factor, allowing RNAPII to proceed with elongation (Core & Adelman, 2019). The RNAPII CTD acts as a platform for the recruitment of various RNA processing factors, coordinating

transcription with co-transcriptional processes such as capping, splicing, and 3′ end processing (Harlen & Churchman, 2017).

Termination of transcription involves the recognition of polyadenylation signals and the cleavage of the nascent RNA, followed by the release of RNAPII from the DNA template. This process is mediated by the Cleavage and Polyadenylation Specificity Factor (CPSF) complex and other associated factors (Proudfoot, 2016). Precise transcriptional regulation ensures that genes are expressed at the appropriate time, location, and levels, which is crucial for development and adaptation to environmental changes. Dysregulation of transcription can lead to aberrant gene expression patterns associated with diseases such as cancer, neurodevelopmental disorders, and metabolic syndromes (Lee & Young, 2013).

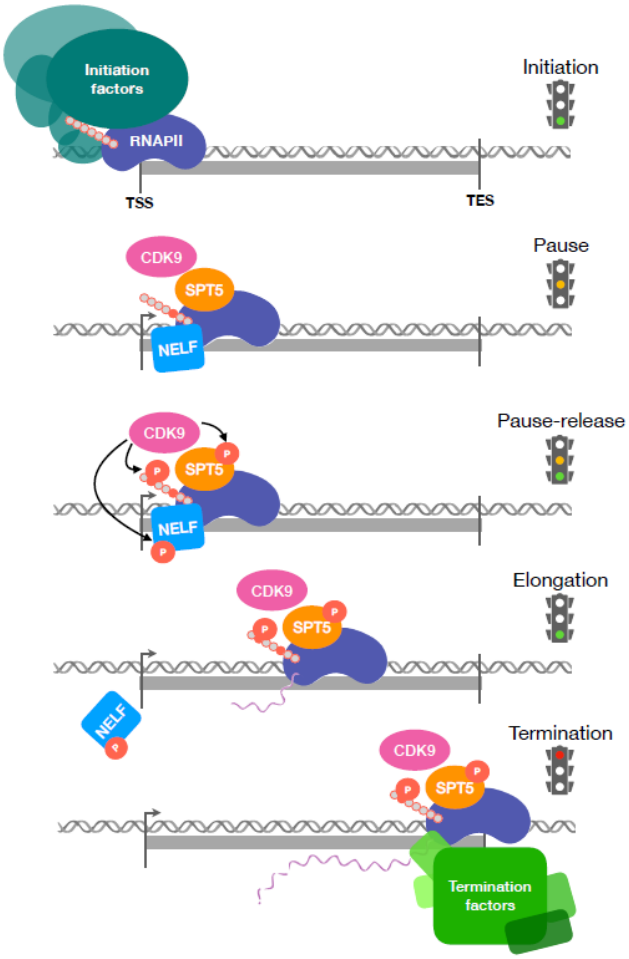


Fig. 1 | Simplified schematic of the steps and checkpoints in the RNAPII transcription cycle. Traffic lights represent the signal received by RNAPII at each labeled step. “Initiation factors” represent general transcription factors and the Mediator complex. The grey circles trailing behind RNAPII represent one heptad repeat (YSPTSPS) of the C-terminal domain (CTD) of RNAPII, with red-orange circles indicating phosphorylation of that residue. At the pause site, CDK9 (pTEFb), SPT5 (DSIF), and NELF are present with RNAPII, the CTD is phosphorylated at Serine 5 (Ser5). During the pause-release phase, CDK9 phosphorylates SPT5, NELF, and the C-terminal domain (CTD) of RNAPII at Serine 2 (Ser2). Phosphorylated NELF dissociates from the holoenzyme, and RNAPII is released into elongation. “Termination factors” represent CPSF and the exonuclease XRN2, among others.

1.2 The Integrator complex

The Integrator complex was first identified by Baillat et al. (2005) as a multi-subunit complex that associates with the RNAPII CTD and mediates the 3′ end processing of small nuclear RNAs (snRNAs). Unique to metazoans and absent in yeast, Integrator is believed to have evolved to facilitate complex gene regulatory mechanisms in higher organisms (Gardini et al., 2014).

Composition and Modular Organization

The Integrator complex comprises 15 subunits (INTS1–INTS15), organized into distinct modules with specialized functions (Offley et al., 2023; Skaar et al., 2015) (**Fig.2**). The endonuclease module consists of INTS9 and INTS11, which form a heterodimer analogous to CPSF100 and CPSF73 in the CPSF complex. INTS11 provides the catalytic activity for RNA cleavage (Albrecht & Wagner, 2012; Fianu et al., 2021), while INTS4 serves as a scaffold stabilizing the endonuclease module. The phosphatase module, known as the Integrator–PP2A complex (Int-PP2A or INTAC), includes INTS5, INTS6, and INTS8 in association with Protein Phosphatase 2A (PP2A) (Zheng et al., 2020). This module is involved in dephosphorylating RNAPII and associated factors. The enhancer module, comprising INTS10, INTS13, INTS14, and INTS15, is implicated in the regulation of enhancer RNA (eRNA) transcription, associating with active enhancers and influencing gene expression at distal regulatory elements (Lai et al., 2015; Skaar et al., 2015). The central scaffold of Integrator is formed by INTS1, INTS2, and INTS7, facilitating the assembly and structural integrity of the complex (Welsh & Gardini, 2023). Additionally, INTS3 associates with the Sensor of Single-Stranded DNA (SOSS) complex involved in DNA damage response and genome stability (Welsh & Gardini, 2023).

Functional Roles

Integrator mediates the 3′ end processing of U snRNAs, essential components of the spliceosome (Baillat et al., 2005). The endonuclease activity of INTS11 cleaves the nascent snRNA transcripts, a critical step distinct from the cleavage of polyadenylated mRNAs. Beyond snRNA processing, Integrator regulates RNAPII transcription of protein-coding genes. Recruited to gene promoters and enhancers in a transcription-dependent manner, Integrator influences RNAPII pause-release and elongation (Gardini et al., 2014; Skaar et al., 2015). Loss of Integrator subunits leads to defects in transcriptional responses, reduced RNAPII occupancy along gene bodies, and decreased mRNA levels, indicating its role in facilitating transcription elongation (Gardini et al., 2014). Integrator also modulates the transcription and processing of eRNAs at active enhancers. The endonuclease activity of INTS11 is required for proper termination of eRNA transcription, preventing interference with downstream genes and maintaining transcriptional fidelity (Lai et al., 2015). Through its association

with the SOSS complex, Integrator participates in DNA damage response pathways, contributing to genome stability (Welsh & Gardini, 2023). INTS3 and associated subunits help prevent the accumulation of R-loops and facilitate repair processes.

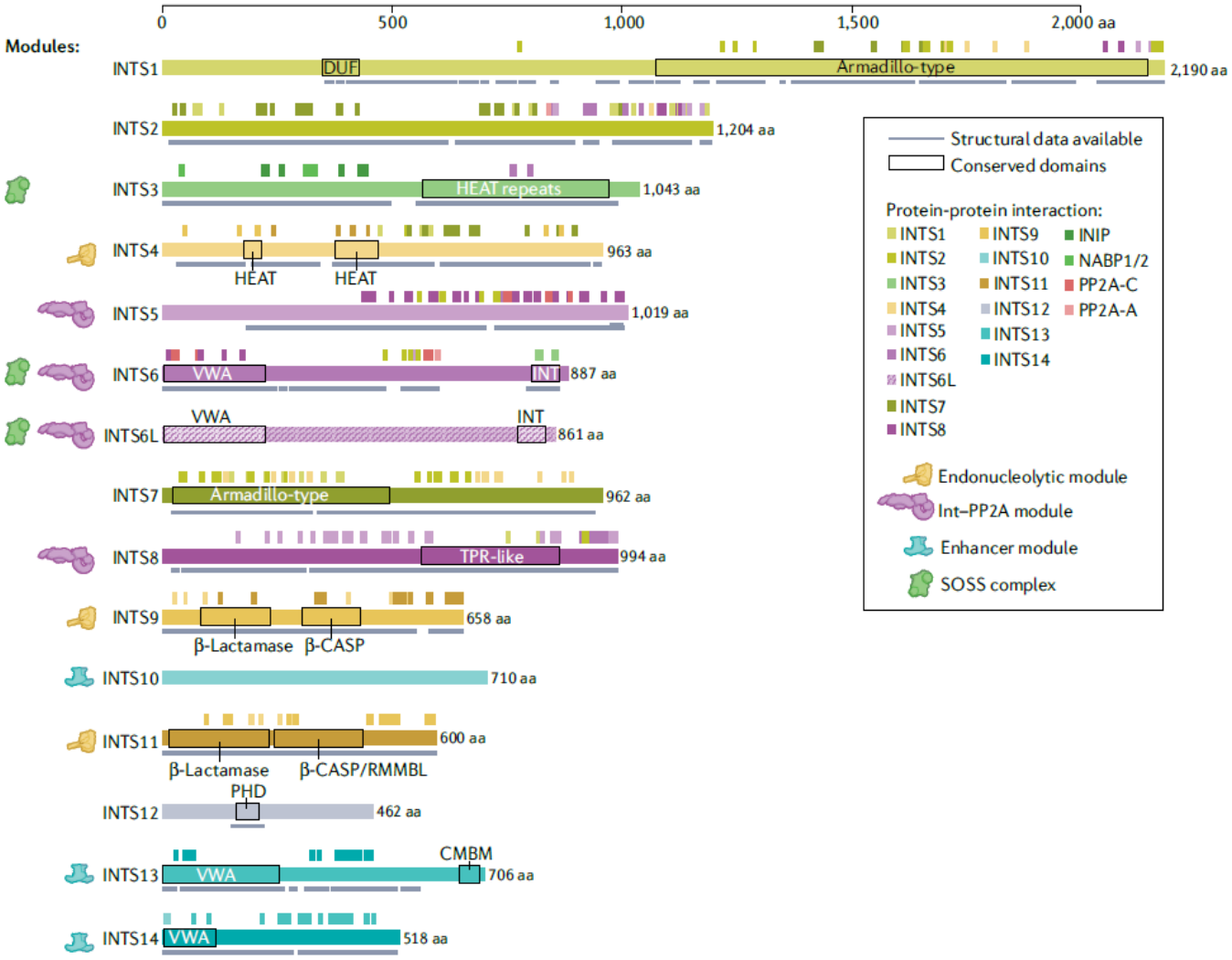


Fig. 2 | The 15 known subunits of the mammalian Integrator complex are depicted to scale. (Welsh & Gardini, 2023) Both Integrator complex Subunit 6 (INTS6) and its mutually exclusive paralogue INTS6L are shown. All annotated protein motifs and domains are named and boxed. The availability of published structural data (cryogenic electron microscopy or X-ray crystallography) is marked by an underline. Colored bars on top of each subunit diagram denote a mapped interaction surface with another member of Integrator, serine/threonine-protein phosphatase 2A (PP2A) or the sensor of single-stranded DNA (SOSS) complex, according to the specific color code. A symbol at the left indicates whether the corresponding subunit is a component of the endonuclease module, the Integrator–PP2A (Int–PP2A) module or the enhancer module, or whether it is found in the SOSS DNA repair complex. aa, amino acids; CMBM, cleavage module-binding motif; DUF, domain of unknown function; VWA, von Willebrand factor type A domain.

1.3 The Integrator-PP2A complex (INTAC)

The Integrator–PP2A complex (INTAC) integrates the phosphatase activity of PP2A with the transcriptional regulatory functions of Integrator (Vervoort et al., 2021; Zheng et al., 2020). INTAC comprises the phosphatase module of Integrator, consisting of INTS5, INTS6, and INTS8, in association with the PP2A core dimer composed of the scaffold subunit PPP2R1A and the catalytic subunit PP2A-C (**Fig.3**). Structural studies have revealed that INTS8 binds to the scaffold subunit PPP2R1A, while INTS6 interacts directly with the catalytic subunit PP2A-C, effectively anchoring PP2A within the complex (Zheng et al., 2020). This arrangement positions the phosphatase activity in proximity to RNAPII and its associated factors, enabling precise regulation of transcription.

Functional role in transcriptional regulation

INTAC plays a critical role in enforcing RNAPII promoter-proximal pausing by opposing the phosphorylation of the RNAPII CTD and elongation factors (Vervoort et al., 2021). PP2A within INTAC dephosphorylates serine 2 and serine 5 residues of the RNAPII CTD, as well as DSIF and NELF. This dephosphorylation maintains RNAPII in a paused state, preventing premature elongation and ensuring proper regulation of gene expression; this regulatory mechanism allows cells to modulate gene expression in response to developmental cues and environmental signals, contributing to processes such as differentiation, stress responses, and cell cycle progression. Loss of INTAC components results in increased phosphorylation of RNAPII and elongation factors, leading to enhanced transcriptional elongation and potential aberrant gene expression (Shi, 2009; Vervoort et al., 2021).

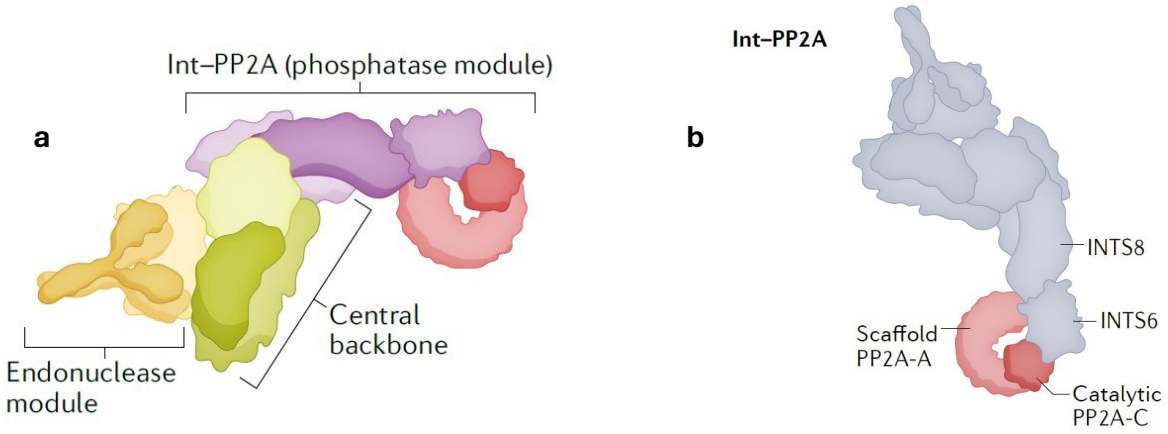


Fig. 3 | The Int-PP2A or INTAC complex. (Welsh & Gardini, 2023) **(a)** overview of the fully assembled INTAC, which is composed of the phosphatase module, the central backbone, and the enhancer module. SOSS is not depicted, despite INTS3 showing evidence of interacting with the C-terminus of INTS6. **(b)** PP2A core, consisting of a scaffold A subunit and a catalytic C subunit, binding Integrator (Int-PP2A) instead of a regulatory B subunit.

1.4 PP2A structure and general functions

The Protein Phosphatase 2A (PP2A) is a highly conserved serine/threonine phosphatase accounting for a significant portion of cellular phosphatase activity (Wlodarchak & Xing, 2016). PP2A is a heterotrimeric holoenzyme composed of a catalytic subunit (PP2A-C), a scaffold subunit (PP2A-A or PPP2R1A/B), and a regulatory B subunit (**Fig.4**). The catalytic subunit exists in α and β isoforms and is responsible for the enzyme's active site. The scaffold subunit anchors the catalytic and regulatory subunits, forming an elongated horseshoe-shaped structure that provides binding interfaces (Xu et al., 2006). The regulatory B subunits confer substrate specificity and localization, classified into four families: B/B55/PR55, B′/B56/PR61, B″/PR72/PR130, and B‴/Striatins. The combinatorial assembly of these subunits results in a diverse array of PP2A holoenzymes with distinct functions (Sents et al., 2013).

Functional roles

PP2A regulates numerous cellular processes, including cell cycle progression, signal transduction, apoptosis, and cytoskeletal dynamics. In cell cycle progression, PP2A controls the activity of cyclin-dependent kinases and other cell cycle regulators, influencing transitions between cell cycle phases (Ruvolo, 2016). In signal transduction, PP2A modulates pathways such as MAPK, Wnt, and Akt by dephosphorylating key components (Perrotti & Neviani, 2013). PP2A also influences apoptotic pathways by regulating the phosphorylation state of proteins like Bcl-2 and BAD, thereby affecting cell survival (Janghorban et al., 2014).

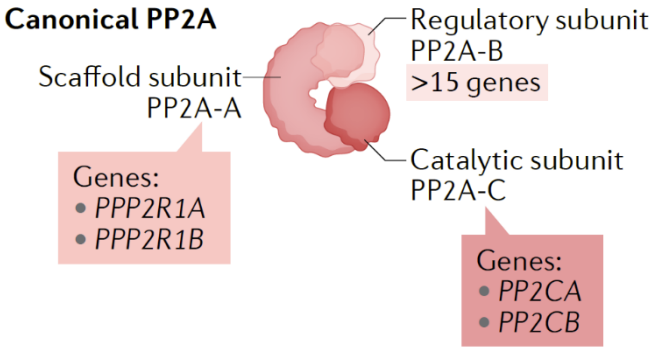


Fig. 4 | PP2A canonical complex. (Welsh & Gardini, 2023) The structural scheme of the canonical PP2A complex, composed of one scaffold subunit (A), one catalytic subunit (C) and one regulatory subunit (B).

1.5 PPP2R1A structure and the R182W mutation

PPP2R1A encodes the α isoform of the PP2A scaffold subunit A, essential for holoenzyme assembly and stability. The scaffold subunit comprises 15 tandem HEAT repeats, forming the elongated horseshoe-shaped structure (Xu et al., 2006) previously mentioned. Each HEAT repeat consists of two α-helices connected by a short loop, providing a modular framework for protein-protein interactions. The inner and outer surfaces of the HEAT repeats provide binding interfaces for the catalytic and regulatory subunits of PP2A (Groves et al., 1999), and this structural arrangement facilitates the assembly of diverse PP2A holoenzymes with specific regulatory subunits, determining substrate specificity and localization within the cell.

The R182W mutation

The R182W mutation in PPP2R1A involves a missense substitution of arginine to tryptophan at position 182, located within the HEAT repeat 5. This residue is critical for binding regulatory B subunits, particularly those in the B56 family (Haesen et al., 2016). The mutation disrupts the scaffold subunit's ability to interact with specific B subunits, leading to altered substrate specificity and impaired PP2A holoenzyme assembly: this can result in hyperphosphorylation of PP2A targets and dysregulation of signaling pathways crucial for cellular function.

The R182W mutation has been identified in various diseases. In cancer, for example, is has been found in endometrial, ovarian, and lung cancers, contributing to oncogenesis through PP2A inactivation (Haesen et al., 2016), because the loss of PP2A tumor suppressor activity leads to enhanced cell proliferation and survival. In neurodevelopmental disorders, germline mutations like R182W are associated with intellectual disability and developmental delays, highlighting the importance of PP2A in neuronal development and function (Sanders et al., 2018). The dominant-negative effect of the mutant scaffold subunit interferes with normal PP2A function, exacerbating the impact on cellular signaling pathways involved in neurodevelopment.

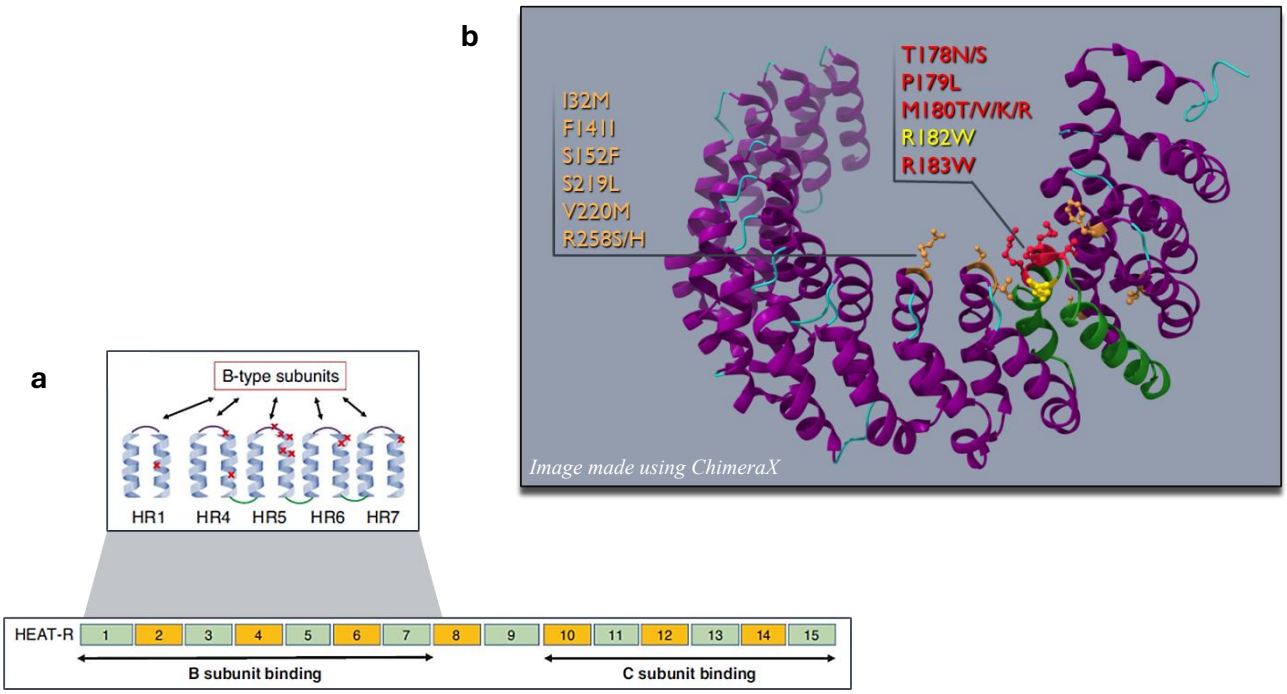


Fig. 5 | PPP2R1A structure and recurring mutations. Image adapted from (Lenaerts et al., 2021) **(a)** The structure of PPP2R1A, consisting of 15 HEAT repeats (HR), is shown. Several recurrent mutations identified in cases of neurodevelopmental disorders are shown as red Xs. Many of them, including R182W, are concentrated on HR5, which represents a hotspot for mutations. **(b)** Structural image of PPP2R1A. The HEAT repeats are shown in purple, except for HR5, which is shown in green; the R182W mutation is highlighted in yellow; other recurrent mutations that are part of the same hotspot are shown in red. Lastly, other recurrent mutations that are not part of HR5 are shown in orange. Note how, due to the three-dimensional conformation of PPP2R1A, all mutations, including those outside of HR5, are found to be part of the same interacting surface.

1.6 Neurodevelopmental disorders and the R182W mutation of PPP2R1A

The neurodevelopmental disorders (NDDs) are a group of conditions characterized by impairments in cognition, behavior, and motor skills, arising from atypical brain development. They include disorders such as autism spectrum disorder (ASD), intellectual disability (ID), attention-deficit/hyperactivity disorder (ADHD), and cerebral palsy (Thapar et al., 2017). NDDs result from genetic, environmental, and epigenetic factors that disrupt neuronal development. Genetic mutations

affecting neurodevelopmental genes play a significant role, often involving genes crucial for neuronal migration, synaptic function, and signal transduction (Rodger et al., 2020).

The R182W mutation in PPP2R1A and NDDs

Individuals harboring the R182W mutation in PPP2R1A exhibit a range of clinical manifestations, including intellectual disability ranging from mild to severe cognitive impairment, developmental delays in motor milestones and speech development, and neurological features such as hypotonia, seizures, and structural brain anomalies like corpus callosum hypoplasia (Y. Zhang et al., 2020). The mutation leads to impaired PP2A assembly due to reduced binding of B56 regulatory subunits, affecting holoenzyme formation and function (Haesen et al., 2016). Dysregulated signaling pathways, such as hyperactivation of the mTOR and MAPK pathways, which are critical for neuronal growth and synaptic plasticity, result from the impaired dephosphorylation of key substrates (Lenaerts et al., 2021).

1.7 Integrator complex Subunit 6 (INTS6) and prostate cancer

The Integrator complex Subunit 6 (INTS6), also known as Deleted in Cancer 1 (DICE1), as previously described, is a core component of the Integrator complex's phosphatase module (INTAC) because it interacts directly with the catalytic subunit PP2A-C and with NELF-B, anchoring PP2A to the complex and facilitating the dephosphorylation of RNAPII and associated factors (Zheng et al., 2020). Beyond transcriptional regulation, INTS6 contributes to genome stability through interactions with the SOSS complex, participating in DNA damage response and repair mechanisms (Skaar et al., 2015). INTS6 has also been implicated in cell cycle checkpoints and apoptosis, influencing cellular proliferation and survival (Lui et al., 2017).

Tumor Suppressor Properties

INTS6 is considered a tumor suppressor gene due to its frequent loss of function in various cancers. Deletions and reduced expression of INTS6 are observed in lung, breast, prostate, and other cancers, correlating with disease progression and poor prognosis (Wieland et al., 2001). Restoration of INTS6 expression in cancer cell lines, on the other hand, suppresses tumor cell proliferation, induces apoptosis, and inhibits colony formation, supporting its role in tumor suppression (Kapp et al., 2013).

INTS6 suppresses tumor growth through multiple mechanisms. In the context of transcriptional regulation, loss of INTS6 leads to hyperphosphorylation of RNAPII and increased transcriptional elongation of oncogenic genes, resulting in aberrant gene expression profiles that promote

tumorigenesis (Vervoort et al., 2021). INTS6 also induces cell cycle arrest at the G1 phase by modulating cell cycle regulators such as cyclin D1 and p21, thereby inhibiting cellular proliferation (Kapp et al., 2013). Additionally, INTS6 promotes apoptosis by upregulating pro-apoptotic factors and downregulating survival signals, enhancing the sensitivity of cancer cells to apoptotic stimuli (Lui et al., 2017).

INTS6 deletion in prostate cancer

In prostate cancer, INTS6 frequently exhibits deletions or reduced expression, which correlates with disease progression and poor clinical outcomes (Kapp et al., 2013). The loss of INTS6 function contributes to the malignant phenotype of prostate cancer cells, promoting resistance to apoptosis and uncontrolled proliferation through the hyperphosphorylation of RNAPII resulting in aberrant gene expression profiles that promote tumorigenesis (Vervoort et al., 2021).

1.8 DDX26B and INTS6: structural similarities and potential functional overlap

DDX26B, also known as INTS6L (Integrator complex Subunit 6-Like), is a homolog of INTS6 and belongs to the DEAD-box helicase family. DEAD-box proteins are involved in various aspects of RNA metabolism, including transcription, splicing, ribosome biogenesis, and RNA decay (Linder & Jankowsky, 2011). INTS6 and DDX26B share a substantial degree of sequence and structural similarity, with about 63% of amino acid identity. This similarity is reflected in their shared domains, particularly the DEAD-box helicase motifs, which are essential for their interactions with RNA and involvement in transcriptional regulation. While the specific functions of DDX26B within the Integrator complex are less characterized, it's known that it can associate with the complex and perhaps contribute to its roles in transcriptional regulation and RNA processing. Also, the potential redundancy or complementary functions between DDX26B and INTS6 could provide robustness in critical regulatory pathways, ensuring proper gene expression control.

AIMS OF THE STUDY

The overarching goal of this thesis is to elucidate the critical roles of the PP2A-Integrator axis in cellular function and disease, focusing on two distinct but mechanistically connected projects: the impact of the R182W mutation in PPP2R1A on neurodevelopment, and the role of INTS6 in prostate cancer progression. By investigating these areas, we aim to deepen our understanding of how disruptions in transcriptional regulation contribute to disease pathology.

For the PPP2R1A project, we initially aim to observe and characterize the morphological and molecular defects caused by the R182W mutation of PPP2R1A: by generating Induced Pluripotent Stem Cells (iPSCs) harboring heterozygous or homozygous R182W mutation, we model the mutation's impact on neural differentiation processes. Then, our objective becomes to determine whether the defects that we observed are caused by an impairment in transcriptional regulation resulting from the loss of PP2A's transcriptional functions, or if they arise from other disrupted PP2A activities.

In the INTS6 project, we explore the consequences of deleting INTS6 in prostate cancer cells, expecting an increase in tumour growth speed and aggressiveness. Contrary to our expectations, preliminary data suggest a decrease in proliferation upon INTS6 deletion, prompting a deeper investigation into the underlying mechanisms. Additionally, we are interested in understanding how alternative splicing of exon 11 affects DDX26B's ability to interact with other proteins, in particular with the Integrator complex, and whether it can functionally substitute INTS6 in prostate cancer cells.

By examining these two systems, we hope to uncover common mechanisms by which disruptions in the PP2A-Integrator axis contribute to disease.

RESULTS

2 – Recurring mutation of PP2A impairs neurogenesis

2.1 Generating a model for neurodevelopmental disorders in iPSCs

In order to create a model for the study of neurodevelopmental disorders, the iPSC Core at the University of Pennsylvania was tasked with generating two Induced Pluripotent Stem Cell (iPSC) SV20 lines that were heterozygous or homozygous for the R182 mutation of PPP2R1A. To achieve this knock-in (KI) it was decided to use a strategy based on the CRISPR/Cas9 system and Homology-Directed Repair (HDR), which is a mechanism that allows cells to repair double-stranded DNA damage by using a portion of the genome homologous to the damaged one as a template (**Fig.6**).

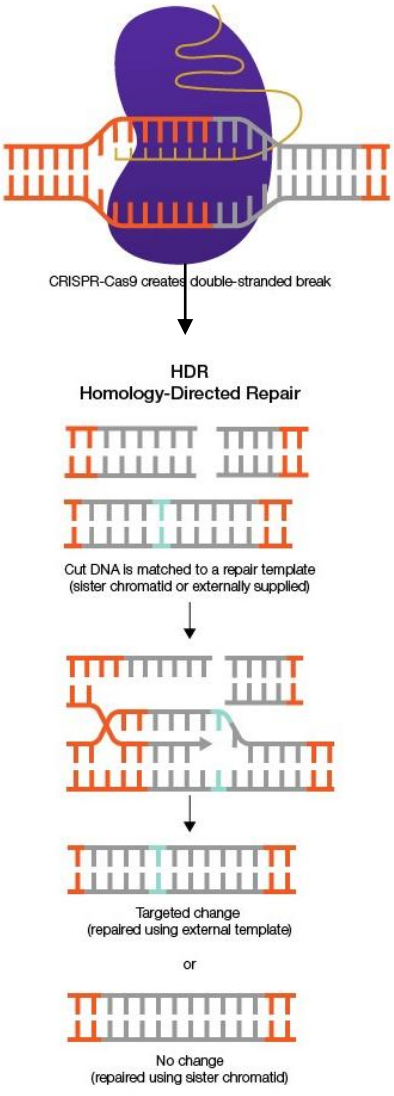


Fig. 6 | The mechanism of Homology-Directed Repair (HDR). Once the caspase has cleaved the DNA, thus introducing a double-strand break, the cell attempts to repair the damage by HDR, that is, by using a homologous region of the genome as a template for repair. In most cases, the homologous region in question is that of the sister chromatid. However, it is also possible for an exogenous template to be utilized. In such instances, a template containing sequence modifications can be employed for the purpose of integrating them into the cell's genome. In the event that HDR is not a viable option, the cell will proceed with the restoration of chromosome structure through Non-Homologous End Joining (NHEJ).

First, a plasmid was constructed to express the CRISPR/Cas9 system, alongside the gRNA to target the region of interest corresponding to R182, and GFP to enable cell sorting. Subsequently, two single-stranded oligodeoxynucleotides (ssODNs), one Wild-Type (WT) and one carrying the R182W substitution (corresponding to the transition from **C**GG to **T**GG triplet), were generated for use as templates for Homology-Directed Repair (**Fig.7**). Additionally, both ssODNs exhibited a silent mutation at the PAM site recognized by the gRNA of the CRISPR/Cas9 machinery (from **C**CC to **T**CC), which was introduced to prevent further cleavage of the same region following the initial cut. A second silent mutation was incorporated approximately 30 bp downstream, generating the restriction enzyme sequence targeted by the *HindIII* enzyme. This was intended to facilitate the distinction of knock-in-bearing cells via PCR (AAGCT**G** to AAGCT**T**).

At this point, the plasmid and the mutant ssODN (in combination with the WT ssODN, if the objective was to obtain a heterozygous mutant) were co-transfected within iPSCs, which were subsequently sorted according to GFP positivity and plated at single-cell densities for clonal isolation. Subsequently, the isolated clones were genotyped by restriction enzyme digestion and PCR, and KI was confirmed by Sanger sequencing (**Fig.8**). The final outcome of the procedure was the generation of two iPSC lines: one homozygous (8B or -/-) and one heterozygous (5A3 or +/-) for the R182W mutation of PPP2R1A.

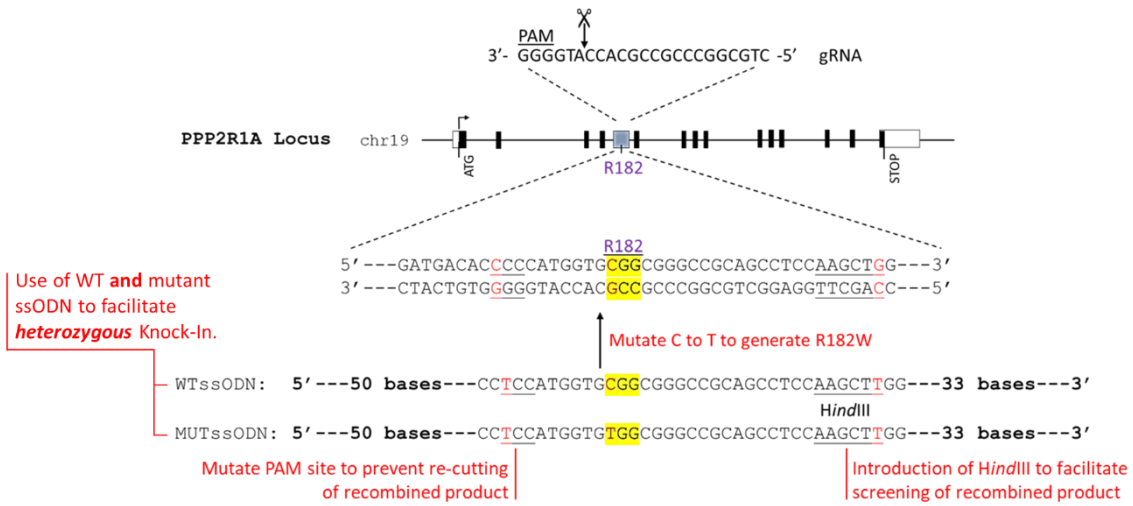


Fig. 7 | Design of the R182W CRISPR/Cas9 Knock-In system. The schematic illustrates the experimental design to generate R182W-positive clones. Of the 198 clones isolated and tested, only 5A3 and 8B were suitable for further analysis, representing heterozygous and homozygous mutations, respectively.

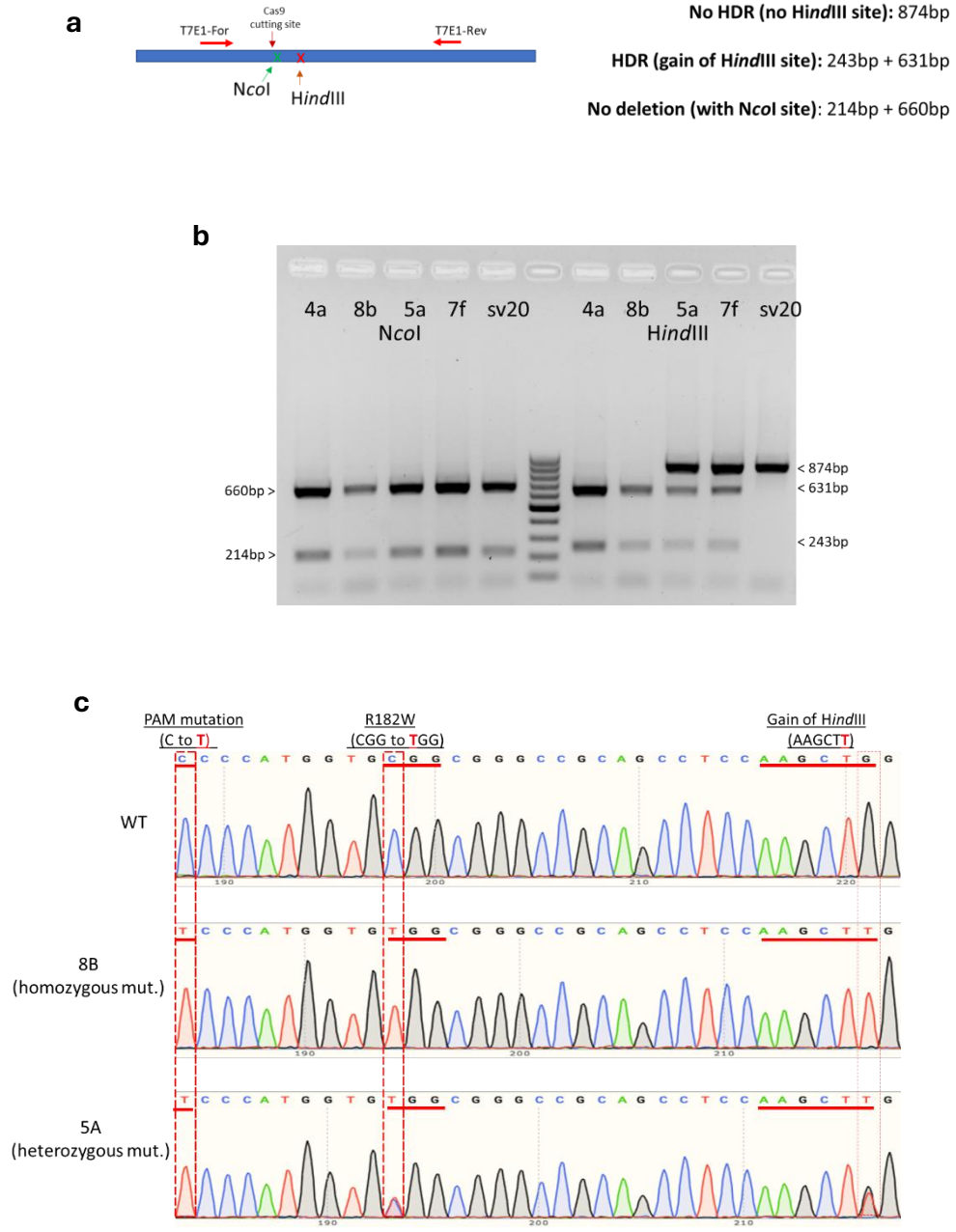


Fig. 8 | Selection and validation of clones carrying the R182W mutation. Panel **(a)** provides a summary of the expected size of the DNA fragments that would be produced following the PCR amplification of the R182-containing locus and subsequent digestion with the restriction enzymes *NcoI* and *HindIII*. The presence of fragments of varying sizes is indeed associated with the status of the mutation, whether inserted or not, and with whether the mutation is present in a homozygous or heterozygous state within the cell. Panel **(b)** depicts the outcome of the electrophoretic run, which was instrumental in identifying clones 8B and 5A. As can be seen in the former, there is an absence of fragments at 214 bp and 874 bp (a combination that indicates the presence of the R182W mutation in homozygosity). In contrast, the latter shows the presence of fragments at 214 and 874 bp, as well as fragments at 243 and 631 bp, indicative of heterozygosity. Panel **(c)** depicts a Sanger sequencing result, which served to confirm the correct insertion of all desired elements.

2.2 Neural differentiation of iPSCs

The iPSC lines obtained from the iPSC Core were subjected to a differentiation process in accordance with the steps illustrated in **Fig.9**: the process commences by differentiating the iPSCs into neural progenitor cells (NPCs) and culminates in the generation of cortical dopaminergic neurons. The entire induction process, from start to finish, takes approximately 42 days; however, this timeframe may be extended. At the conclusion of the protocol presented here, the neurons are not yet fully mature, as the primary distinctions between WT and mutant cells, the focus of our study, emerge during the initial stages of differentiation. Attaining full maturity could be expected around day 63.

It is important to note that the homozygous 8B mutant does not represent a physiological model, as the R182W mutation has never been observed in nature. This may be due to the fact that iPSCs carrying a homozygous R182W mutation are unable to survive differentiation into mature neurons, with lethality exhibited at the very early induction stages. This would likely translate to a lethal scenario *in vivo*, thereby precluding their consideration for any neural induction experiment.

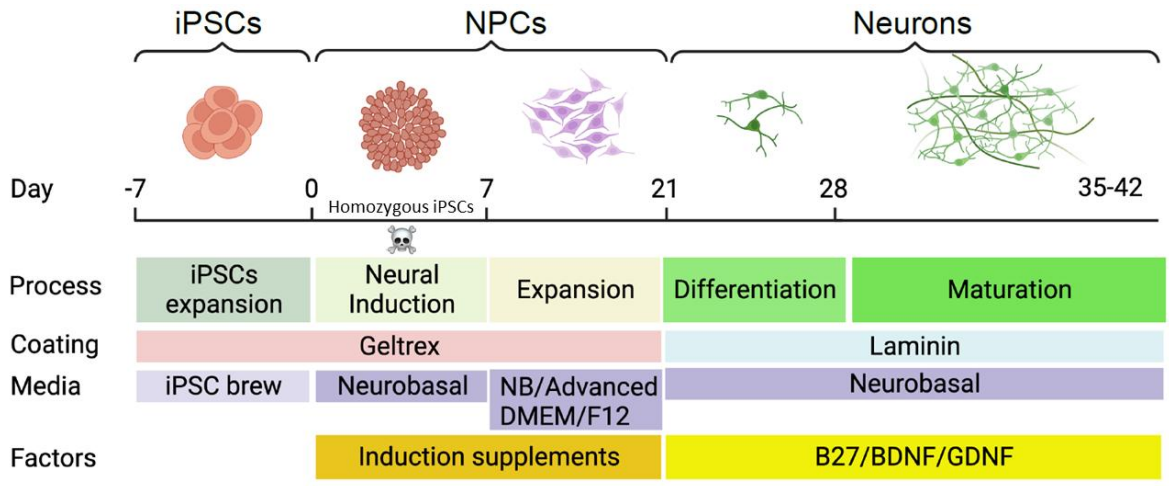
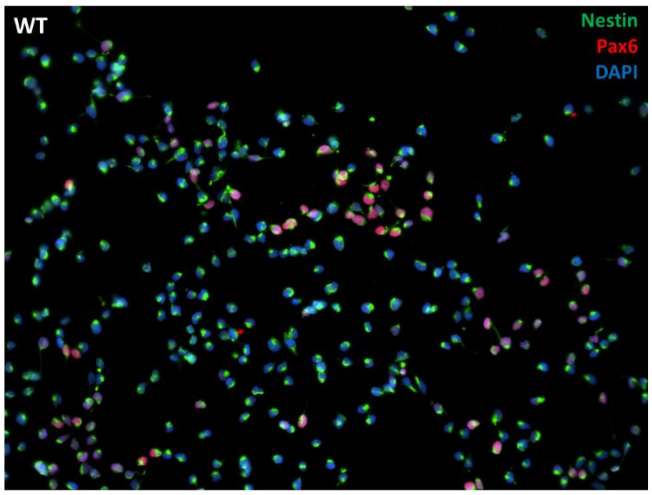


Fig. 9 | Differentiation protocol for iPSCs. The differentiation of Induced Pluripotent Stem Cells (iPSCs) into neural cells commences when the maintenance media is replaced with a medium containing the induction factors. This stage also coincides with the death of iPSCs that are homozygous for R182W. Subsequently, the process progresses through a phase of NPC expansion, followed by a phase of actual neural differentiation (which also necessitates a transition to a distinct extracellular matrix) and ultimately the prolonged maturation of the generated neurons into cortical neurons.

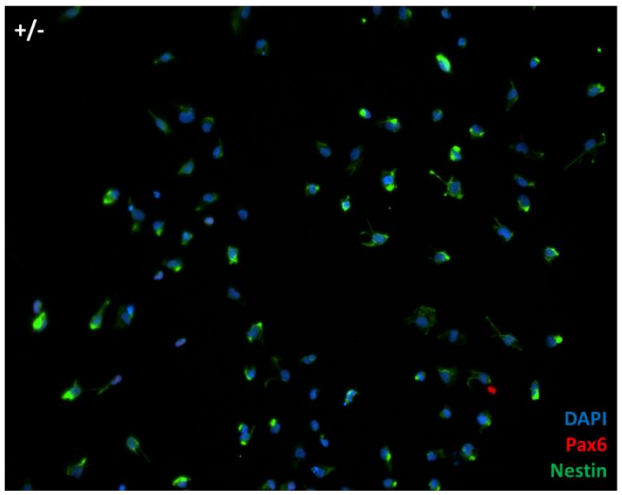
2.3 Morphological differences between WT and 5A3 cells during neural differentiation

The WT and 5A3 iPSCs were initially differentiated into NPCs, and subsequently, after an additional 21 days, into early cortical neurons. The cells were immunostained with various antibodies at the outset of the procedure, specifically at the NPC stage, and also after 7, 14, and 21 days of differentiation. Images were captured at each of these stages using an epifluorescence microscope. As illustrated in **Fig.10a** and **b**, 5A3 NPCs exhibit a notable increase in size and a loss of roundness, accompanied by the presence of cytoplasmic projections. In these panels, the nuclei of the cells were stained with 4',6-diamidino-2-phenylindole (DAPI, in blue), while the cytoplasm was revealed through the use of an antibody targeting Nestin (in green). Pax6, an important neural marker, was stained red instead. A comparison of WT and mutant cells revealed that none of the NPCs 5A3 expressed Pax6.

10a



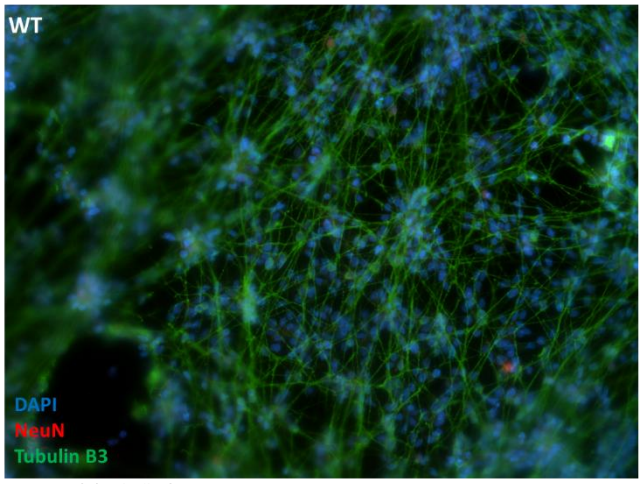
10b



The next panels (**Fig.10c** to **h**) illustrate neurons at varying stages of differentiation. In addition to DAPI staining of the nuclei (blue), the cytoplasm of neurons was indicated by the specific marker Tubulinβ3 (green), while NeuN (red) was used as a marker for the nucleus of later-stage maturation neurons.

The panels **c** and **d** of **Fig.10** illustrate the cellular morphology after seven days of neural differentiation. The morphological differences between the two clones are striking and immediately appreciable. However, it is crucial to highlight that the WT cells have already established a primordial neural network, as evidenced by the sprawling Tubβ3 staining, while the mutant cells appear to have been unable to do so.

10c



10d

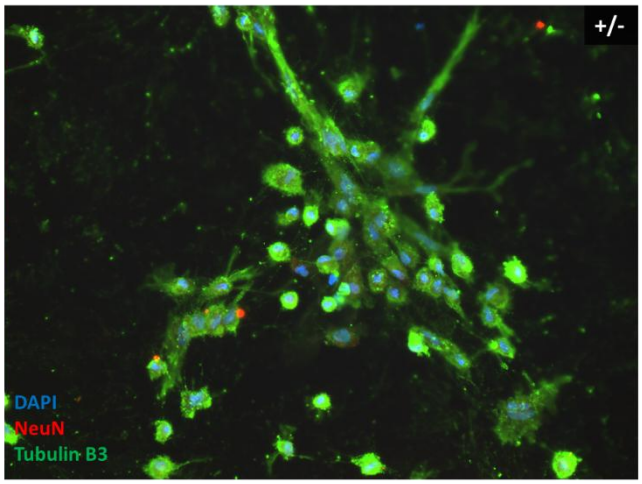
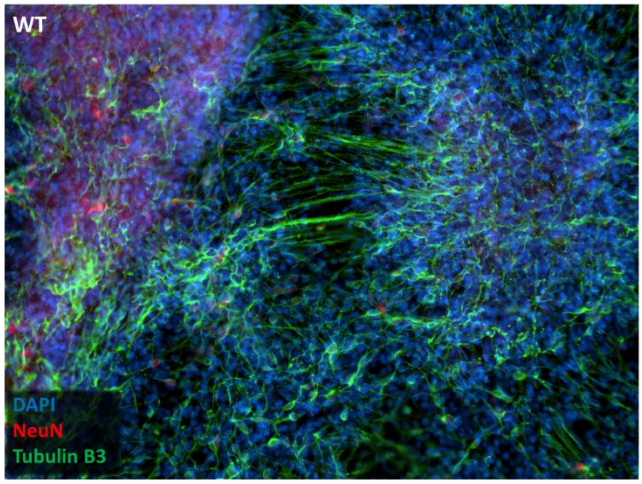


Image made by Yingjie Zhang

After a 14-day differentiation period (illustrated in **Fig.10e - f**), wild-type (WT) neurons begin to organize spontaneously into structures that resemble nuclei (clusters of neurons with similar functions and connections, which in an *in vivo* context would be located inside the central nervous system), and their connections appear more refined. In contrast, the 5A3 cells exhibit only minimal indications of axon and dendrite formation, which are barely visible in the presented images.

10e



10f

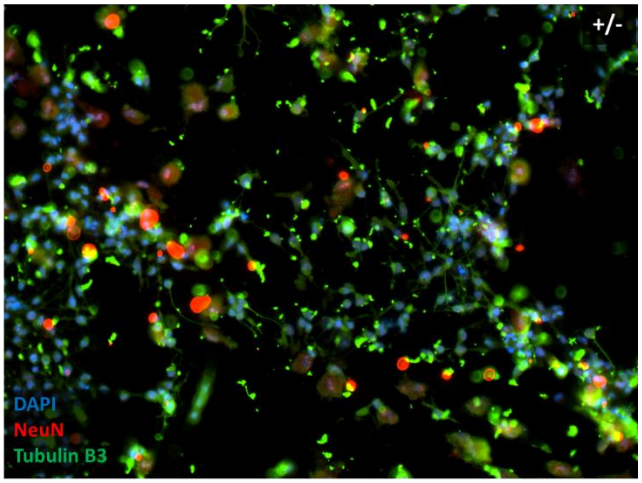
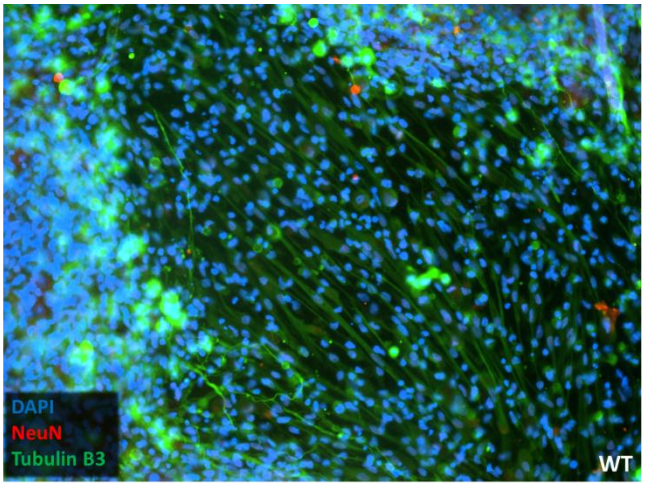


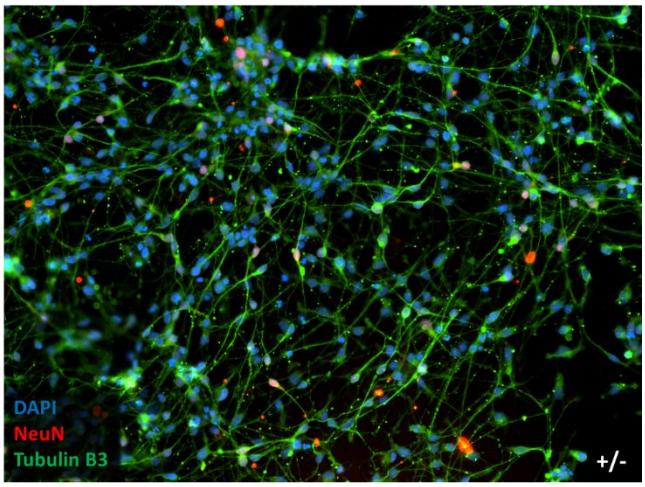
Image made by Yingjie Zhang

At the conclusion of the 21-day process of differentiation of NPCs into early neurons, it is evident that the WT cells have undergone further reorganization into nuclei. Of particular significance is the observation of the orderly directionality of the neural network that has been established, which serves as an indicator of successful differentiation. In contrast, the 5A3 neurons are only now beginning to organize into a less mature and organized network than the WT neurons (**Fig.10g - h**).

10g



10h



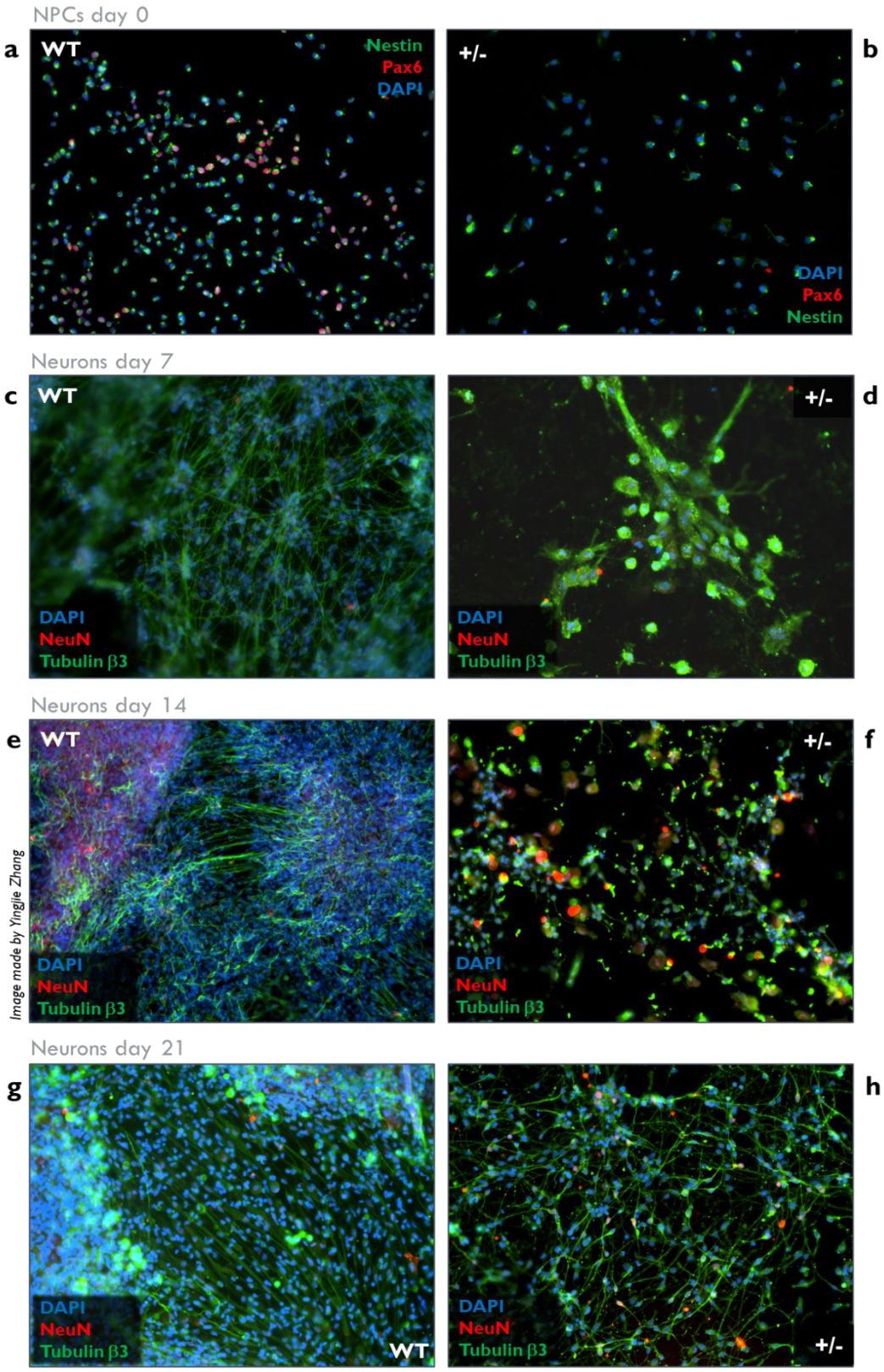


Fig. 10 | Differentiation stages of WT and Heterozygous mutant (+/-) cells. Panels depict WT (left) and R182W heterozygous (+/-, right) cells at various differentiation stages, highlighting morphological differences in neuron formation and neural network development as previously described.

2.4 Differences in gene expression between WT and mutant iPSCs

Given the notable morphological discrepancies observed between WT and 5A3 cells during their differentiation, we proceeded with an analysis of their gene expression. To this end, we initiated our investigation at the fundamental level, conducting total RNA-seq on all three iPSCs lines in our possession, including the homozygous mutant 8B. The results demonstrate significant distinctions between the various lines and the WT line, with the 5A3 iPSCs exhibiting 560 downregulated genes and 493 upregulated genes in comparison to the parental line (**Fig.11a**). This trend is further accentuated when examining the 8B iPSCs, which exhibit a total of 1,572 downregulated genes and 1,053 upregulated genes in comparison to WT (**Fig.11b**). This considerable discrepancy is likely to be a primary factor contributing to their inability to survive even the initial stage of neural differentiation.

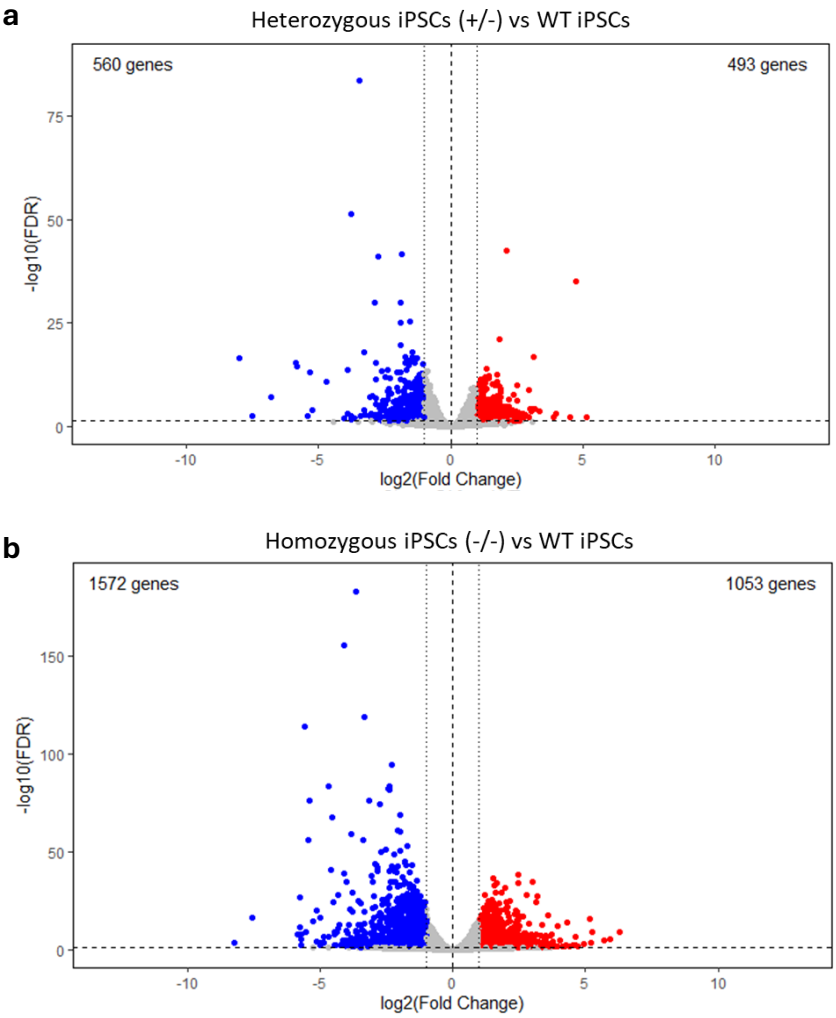


Fig. 11 | Total RNA-seq of WT, Heterozygous and Homozygous iPSCs. Volcano plots show gene expression differences between WT cells and the heterozygous clone 5A3 **(a)** or the homozygous clone 8B **(b)**. The data indicate that the number of dysregulated genes, including both up- and downregulated genes, in clone 8B is more than twice as high as in clone 5A3.

A more detailed examination of the data revealed that only a small proportion of the dysregulated genes were shared between the two groups. Indeed, an examination of the combined set of upregulated and downregulated genes in the two clones reveals that only 9.3% of them are dysregulated in both (**Fig.12a**). This percentage remains relatively consistent when the total up- and downregulated genes are considered separately **(Fig.12b-c)**. However, if we consider each clone individually, the data indicate that the 312 shared genes still account for only the 11.9% of the total dysregulated genes (2625) in clone 8B, but they represent a much more impactful proportion (29.6%) of the total dysregulated genes (1053) found into the heterozygous 5A3 line (a proportion that remains similar when considering the individual categories: namely, 24.1% of upregulated genes and 26.6% of downregulated genes for the 5A3 clone).

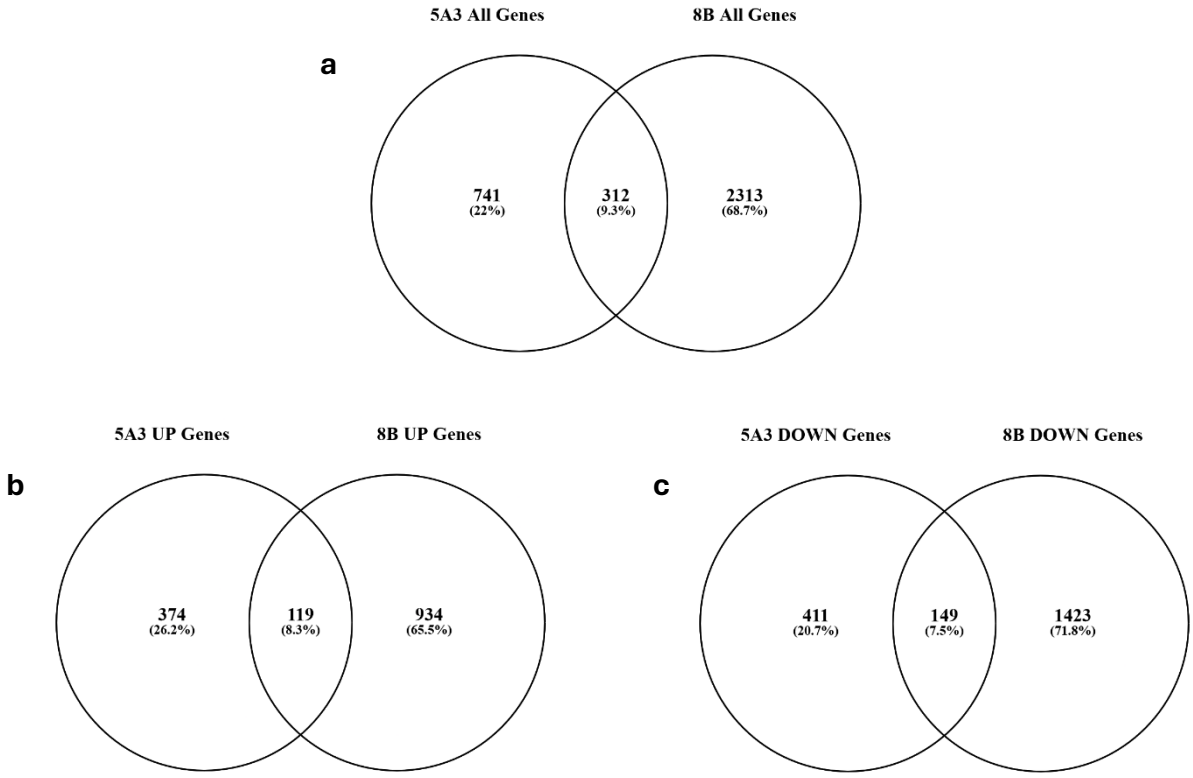


Fig. 12 | Distribution of the dysregulated genes among clones. Venn diagrams display the distribution of shared and unique dysregulated genes in clones 5A3 and 8B, as identified by total RNA-seq. The graphs illustrate the number and respective percentage of shared and unique genes in relation to the total number of upregulated and downregulated genes **(a)**, or only to one of the two conditions, nominally upregulated **(b)** or downregulated **(c)** genes.

The data obtained from total RNA-seq were also analyzed using the g:Profiler software to perform a Gene Ontology Enrichment Analysis (**Fig.13**). The results of this analysis demonstrate that the upregulated genes identified in 5A3 cells predominantly belong to diverse metabolic pathways, primarily associated with the biosynthesis of peptides, organic compounds, and amides. Additionally, they are involved in translation and ribosome biogenesis (**Fig.13a**). Conversely, the downregulated genes in these cells are linked to neurogenesis, systemic development, and the regulation of cellular component organization (**Fig.13b**).

This information thus indicates a general impairment of the neuronal differentiation capabilities of 5A3 iPSCs, and at the same time suggests an increased drive for protein synthesis within them. This suggests that these cells may have an incomplete ability to differentiate into neurons, which is consistent with the results of the immunofluorescence experiments.

In the case of 8B iPSCs, however, the situation appears to be more complex. The upregulated pathways appear to be associated with developmental processes and not only with strictly metabolic or biosynthetic processes, which nevertheless appear to be significantly involved (**Fig.13c**). Conversely, the downregulated pathways appear to be associated with processes controlling and regulating cell and tissue growth and organization, with cell-to-cell communication and signaling being particularly affected, as well as their ability to transport and localize (**Fig.13d**).

The data indicate a markedly strong drive toward individual growth and differentiation of these cells, devoid of control and coordination systems typical of a tissue. This may explain why 8B iPSCs are unable to survive even the earliest phase of neural differentiation (and possibly of any other differentiation in cells belonging to a tissue).

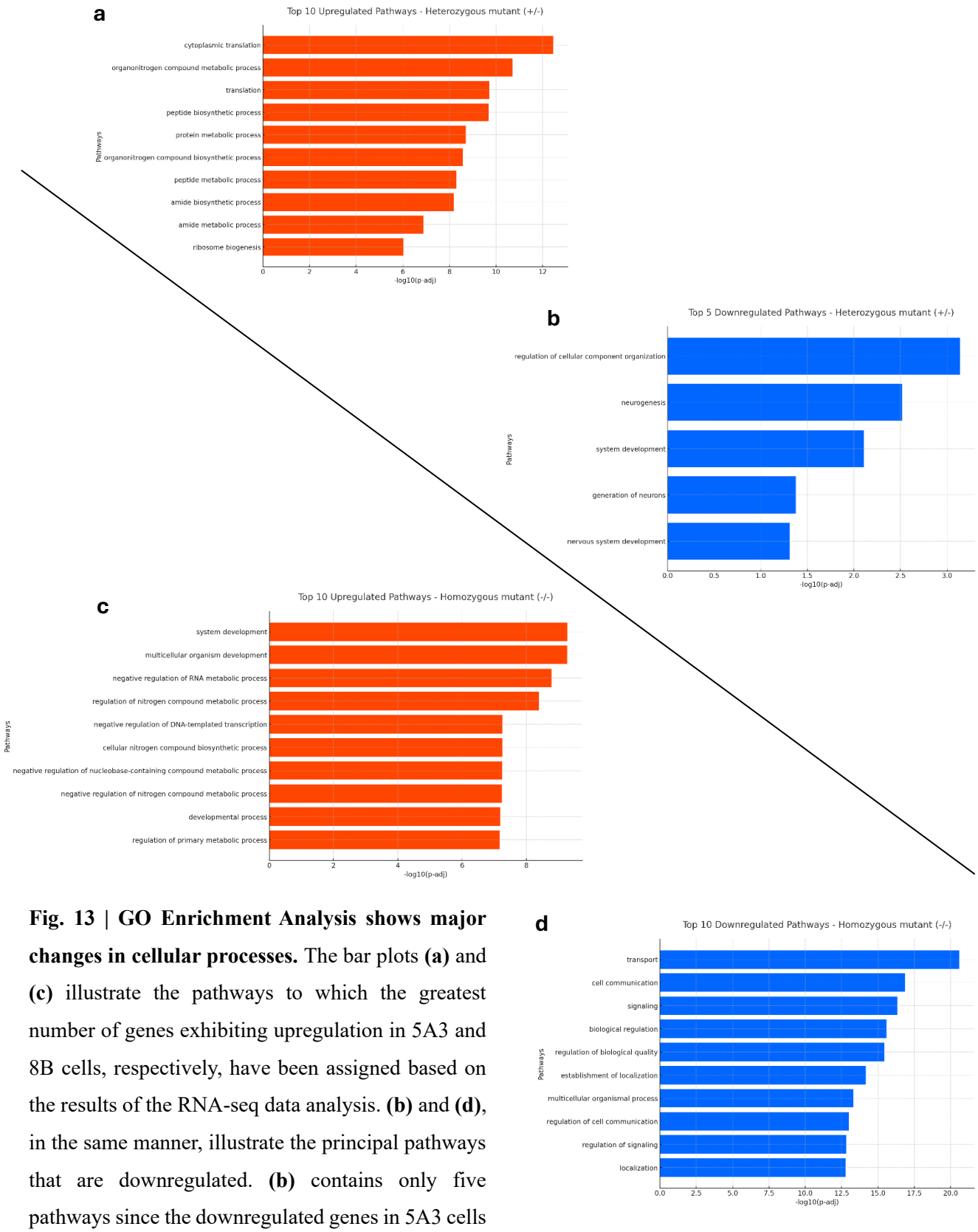


Fig. 13 | GO Enrichment Analysis shows major changes in cellular processes. The bar plots **(a)** and **(c)** illustrate the pathways to which the greatest number of genes exhibiting upregulation in 5A3 and 8B cells, respectively, have been assigned based on the results of the RNA-seq data analysis. **(b)** and **(d)**, in the same manner, illustrate the principal pathways that are downregulated. **(b)** contains only five pathways since the downregulated genes in 5A3 cells are distributed across numerous pathways, and thus, none of them appear to have been significantly affected. The five pathways shown in the figure exceeded the significance threshold by a considerable margin, suggesting that these are the only ones to have been profoundly affected by the mutation. All graphs are based on the adjusted p-value (p-adj).

2.5 ChIP-seq analysis revealed no significant alterations in RNAPII binding

Given the observed alterations in gene regulation among iPSCs bearing the R182W mutation, we proceeded to investigate the behaviour of the RNA Polymerase II (RNAPII) in these cells by ChIP-seq. This experiment allowed us to compare the binding of RNAPII on the genome of WT and 5A3 iPSCs in correspondence of the 3,000 protein-coding genes most transcribed by RNAPII itself.

The results demonstrated that RNAPII exhibited a similar association with DNA in both WT and mutant cells at the transcription start site (TSS **Fig.14b**), with no statistically significant difference between the two (**Fig.14d**). Along the gene body and at the Transcription End Site (TES), however, RNAPII appears to be more present on the genome in WT cells (**Fig.14a-c**). The data collected confirm the general trend previously observed, namely that gene expression in 5A3 cells is partially altered. Furthermore, the unchanged presence of RNAPII at the TSS indicates that recruitment occurs as expected, and that the R182W mutation potentially affects the later stages of transcription.

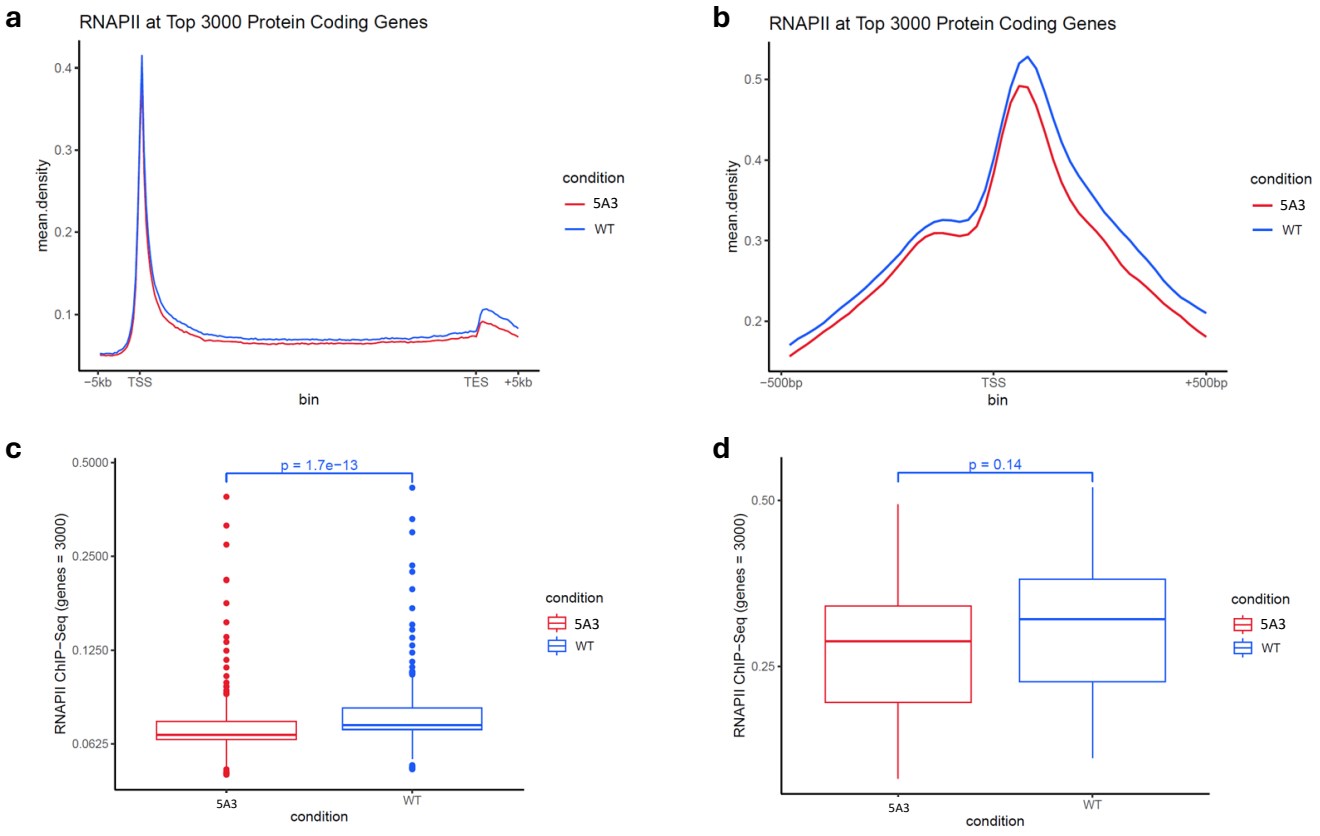


Fig. 14 | ChIP-seq data indicates a change in RNAPII binding to the genome. (a) Average profile of RNAPII binding to genes. The X-axis represents the structure of the genes, with the Transcription Start Site (TSS) at the beginning and the Transcription End Site (TES) at the end, showing also the regions 5kb upstream and downstream of them. **(b)** A closer examination of the peak at the transcription start site (TSS). **(c)** and **(d)** depict box plots quantifying the differences between RNAPII binding in wild-type cells and in 5A3 cells, as illustrated in graphs **(a)** and **(b)**, respectively.

2.6 Dysregulation of gene expression increases dramatically at NPC stage

Following the analysis of gene expression in iPSCs, the neural differentiation process was initiated and a total RNA-seq was made using, this time, NPCs. As previously described, only heterozygous mutants are capable of surviving the differentiation process: therefore, homozygous 8B cells will not be mentioned from this point onwards.

The total RNA-seq analysis revealed that as differentiation proceeded, transcriptional dysregulation in 5A3 mutant cells increases exponentially, reaching a staggering amount of 3621 downregulated and 2872 upregulated genes (**Fig.15a**). A Gene Ontology Enrichment Analysis of the data obtained reveals that seven of the ten pathways most affected by the downregulation of its member genes are related to neurogenesis and development (**Fig.15c**), further confirming the significant impact of the R182W mutation on neuronal differentiation. Regarding the most significantly upregulated pathways, however, the top ranks indicate the presence of systemic developmental processes and anatomical structure formation (**Fig.15b**), suggesting that the propensity for differentiation observed in iPSCs is not only still present but has even become much stronger. Furthermore, subsequent analysis revealed that one of the primary involved pathways appears to be the one responsible for the development of vascular endothelium.

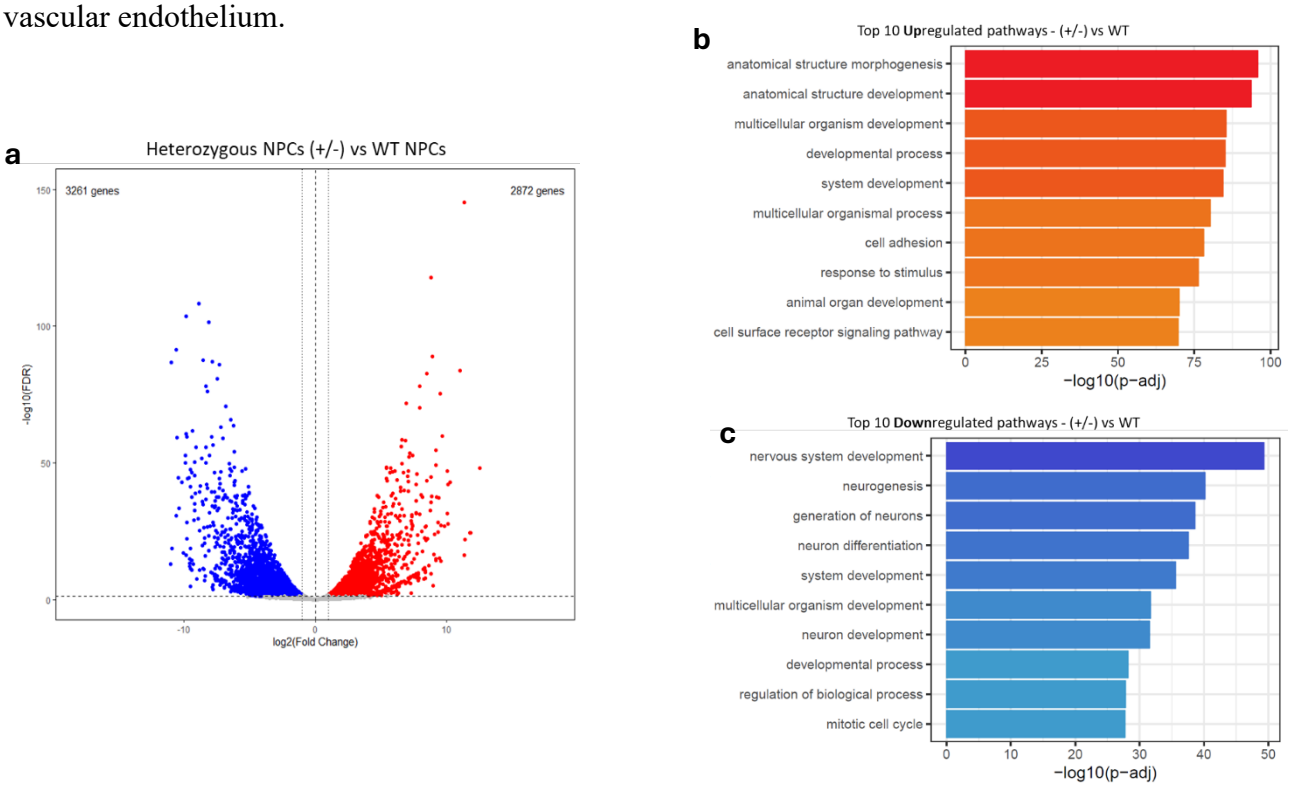


Fig. 15 | Total RNA-seq on NPCs shows heavy gene expression dysregulation. (a) A volcano plot illustrating the distribution of genes that are both upregulated and downregulated in 5A3 NPCs. **(b)** and **(c)** illustrate the ten pathways to which the majority of the up- and downregulated genes in NPCs 5A3 belong, respectively, with the data organized according to their adjusted p-value (p-adj).

2.7 Mass Spectrometry on NPCs reveals alterations in the interactions between Integrator and PP2A

Given that the R182W mutation of PPP2R1A has been demonstrated to have a significant impact on gene expression, we proceeded by investigating the interactome of PPP2R1A itself. The rationale behind this decision is to verify whether the observed changes in the transcriptome are due to an impairment of the function of PPP2R1A and, consequently, of PP2A, which is involved in a multitude of cellular processes.

To this end, a nuclear extract was obtained from WT and 5A3 NPCs, and immunoprecipitation was performed using antibodies specific for PP2CA, which is the major catalytic subunit of PP2A. The reason this antibody was selected in preference to one specific for PPP2R1A is that the latter performs poorly, and also, based on our previous data, the immunoprecipitations conducted on PP2CA appear to pull down a panel of proteins that is highly similar to those obtained by a PPP2R1A immunoprecipitation. Following the immunoprecipitation with PP2CA antibodies, the resulting products were analysed via mass spectrometry, which yielded intriguing results (**Fig.16**).

The differences between the nuclear extracts from WT and 5A3 NPCs are readily apparent, with hundreds of proteins exhibiting higher abundance in the WT extracts. Additionally, we observed that certain protein classes, including zinc fingers and components of the splicing mechanism, as well as transcriptional factors, were all present in higher quantities in the WT fraction, with no proteins belonging to these classes predominantly detected in the 5A3 fraction. What is particularly noteworthy is that all of the immunoprecipitated Integrator and RNAPII subunits are present in extracts from WT NPCs in overwhelming quantities. In contrast, in the fraction belonging to 5A3 NPCs, some regulatory B subunits of PP2A appear in greater amounts, suggesting a change in the PP2A complex composition (and thus in its interactions) in 5A3 cells.

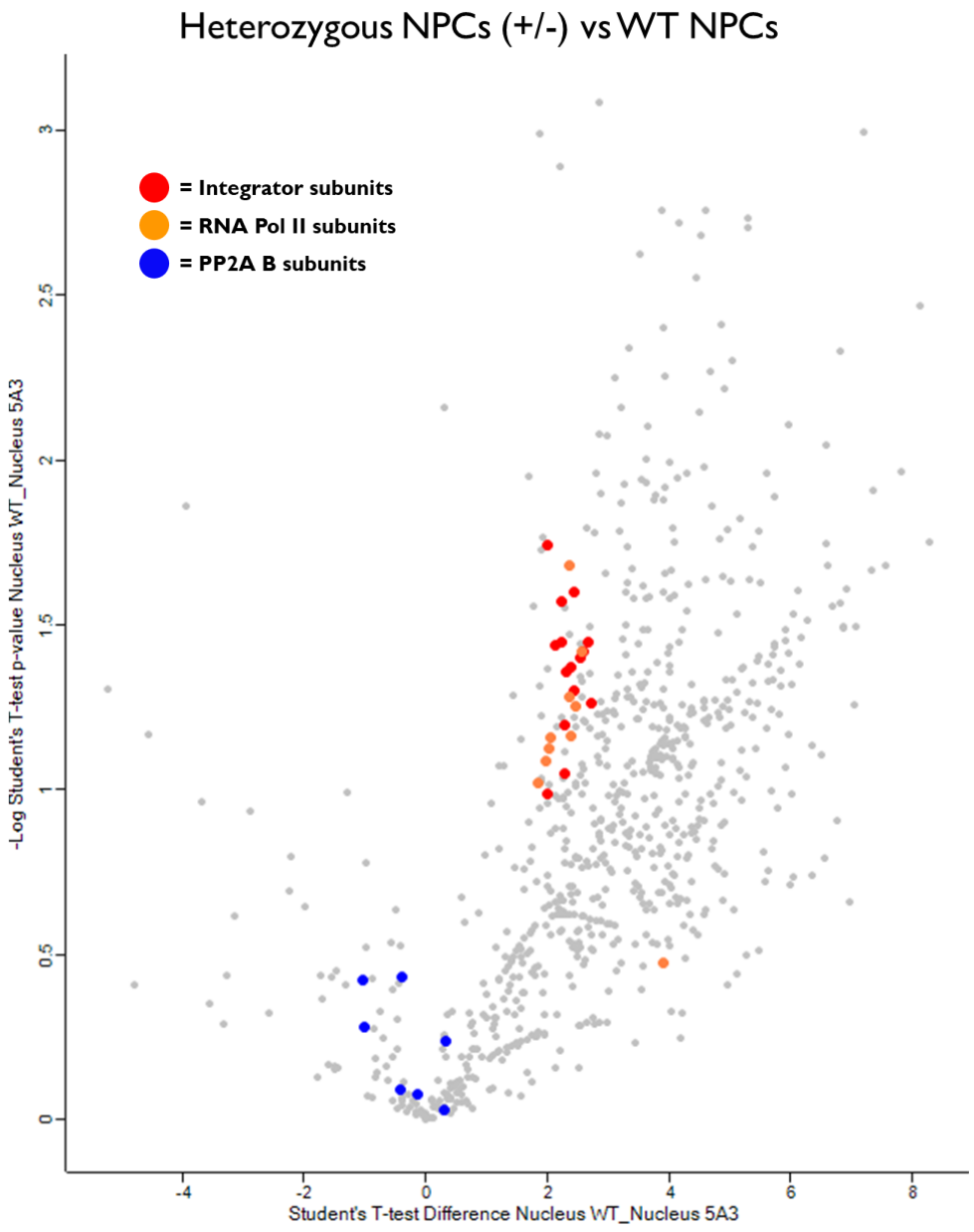


Fig. 16 | Mass spec data suggest a preference for the PP2A canonical complex in 5A3 NPCs. The presented volcano plot illustrates the distribution of proteins that were pulled down through immunoprecipitation using anti-PP2CA antibodies on nuclear extracts derived from NPCs. The positive values on the X-axis correspond to proteins that are exclusively (or at least in greater amounts) present in WT NPCs, while negative values indicate the same but for 5A3 NPCs. All Integrator subunits identified by mass spectrometry are highlighted in red, while the RNAPII subunits are highlighted in orange. The regulatory B subunits of PP2A are highlighted in blue.

3 – INTS6 role in prostate cancer progression

3.1 Modelling INTS6 role in prostate cancer cells

The initial step in our characterization of INTS6 as a tumor suppressor was the creation of a cellular model lacking INTS6 itself. To this end, we selected the DU145 cell line, derived from prostate cancer and expressing INTS6, and performed a CRISPR/Cas9 knockout of INTS6. To this end, a plasmid expressing both the CRISPR apparatus and the necessary gRNA was generated and the construct was introduced into DU145 cells via lentiviral infection. Following clonal selection, three sets of Western blots were performed on samples collected in consecutive weeks to identify clones that showed no trace of INTS6. At the end of this process, three clones were identified, namely clone 7(S), clone 34, and clone 48, also named C7(S), C34 and C48 (**Fig.17**).

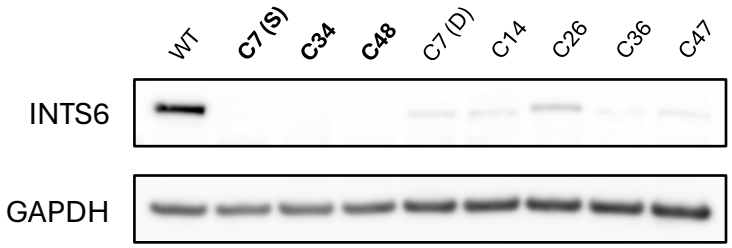


Fig. 17 | INTS6 Knockout in DU145 cells. This Western blot demonstrates the successful knockout of INTS6 in three clones derived from DU145 cells: clone 7(S), clone 34, and clone 48. GAPDH was utilized as the loading control.

Subsequently, a new Western blot was conducted on these cells to ascertain whether any alterations in the levels of certain major INTS6-interacting proteins within the Integrator complex and PP2A phosphatase complex could be observed (**Fig. 18**). Specifically, we deemed it appropriate to test: INTS1, because it is the core subunit of the central backbone of Integrator; INTS3, because it is directly associated with INTS6 (SOSS complex); INTS5 and PPP2R1A, because they are components of the phosphatase module (Int-PP2A) and interact directly with INTS6; INTS10, to control the status of the enhancer module of Integrator. The results of this experiment demonstrate that the levels of most of the proteins examined do not exhibit significant alterations in comparison to those of the wild-type (WT) cells or even among individual clones, with a few exceptions: specifically, the bands

corresponding to INTS1 appear to be less intense in the clones than in the WT cell, and regarding INTS10, slight differences are observable between clone 34 (which is more similar to WT) and clones 7(S) and 48, whose bands appear to be less intense; additionally, clone 34 is the only one that shows a significantly less strong band for INTS5 compared to the other clones.

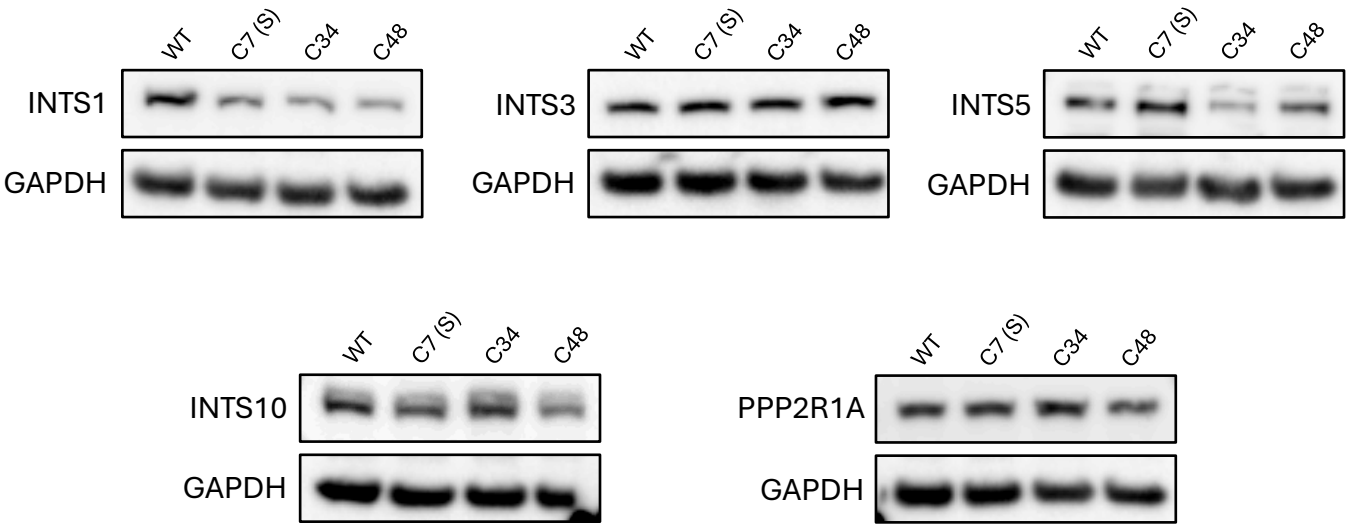


Fig. 18 | Main INTS6 interactors are not severely impacted by its Knockout. This Western blot demonstrates that the knockout of INTS6 does not markedly influence its primary interactors, although there are some minor exceptions (INTS1, and to a lesser extent, INTS10 and INTS5, but only in certain clones).

3.2 Quantification of INTS6 KO clones' growth rate

To test the hypothesis that the loss of INTS6, and thus of its tumor-suppression function, would lead to an increase in replication rate and tumor aggressiveness, we created a growth curve using a MTT assay (**Fig. 19**). To this end, a series of decreasing concentrations of DU145 cells, comprising both wild-type (WT) and INTS6 KO variants, were plated and their growth was measured over a period of five days. The results indicate that clone 48 exhibits a significantly faster growth rate than the wild-type (WT) cells, while clone 7(S) and clone 34 demonstrate a significantly slower growth rate (**Fig. 19a-b**). However, an analysis of the difference between the absorbances ($\Delta_{absorbance}$) measured on Day 1 and those measured at the conclusion of the experiment (Day 3) reveals that it is, in fact, the WT cells that exhibited the greatest growth, demonstrating an increase in absorbance that exceeded that of all other clones, including clone 48 (**Fig. 19c-d**). This trend is consistent across all cell concentrations, with the exception of the most dilute sample, which demonstrates that clone 48 exhibits a slight increase in growth rate compared to the WT cells in these conditions. Therefore, the growth curve provides evidence that partially challenges the initial hypothesis.

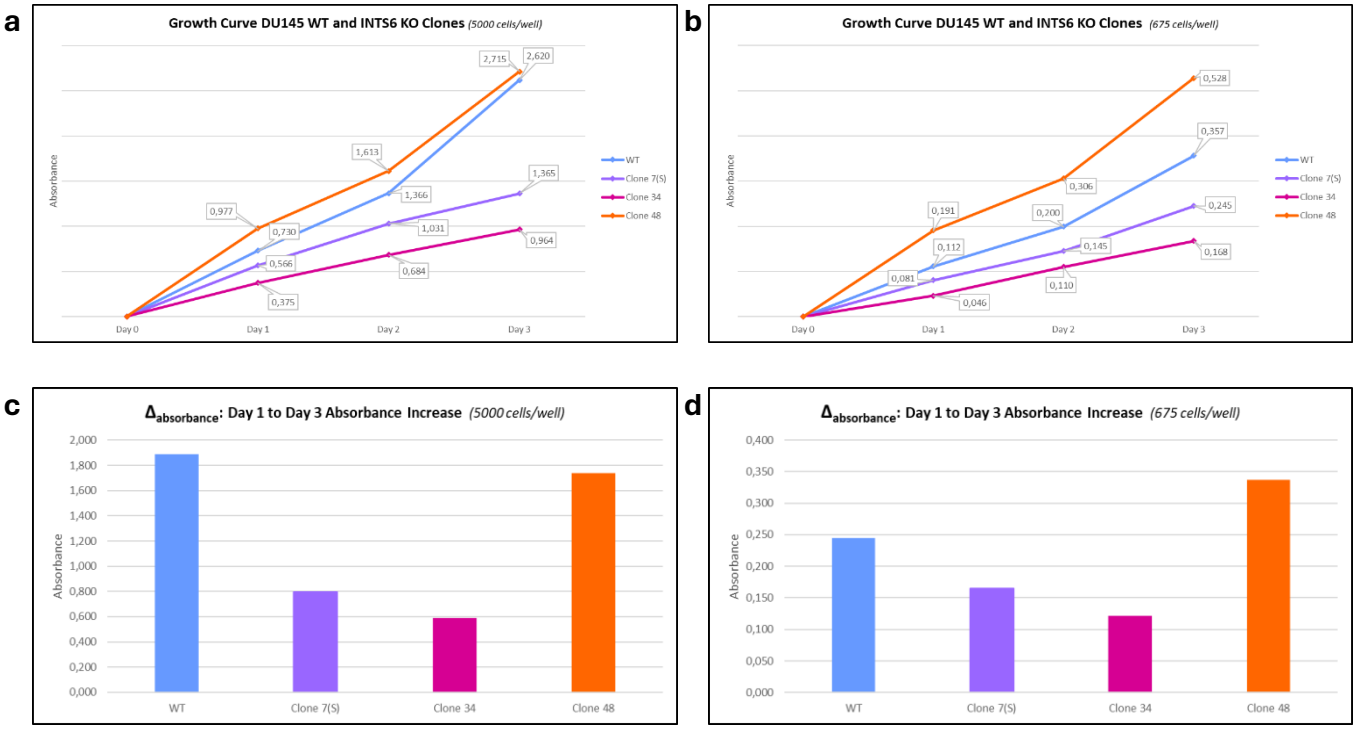


Fig. 19 | Growth curves reveal slower replication rate for INTS6 KO clones. The growth curves **(a)** and **(b)**, which were created starting from different plating dilutions, demonstrate that clone 48 is the sole INTS6 KO line that exhibits a growth rate that is faster than that of the WT cells. In contrast, the graph **(c)** challenges this hypothesis, demonstrating that the greatest increase in absorbance ($\Delta_{absorbance} = Abs_{day3} - Abs_{day1}$) during the measurement period was observed in WT cells, with the exception of the condition with the lowest plating dilution **(d)**.

3.3 Colony Formation Assay confirms differences in growth speed

In order to further elucidate the discrepancies between WT and INTS6 KO cells, a Colony Formation Assay (CFA) with staining via Crystal Violet was conducted. As illustrated in **Fig.21**, the outcome of the MTT assay is corroborated by this CFA, which demonstrates that clones 7(S) and 34 exhibit a significantly slower growth rate compared to the WT cells. Conversely, clone 48 displays a notable capacity for faster growth, as evidenced by its superior colony formation capacity relative to the other clones. Colonies count indicate that clone 48 is capable of generating a greater number of colonies than the WT cell line (**Fig. 20a**). However, the data is not sufficiently reliable on its own due to the difficulty in accurately distinguishing individual colonies when certain confluence thresholds are reached, particularly if they begin to fuse with adjacent colonies. For this reason, we also measured the surface area of the plates actually occupied by the cells. In this case, a marked superiority of the DU145 WT cells is clearly observed (**Fig. 20b**). In conclusion, the results of both the CFA and the MTT assay indicate that INTS6 KO cells exhibit a general reduction in growth rate compared to WT cells, which is contrary to our initial hypothesis.

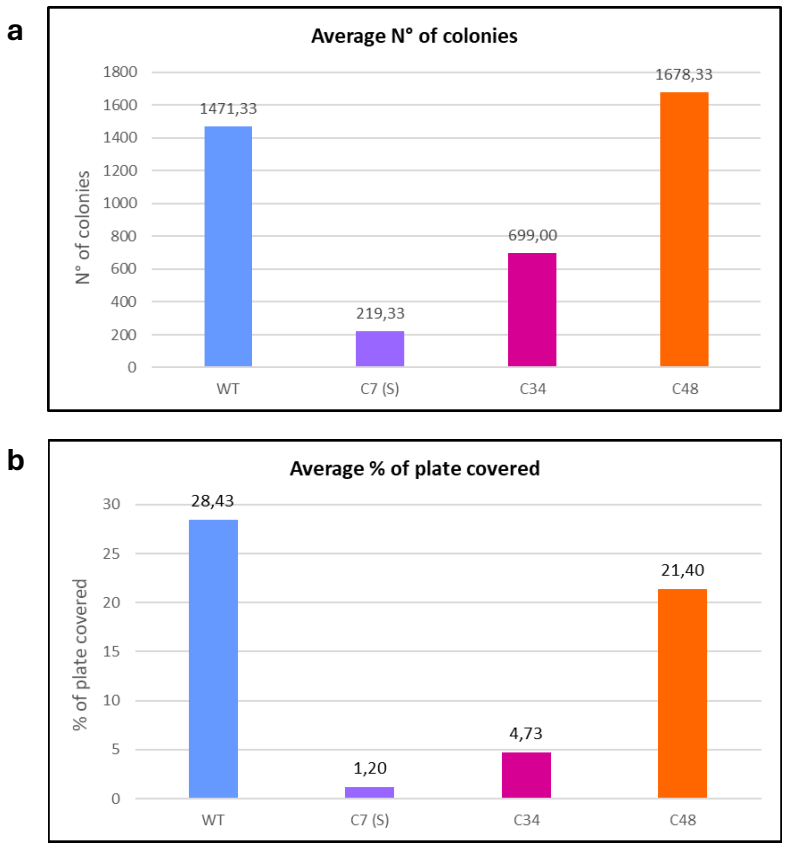
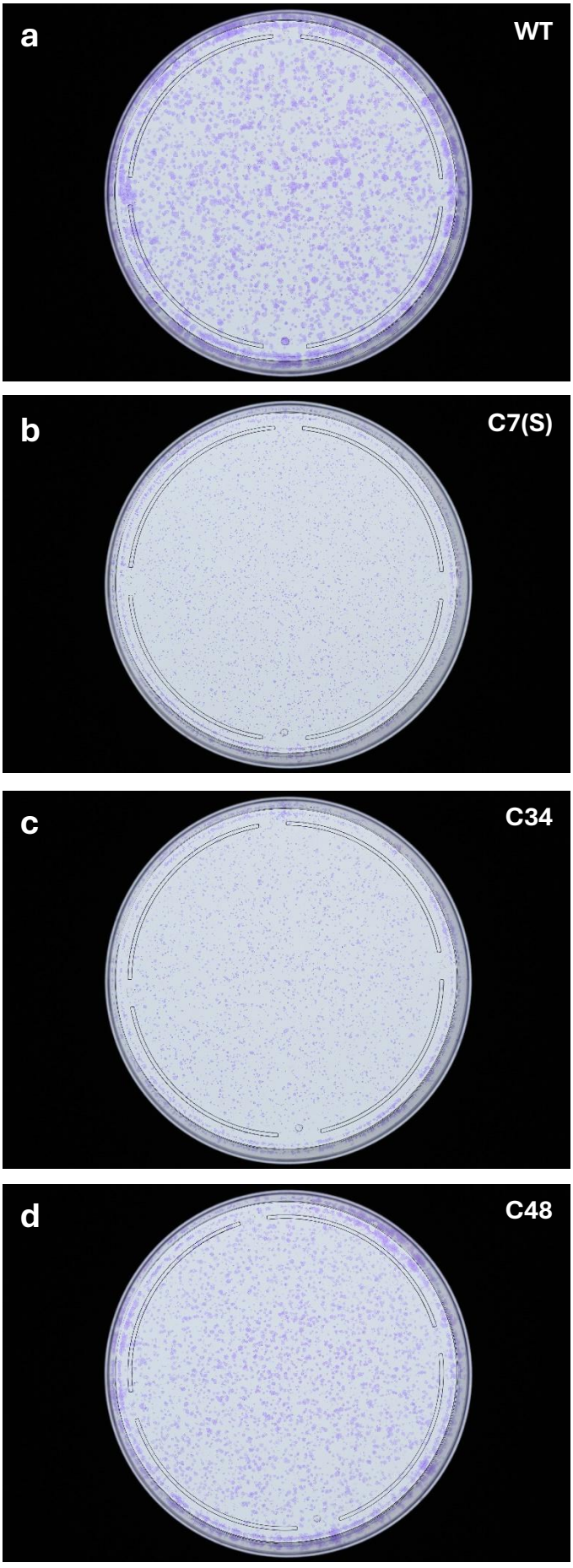


Fig. 20 | Colony Formation Assay data confirms slower growth rate for INTS6 KO clones. (a) Bar plot illustrating the average number of colonies counted per 10 cm plate. **(b)** Bar plot showing the average percentage of plate surface area occupied by colonies. Due to the inherent challenges associated with colony counting, this remains the most reliable methodology for evaluating cell growth.

Fig. 21 | Colony Formation Assay pictures show differences between WT and KO cells. Pictures of WT cells **(a)** and clone 7(S) **(b)**, 34 **(c)**, 48 **(d)** allow to visualize the disparate growth rates exhibited by each line. While the WT and C48 cells are unquestionably the fastest, with the former exhibiting colonies that are significantly larger than the latter, it is noteworthy that a considerable number of stained cells are present in both C7(S) and C34. The reason for their low colony and plate coverage numbers is that the vast majority of these cells did not form colonies, resulting in their exclusion from the analysis by the software that was used for counting them.



3.4 DU145 cells xenograft in mice

Given the *in vitro* observation that DU145 INTS6 KO cells grow at a significantly slower rate than WT cells, we sought to determine whether this growth differential also manifests *in vivo* or whether the cells require a physiological environment to demonstrate their increased tumorigenicity. To this end, a mouse xenograft experiment was designed and conducted by the Asangani Lab at the University of Pennsylvania. For reasons of cost and implementation, only three cell lines were selected for analysis: the WT line and clones 7(S) and 34. The latter were chosen over clone 48 because they exhibited greater similarity to each other and were therefore expected to yield comparable results.

A total of 12 mice were utilized in the study, with four mice per cell line. Two xenografts were performed in each mouse, one on the right flank and one on the left. This resulted in a total of eight tumors per cell line (24 tumors in total). Following the sacrifice of the animals and the subsequent extraction of the tumors, six samples were processed for each cell line to extract individual cells for genomics studies. The remaining two samples were retained by the Asangani Lab for future immunohistochemistry studies. Of the samples designated for genomics, three from each line (specifically, all those injected into the right flank of the animals) were utilized for total RNA-seq.

The entire study spanned 94 days, with tumor volume measurements taken every four days. The growth data from these measurements (**Fig. 22**) clearly demonstrate that, as observed *in vitro*, *in vivo* INTS6 KO cell lines exhibit a markedly slower growth rate than WT cell lines, although there is some variation between clone 7(S) and clone 34. In particular, clone 34 exhibits a comparable growth profile across all xenografts, with an average delay of approximately two weeks compared to the WT group in reaching equivalent volumes. It is also noteworthy that the C34 mice were sacrificed at day 66 (and not later) due to the emergence of signs of necrosis and infection in the lower abdomen. In contrast, the C7(S) mice exhibited an extremely low growth rate, with the tumor masses barely reaching a size that made them removable by the time of sacrifice (day 94).

In general, the results of this experiment corroborate those obtained from the MTT assay and the CFA, demonstrating that the clones exhibit a slower rate of cell replication than WT cells and that clone 7(S) displays an even more pronounced lag in growth compared to clone 34. It is also noteworthy that two WT tumors exhibited a markedly slower growth rate than the others in the same line. This discrepancy is likely attributed to an error that occurred during the transplantation process or to inherent mouse-to-mouse variability.

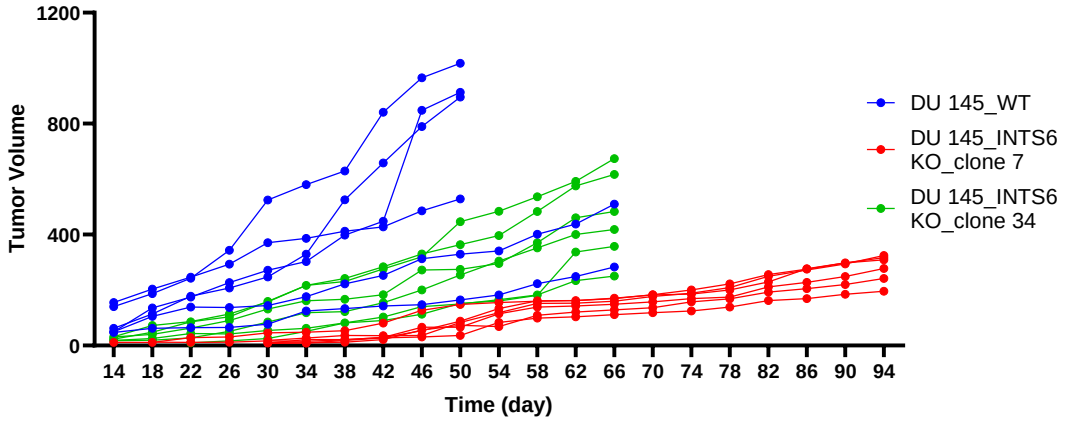


Fig. 22 | Xenograft experiment confirms previous observations regarding KO clones' growth speed. The graph illustrates the growth of tumors within the mice. The volume of the tumors is clearly indicative of a markedly slower growth rate for the C7(S) and C34 cells, particularly for C7(S), in comparison to their WT counterparts. It is noteworthy that two WT tumors exhibited a significantly slower growth rate than the other replicates. In general, clone 34 exhibits a delay of approximately one to two weeks in reaching a given volume compared to WT lines, while for clone C7(S), this delay extends to nearly two months.

3.5 QuantSeq 3' mRNA-seq on xenograft samples

As previously indicated, a total of nine samples (three per cell line) derived from xenografts in mice were utilized for a QuantSeq 3' mRNA-seq. Subsequent to the preliminary analyses, the WT sample No. 3 (corresponding to the tumor that exhibited a markedly slower growth rate than the others) and the C34 sample No. 3 were excluded because two Principal Component Analyses (PCA) conducted in succession on the group unequivocally classified them as outlier data (**Fig. 23**).

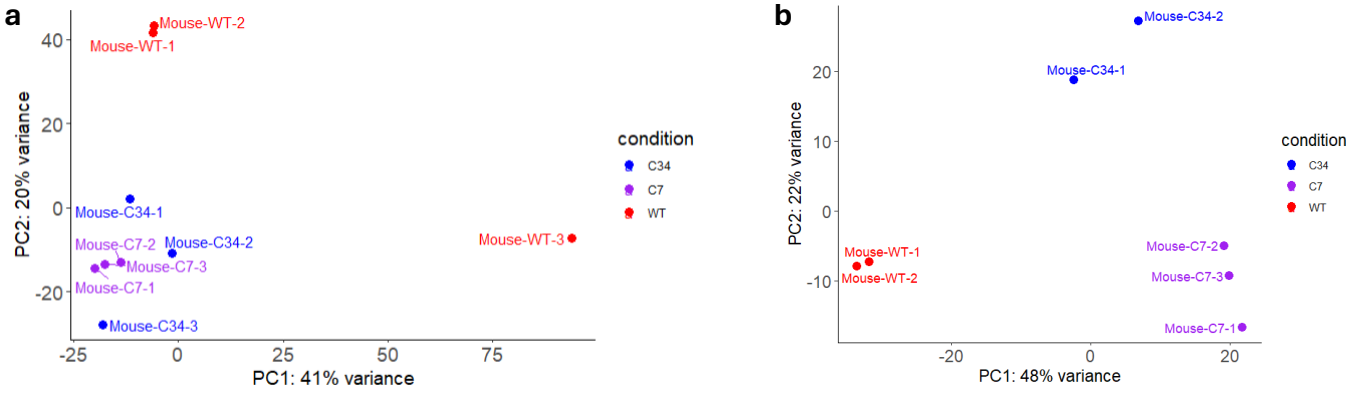


Fig. 23 | Principal Component Analysis identifies outlier data. A comparison of the initial Principal Component Analysis (PCA) of the entire data set **(a)** with the final PCA after the removal of the WT-3 and C34-3 samples **(b)** is presented here. The distribution and clustering of the samples appear significantly improved, which demonstrates the importance of identifying and eliminating outlier data.

A comparison of the differential gene expression between INTS6 KO clones and WT cells immediately reveals significant differences in gene expression, with more than 1600 up- and downregulated genes in clone 7(S) and approximately 1000 in clone 34 (**Fig. 24a-b**). The evident disparity in the extent of gene dysregulation as evidenced by these analyses could thus account for the observed discrepancies between the two clones.

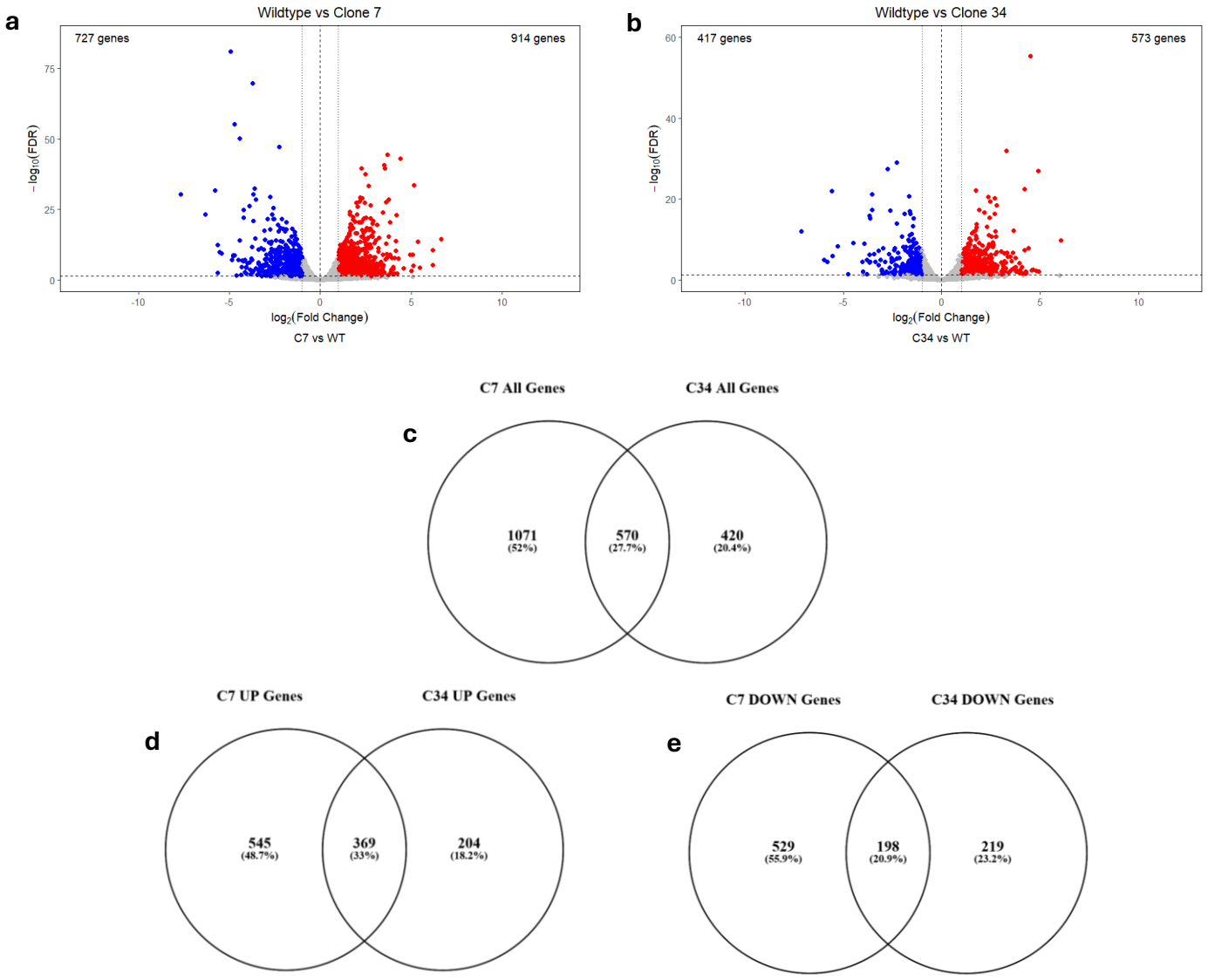


Fig. 24 | Differential gene expression analysis reveals significant differences between clones. Volcano plots illustrate the differential gene expression of clones 7(S) **(a)** and 34 **(b)** in comparison with WT cells. Venn diagrams, in contrast, illustrate the distribution of identified genes as a whole **(c)** or by separated into upregulated **(d)** and downregulated **(e)** categories.

It is also noteworthy that the genes identified as dysregulated in the two clones do not overlap entirely. In fact, when considering the sum of up- and downregulated genes, only 27.7% of the genes (i.e., corresponding to a total of 570 genes, representing 34.7% of the genes detected in clone 7(S) and 57.6% of those dysregulated in clone 34), were found to be in common (**Fig. 24c**). If we consider only the upregulated genes in common (**Fig. 24d**), this percentage is 33% (equivalent to 369 genes, representing 40.4% of the upregulated genes in clone 7(S) and 64.4% of those in clone 34). With regard to the downregulated genes (**Fig. 24e**), it is observed that only 20.9% of the genes are shared between the two clones (corresponding to 198 genes, representing 27.2% of the genes in clone 7(S) and 47.5% of those in clone 34). The discrepancy in gene expression between clone 7(S) and clone 34 is not merely quantitative, attributable to a disparity in the number of genes involved: it is also qualitative, as evidenced by the significant divergence in the dysregulated genes between the two clones.

Subsequently, a Gene Ontology Enrichment Analysis was conducted on the data derived from the 3' mRNA-seq, which revealed that a significant proportion of the upregulated genes belong to the same pathways in both clone 7(S) and clone 34 (**Fig. 25a-c**). Notably, the first three most significantly impacted pathways are identical between the two clones and pertain to developmental processes in general, as well as those specific to tissues and anatomical structures. However, some differences between the two clones are also evident. In particular, clone 34 displays a greater overall upregulation of cellular processes related to extracellular mechanisms, adhesion, and cell migration. In contrast, clone 7(S) appears to be significantly perturbed in processes related to cell differentiation and the development of multicellular structures. In general, these processes appear to be oriented toward faster cell growth and replication than in WT cells. However, our experimental observations provide evidence that contradicts this hypothesis.

A detailed examination of the downregulated genes and the pathways associated with them reveals that, in both clone 7(S) and clone 34, cellular signaling, including transduction, the ability to receive and respond to stimuli, and the capacity to regulate biological processes in a positive or negative manner, are the most significantly impacted (**Fig. 25b-d**). In clone 7(S), numerous genes associated with the same pathways that exhibited upregulation in the previously presented graphs are also significantly downregulated. This suggests that different components of the same pathways may undergo both upregulation and downregulation, which could explain why these cells exhibit significantly slower growth than those in clone 34: the growth potential may be blocked by the downregulation of other essential components involved in these processes.

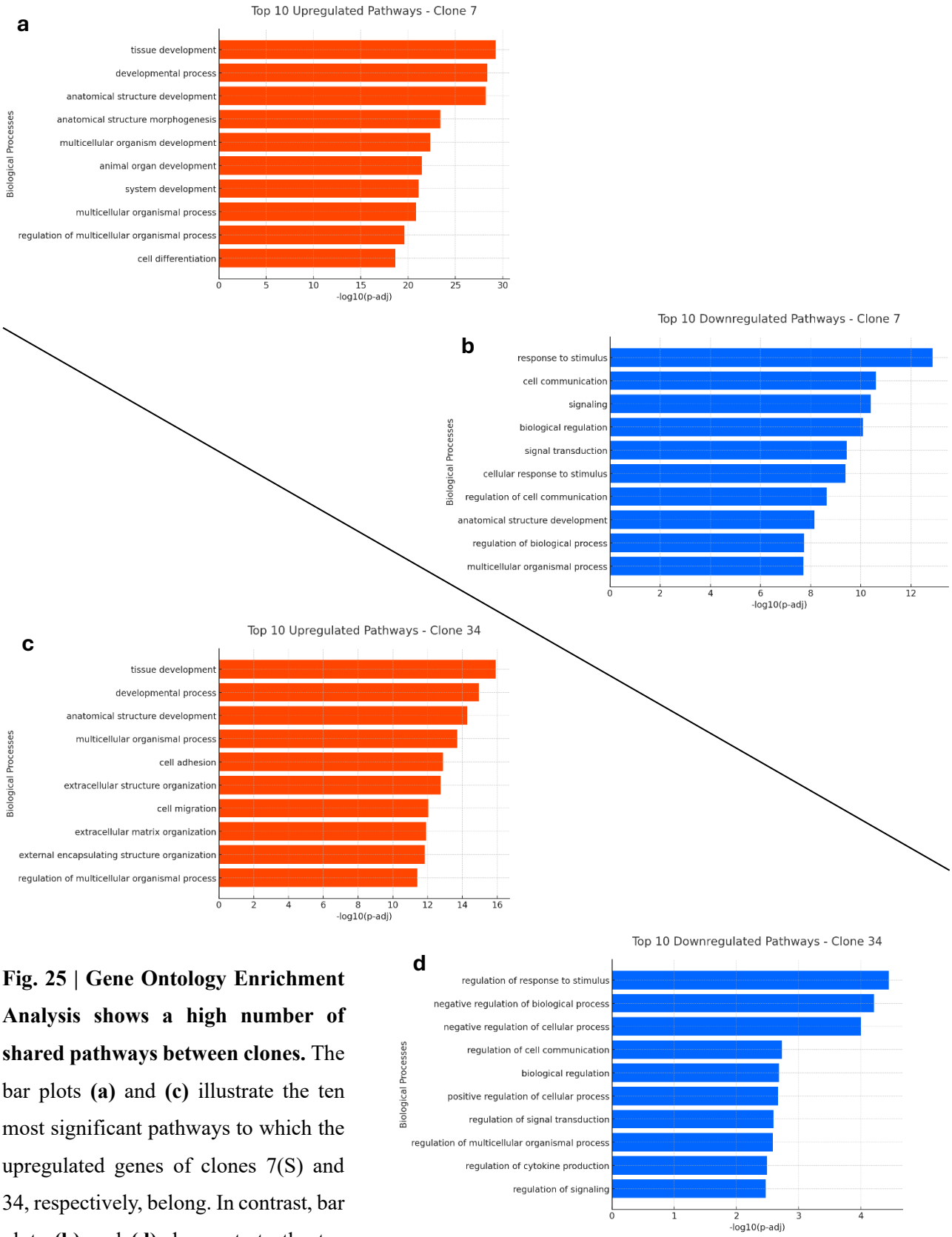


Fig. 25 | Gene Ontology Enrichment Analysis shows a high number of shared pathways between clones. The bar plots **(a)** and **(c)** illustrate the ten most significant pathways to which the upregulated genes of clones 7(S) and 34, respectively, belong. In contrast, bar plots **(b)** and **(d)** demonstrate the ten pathways most significantly impacted by the downregulation of the genes in clones 7(S) and 34, respectively.

In general, both clones appear to have difficulty in properly processing stimuli or receiving and sending signals, which are processes that underlie the cell growth: therefore, the deletion of INTS6 may have caused an increase in the ability of cells to replicate and grow, as predicted by our hypotheses. However, these changes were counterbalanced by the loss of the ability to develop multicellular clusters or colonies, which is consistent with the observations made in the Colony Formation Assay, where almost all C7(S) and C34 cells did not form colonies.

The validity of this hypothesis is further supported by the data obtained from a GO Enrichment Analysis performed on the overlapping genes between the two clones (**Fig.26**). The result corroborates the hypothesis that the most upregulated pathways are related to organ and tissue growth, as well as to developmental processes in general (**Fig.26a**). A closer examination also reveals that among the most significantly impacted pathways there are two associated with blood vessels. The downregulated pathways also align with previously observed trends, including a notable impact on cytokine production (**Fig.26b**). In conclusion, these discrepancies in gene expression may elucidate why the knockout of INTS6 exhibits diminished cell proliferation when compared to that of wild-type tumor cells, despite its predicted tumor-suppressive function.

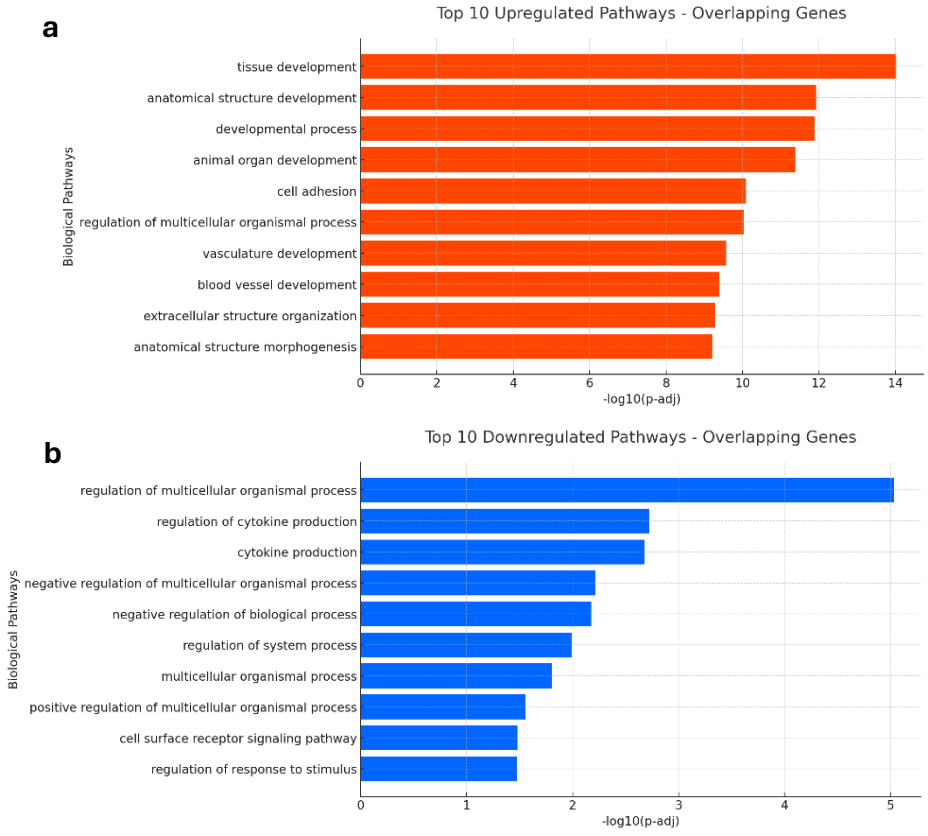


Fig. 26 | Gene Ontology Enrichment Analysis highlights clones' shared dysregulated pathways. The bar plots illustrating the up- **(a)** and downregulated **(b)** pathways to which the overlapping genes between clone 7(S) and clone 34 belong demonstrate contrasting patterns of up- and downregulation in the same cell growth-related pathways, as well as a notable impairment of signaling and stimulus response.

3.6 Characterisation of the canonical and the long isoforms of DDX26B

In parallel with the previously described experiments, our research also focused on DDX26B, the homologous gene of INTS6. This was done with the aim of understanding whether, following the deletion of INTS6 by CRISPR/Cas9 KO, this protein was able to replace it by taking over its functions entirely or at least partially. To achieve this, we initially constructed plasmids expressing tagged versions of INTS6 and DDX26B (utilizing FLAG, HA, or a combination of both, at either the C-terminus or N-terminus of the protein) and subsequently introduced them into HEK 293T cells, conducting immunoprecipitations (IPs) for the tags.

Our goal was to investigate the interactome of DDX26B, and to achieve this, the products of the IPs were analysed by mass spectrometry (IP-MS). However, an issue arose immediately: up to this point, in IPs targeting Integrator subunits, DDX26B was frequently co-immunoprecipitated; conversely, in IPs targeting the tagged version of DDX26B, almost no Integrator subunit was pulled down. **Fig. 27a** illustrates that, when comparing the IP-MS conducted on INTS6 with those on DDX26B, the majority of the Integrator subunits identified are exclusive to the INTS6 IPs. Furthermore, INTS5, INTS8, INTS12 and a few subunits of PP2A were also identified in the IPs for DDX26B, but only in trace amounts compared to the IPs for INTS6; INTS3 too was detected in the IPs targeting DDX26B, and it even appears to have been pulled down in bigger quantities by DDX26B than by INTS6. This result is not unexpected, however, as both proteins are known to bind INTS3 tightly.

In order to ascertain the cause of this anomaly, it was discovered that the two principal isoforms of DDX26B exhibit a single difference, namely the presence or the absence of exon 11 (**Fig. 28**), and apparently the absence of this exon is responsible for preventing DDX26B “short” (henceforth referred to simply as "DDX26B" or "DDX26B canonical") from binding to Integrator and a multitude of other proteins that are considered to be part of the DDX26B interactome. Apparently, in a fortuitous turn of events, the cDNA of DDX26B that had been selected for cloning and tagging belonged to the "short" version, so we designed new plasmids expressing this time the “long” version of DDX26B (henceforth referred to as “DDX26B**L**” or “DDX26B long”) tagged with HA and FLAG, and finally we repeated the IP-MS experiments. This time, the results showed that all the Integrator subunits and the rest of the interactome that had been expected to be observed were identified once more (**Fig. 27b**), thus confirming that the presence of the exon 11 is critical in enabling DDX26B to interact with a wide range of proteins. This conclusion was further supported by a comparison of the IP-MSs of DDX26B canonical and DDX26B long, which revealed that the latter has a significantly larger interactome and binds more effectively than DDX26B canonical to the majority of proteins they have in common (**Fig. 27c**).

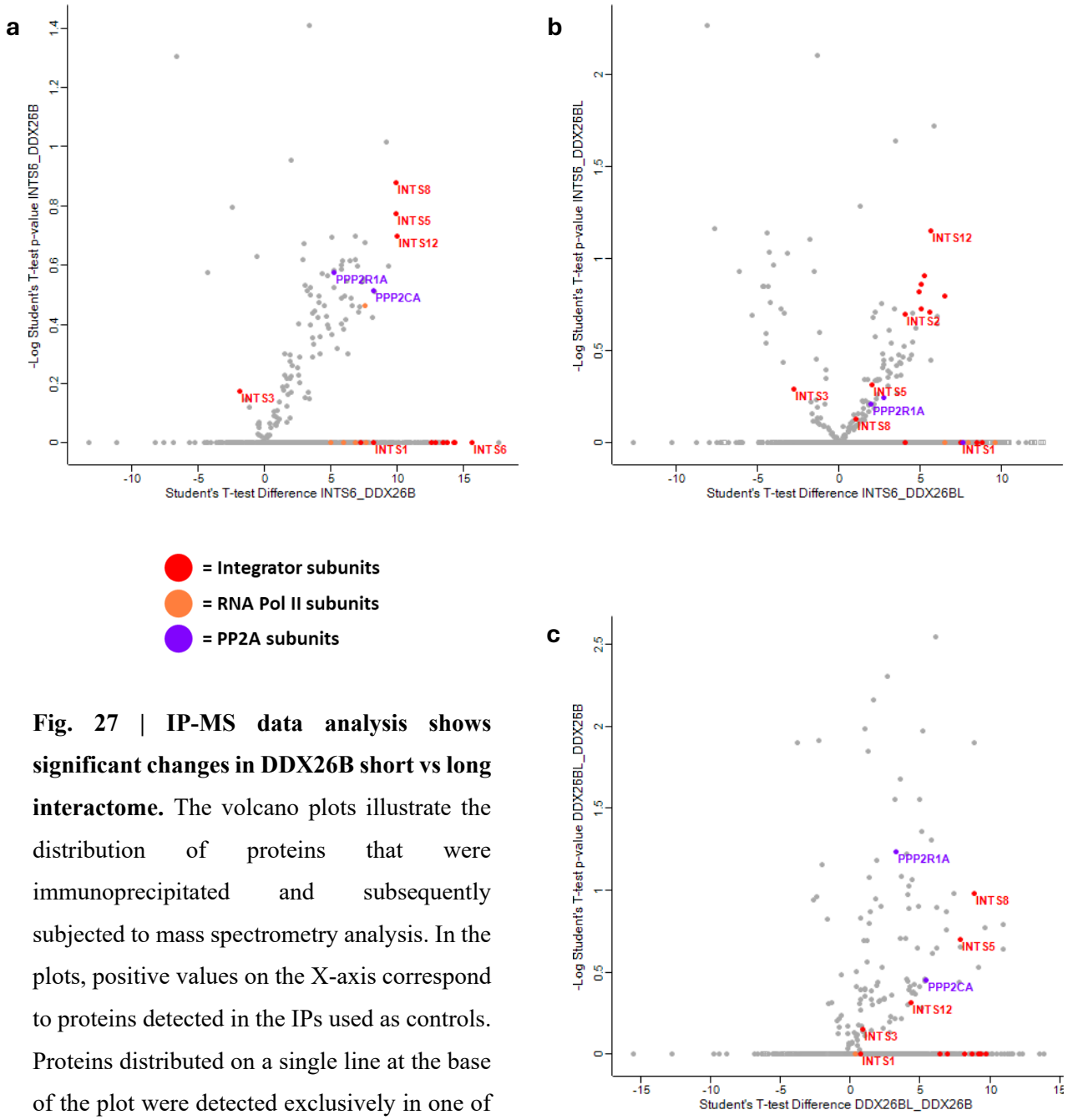


Fig. 27 | IP-MS data analysis shows significant changes in DDX26B short vs long interactome. The volcano plots illustrate the distribution of proteins that were immunoprecipitated and subsequently subjected to mass spectrometry analysis. In the plots, positive values on the X-axis correspond to proteins detected in the IPs used as controls. Proteins distributed on a single line at the base of the plot were detected exclusively in one of the two conditions being compared, while those distributed in the body of the plot were detected in both IPs but in different amounts between the two samples. **(a)** Comparison between INTS6 (control) and canonical DDX26B. The IPs for DDX26B immunoprecipitated only INTS3/5/8/12 and two subunits of PP2A. **(b)** Comparison between INTS6 (control) and DDX26B long reveals the detection of numerous additional Integrator subunits in the IPs for DDX26B. **(c)** Comparison of the canonical and long (used as a control) forms of DDX26B. The long isoform of DDX26B has a considerably larger interactome and is markedly more efficient than the canonical isoform in binding most of the proteins they have in common. It is noteworthy that only the long isoform of DDX26B interacts with other Integrator subunits in addition to INTS3/5/8/12.

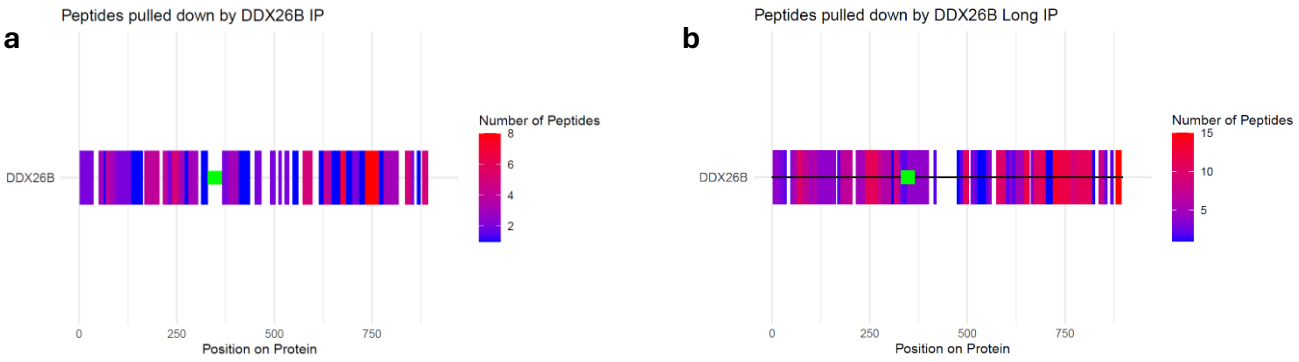


Fig. 28 | IP-MS coverage of DDX26B canonical vs long isoform. In these graphs, the colored segments correspond to the peptide fragments of DDX26B that were detected by the mass spectrometer. The fragments have been aligned to the sequence of DDX26B long (898aa) and the area corresponding to exon 11, spanning 36aa (position 330-366), is highlighted in green. It is evident that the canonical isoform of DDX26B **(a)** lacks fragments corresponding to this region, as it is devoid of the relevant exon. Conversely, the long isoform of DDX26B **(b)** displays coverage of this area.

3.7 Development of an exon skipping/inclusion strategy

Following the discovery of the impact that exon 11 of DDX26B has on its ability to interact with other proteins, a system was created to artificially induce the inclusion or skipping of this exon during DDX26B's transcript splicing events. The goal of this system was to create efficient models to study this exon in more detail.

To this end, we followed the guidance of the Guccione Lab (Icahn School of Medicine at Mount Sinai, New York) to design three Splice-Switching antisense Oligonucleotides (SSOs): these are synthetic molecules comprised of modified nucleotides (or nucleotide analogues) that bind to a complementary sequence and are designed to base-pair and create a steric block to the binding of splicing factors to the pre-mRNA. Thus, the base pairing of SSOs to target RNA results in the alteration of splice site recognition by the spliceosome, which in turn affects the splicing of the targeted transcript (**Fig. 29**). It is crucial to note that the nucleotides of an SSO must undergo chemical modification to prevent the recruitment of RNase H, which would otherwise result in the degradation of the pre-mRNA-SSO complex. It is noteworthy that the SSOs designed by us possess a phosphorothioate (PS) backbone modification, which serves to enhance their stability and resistance to nuclease; however, to prevent the recruitment of RNase H, it was also necessary to introduce 2′-O-methyl (2′-OMe) alterations at each 2′ sugar position.

Two of our SSOs induce exon inclusion or skipping of exon 11 of DDX26B, while the remaining one serves as a control and allows the exon skipping of exon 11 of INTS6. The system is currently undergoing testing in DU145 WT and INTS6 KO cells (all clones, hence C7(S), 34 and 48), and the results of a QuantSeq 3' mRNA-seq are awaited to ascertain the efficacy of the strategy.

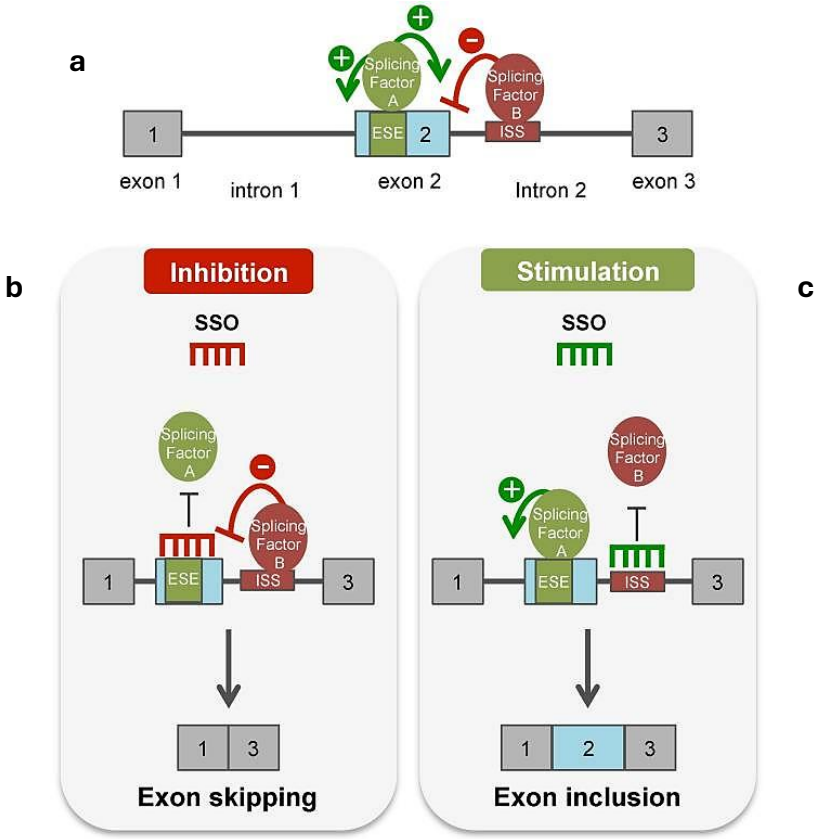


Fig. 29 | Mechanism of SSO-induced exon skipping/inclusion. (a) Diagram of a pre-mRNA transcript with exons depicted as grey boxes and introns as lines. The figure depicts the binding of an intronic splicing silencer (ISS, red) and an exonic splicing enhancer (ESE, green) by a trans-acting inhibitory splicing factor protein (red oval) or a stimulatory splicing factor (green oval). These proteins function to either block or promote splicing at splice sites bordering the surrounding exons. **(b)** An SSO that base-pairs to a splicing enhancer sequence creates a steric block to the binding of the stimulatory splicing factor to its cognate enhancer binding site. This block thus disrupts splicing and results in exon skipping. **(c)** Conversely, an SSO that base-pairs to a splicing silencer sequence element blocks splicing silencer activity by preventing binding of a negatively acting splicing factor. The disruption of the binding of splicing inhibitory proteins to their cognate binding sequence activates splicing at the splice site that is negatively regulated by the silencer element, resulting in exon inclusion.

DISCUSSION AND CONCLUSIONS

Recurring mutation of PP2A impairs neurogenesis

This study provides a comprehensive investigation into the biological effects of the R182W mutation in PPP2R1A, the scaffold subunit of protein phosphatase 2A (PP2A), within the context of neurodevelopmental disorders. Our work utilized an *in vitro* model involving Induced Pluripotent Stem Cells (iPSCs) carrying the R182W mutation, both in heterozygous (5A3) and homozygous (8B) forms, to explore the mutation's influence on neural differentiation and gene expression. Our findings reveal that the mutation compromises neurogenesis and impairs crucial developmental pathways, offering new insights into the molecular mechanisms underlying neurodevelopmental disorders.

Impaired differentiation and morphological disruptions

One of the primary observations in this study was the significant impact of the R182W mutation on iPSC differentiation into Neural Progenitor Cells (NPCs) and, ultimately, cortical neurons. Starting from the beginning, we immediately observed that the homozygous (8B) cells, although generated successfully, were unable to progress through the initial stages of differentiation, indicating the severe functional disruptions induced by the mutation; in contrast, the heterozygous (5A3) line exhibited delayed and incomplete differentiation, but were able to survive until further stages of the differentiation process.

Morphological analyses revealed clear differences in cell structure between wild-type and mutant cells, with 5A3 cells displaying abnormal cytoplasmic projections and an absence of the structural organization typical of the neurons. This morphological divergence suggests that the R182W mutation hinders cells from undergoing the fundamental changes required for effective neurogenesis, a finding that aligns with neurodevelopmental dysfunction in affected individuals.

Transcriptomic analysis and dysregulation of neural pathways

The mutation's impact on cellular function became further evident through transcriptomic analysis. Total RNA sequencing demonstrated extensive gene expression dysregulation in both 5A3 and 8B cells, with the effect intensifying as differentiation advanced to the NPC stage, indicating significant transcriptional shifts compared to wild-type cells. This dysregulation affects crucial pathways, particularly those related to neurogenesis, and this transcriptional imbalance may stem from disruptions in the regulatory roles of PP2A and its complex interactions, as the R182W mutation is

known to compromise PP2A's structural stability, likely affecting its function in numerous pathways integral to neurodevelopment.

The Gene Ontology Enrichment Analysis provided additional context to the transcriptomic findings. Dysregulated genes in mutant lines were indeed enriched in pathways related to cellular differentiation, cytoskeletal organization, and neural development, reinforcing the mutation's impact on critical neurogenic processes. Upregulation of genes associated with biosynthetic pathways in mutant cells also suggests a cellular response to stress or compensatory demand for protein synthesis, possibly triggered by the impaired cellular maturation. Conversely, downregulation in pathways central to neurogenesis and cellular structure may explain the failure of mutant iPSCs to differentiate effectively, supporting the notion that PPP2R1A mutations disrupt neurogenesis.

RNAPII binding and potential impact on transcriptional regulation

In exploring the mutation's influence on transcription, our ChIP-seq analysis showed that while the RNAPII binding at the Transcription Start Site (TSS) remains largely unaffected, there is a reduced RNAPII presence along the gene body and at the Transcription End Site (TES) in mutant cells. This finding suggests that while RNAPII recruitment may occur normally, the elongation or termination phases could be impacted by the mutation. Given PP2A's role in dephosphorylating RNAPII, the R182W mutation might lead to stalled transcription or improper termination, possibly resulting in the partial transcriptional dysregulation observed. This potential elongation or termination delay provides a new angle for understanding how the R182W mutation may subtly affect transcription, in particular in the context of these intricated neurodevelopmental processes.

Altered interactions within the PP2A complex and implications for cellular function

Mass spectrometry analysis added a new layer to our understanding of the R182W mutation's impact by revealing alterations in the PP2A complex composition, especially in mutant cells. It should be recalled that within cells, the predominant form of the Integrator complex is the one in which the regulatory B subunit of PP2A is absent and in its place there is the interaction between PPP2R1A and the Integrator complex (Int-PP2A or INTAC) through binding to INTS6.

The data derived from mass spectrometry indicate that, in 5A3 cells, PPP2R1A appears to lose the ability to interact with Integrator, but not with some regulatory B subunits. This suggests that, due to the R182W mutation, PP2A is unable to bind Integrator, thereby reverting to the canonical form composed of the three subunits A (scaffold), B (regulatory), and C (catalytic) (**Fig.30**).

The ramifications of this shift towards the canonical PP2A complex in mutant cells are profound, as they are indicative of a disruption in essential PP2A interactions with the transcriptional machinery. This could hinder the PP2A complex's role in regulating RNAPII pausing and release, thereby impacting the gene expression patterns necessary for proper neuronal maturation, and could demonstrate the link between neurodevelopmental disorders and transcription providing a potential molecular basis for the neurodevelopmental deficiencies associated with PPP2R1A mutations.

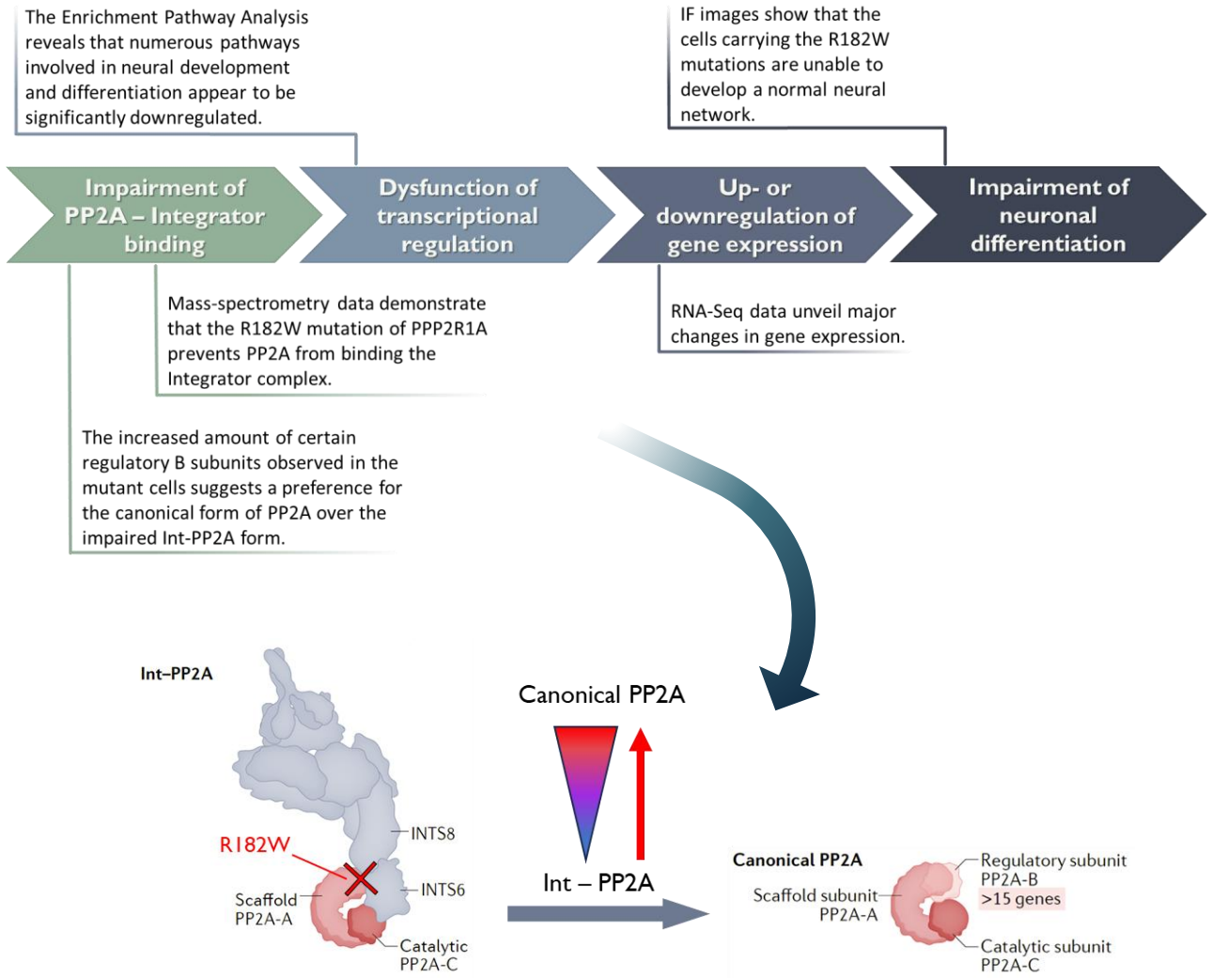


Fig. 30 | Proposed model for the effects of the R182W mutation on PP2A. As a result of the R182W mutation, PPP2R1A loses its ability to bind the Integrator complex and PP2A reverts to its canonical form consisting of the A, B, and C subunits. This change, and in particular the loss of PP2A-mediated transcriptional regulation, would explain the reasons behind the enormous gene dysregulation observed in our experiments on 5A3 cells and would demonstrate the link between neurodevelopmental disorders and transcription.

INTS6 role in prostate cancer progression

Our study aimed to explore the function of Integrator Subunit 6 (INTS6) in prostate cancer, focusing on its elimination in DU145 cells. Unlike previous findings that associate INTS6 with tumor suppression, our experiments revealed an unexpected outcome: DU145 cells with INTS6 Knockout showed slower growth compared to their wild-type counterparts. This observation suggests a more intricate role for INTS6, challenging the classical tumor-suppressor paradigm and revealing complex interactions with cellular pathways that govern proliferation and survival.

Re-evaluating INTS6's role in tumorigenesis

Our hypothesis was based on INTS6's established function as a tumor suppressor and its impact on cell cycle regulation and proliferation. Previous studies have demonstrated that INTS6 deletion or downregulation is often associated with enhanced proliferation in various cancers, including lung and esophageal carcinoma. Given this background, we anticipated that INTS6 knockout in DU145 cells would increase growth rates. However, contrary to this expectation, the INTS6 KO clones in our study demonstrated a significant reduction in replication rate and colony formation, both in vitro and in vivo. This reduction was particularly noticeable in clone 7(S), which exhibited a delay of nearly two months in reaching equivalent tumor volumes to WT cells in xenograft models.

The unexpected slower growth rate in INTS6 KO cells may be explained by disruptions in essential signaling pathways. Our RNA sequencing and gene ontology analyses revealed significant upregulation of pathways linked to cell differentiation, development, and structural formation in the KO clones. However, these growth-promoting pathways were counterbalanced by downregulation in parallel components critical to cell cycle progression and signaling reception. This dual impact could lead to an overall inhibition of growth, as cells attempt to proliferate but are restricted by concurrent pathway dysregulation. This suggests that INTS6 may facilitate prostate cancer cell growth under certain conditions, operating through more complex regulatory interactions than previously understood.

Pathway-specific effects and cellular implications

Both clones 7(S) and 34 showed upregulation in developmental pathways associated with tissue organization and growth; however, they simultaneously displayed significant downregulation in signal transduction and response mechanisms, particularly those related to cytokine activity and receptor signaling. This implies a paradoxical state where cells may gain proliferative signals but are unable to fully translate them into effective cell cycle progression due to downregulated feedback or

signaling mechanisms. The slower growth observed in INTS6 KO clones could result from this imbalance, as downregulated components of the growth pathways may act as bottlenecks, preventing the cells from effectively transitioning through critical cell cycle stages.

In addition, our colony formation assays illustrated a unique phenomenon where INTS6 KO cells showed impaired capacity for forming multicellular clusters, a trait essential for aggressive proliferation and metastasis. This reduction in colony-forming ability, corroborated by the in vivo data from the xenograft models, further highlights that the removal of INTS6 affects the DU145 cell's structural and signaling integrity, which may impair its capacity to grow as an unregulated tumor.

Functional potential of DDX26B as a partial analog

Parallel to our investigations of INTS6, we explored the role of its homologous gene, DDX26B, to determine if it could compensate for the loss of INTS6 in DU145 cells. The IP-MS and interactome analyses revealed that DDX26B, especially its long isoform, shared considerable overlap with INTS6 in terms of its interactions within the Integrator complex. The long isoform, due to the presence of the exon 11, displayed a more extensive interactome, including binding with multiple Integrator subunits and PP2A phosphatase components, suggesting it may indeed assume some of INTS6's roles in the cellular context.

However, despite this overlap, DDX26B could not fully replicate INTS6's regulatory control in DU145 cells, as evidenced by the growth phenotype observed in the KO clones. Our findings show that while DDX26B possesses a similar interactome, it lacks the functional impact necessary to maintain the same growth-promoting effects of INTS6, particularly within the context of signaling pathways that influence the cell cycle and multicellular cluster formation. These insights into DDX26B's limitations may indicate that it cannot fully substitute for INTS6, although it may offer partial compensatory effects under certain cellular conditions.

Finally, from a more practical point of view, the duality of the effects of INTS6 loss in prostate cancer suggests that it would be extremely difficult to use this protein as a target for potential oncological therapies. Therefore, an important alternative could come from studying INTS6's interactome or DDX26B, its potential replacement: the characterisation of these elements could provide targets to act on, in particular proteins with a more consistent behaviour which are suitable for use as therapeutic targets.

Future directions

Future developments for the projects discussed here would certainly benefit, starting from the beginning, from further investigation of the defects caused by the R182W mutation of PPP2R1A. This could be achieved by using iPSCs to create organoids, thereby enabling the study of the mutation in a three-dimensional neural context. This would provide a more accurate simulation of the physiological conditions in which neurons are found within the human body. While this condition is more complex to generate, it has the potential to provide more morphological and structural information, as well as revealing new differences in the creation of nuclei by WT and mutant cells.

From a molecular perspective, it is imperative to perform ChIP-seq experiments on NPCs, potentially using PPP2R1A itself (or functionally comparable targets, such as PP2CA) as a target, in addition to RNAPII. Additionally, conducting genomic analyses on early neurons derived from stem cells would be ideal. Given the cost-effectiveness ratio, the most efficient strategy would be to begin with a total RNA-seq.

Regarding the study of INTS6 and its homologue DDX26B, it would be necessary to perform ChIP-seq on DU145 WT and INTS6 KO cells using RNAPII as a target, but also INTS6 and DDX26B themselves, in particular to characterise in detail the consequences that the absence of INTS6 causes in the activity and binding of DDX26B and to what extent this protein is able to replace INTS6.

To further expand our understanding of the distinctions between the canonical and long form of DDX26B, it is essential to first refine the system to induce exon skipping and/or exon inclusion by using SSOs. Subsequent to this, we could undertake high-throughput genomic analyses to identify the binding and functional disparities between the two DDX26B isoforms.

It is also crucial to characterise the reasons behind the profound impact of exon 11 on DDX26B's binding capacity to the Integrator complex and PP2A phosphatase. This may be achieved through structural analysis of the Integrator and DDX26B binding sites and the exon 11 region.

Final remarks

The two studies presented in this thesis emphasize the pivotal function of the Integrator-PP2A complex in regulating transcription and sustaining cellular homeostasis. Disruption of this axis, whether through the R182W mutation in PPP2R1A or loss of INTS6, results in substantial alterations in gene expression that exert a profound impact on differentiation and proliferation.

In neural cells, the R182W mutation has been demonstrated to impair Integrator-PP2A interactions, which in turn has been linked to defective neural differentiation and neurodevelopmental disorders. In prostate cancer cells, the loss of INTS6 results in the disruption of INTAC function, which in turn leads to transcriptional dysregulation and a reduction in cell proliferation that is contrary to expectations.

These findings confirm the critical role played by the Integrator-PP2A complex as a key regulator of transcriptional homeostasis and, thus, of gene expression. Furthermore, our research emphasizes the importance of post-translational modifications, such as phosphorylation, in controlling RNAPII activity and transcriptional elongation: the balance between kinase and phosphatase activities, which is mediated by complexes like INTAC, is indeed crucial for fine-tuning RNAPII functions.

MATERIAL AND METHODS

Cell Lines

DU145 (ATCC HTB-81) cells were cultured at 37°C and 5% carbon dioxide (CO2) in Eagle's Minimum Essential Medium (MEM; Corning) supplemented with 10% fetal bovine serum (FBS; Peter MacCallum Cancer Centre). DU145 INTS6 KO clones were supplemented with 2 µg/mL of Puromycin for selection. Human Embryonic Kidney (HEK) 293T cells (ATCC CRL-11268; used also for generation of lentivirus) were cultured at 37°C and 5% CO_2 in Dulbecco's Modified Eagle Medium (DMEM: ThermoFisher Scientific) supplemented with 10% FBS, 100 U/mL penicillin; 100 µg/mL streptomycin and 2mM L-glutamine (Corning).

iPSC culture

iPSC-SV20 cells were provided by the University of Pennsylvania iPSC Core. The cells were cultured on Geltrex-coated plates, maintained in StemMACS™ iPS-Brew XF (Miltenyi Biotec). iPSCs were split to 1:6 or 1:10 every 4 to 5 days using ReleSR (STEMCELL Technologies) and fed with fresh medium every other day. 10 µM ROCK inhibitor (Y-27632) was added to the media for 24 h following splitting to prevent cell death.

Neural induction of iPSCs

The neural induction protocol is based on the manufacturer's instructions (Gibco: "*Induction of Neural Stem Cells from Human Pluripotent Stem Cells Using PSC Neural Induction Medium*") and summarized in **Fig.9**. To differentiate iPSCs into NPCs, we used Neural Induction Medium following manufacturer's instructions and started with high quality iPSCs with minimal differentiated colonies. Briefly, iPSCs were cultured to reached 70-80% confluency, incubated with pre-warmed Accutase (STEMCELL Technologies) at 37°C for 5 minutes and passed through a 40 µM cell strainer to generate single cell suspension in Neural Induction Medium (98% Gibco™ Neurobasal™ Media and 2% Gibco™ Neural Induction Supplement). 2.5 x 10^5 iPSCs per well were seeded in Geltrex-coated 6 well plates and incubated with 10 µM of ROCK inhibitor overnight. Cells were fed with fresh Neural Induction Medium on Day 2 (2.5 ml per well), Day 4 (2.5 ml per well) and Day 6 (5 ml per well). On Day 7 of neural induction, NPCs (P0) were harvested by incubating with pre-warmed Accutase for 8 minutes at 37 °C. Dissociated cells were resuspended with DPBS and passed through

a 100 μM cell stainer. After centrifuged at 300 × g for 4 minutes and washed once with DPBS, P0 NPCs were resuspended with pre-warmed complete Neural Expansion Medium (49% Neurobasal Media, 49% Gibco™ Advanced DMEM/F12 and 2% Neuronal Induction Supplement and plated on Geltrex-coated dishes at a density of 1 X 10^5 cells per cm^2. NPCs were fed with fresh Neuronal Expansion Media every other day for 6 - 7 days until NPCs reached 80-90% confluence. 10 μM ROCK inhibitor was added to the Neural Expansion Medium to treat the new plated NPCs overnight to prevent cell death. Dissociated NPCs were cryopreserved in Neural Expansion Medium with 10% dimethyl sulfoxide (DMSO). Expanded NPCs from passage 4 were used for characterization and neuronal differentiation.

To differentiate NPCs into cortical neurons, we followed the protocol from the article S. C. Zhang et al. (2001) with slight modification. Briefly, Accutase-dissociated NPCs were plated onto laminin (10 μg/ml) coated 6 well plates at a density of 5 × 10^4 cells per cm^2 in a neuronal differentiation medium, B27, GlutaMAX, nonessential animal acids, 20 ng/ml brain-derived neurotrophic factor (BDNF) and 20 ng/ml glial cell-derived neurotrophic factor (GDNF) for 21 days. Medium was changed every other day before harvesting.

Immunocytochemistry (ICC)

Immunofluorescence experiments were performed as described (Zucco et al., 2018). Briefly, cells were fixed with 4% formaldehyde for 15 min at room temperature, washed three times with PBS for 15 min and incubated with 1% goat serum in PBST containing 0.1% Triton X-100 for 30 min at room temperature. Cells were incubated with primary antibodies at 4°C overnight, washed three times with PBS for 10 min at room temperature and incubated with secondary antibodies for 1 hour at room temperature. After that, cells were incubated with 1 μg/ml of DAPI for 15 min and mounted in SlowFade™ Gold Antifade Mountant (ThermoFisher Scientific, Cat#S36938) and imaged using a Nikon 80i Upright Microscope.

Lentiviruses packaging

Lentiviruses were produced in HEK293T cells by co-transfection 8 μg of plasmid DNA with three packaging plasmids (2.5 μg of pRSV-REV, 5 μg of pMDLg/pRRE and 3 μg of pMD2.G per 10 cm cell culture dish) using calcium phosphate transfection (Chen & Okayama, 1987). Lentiviruses were harvested 48 and 72 hours after transfection and the cell debris pelleted by centrifugation at 2500 rpm

for 5 minutes. Then the supernatant was filtered and the viruses immediately used or stored at 4°C for 2-4 days.

Generating Knockout Clones with CRISPR/Cas9

DU145 cells were plated into a 6-well plate at 30-40% confluence. When the cells reached 60-70% confluence, 2.5-3 ml of medium with lentiviruses per well plus 2 µg/ml polybrene (Thermo Fisher, cat# TR1003G) was added to replace the old medium. 24 hours after induction, the virus medium was removed and replaced with fresh cell culture medium for another 24 hours. After that, cells were selected with 0.2 µg/ml of Puromycin (InvivoGen, cat# ant-pr-1) in fresh medium. Following selection with puromycin for at least 72 hours, cells were trypsinized and plated into 15 cm tissue culture dishes. Single cells were maintained in the basal medium supplemented with 0.5 ug/ml of puromycin for the next 2-3 weeks until cell colonies appeared. Individual microcolonies were moved to a 24-well plate and disrupted for clonal expansion. Clones were verified by Western Blot and, later, by ChIP-seq and RNA-seq performed on the most promising ones.

Western blot

Collected cells were washed twice with cold PBS and lysed in ChIP lysis buffer (150 mM NaCl, 1% Triton X-100, 0.7% SDS, 500 mM DTT, 10 mM Tris-HCl, 5 mM EDTA) supplemented with 1 µg/mL each of aprotinin, leupeptin and pepstatin for 30 min. Lysed cells were centrifuged for 10 min at 21300g at 4°C and the supernatant collected. Protein concentration was determined using the Pierce™ Bradford Protein Assay Kit (Thermo Fisher). 4X Bolt™ LDS Sample Buffer and 10X Bolt™ Sample Reducing Agent (Invitrogen) were added to protein extracts for 5 minutes at 95°C. Proteins were loaded into Bolt™ 4-12% Bis-Tris Plus gels (Invitrogen) and separated through gel electrophoresis (SDS-PAGE) in Bolt™ MES running buffer (Invitrogen). Separated proteins were transferred to Immunoblot PVDF membranes (Bio-Rad) in Tris-Glycine buffer (Bio-Rad). The membranes were incubated with 10% BSA in 1X TBST for 1h at room temperature. Primary incubation (suitable antibodies were diluted in 5% BSA in 1X TBST) occurred at room temperature for 2 hours, or at 4°C overnight. Membranes were washed twice with TBST for 10 minutes and then incubated with HRP-linked anti-mouse or anti-rabbit secondary antibody (Cell Signaling; 1:10000 in 5% BSA in 1X TBST) for 1 hour at room temperature. Proteins were detected using Clarity Western ECL substrate (Bio-Rad) and imaged with ImageQuant LAS 4000 (GE healthcare).

Growth curve by MTT Assay

To generate the growth curve, the Cell Proliferation Kit I (MTT) (Roche, 11465007001) was used according to the manufacturer's instructions. On Day 1 the cells were plated in quadruplicate into three 96-wells plates (Plate 1, 2 and 3) in different dilutions (5000, 2500, 1250 and 675 cells/well), 100 µl total volume of media per well. On Day 2, 10 µl of MTT labeling reagent per well were added to Plate 1 and then it was left into the incubator for 4 hours; then, 100 µl of Solubilization buffer per well were added and the plate was left overnight into the incubator. On Day 3, Plate 1 wells' spectrophotometrical absorbance was measured using a microplate (ELISA) reader. The wavelength to measure absorbance of the formazan product is between 550 and 600 nm according to the filters available for the ELISA reader used. Use a reference wavelength >650 nm. Also, on Day 3, repeat the Day 2 procedure but on Plate 2 (incubation with MTT labeling reagent and then with the Solubilization buffer). On Day 4 measure the Plate 2 absorbance and prepare the Plate 3 in the same manner. On Day 5, measure Plate 3 absorbance. The blank values for the absorbance was given by four wells filled with media but no cells.

Colony Formation Assay (CFA)

Cells were plated in 10 cm dishes at a concentration of 6000 cells/plate, then they were left growing for 7 days. On Day 7, the media was removed, the cells washed once with DPBS and then stained by adding directly into the plates 4 – 5 ml of 0.5% w/v Crystal Violet (Sigma Aldrich # C6158-50G) in 20% methanol. After 30 minutes, the Crystal Violet solution was discarded and the plates washed by gently submersion into a tank full of distilled water. After that, the plates were left open at room temperature, upside down and slightly inclined, to let them dry overnight. On the day after images were taken using a Nikon Z6II digital camera. The analysis (colonies count and area coverage) were performed using the NIS-Element software.

RNA extraction

RNA was extracted from cells using TRI reagent and purified with the Zymo Direct-zol RNA miniprep kit (Zymo Research, R2050) following manufacturer's instructions. Briefly, media was removed from the wells and cells were washed with PBS once. 0.5 - 1 ml of dissociating agent (Trypsin for DU145 or HEK 293T, ReleSR for iPSCs, Accutase for NPCs) was added to each well of a 6 well plate for 3 to 8 min. Cell suspension were resuspended in media or PBS, collected and

centrifuged at 300 g for 5 min. The pellets were then washed once with PBS and centrifugated again. After removing the PBS, 300 - 600 µl TRI reagent was added to each pellet and either frozen at -80°C or processed for RNA purification following the Zymo Directed protocol. RNA concentration was quantified by Nanodrop.

QuantSeq 3' mRNA-seq

Cells were lysed with TRI reagent and total RNA was purified using Zymo Research Direct-zol RNA mini-prep kit (R2050) following the manufacturer's manual, as previously described. Libraries were generated using the QuantSeq 3' - mRNA Seq Library Prep Kit for Illumina (Lexogen). 75 base-pair single-end reads were sequenced on the Illumina NextSeq 2000. Reads were aligned to hg19 human reference genome using STAR v2.5. FeatureCounts (Liao et al., 2014) was used for counting reads to the genes. Data were normalized using Voom and differential gene expression analysis was performed using DEseq2 in R (v1.38.3) unless otherwise noted. Data was visualized using ggplot2 (3.3.6). GO enrichment analysis was done using gprofiler2 package in R (v 0.2.1). Gene set enrichment analysis (GSEA) was 500 randomly selected genes from the select data set using the clusterProfiler package in R (v4.6.2).

Total RNA-Seq

Cells were lysated with TRI reagent and total RNA was purified using Zymo Research Direct-zol RNA mini-prep kit (R2050) following manufacturer's manual, as previously described. Libraries were generated using the Illumina Stranded Total RNA Prep with Ribo-Zero Plus (20040525). 40 base-pair single-end reads were sequenced on the Illumina NextSeq 2000 sequencer. Reads were aligned to hg19 human reference genome using STAR v2.5. FeatureCounts (Liao et al., 2014) was used for counting reads to the genes. Data were normalized using Voom and differential gene expression analysis was performed using DEseq2 in R (v1.38.3) unless otherwise noted. Data was visualized using ggplot2 (3.3.6). GO enrichment analysis was done using gprofiler2 package in R (v 0.2.1). Gene set enrichment analysis (GSEA) was 500 randomly selected genes from the select data set using the clusterProfiler package in R (v4.6.2)

Chromatin Immunoprecipitation Sequencing (ChIP-seq)

Cells were resuspended at a concentration of 10 million cells per 10 ml in fresh media at room temperature in 50 ml Falcon tubes. For each replicate, 15 - 25 million cells were harvested for cross-linking. 15 – 25 million cells were moved to a 15 ml Falcon tube in a total volume of 10 ml of media, then the tube was placed on a rocker for 5 min at room temperature with 1% of formaldehyde (Sigma; Cat#252549). To quench the cross-linking reaction, 560 µl of 2.5 M Glycine per 10 ml media was added and cells were incubated at room temperature with rotation for 5 min. After being washed twice with cold PBS and spun at 2500 rpm for 10 min at 4℃, the pellet was aspirated dry and finally frozen at -80 ℃. On the Day1 of the ChIP, cells were then resuspended in 900 - 950 µl of ChIP lysis buffer (150 mM NaCl, 1% TritonX-100, 5mM EDTA, 10mM TrisCl, 500 uM DTT, 0.4% SDS) and sonicated to an average length of 200-250 bp using a Covaris S220 Ultrasonicator. Fragmented chromatin was cleared at 21300 g for 10 min and diluted with SDS-free ChIP lysis buffer. For each immunoprecipitation, cleared fragmented chromatin was incubated with 15 µg of antibody, and Protein A or Protein G Dynabeads (Invitrogen) at 4 ℃ overnight. On Day 2, after the incubation, the beads were washed twice with each of the following buffers: Mixed Micelle Buffer (150 mM NaCl, 1% Triton X-100, 0.2% SDS, 20 mM Tris-HCl (pH 8.0), 5 mM EDTA, 65% sucrose), Buffer 500 (500mM NaCl, 1% Triton X-100, 0.1% Na-deoxycholate, 25mM HEPES, 10mM Tris-HCl (ph 8.0), 1mM EDTA), LiCl/detergent wash buffer (250 mM LiCl, 0.5% Na-deoxycholate, 0.5% NP-40, 10 mM Tris-HCl (pH 8.0), 1 mM EDTA), followed by a final wash with 1X TE. To elute samples, beads were then resuspended with 1X TE supplemented with 1% SDS and incubated at 65 ℃ for 10 min. After eluted twice, samples and the untreated input (5% of the total sheared chromatin) were incubated at 65 ℃ overnight to reverse cross-link. On Day 3, after reverse cross-linking, samples were treated with 0.5 mg/ml proteinase K at 65 ℃ for one hour and purified using the Zymo ChIP DNA Clean Concentrator kit (Zymo Research D5205) as the manufacturer's manual and quantified by QUBIT. On Day 4, Barcoded libraries were made using NEB Ultra II DNA Library Prep Kit for Illumina following manufacturer's instructions and quantified by QUBIT and Agilent TapeStation. Libraries were sequenced on Illumina NextSeq 2000 with 40 base pair paired-end reads. Sequences were aligned to human reference hg19. Samtools (1.9.0) was used to remove the PCR duplicates (rmdup) and the reads with a mapping quality score of less than 10 from the aligned reads. Bigwig files of the data generated with deeptools (v2.4.2, bamCoverage–binSize 10–normalizeTo1× 3137161264–extendReads 150–ignoreForNormalization chrX) and visualized on the WashU Epigenome Browser (https://epigenomegateway.wustl.edu/browser/) or the USCS Genome Browser (https://genome.ucsc.edu/). For normalization of the data, each number of the filtered reads was divided by the lowest number of the filtered reads in the same set of experiments, generating a

downsampling factor for each sample. Normalized BAM files were generated using samtools view -s with the above downsampling factors and further converted to normalized BAM files using bamCoverage–binSize 10–extendReads 150. Peaks were called by MACS2 (Dobbs et al., 2008).

Sample preparation for IP-MS

Cells were collected and centrifuged at 3000 RPM for 8 minutes at 4°C. The pellet was washed twice in cold PBS. Cell pellets were resuspended in 5 packed volumes (PCV) of Buffer A (10mM Tris pH 7.9, 1.5mM MgCl2, 10mM KCl, 0.5mM DTT, and 1mg/ml each of protease inhibitors aprotinin, leupeptin, and pepstatin). Resuspended cells were mixed with rotation for 10 minutes at 4°C. Cells were then resuspended and homogenized in 2 PCV of Buffer A after pelleting at 1200 rpm for 10 minutes at 4°C. Homogenized cells were pelleted and the supernatant was saved as the cytoplasmic fraction. Pellets from previous spin were resuspended in 1 PCV of Buffer C (20mM Tris pH 8.0, 1.5mM MgCl2, 0.42M NaCl, 25% glycerol, 0.2mM EDTA, 0.5mM DTT and 1mg/ml each of protease inhibitors aprotinin, leupeptin, and pepstatin) and homogenized. The resuspended extract was incubated for 30 minutes at 4°C while rotating. After one more spinning, supernatant was saved as the nuclear extract. All the saved extracts were dialyzed overnight in BC80 (20mM Tris pH 8.0, 80mM KCl, 0.2mM EDTA, 10% glycerol, 1mM B-mercaptoethanol, 0.2mM phenylmethylsulfonyl fluoride (PMSF)) at 4°C. Dialyzed samples were spun down at 1500 RPM for 20 minutes and stored at -80°C after being snap-frozen in liquid nitrogen.

Immunoprecipitation-mass spectrometry (IP-MS)

For each IP, 1.5 mg of nuclear extract, 6 μg of antibody, 70 μL of Dynabeads Protein A or G were mixed in co-IP buffer (20mM Tris pH 8.0, 100mM NaCl, 0.1% NP-40, 20mM Tris pH 8.0, 1.5mM MgCl2, 0.42M NaCl, 25% glycerol, 0.2mM EDTA, 0.5mM DTT and 1mg/ml each of protease inhibitors aprotinin, leupeptin, and pepstatin) to make a 500 μL volume of reaction. IPs were incubated at 4 °C for 2 hours with rotation and then were washed 3 times with 1 volume of Co-IP buffer and once with PBS + 0.05% NP-40. Proteins bound to the beads were eluted with 0,1 M glycine pH 3 (IgG elution Buffer) while shaking on the thermocycler at 1200 RPM at room temperature for 2 minutes. Tris-HCl pH 9.5 was added to eluate to neutralize pH. Eluates were prepared for Western Blot and run on a 12-well Novex WedgeWell 10% Tris-Glycine Gel (Invitrogen) with Tris-Glycine-SDS buffer (Bio-Rad), at 100V for 10-15 minutes. The gel was stained with Colloidal Blue staining kit (Invitrogen) overnight and processed at the proteomics facility at the Wistar Institute. The gel

lanes were excised and digested by trypsin and analyzed by LC-MS/MS on a Q-Exactive Plus mass spectrometer. The data were searched with full tryptic specificity against the UniProt human database using MaxQuant (false discovery rates for protein and peptide identifications were set at 1%). The iBAQ values present in the files generated by MaxQuant were normalized to the total levels of the protein that was pulled down. Then, the data analysis was performed using Perseus (iBAQ values were filtered based on contaminants and then transformed into log10; a two-sample t-test was used to test if samples were equal or not, and a table with –Log Student's T-test p-values and Student's T-test differences was obtained as final output and eventually transformed in graphs).

Antibodies	Manifacturer	Cat. #
Anti-rabbit IgG (HRP)	Cell Signaling	7074
Anti-mouse IgG (HRP)	Cell Signaling	7076
Anti-GAPDH rabbit polyclonal	Cell Signaling	2118
Anti-DICE1 (INTS6) clone H-6, mouse monoclonal	Santa Cruz	sc-56767
Anti-INTS1 mouse monoclonal	MilliporeSigma	MABS1984
Anti-INTS3 rabbit polyclonal	Proteintech	16620-1-AP
Anti-INTS5 rabbit polyclonal	Proteintech	14069-1-AP
Anti-INTS10 rabbit polyclonal	Proteintech	15271-1-AP
Anti-RNAPII, raised against the N-terminus of subunit B1, rabbit polyclonal	Gardini Lab, The Wistar Institute	(Barbieri et al, 2018)
Anti-FLAG clone M2, mouse monoclonal	MilliporeSigma	F1804
Anti-HA rabbit polyclonal	MilliporeSigma	H6908
Anti-PPP2R1A rabbit polyclonal	Proteintech	15882-1-AP
Anti-PP2CA (clone 1D6) mouse monoclonal	MilliporeSigma	824101
Anti-PAX6 rabbit polyclonal	Proteintech	12323-1-AP
Anti-Nestin mouse monoclonal	MilliporeSigma	ZMS1022
Anti-TUBB3 mouse monoclonal	Proteintech	66375-1-I
Anti-NeuN rabbit polyclonal	Proteintech	26975-1-AP
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 594	ThermoFisher	A-11012
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	ThermoFisher	A-11008

REFERENCES

Adelman, K., & Lis, J. T. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nature Reviews. Genetics*, *13*(10), 720–731. https://doi.org/10.1038/NRG3293

Albrecht, T. R., & Wagner, E. J. (2012). snRNA 3' end formation requires heterodimeric association of integrator subunits. *Molecular and Cellular Biology*, *32*(6), 1112–1123. https://doi.org/10.1128/MCB.06511-11

Baillat, D., Hakimi, M. A., Näär, A. M., Shilatifard, A., Cooch, N., & Shiekhattar, R. (2005). Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell*, *123*(2), 265–276. https://doi.org/10.1016/J.CELL.2005.08.019

Chen, C., & Okayama, H. (1987). High-efficiency transformation of mammalian cells by plasmid DNA. *Molecular and Cellular Biology*, *7*(8), 2745–2752. https://doi.org/10.1128/MCB.7.8.2745-2752.1987

Core, L., & Adelman, K. (2019). Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes & Development*, *33*(15–16), 960–982. https://doi.org/10.1101/GAD.325142.119

Dobbs, S., Metreveli, Z., Seth, K. K., Tomaradze, A., Libby, J., Powell, A., Wilkinson, G., Ecklund, K. M., Love, W., Savinov, V., Lopez, A., Mendez, H., Ramirez, J., Ge, J. Y., Miller, D. H., Shipsey, I. P. J., Xin, B., Adams, G. S., Anderson, M., … Zweber, P. (2008). Precision measurement of the mass of the hc(1P1) state of charmonium. *Physical Review Letters*, *101*(18). https://doi.org/10.1103/PHYSREVLETT.101.182003

Fianu, I., Chen, Y., Dienemann, C., Dybkov, O., Linden, A., Urlaub, H., & Cramer, P. (2021). Structural basis of Integrator-mediated transcription regulation. *Science (New York, N.Y.)*, *374*(6569), 883–887. https://doi.org/10.1126/SCIENCE.ABK0154

Gardini, A., Baillat, D., Cesaroni, M., Hu, D., Marinis, J. M., Wagner, E. J., Lazar, M. A., Shilatifard, A., & Shiekhattar, R. (2014). Integrator regulates transcriptional initiation and pause release following activation. *Molecular Cell*, *56*(1), 128–139. https://doi.org/10.1016/J.MOLCEL.2014.08.004

Groves, M. R., Hanlon, N., Turowski, P., Hemmings, B. A., & Barford, D. (1999). The structure of the protein phosphatase 2A PR65/A subunit reveals the conformation of its 15 tandemly repeated HEAT motifs. *Cell*, *96*(1), 99–110. https://doi.org/10.1016/S0092-8674(00)80963-0

Haesen, D., Asbagh, L. A., Derua, R., Hubert, A., Schrauwen, S., Hoorne, Y., Amant, F., Waelkens, E., Sablina, A., & Janssens, V. (2016). Recurrent PPP2R1A Mutations in Uterine Cancer Act through a Dominant-Negative Mechanism to Promote Malignant Cell Growth. *Cancer Research*, *76*(19), 5719–5731. https://doi.org/10.1158/0008-5472.CAN-15-3342

Hahn, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. *Nature Structural & Molecular Biology*, *11*(5), 394–403. https://doi.org/10.1038/NSMB763

Harlen, K. M., & Churchman, L. S. (2017). The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nature Reviews. Molecular Cell Biology*, *18*(4), 263–273. https://doi.org/10.1038/NRM.2017.10

Janghorban, M., Farrell, A. S., Allen-Petersen, B. L., Pelz, C., Daniel, C. J., Oddo, J., Langer, E. M., Christensen, D. J., & Sears, R. C. (2014). Targeting c-MYC by antagonizing PP2A inhibitors in breast cancer. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(25), 9157–9162. https://doi.org/10.1073/PNAS.1317630111/-/DCSUPPLEMENTAL/PNAS.201317630SI.PDF

Kapp, L. D., Abrams, E. W., Marlow, F. L., & Mullins, M. C. (2013). The integrator complex subunit 6 (Ints6) confines the dorsal organizer in vertebrate embryogenesis. *PLoS Genetics*, *9*(10). https://doi.org/10.1371/JOURNAL.PGEN.1003822

Lai, F., Gardini, A., Zhang, A., & Shiekhattar, R. (2015). Integrator mediates the biogenesis of enhancer RNAs. *Nature*, *525*(7569), 399–403. https://doi.org/10.1038/NATURE14906

Lee, T. I., & Young, R. A. (2013). Transcriptional regulation and its misregulation in disease. *Cell*, *152*(6), 1237–1251. https://doi.org/10.1016/J.CELL.2013.02.014

Lenaerts, L., Reynhout, S., Verbinnen, I., Laumonnier, F., Toutain, A., Bonnet-Brilhault, F., Hoorne, Y., Joss, S., Chassevent, A. K., Smith-Hicks, C., Loeys, B., Joset, P., Steindl, K., Rauch, A., Mehta, S. G., Chung, W. K., Devriendt, K., Holder, S. E., Jewett, T., … Janssens, V. (2021). The broad phenotypic spectrum of PPP2R1A-related neurodevelopmental disorders correlates with the degree of biochemical dysfunction. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, *23*(2), 352–362. https://doi.org/10.1038/S41436-020-00981-2

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics (Oxford, England)*, *30*(7), 923–930. https://doi.org/10.1093/BIOINFORMATICS/BTT656

Linder, P., & Jankowsky, E. (2011). From unwinding to clamping - the DEAD box RNA helicase family. *Nature Reviews. Molecular Cell Biology*, *12*(8), 505–516. https://doi.org/10.1038/NRM3154

Lui, K. Y., Zhao, H., Qiu, C., Li, C., Zhang, Z., Peng, H., Fu, R., Chen, H. an, & Lu, M. qiang. (2017). Integrator complex subunit 6 (INTS6) inhibits hepatocellular carcinoma growth by Wnt pathway and serve as a prognostic marker. *BMC Cancer*, *17*(1). https://doi.org/10.1186/s12885-017-3628-3

Offley, S. R., Pfleiderer, M. M., Zucco, A., Fraudeau, A., Welsh, S. A., Razew, M., Galej, W. P., & Gardini, A. (2023). A combinatorial approach to uncover an additional Integrator subunit. *Cell Reports*, *42*(3). https://doi.org/10.1016/J.CELREP.2023.112244

Perrotti, D., & Neviani, P. (2013). Protein phosphatase 2A: a target for anticancer therapy. *The Lancet. Oncology*, *14*(6). https://doi.org/10.1016/S1470-2045(12)70558-2

Peterlin, B. M., & Price, D. H. (2006). Controlling the elongation phase of transcription with P-TEFb. *Molecular Cell*, *23*(3), 297–305. https://doi.org/10.1016/J.MOLCEL.2006.06.014

Proudfoot, N. J. (2016). Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. *Science (New York, N.Y.)*, *352*(6291). https://doi.org/10.1126/SCIENCE.AAD9926

Rodger, C., Flex, E., Allison, R. J., Sanchis-Juan, A., Hasenahuer, M. A., Cecchetti, S., French, C. E., Edgar, J. R., Carpentieri, G., Ciolfi, A., Pantaleoni, F., Bruselles, A., Onesimo, R., Zampino, G., Marcon, F., Siniscalchi, E., Lees, M., Krishnakumar, D., McCann, E., … Reid, E. (2020). De Novo VPS4A Mutations Cause Multisystem Disease with Abnormal Neurodevelopment. *American Journal of Human Genetics*, *107*(6), 1129–1148. https://doi.org/10.1016/J.AJHG.2020.10.012

Ruvolo, P. P. (2016). The broken "Off" switch in cancer signaling: PP2A as a regulator of tumorigenesis, drug resistance, and immune surveillance. *BBA Clinical*, *6*, 87–99. https://doi.org/10.1016/J.BBACLI.2016.08.002

Sanders, S. J., Campbell, A. J., Cottrell, J. R., Moller, R. S., Wagner, F. F., Auldridge, A. L., Bernier, R. A., Catterall, W. A., Chung, W. K., Empfield, J. R., George, A. L., Hipp, J. F., Khwaja, O.,

Kiskinis, E., Lal, D., Malhotra, D., Millichap, J. J., Otis, T. S., Petrou, S., … Bender, K. J. (2018). Progress in Understanding and Treating SCN2A-Mediated Disorders. *Trends in Neurosciences*, *41*(7), 442–456. https://doi.org/10.1016/J.TINS.2018.03.011

Sents, W., Ivanova, E., Lambrecht, C., Haesen, D., & Janssens, V. (2013). The biogenesis of active protein phosphatase 2A holoenzymes: a tightly regulated process creating phosphatase specificity. *The FEBS Journal*, *280*(2), 644–661. https://doi.org/10.1111/J.1742-4658.2012.08579.X

Shi, Y. (2009). Serine/threonine phosphatases: mechanism through structure. *Cell*, *139*(3), 468–484. https://doi.org/10.1016/J.CELL.2009.10.006

Skaar, J. R., Ferris, A. L., Wu, X., Saraf, A., Khanna, K. K., Florens, L., Washburn, M. P., Hughes, S. H., & Pagano, M. (2015). The Integrator complex controls the termination of transcription at diverse classes of gene targets. *Cell Research*, *25*(3), 288–305. https://doi.org/10.1038/CR.2015.19

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

Thapar, A., Cooper, M., & Rutter, M. (2017). Neurodevelopmental disorders. *The Lancet. Psychiatry*, *4*(4), 339–346. https://doi.org/10.1016/S2215-0366(16)30376-5

Vervoort, S. J., Welsh, S. A., Devlin, J. R., Barbieri, E., Knight, D. A., Offley, S., Bjelosevic, S., Costacurta, M., Todorovski, I., Kearney, C. J., Sandow, J. J., Fan, Z., Blyth, B., McLeod, V., Vissers, J. H. A., Pavic, K., Martin, B. P., Gregory, G., Demosthenous, E., … Johnstone, R. W. (2021). The PP2A-Integrator-CDK9 axis fine-tunes transcription and can be targeted therapeutically in cancer. *Cell*, *184*(12), 3143-3162.e32. https://doi.org/10.1016/J.CELL.2021.04.022

Vos, S. M., Farnung, L., Boehning, M., Wigge, C., Linden, A., Urlaub, H., & Cramer, P. (2018). Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. *Nature*, *560*(7720), 607–612. https://doi.org/10.1038/S41586-018-0440-4

Welsh, S. A., & Gardini, A. (2023). Genomic regulation of transcription and RNA processing by the multitasking Integrator complex. *Nature Reviews Molecular Cell Biology*, *24*(3), 204–220. https://doi.org/10.1038/s41580-022-00534-2

Wieland, I., Röpke, A., Stumm, M., Sell, C., Weidle, U. H., & Wieacker, P. F. (2001). Molecular characterization of the DICE1 (DDX26) tumor suppressor gene in lung carcinoma cells. *Oncology Research*, *12*(11–12), 491–500. https://doi.org/10.3727/096504001108747503

Wlodarchak, N., & Xing, Y. (2016). PP2A as a master regulator of the cell cycle. *Critical Reviews in Biochemistry and Molecular Biology*, *51*(3), 162–184. https://doi.org/10.3109/10409238.2016.1143913

Xu, Y., Xing, Y., Chen, Y., Chao, Y., Lin, Z., Fan, E., Yu, J. W., Strack, S., Jeffrey, P. D., & Shi, Y. (2006). Structure of the protein phosphatase 2A holoenzyme. *Cell*, *127*(6), 1239–1251. https://doi.org/10.1016/J.CELL.2006.11.033

Zhang, S. C., Wernig, M., Duncan, I. D., Brüstle, O., & Thomson, J. A. (2001). In vitro differentiation of transplantable neural precursors from human embryonic stem cells. *Nature Biotechnology*, *19*(12), 1129–1133. https://doi.org/10.1038/NBT1201-1129

Zhang, Y., Li, H., Wang, H., Jia, Z., Xi, H., & Mao, X. (2020). A De Novo Variant Identified in the PPP2R1A Gene in an Infant Induces Neurodevelopmental Abnormalities. *Neuroscience Bulletin*, *36*(2), 179–182. https://doi.org/10.1007/S12264-019-00430-4

Zheng, H., Qi, Y., Hu, S., Cao, X., Xu, C., Yin, Z., Chen, X., Li, Y., Liu, W., Li, J., Wang, J., Wei, G., Liang, K., Chen, F. X., & Xu, Y. (2020). Identification of Integrator-PP2A complex (INTAC), an RNA polymerase II phosphatase. *Science (New York, N.Y.)*, *370*(6520). https://doi.org/10.1126/SCIENCE.ABB5872

Zucco, A. J., Pozzo, V. D., Afinogenova, A., Hart, R. P., Devinsky, O., & D'Arcangelo, G. (2018). Neural progenitors derived from Tuberous Sclerosis Complex patients exhibit attenuated PI3K/AKT signaling and delayed neuronal differentiation. *Molecular and Cellular Neurosciences*, *92*, 149–163. https://doi.org/10.1016/J.MCN.2018.08.004