Role of Nuclear Structure-Function Relationships in Epithelial to Mesenchymal Transitions

A thesis

submitted in the partial fulfilment of the requirements

for the degree of

Doctor of Philosophy

By
Maithilee Sanjay Khot
20133254

Under the Guidance of
Dr Kundan Sengupta

IISER PUNE

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH, PUNE
2021

CERTIFICATE

Certified that the work incorporated in the thesis entitled "Role of Nuclear Structure-Function Relationships in Epithelial to Mesenchymal Transitions" submitted by Maithilee Khot was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other University or institution.

डॉ. कुंदन सेनगुप्ता / Dr. Kundan Sengupta
सहयोगी प्राध्यापक / Associate Professor
भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान
Indian Institute of Science Education & Research
पुणे / Pune – 411 008, India

(Supervisor)

Date:14.01.2021

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH (IISER), PUNE
(An Autonomous Institution, Ministry of Human Resource Development, Govt. of India)
Dr. Homi Bhabha Road, Pashan, Pune – 411 008

DOCTORAL STUDIES OFFICE

Form for submission of Ph.D. Thesis
<u>Declaration by student</u>

Name of Student: <u>Maithilee Khot</u>   Reg. No.:<u> 20133254</u>

Thesis Supervisor(s):<u> Dr Kundan Sengupta</u>   Department: <u>Biology</u>

Date of joining : <u> 01/08/2013</u>   Date of Pre-Synopsis Seminar : <u>16/07/2020</u>

Title of Thesis : <u>Role of Nuclear Structure-Function Relationships in Epithelial to Mesenchymal Transitions.</u> I declare that this written submission represents my idea in my own words and where others' ideas have been included; I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Inst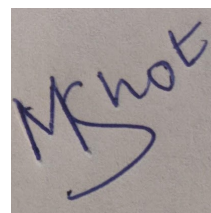itute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

The work reported in this thesis is the original work done by me under the guidance of <u>Dr Kundan Sengupta.</u>

Date:13/01/2021

Signature of the student

Ph: +91-20-25908122 | Fax: +91-20-20251566   *Email:* *doctoraloffice@iiserpune.ac.in*

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH (IISER), PUNE
(An Autonomous Institution, Ministry of Human Resource Development, Govt. of India)
Dr. Homi Bhabha Road, Pashan, Pune – 411 008

DOCTORAL STUDIES OFFICE
Form for submission of Ph.D. Thesis
<u>Certificate by Supervisor</u>

I certify that the thesis entitled

<u>Role of Nuclear Structure-Function Relationships in Epithelial to Mesenchymal Transitions</u>
presented by ~~Mr~~/Ms <u>Maithilee Khot</u> represents ~~his~~/her original work which was carried out
by ~~him~~/her at IISER, Pune under my guidance and supervision during the period from
<u>1/08/2013</u> to <u>31/07/2020.</u> The work presented here or any part of it has not been included
in any other thesis submitted previously for the award of any degree or diploma from any
other University or institutions. I further certify that the above statements made by ~~him~~/her
in regard to ~~his~~/her thesis are correct to the best of my knowledge.
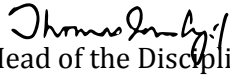
Date :14/01/2021

डॉ. कुंदन सेनगुप्ता / Dr. Kundan Sengupta
सहयोगी प्राध्यापक / Associate Professor
भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान
Indian Institute of Science Education & Research
पुणे / Pune - 411 008, India

(Supervisor)

Recommended and Forwarded

Date: 15.12.2020

Head of the Discipline

Received the Thesis

Date:

Dean, Doctoral studies

<u>**Enclosures**</u> **(**check list)

1. Abstract of Thesis
2. Electronic copy of the Thesis and synopsis separately
3. Consolidated Progress Report as per funding agencies format
4. Copy of online generated receipt of Thesis Submission fees (through SBI I collect)

Ph: +91-20-25908122 | Fax: +91-20-20251566   *Email: doctoraloffice@iiserpune.ac.in*

Acknowledgements

I would like to begin by thanking my PhD Supervisor Dr. Kundan Sengupta for his constant support and guidance throughout the course of my PhD. While he helped me plan and execute experiments, he also trained me to think independently and explore my own ideas. He is extremely approachable and was always available to discuss and debate data and experiments. He has been extremely encouraging and motivating. He has also helped me hone my skills of scientific communication with critical inputs during presentations and manuscript and thesis writing. I am certain that the training I have received in the Chromosome Biology Lab (CBL) will help me in all my future endeavours.

I am grateful to my Research Advisory Committee (RAC) members- Dr. Anuradha Ratnaparkhi (Agharkar Research Institute, Pune), Dr Girish Deshpande (IISER-Pune) and Dr. Mayurika Lahiri (IISER-Pune) for critically evaluating my work and their valuable inputs during my RAC meetings. I would like to thank Kishore Hari and Mohit Kumar Jolly (Center for BioSystems Science and Engineering, Indian Institute of Science, Bengaluru) for the mathematical simulations of our work and Elangoli Ebrahimkutty Faseela and Radhakrishnan Sabarinathan (National Centre for Biological Sciences, Tata Institute of Fundamental Research) for analyses of the TCGA data. These collaborations have helped us understand our experimental data from a new perspective. Discussions with them have always been very thought-provoking and useful. I thank Genotypic Technology Private Limited, Bengaluru, for sample processing and analyses of array CGH data.

I would especially like to thank CBL members, present and past, for making working in the lab enjoyable. I thank Apoorva, for initiating the study on Twist1 overexpression in colorectal cancer cells. I cannot thank Dyuthi and Sanika enough who have helped me develop this work. They have been a constant source of support I would not have been able to do this without them. I thank Vani Pande for help with the gene loci experiments. I thank Ajay, Roopali, Shalaka, Ayantika, Devika, Jiffin and Adwait for discussions and critical suggestions throughout the years.

I am extremely grateful to IISER, Pune for this opportunity and my fellowship and thank IISER for the support in the form of equipment, infrastructure, and management. I would especially like to also thank the IISER microscopy and flow cytometry facilities. I am grateful to the open lab policy of IISER, Pune, which allows sharing of equipment and reagents between the labs.

Lastly, I would like to thank my friends and family for their support throughout my PhD. My parents have been my pillars of strength and have always been supportive and encouraging of my dreams. I am grateful to my husband Bhushan for his support and understanding through these years. I thank Neeraja for her patience and motivation. I thank my in-laws for being supportive and wishing me well.

Table of Contents

Abstract

Epithelial to mesenchymal transition (EMT) is a process essential for morphogenesis during embryonic development. The process of EMT is reactivated in epithelial tumors and is implicated in cancer progression and metastasis. EMT is associated with large-scale cellular and transcriptomic changes. However, how EMT impinges on nuclear structure and function remains unclear. Here we examined the role of two inducers of EMT - (a) TGF-β1 and (b) Twist1 and their impact on genome organization and function in cancer cell lines. We investigated the effect of TGF-β induced EMT by studying the spatial localization of chromosome territories, gene loci and its correlation with gene expression in human lung adenocarcinoma cells (A549). We also examined the regulatory role of components of the nuclear envelope and nuclear landmarks in EMT. Furthermore, we investigated the role of global genome organizers Lamins and CTCF in TGF-β1 induced EMT. Interestingly, while the loss of CTCF did not alter EMT, CTCF knockdown significantly deregulated TGF-β1 induced gene expression changes. In an independent study, we also examined the consequences of overexpressing Twist1 in colorectal cancer cells. Twist1 overexpression was accompanied by an increase in chromosomal instability (CIN). We uncovered the mechanistic involvement of Twist1 in the deregulation of factors that regulate the cell cycle. Twist1 overexpression was accompanied by copy number alterations and an increase in DNA Double Strand Breaks (DSBs). In addition, analyses of gene expression profiles of patient derived datasets from The Cancer Genome Atlas (TCGA) revealed a positive correlation between Twist1 and EMT associated genes across cancers, whereas the correlation of TWIST1 with CIN or DNA DSB repair genes is cancer subtype specific. Interestingly, Twist1 overexpression also downmodulates levels of nuclear lamins, likely to alter spatiotemporal organization of the genome. Taken together, these studies reveal an overarching impact of EMT on genome organization and function, potentially contributing to an increase in aggressive sub-populations of cancer cells.

Abbreviations

EMT: Epithelial to Mesenchymal Transition

MET: Mesenchymal to Epithelial Transition

CIN: Chromosomal Instability

TGFβ: Transforming Growth Factor beta

NB: Nuclear Body

NPC: Nuclear Pore Complex

LINC: Linker of Nucleoskeleton and Cytoskeleton

INM: Inner Nuclear Membrane

ONM: Outer Nuclear Membrane

SAC: Spindle Assembly Checkpoint

CT: Chromosome Territory

TAD: Topologically Associating Domain

CTCF: CCCTC-binding factor

ChIP: Chromatin Immunoprecipitation

FISH: Fluorescence *in-situ* Hybridization

ATAC-seq: Assay for Transposase-Accessible Chromatin using sequencing

array CGH: array Comparative Genomic Hybridization

TCGA: The Cancer Genome Atlas

CNA: Copy Number Alterations

DSB: Double Strand Break

siRNA: small interfering RNA

Chapter 1 Introduction and Review of Literature

1.1 The Nucleus

Robert Brown during 1829–1832 observed a variety of plant cells and found a centrally located entity in all cells that he referred to as the 'nucleus' (Pederson, 2011). No function was attributed to the nucleus at that time. The nucleus was identified as the organelle that contained the genome. The nucleus is devoid of membrane bound sub-compartmentalization. Chromatin and nuclear bodies are dynamic within the nucleus. However, structure-function relationships of the nucleus remain unclear.

1.2 3D organization of the genome

One of the primary levels of organization within the nucleus is compartmentalization of the chromatin into euchromatin and heterochromatin **(Fig.1.1)**. Euchromatin is relatively more open DNA and is enriched for histone modifications such as H3K4me3 and acetylated H3 (Li et al., 2008). Heterochromatin is more condensed chromatin and is largely composed of repetitive DNA sequences. Constitutive heterochromatin is enriched in the histone modifications, namely H3K9me3, while H3K27me3 is often associated with facultative heterochromatin (Trojer and Reinberg, 2007). Condensed state heterochromatin has reduced accessibility to transcription factors and is also repressed in gene expression. Therefore, the extent of chromatin condensation regulates transcriptional activity of genes (Tumbar et al., 1999).

Figure 1.1
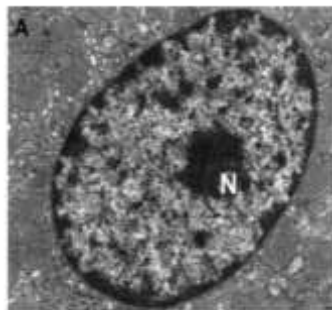
Fig.1.1: Electron micrograph of the nucleus Electron dense heterochromatin is enriched at the nuclear periphery and at the nucleolus (N). Euchromatin is lightly stained as compared to heterochromatin. Reprinted with permission. (Fedorova and Zink, 2008)

Carl Rabl in 1885 suggested the ordered arrangement of chromosomes in the interphase nucleus. He described the non-random organization of centromeric foci at the nuclear periphery, in the nuclei of salamander larvae, which is referred to as the ''Rabl configuration'' (Rabl and C., 1885) **(Fig.1.2A)**. In 1909, Theodor Boveri observed that each chromosome occupies a distinct region in the blastomere, which he referred to as chromosome territories (CTs) (Boveri, 1909) **(Fig.1.2B)**.

Figure 1.2
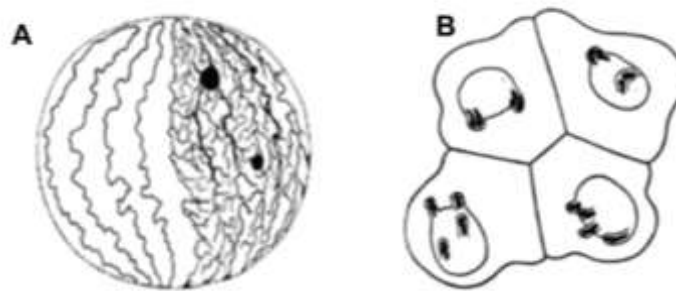
Fig.1.2: Early drawings of non-random organization of the genome in interphase (A): Rabl configuration where the genome is organized primarily as chromatin threads (left side), from which secondary and tertiary threads branch out and form a chromatin network (right side). Centromeres congress at one side of the nucleus (top) whereas the telomeres cluster at the opposite side (bottom). **(B)** Boveri's drawing of a fixed four-cell embryo of *Ascaris megalocephala univalens* shows two pairs of cells with a distinctly different arrangement of nuclear protrusions which contain the distal parts of the two germ line chromosomes in the interphase nuclei. Reprinted with permission. (Cremer and Cremer, 2010)

One of the early evidences of chromosome territories originated from studies on Chinese hamster cells. In these experiments, a laser-UV-microbeam was used to damage a small part of the nucleus. The rationale was that if distinct chromosomes occupy specific nuclear volumes, damaging a small portion of the nucleus would damage only a small subset of chromosomes. A subset of chromosomes was damaged, suggesting that chromosomes indeed occupy distinct regions inside the nucleus (Zorn et al., 1976) **(Fig.1.3)**

Figure 1.3
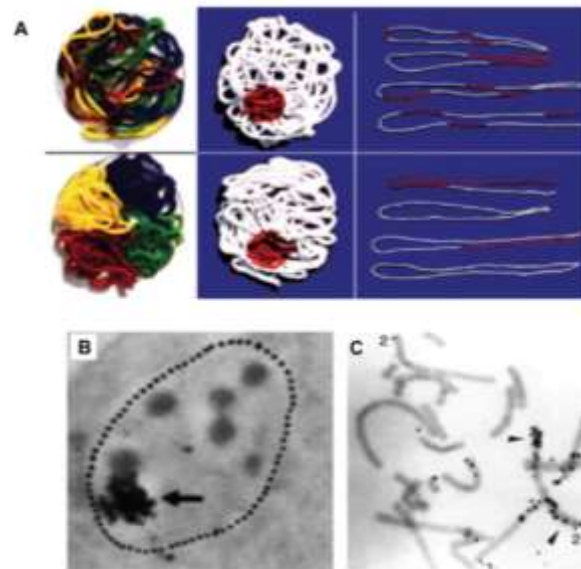
Fig.1.3: Experimental evidence for chromosome territories (A) Experimental rationale of laser-UV-microbeam experiments to distinguish between a non-territorial (upper row) and a territorial (bottom row) chromosome arrangement **(B)** Autoradiograph of a diploid Chinese Hamster cell. The nucleus of a living cell was micro-irradiated in G1, pulse-labelled with ^{3}H thymidine and fixed immediately thereafter. The arrow points to a cluster of silver grains detected over the site of microirradiation. **(C)** Metaphase spread from the same experiment obtained about 40 hours after microirradiation. One chromosome 1 and one chromosome 2 are intensely marked with silver grains, indicating that the microbeam hit the respective territories during interphase. Reprinted with permission. (Cremer and Cremer, 2010)

In the 1980s, multi-colour whole chromosome Fluorescent *in situ* hybridization (FISH) was developed, which enabled visualization of every chromosome inside the nucleus. To study the spatial arrangement of CTs in three-dimensions, 3D-FISH was developed (Cremer et al., 2008). This enabled the understanding of higher-order chromatin organization in the nucleus. CTs show a characteristic radial distribution (Cremer et al., 2001). The non-random radial organization of chromosomes positively correlates with their gene densities. The gene-rich chromosomes are localized towards the centre of the nucleus, while gene poor chromosomes are closer towards the periphery. Human chromosome 18 and 19 territories have been extensively studied owing to their comparable DNA content but divergent gene densities. These chromosome territories occupy a peripheral (CT18) and more central (CT19) localization in the interphase nucleus (Croft et al., 1999). Non-random radial arrangement of chromosomes shows a size dependence in the mouse genome. Larger chromosomes are closer to the nuclear periphery whereas smaller

chromosomes are closer to the nuclear interior (Sun et al., 2000). Chromosome conformation capture assays, especially its variant which is coupled with high-throughput sequencing (Hi-C) (Lieberman-Aiden et al., 2009), reveals pair-wise DNA-DNA contact frequencies across the entire genome (Dekker et al., 2013). Hi-C data reveals that the genome is partitioned into open and active 'A' compartments, and more closed, inactive 'B' compartments, at a megabase (Mbp) scale. Within these compartments are regions of high frequency of contacts, referred to as Topologically Associating Domains (TADs) (Rao et al., 2014). TADs contain chromatin loops that are stabilized by CTCF- a chromatin organizer that binds to the CCCTC motif. CTCF also binds TAD boundaries (Rao et al., 2014) **(Fig.1.4)**. While TAD boundaries are typically conserved, interactions within TADs are variable and mainly specify regulatory contacts that are cell type specific (Smith et al., 2016). Linearly non-adjacent TADs can also contact each other, giving rise to long-range interactions. Long-range interactions may vary between cell types and change during physiological states of a cell, like differentiation (Paulsen et al., 2019; Quinodoz et al., 2018; Szabo et al., 2018). Local decondensation and loop extrusion are required to allow gene expression. Cohesin facilitates chromatin loop extrusion (Davidson et al., 2019). However, cohesin-independent mechanisms of loop extrusion also exist (Brackley et al., 2018).

Figure 1.4
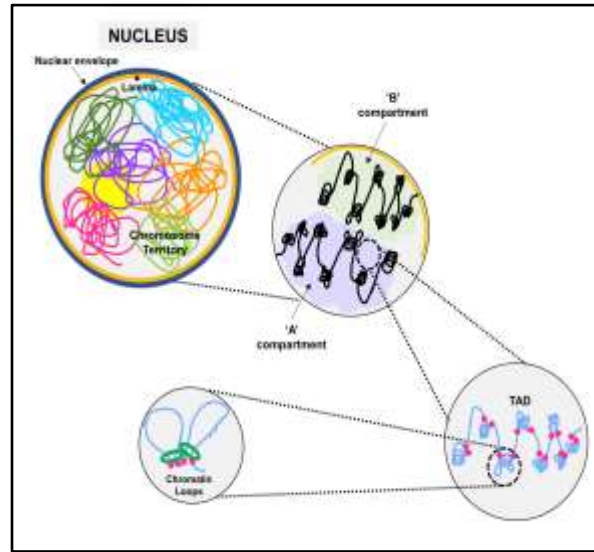
Fig.1.4: Hierarchical three-dimensional arrangement of the genome In the interphase nucleus, each chromosome occupies a distinct sub-volume referred to as 'Chromosome Territory (CT), represented here in different colours. Transcriptionally active regions tend to cluster together usually in the interior of the nucleus to form 'A' compartment while inactive regions form 'B' compartment that is found more towards the nuclear periphery. Each of these compartments is further organized into self-interacting genomic domains known as Topologically Associated Domains (TADs). Each TAD consists of chromatin loops which regulate gene expression.

Additional factors such as nuclear landmarks, like the nucleoli (Quinodoz et al., 2018), splicing speckles (Chen et al., 2018), the nuclear envelope (Buchwalter et al., 2019) and the nuclear pore complex (Labade et al., 2016), also facilitate genome organization by serving as regions for tethering chromatin.

1.2.1 Nuclear bodies and genome organization

The nucleus comprises a number of membrane-less sub-nuclear structures composed of protein and RNA aggregates, known as nuclear bodies (NBs) (Mao et al., 2011). NBs are highly dynamic structures that can self-assemble and disassemble, while its constituents exchange with the surrounding nucleoplasm. NBs are repositories of substrates, enzymes and other intermediates within a confined space, that enhance reaction rates of enzymatic reactions (Mao et al., 2011). The composition of nuclear bodies and their numbers is cell type specific (Förthmann et al., 2013). Nuclear bodies include the nucleolus, Cajal bodies, paraspeckles, and PML bodies, with different functions. Specific gene loci associate with NBs that regulate gene activity. For

example, immuno-FISH analysis of Jurkat T cells identified *TP53* gene loci in proximity to a PML body in ~50% of cells (Sun et al., 2003). Heat-shock associated gene Hsp70 shows enhanced contact with nuclear speckles upon heat shock, leading to transcriptional upregulation (Khanna et al., 2014). The nucleolus, the most prominent nuclear body, is involved in ribosome biogenesis. The rDNA is present as tandem repeats on Nucleolar Organizer Regions (NOR) of human chromosomes 13, 14, 15, 21 and 22. After mitosis, the nucleolus forms around active NORs and is thus involved in the structural and functional organization of these chromosomes (Fulka and Aoki, 2016). Regions across the genome contact the nucleolus at 'Nucleolus Associated Domains' (NADs) (Németh et al., 2010). NADs are regions of low gene densities, low transcriptional levels and repressive histone modifications (H4K20me3, H3K27me3, and H3K9me3). Centromeric and pericentromeric satellite repeats and subtelomeric regions were also identified as NADs (van Koningsbruggen et al., 2010). In summary, nuclear bodies organize the genome and regulate gene activity.

1.2.2 Nuclear envelope and genome organization

In addition to nuclear bodies, the nuclear envelope is also involved in chromatin organization. The nuclear envelope is composed of the inner nuclear membrane (INM) and the outer nuclear membrane (ONM). The outer nuclear membrane is continuous with the endoplasmic reticulum (ER) (Watson, 1955). The INM and the ONM have a distinct subset of proteins rarely found in the ER (Hetzer et al., 2005). The nuclear envelope is intermittently perforated by the nuclear pore complex (NPC). The primary role of the NPC is the regulated entry and exit of molecules between the nucleus and the cytoplasm. Transport independent functions such as chromatin organization and gene regulation have now been attributed to nucleoporins (NUPs) (Hetzer et al., 2005; Labade et al., 2016) **(Table 1.1).**

The nuclear envelope has over 60 different associated proteins. SUN-domain containing proteins are present in the INM, while KASH family members (Nesprin 1/2) are present in ONM. These two sets of proteins interact in the interstitial space of the two membranes and together form the LINC (Linker of Nucleoskeleton and Cytoskeleton) complex. The LINC complex physically connects the nucleus to the cell cytoskeleton (Crisp et al., 2006). The LINC complex functions as a mechanotransducer, which modulates cell migration, chromosome territory organization and transcription (Rothballer et al., 2013) **(Fig.1.5).**

Figure 1.5
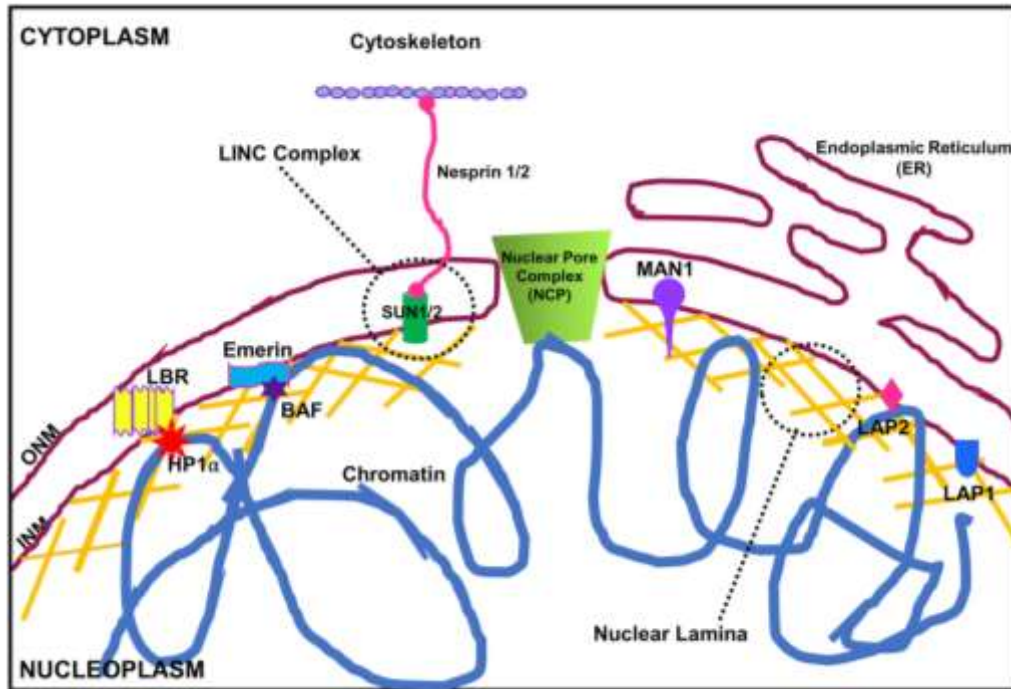
Fig.1.5: Schematic model of the nuclear envelope proteins The nuclear envelope consists of the inner nuclear membrane (INM) and outer nuclear membrane (ONM) separated by the perinuclear space. Various nuclear transmembrane proteins are embedded into the INM that interact with nuclear lamina and chromatin. The LINC complex proteins connect the nucleus with the cell cytoskeleton. The nuclear pore complex (NPC) is involved in nuclear transport as well as genome organization.

Another set of proteins, the nuclear lamins, line the inner nuclear membrane (Aaronson and Blobel, 1975). The lamin meshwork or the nuclear lamina, composed of Lamin A/C, Lamin B1 and Lamin B2 provide structural rigidity to the nucleus (Ho and Lammerding, 2012). Various proteins interact with lamins, such as the lamin B Receptor (LBR) (Pyrpasopoulou et al., 1996), Emerin (Clements et al., 2000), MAN1 (Paulin-Levasseur et al., 1996) and Lamina-Associated Polypeptide 1 & 2 (LAP1 & LAP2) (Foisner and Gerace, 1993). The fact that defects in nuclear envelope proteins lead to 'Nuclear Envelopathies', highlights the importance of these proteins (Taimen et al., 2009). Some of these proteins are cell type specific. As a result, different proteins contribute to genome organization through mechanisms unique to the cell type (Peric-Hupkes et al., 2010). For example, Lamin A/C and Emerin mutations specifically lead to muscular dystrophies, suggestive of cell type specific functions (Gotzmann and Foisner, 2013)**.**

Chromatin association with the nuclear envelope proteins contributes to the spatio-temporal organization of the genome **(Table 1.1)**. While certain active regions of the genome interact with NPCs, inactive or silent regions interact with the INM proteins and the nuclear lamina, forming lamina-associated domains (LADs) (Guelen et al., 2008). Lamin B receptor (LBR), a lamin interacting protein also contributes to tethering heterochromatin to the nuclear periphery (Briand and Collas, 2020). LAP2α, LAP2β and BAF are also involved in chromatin interactions (Dorner et al., 2007). Other INM proteins containing a LEM domain, such as Emerin, LEMD2 and LEMD3/MAN1 also bind lamins and chromatin (Pradhan et al., 2018). LAP2β and Emerin interact with histone deacetylase HDAC3, and contribute to the repressive environment of the nuclear periphery (Demmerle et al., 2012). Dermal fibroblast cell lines derived from laminopathy patients showing LMNA mutations, and X-EDMD patients with Emerin mutation show mislocation of gene poor chromosomes 13 and 18 away from the nuclear periphery, emphasizing the role of nuclear envelope proteins in genome organization (Meaburn et al., 2007).

Table 1.1: Nuclear envelope proteins as genome organizers

Nuclear envelope protein	Model	Method	Highlights	Reference
Lamin A/C	HeLa	ChIP seq	Lamin interacting domains (LiDs) interact with both peripheral and internal lamins, LiDs contain euchromatic and heterochromatic regions	(Lund et al., 2015)
Lamin B1	HT1080, TIG3	ChIP seq, DamID	Lamina-associated domain (LADs), repressive environment at the periphery, LAD boundaries marks by CTCF Constitutive LADs (cLADs) and facultative LADs (fLADs)	(Meuleman et al., 2013) (Guelen et al., 2008)
Emerin	*C. elegans*	DamID	Binds transcriptionally inactive regions, enriched on genes involved in muscle and neuronal function	(González-Aguilera et al., 2014)
LAP2a	Immortalized murine dermal fibroblasts	ChIP	Overlap with Lamin A associated domains, binds euchromatin	(Gesson et al., 2016)
LBR	U2OS	DamID	Binds heterochromatin, enriched for H3K9me2, overlaps with LMNB1 LADs	(Ibarra et al., 2016)
Nup98	IMR90,ESC, HeLa	ChIP seq	Associates with developmentally associated genes- neuronal development	(Liang et al., 2013)
Nup93	U2OS, IMR90	DamID	Enriched for H3K4me3 and H3K27ac, super enhancers, regulate cell identity genes, shows overlap with CTCF binding	(Ibarra et al., 2016)
Nup153	U2OS, IMR90	DamID	Enriched for H3K4me3 and H3K27ac, super enhancers, regulate cell identity genes	(Ibarra et al., 2016)

Dedicated mechanisms exist in order to re-establish chromatin organization after every cell division. During cell division, the nuclear envelope disassembles during early prophase and re-assembles during early/late telophase. Simultaneously, chromosomes condense and segregate to each daughter cell. After cell division, the interphase genome organization is restored in a cell type specific manner by nuclear envelope proteins (Gerlich et al., 2003). For example, in late anaphase, BAF bridges chromosomes to form a single chromatin mass that serves as a surface

for nuclear envelope assembly (Samwer et al., 2017). BAF recruits other LEM proteins of the nuclear envelope and nuclear lamins to the chromatin (Wandke and Kutay, 2013). Nucleoporins bind to histones and act as seeding sites for future NPC assembly (Zierhut et al., 2014). Heterochromatin protein 1 (HP1) and LBR bind heterochromatin to position it at the nuclear periphery (Solovei et al., 2013). Furthermore, nuclear envelope proteins like B-type lamins and SUN1 bind to spindle apparatus during mitosis. In summary, nuclear envelope proteins are required for faithful cell division and re-establishing genome organization after every cell division.

Genome organization is cell type specific and in order to understand the functional relevance of nuclear organization, it is important to examine genome organization across cell types. Croft *et al*, studied CT organization in two cell types: primary lymphocytes and lymphoblastoid cell line (Croft et al., 1999). In spite of cell type specific organization observed in the two cells, its functional relevance was unclear. Studying nuclear architecture changes that accompany functional changes in dynamic processes, provides crucial insights in understanding nuclear structure-function relationships. Changes in nuclear organization have also been studied in the context of cellular differentiation. Furthermore, CT repositioning accompanies differentiation of mesenchymal stem cells (MSCs) into adipocytes (Szczerbal et al., 2009).

In this thesis my aim was to examine nuclear structure-function relationships in the dynamic process of epithelial to mesenchymal transition (EMT). EMT is a process in which epithelial cells trans-differentiate into mesenchymal cells. EMT offers a tractable system *in-vitro* to manipulate cell phenotypes. In addition, EMT is a reversible process. We examined nuclear architecture changes associated with change in cell phenotypes during EMT. Furthermore, EMT is involved in cancer metastasis, making it a relevant paradigm to study nuclear structure-function relationships.

1.3 Epithelial to Mesenchymal Transition (EMT)

The epithelial to mesenchymal transition (EMT) is a cellular program involving conversion of epithelial cells to mesenchymal cells. It was described for the first time by Frank Lillie in 1908 (Lillie, 1908). In the 1960s Elizabeth Hay studied formation of the primitive streak in Chick

embryo, that requires the conversion of epithelial cells to mesenchymal cells (Trelstad et al., 1967). This was the first comprehensive report of Epithelial to Mesenchymal "Transformation". However, the term transformation was replaced with transition, to distinguish it from neoplastic transformation (Kalluri and Neilson, 2003). In 1982, Garry Greenburg and Elizabeth Hay demonstrated that epithelial cells of the eye lens of embryonic and adult chick, cultured in 3D collagen gels, show cell elongation, detachment from explants, and migration of individual cells (Greenburg and Hay, 1982)), in a process termed as Epithelial to Mesenchymal Transition (EMT) (Yang et al., 2004)). EMT was recognized during various stages of embryonic development and is involved in tissue remodelling events like mesoderm formation, neural crest development, heart valve development, secondary palate formation, and male Mullerian duct regression (Yang et al., 2004).

Michael Stocker and Michael Perryman discovered that the epithelial cell line- Madin–Darby canine kidney (MDCK) could be converted into migratory fibroblasts when cultured with conditioned medium, from fibroblasts. The factor responsible was termed the 'scatter factor' (Stoker et al., 1987), and was later identified as hepatocyte growth factor (HGF) (Weidner et al., 1993). The discovery that EMT could be achieved *in vitro*, thus enabled experimental manipulations into the process of EMT.

EMT is broadly characterized by three major changes in cellular phenotype (i) morphological changes from a cobblestone-like monolayer of epithelial cells to dispersed, spindle-shaped mesenchymal cells (ii) Decrease in expression of epithelial markers like E-cadherin, ZO1, Occludin and increase in mesenchymal markers like Vimentin, N-cadherin, α-SMA and (iii) the functional changes associated with the conversion of immotile cells to motile cells. Not all three changes are always observed during an EMT; however, acquisition of the ability to migrate and invade ECM as single cells is considered a functional hallmark of EMT (Lamouille et al., 2014). EMT is also a reversible program; certain epithelial cells that have undergone EMT to become mesenchymal cells, may revert back to its epithelial state through Mesenchymal to Epithelial Transition (MET) (Bendinelli et al., 2015). However, recent studies do not support a binary transition from E to M states. Epithelial cells may show both epithelial and mesenchymal properties and exist as E/M hybrids. Therefore EMT is a continuum of phenotypic cell states. 'EMT hybrid' states could range from co-expression of epithelial and mesenchymal markers to

cells that have dampening of epithelial characteristics, such as remodelling of junction proteins or loss of apicobasal polarity without gain of mesenchymal properties (Jolly et al., 2015). Notwithstanding gain of mesenchymal properties, EMT is characterized by considerable cell-cell heterogeneity, qualifying EMT as more of a plastic process rather than a transition- called epithelial to mesenchymal 'plasticity' (EMP) (Nieto et al., 2016) **(Fig.1.6)**.

Figure 1.6
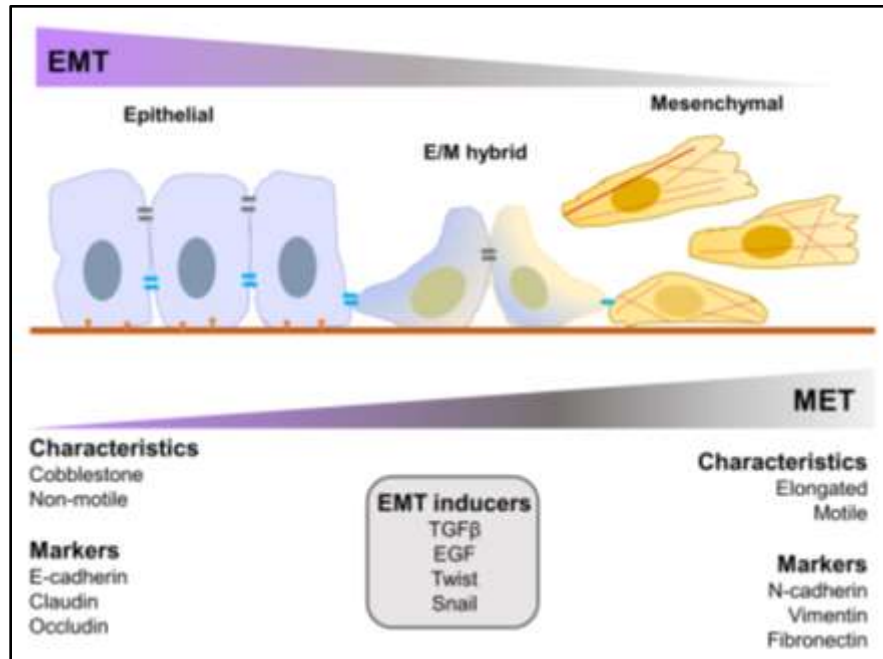
Fig.1.6: Epithelial to Mesenchymal Plasticity Epithelial to Mesenchymal Transition is a continuum of phenotypes and the intermediate stages are termed as E/M hybrids. During this transition, depending on the stage, cells demonstrate a wide spectrum of epithelial and mesenchymal characteristics.

EMT is central to both physiological and pathological processes. EMT is central to both physiological and pathological processes. EMT is observed in three different biological settings and each has a distinct functional outcome. Based on biological context, EMT is classified into three types. EMT involved in-Development (Type I) ,Wound healing and Fibrosis (Type II) and Cancer metastasis (Type III) (Kalluri and Weinberg, 2009) .

Early stages of embryogenesis, implantation of the embryo and formation of the placenta involve EMT. This is the first time that EMT occurs during development (Vićovac and Aplin, 1996) . Later during embryogenesis, the progress of gastrulation generates three germ layers-endoderm,

mesoderm and ectoderm. Gastrulation involves EMT. During gastrulation, the epithelial cells of the epiblast layer undergo EMT and give rise to a furrow or invagination known as the primitive streak. Cells of the primitive streak gain the property of migration and form the meso-endoderm that later separates to form the mesoderm and the endoderm through EMT (Hay, 1990). As embryogenesis progresses, EMT is also involved in the formation of the neural crest, palatal roof closure and cardiac septal formation. These processes involving EMT also show MET which give rise to new epithelia during development (Thiery et al., 2009).

EMT is also an essential process in wound healing. Wound healing involves three phases: inflammatory, proliferative and maturation phases. The inflammatory phase limits tissue damage through phagocytosis. The proliferative phase leads to formation of granulation tissue, angiogenesis, deposition of new ECM and then re-epithelialization (Haensel and Dai, 2018). During wound healing, keratinocytes at the border of the wound undergo EMT and re-epithelialization or MET when the wound is closed (Haensel and Dai, 2018). During wound healing, contractile myofibroblasts secretes large amounts of ECM proteins and aids the wound closure (Hinz and Gabbiani, 2003). In the normal wound healing process, many myofibroblasts undergo apoptosis and disappear once re-epithelialization is complete (Hinz and Gabbiani, 2003). However, prolonged myofibroblast activity results in fibrogenesis. Persistent myofibroblast activation is a common characteristic of fibrotic diseases. Thus, dysregulation of injury-triggered EMT is believed to contribute to fibrosis of multiple organs (Stone et al., 2016).

Cancers of epithelial origin employ EMT to metastasize. Metastasis is a hallmark of cancer progression- the process by which cells from tumors migrate from their primary site of origin to distant sites where they form secondary tumors. Epithelial cells in the primary tumor escape the confines of cell-cell junctions, facilitating cell migration during EMT (Thiery, 2002). These cells gain stem-like properties, chemoresistance, radioresistance and resistance to anoikis (Tiwari et al., 2012).

While the three types of EMTs represent distinct biological processes, a common set of genetic and biochemical elements appears to regulate these processes. As research on these is progressing, we are discovering the similarities and differences among the three types of EMTs.

Work in this thesis focuses on type III EMT and subsequent literature is discussed more in the context of type III EMT.

Various extracellular signals and factors induce EMT. However, EMT induction is cell type and inducer specific (Thiery et al., 2009). EMT is mediated by signaling pathways such as Transforming Growth Factor β (TGF-β) and Bone Morphogenetic Protein (BMP), Wnt–β-catenin, Notch and Hedgehog (Gonzalez and Medici, 2014). Signaling pathways of EMT converge on a subset of effector molecules- EMT transcription factors (EMT-TFs). EMT-TFs induce EMT by enhancing mesenchymal properties. E-cadherin, a pluripotent calcium-dependent adhesion molecule, is expressed in most epithelial tissues to connect adjacent epithelial cells. Loss of E-cadherin is a central event in EMT and EMT-TFs are key repressors of E-cadherin. E-cadherin decrease is associated with induction of EMT. However, E-cadherin knockdown is not always sufficient to induce EMT. Similarly, restoring E-cadherin is insufficient to revert EMT. EMT-TFs such as Snail1, Snail2, Zeb1, Zeb2, Twist1 and Twist2 regulate E-cadherin transcription. EMT-TFs are common to developmental EMT and pathological EMT (Nieto and Cano, 2012). EMT-TFs show functional overlap with a temporal hierarchy, and cooperation in the activation of these transcription factors. For example, in TGF-β1 induced EMT, Snail is expressed first followed by Twist expression (Tran et al., 2011).

In conclusion, EMT is a developmental process that is also utilized by cells of epithelial cancer for metastasis. EMT is an extremely regulated, multistage process, involving a high degree of cellular plasticity.

1.4 Epigenetic regulation of EMT

Studies in the past decade have revealed that along with EMT-TFs, different epigenetic regulators are also essential mediators of EMT. They modulate the activity of EMT-TFs and thus control the expression of various epithelial and mesenchymal genes. Epigenetic modifications such as DNA methylation, histone modifications and miRNAs also regulate and fine tune epithelial-mesenchymal plasticity.

1.4.1 DNA methylation

DNA methylation is a covalent modification that usually occurs at the 5'-position of the cytosine ring (5mC) within CpG dinucleotides, repressing transcription. During EMT, genome-wide CpG methylation pattern remains unchanged. However, certain CpG sites associated with transcriptional regulation of EMT-associated genes show altered methylation (Carmona et al., 2014; McDonald et al., 2011). Promoter hypermethylation silences E-cadherin in a wide range of cancer cells (Lombaerts et al., 2006; Reinhold et al., 2007). Several EMT-TFs like Snail1 and Zeb1 recruit DNA methyltransferases (DNMT1) to E-cadherin promoter for CpG methylation and thus repress E-cadherin (Fukagawa et al., 2015). Methyl-DNA-binding domain (MBD) proteins, such as MeCP2, MBD1-4 and Kaiso are involved in DNA methylation mediated silencing. Twist1 recruits MBD1 on E-cadherin promoter and silences gene expression in pancreatic cancer cells (Xu et al., 2013). Furthermore, promoter CpG methylation mediated downregulation of E- cadherin, correlates with elevated MeCP2 expression in colorectal cancer cells (Darwanto et al., 2003), and upon overexpression of Kaiso during prostate and breast carcinogenesis (Jones et al., 2012, 2014). In summary, DNA methylation of EMT associated genes function as an important regulator of EMT.

1.4.2 Histone modifications

1.4.2.1 Histone acetylation

Transcriptionally permissive euchromatin usually is hyperacetylated by histone acetyltransferases (HATs) such as PCAF, p300/CBP, TIP60 and hMOF (Lee and Workman, 2007). In Wnt and β-catenin dependent EMT pathway, β-catenin is translocated to the nucleus and recruits p300/CBP on target gene promoters and up-regulates their expression (Hao et al., 2011; Zhou et al., 2012). Furthermore, p300 and PCAF interact with ZEB1 and Twist1 (Hamamori et al., 1999)(Mizuguchi et al., 2012)(Hamamori et al., 1999) to regulate EMT. Over-expression of Snail1 increases TGFβ1 signaling by up-regulating TGFBR2 expression. Increased H3K9Ac was detected on TGFBR2 promoter upon Snail1/2 overexpression (Dhasarathy et al., 2011). Histone deacetylases (HDACs) counteract HATs and predominantly function as transcriptional repressors. Snail1 recruits HDAC1/2-containing SIN3A complex to the E-cadherin promoter for gene silencing (Peinado et al., 2004). HDAC1/2 is recruited by Zeb1 to the

E-cadherin promoter for repression in pancreatic cancer cells. In prostate cancer cells, Zeb1 recruits SIRT1, a HDAC, for E-cadherin silencing and to promote EMT and metastasis (Aghdassi et al., 2012). ZEB1 induced EMT also showed a global H3K27 deacetylation, suggesting that histone acetylation is a key epigenetic modification in transcriptional reprogramming of EMT (Roche et al., 2013).

1.4.2.2 Histone methylation

Arginine methyltransferases (PRMTs) can mono- and di-methylate arginines. Similarly, lysines can be mono-, di-, and tri- methylated by histone methyltransferases (HMTases). Euchromatin is marked by H3K4me2/3, especially in the promoter and enhancer regions. H3K36me3 is associated with transcriptional elongation. TGF-β induced EMT in mouse AML12 cells shows a global increase of H3K4me3 and H3K36me3 marks. Thus, there is a genome-wide alteration in these two permissive methylation marks upon EMT (McDonald et al., 2011). Along with HMTases, histone methylation is also regulated by histone demethylases (HDMs). Overexpression of LSD1, a histone demethylase is observed across tumors, including colorectal, breast and prostate (Højfeldt et al., 2013). Taken together, regulation of the permissive H3K4 methylation at different regions across the genome regulates EMT. ChIP-chip analyses upon TGF-β induced EMT showed that reduction in H3K9me2 levels are mainly specific to large organized heterochromatin K9 modifications (LOCKs) (McDonald et al., 2011). Interestingly, LOCKs overlap with LADs. This suggests that EMT may lead to global reorganization of the genome.

Transcriptionally repressed genes are labelled with histone modification such as H3K9, H3K27, H4K20 and H4R3. G9a (a HMT), catalyzes the addition of H3K9me2 modification on the promoter of EPCAM, which is a cell adhesion molecule. This repressses EpCAM expression and leads to EMT and metastasis, both *in vitro* and *in vivo,* in lung cancer (Chen et al., 2010). In breast cancer cells, Snail1 recruits G9a and DNA methyltransferase (DNMT1, DNMT3a, and DNMT3b) to the promoter of CDH1 and silences it. Inhibiting G9a decreases H3K9me2, curtailing EMT and tumor metastasis (Dong et al., 2012). G9a also represses E-cadherin in head and neck squamous cell carcinoma (Liu et al., 2015). This suggests that Snail1-G9a mediated repression of E-cadherin is potentially a common mechanism of E-cadherin downregulation during EMT. In addition, Snail1recruits SUV39H1, another HMT, to repress E-cadherin

promoter via H3K9me3 (Dong et al., 2013). Histone methyltransferase, SETDB1 is down-regulated in breast cancer cells and correlates with reduced EMT and CSC. Methylation of H3K27 is another important histone modification that represses gene expression and is catalyzed by EZH2 or EZH1- containing Polycomb repressive complex 2 (PRC2) (Di Croce and Helin, 2013). Upon Twist1 overexpression in epithelial breast cell line, HMLE, H3K4me3 to H3K27me3 switch was observed on EMT marker gene promoters, including E-cadherin. On the other hand, genes such as ZEB2, N-cadherin, PDGFRα and ESRP1 that are upregulated upon EMT, show H3K27me3 to H3K4me3 switch (Malouf et al., 2013). In breast cancer cells, the bivalent chromatin state was found on ZEB1 promoter, which facilitates the rapid switch to a CSC-like state upon different EMT signals (Chaffer et al., 2013). In summary, histone modification alters the state of chromatin compaction, which in turn affects genome organization and regulates the EMT programme.

1.4.3 miRNA

EMT is controlled by a complex regulatory network. miRNAs have emerged as important post-translational regulators of EMT. miRNAs are small (19–25 nucleotides long) noncoding, single-stranded RNAs that control gene expression by targeting mRNA transcripts and leading to their translational repression or degradation, according to the level of complementarity (Tornesello et al., 2020). miR-200 (Mongroo and Rustgi, 2010) and miR-34a (Nie et al., 2019) family are important players in the EMT circuit that promote MET, and protect the epithelial phenotype. The miR-200 family members directly degrade mRNAs encoding ZEB1 and ZEB2 (Park et al., 2008). Similarly, Zeb1 binds to and represses the promoter of miR-200 associated genes, forming a double-negative feedback loop (Burk et al., 2008; Gregory et al., 2011). EMT-TFs like Snail, Zeb, Gata3 and several miRNAs are also implicated in forming a double-negative feedback loop and regulating EMT (Díaz-López et al., 2014) .

Over time, a large number of miRNAs have been added to the repository of EMT regulating miRNAs. miR-9, a MYC-induced miRNA, directly targets the E-cadherin-encoding mRNA and causes its downregulation, leading to increased cell migration and EMT-like phenotype [29]. The miR-221/222 target a set of factors like ESR1 (Di Leva et al., 2010), Dicer (Cochrane et al., 2010), and TRPS1 and induce EMT induction in breast cancer cells (Stinson et al., 2011). TGF-β

induced EMT in NMuMG mammary epithelial cells leads to expression of miR-155, while knockdown of miR-155 inhibits TGF-β induced EMT, migration, and invasion, potentially via miR-155 mediated downregulation of RHOA (Kong et al., 2008). miR-30a targets SNAI1 in non-small cell lung cancer cells and thus inhibits EMT (Kumarswamy et al., 2012). Taken together, miRNAs have emerged as important regulators of EMT **(Table 1.2)**.

Table 1.2: miRNAs that regulate EMT (Reviewed in (Zhang and Ma, 2012). Reprinted with permission)

Expression miRNA	Effect on EMT	Target Gene	Reference
miR-9	+	CDH1	(Ma et al., 2010)
miR-15b	-	BMI1	(Sun et al., 2012)
miR-27	+	APC	(Zhang et al., 2011)
miR-29a	+	TTP	(Gebeshuber et al., 2009)
miR-30a	-	Snail	(Kumarswamy et al., 2012)
miR-103/107	+	DICER1	(Martello et al., 2010)
miR-155	+	RHOA	(Chaffer et al., 2013)
miR-194	-	BMI1	(Dong et al., 2011)
miR-200 family	-	ZEB1/ZEB2, Sec23a	(Park et al., 2008), (Gregory et al., 2011) (Korpal et al., 2011), (Kim et al., 2011)
miR-205	-	ZEB1/ZEB2	(Gregory et al., 2008)
miR-204	-	TGFβR2, SNAIL2	(Wang et al., 2010)
miR-221/222	+	TRPS1, ESR1, DICER1	(Cochrane et al., 2010; Di Leva et al., 2010; Stinson et al., 2011)

+ Promotes

- Suppresses

1.5 Impact of EMT on nuclear organization

1.5.1 Nuclear envelope and signaling

The nuclear lamina serves as a signaling hub and tethers signaling molecules, c-Fos and ERK, that both bind to Lamin A/C where ERK phosphorylates c-Fos and activates c-Fos/AP1 driven transcription downstream (Rodríguez et al., 2010). Furthermore, the Lamin scaffold can recruit PP2A that de-phosphorylates pRb and regulates TGF-β signaling (Van Berlo et al., 2005). Nucleoplasmic pool of A-type lamins along with LAP2α regulates pRb/E2F signaling, fine tuning cell proliferation and differentiation. Loss of LAP2α shows increased cell proliferation (Gesson et al., 2014). MAN1- an inner nuclear membrane protein, interacts with Smad2/3, which are effector molecules of TGF-β signaling, and antagonize TGF-β signaling (Lin et al., 2005). Similarly, Emerin regulates the nuclear flux of β-catenin, an effector molecule of Wnt signaling (Markiewicz et al., 2006). Emerin also regulates mechanosignaling and enhances actin polymerization (Holaska et al., 2004), which regulates MKL1 (Ho et al., 2013) which in turn upregulates genes regulated by serum response factor (SRF), like actin and vinculin (Miralles et al., 2003). Thus, nuclear envelope proteins regulate cell signaling pathways including pathways such as TGF-β and Wnt signaling involved in EMT induction .

1.5.2 Effect of EMT on genome

1.5.2.1 Genome instability

Studies in breast epithelial cell line, MCF10A show that cells with persistent proliferation during TGF-β induced EMT show failed cytokinesis. They sustain mitotic abnormalities, giving rise to aneuploidy. Furthermore, along with genomic instability, these cells show downregulation of multiple nuclear envelope proteins that are known to be involved in mitotic regulation. Lamin B1 downregulation phenocopied these effects. In metastatic breast cancer patients, increased mesenchymal marker expression within single circulating tumor cells is correlated with genomic instability (Comaills et al., 2016). Similarly, Twist1 overexpression also leads to aneuploidy in the MCF7 cell line (Vesuna et al., 2006).

1.5.2.2 Genome organization

Another aspect of genome organization that is altered by EMT are chromatin modifications, which may in turn lead to changes in genome organization. As mentioned previously, studies have focused on genome-wide changes in DNA methylation and histone modifications associated with EMT. How this correlates with genome organization and impinges on transcriptional changes in EMT is largely unknown.

The focus of this project is to understand how epithelial to mesenchymal transition, which is associated with morphological changes of cells, impinges on the structure and function of the nucleus. Further, this project aims to unravel the mechanisms by which components of the nuclear membrane and other nuclear landmarks, regulate the dynamic process of epithelial to mesenchymal transition.

Chapter 2 Materials and Methods

2.1 Cell culture

Cell lines used in this study were maintained in appropriate growth medium supplemented with antibiotics: penicillin (100 units/ml) & streptomycin (100 μg/ml) (Gibco, 15070-063) and 10% heat-inactivated fetal bovine serum (FBS) (Gibco, 6140). Cells were cultured at 37°C in the presence of 5% CO_2. Cells were sub-cultured when ~60-70% confluent and early passage cells (up to ~10-15 passages) were used for all experiments. We ensured that cultures were free of *Mycoplasma* contamination by DAPI staining the cells periodically. Human colorectal adenocarcinoma cell lines DLD1 and SW480 were a kind gift from the laboratory of Thomas Ried (NCI/NIH, Bethesda, USA). HCT116 colorectal cancer cell line was received from Dr Mayurika Lahiri, IISER-Pune. A549 lung adenocarcinoma cell line (ATCC® CCL-185™) was procured from ATCC. DLD1 cells were cultured in RPMI1640 (Gibco, 11875), while HCT116, SW480 and A549 cells were cultured in DMEM (Gibco, 11995). The karyotypes of these cell lines were validated using DAPI-stained metaphase spreads. The karyotypes of these cell lines were stable and did not vary across passages as shown by analyses of metaphase spreads that showed a consistent modal number for DLD1 (45–46 chromosomes), HCT116 (42–43 chromosomes), SW480 cells (56–57 chromosomes) and A549 cells (63-64 chromosomes).

2.2 Western blotting

Protein lysates were prepared by scraping cells in ice-cold RIPA buffer (pH=7.2, 50 mM Tris Cl, 150 mM NaCl, 0.1% SDS, 0.01% sodium azide, 0.5% sodium deoxycholate, 1 mM DTT, 1% NP40) containing 1X protease inhibitor cocktail (Roche, 4693116001). This was followed by centrifugation at 300g at 4°C for 10 min. Protein estimation was performed using BCA kit (Thermo Fischer Scientific, 23225) and an equal amount of protein was loaded onto an SDS-PAGE gel. Proteins thus resolved were transferred onto a PVDF membrane (Millipore, IPVH00010). Immunoblots were blocked using 5% non-fat milk prepared in 1X TBST (pH 7.4). Immunodetection was performed by adding primary antibodies against the protein of interest and incubated overnight at 4°C. The antibodies used in this study and the dilutions at which they were used are enlisted in **Table 2.1**. Following primary antibody, the blots were incubated with appropriate secondary antibodies. Secondary antibodies used were sheep anti-mouse-HRP

(Amersham, NA9310V), donkey anti-rabbit-HRP (Amersham, NA9340V) and goat anti-rat-HRP (Amersham, NA935) at 1:10,000 dilution for 1 h at RT. Between incubation, blots were rinsed thrice with 1X TBST for 10 min each at RT. Chemiluminescent substrate ECL Prime (Amersham, 89168-782) was used to develop immunoblots and imaged with ImageQuant LAS 4000.

Table 2.1: List of primary antibodies with their dilutions used for western blotting

S.No.	Antibody	Dilution
1	Twist1 (ab50887)	1:500
2	E-cadherin (ab1416)	1:1000
3	Vimentin (Sigma, V2258)	1:500
4	Lamin (A + C) (ab108595)	1:1000
5	Lamin B1 (ab16048)	1:1000
6	Lamin B2 (ab8983)	1:500
7	Bub1 (ab54893)	1:1000
8	BubR1 (ab54894)	1:1000
9	Mad1 (ab126148)	1:3000
10	Mad2 (ab24588)	1:500
11	Aurora B Kinase (CST3094) (Dr Balasubramanian, IISER-Pune)	1:1000
12	Aurora B Kinase (ab2254)	1:1000
13	GAPDH (Sigma, G9545)	1:10 000
14	N-cadherin (ab76057)	1:1000
15	Lamin A (ab26300)	1:1000
16	Emerin (ab40688)	1:500
17.	Lamin B Receptor (LBR) (ab32535)	1:500

2.3 Immunofluorescence assay

Cells grown on coverslips were washed with 1X PBS and fixed with 4% paraformaldehyde (Sigma, P6148) prepared in 1X PBS, pH 7.4 at RT for 10 min, washed thrice in 1X PBS (5 min each). Fixation was followed by permeabilization in 0.5% Triton-X-100 prepared in 1X PBS at RT for 10 min. Cells were blocked in 1% Bovine Serum Albumin (BSA) (Sigma, A2153) prepared in 1X PBS, for 30 min and washed three times with 1X PBS. Incubation with primary antibodies was performed in 0.1% BSA for 90 min at RT and with secondary antibodies for 60 min at RT, with washes in between using 1X PBS. Primary antibodies with the dilutions used are listed in **Table 2.2.** Secondary antibodies used were Alexa 488 (Invitrogen, A11034), 1:1000; Goat anti-Rabbit-Alexa 568 (Invitrogen, A11011), 1:1000; Goat anti-mouse-Alexa 488 (Invitrogen, A11029), 1:1000; Goat anti-mouse-Alexa 568 (Invitrogen, A11004), 1:1000. Cells were washed thrice in 1X PBST. Cells were counterstained with 4′,6-diamidino-2-phenylindole (DAPI) (Invitrogen, D1306) for 2 min at RT, washed with 1X PBS for 5 min and mounted in Slowfade Gold Antifade (Invitrogen, S36937). A positive control, no primary antibody with secondary antibody only controls and protein knockdown controls were used whenever necessary. Cells were imaged on a Zeiss LSM710 confocal microscope with 405, 458 and 561 nm laser lines, using a 63X oil immersion objective, NA 1.4 at 1X or 2.5X digital zoom. X–Y resolution was 512 X 512. Confocal z-stacks were collected at intervals of 0.34 μm.

Table 2.2: List of primary antibodies with their dilutions used for Immunofluorescence assay

S.No.	Antibody	Dilution
1	E-cadherin (ab1416)	1:500
2	Phalloidin Alexa Fluor 488 (A12379)	1:100
3	Lamin A (ab26300)	1:1000
4	γH2AX (ab26350)	1:750
5	N-cadherin (ab76057)	1:1000
6	Vimentin (Sigma, V2258)	1:500

7	Lamin A (ab26300)	1:1000
8	Lamin B1 (ab16048)	1:500
9	Lamin B2 (ab8983)	1:400
10	Emerin (ab40688)	1:500
11	Lamin B Receptor (LBR) (ab32535)	1:500
12	Nucleolin (ab13541)	1:500
13.	H3K4me3 (Sigma 07-473)	1:1000
14	H3K27me3 (Sigma 07-449)	1:1000
15	CTCF (ab70303)	1:500

2.4 RT-PCR

Total RNA was extracted using the Trizol method (Rio et al., 2010) from DLD1 and SW480 cells transfected independently with vector control and Twist1. cDNA was synthesized from 1 µg of total RNA with the Verso cDNA kit (AB-1453/B) using Oligo(dT) primers. cDNA was used as a template, and RT-PCR was carried out using primers designed to span intron-exon junctions **(Table 2.3)**. GAPDH was used as an internal control. Real-time quantitative PCR was performed in a 5 µl reaction mixture containing KAPA SYBR FAST qPCR Master Mix (2X) (KK4602, Merck) and 2 µM each of the forward and reverse primer using the Bio-Rad RT-PCR instrument (CFX96 Touch). Fold change in expression was calculated by double normalization of Ct values to the internal control (GAPDH) and empty vector control by the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001).

Table 2.3: List of RT-PCR primers with their sequences used in this study

S.No.	Gene Name	Primer sequence
1	TWIST1	Forward- 5'GCGCTGGGGAAGATCATC3' Reverse- 5'GGTCTGAATCTTGCTCAGCTT3'
2	BUB1	Forward- 5' TGTTGAGCAGGTTGTTATGTATTG3' Reverse- 5' GTCTGTCTTCATTTACCCATTGC3'
3	BUBR1	Forward- 5'CAGCCAGTTATGACACCATGTA3' Reverse- 5'TGATGGCTCTGAACCCTTTG3'
4	MAD1L1	Forward- 5' CCTTCAGACTTGGACTGTGT3' Reverse- 5' CATGGTTGCTTTCGCGATTAC3'
5	MAD2L1	Forward- 5' ACAGCTACGGTGACATTTCT3' Reverse- 5' GTCCCGACTCTTCCCATTT3'
6	AURKB	Forward- 5' CATCGTCAAGGTGGACCTAAAG3' Reverse- 5' GGGTTATGCCTGAGCAGTTT3'
7	GAPDH	Forward- 5' CGAGATCCCTCCAAAATCAAG3' Reverse- 5'GCAGAGATGATGACCCTTTT3'
8	CDH1	Forward- 5'CCAGTGAACAACGATGGCATT3' Reverse- 5'TGCTGCTTGGCCTCAAAAT3'
9	ZO1	Forward- 5'GAGGCAGCTCACATAATGCTA3' Reverse- 5'TTCCTCGGGATATGGATCCTT3'
10	OCLN	Forward- 5'CACGCCGGTTCCTGAAGT3' Reverse- 5'CGAGGCTGCCTGAAGTCATC3'
11.	CLDN1	Forward- 5'GCACCGGGCAGATCCA3' Reverse- 5'TTGCAATGTGCTGCTCAGATT3'
12.	CDH2	Forward- 5'CAGCAACGACGGGTTAGTC3' Reverse- 5'TGCAGCAACAGTAAGGACAAA3'
13.	VIM	Forward- 5'TCCAAACTTTTCCTCCCTGAA3' Reverse- 5'GGGTATCAACCAGAGGGAGTG3'
14.	FN1	Forward- 5'ATGGGAGAAGTATGTGCATGC3' Reverse- 5'TGGAAATGTGAGATGGCTGTG3'
15.	SNAIL1	Forward- 5'AATCGGAAGCCTAACTACAGC3' Reverse- 5'CAGAGTCCCAGATGAGCATTG3'
16.	ZEB1	Forward- 5'- GCACCTGAAGAGGACCAGAG-3' Reverse- 5' - TGCATCTGGTGTTCCATTTT-3'
17	LMNA	Forward-5'CCGCAAGACCCTTGACTCA-3' Reverse-5'TGGTATTGCGCGCTTTCAG-3'

18	LMNB1	Forward-5'CGACCAGCTGCTCCTCAACT-3' Reverse-5'CTTGATCTGGGCGCCATTA-3'
19	LMNB2	Forward-5'AGTTCACGCCCAAGTACATC-3' Reverse-5'CTTCACAGTCCTCATGGCC-3'
20	SUN1	Forward-5'CCGAGGGAGACTGACTTTATG-3' Reverse-5'CTTCTGGATGGCCTCAGATT-3'
21	SUN2	Forward-5'AGTCCTCTCAGGACCTTGAA-3' Reverse-5'ACCAGCGACTCACTGTAGTA-3'
22	EMD	Forward-5'CAGAGCAAGGGCTACAATGA-3' Reverse-5'CGTCAGCATCTGGGAATGAA-3'
23	MAN1	Forward- 5'TGCTTAGGTGTAGTGATGTTT3' Reverse- 5'CTGCCTTGTTTCCCTCCTCTT3'
24	NESP2	Forward-5'GTGAGTGTGGTTGACTCATCTC-3' Forward-5'CTCATAGGTGGCGCATTGT-3'
25	LBR	Forward- 5'CACAGTATAGCCTTCGTCCAA3' Reverse- 5'CAACAGGAAGAGGAACACAGG3'
26	CTCF	Forward- 5' F-CGTTACTGTGATGCTGTGTTTC-3' Reverse- 5' TCATGTGCCTCTCCTGTCTA-3'
27	IGFL1	Forward-5'TGTCGCAGTGTCAGCTAATG-3' Reverse-5'GGTAACACCAGCCTCTTTCT-3'
28	MAF	Forward-5' CAGGAGCAAAGCCATCCATA-3' Reverse-5'GCAAGCGCTGTTTCTCTTTAC-3'
29	COL4A1	Forward-5'TCTATGCACCGCTTCATCTC-3' Reverse-5'TCCTTCTTTCTCACCTCTTTCC-3'
30	IGBP5	Forward-5' GAGCAAGTCAAGATCGAGAGAG-3' Reverse-5'GGAGATGCGGGTGTGTTT-3'
31	IGBP7	Forward-5' GGGTGCTGGTATCTCCTCTA-3' Reverse-5'TGTAAGGCATCAACCACTGTAA-3'
32	SERPINE1	Forward-5'ACAGCTGAGGGACAAATTCC-3' Reverse-5'ACACCTGGGAGCTGTAGA-3'
33	IL11	Forward-5'ACAGCTGAGGGACAAATTCC-3' Reverse-5'ACACCTGGGAGCTGTAGA-3'
34	SPOCK1	Forward-5'CCGCTTTCGAGACGATGATTA-3' Reverse-5'CCTGGGTCACACACACTTT-3'
35	COL5A1	Forward-5'GCTCTCTTGTGGTGCTATCTATC-3' Reverse-5'ACAGACTGTGGAGGCAATAAC-3'

2.5 Flow cytometry

Cells were trypsinized, washed with 1X PBS and then fixed in chilled 70% ethanol. Ethanol was added dropwise to the pellet while vortexing. This ensured the fixation of all cells and minimized clumping. After chilling on ice for 15 min, the cells were centrifuged at 200g for 10 min. The pellet was resuspended in 1X PBS, subjected to RNaseA (Sigma, R6513-10MG) (10μg) treatment at 37°C for 45 min. Further, propidium iodide (Sigma P4170-10MG) (10 μg) was added to the samples. Cell suspensions were subsequently run on BD FACS Calibur^{TM} (BD Biosciences) and analyzed using Cell Quest Pro^{TM} software. Unstained cells (cells without propidium iodide) were used as controls to identify the cell population, determine the level of background fluorescence or autofluorescence and set the voltages and negative gates appropriately.

2.6 Metaphase spread preparation

Colcemid (Roche 10295892001) (1% v/v) was added to cells at ~60-70% confluency and incubated for 90 min at 37°C. Cells were harvested by trypsinization and centrifuged at 200g at 4°C for 10 min. The pellet was resuspended in 4 ml pre-warmed 0.075 M KCl and incubated at 37°C for 30 min. 4-5 drops of fixative [methanol:acetic acid (3:1)] were added and cells were centrifuged at 200g at 4°C for 10 min. The cell pellet was washed three times with a fixative solution. After the washes cells were suspended in fixative solution at an appropriate dilution and dropped onto clean glass slides. Metaphases were stained with DAPI (0.05 μg/ml in 2X SSC, pH 7.4).

2.7 Gene ontology

Gene lists were submitted to the online database DAVID (https://david.ncifcrf.gov/) to perform gene annotation enrichment and functional annotation clustering analyses. The output from the analysis consisted of GO categories from the Biological Process (BP), Molecular Function (MF) and Cellular Components (CC) sections. A graph was plotted for the $-\log_{10}$(p-value) of enrichment for all GO categories obtained.

2.8 STRING

A list of proteins was submitted to the online database STRING (https://string-db.org/) which predicts protein-protein interaction networks. Protein-protein interactions in *Homo Sapiens* for which there is experimental evidence was used as a source and a visual output of the network was generated.

2.9 Microscopy

Cells were imaged on Zeiss LSM700, LSM710 and LSM780 confocal microscopes with 405nm, 488 nm, 561 nm and 594 nm laser lines using a 63X Plan-Apochromat 1.4-numerical-aperture (NA) oil immersion objective at 1.0 to 2.5X digital zoom. Scanning was performed sequentially (*x-y*, 512 pixels by 512 pixels [1 pixel = 0.105 µm]), and z-stacks were collected at a step size of 0.34 µm and a pinhole size of 0.7µm (1 arbitrary unit [AU] for 405 nm laser line). The pixel depth was 8 bits, the line averaging was 4, and the scan speed was 10.

2.10 Image processing and analyses

Images were quantified using ImageJ software. E-cadherin levels were measured by tracing out E-cadherin-staining manually and intensity was measured along the traced line. For actin staining, the aspect ratio was calculated as a ratio of major axis/minor axis.

2.11 Statistical analysis and graphs

The number of technical replicates (n) and biological replicates (N) differs for each experiment. Statistical analysis was performed, and graphs were plotted with GraphPad Prism 6 software or Microsoft Excel. Statistical tests used have been mentioned for each experiment in the figure legends.

2.12 Methods specific to Chapter 3

2.12.1 TGF-β1 treatment for EMT induction

A549 cells (~0.2 X 10^6) were seeded into 35 mm dishes. Upon attachment, the medium was changed to medium containing DMEM with 1% FBS and penicillin (100 units/ml)-streptomycin (100 μg/ml). TGFβ1 (PeproTech,100-21C) was added at a final concentration of 5 ng/ml in this reduced serum medium and incubated for 48h. EMT induction was assessed and further experiments were conducted.

2.12.2 Reversal of EMT

TGF-β1 containing medium was removed from cells which had undergone 48h of TGF-β1 induced EMT and medium was changed to DMEM containing 1% FBS and penicillin (100 units/ml)-streptomycin (100 μg/ml). Media was replenished with the reduced serum. Cells were collected at 24h, 48h, 72h and 96h and reversal of EMT (MET) was assessed by western blotting.

2.12.3 Inhibition of EMT

A549 cells (~0.2 X 10^6) were seeded into each well of a 6-well plate. After attachment of cells overnight, the medium was changed to DMEM containing 1% FBS and penicillin (100 units/ml)-streptomycin (100 μg/ml). SB 431542 **i**nhibitor was added at 10 uM to the medium and cells were incubated at 37°C for 1h. Following this, TGF-β1 was added at a final concentration of 5 ng/ml and incubated for 48h. Untreated, TGF-β1 treated, DMSO (vehicle control), DMSO+TGF-β1 treated and SB431542 only served as controls. Cell morphology and EMT marker expression were examined as a readout of EMT induction.

2.12.4 Scratch assay

Cells (~0.3X10^6) were seeded in each well of a 6-well plate. TGF-β1 treatment was performed as mentioned above. After 48h of TGF-β1 treatment, a scratch was created in the monolayer of cells. Cells were washed twice with 1X PBS to remove cells dislodged due to the scratch. Plain medium and TGF-β1 containing medium were replenished in control and test respectively.

Bright-field images were acquired along the length of the scratch at this 0h time point. Three images were acquired for each scratch at every time point. In order to monitor the migration of cells and wound healing, the scratched was imaged after every 2h up to 24h. The area occupied by the cells was analysed from the images of the scratch at every time point and represented as percentage of wound closure. Image analysis was done using Image-Pro Plus Software.

2.12.5 Three-dimensional fluorescence in situ hybridization (3D-FISH)

Probe preparation

Bacterial artificial chromosome (BAC) DNA for *E-cadherin* (RP11- 354M1) and *N-cadherin* (RP11-643F8) was purchased from CHORI BAC-PAC Resources. BAC DNA extraction was done using Hi Pure Plasmid DNA Extraction Kit (Invitrogen K210017). Nick translation of BAC DNA was performed using the Nick Translation Kit (Roche 11 745 808 910). Nick translation reaction was carried out at 15°C for 90 min and the reaction was stopped by adding 2 µl of 0.5M EDTA (pH 8.0) at 65°C for 10 min. Nick-translated DNA was precipitated using ethanol precipitation method at -80° C overnight followed by centrifugation at 300g for 45min. Precipitated DNA was resuspended in hybridization mix (50% deionized Formamide + Master mix containing 20% Dextran Sulphate and 4X SSC, 0.1 mg Salmon Sperm DNA in 2X SSC solution, pH 7.4)

Fixation & Hybridization

Cells (~0.2X10^6) were seeded on 22 X 22 mm^2 coverslips. After 48h of TGFβ1 treatment, the cells were washed with ice-cold 1X PBS and treated with cytoskeletal (CSK) digestion buffer (0.1 M NaCl, 0.3 M sucrose, 3 mM MgCl$_2$, 10 mM PIPES (pH 7.4), 0.5% Triton X-100) for 5 min followed by fixation with 4% PFA in 1X PBS (pH 7.4) for 10 min at RT. The cells were permeabilized in 0.5% Triton X-100 (prepared in 1X PBS) for 10 min and incubated in 20% glycerol (prepared in 1X PBS) for 60 min followed by four freeze-thaw cycles in liquid nitrogen. The cells were washed three times with 1X PBS and treated with 0.1 N HCl for 10 min followed by three washes in 1X PBS for 5 min each. The cells were incubated in 50%

formamide (FA)/ 2X saline sodium citrate (SSC) (pH 7.4) overnight at 4°C or until used for hybridization. Cells were hybridized with a nick-translated BAC DNA probe (3 µl) and/or 3 µl of human whole chromosome paint (Applied Spectral Imaging (ASI, Israel). Post hybridization, coverslips were washed in 50% FA/2X SSC (pH 7.4) thrice for 5 min each at 45°C, followed by three washes for 5 min each in 0.1X SSC at 60°C. Coverslips were then counterstained with DAPI for 2 min, washed in 2X SSC and mounted in Slowfade Gold antifade (Invitrogen, S36937).

Image analysis

(i) Radial distance measurements of chromosome territories

Image-Pro Plus (v 7.1), Media Cybernetics, USA was used to perform 3D radial position measurements for CTs. Optical sections (z=0.34 µm) of confocal microscopy images of hybridized nuclei were reconstructed into 3D rendering for each nucleus independently. The acquired images were thresholded and surface rendered for each of the channels - red, green, and blue. Three-dimensional distance measurements of CTs were performed using the geometric center of the DAPI-stained nucleus (blue channel) and the CTs (red and green channels), determined using plugins from the software. The distance between the geometric center of the nucleus (A) and that of the territory (B) was measured (X). The vector from the geometric center of the nucleus (A) to the geometric center of the CT (B) was extended to a third collinear point at the nuclear periphery (C). The distance between the geometric center of the nucleus (A) and (C) was calculated (Y). The relative distance of a CT from the center of the nucleus was calculated as a percentage of the total distance from the center of the nucleus to the nuclear periphery (Y), %radial distance $(RD) = (X/Y) * 100$ (Tanabe et al. 2002). The radial distance measurements have been represented as (i) dot scatter plots of radial distances and (ii) binned into five shells of ~20% of the nuclear sub-volume (0% - nuclear center, 100% - nuclear periphery).

(ii) Gene loci measurement from the nuclear periphery

Distances of gene loci from the nuclear periphery were measured (in µm) in 3D using the boundary of the DAPI signal as a marker of the nuclear periphery (Shachar et al., 2015). Huygens Professional software was used for distance measurement. 3D reconstruction was

performed using surface rendering for the nucleus (blue channel), gene locus and CT. Identical threshold and seeding levels were used for surface rendering for all images. The nucleus (DAPI) was selected as an anchor. Center of mass (CM) was determined for the gene locus signal of interest, and the shortest distance between the CM and surface of the anchor (DAPI) was measured using the 'object analyzer advanced tool' in Huygens software.

2.12.6 siRNA mediated knockdown

Transient siRNA transfections were performed using Lipofectamine RNAimax reagent (Invitrogen, 13778) in reduced serum Opti-MEM (Gibco, 31985) for 6h, after which cells were transferred to complete medium and incubated for 48h. siLacZ, siNeg (Dharmacon ON_TARGET PLUS Non-targeting control pool) or respective scrambled siRNAs were used as negative controls.

Table 2.4: List of siRNAs and their sequences used in this study

S.No.	siRNA	siRNA sequence
1	siLMNA/C	5'-CAGUCUGCUGAGAGGAACA-3'
2	siLMNA-scramble (control)	5'-GAUGAGGCGGUUAGAUGUA-3'
3	siLMNB1	5'-AGACAAAGAGAGAGAGAUG-3'
4	CTCF	5'ACAAGAATGAGAAGCGCTT3'
5	siLMNB2	5'-GAGCAGGAGAUGACGGAGA-3'
6	siLacZ (control)	5'-CGUACGCGGAAUACUUCGA-3'
7	siNeg (control)	ON-TARGETplus Non-targeting control pool (D-001810-02-20 DHARMACON)

2.12.7 Image processing and analyse

Anisotropy score for the actin organization was calculated using the 'FibrilTool', an ImageJ plug-in (Boudaoud et al., 2014). A score ~0 denotes isotropy while ~1 denotes anisotropy (parallel arrangement of actin filaments). Nuclear topology analyses and measurements like nuclear volume (object counter 3D plugin) and area were performed using ImageJ software. Classification of nuclear morphologies was done manually. Levels of modifications were quantified as total fluorescence intensity in the nucleus using ImageJ. Line-scan analyses for examining the expression of nuclear envelope proteins was also done using ImageJ software. The scheme of analyses is represented in **Figure 3.11.**

2.12.8 Analyses of data from The Cancer Genome Atlas (TCGA)

miRNA expression levels across cancer subtypes from TCGA data were analysed using online portal FireBrowse (http://firebrowse.org/). UCSC Xena Browse was used to examine correlations between expression levels of a set of genes across cancer datasets. We used the Pan-Cancer (Pan_Can) TCGA dataset for our analyses (http://xena.ucsc.edu/).

2.12.9 HiC data analyses

3D-genome Interaction Viewer and database (3DIV) was used to access and visualize TAD organizations at specific genome sites (http://kobic.kr/3div/). Default parameters like interaction range=2Mbp, TopDom w=20 and resolution=20,000bp were used.

2.13 Methods specific to Chapter 4

2.13.1 DNA transfection

For plasmids midipreps ~50 ml bacterial cultures were grown overnight at 37°C at 180 rpm shaking. Extraction of DNA was performed using Exprep plasmid SV kit (GeneAll), eluted in nuclease-free water (NFW) and used for transfection. All plasmids used in this study were confirmed by sequencing.

DLD1 and SW480 cells were seeded at ~70% confluency and transfected with ~2 μg of pBp-mTwist1 vector and pBp-Empty vector as control (Gift from Annapoorni Rangarajan, IISc Bengaluru, India and Robert Weinberg, MIT, USA) using LTX and PLUS (Invitrogen 15338-100) reagents in OptiMEM (Gibco-31985070) for 6h after which cells were cultured in complete growth medium. After 24h of transfection, 1 μg/ml and 0.8 μg/ml puromycin (Gibco A11138) were added to DLD1 and SW480 cells respectively to select for transfected cells and cultured for another 48h.

2.13.2 Array comparative genomic hybridisation (array CGH)

The DNA quality, purity and quantity control test was determined using Nanodrop-1000 (JH Bio, USA) followed by a gel QC. For sample labelling, to ensure high quality of aCGH data, Agilent Direct method has been used for the sample processing. About 1 ug of control and test DNA was used for the restriction digestion in the master mix containing *Alu* I and *Rsa* I restriction enzymes, as per manufacturer's recommendation. The samples were incubated at 37°C for 2 hours followed by heat inactivation of enzymes at 65°C for 20 minutes. To confirm the efficiency of restriction enzymes to obtain fragments of size 200-500 bp, about 2 μL of the digested gDNA was tested on a 0.8% agarose gel. Labelling of samples was done by random priming method, in which the random hexamers, Cy3-dUTP & Cy5-dUTP, dNTP, Buffer and Klenow enzyme were used. Briefly, 1X Random primer mix was added to each of 26 μl digested control and test samples. The DNA was denatured at 95°C for 3 minutes followed by snap chill on ice for 5 minutes. Master mix for Cy3 and Cy5 dNTPs were done separately to ensure that the control sample was labelled with Cy3 and test sample with Cy5 respectively. About 19 ul of labelling master mix prepared as per manufacturer's recommendation was added to the denatured

control and test DNA sample and incubated at 37°C for 2 hours followed by enzyme heat inactivation at 65°C for 10 minutes. The labelled samples were cleaned up by an Amicon 30kDa filter size-exclusion filter. The sample volume was adjusted with respect to array format. The specific activity and yield were optimum to proceed for the hybridization. Equal amount of labelled Test & Control DNA samples was added into a fresh tube containing 50 µl of Human Cot-1 DNA (1mg/ml), 52 ul of Agilent 10X blocking agent and 260 ul Agilent 2X hybridization buffer. The total hybridization volume was 520 ul. The above hybridization mix was denatured at 95°C for 3 minutes and incubated the microfuge tubes at 37°C for 30 minutes. The samples were hybridized at 65°C for 40 hours in the hybridization chamber. After hybridization, the slides were washed using aCGH Wash Buffer 1 (Agilent Technologies, Part Number 5188-5221) at room temperature for 5 minutes and aCGH Wash Buffer 2 (Agilent Technologies, Part Number 5188-5222) at 37°C for 1 minute. The slides were then washed with Acetonitrile for 10 seconds. The microarray slide was scanned using Agilent Scanner (Agilent Technologies, Part Number G2565CA). Image analysis was performed using Agilent Feature Extraction software. Feature extracted raw data was normalized by applying LOWESS normalization method & further data analysis was carried out using Agilent CytoGenomics 3.01.1 software and Excel. Agilent CytoGenomics supports CGH arrays, which allows the detection of regions of loss or gain. Aberration Detection Method II (ADM-2) algorithm was applied to identify significant regions having amplifications and deletions among each of the samples. GC Correction algorithm was then applied to correct aCGH log-ratio data for the presence of "wavy" artefacts. Penetrance analysis was performed to find the percentage of samples that share aberrations in a particular genomic region among multiple samples (Amplification and deletions are considered separately). Common aberrations among the samples were identified. Differential aberration analysis for two groups was performed and specific aberration for each group was determined. Graphical representation has been done using the Human UCSC genome browser by loading the data in wiggle file format. Various genome view, chromosome view, gene view plots were generated for the amplification and deletions data with respect to all samples. We acknowledge Genotypic Technology Private Limited, Bengaluru, India for sample processing and data analysis.

2.13.3 RAndomized CIrcuit PErturbation (RACIPE)

The network was simulated using the tool "RAndomized CIrcuit PErturbation (RACIPE)" (Huang et al., 2017), available for download at https://github.com/simonhb1990/RACIPE-1.0. Briefly, RACIPE models a given regulatory network using a system of Ordinary Differential Equations. Each equation in the system represents the dynamics of one node in the network and is of the following form:

$$\frac{dX_i}{dt} = g_{X_i} * \prod_{j=1}^{n} H^S(X_j, X_i) - k_{X_i} * X_i$$

Where g and k represent the production and degradation of a node and $H^S(X_j, X_i)$ is a modified Hill function that represents the regulation of X_i by X_j. Further details are available in (Huang et al., 2017). The tool then samples multiple parameter sets randomly via a uniform distribution from a pre-defined range of parameters. As the exact kinetic parameters are often not available for biological systems, this sampling allows us to obtain a generic behaviour of the network, accounting for cell-cell variability in kinetic parameters. For each such parameter set, the system is simulated at multiple initial conditions to identify the number of steady states. For the current analysis, 10000 parameter sets were sampled, and 100 random initial conditions were chosen for each parameter set. The ODE's were integrated using Euler's method of numerical integration. All parameter ranges used in this study are defaults of the tool. Linear regression was used to fit coupled gene expression data obtained from RACIPE to a line. Corresponding p-value ranges are reported.

2.13.4 TCGA expression analysis

Gene expression (RSEM gene-normalized, version 2016_01_28) and somatic mutation data (MC3) of TCGA samples ($n = 8657$) across 30 tumor types were downloaded from Firebrowse server (http://firebrowse.org). The correlation coefficient between TWIST1 expression and other gene expressions/mutation burden/copy number alterations and its significance were computed using iteratively reweighted least-squares approach. To adjust for multiple hypothesis testing, Bonferroni correction on P values per gene set was performed, and $q < 0.01$ was considered as significant. All plots were generated using the Seaborn package in Python.

The results shown here are in whole or part based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

2.13.5 Image Processing and analyses

For analyses of γH2AX foci, immunofluorescence images were thresholded for each nucleus counterstained with DAPI, and the 'find maxima' function was used to enumerate the number of γH2AX foci per nucleus.

Chapter 3 Effect of TGF-β1 induced EMT on nuclear & genome organization

3.1 Introduction

Growth factors such as Fibroblast Growth Factor (FGF), Hepatic Growth Factor (HGF), Platelet Derived Growth Factor (PDGF) and Transforming Growth Factor-β (TGF-β) induce EMT (Moustakas and Heldin, 2007). TGF-β is an important EMT inducer during embryogenesis, fibrosis and cancer progression (Miyazono, 2009). TGF-β induced EMT was first identified in the normal murine mammary gland epithelial cell line NMuMG (Miettinen et al., 1994). TGF-β has three isoforms TGF-β1, TGF-β2 and TGF-β3, each of which binds to the same cell surface receptors and induces EMT (Moustakas and Heldin, 2007), (Piek et al., 1999) (Valcourt et al., 2005).

TGF-β exerts its functions via two cell surface receptors - TGF-β type I receptor (TβRI) and TGF-β type II receptor (TβRII) (Massagué, 1992). TβRI and TβRII are structurally similar, consisting of an extracellular domain, a single transmembrane domain and an intracellular domain with kinase activity. TGF-β first binds to TβRII, which then recruits TβRI, forming a hetero-tetrameric complex - composed of two TβRIIs and two TβRIs. TGF-β binds to and activates TβRII by phosphorylation. Activated TβRII further activates TβRI by phosphorylation. This receptor complex in turn activates Smad proteins, which are intracellular mediators of the TGF-β signaling cascade (Heldin and Moustakas, 2016).

The TGF-β signaling pathway elicits (i) The Smad-dependent signaling and (ii) The Non-Smad/Smad independent signaling. Smads are intracellular transcriptional effectors of TGF-β signaling. The Smad family are categorized as (i) Receptor-regulated or (R-) Smads (R-Smad1, -2, -3, -5 and -8) that are phosphorylated by activated TβRI kinases (ii) common (Co-) Smad (Smad4) that forms heteromeric complexes with activated R-Smads and (iii) inhibitory I-Smads (I-Smads 6/7), which antagonize canonical Smad signaling (Heldin and Moustakas, 2016).
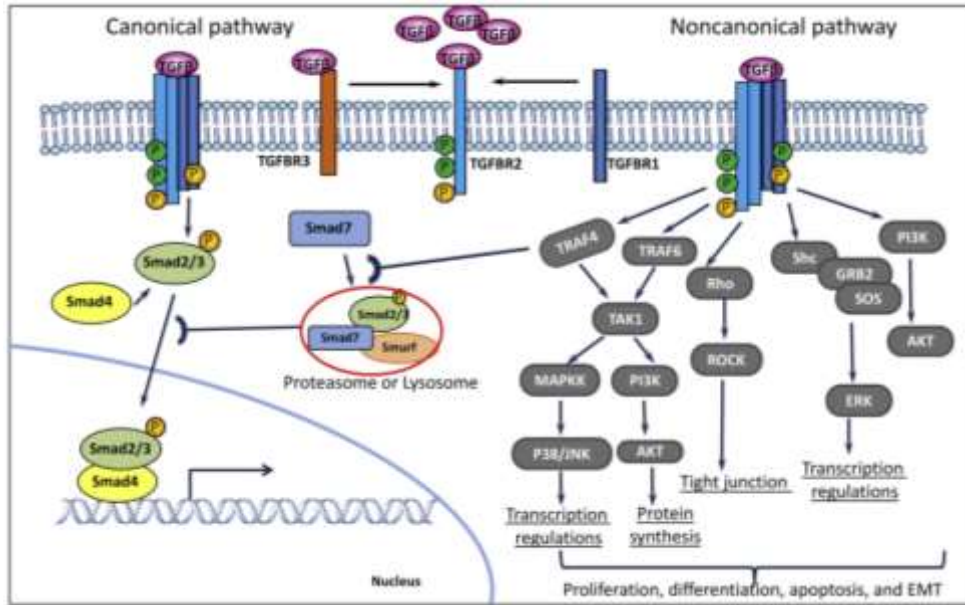
Figure 3.1

Fig.3.1: Canonical (Smad dependent) and Non-Canonical (Smad independent) TGF-β signaling
Reprinted with permission from (Vander Ark et al., 2018).

In the Smad dependent pathway, the activated receptors phosphorylate Smad2 and Smad3. Activated Smads are transported into the nucleus where they form a trimer with Smad4 (Shi and Massagué, 2003). The complex then associates with different transcription factors and cofactors, and selectively binds certain genes, in a context-specific manner (Feng and Derynck, 2005). Each Smad2-Smad3-Smad4-Transcription factor complex binds to specific sets of genes. This complex binds to a cognate sequence in the regulatory region of the target genes. Through this combinatorial interaction with different transcription factors, a common TGF-β stimulus can activate or repress hundreds of target genes (Massagué et al., 2005).

Activated TGF-β receptor directly activates the MAP kinase (MAPK) pathway, Rho-like GTPase signaling pathways and the phosphatidylinositol-3-kinase (PI3K)/AKT pathway in a Smad-independent manner. Activation of non-Smad signaling is context-dependent and these pathways also cross-talk with the canonical Smad pathway (Zhang, 2009).

TGF-β is a tumor suppressor in normal and premalignant epithelial cell types and can regulate cell cycle genes. For example, TGF-β arrests cell cycle at G1 by inducing cyclin-dependent kinase (CDK) inhibitors (p15INK4b, p21WAF1 and p27KIP1) (Baghdassarian and Ffrench, 1996). Additionally, TGF-β signaling is also associated with apoptosis through the induction of

reactive oxygen species (ROS), (Yan et al., 2014), modulating epigenetic regulators (DNMTs) (Cardenas et al., 2014) and inducing telomere shortening via human telomerase reverse transcriptase (hTERT) (Cassar et al., 2017) to regulate cell proliferation.

Mutations in TGF-β signaling components are associated with loss of function. SMAD4 is frequently mutated in gastrointestinal cancers (Hahn et al., 1996). TGF-β1 signaling components such as the TGF-β receptors (Markowitz et al., 1995) and other Smads (Smad2 and Smad3) are mutated in bladder, colon, breast, esophageal, stomach, brain, liver, and lung cancers (Macías-Silva et al., 1996). Epigenomic alterations also promote the oncogenic potential of TGF-β1. For example, the increased methylation of the TβRI promoter decreases TβRI activity in gastrointestinal tumors (Pinto et al., 2003). As cancer progresses, cancer cells remain responsive to TGF-β but acquire resistance to its cytostatic effects. Furthermore, culturing transformed epithelial cell lines in the presence of TGF-β promotes EMT (Tretbar et al., 2019). The reason for this switch from TGF-β1 being a tumor suppressor to being a promoter of cancers remains elusive.

Nuclear envelope proteins regulate TGF-β signaling. MAN1 – an inner nuclear membrane protein binds to and sequesters Smad2/3 at the nuclear envelope. This prevents Smad from binding to other transcription factors. MAN1 also binds a de-phosphorylase-'PPM1A', dephosphorylating pSmad2/3, thereby inhibiting TGFβ signaling. Furthermore, overexpression of MAN1 in mink lung epithelial cell line (Mv1Lu) prevents TGFβ mediated regulation of cell proliferation (Lin et al., 2005). Another LINC complex protein of the outer nuclear membrane, Nesprin2, affects the translocation kinetics of pSmad2/3 into the nucleus. Nesprin2 knockdown increases the time required for the translocation of pSmad2/3 from 5 minutes to 10 minutes in HaCaT cells (Rashmi et al., 2012). Furthermore, LMNA[-/-] MEFs show increased proliferation upon TGF-β1 treatment. Rescue assays with full-length LMNA in a LMNA[-/-] background, restores cell cycle arrest by TGF-β1. LMNA interacts with PP2A and hypo-phosphorylates retinoblastoma protein (pRB) to regulate cell cycle arrest (Van Berlo et al., 2005). Thus, it is evident that nuclear architecture proteins also impact and regulate TGF-β signaling pathway. However, these studies have focused on the role of nuclear architecture proteins in regulating the TGF-β mediated cell cycle in non-transformed cells. The regulatory role of nuclear architecture proteins in TGF-β induced EMT remains unclear.

Nuclear envelope proteins function as a signaling hub and are involved in genome organization. Both these roles of the nuclear envelope are less studied in the context of TGF-β1 induced EMT. Here we investigated the mechanisms of molecular cross-talk between nuclear envelope factors and genome organizers in the regulation of EMT and examined the impact of EMT on the nuclear landscape.

3.2 Results

3.2.1 TGF-β1 induced Epithelial to Mesenchymal Transition (EMT) in lung adenocarcinoma

3.2.1.1 Impact of TGF-β1 treatment on cell morphology & actin organization

Cells were cultured with reduced serum (1% FBS) and treated with TGF-β1 (5ng/ml) for 48h. While cells cultured in 1% FBS medium served as controls. Consistent with EMT induction, the majority of TGF-β1-treated A549 cells (~80%) showed an elongated, spindle-shaped morphology, while control A549 cells had predominantly cobble-shaped morphology **(Fig.3.2A)**. Independent staining with actin also showed a cobblestone morphology in control cells, while TGF-β1 treatment showed spindle-shaped cells, corroborating EMT induction **(Fig.3.2B)**. The elongated morphology of cells was quantified by measuring the aspect ratio of cells (length/breadth). TGF-β1 treatment showed a marked increase in the aspect ratio of cells (Untreated cells M=1.5 & TGF-β1 treated cells M=2.9) **(Fig.3.2C)**.
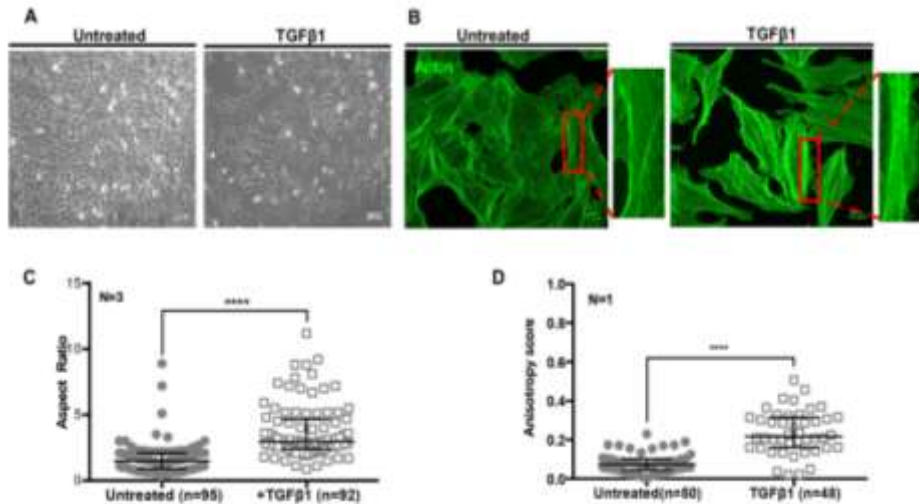
Figure 3.2

Fig.3.2: Effect of TGF-β1 treatment on cell morphology (A) Bright-field images of untreated A549 cells showing cobble-stone morphology and TGF-β1-treated cells showing spindle-shaped morphology. Scale bar ~200μm **(B)** Immunofluorescence images of phalloidin-stained actin in untreated and TGF-β1-treated A549 cells showing actin stress fibres. Scale bar ~10μm **(C)** Quantification of aspect ratio of untreated (median=1.5) and TGF-β1-treated (median=2.9), Median & IQR **(D)** Quantification of anisotropy score for actin organization in untreated (median=0.1), and TGF-β1-treated (median=0.2) A549 cells which exhibit parallel arrangement of actin stress fibres upon TGF-β1 treatment, Median & IQR (Mann Whitney test, two-tailed, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates, n:number of cells examined.

Interestingly, actin organization was altered upon TGF-β1 treatment wherein, actin assembled into thick parallel bundles throughout the cell in the form of stress fibres. In contrast, in untreated (control cells), filamentous actin is largely localized at cell-cell junctions **(Fig.3.2B)**. The altered organization of actin was also recapitulated as an increase in anisotropy score upon TGF-β1 treatment **(Fig.3.2D)**. Thus, TGF-β1 treatment-induced cell elongation, and actin reorganization, characterises mesenchymal cells.

3.2.1.2 Effect of TGF-β1 treatment on expression profile of EMT markers

Epithelial and mesenchymal cells are characterized by the expression of a distinct subset of proteins that are employed as markers for each of these cell types (Mani et al., 2008). We examined gene expression profiles of epithelial and mesenchymal markers upon TGF-β1 treatment of A549 cells. E-cadherin - a bona fide epithelial marker along with Claudin1 was downregulated upon TGF-β1 treatment (~0.2 ± 0.06 and ~0.5 ± 0.1 respectively) **(Fig.3.3A)**.

Amongst the mesenchymal markers, N-cadherin showed a concomitant increase in transcript levels upon TGF-β1 treatment (~2.3 ± 0.4) **(Fig.3.3B)**.

Figure 3.3

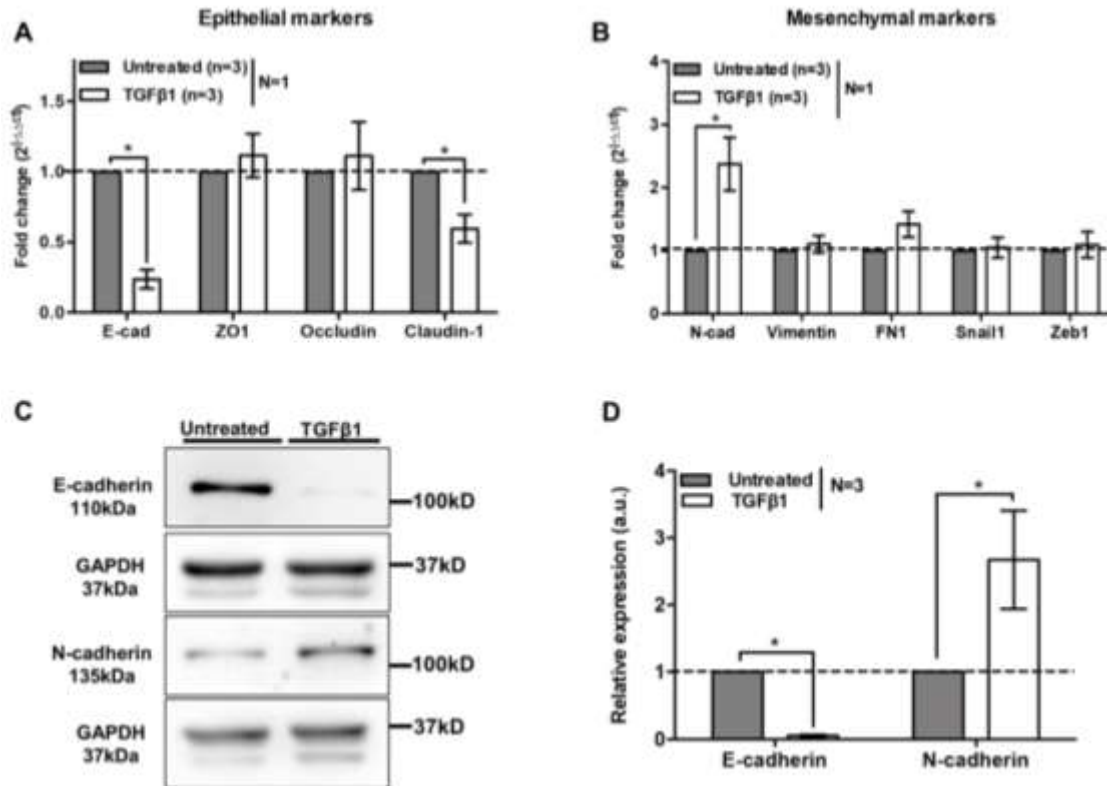**Fig.3.3: Expression of EMT markers upon TGF-β1 treatment of A549 cells (A)** qRT-PCR of epithelial markers in A549 cells upon TGF-β1 treatment shows downregulation of E-cadherin & Claudin **(B)** qRT-PCR of mesenchymal markers in A549 cells upon TGF-β1 treatment shows upregulation of N-cadherin. (unpaired t-test, Mean±SD, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates, n: technical replicates **(C)** Representative immunoblot for epithelial and mesenchymal markers upon TGF-β1 treatment. **(D)** Quantification of western blot shown in (C) shows downregulation of epithelial marker (E-cadherin) and upregulation of mesenchymal marker (N-cadherin) (unpaired t-test, Mean± SD, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates.

Similar expression changes were also detected at the protein level. E-cadherin levels (Mean= 0.05 ± 0.001) showed a decrease in protein levels after 48h of TGF-β1 treatment in A549 cells **(Fig.3.3C&D)**. Also, TGF-β1 treatment significantly increased the N-cadherin protein levels (Mean = 2.6 ± 0.4) **(Fig.3.3C&D)**. In summary, EMT-associated 'cadherin switch' was induced in A549 lung cancer cells with 48h of TGF-β1 treatment- a key hallmark of epithelial to mesenchymal transitions.

We also examined expression levels of the epithelial marker E-cadherin and mesenchymal markers N-cadherin and Vimentin by immunostaining at the single cell level **(Fig.3.4)**. Upon TGF-β1 treatment, E-cadherin levels were hardly detectable, while N-cadherin showed an increase **(Fig.3.4C&D)**. Along with being present at cell-cell junctions, N-cadherin staining was also observed in nuclei of A549 cells. Aberrant nuclear staining for N-cadherin has been observed in nasopharyngeal carcinoma and is associated with poor prognosis (Luo et al., 2012). The 'no primary antibody' controls do not show any non-specific binding of the fluorescently-tagged secondary antibody, demonstrating the specificity of the N-cadherin antibody. However, to validate the nuclear localization of N-cadherin in A549 cells, we need to perform N-cadherin staining in another cell line or use another N-cadherin antibody for staining A549 cells.

Furthermore, Vimentin - another mesenchymal marker, showed very little change in expression at the single cell level **(Fig.3.4E)**.
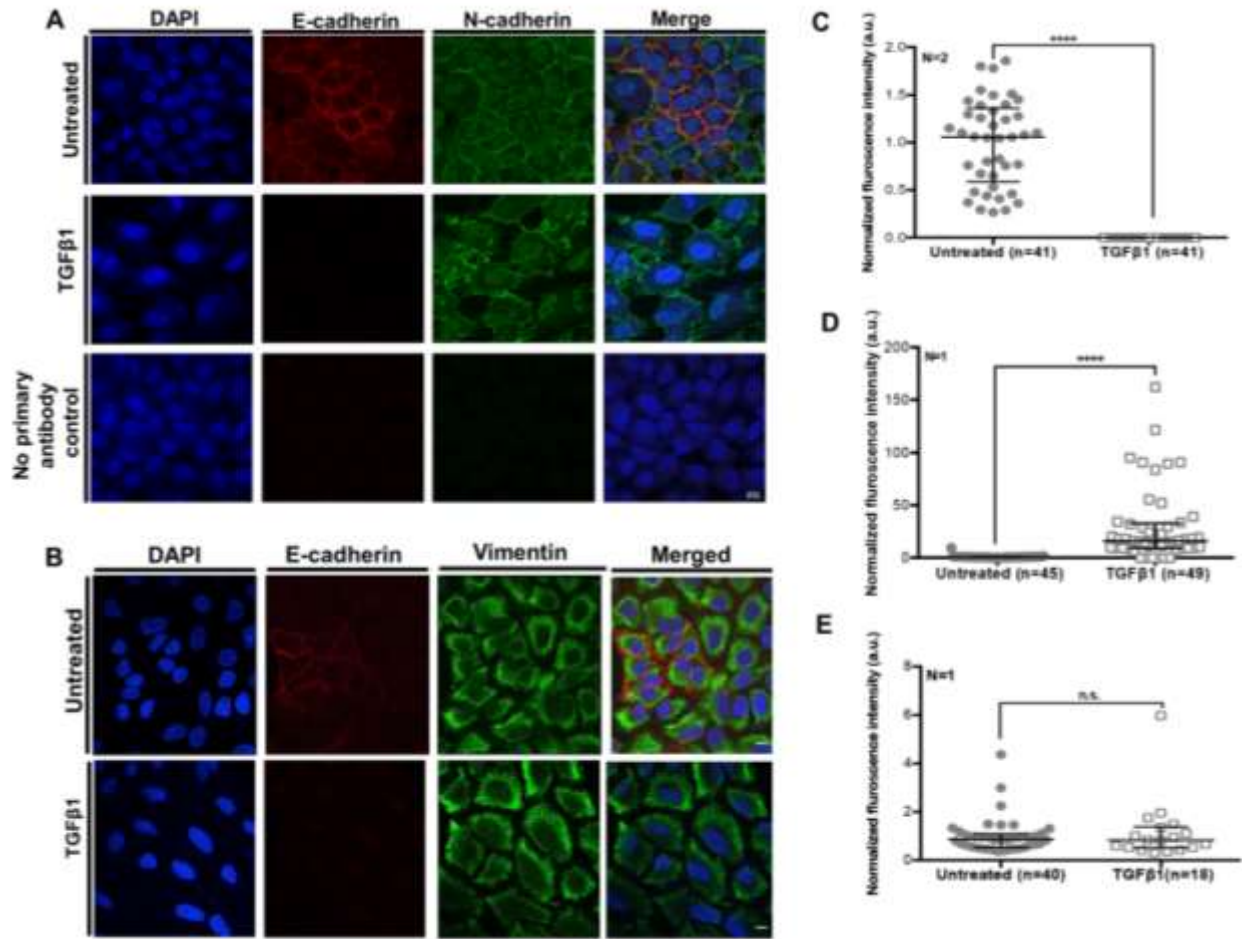
Figure 3.4

Fig.3.4: Expression levels of epithelial and mesenchymal markers in A549 cells ±TGF-β1 treatment (A) Immunofluorescence staining for E-cadherin (red) and N-cadherin (green) in untreated and TGF-β1-treated A549 cells **(B)** Immunofluorescence staining for E-cadherin (red) and Vimentin (green) in untreated and TGF-β1-treated A549 cells ~10μm **(C)** Quantification of immunofluorescence intensities of E-cadherin **(D)** Quantification of immunofluorescence intensities of N-cadherin **(E)** Quantification of immunofluorescence intensities of Vimentin ±TGF-β1 (Mann Whitney test, Median & IQR *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates, n: number of cells.

Out of the total population of cells, ~56% of the cells were positive for both E-cadherin and N-cadherin, whereas the remaining ~44% cells stained positive for N-cadherin alone **(Fig.3.5)**. Furthermore, in addition to the marginal membrane localization of N-cadherin in A549 cells, it also showed nuclear localization **(Fig.3.5A)**.
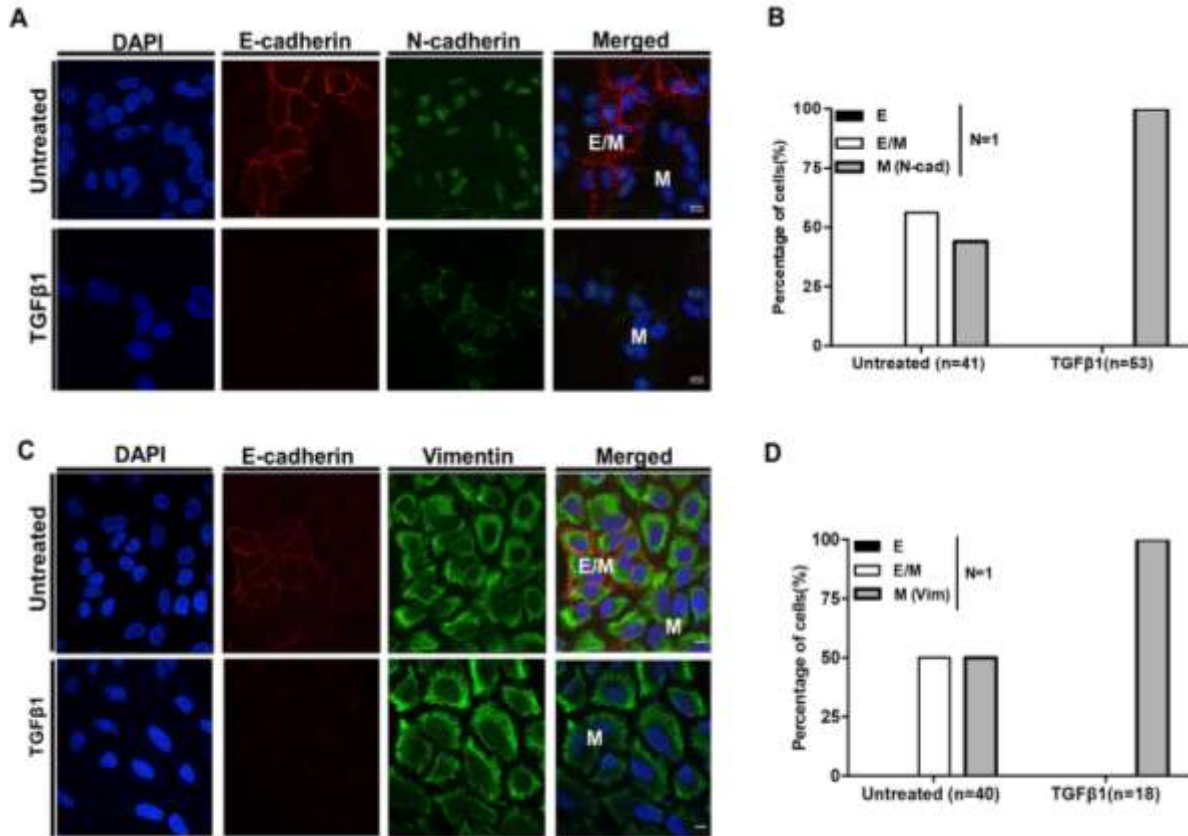
Figure 3.5

Fig.3.5: Distribution of epithelial and mesenchymal cells in A549 cells ±TGF-β1 treatment (A) Immunofluorescence staining for E-cadherin (red) and N-cadherin (green) in untreated and TGF-β1-treated A549 cells (B) Quantification of percentage of cells ±TGF-β1 expressing E-cadherin & N-cadherin shown in (A) (C) Immunofluorescence staining for E-cadherin (red) and Vimentin (green) in untreated and TGF-β1-treated A549 cells ~10μm (D) Quantification of percentage of cells ±TGF-β1 expressing E-cadherin & N-cadherin shown in (C) E/M=EMT hybrid cells, M=mesenchymal cells. (Mann Whitney test, Median & IQR, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates, n: number of cells.

It is unclear if, like E-cadherin, the aberrant nuclear localization of N-cadherin is also associated with lung tumorigenesis (Su et al., 2015). A549 cells stained positive for both E-cadherin and Vimentin (~50%), while ~50% cells stained positive for Vimentin alone. Therefore A549 is a hybrid cell line expressing both epithelial and mesenchymal markers. While TGF-β1 treatment converted these cells into mesenchymal cells with hardly any detectable levels of an epithelial marker, E-cadherin.

3.2.1.3 Effect of TGF-β1 treatment on cell migration

EMT is associated with the functional acquisition of enhanced cell migration. We investigated the impact of TGF-β1 treatment on cell migration using the scratch wound assay, for which we created a scratch, within a confluent monolayer of TGF-β1-treated cells. Wound healing was recorded for ~24h **(Fig.3.6A)**. Wound healing assay demonstrated that TGF-β1 treatment showed enhanced wound healing as a consequence of cell migration as compared to untreated control cells **(Fig.3.6B)**.
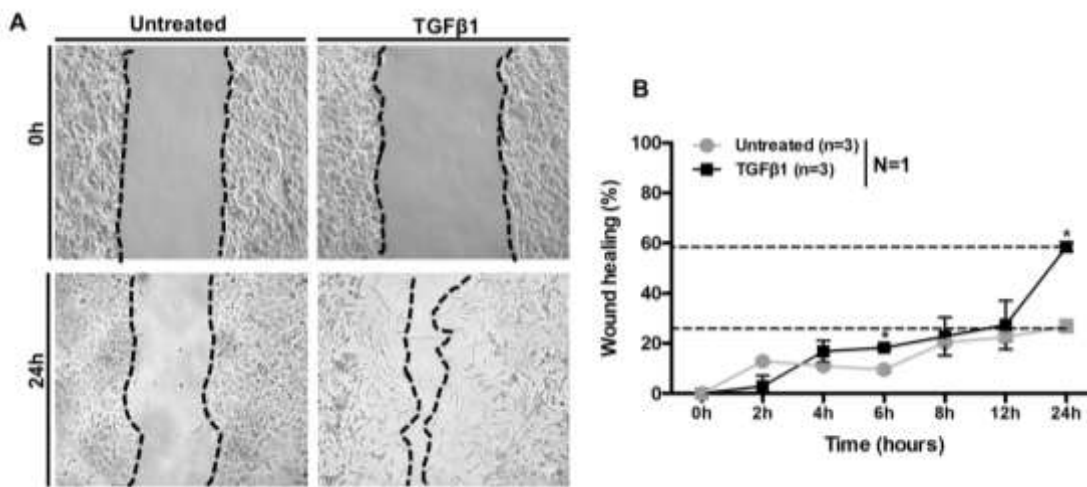
Figure 3.6

Fig.3.6: Effect of TGF-β1 treatment on cell migration (A) Bright-field images for scratch wound assay for untreated and TGF-β1-treated A549 cells at 0h & 24h time points after creating the scratch ~200μm **(B)** Quantification of wound healing across time for untreated and TGF-β1-treated A549 cells show enhanced cell migration upon TGF-β1-treatment. (Mean ±SD, unpaired t-test performed at every time point n: number of fields imaged at every time point, N: independent biological replicates).

Therefore, a range of attributes such as (i) spindle-shaped cells (ii) actin stress fibre formation (iii) downregulation of epithelial markers (E-cadherin & Claudin1) (iv) upregulation of mesenchymal marker (N-cadherin) (v) enhanced cell migration, consistent with EMT, were observed in TGF-β1 treated A549 lung cancer cells.

Of note, withdrawal of TGF-β1 restored E-cadherin levels, with a decrease in N-cadherin levels, revealing mesenchymal to epithelial transition (MET) **(Fig.3.7A)**. Furthermore, treatment with TGF-β receptor kinase inhibitor SB-431542 showed an inhibition of transitions in cell morphology and cadherin switch **(Fig.3.7B&C)**.
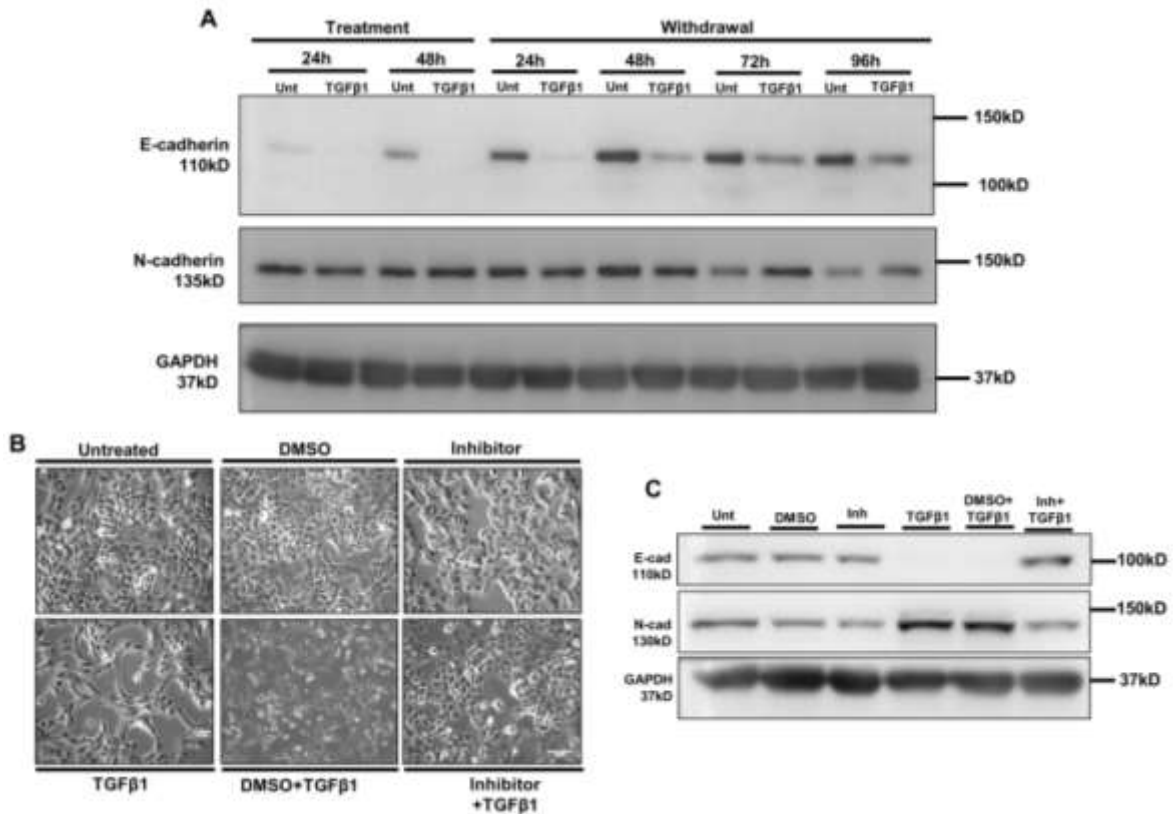
Figure 3.7

Fig.3.7: Reversibility & inhibition of TGF-β1 induced EMT. (A) Western blot of EMT markers upon withdrawal of TGF-β1 across time points shows MET N=1**(B)** Bright-field images of A549 cells upon TGF-β1-treatment in the presence and absence of Alk5 receptor inhibitor SB-431542 **(C)** Western blot for EMT markers upon TGF-β1 treatment in the presence and absence of Alk5 receptor inhibitor SB-431542 shows inhibition of EMT induction N=1.N: number of independent biological replicates

In summary, TGF-β1 treatment of A549 cells reveals key characteristics associated with EMT. Furthermore, TGF-β1 induced EMT in lung adenocarcinoma cells was reversible, as it showed MET upon withdrawal of TGF-β13

3.2.2 Effect of EMT on nuclear landscape

3.2.2.1 Nuclear morphometry/topology in EMT

EMT alters cell morphology, along with enhanced actin stress fibre formation. The nucleus is connected with the cell cytoskeleton via the LINC complex. However, the effect of EMT on nuclear structure and function remains unclear. DAPI stained nuclei were examined for the effect

of EMT on nuclear topology by quantifying area, volume and aspect ratio of the nucleus (**Fig.3.8A**). There was no change in the volume of the nucleus (**Fig.3.8B**). Nuclear area also increased (Epithelial cells, M=180.1μm^2) & Mesenchymal cells (M=185.8μm^2) (**Fig.3.8C**). Interestingly, nuclear morphology was elongated as seen by the increased aspect ratio of nuclei (Epithelial cells, M= 1.4) & (Mesenchymal cells, M=1.6) (**Fig.3.8D**). This is consistent with an elongated and elliptical shape of the nucleus in AML-12 cells upon TGF-β1 induced EMT (McDonald et al., 2011). In summary, EMT induction is also accompanied by significant changes in nuclear morphology of lung cancer cells.
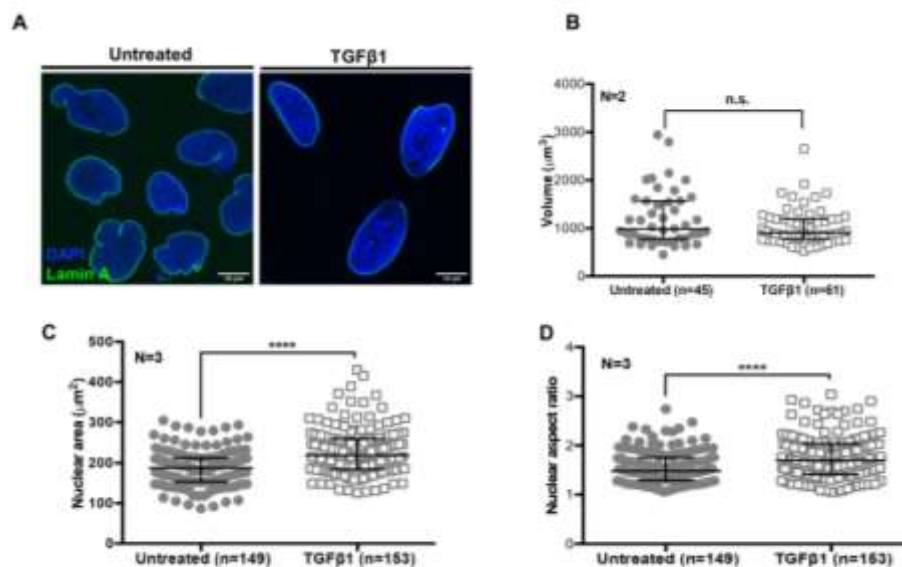
Figure 3.8

Fig.3.8: Morphometric analysis of the nucleus upon TGF-β1 induced EMT in A549 cells (A) Immunofluorescence images of untreated and TGF-β1-treated A549 cells stained with DAPI (blue) and Lamin A (green). Scale bar ~5μm (**B**) Quantification of nuclear volume upon EMT (**C**) Quantification of nuclear area upon EMT (**D**) Quantification of nuclear aspect ratio upon EMT. (Mann Whitney test, Median & IQR, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates, n: number of nuclei.

3.2.2.2 Effect of EMT on nuclear morphology

Lung cancer A549 cell line was immunostained for Lamin A/C, which revealed diverse aberrant nuclear morphologies (**Fig.3.9A**). We asked if TGF-β1-induced EMT affects nuclear morphologies. Regardless of EMT induction, A549 cells showed aberrant nuclear morphologies

(~35%) **(Fig.3.9B)**. Furthermore, we investigated if specific types of nuclear aberrations were enhanced upon EMT. However, EMT does not significantly alter the distribution of such nuclear aberrations **(Fig.3.9C)**.
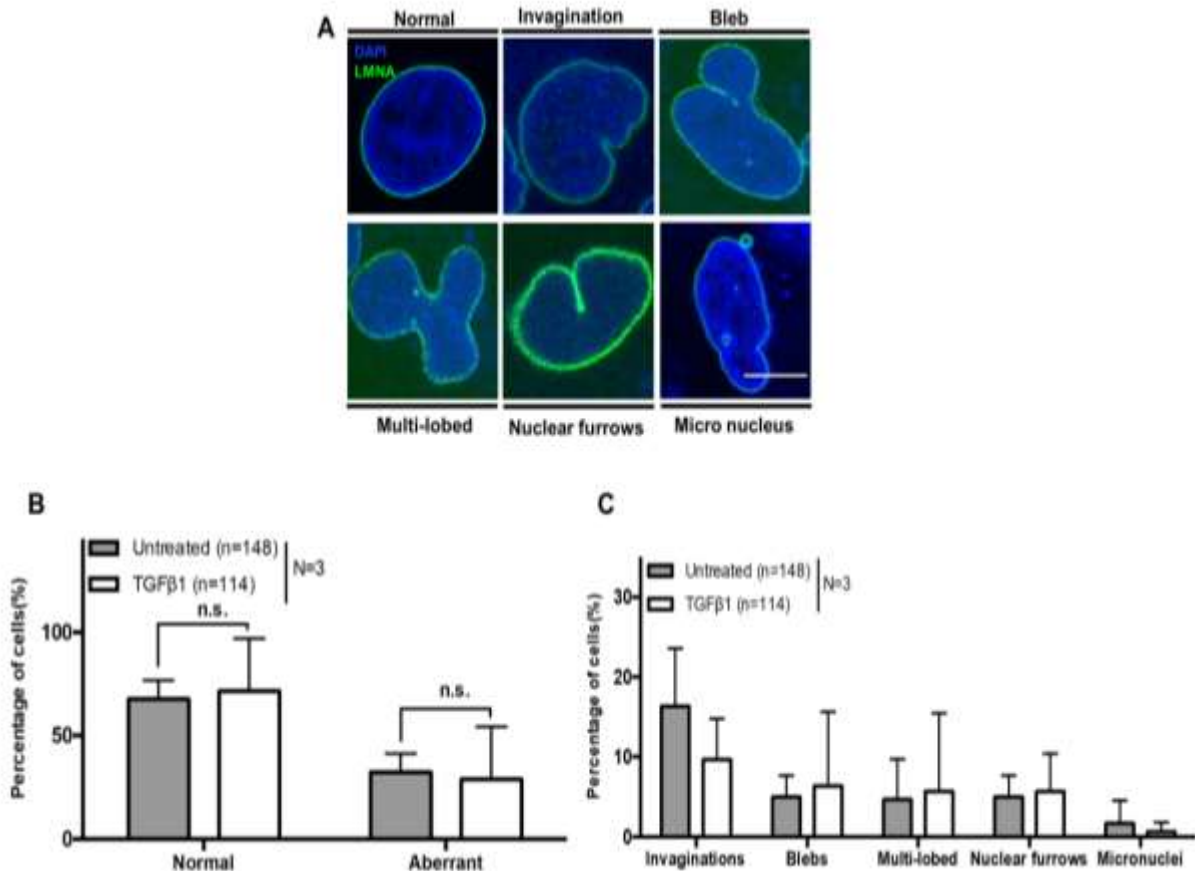
Figure 3.9

Fig.3.9: Analysis of the nuclear morphologies upon TGF-β1 induced EMT in A549 cells (A) Immunofluorescence images of untreated and TGF-β1-treated A549 cells stained with DAPI (blue) and Lamin A (green). Scale bar~5μm **(B)** Quantification of percentage of cells with nuclear aberrations (unpaired t-test, Mean ±SD, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001) **(C)** Classification of nuclear aberration ±TGF-β1-treatment, shows unaltered distribution upon TGF-β1-treatment (unpaired t-test, Mean ±SD, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates, n: number of nuclei.

3.2.2.3 Effect of EMT on nuclear envelope factors

Aberrant nuclear morphologies indicate altered expression of nuclear envelope proteins. We examined the levels of nuclear envelope proteins upon TGF-β1 induced EMT. Interestingly, qRT-PCR analyses revealed the specific downregulation of Nesprin2 levels upon TGF-β1
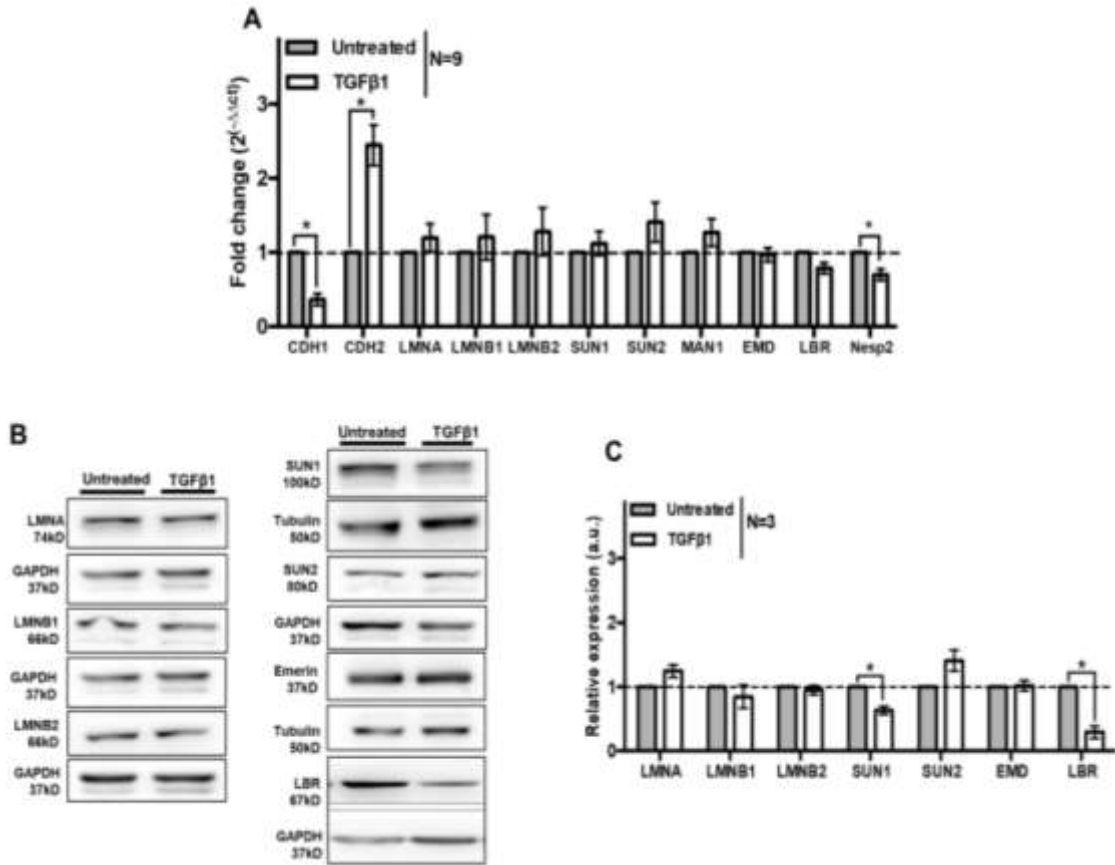
induced EMT **(Fig.3.10A)**.

Figure 3.10

Fig.3.10: Effect of TGF-β1 treatment on expression of nuclear envelope factors (A) qRT-PCR for expression of nuclear envelope factors **(B)** Western blot for nuclear envelope proteins with and without TGF-β1 treatment **(C)** Quantification of western blots of nuclear envelope proteins (unpaired t-test, Mean ±SD, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates.

Furthermore, SUN1 and LBR were downregulated, indicating that certain LINC complex proteins are responsive and indeed altered during EMT **(Fig.3.10B&C)**. In addition, immunofluorescence staining also revealed a downregulation of nuclear envelope proteins at the single cell level **(Fig.3.11).**
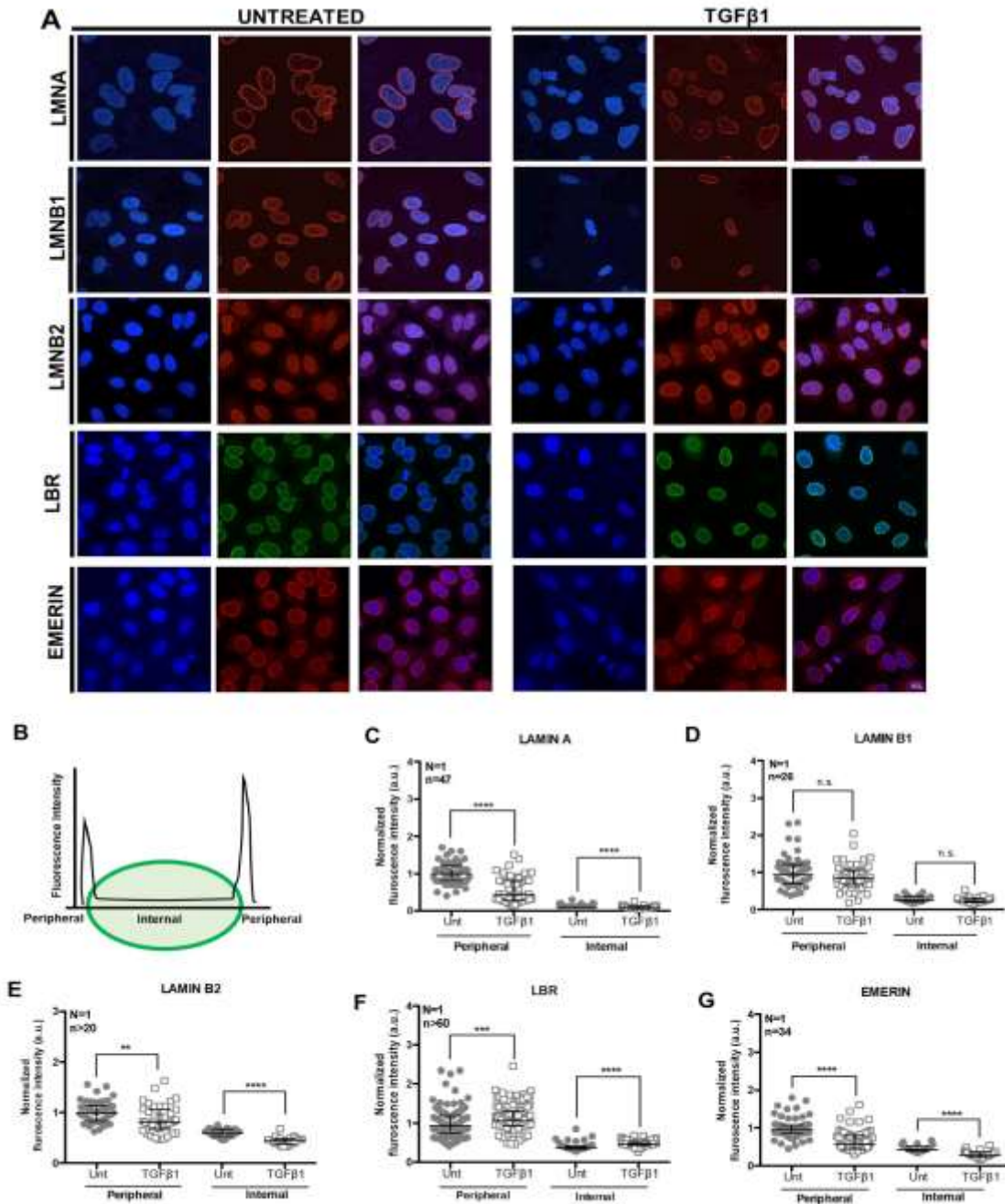
Figure 3.11

Fig.3.11: Expression of nuclear envelope factors (A) Immunofluorescence for nuclear envelope proteins of A549 cells with and without TGF-β1 treatment. Scale bar ~10μm **(B)** Schematic representation of fluorescence intensity quantification done for each nucleus using line-scan analysis. Normalized average fluorescence intensity from line-scans across nuclei for **(C)** Lamin A, **(D)** Lamin B1 **(E)** Lamin B2 **(F)** LBR and **(G)** Emerin. (Mann Whitney test, Median & IQR, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates, n: number of nuclei.

3.2.2.4 Effect of TGF-β1 on chromosomal stability

Nuclear envelope proteins regulate cell ploidy (Kuga et al., 2014). As TGF-β1 induced EMT alters the expression of nuclear envelope proteins, we studied the effect of EMT on chromosomal stability. We performed flow cytometry to examine the overall ploidy of A549 lung cancer cells **(Fig.3.12A).** A549 cells did not show any changes in cellular ploidy upon TGF-β1 induced EMT. Furthermore, the numbers of cells in various phases of the cell cycle were unaffected **(Fig.3.12B)**

Figure 3.12

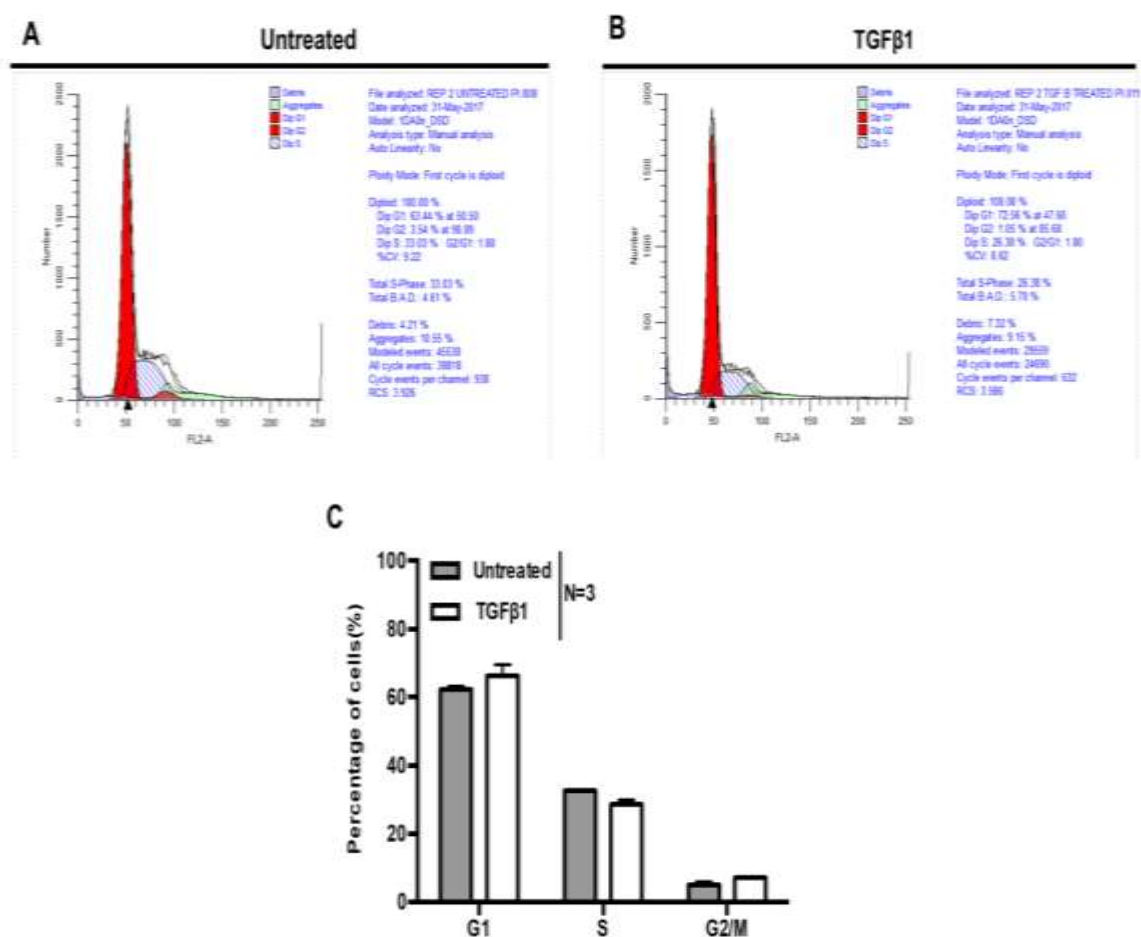**Fig.3.12: Effect of TGF-β1 treatment on ploidy (A)** Cell cycle profile by flow cytometry of untreated and TGF-β1 treated A549 cells **(B)** Distribution of A549 cells in different phases of cell cycle with and without TGF-β1, N~10,000 (Fischer's exact test, Mean±SD, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates, n: number of cells.

In addition, we also examined the ploidy at the single cell level from metaphase spreads (**Fig.3.13A**). EMT induction also did not show any changes in the modal number of chromosomes of A549 cells upon TGF-β1-treatment (Modal chromosome number: 63-64) (**Fig.3.13B**).

Figure 3.13
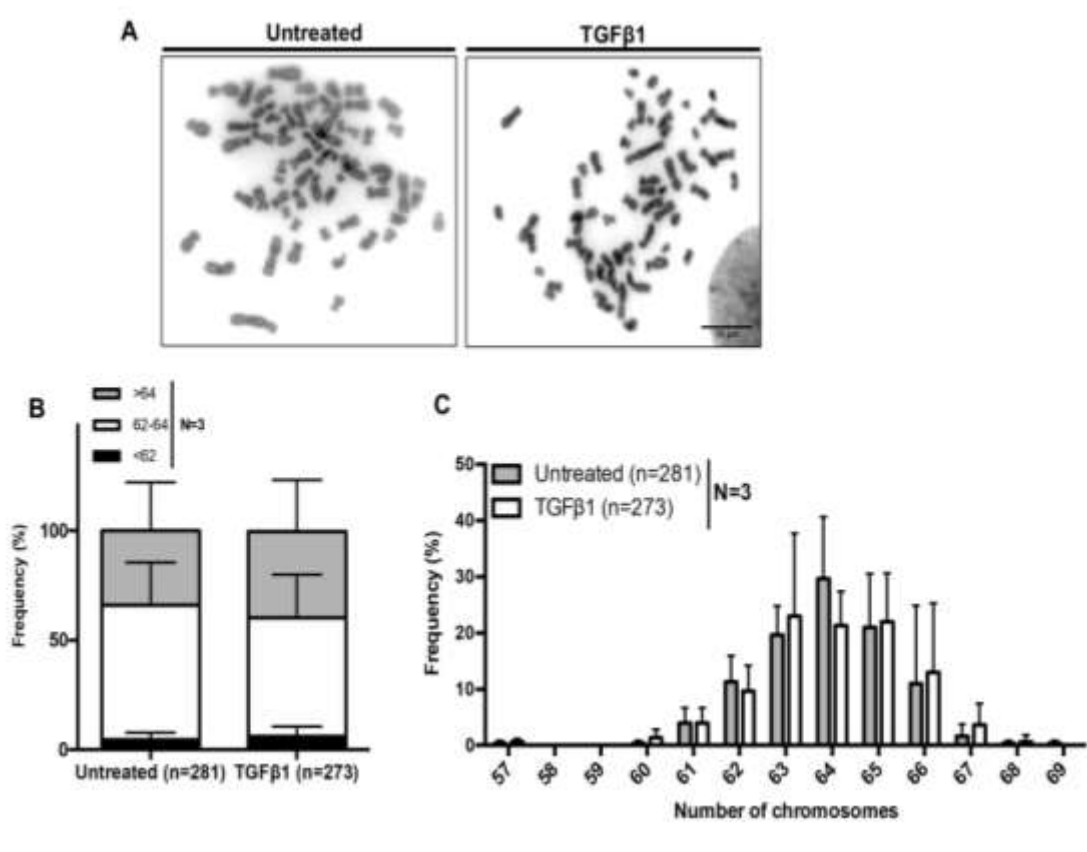
Fig.3.13: Effect of TGF-β1 treatment on chromosome stability (A) Representative metaphase spreads from untreated and TGF-β1 treated A549 cells. Scale bar ~10μm. **(B)** Quantification of chromosome numbers from metaphase spreads of untreated and A549 treated cells. (Chi-square test, Mean±SEM, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates, n=number of metaphase spreads.

In summary, TGF-β1 induced EMT in A549 cells impinges on the nuclear topology giving rise to elongated nuclei. It affects the expression of certain nuclear envelope factors without altering the ploidy of cells.

3.2.3 Effect of Epithelial to Mesenchymal Transition (EMT) on genome organization

3.2.3.1 Histone modification

Euchromatin is marked by active histone mark H3K4me3 while inactive heterochromatin is marked by H3K27me3 (Kimura, 2013). We examined the effect of EMT on these histone modifications. We investigated total levels of the active mark H3K4me3 and inactive mark H3K27me3 at the single cell by immunofluorescence assay **(Fig.3.14A&C)**. H3K4me3 showed a decrease with a concomitant increase in H3K27me3 upon TGF-β1 induced EMT in A549 cells **(Fig.3.14B&D)**. These epigenetic changes are indicative of chromatin reorganization and altered gene expression changes in EMT.

Figure 3.14
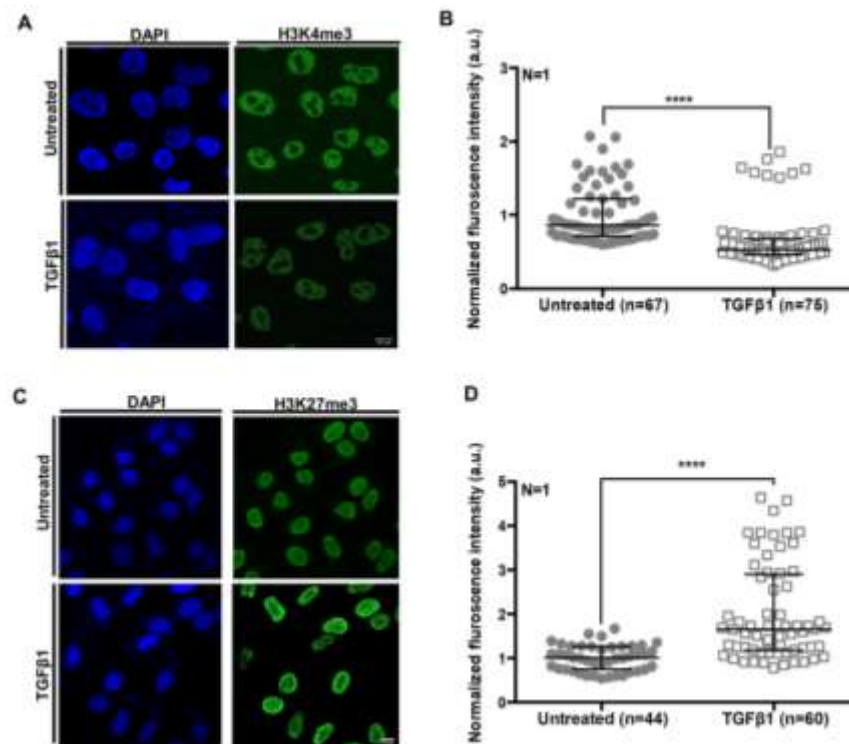
Fig.3.14: Histone modifications in A549 cells ±TGF-β1 treatment (A) Representative mid optical sections of immunofluorescence staining for H3K4me3 (green) in untreated and TGF-β1-treated A549 cells **(B)** Quantification of normalized total fluorescence intensity in the nucleus for H3K4me3 **(C)** Representative mid optical sections of immunofluorescence staining for H3K27me3 (green) in untreated and TGF-β1-treated A549 cells. Scale bar ~10μm **(D)** Quantification of normalized total fluorescence intensity in the nucleus for H3K27me3 (Mann Whitney test, Median & IQR, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates, n: number of nuclei.

3.2.3.2 Chromosome territories

We next examined the effect of TGF-β1 induced EMT on the radial position of chromosome territory 18 (CT18) and chromosome 19 territory (CT19) respectively. Of note, human chromosomes 18 & 19 are of comparable sizes (80.37 and 58.62 Mbp respectively) but divergent in their gene densities (~12 genes/Mbp and ~42 genes/Mbp respectively). It is well established that chromosomes are radially positioned, in the interphase nucleus, in a gene density dependent manner (Croft et al., 1999). CT18 (gene poor) is localized at ~66.3%, while CT19 (gene rich) is localized at ~58.4% in diploid colorectal cancer cell line DLD-1 (Cremer et al., 2003).

We performed 3D-FISH for CT18 and CT19 upon TGF-β1 treatment **(Fig.3.15A)**. We examined the relative distributions of the CTsin each of the nuclear sub-shells. While it remained unchanged upon TGF-β1 induced EMT, CT19 showed repositing towards the interior. **(Fig.3.15B&C)**. Furthermore, the median radial distributions of chromosome 18 and 19 territories were unaltered **(Fig.3.15D&E)**. Interestingly, CT18 or CT19 hardly show any spatial separation in untreated A549 cells, since both CT18 and CT19 showed a comparable radial distance (CT18, M= 61.1%, CT19, M= 56.7%).
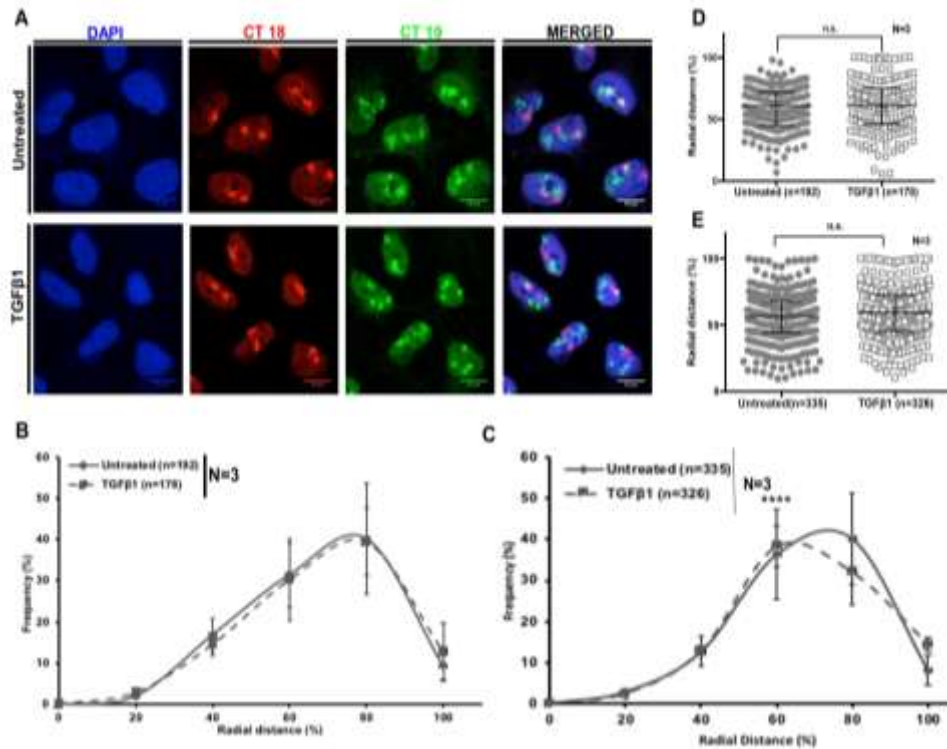
Figure 3.15

Fig.3.15: Radial positioning of CT18 & CT19 upon EMT. (A) 3D Fluorescence in-situ hybridization for CT18 & CT19 in untreated and TGF-β1-treated A549 cells. Images here show maximum intensity projections. Scale bar ~10 μm **(B)** Distribution profiles of radial distance of CT18 in untreated and TGF-β1-treated A549 cells binned into five sub-shells of ~20 % radial distance each. n: number of CTs analyzed **(C)** Distribution profiles of radial distance of CT19 in untreated and TGF-β1-treated A549 cells binned into five sub-shells of ~20 % radial distance each **(D)** Radial distance of CT18 in untreated and TGF-β1-treated A549 cells **(E)** Radial distance of CT19 in untreated and TGF-β1-treated A549 cells. (Mann Whitney test, Median & IQR, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates, n: number of chromosome territories.

3.2.3.3 Gene loci dynamics of EMT associated genes

Gene loci are also positioned in a non-random manner and show a correlation between their expression status and spatial localization in the interphase nucleus (Meaburn and Misteli, 2008). Gene loci positioned closer to the nuclear envelope are typically repressed, while gene loci closer to the nuclear interior are active (Ballabio et al., 2009). We asked if the expression status of EMT associated gene loci correlate with their spatial locations in the nucleus during EMT, for which we examined the expression and locations of EMT associated genes (i) *E-cadherin (CDH1)*, down-regulated and (ii) *N-cadherin (CDH2)*, up-regulated upon TGF-β1 induced EMT **(Fig.3.16A&B)**. Interestingly, despite showing changes in their gene expression levels, neither

E-cadherin nor N-cadherin genes showed any change in their locations in the interphase nucleus upon EMT (**Fig.3.16D&E**). In summary, 3D FISH did not capture any changes in the relative locations of chromosome 18, 19 territories or EMT associated gene loci, suggesting alternative mechanisms regulating the expression of gene loci during EMT.
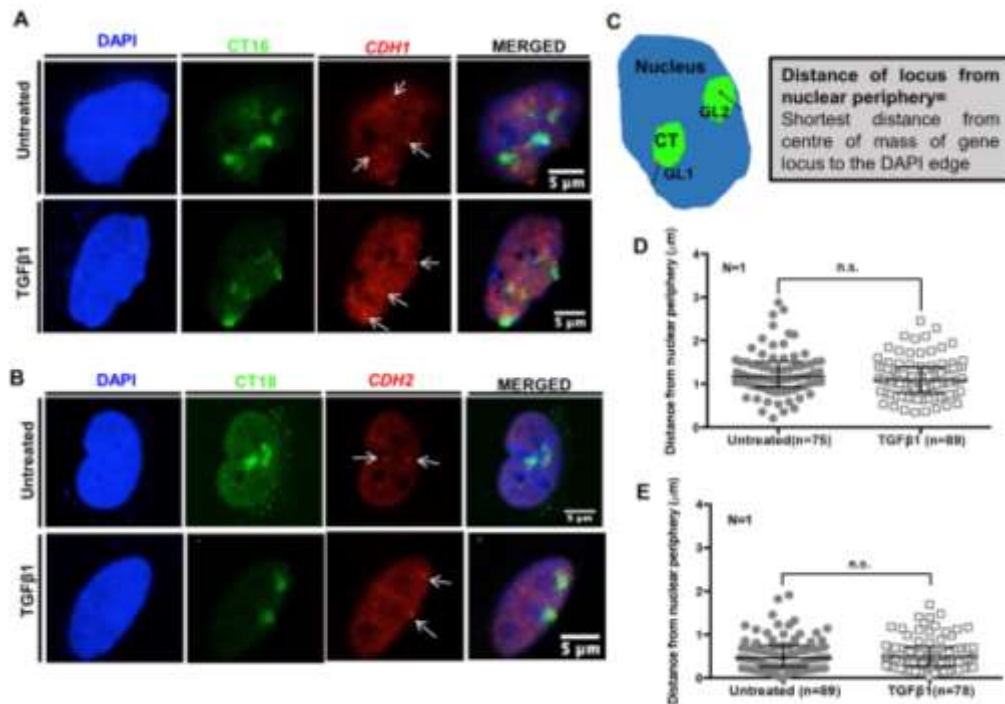
Figure 3.16

Fig.3.16: Spatial organization of gene loci in EMT. (A) 3D FISH images (Maximum intensity projection) of CT16 (green), CDH1 gene loci (red) and DAPI (blue) (**B**) 3D FISH images (Maximum intensity projection) of CT16 (green), CDH2 gene loci (red) and DAPI (blue). Scale bar ~5 μm (**C**) Scheme of 3D evaluation of spatial positioning of gene loci in the interphase nucleus. (**D**) Dot scatter plot shows the shortest distance of CDH1 gene locus from the DAPI edge (median with IQR) (**E**) Dot scatter plot shows the shortest distance of CDH2 gene locus from the DAPI edge (Mann Whitney test, Median & IQR, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates, n: number of chromosome territories.

3.2.4 Role of genome organizers in EMT

The eukaryotic genome folds into a hierarchical three-dimensional (3D) structure. Genome organization is mediated by proteins involved in chromatin folding (Rowley and Corces, 2018). 3D genome organization is critical for gene regulation and potentially altered in diseases like cancer (Qiu and Huang, 2020). Here, we examined the role of genome organizers- Lamins and CTCF, in EMT.

3.2.4.1 Lamins as genome organizers

Lamins impart rigidity to the nucleus. Softer nuclei traverse through narrow spaces relatively easily that further contributes to cell migration during metastasis (Friedl and Alexander, 2011). Additionally, the nuclear lamina is a genome organizer and contacts chromatin through 'Lamina-Associated Domain'(LADs). LADs tether DNA to the nuclear periphery and are primarily involved in transcriptional repression (Guelen et al., 2008).

3.2.4.1.1 Expression of Lamins in cancer

Gene expression datasets derived across cancer subtypes from The Cancer Genome Atlas (TCGA) were analyzed which showed an alteration in lamin levels across cancers **(Fig.3.17A-C)**. We examined the Pan Cancer dataset (PAN-CAN) from TCGA for Lamin expression across cancers and its correlation with expression levels of EMT markers; E-cadherin and N-cadherin. Interestingly, all three lamin subtypes show a negative correlation with N-cadherin expression in a cohort of cancer patients **(Fig.3.17D-F)**. Downregulation of Lamin A/C is associated with disease recurrence in stage II and III colon cancer types (Belt et al., 2011).

Figure 3.17
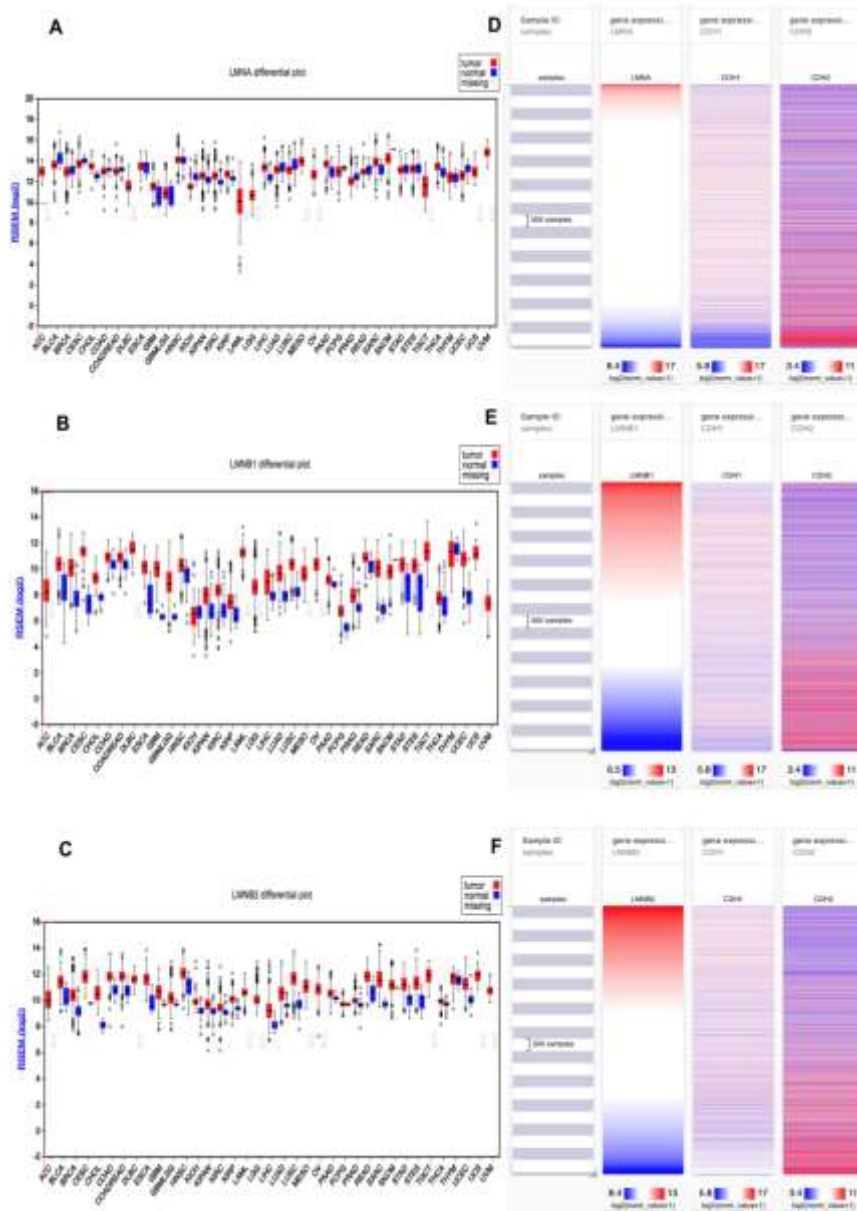
Fig.3.17: TCGA analyses for Lamin expression levels across cancer subtypes. (A) Lamin A gene expression (RSEM log2) level within and across cancers from TCGA **(B)** Lamin B1 gene expression (RSEM log2) level within and across cancers from TCGA **(C)** Lamin B2 gene expression (RSEM log2) level within and across cancers from TCGA. Each dot represents a tumor sample and the horizontal red bar indicates the median expression value within that cancer cohort. The cancer subtype abbreviations are shown below **(D)** Correlation of Lamin A gene expression with E-cadherin and N-cadherin expression across cancers **(E)** Correlation of Lamin B1 gene expression with E-cadherin and N-cadherin expression across cancers **(F)** Correlation of Lamin B2 gene expression with E-cadherin and N-cadherin expression across cancers. Samples are on the y-axis and columns on the x-axis represent the gene of interest. Each row represents the same sample.

3.2.4.1.2 Role of Lamins in TGF-β1 induced EMT

Next, to address the role of lamins in EMT, each of the lamins was downregulated using siRNA mediated knockdown, followed by TGF-β1 treatment **(Fig.3.18A)**. The extent of knockdown was assessed by western blotting. We examined if TGF-β1 treatment in the background of Lamin knockdown altered expression of EMT markers **(Fig.3.18B-D)**. However, neither of the lamin knockdowns affected TGF-β1 induced EMT in A549 cells **(Fig.3.18E-G)**.

In summary, although Lamin levels are altered across cancer subtypes, knockdown of either Lamin A/C, B1 or B2 does not alter TGF-β1 induced EMT in lung adenocarcinoma cell line A549.

Figure 3.18
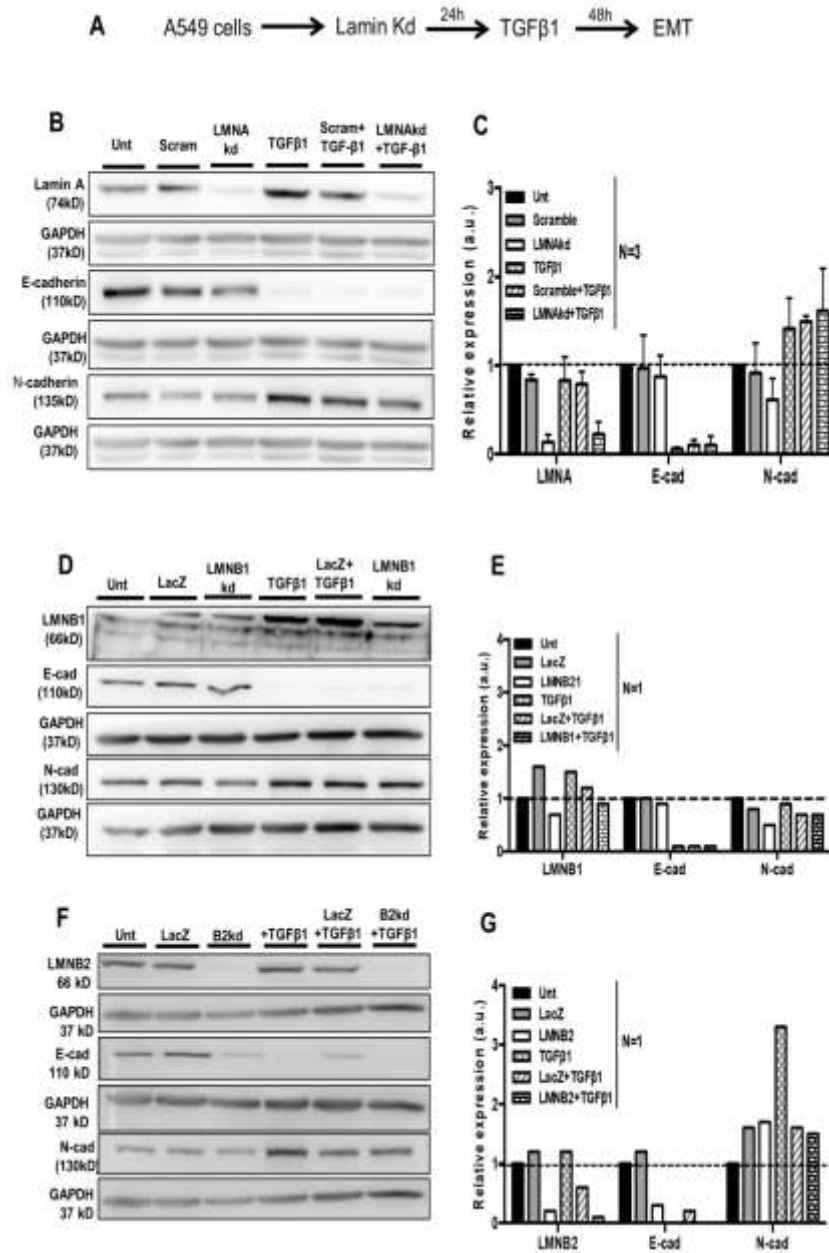
Fig.3.18: **Effect of knockdown of Lamins on TGF-β1 induced EMT. (A)** Experimental scheme to examine the effect of Lamin Kd on TGF-β1 induced EMT **(B)** Representative immunoblot for epithelial and mesenchymal markers of TGF-β1 induced EMT upon Lamin A/C Kd (Scramble RNA used as control) **(C)** Representative immunoblot for epithelial and mesenchymal markers of TGF-β1 induced EMT upon Lamin B1 Kd (siLacZ used as control) **(D)** Representative immunoblot for epithelial and mesenchymal markers of TGF-β1 induced EMT upon Lamin B2 Kd (siLacZ used as control) **(E)** Quantification of western blot shown in (B) **(F)** Quantification of western blot shown in (C) **(G)** Quantification of western blot shown in (D). (unpaired t-test, Mean± SD, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates.

3.2.4.2 CTCF as a genome organizer

Recent technological advancements and genome-wide studies show that the genomes are organized into small functional domains termed Topologically Associating Domains (TADs). TADs are organized by CCCTC-binding factor (CTCF) protein. Altered levels of CTCF may result in altered genome organization leading to deregulated gene expression (Braccioli and de Wit, 2019).

3.2.4.2.1 Expression of CTCF in cancer

We examined the gene expression levels of CTCF across cancer subtypes from TCGA. We observed that CTCF levels are variable across cancer subtypes **(Fig.3.19A).** We also examined the correlation between CTCF expression levels and expression of EMT markers. However, CTCF levels do not correlate with EMT in the Pan-Can dataset examined here **(Fig.3.19B).**

Figure 3.19
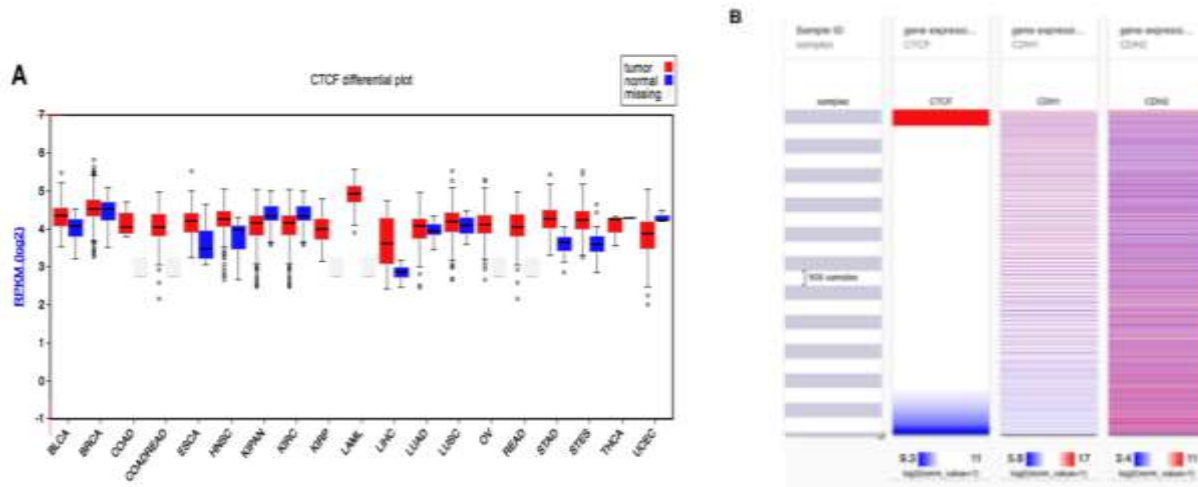
Fig.3.19: TCGA analyses for CTCF expression levels across cancer subtypes. (**A**) CTCF gene expression (RSEM log2) level within and across cancers of TCGA (**B**) Correlation of CTCF gene expression with E-cadherin and N-cadherin expression across cancers. Samples are on the y-axis and columns on the x-axis represent the gene of interest. Each row represents the same sample.

We also investigated the effect of TGF-β1 induced EMT on the expression of CTCF. Immunofluorescence followed by imaging did not show any changes in CTCF levels upon EMT **(Fig.3.20A&B).**

Figure 3.20
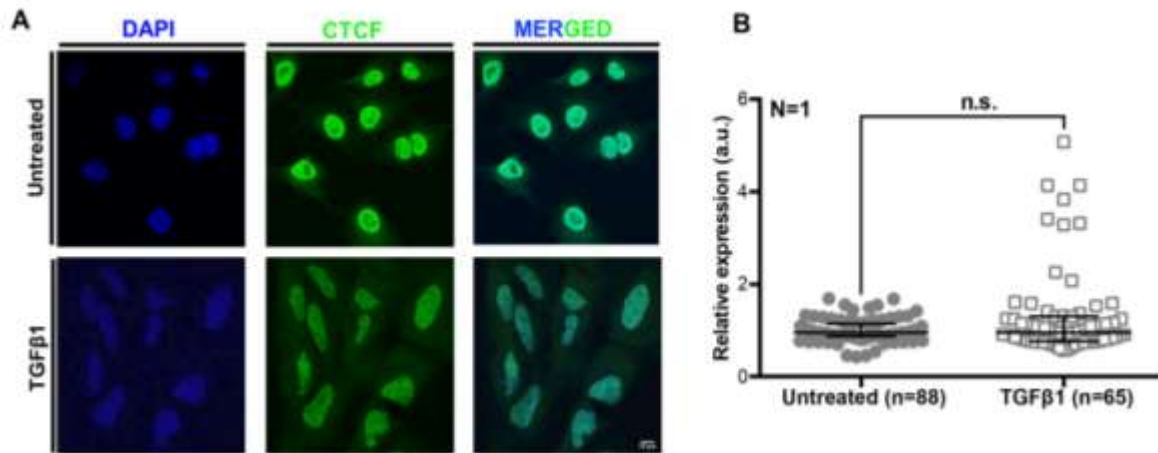
Fig.3.20: Effect of TGF-β1 induced EMT on CTCF expression (A) Mid-optical sections of cells stained for CTCF (green) with and without TGF-β1 treatment **(B)** Quantification of normalized total fluorescence intensities (Mann Whitney test, Median & IQR, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$) N: number of independent biological replicates, n: number of nuclei.

3.2.4.2.2 Role of CTCF in eliciting TGF-β1 induced transcriptional response

It is well established that EMT significantly alters the transcriptome (Du et al., 2016). We shortlisted the top up-regulated genes upon TGF-β1 treatment from previously curated RNA Seq data sets (GSE69667) of A549 cells treated with TGF-β1. In order to investigate the involvement of CTCF in regulating TGF-β1 responses, we examined whether TGF-β1 exerts its effect on these genes in the background of CTCF knockdown. CTCF Kd for ~24h hours was followed by TGF-β1 treatment **(Fig.3.21A)**. First, we validated the over-expression of these genes upon TGF-β1 treatment using qRT-PCR **(Fig.3.21C)**.

Figure 3.21
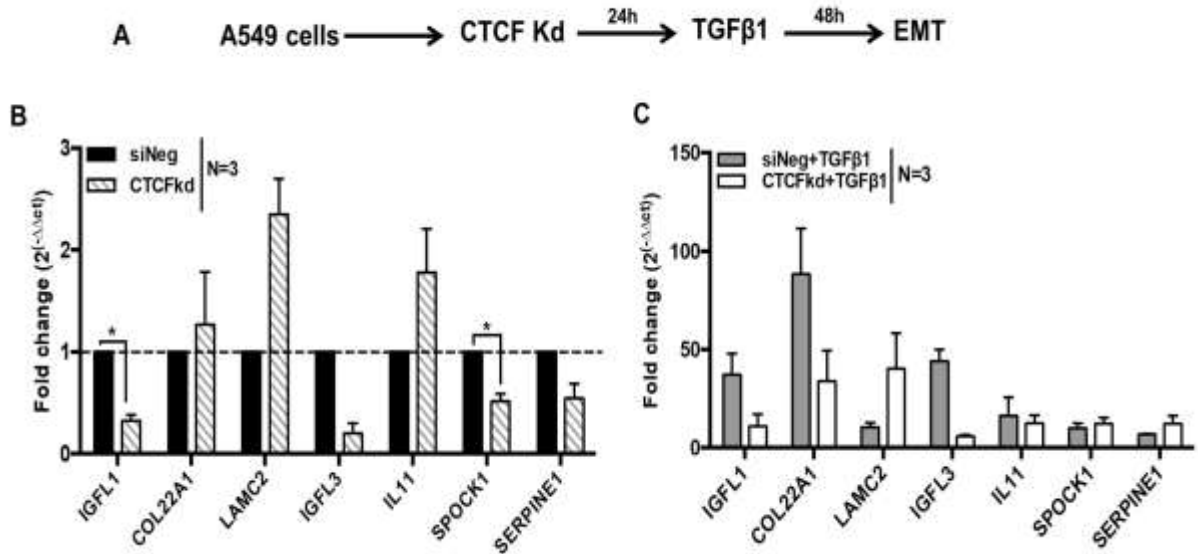
Fig.3.21: Effect of CTCFKd on TGF-β1 induced transcriptional response. (A) Experimental scheme to examine the effect of Lamin Kd on TGF-β1 induced EMT **(B)** qRT-PCR for TGF-β1 regulated genes upon CTCF Kd. siNeg served as control **(C)** qRT-PCR for TGF-β1 regulated genes upon TGF-β1 treatment alone and CTCF Kd +TGF-β1. (unpaired t-test, Mean±SEM, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates, n: technical replicates.

Next, we examined the effect of CTCF Kd on the expression of these genes. When TGF-β1 treatment was performed in the background of CTCF Kd, LAMC2 and SERPINE1 were upregulated as compared to TGF-β1 treatment alone whereas IGFL1, COL22A1 and IGFL3 were downregulated as compared to TGF-β1 treatment alone. A small fraction of genes were unaffected by TGF-β1 treatment in the background of CTCF Kd (IL11 and SPOCK1) **(Fig.3.21B)**. Interestingly, CTCF Kd alone could alter the basal expression of most genes (Downregulation-IGFL1, IGFL3, SPOCK1 and Serpine1, Upregulation-IL11) **(Fig3.21D)**. This suggests that CTCF may be required for the expression of these genes in A549 cells **(Table 3.1).**

Table 3.1: Effect of CTCF Kd on TGF-β1 induced transcriptional response

Gene	CTCF Kd wrt siNeg (Mean±SEM Normalized to siNeg)	siNeg+TGF-β1 wrt siNeg (Mean±SEM Normalized to siNeg)	CTCF Kd+TGF-β1 wrt TGF-β1 (Mean±SEM Normalized to siNeg)
IGFL1	down (0.3±0.06)	up (37.3±10.70)	dampened (10.9±4.32)
COL22A1	unaffected (1.2±0.52)	up (88.3±23.24)	dampened (33.9±8.94)
LAMC2	up (2.3±0.29)	up (10.6±1.80)	enhanced (40.1±10.49)
IGFL3	down (0.2±0.08)	up (44.0±4.94)	dampened (5.7±0.45)
IL11	up (1.7±0.42)	up (16.2±9.47)	unaffected (12.5±2.95)
SPOCK1	down (0.5±0.06)	up (9.9±2.60)	unaffected (12.2±0.28)
SERPINE1	down (0.5±0.18)	up (6.7 ±0.34)	enhanced (12.2±2.29)

In order to understand the molecular basis of CTCF mediated regulation of the transcriptional responses of TGF-β1, we tested the role of CTCF in genome organization in the context of EMT by analyzing Hi-C data sets. While Hi-C data is available from A549 cells, Hi-C data upon CTCF Kd from A549 cells is not available. Therefore we examined Hi-C data from HEK293 cells upon CTCF Kd and compared it with data from control HEK293 cells. We examined the Hi-C data for two candidate genes COL22A1 and LAMC2 upon CTCF Kd in HEK cells and observed change in TAD organization around these genes **(Fig.3.22)**. Therefore, we speculate that CTCF Kd may alter chromatin architecture for these genes altering their transcriptional responses.

Figure 3.22
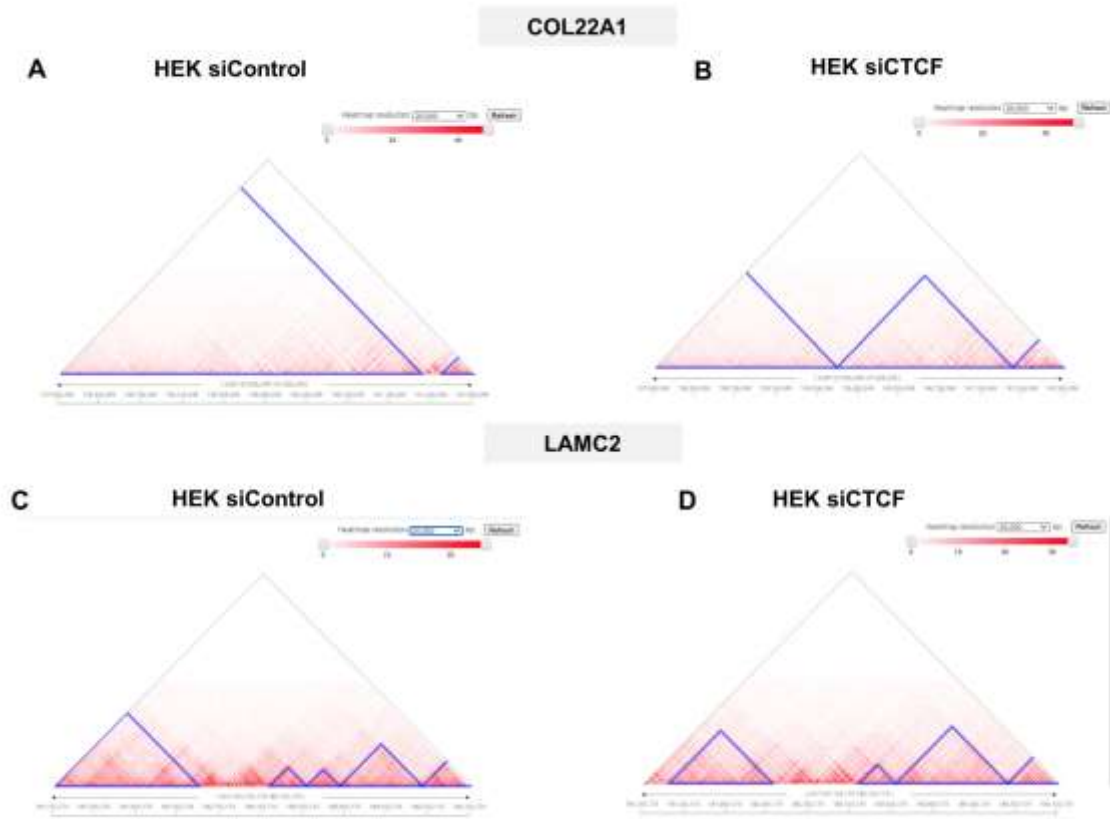
Fig.3.22: 3D gene organization of TGF-β1 responsive gene. (A) Heat map of chromatin contacts around COL22A1 gene in HEK siControl cells **(B)** Heat map of chromatin contacts around COL22A1 gene in HEK siCTCF cells **(C)** Heat map of chromatin contacts around LAMC2 gene in HEK siControl cells **(D)** Heat map of chromatin contacts around LAMC2 gene in HEK siCTCF cells. Interaction range=2Mbp. TopDom w=20. Resolution=20,000bp.

3.2.4.2.3 Role of CTCF in TGF-β1 induced EMT

As CTCF Kd altered expression of TGF-β1-overexpressed genes, we further investigated if altered levels of CTCF affect TGF-β1 induced EMT. We performed siRNA mediated Kd of CTCF in A549 cells and treated CTCF Kd cells with TGF-β1 **(Fig.3.24).** We examined the effect of CTCF Kd on the expression of EMT markers-Ecadherin and N-cadherin both at the transcript and protein level **(Fig.3.24A-C)**. CTCF Kd does not affect TGF-β1 induced 'cadherin-switch'**.** We also analyzed the aspect ratio of cells upon TGF-β1 treatment with and without CTCF depletion. We observed that cells gained spindle-shaped morphology TGF-β1 treatment even with CTCF Kd **(Fig.3.24D)**. Taken together, although CTCF Kd alters TGF-β1-mediated transcriptional responses, CTCF Kd does not perturb TGF-β1 induced EMT.

Figure 3.23
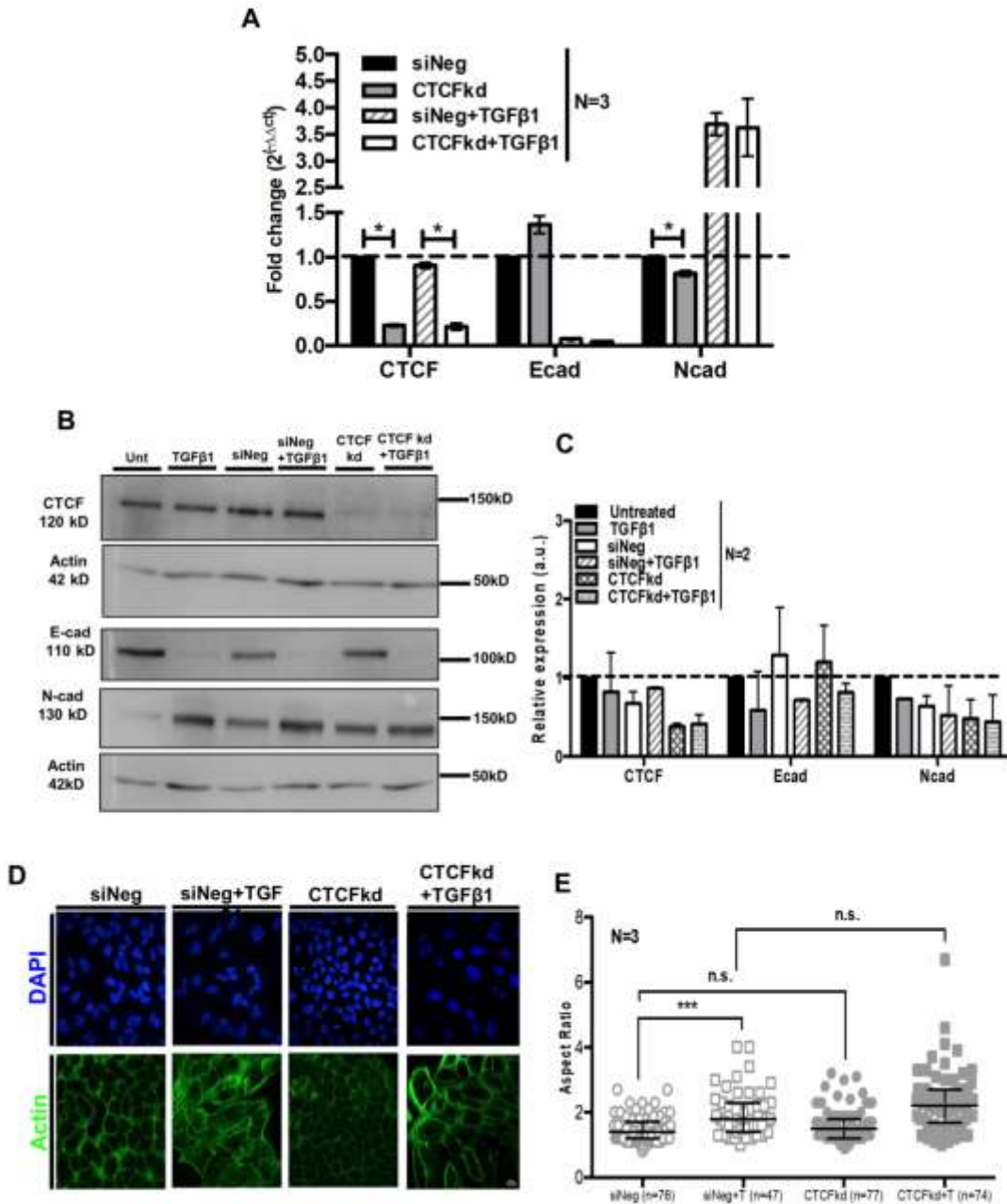
Fig.3.23: Effect of CTCF knockdown on TGF-β1 induced EMT (A) qRT-PCR of epithelial marker (E-cadherin) and mesenchymal marker (N-cadherin) in A549 cells of TGF-β1 treatment upon CTCF Kd (unpaired t-test, Mean ± SEM, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: independent biological replicates **(B)** Representative immunoblot for epithelial and mesenchymal markers of TGF-β1 induced EMT upon CTCF Kd **(C)** Quantification of western blot shown in (B) (unpaired t-test, Mean± SD, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: independent biological replicates **(D)** Immunofluorescence images of phalloidin stained actin of TGF-β1-treated A549 cells ± CTCF Kd showing actin stress fibres. Scale bar ~10μm **(E)** Quantification of aspect ratio TGF-β1-treated A549 cells ± CTCF Kd shown in (D). (Mann Whitney test, Median & IQR, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001), N: number of independent biological replicates, n: number of nuclei.

Figure 3.24
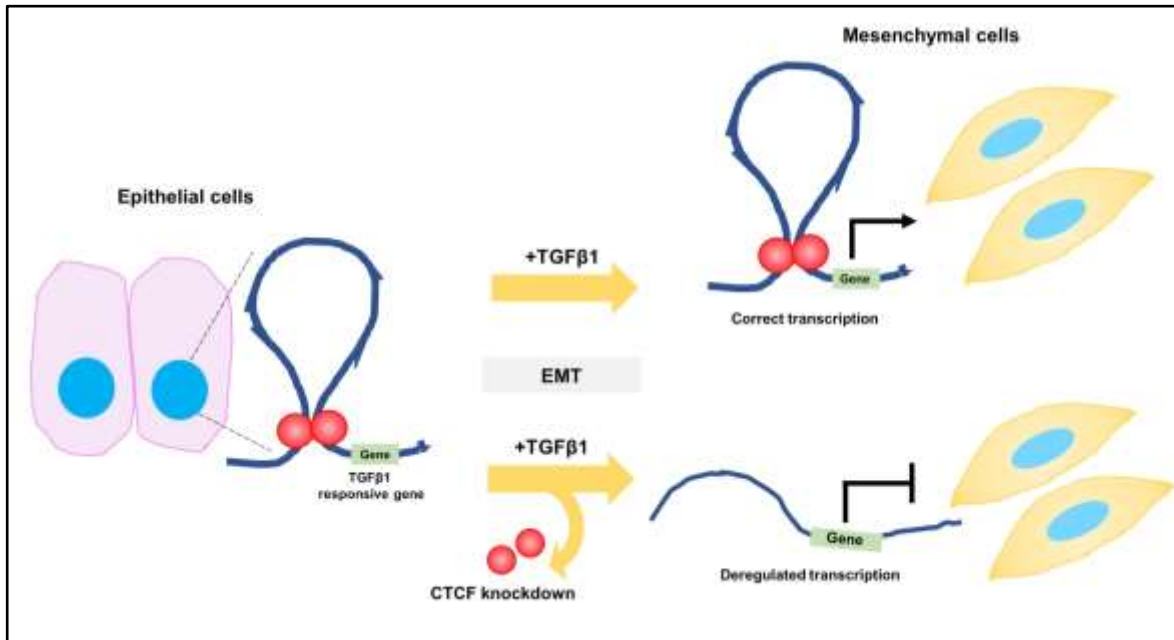
Fig.3.25: Effect of CTCF knockdown on TGF-β1 induced EMT. CTCF maintains the 3D organization in cells. Although cells can undergo EMT upon CTCF Kd, there is transcriptional deregulation of TGF-β1 responsive genes.

3.3 Discussion

Epithelial to Mesenchymal transition (EMT) is involved in cancer metastasis. Recent evidence suggests 'hybrid epithelial-mesenchymal' (hybrid E/M) phenotypes. Hybrid E/M phenotype offers more plasticity to cancer cells to adapt to stressful conditions and aids metastasis (Jolly et al., 2015). For decades pathologists have relied on aberrant nuclear morphologies for cancer diagnosis and prognosis (Chow et al., 2012). We observed that A549 which is an EMT hybrid cell line shows various types of aberrant nuclear morphologies. Classifying nuclear morphologies based on the EMT status of the cell will improve the prognostic value of nuclear morphology in cancers.

Furthermore, aberrant nuclear morphologies are indicators of deregulation of nuclear envelope proteins (Chow et al., 2012). Nuclear envelope proteins are known to regulate TGF-β1 signaling. Aberrant nuclear morphologies of A549 cells may therefore show defects in TGF-β1 signaling. Altered levels of nuclear envelope proteins may also impact genome organization in these cells that may in turn lead to transcriptional deregulation.

We examined genome organization at the levels (i) chromatin compartmentalization-euchromatin and heterochromatin, (ii) chromosome territory organization and (iii) gene loci dynamics upon EMT. We detected a downregulation of active euchromatic mark H3K4me3 and upregulation of inactive heterochromatic mark H3K27me3. There was no change in the radial position of CT18 & CT19 upon EMT. However, CT18 and CT19 show reduced spatial segregation in untreated A549 cells. It is important to note that A549 cells are aneuploid. Accommodating extra copies of chromosomes within the confines of the nucleus may lead to altered radial positions. A study comparing radial positioning of CT18 and CT19 in normal versus transformed cell lines showed that a majority of transformed cells show inverted CT positioning, i.e. CT18 is more internally located than CT19 (Cremer et al., 2003). We also examined the correlation between change in the expression status of genes and their position in the 3D interphase nucleus. For EMT associated genes - CDH1 and CDH2, we did not observe any repositioning concomitant with a change in expression upon TGF-β1 induced EMT. Gene loci repositioning studies of cancer-associated genes in early breast cancer tumorigenesis and prostate cancer show that although these genes may show repositioning they may not show a

transcriptional change and vice versa, a gene that shows transcriptional change may not show repositioning (Leshner et al., 2016; Meaburn and Misteli, 2008).

Early work studying the impact of TGF-β1 induced EMT on the genome focused on understanding associated epigenetic modifications. Data from AML-12, mouse hepatocytes showed that overall DNA methylation is unaffected. Activating histone modifications H3K4me3 and H3K36me3 increased while heterochromatin marker H3K9me2 showed a downregulation. Interestingly, the downregulation of H3K9me2 mapped predominantly to LOCKs. LOCKs are large (100 kb–5 Mb), non-repetitive heterochromatin domains which are enriched for H3K9Me2 and they overlap with nuclear lamina-associated domains (LADs) (McDonald et al., 2011). This suggests that TGF-β1 induced EMT potentially reorganizes the genome. A study in normal murine mammary gland epithelial (NMuMG) cells examined whether Lamin B1 contacts are altered upon TGF-β1 induced EMT. Lamin B1 ChIP-seq from an euchromatin enriched fraction and assay for transposase-accessible chromatin using sequencing (ATAC-seq), whole-genome chromatin conformation capture (Hi-C) was carried out to detect chromatin organization to analyze their dynamism in the context of EMT. This study demonstrated that Lamin B1 is associated with euchromatin in Euchromatic LADs (eLADs). eLADs are present in the active 'A' compartment. At the onset of EMT, Lamin B1 is enriched at TAD borders. As EMT progressed, over time, additional eLADs were formed involving genes that belong to the EMT pathway. Furthermore, depletion of Lamin B1 altered gene expression profile and impaired TGF-β1 induced EMT (Pascual-Reguant et al., 2018).

Another study in mouse lung epithelial (MLE12) cells showed the reverse, where silencing Lamin B1 promoted EMT, tumor growth and metastasis. Lamin B1 recruits polycomb repressive complex 2 (PRC2) and represses cell migration genes. Lamin B1 knockdown re-activates RET/p38 pathway and causes EMT. Thus Lamin B1 shows a tumor suppressor role in lung cancer and establishes a link between aberrant nuclear structure and epigenetic patterning with cancer (Jia et al., 2019).

Intriguingly, knockdown of either Lamin B1, Lamin A/C or Lamin B2 does not alter EMT induction in A549 human lung adenocarcinoma cells. However, the contribution of lamins in normal versus non transformed cells remains unclear.

A study of TGF-β1 induced EMT in MCF10A also showed an increase in chromatin accessibility upon EMT which was reversed upon MET. Interestingly, these regions show increased accessibility corresponding to AP1 and Smad binding sites and reduction in the number of CTCF binding sites. CTCF binding was altered between +10Kbp and +1Mbp from TSS. Thus, CTCF binding may regulate long-distance interactions. CTCF knockdown did not impair EMT induction but showed a global increase in covalent histone modification in H3K4me3, H3K27me3 and H3K9me3. These findings are suggestive of a role for chromatin looping and reorganization in EMT phenotypes (Johnson et al., 2020). Here we show that, even in A549 cells, CTCF knockdown does not alter TGF-β1 induced EMT, but alters the expression of certain genes upregulated upon TGF-β1 treatment. Thus, CTCF may alter chromatin organization, histone marks and deregulate expression of TGF-β1 responsive genes.

Furthermore, TGF-β1 increases chromatin accessibility of ~80% enhancers in NMuMG cells. Interestingly, most of the TGFβ-regulated genes are located proximal to enhancers, resulting in the deregulation of TGFβ co-regulated genes i.e. TGFβ regulatory domains (TRDs). CRISPR-mediated inactivation of enhancers within TRDs alters TGFβ-dependent regulation of all co-regulated genes. The area of TRD influence is restricted by TADs borders. Thus, TGFβ co-regulated genes that depend on a single enhancer are located within a single TAD (Guerrero-Martínez et al., 2020).

A study in a genetic mouse model of skin squamous cell carcinoma (SCC) employed ATAC-seq combined with RNA seq to define transcriptional and chromatin landscapes in different epithelial-mesenchymal hybrid states. ATAC-seq peaks enriched motifs for AP1, Ets, Tead, and Runx motifs at transition states. This suggests that a preserved, core set of transcription factors are required to induce chromatin remodeling of the intermediate state of EMT (Latil et al., 2017; Pastushenko et al., 2018). Currently, there is limited understanding on how the genome is organized in EMT hybrid states. Our studies highlight the importance of analyzing E/M hybrid cell types such as A549. It is of crucial relevance to address if chromosome territory organization is affected in E/M hybrid state. Such studies aim to provide a comprehensive understanding of

the functional relevance of nuclear structure and genome organization in the dynamic process of epithelial to mesenchymal plasticity.

Chapter 4 Role of EMT inducer Twist1 in genome instability

Results from this chapter are published as a part of

Khot, M., Sreekumar, D., Jahagirdar, S., Kulkarni, A., Hari, K., Faseela, E.E., Sabarinathan, R., Jolly, M.K., and Sengupta, K. (2020). Twist1 induces chromosomal instability (CIN) in colorectal cancer cells. Hum. Mol. Genet. *29*, 1673–1688

4.1 Introduction

Twist1 is a class II basic helix–loop–helix (bHLH) transcription factor, structurally characterized by the presence of two distinct and conserved domains- (i) a domain containing a stretch of basic amino acids (ii) two amphipathic α-helices separated by an inter-helical loop (HLH) (Jan and Jan, 1993; Murre et al., 1989) **(Fig.4.1).**

Figure 4.1
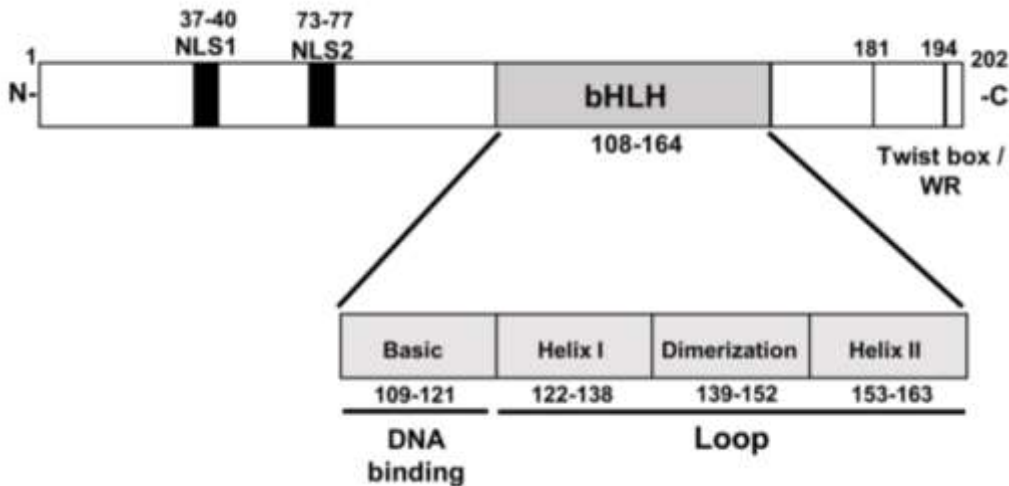
Fig. 4.1: Domain organization of human Twist1 protein

The basic domain is the DNA binding domain while the HLH domain interacts with another bHLH factor to form homo or heterodimer. All bHLH proteins recognize a hexanucleotide consensus sequence known as the E-box (CANNTG) (Jan and Jan, 1993; Murre et al., 1989). The bHLH proteins are divided into two sub-classes- (i) Class I bHLH proteins, or E proteins (eg. E12, E47). E proteins are expressed in most tissues and can form homodimers or heterodimers. (ii) Class II bHLH proteins (eg. Twist1, NeuroD) are expressed only in certain specific tissues and heterodimerize with E proteins to bind E-boxes (Murre et al., 1994). Binding

of bHLH proteins to the E-box is required for the transcriptional regulation of target genes such as *CDH1, PER1, PER2 and HPCA*, (Wong et al., 2014), (Panda et al., 2002) (Massari and Murre, 2000). E-boxes are present in the regulatory elements of many genes involved in organogenesis (Jan and Jan, 1993).

Twist1 was first identified in *Drosophila* and is involved in mesoderm-associated embryonic development through the process of EMT (Simpson, 1983). Twist1 is required in dorsoventral patterning of *Drosophila* embryos. Embryos with homozygous deletion of *Twist1* show incomplete dorsal patterning, abnormal gastrulation and fail to differentiate into the mesoderm layer (Thisse et al., 1987). The name 'Twist' was coined, as loss of this gene in *Drosophila* results in dead embryos showing a twisted appearance **(Fig.4.2)** (Nüsslein-Volhard et al., 1984).

Figure 4.2
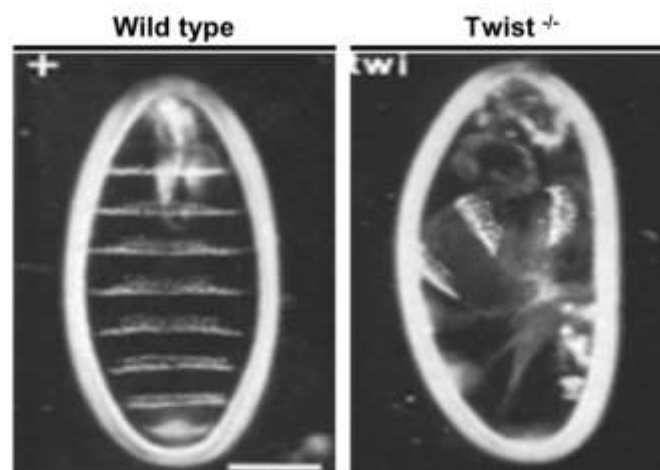
Fig.4.2: Dark-field images of cuticle preparations of normal and homozygous mutant *Drosophila* embryos (Scale bar ~ 0.1 mm. Reprinted with permission from (Jürgens et al., 1984)

Studies in mice reveal that Twist1 is essential for mesenchyme development. Mice lacking Twist1, show failure of neural tube closure and are embryonically lethal at E11.5. *Twist1* heterozygous mice show craniofacial and limb abnormalities (Chen and Behringer, 1995). Consistently, in humans, 16 of the 22 patients of Saethre-Chotzen syndrome carried *Twist1* mutations. Saethre-Chotzen syndrome is a genetic condition characterized by the premature fusion of bones of the skull (craniosynostosis). This early fusion prevents the skull from growing normally and affects the shape of the head and face (El Ghouzzi et al., 1999) . Thus, Twist1 has a

critical function in early development, and mutations or reduced levels of Twist1 show drastic developmental defects in higher organisms.

Twist1 is a known regulator of EMT in development and is aberrantly activated in epithelial tumors. The untimely expression of Twist1 leads to EMT of tumor cells, which enables cells to migrate, thereby promoting metastasis (Lamouille et al., 2014). Twist1 overexpression induces dissemination of tumorigenic cells, stemness and chemoresistance (Zhao et al., 2017). Twist1 is overexpressed across cancers of the breast, prostate, stomach, colorectal and lung, amongst others (Zhao et al., 2017). In summary, Twist1 overexpression is a major driver of cancer progression.

Seminal work by Hanahan and Weinberg on the hallmarks of cancer, has shown that genome instability and mutations characterize cancer progression (Hanahan and Weinberg, 2011)). Chromosomal instability (CIN) is primarily a result of errors in chromosome segregation during mitosis, which lead to structural and numerical chromosomal abnormalities (Vargas-Rondón et al., 2017). Various lines of evidence suggest that chromosomal instability is associated with progression of colon, bladder, prostate and breast cancers (Vargas-Rondón et al., 2017).

Since Twist1 overexpression contributes to cancer progression by inducing EMT, we asked if Twist1 overexpression is also associated with CIN. For instance, Twist1 overexpression induces chromosomal instability (CIN) in breast cancers (Mironchik et al., 2005). Spectral karyotyping (SKY) analyses of metaphases derived from Twist1 overexpressing MCF-7 (breast cancer cell line), showed an increase in chromosomal aberrations such as aneuploidy and translocations (Vesuna et al., 2006). Consistent with this observation, the stroma of colorectal tumors shows a positive correlation between Twist1 positive cells and CIN (Vesuna et al., 2006). However, the underlying mechanisms of Twist1-induced CIN are unclear.

Colorectal cancers show microsatellite instability (MSI), characterized by the insertion of repetitive nucleotide stretches, typically corrected by proteins of the mismatch repair system (MMR) (Nojadeh et al., 2018). Colorectal cancers that are mismatch repair-deficient (MMR−) show high microsatellite instability (MSI+), while mismatch repair-proficient (MMR+)

colorectal cancers hardly show microsatellite instability, but have elevated levels of CIN (Nojadeh et al., 2018). Here we examined the effect of Twist1 overexpression on chromosomal stability in two colorectal cancer cell lines, DLD1 (MSI+) and SW480 (CIN+)

4.2 Results

4.2.1 Twist1 overexpression induces epithelial to mesenchymal transition (EMT) in colorectal cancer cells

4.2.1.1 Effect of Twist1 on EMT markers

The role of Twist1 is well established in EMT during early development and cancer progression. Transient overexpression was preferred over stable expression as a model to mimic the heterogeneous increase in the levels of Twist1 during cancer progression (Zhao et al., 2017). We studied the effect of transiently overexpressing Twist1 in two colorectal cancer cell lines, of contrasting genetic background (i) DLD1—a near diploid, mismatch repair-deficient cell line and (ii) SW480—aneuploid, mismatch repair proficient cell line **(Table 4.1)**. We transfected each of the cell lines with Twist1 and examined Twist1 protein levels by immunoblotting. Both cell lines showed Twist1 overexpression **(Fig.4.3A)**.

Table 4.1: Genetic background of colorectal cancer cell lines used in this study

Cell Line	Microsatellite Instability (MSI) status	Chromosomal Instability (CIN)	APC	KRAS	p53
DLD1	MSI[1]	Negative[1]	Truncation-Amino acid 1417 [2]	Mutant G13D[1]	Mutant-S241F[2]
SW480	MSS[1]	Positive[1]	Truncation-Amino acid 1318[3]	Mutant G12V[1]	Mutant-R273H;P309S

[1] (Ahmed et al., 2013)
[2] (Chandra et al., 2012)
[3] (Liu and Bodmer, 2006)

Next, we examined the effect of Twist1 overexpression on EMT markers namely E-cadherin and Vimentin. Twist1 overexpression showed a significant decrease in the levels of the epithelial marker- E-cadherin (~30%), and an increase in the expression levels of the mesenchymal marker- Vimentin (~43%) in DLD1 cells **(Fig.4.3A&B).** However, in SW480 cells, although there was a downregulation of E-cadherin (~64%) there was only a marginal increase in Vimentin levels **(Fig.4.3A&C).**

Figure 4.3
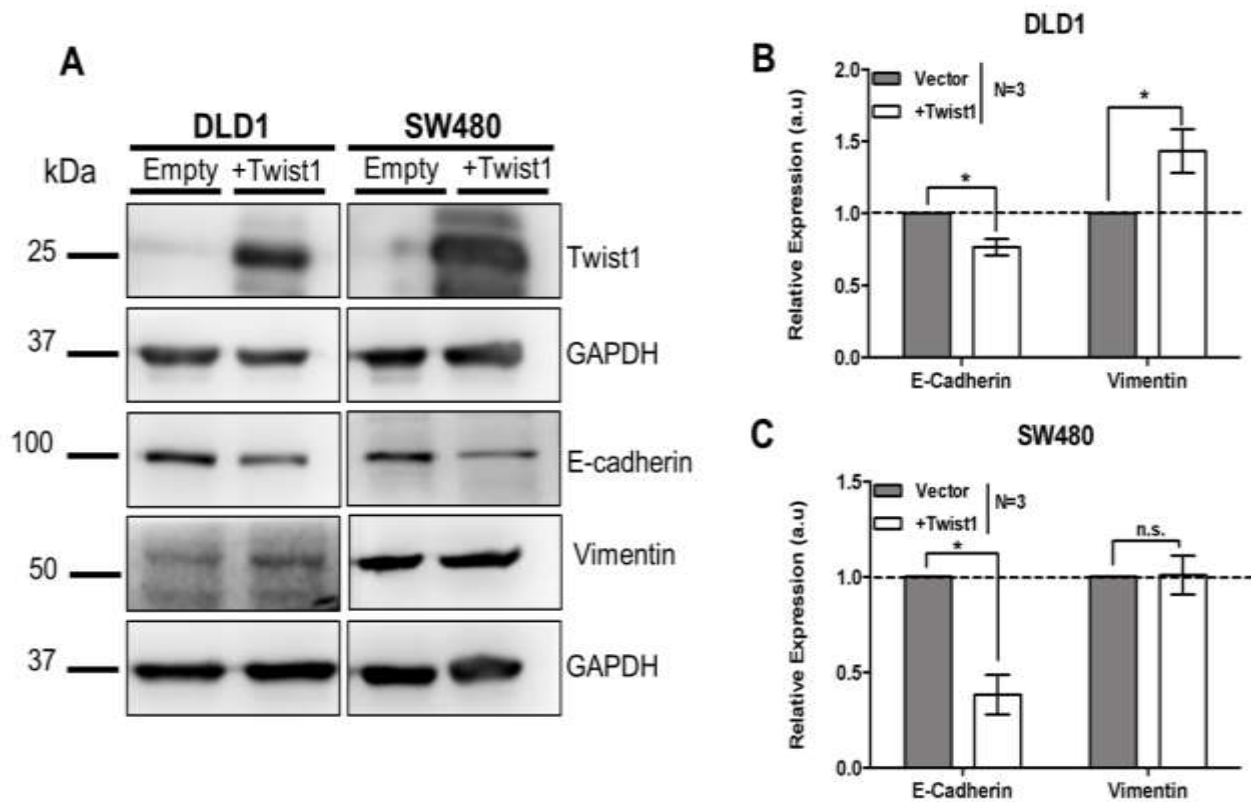
Fig.4.3: Expression of EMT markers upon Twist1 overexpression in colorectal cancer cells (A) Representative immunoblot showing Twist1 overexpression and expression of EMT markers, E-cadherin and Vimentin in DLD1 and SW480 cell lines **(B)** Quantification of band intensities of E-cadherin and Vimentin protein levels in DLD1 cell line upon Twist1 overexpression **(C)** Quantification of band intensities of E-cadherin and Vimentin protein levels in SW480 cell line upon Twist1 overexpression. (unpaired t-test, N=3, Mean±SD, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates.

We also examined the status of EMT induction at the single cell level. Cells overexpressing Twist1 and cells expressing the empty vector alone, were immunostained for E-cadherin

(**Fig.4.4A**). E-cadherin levels were examined at the single cell level in both these conditions. E-cadherin levels showed a significant reduction in both DLD1 (~50%) and SW480 cells (~45%) (**Fig.4.4A–C**). In conclusion, Twist1 overexpression causes a change in expression of EMT markers in the two cell lines: DLD1 and SW480.

Figure 4.4
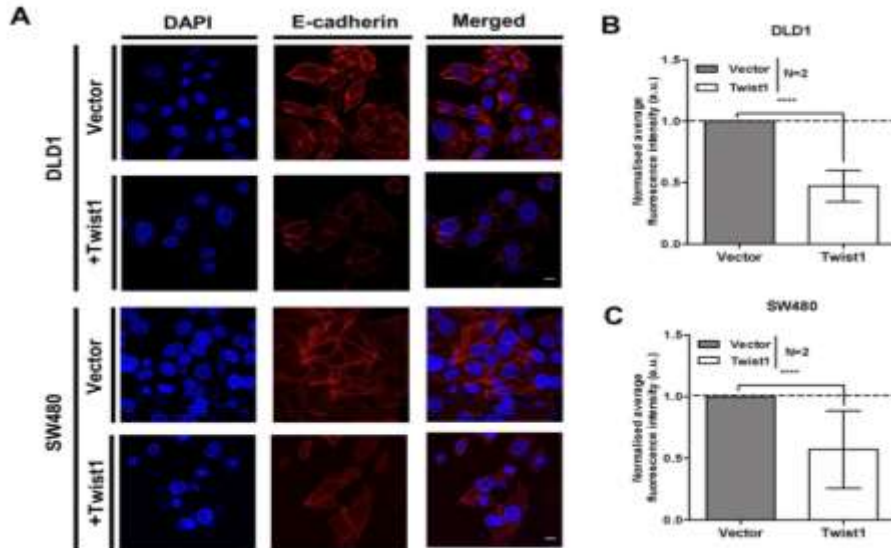
Fig.4.4: Single cell analysis of E-cadherin expression upon Twist1 overexpression in colorectal cancer cells (**A**) Representative mid-optical sections from confocal z-stacks of DLD1 and SW480 cells, immunostained for E-cadherin. Scale bar ~10 μm. (**B & C**) Normalized fluorescence intensity of E-cadherin, for vector and Twist1 overexpressing cells. (Mann-Whitney test, N=2, n>60, Mean±SD, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates, n: number of cells.

4.2.1.2 Effect of Twist1 overexpression on cell morphology

EMT induction is accompanied by a characteristic transition in cell morphology from the typical cobblestone morphology of epithelial cells to a relatively more elongated and spindle-shaped morphology of mesenchymal cells (Leggett et al., 2016). To examine the effect of Twist1 overexpression on cell morphology, we immunostained cells for actin, using phalloidin (**Fig.4.5A**).

Figure 4.5
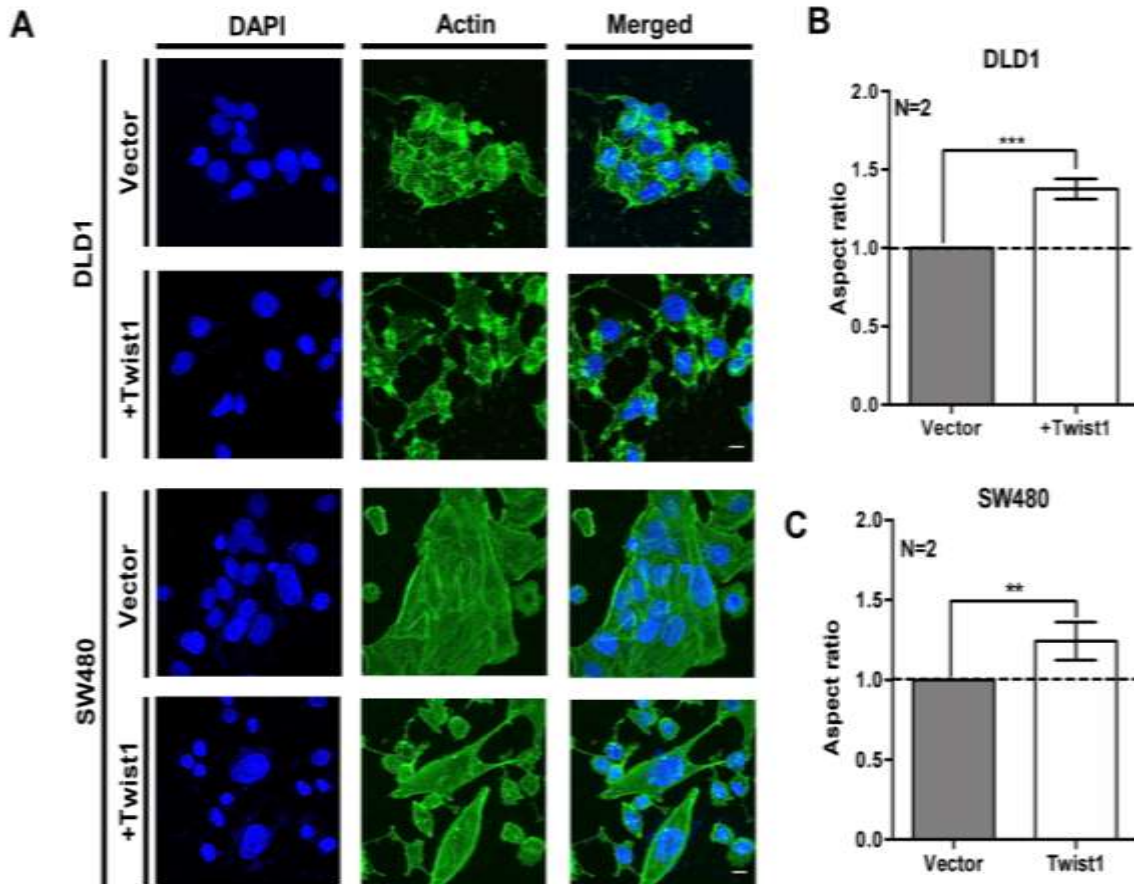
Fig.4.5: Actin organization upon Twist1 overexpression in colorectal cancer cells (A) Immunostaining for Actin shows an elongated and spindle-shaped morphology upon Twist1 overexpression. Scale bar ~10 μm. **(B & C)** Quantification of aspect ratio of DLD1 and SW480 cells, respectively (Mann-Whitney test, N=2, n>40, Mean±SD, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates, n: number of cells.

We next calculated the aspect ratio of each cell to quantify the extent of cell elongation as a function of EMT, since EMT is characterized by an increase in elongated mesenchymal cells. Interestingly, both the cell lines showed a comparable increase in the aspect ratio of cells, wherein DLD1 showed ~30% increase, and SW480 showed ~24% increase. **(Fig.4.5B&C)**. This data suggests that a distinctive change in morphology involving cell elongation accompanies EMT upon Twist1 overexpression in colorectal cancer cells.

In conclusion, Twist1 overexpression in DLD1 and SW480 cells shows EMT induction characterized by downregulation of an epithelial marker, E-cadherin and upregulation of a

mesenchymal marker Vimentin, and is accompanied by elongated cell morphology: the two hallmarks of EMT.

4.2.2 Twist1 overexpression enhances nuclear aberrations in colorectal cancer cells

Aberrant nuclear morphologies such as nuclear blebs and micronuclei are enhanced in cancers (Smith et al., 2018). The frequency of such aberrant nuclear morphologies serve as diagnostic features, routinely quantified in histopathological analyses of tissue biopsy samples (Zink et al., 2004). As Twist1-induced EMT is associated with cancer progression, we asked if Twist1 overexpression alters nuclear morphologies of cancer cells. We therefore determined the number of nuclear blebs and micronuclei upon Twist1 overexpression in colorectal cancer cells. The cells were immunostained for Lamin A as a marker of the nuclear envelope to examine nuclear morphologies **(Fig.4.6A)**.

Figure 4.6
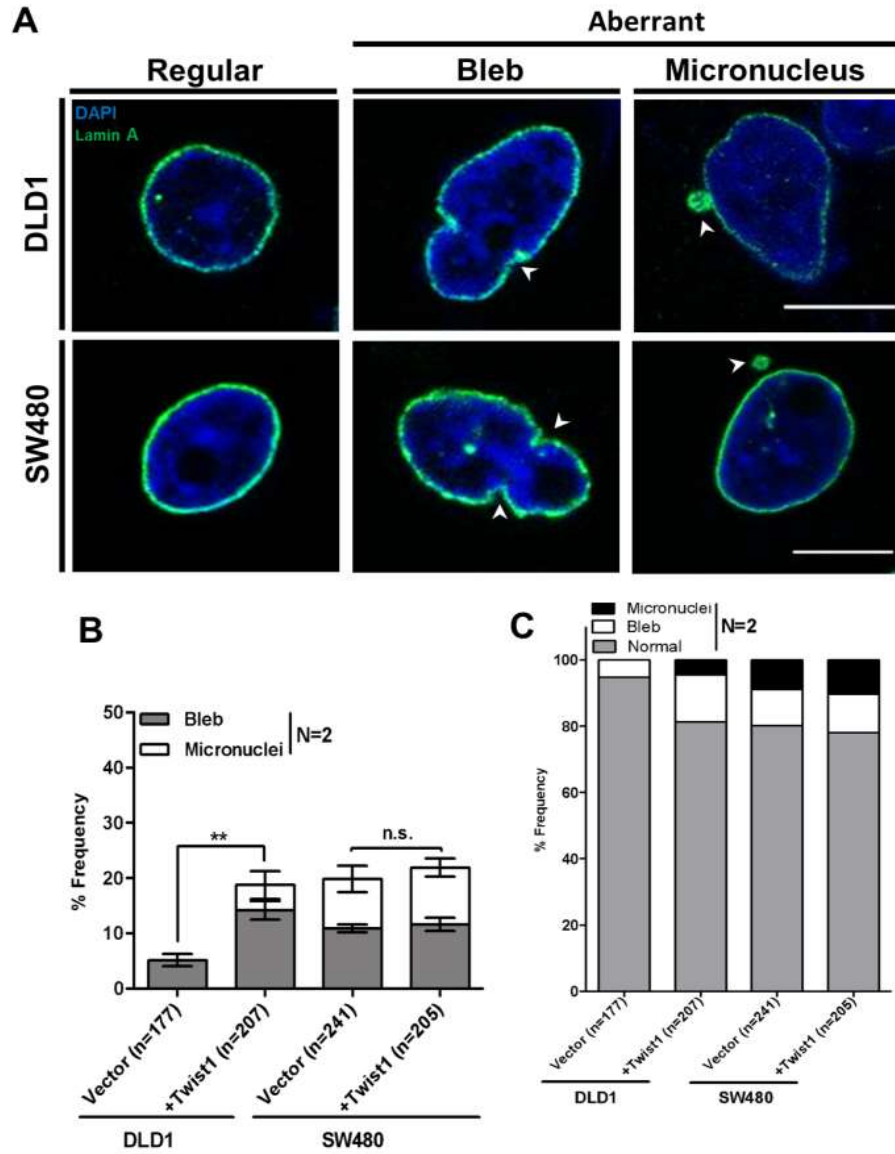
Fig.4.6: Nuclear morphologies upon Twist1 overexpression in colorectal cancer cells (A) Representative confocal images of nuclei upon Twist1 overexpression in DLD1 and SW480 cells immunostained with Lamin A, showing nuclear blebs and micronuclei. Scale bar ~10 μm **(B)** Quantification of number of cells showing aberrant nuclei upon Twist1 overexpression. (Chi-square test, N=2, Mean with Range, *P < 0.05, **P < 0.01, ***P< 0.001 and ****P < 0.0001) **(C)** Quantification of number of cells showing aberrant nuclei upon Twist1 overexpression, with frequency of regular nuclei observed across both cell lines. N: number of independent biological replicates, n: number of nuclei.

The nuclear bleb is a spherical or oval protrusion of the nuclear membrane connected to the nucleus by a narrow chromatin segment (Capell and Collins, 2006). In contrast, a micronucleus is detached from the main nucleus, but not exceeding one third of the diameter of the main nucleus (Grover and Mujib, 2017). Twist1 overexpression shows an increase (~5%) in the frequency of micronuclei and nuclear blebs (~9%) in DLD1 cells, while SW480 cells hardly showed an increase in these aberrations **(Fig.4.6B)**. Taken together, these assays suggest that Twist1 overexpression induces increased nuclear aberrations, characterized by nuclear blebs and micronuclei **(Table 4.2)**.

Table 4.2: Nuclear aberrations in colorectal cancer cell lines upon Twist1 overexpression

	DLD1		SW480	
	Vector	**+Twist1**	**Vector**	**+Twist1**
Normal	95%	81%	80%	78%
Blebs	5%	14%	11%	12%
Micronuclei	0%	5%	9%	10%

4.2.3 Twist1 overexpression enhances mitotic aberrations in colorectal cancer cells

Since an increase in aberrant nuclear morphologies are precursors to mitotic defects, we determined if Twist1 overexpression enhances mitotic aberrations (Capo-chichi et al., 2011; Hatch et al., 2013). We enumerated the number of mitotic aberrations by scoring for anaphase bridges, lagging chromosomes and tripolar spindles upon Twist1 overexpression **(Fig.4.7A)**. Near-diploid DLD1 cells showed a significant increase in the extent of mitotic defects, as compared to SW480 cells **(Fig.4.7B&C)**.

Figure 4.7
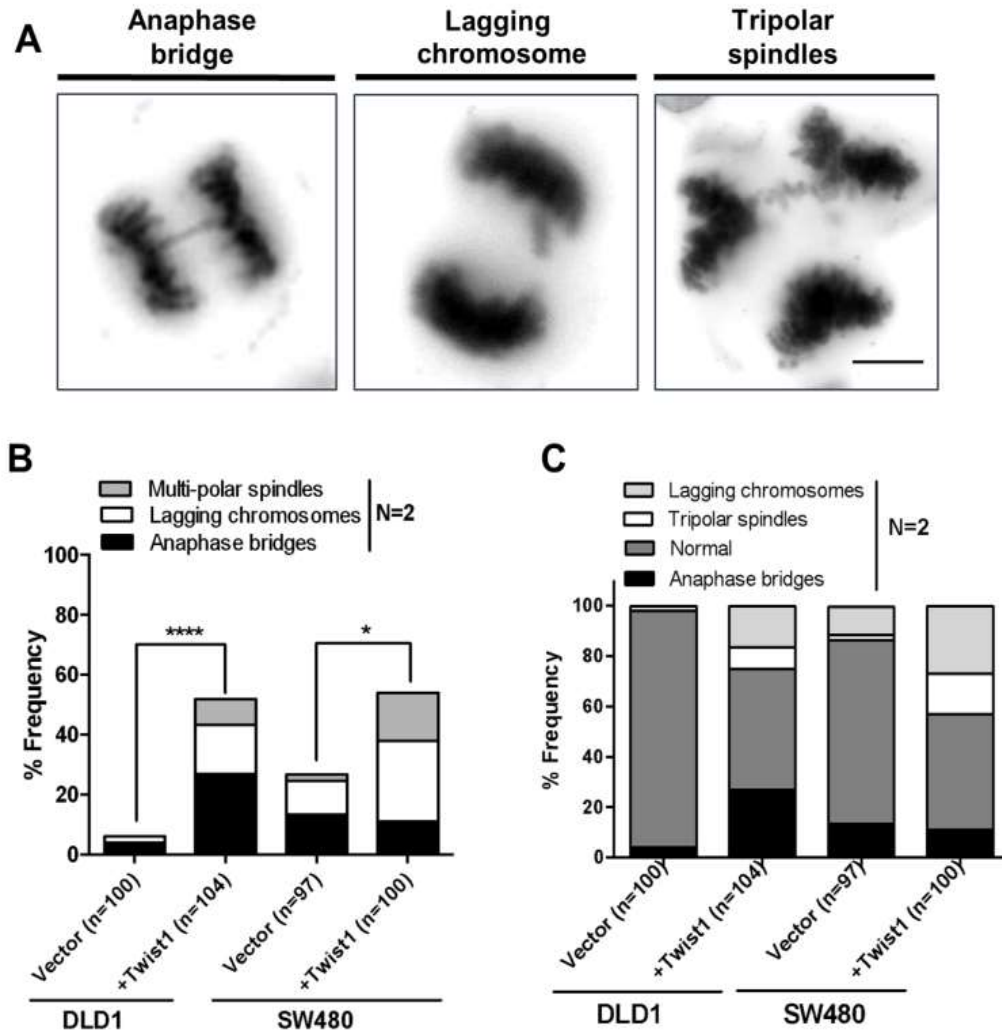
Fig.4.7: Mitotic aberrations upon Twist1 overexpression in colorectal cancer cells (A) Representative images of mitotic aberrations showing anaphase bridges, lagging chromosomes, and tripolar spindles. Scale bar ~10 μm **(B)** Quantification of the number of mitotic aberrations upon Twist1 overexpression. (Chi-square test, N=2, Mean, *P < 0.05, **P < 0.01, ***P< 0.001 and ****P < 0.0001). N: number of independent biological replicates, n: number of nuclei **(C)** Quantification of the number of mitotic aberrations upon Twist1 overexpression plotted with frequency of normal mitotic divisions.

Furthermore, DLD1 cells showed an increase in anaphase bridges (~23%), lagging chromosomes (~14%) and tripolar spindles (~9%) **(Fig.4.7B)**, while SW480 cells showed a decrease (~3%) in anaphase bridges, accompanied by an increase in lagging chromosomes (~16%) and tripolar spindles (~14%) respectively **(Fig.4.7C** and **Table 4.3)**. In summary, these studies show a distinctive increase in the extent of mitotic aberrations upon Twist1 overexpression in colorectal cancer cells.

Table 4.3: Nuclear aberrations in colorectal cancer cell lines upon Twist1 overexpression

	DLD1		SW480	
	Vector	**+Twist1**	**Vector**	**+Twist1**
Lagging chromosome	2%	16%	11%	27%
Anaphase bridge	4%	27%	14%	11%
Multipolar spindle	0%	9%	2%	16%
Normal	94%	48%	73%	46%

4.2.4 Twist1 overexpression induces chromosomal instability in colorectal cancer cells

4.2.4.1 Numerical chromosomal aberrations

Having found a significant increase in mitotic aberrations associated with Twist1 overexpression, we asked if Twist1 induces chromosomal instability (CIN) in colorectal cancer cells. We first analyzed the ploidy of cells upon Twist1 overexpression by flow cytometry. Neither DLD1 nor SW480 cells showed any change in their overall ploidy, upon ~72 hours of Twist1 overexpression **(Fig.4.8A-D)**. This data suggests that colorectal cancer cells may have inherent mechanisms that protect them from gross changes in overall chromosomal ploidy, even upon Twist1 overexpression.

Figure 4.8
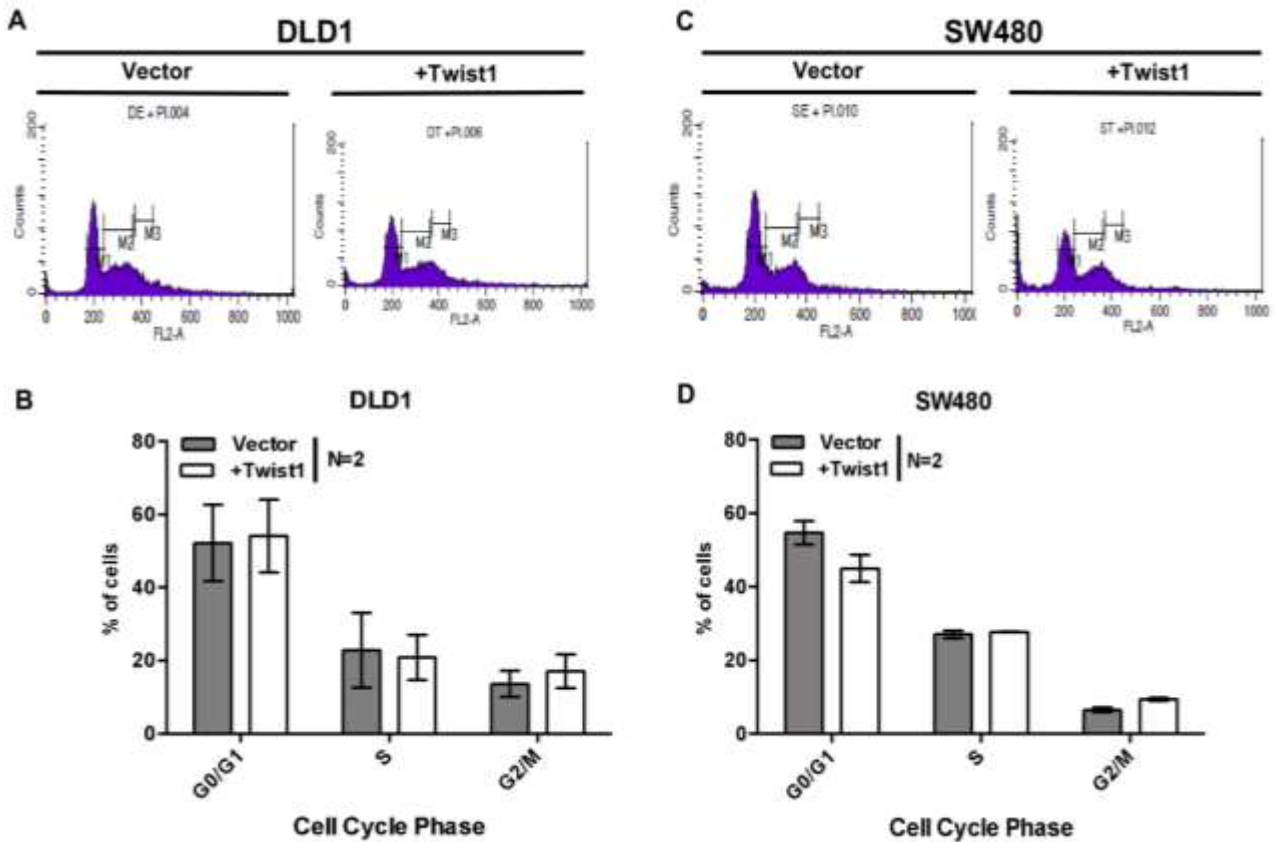
Fig.4.8: Effect of Twist1 overexpression on ploidy and cell cycle of colorectal cancer cells (**A**) Representative flow cytometry profiles for ploidy analysis of vector and Twist1 overexpressing cells for DLD1 (**B**) Quantification of cells from flow cytometry analysis for DLD1 across cell cycle phases upon Twist1 overexpression compared to vector control (**C**) Representative flow cytometry profiles for ploidy analysis of vector and Twist1 overexpressing cells for SW480 (**D**) Quantification of cells from flow cytometry analysis for SW480 across cell cycle phases upon Twist1 overexpression compared to vector control. (Mann-Whitney test, N=2, n>10,000, Mean±SD, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates, n: number of cells.

We next asked if Twist1 overexpression alters chromosome numbers. We counted the number of chromosomes from individual metaphase spreads, prepared from cells overexpressing Twist1 for 72h. Metaphase spreads of cells expressing the empty vector served as control (**Fig.4.9A-B**). We detected a significant increase in the number of cells showing whole chromosomal gains (~23%) and losses (~16%), upon Twist1 overexpression in the DLD1 cell line (**Fig.4.9C**). In contrast, there was a significant increase in whole chromosomal losses (~32%), and a decrease in whole chromosome gains (~7%) upon Twist1 overexpression in SW480 cells (**Fig.4.9D**).

Figure 4.9
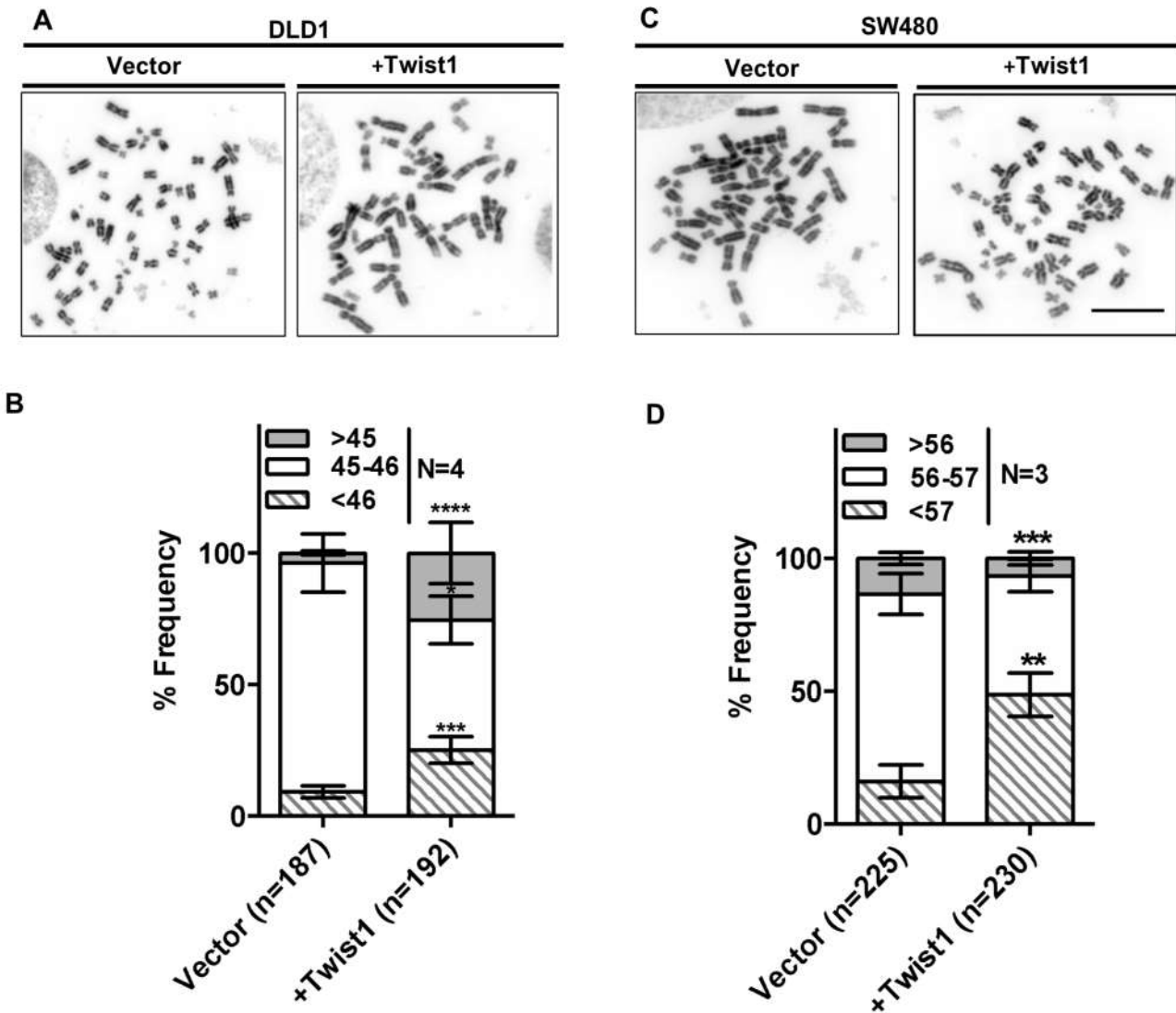
Fig.4.9: Effect of Twist1 overexpression on chromosomal stability in colorectal cancer cells (A) Representative images of metaphase chromosome spreads derived from DLD1 upon Twist1 overexpression. Scale bar ~10 μm **(B)** Quantification of whole chromosomal gains and losses in DLD1 from metaphase spreads **(C)** Representative images of metaphase chromosome spreads derived from SW480 upon Twist1 overexpression. Scale bar ~10 μm **(D)** Quantification of whole chromosomal gains and losses in SW480 from metaphase spreads. (Z-test of proportions, N=3, n>180, Mean±SEM, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$). N: number of independent biological replicates, n: metaphase spreads.

4.2.4.2 Sub-chromosomal aberrations

We performed array-Comparative Genomic Hybridization (array CGH) as an independent approach to determine if Twist1 induces amplifications and deletions at the sub-chromosomal

level, across the genome **(Fig.4.10A-F)**. EMT was induced upon Twist1 overexpression followed by array CGH analyses. Cells transfected with the corresponding empty vector served as reference. Analysis of array CGH data revealed sub-chromosomal amplifications and deletions across the genome **(Fig.4.10C&D)**.

Figure 4.10
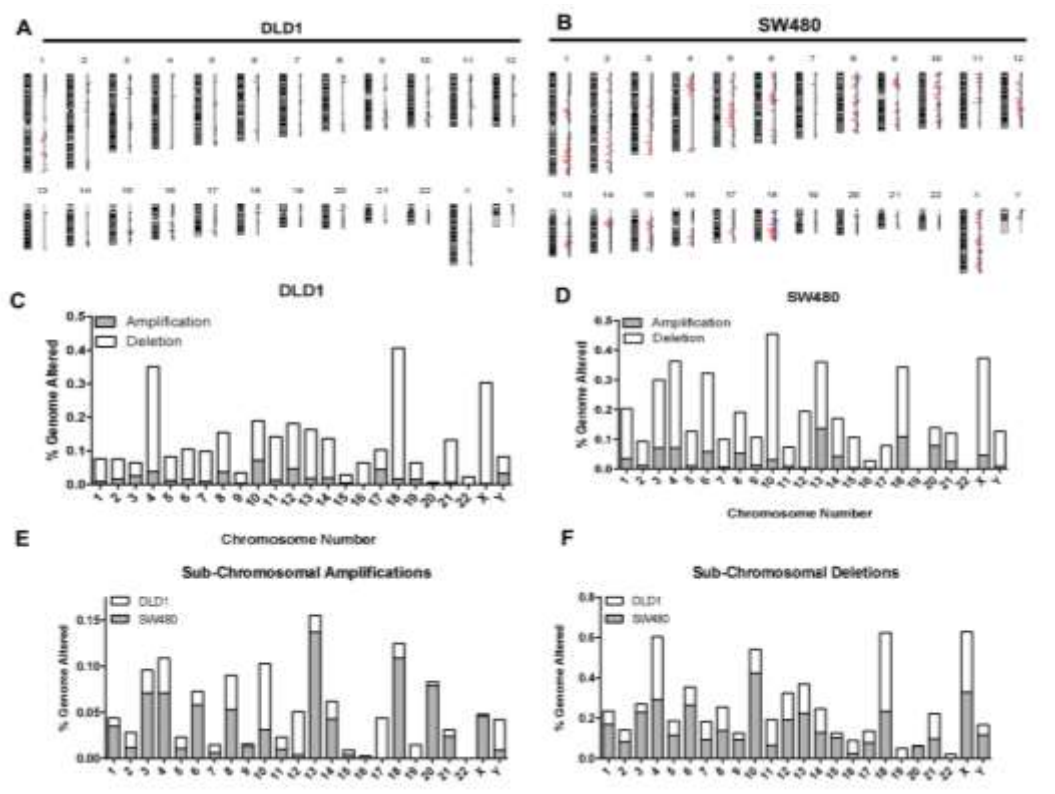
Fig.4.10: Effect of Twist1 overexpression on copy number alterations in colorectal cancer cells (A) Representative chromosomal ideograms showing sub-chromosomal deletions and amplifications in DLD1 upon Twist1 overexpression **(B)** Representative chromosomal ideograms showing sub-chromosomal deletions and amplifications in SW480 upon Twist1 overexpression. Red indicates deletion and Blue indicates amplification. **(C)** Sub-chromosomal amplifications and deletions, quantified for each chromosome, normalized to its total DNA content for DLD1 **(D)** Sub-chromosomal amplifications and deletions, quantified for each chromosome, normalized to its total DNA content for SW480. **(E)** Sub-chromosomal amplifications, quantified for each chromosome, normalized to its total DNA content for DLD1 and SW480 cells **(F)** Sub-chromosomal deletions, quantified for each chromosome, normalized to its total DNA content for DLD1 and SW480 cells. The array CGH was from two independent biological replicates (N = 2, mean).

Sub-chromosomal deletions were more prevalent upon Twist1 overexpression (**Fig.4.10E&F**). Surprisingly, human Chr.4, Chr.10, Chr.18 and Chr.X showed a significantly greater extent of sub-chromosomal deletions in DLD1 cells as compared to other chromosomes **(Fig.4.10C)**. SW480 cells on the other hand, showed a larger repertoire of sub-chromosomal deletions that predominantly map to human Chr.3, Chr.4, Chr.6, Chr.10, Chr.13, Chr.18 and Chr.X

(**Fig.4.10D**). Human chromosomes 4, 10, 18 and X consistently showed deletions in both cell lines upon Twist1 overexpression (**Fig.4.10F**). The extent of sub-chromosomal deletions was considerably elevated in SW480 than DLD1 cells. In summary, array CGH analyses revealed a significant increase in the frequency of sub-chromosomal deletions upon Twist1 overexpression- an additional contributor of CIN and genome instability (Rajagopalan et al., 2003).

Taken together, array CGH data suggests that Twist1 overexpression leads to copy number alteration of important cancer associated genes in colorectal cancer cells.

Several studies have demonstrated that copy number alterations are associated with a deregulated transcriptome. We identified the genes closest to the sub-genomic regions that were amplified or deleted upon Twist1 overexpression. Next, we identified the commonly amplified or deleted genes in DLD1 & SW480 (**Fig.4.11A&B**).

Figure 4.11
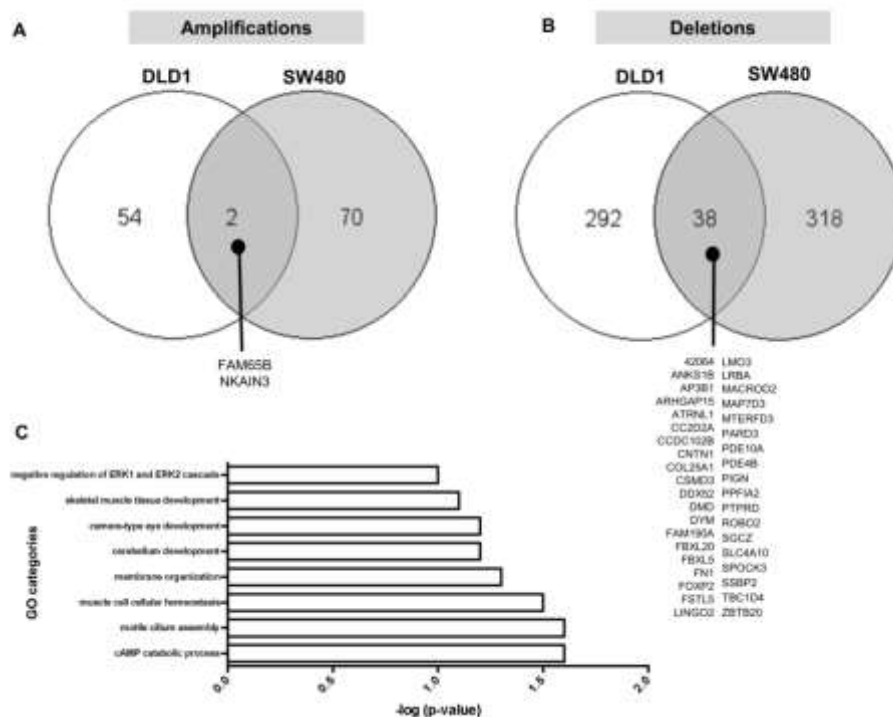
Fig.4.11: Common genes showing copy number alterations in DLD1 and SW480 cells upon Twist1 overexpression (**A**) Venn diagram showing genes amplified in both DLD1 and SW480 upon Twist1 overexpression from array CGH data (**B**) Venn diagram showing genes deleted in both DLD1 and SW480 upon Twist1 overexpression from array CGH data (**C**) Gene ontology (GO) analyses of genes commonly deleted in both the cell lines shown in (B).

Two genes, FAM65B (alias RIPOR2-RHO Family Interacting Cell Polarization Regulator) and NKAIN3 (Sodium/Potassium Transporting ATPase Interacting 3) were amplified in both the cell lines. Interestingly, DLD1 and SW480 cells showed 38 deleted genes common to both cell lines., It is intriguing that Fibronectin1 (FN1), a bonafide marker of EMT and contactin (CNTN1) associated with EMT in cancers (Chen et al., 2015), are deleted in both the cell lines. Gene oncology analysis of the 38 genes revealed enrichment of cancer associated pathways like negative regulation of ERK1 and ERK2 cascade, and membrane organization. We also performed GO analysis for genes that are individually amplified and deleted in the two cell lines **(Fig.4.12 & Fig.4.13)**.

Figure 4.12
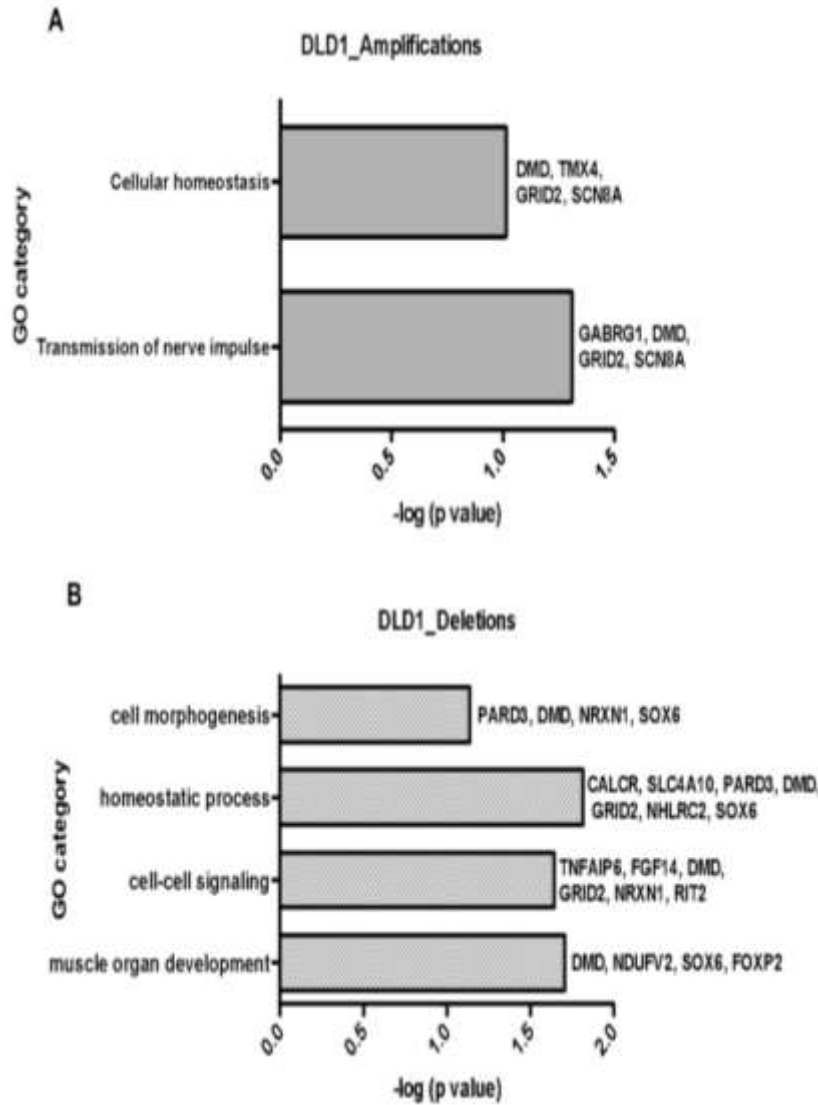
Fig.4.12: Gene ontology analyses of copy number alterations in DLD1 cells upon Twist1 overexpression (A) Gene ontology analysis of genes amplified in DLD1 cells, from array CGH data (B) Gene ontology analysis of genes deleted in DLD1 cells, from array CGH data. While DLD1 shows deletion of genes belonging to cell-cell signaling and cell morphogenesis, SW480 shows important cancer associated pathways being affected, like amplification of genes involved in cytoskeleton organization like LMO7, DLC1, FLNB, TNKS, SGTZ, and deletion of genes involved in cell to cell adhesion, cell projection organization, cytoskeleton reorganization and microtubule-based processes.

Figure 4.13
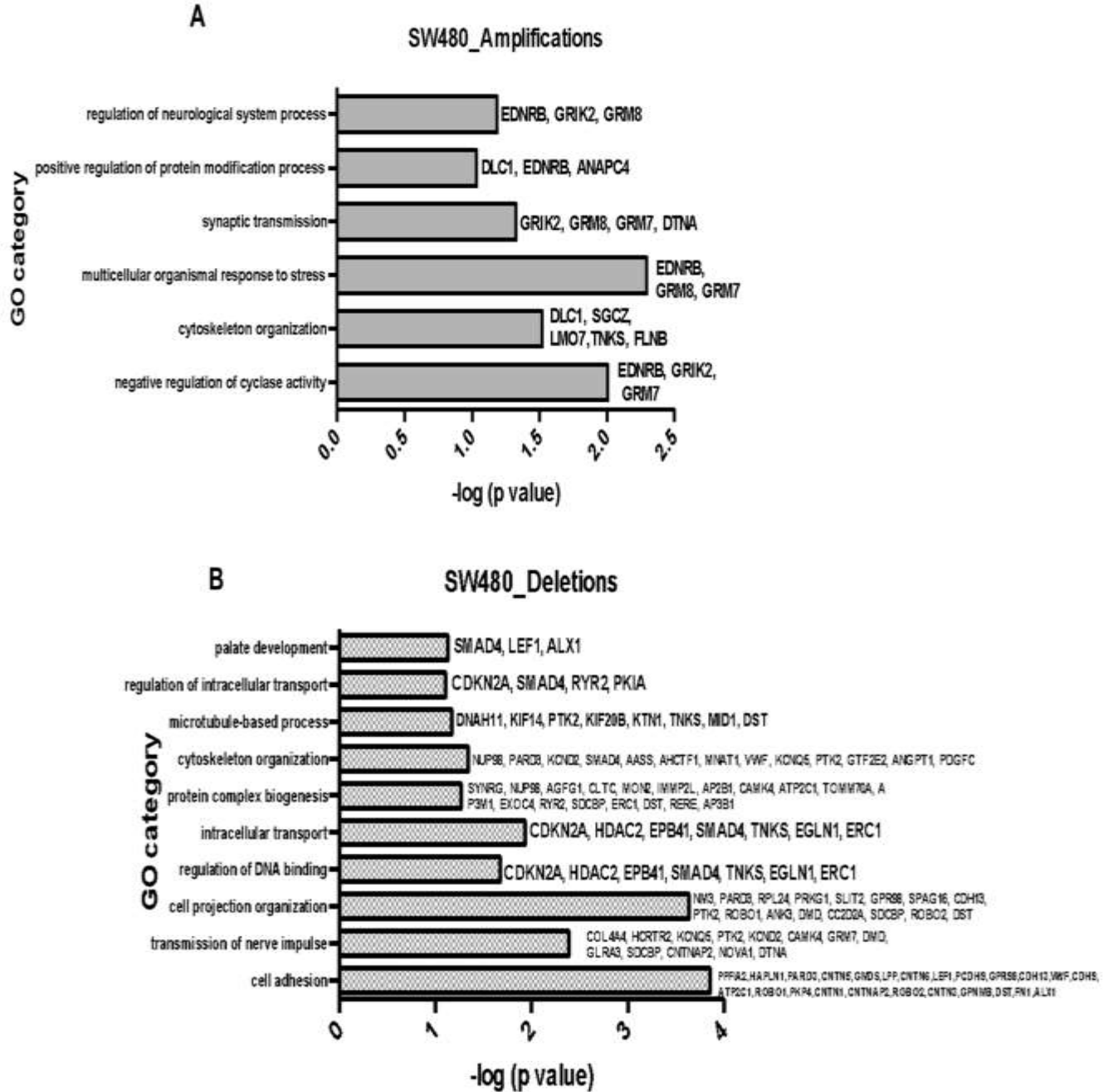
Fig.4.13: Gene ontology analyses of copy number alterations in SW480 cells upon Twist1 overexpression(A) Gene ontology analysis of genes amplified in SW480 cells from array CGH data **(B)** Gene ontology analysis of genes deleted in SW480 cells from array CGH data.

4.2.5 Twist1 overexpression induces DNA damage and downregulates p53

A noteworthy finding from genome wide array CGH analyses was the striking increase in sub-chromosomal deletions across the genome upon Twist1 overexpression. Since sub-chromosomal deletions are a consequence of DNA Double Strand Breaks (DSBs) (Cannan and Pederson, 2016; Varga and Aplan, 2005), we examined whether Twist1 overexpression induces DNA DSBs in colorectal cancer cells. We monitored the number of γH2AX foci as a marker of DSBs upon Twist1 overexpression, in the context of cisplatin-induced DNA damage **(Fig.4.14 A&B)**.

Figure 4.14
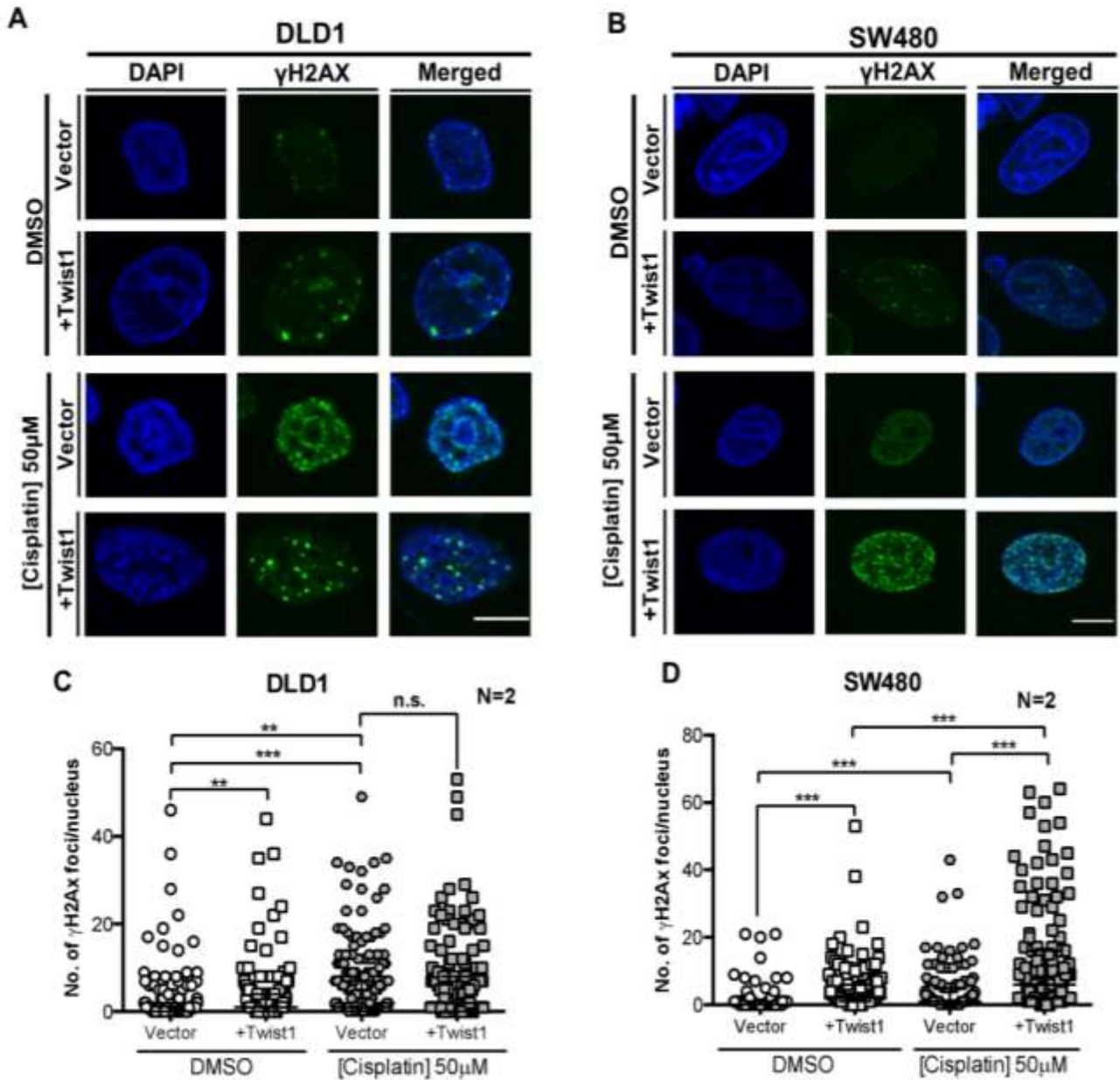
Fig.4.14: **Effect of Twist1 overexpression on DNA damage response** **(A)** Representative mid-optical sections of Cisplatin treated Twist1 overexpressing DLD1 cells, immunostained for γH2AX foci. DMSO (vehicle) treated cells were used as control. Scale bar ~10 μm **(B)** Representative mid-optical sections of Cisplatin treated Twist1 overexpressing SW480 cells, immunostained for γH2AX foci. DMSO (vehicle) treated cells were used as controlScale bar ~10 μm **(C)** Quantification of γH2AX foci in DLD1 **(D)** Quantification of γH2AX foci in SW480 cells. (Mann-Whitney test, N=2, n> 130, Median-IQR, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates, n: number of nuclei.

Interestingly, Twist1 overexpression in DLD1 cells showed a significant increase in the number of γH2AX foci in the interphase nucleus **(Fig.4.14C)**. However, cisplatin treatment in the background of Twist1 overexpression did not alter the number of DNA damage foci **(Fig.4.14C)**. In contrast, Twist1 overexpression in SW480 cells showed a significant increase in γH2AX foci independently, and in the presence of cisplatin **(Fig.4.14D)**. This suggests that Twist1 overexpression enhances DNA double strand breaks in colorectal cancer cells in a differential manner, in the two cell lines.

Since p53 is a master regulator of genome integrity in mammalian cells (Agapova et al., 1996), we determined the effect of Twist1 overexpression on the levels of p53. As previously reported in cancer cell lines (Meng et al., 2018; Piccinin et al., 2012), Twist1 overexpression showed a decrease in p53 levels in DLD1, and a marginal decline in SW480 cells (**Fig.4.15A-B**).

Figure 4.15
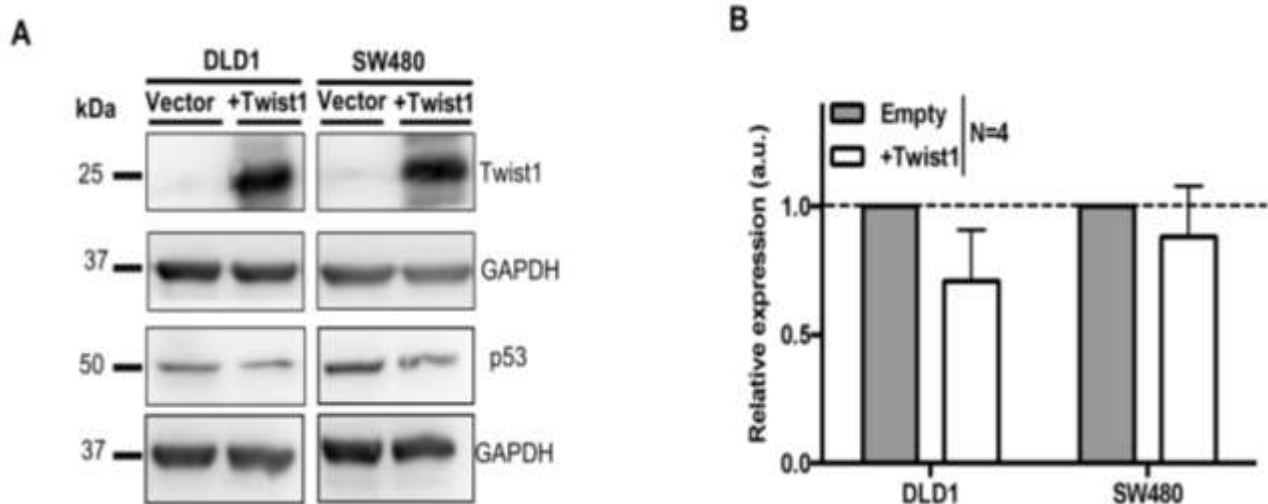
Fig.4.15: Effect of Twist1 overexpression on p53 expression in colorectal cancer cells (A) A representative immunoblot showing p53 levels upon Twist1 overexpression in DLD1 and SW480 cells **(B)** Quantification of p53 protein levels from band intensities normalized to GAPDH (unpaired t-test, N=4, Mean±SD, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates.

Of note, both DLD1 and SW480 cell lines have mutated p53 **(Table 4.1)**. In both cell lines, the mutation renders p53 non-functional. Thus, we do not understand the functional implication of

Twist1 induced downregulation of mutant p53 in the two cell lines, and how it contributes to CIN and/or potentially predisposes cells to elevated levels of DNA damage in cancer cells.

In order to understand the importance of functional p53 in induction of CIN upon Twist1 OE, we overexpressed Twist1 in another near diploid colorectal cancer cell line HCT116, which is (i) wild type for p53 and (ii) shows microsatellite instability (MSI+). Remarkably, HCT116 cells did not show any change in their modal chromosome numbers of 42-43, upon Twist1 overexpression **(Fig.4.16)**. This is consistent with an overarching role for wild type p53 protein in the maintenance of chromosomal stability in colorectal cancer cells. p53 is known to suppress the induction of CIN by inducing growth arrest and apoptosis, in cells with mitotic defects and DNA damage as seen in HCT116 cells (Dalton et al., 2010).

Figure 4.16
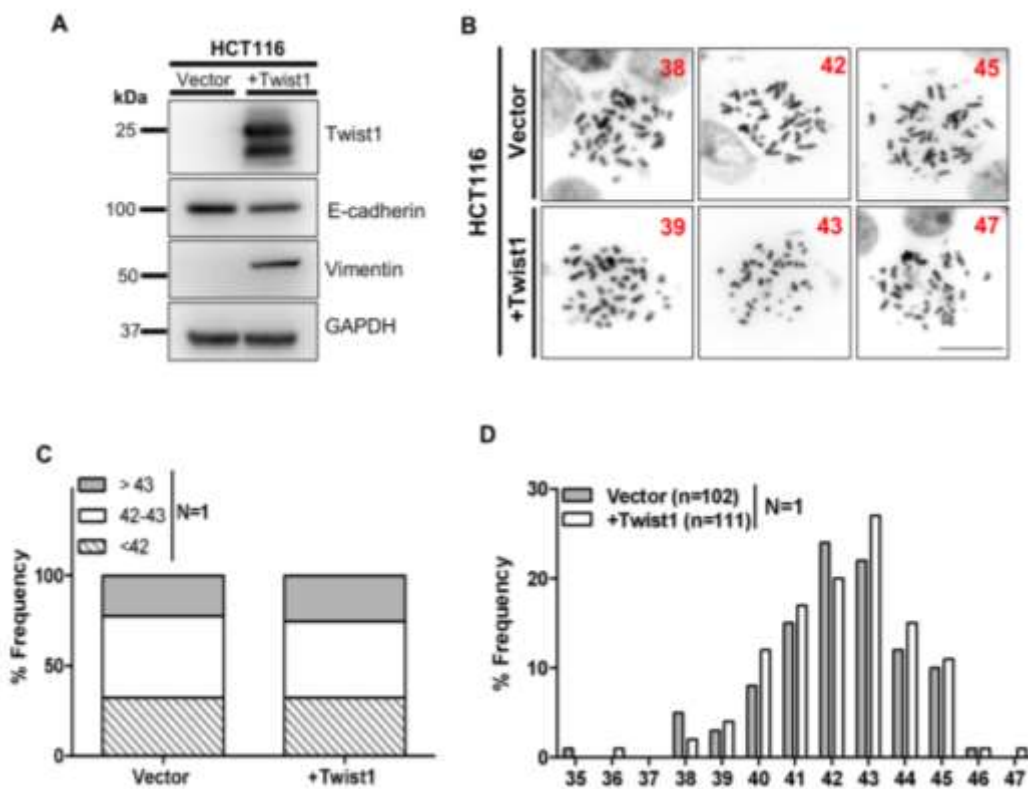
Fig.4.16: Twist1 overexpression in HCT116 colorectal cancer cells (A) Representative immunoblot showing EMT induction upon Twist1 overexpression **(B)** Representative images of metaphase spreads derived from HCT116 upon Twist1 overexpression. Scale bar ~20 μm **(C)** Quantification of whole chromosomal gains and losses in HCT116 cells **(D)** Histogram showing quantification of metaphase spreads derived from cells treated with vector control and Twist1 overexpression (N=1, n>100, N: number of independent biological replicates, n:number of metaphase spreads.

4.2.6 Mechanistic insight into Twist1 induced Chromosomal Instability (CIN)

We sought to address the underlying mechanisms leading to CIN upon Twist1 overexpression. Nuclear lamins that localize at the inner nuclear envelope, maintain the nuclear structure and function (Gerace and Huber, 2012). Lamins also modulate chromosomal stability in colorectal cancer cells (Kuga et al., 2014; Ranade et al., 2017). Therefore, we examined the lamin levels upon Twist1 overexpression. Interestingly, lamin levels decreased in both cell lines, with B-type lamins showing a further decrease in SW480, as compared to DLD1 upon Twist1 overexpression **(Fig.4.17A-C)**. These results suggest that Twist1 overexpression induces a decrease in lamin levels. Along with induction of chromosomal instability in colorectal cancer cells, decrease in lamin levels is also consistent with an increase in aberrant nuclear shapes (Smith et al., 2018).

Figure 4.17
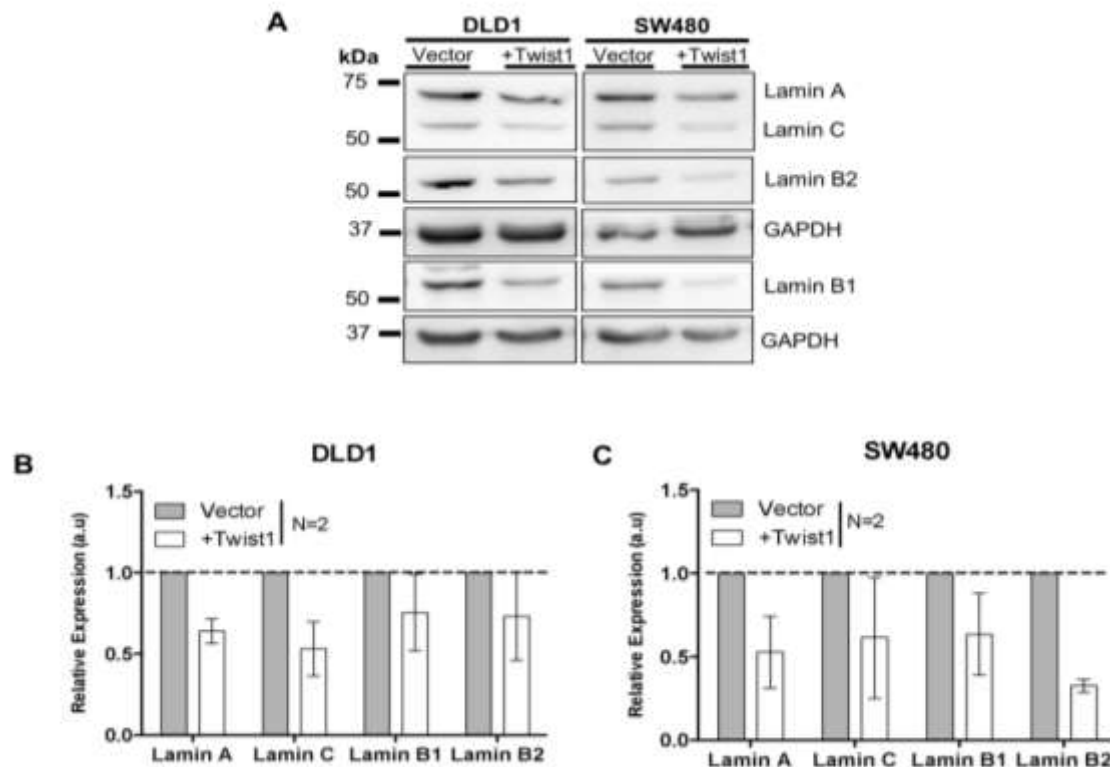
Fig.4.17: Effect of Twist1 overexpression on lamins in colorectal cancer cells (A) Representative immunoblot showing downregulation of lamin levels upon Twist1 overexpression in DLD1 and SW480 cells (N=2) **(B)** Quantification of lamin protein levels from band intensities normalized to GAPDH (unpaired t-test, N=2, Mean±SD, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$). N: number of independent biological replicates.

As we detected a significant increase in mitotic defects and whole chromosomal aberrations, we examined the levels of Spindle Assembly Checkpoint (SAC) proteins namely, Bub1, BubR1, Mad1, Mad2 and Aurora B Kinase upon Twist1 overexpression ((Ditchfield et al., 2003) (Burke and Stukenberg, 2008) (Diogo et al., 2017) (Elowe, 2011) (Logarinho et al., 2004)). We overexpressed Twist1, followed by immunoblotting assays on whole cell extracts derived from these cells. Remarkably, the levels of the Bub1 and BubR1 proteins showed a significant decrease in both colorectal cancer cell lines **(Fig.4.18A-C)**. Mad1 and Mad2- components of the mitotic checkpoint complex- also showed a decrease upon Twist1 overexpression **(Fig.4.19A-C)**.

Figure 4.18
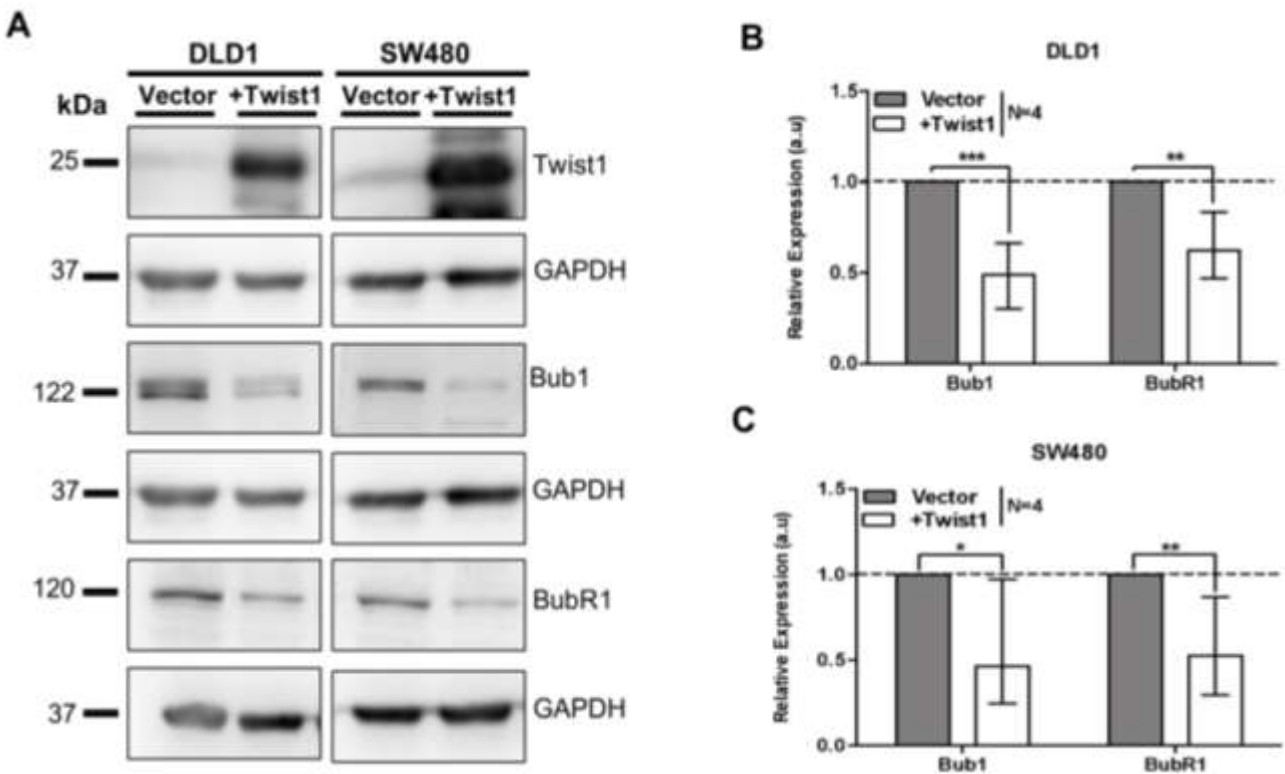
Fig.4.18: Effect of Twist1 overexpression on spindle assembly checkpoint proteins-Bub1 and BubR1 (A) A representative immunoblot showing Twist1 overexpression, accompanied by Bub1 and BubR1 levels in DLD1 and SW480 cells. **(B)** Quantification of Bub1 and BubR1 protein levels from DLD1 cells, calculated from band intensities normalized to GAPDH **(C)** Quantification of Bub1 and BubR1 protein levels from SW480 cells, calculated from band intensities normalized to GAPDH (unpaired t-test, N=4, Mean±SD, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates.

Figure 4.19
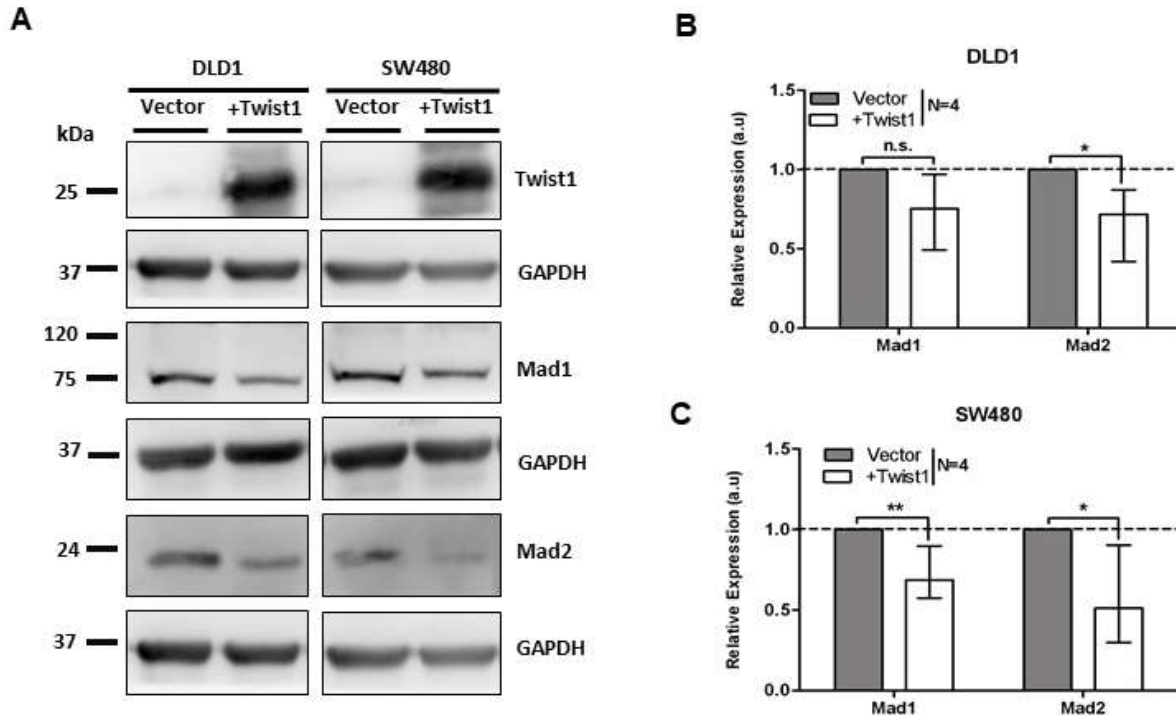
Fig.4.19: Effect of Twist1 overexpression on spindle assembly checkpoint proteins-Mad1 and Mad2 (A) A representative immunoblot showing Twist1 overexpression, accompanied by Mad1 and Mad2 levels in DLD1 and SW480 cells. **(B)** Quantification of Mad1 and Mad2 protein levels from DLD1 cells, calculated from band intensities normalized to GAPDH **(C)** Quantification of Mad1 and Mad2 protein levels from SW480 cells, calculated from band intensities normalized to GAPDH (unpaired t-test, N=4, Mean±SD, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates.

In addition, Aurora B Kinase- a part of the chromosome passenger complex, showed a decrease in protein levels **(Fig.4.20A-C)**. CIN is a consistent feature associated with the deregulation of Bub1/BubR1 levels ((Ditchfield et al., 2003) (Burke and Stukenberg, 2008) (Diogo et al., 2017) (Elowe, 2011) (Logarinho et al., 2004)) and decrease in their levels further affects the levels of downstream proteins such as Mad1/2 ((Ditchfield et al., 2003) (Johnson et al., 2004)). In summary, Twist1 overexpression shows a decrease in the levels of key regulators of chromosomal stability, which further underscores the contribution of Twist1 overexpression to CIN in colorectal cancer cells.

Figure 4.20
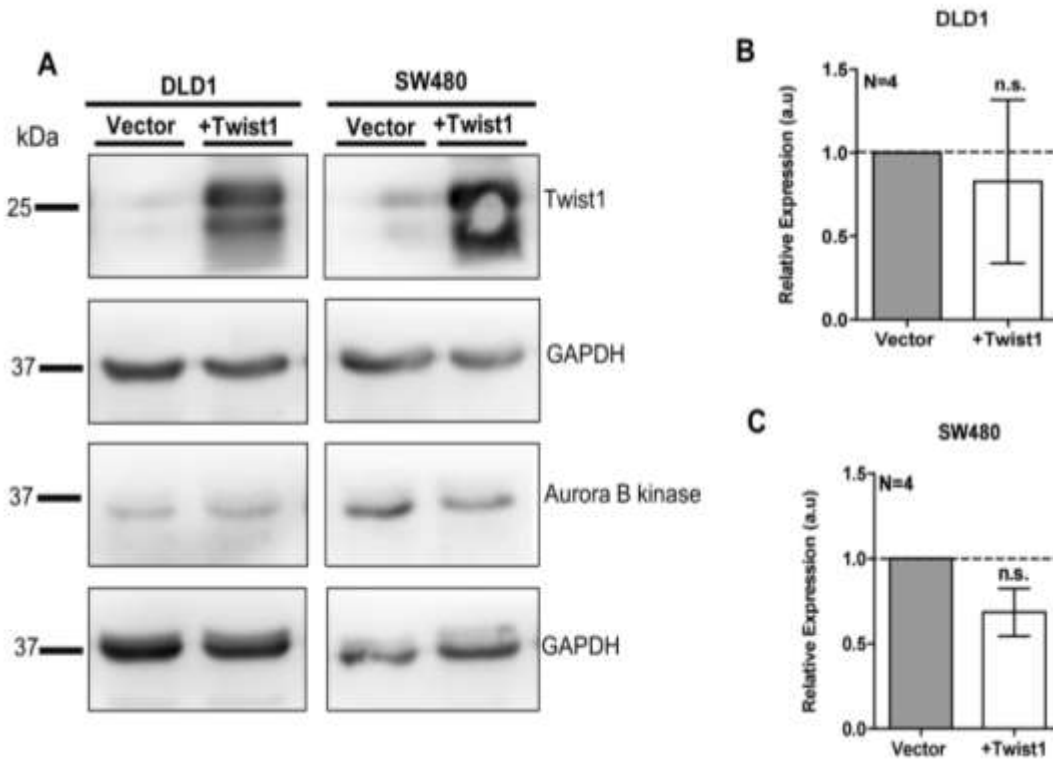
Figure 4.20: Effect of Twist1 overexpression on spindle assembly checkpoint protein-Aurora B kinase (A) A representative immunoblot showing Twist1 overexpression and Aurora B kinase levels in DLD1 and SW480 cells. **(B)** Quantification of Aurora B kinase protein levels from DLD1 cells, calculated from band intensities normalized to GAPDH **(C)** Quantification of Aurora B kinase protein level from SW480 cells, calculated from band intensities normalized to GAPDH (unpaired t-test, N=4, Mean±SD, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates.

Furthermore, we asked if the levels of checkpoint proteins upon Twist1 overexpression were regulated at the transcript level. We therefore performed RT-PCR analyses of checkpoint genes upon Twist1 overexpression. Twist1 showed transcript-level overexpression in both cell lines

Figure 4.21
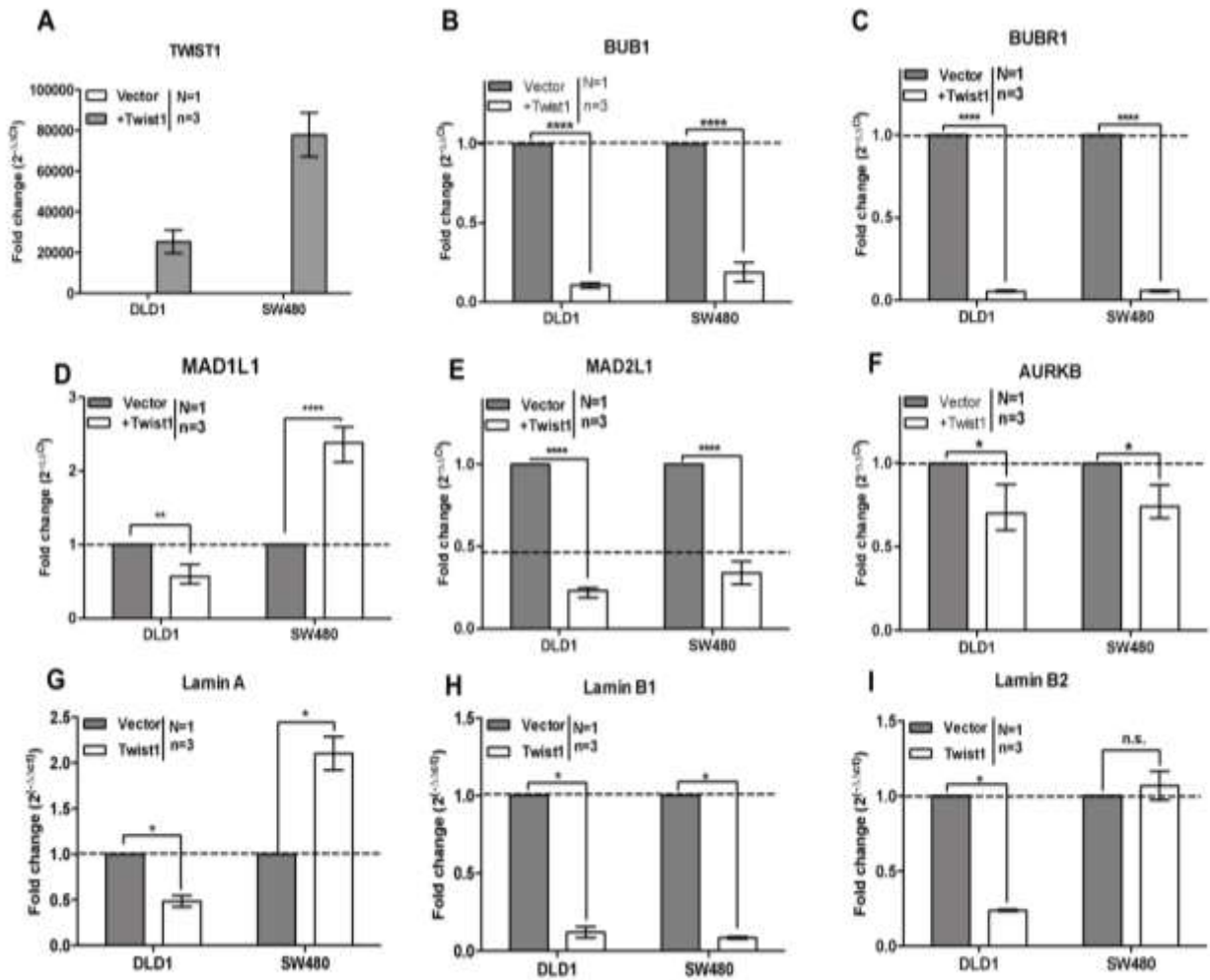
Fig.4.21: Effect of Twist1 overexpression on transcript levels of checkpoint genes and lamins (A) Gene expression levels of Twist1 by RT-PCR upon Twist1 overexpression in DLD1 and SW480 cells, normalized to vector control **(B-F)** Transcript level expression profiling of checkpoint genes determined by RT-PCR in DLD1 and SW480 **(G-I)** Transcript level expression profiling of lamins determined by RT-PCR in DLD1 & SW480. Data normalized to its respective vector control. (Data quantified from N=1, n=3 independent technical replicates, unpaired t-test, Mean±SD, *P < 0.05, **P < 0.01, ***P < 0.001 and ****P < 0.0001). N: number of independent biological replicates.

Interestingly, BUB1, BUBR1, MAD2L1 and AURKB showed a significant decrease in their transcript levels in both DLD1 and SW480 cell lines **(Fig.4.21B-C & E-F)**. In contrast, MAD1L1 showed a differential response as it was downregulated in DLD1, but significantly upregulated in SW480 cells, upon Twist1 overexpression **(Fig.4.21D)**. Additionally, all the three

types of lamins showed a downregulation at the transcript level in DLD1 cells **(Fig.4.21G-I)**. However, in SW480 cells, Lamin A showed upregulation, Lamin B1 showed a downregulation and Lamin B2 was unaltered **(Fig.4.21G-I)**.

In summary, Twist1 overexpression represses checkpoint genes which are required for the maintenance of chromosomal stability, at the transcript level, while lamins are differentially regulated in the two cell lines.

As lamins and SAC genes are regulated at the transcript level and Twist1 is a transcription factor, we examined whether Twist1 binds to the promoters of these genes and regulates their expression. We analyzed ChIP-Seq datasets available from human mammary epithelial cells (HMLE) overexpressing Twist1 (Chang et al., 2015). While Twist1 overexpression induces EMT in HMLE, it is unclear if Twist1 induces CIN in these cells. Analyses of Twist1 occupancy in HMLE cells shows that Twist1 is enriched on E-cadherin (CDH1 gene) at −29 kbp, +20 bp and +49 kbp, with respect to the Transcription Start Site (TSS) **(Table 4.4)**. However, Twist1 does not show promoter occupancy on kinetochore associated genes that we examined. Furthermore, Twist1 showed a relatively proximal occupancy to the TSS of the Lamin A/C (-2184bp and +3720bp) and Lamin B2 (+3756bp) consistent with transcriptional regulation upon Twist1 overexpression **(Table 4.4 & Fig4.21.G-I).**

Table4.4: ChIP-Sequencing analyses of Twist1 occupancy from HMLE (human mammary epithelial) cells upon Twist1 overexpression.

Gene Category	Gene Name	Occupancy	Chromosome	Start	End	Annotation	Distance to TSS (bp)
	CDH1	+	Chr.16	67298818	67299408	Intergenic	-29583
	CDH1	+	Chr.16	67328548	67328884	promoter-TSS (NM_004360)	+20
	CDH1	+	Chr.16	67377982	67378379	intron (NM_004360, intron 2 of 15)	+49484
EMT	VIM	-	Chr.10	N.A.	N.A.	N.A.	N.A.
	BUB1	-	Chr.2	N.A.	N.A.	N.A.	N.A.
	BUBR1	-	Chr.15	N.A.	N.A.	N.A.	N.A.
	MAD1L1	+	Chr.7	2201955	2202527	intron (NM_001013836, intron 10 of 18)	+36868
	MAD1L1	+	Chr.7	2122201	2122623	intron (NM_001013836, intron 11 of 18)	+116697
	MAD2L1	+	Chr.4	121632168	121632386	Intergenic	-424816
	AURKB	-	Chr.17	N.A.	N.A.	N.A.	N.A.
Checkpoint	TP53	-	Chr.17	N.A.	N.A.	N.A.	N.A.
	LMNA	+	Chr.1	154339840	154340272	intron (NM_001282625, intron 3 of 12)	-11029
	LMNA	+	Chr.1	154360004	154360693	intron (NM_005572, intron 1 of 9)	-2184
	LMNA	+	Chr.1	154365724	154366780	intron (NM_001282624, intron 2 of 10)	+3720
	LMNA	+	Chr.1	154355937	154356303	intron (NM_005572, intron 1 of 9)	+5035
	LMNB1	-	Chr.5	N.A.	N.A.	N.A.	N.A.
Nuclear Envelope	LMNB2	+	Chr.19	2404030	2404390	intron (NM_032737, intron 1 of 11)	+3756

To identify other genes that impact CIN, that may be regulated by Twist1, we identified genes that show promoter binding (−1 to +1 Kbp from TSS) of Twist1. Next, we performed Gene Ontology analyses of these genes. This analysis revealed the enrichment of Twist1 on p53-associated genes, that include MDM2, CHEK2 and CCNB1(**Fig.4.22A&B**). Thus, Twist1 may potentially modulate the p53 signaling pathway via MDM2, CHEK2 and CCNB1, suggestive of Twist1-dependent and Twist1-independent transcriptional regulation of genes that maintain chromosomal stability.

Figure 4.22
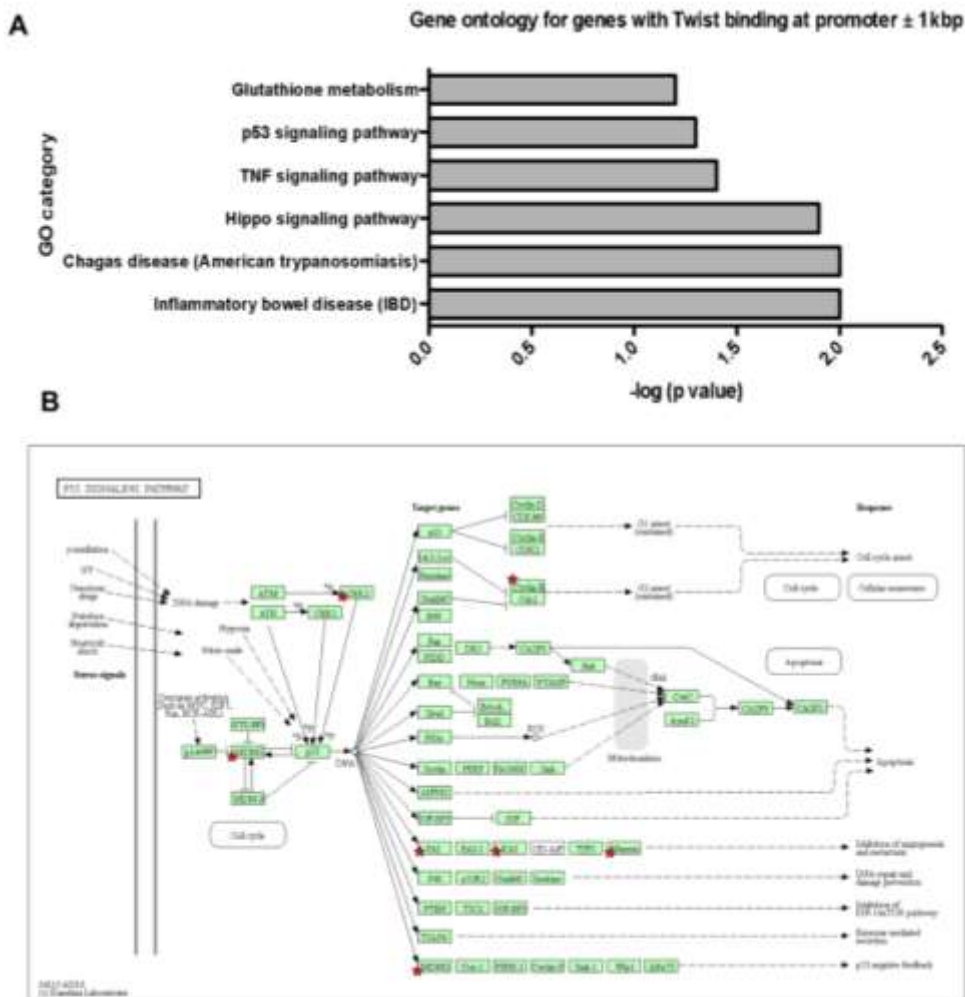
Fig.4.22: Enrichment of Twist1 on gene promoters of the p53 signaling pathway (A) Gene Ontology analyses of genes that show Twist1 occupancy enriched within (-1kbp to +1kbp) of the promoter region **(B)** p53 pathway from KEGG. Red stars indicate genes of the p53 pathway showing Twist1 occupancy in their promoter region.

4.2.7 Twist1 impinges on CIN regulation

To address the potential crosstalk between Twist1 and the regulators of chromosomal stability, we performed network analyses of (i) Generic Protein-Protein Interactions (Xia et al., 2014) (ii) Transcription Factor (TF)-Gene Interactions (Xia et al., 2013) of Twist1 and factors associated with EMT, CIN and DNA damage, using NetworkAnalyst **(Fig.4.23A&B)**. From protein-protein interaction network analyses, p53 emerges as a major hub through which Twist1 regulates CIN factors **(Fig.4.23A).** Twist1 is known to affect the DNA-binding activity of p53, thereby impairing function of p53 (Piccinin et al., 2012; Xia et al., 2014). In addition, p53 directly interacts with Aurora B Kinase and Bub1, while interacting with Mad2 via FZR1 **(Fig.4.23A)**. Interestingly, p53 and LMNA have a common interactor, SUMO1- a post-translation protein modifier, while p53 and LMNB1 may interact via CDK1- a cell cycle regulator **(Fig.4.23A)**.

Figure 4.23
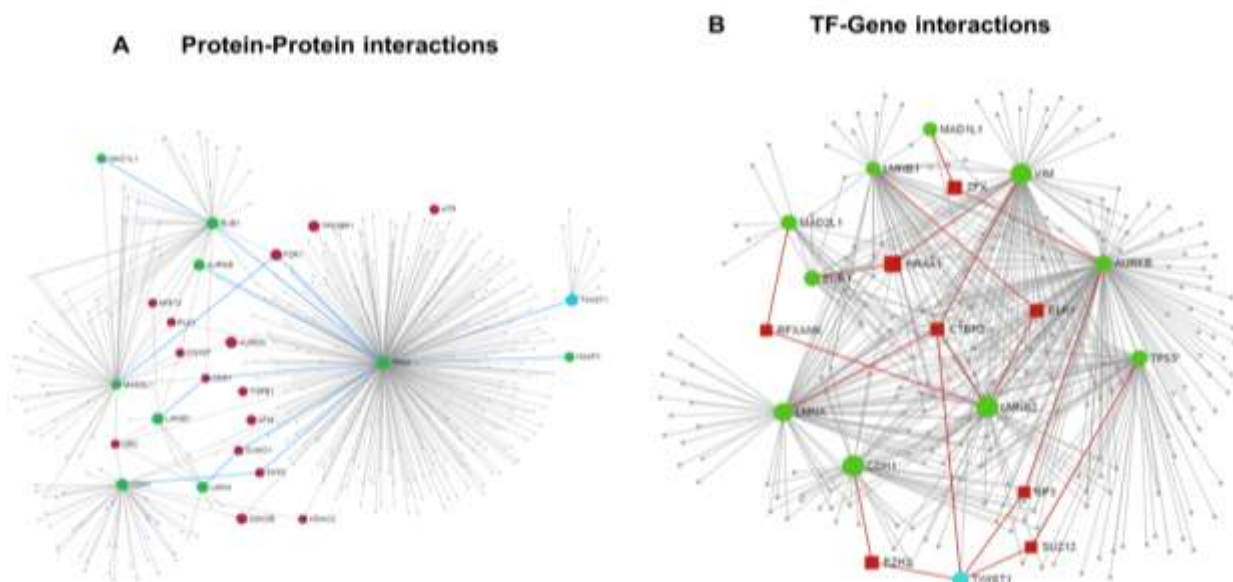
Fig.4.23: Regulatory network between Twist1, EMT and CIN factors- Network Analyst (A) Generic Protein-Protein interactions of Twist1 and other proteins examined in this study (E-cadherin, Vimentin, Bub1, BubR1, Mad1, Mad2, Aurora B kinase, p53, LaminA/C, LaminB1, LaminB2) generated using NetworkAnalyst. Blue line: Candidate and key protein-protein interactions in the context of Twist1, EMT and CIN factors. Gray line: All other protein-protein interactions. Candidate EMT and CIN factors reported in this study **(B)** Transcription factor-Gene interactions between TWIST1 & other target genes (CDH1, VIM, BUB1, MAD1L1, MAD2L1, AURKB, TP53, LMNA, LMNB1 and LMNB2) examined in this study, generated using NetworkAnalyst.

ENCODE ChIP-Seq data for Transcription Factor (TF) enrichment on target genes **(Fig.4.235B)**, shows that Twist1 may modulate TP53 activity via the histone modifier SUZ12- part of the polycomb repressive complex 2 (PRC2). Furthermore, Twist1 indirectly modulates Aurora B kinase activity via SP3 - a transcriptional repressor/activator (Majello et al., 1997). Additionally, Twist1 modulates CDH1 (E-cadherin) via EZH2 (component of the PRC2) (Malouf et al., 2013). Twist1 differentially regulates lamins (LMNA, LMNB2) through the transcriptional repressor CTBP2 (Zhang et al., 2018). RFXANK is enriched on LMNB2 and MAD2L1 genes. Also, NR4A1- a nuclear transcription factor emerges as a modulator of VIM (Vimentin) and BUB1.

In summary, this analysis reveals that Twist1 can regulate CIN, potentially by directly binding and regulating expression of key targets or by indirectly regulating genes that may in turn regulate CIN. Also, p53 came out to be an important node that can mediate Twist1 induced chromosomal instability.

4.2.8 A simulation-based approach shows negative correlation between Twist1 and E-cadherin, BubR1 levels

We next constructed a regulatory network by integrating our experimental data with known interconnections among Twist1, E-cadherin, Vimentin, BubR1, γH2AX and p53. (i) Twist1 overexpression downregulates E-cadherin by binding to its promoter (Vesuna et al., 2008) (ii) Twist1 overexpression upregulates Vimentin and is mediated by CUL2 (Meng et al., 2018) (iii) Twist1 overexpression in colorectal cancer cells downregulates BubR1 **(Fig.4.18&4.21)** (iv) Twist1 downregulates p53 levels (Meng et al., 2018; Piccinin et al., 2012) (v) BubR1 levels positively correlate with levels of p53 and γH2AX in response to DNA damage (Fang et al., 2006) (vi) p53 is required for repair of DNA damage and shows a negative correlation with γH2AX levels (Van Oorschot et al., 2014).

We sought to identify the robust dynamic features emerging from these interconnections. We simulated the network using the RACIPE algorithm (Huang et al., 2017).

Figure 4.24
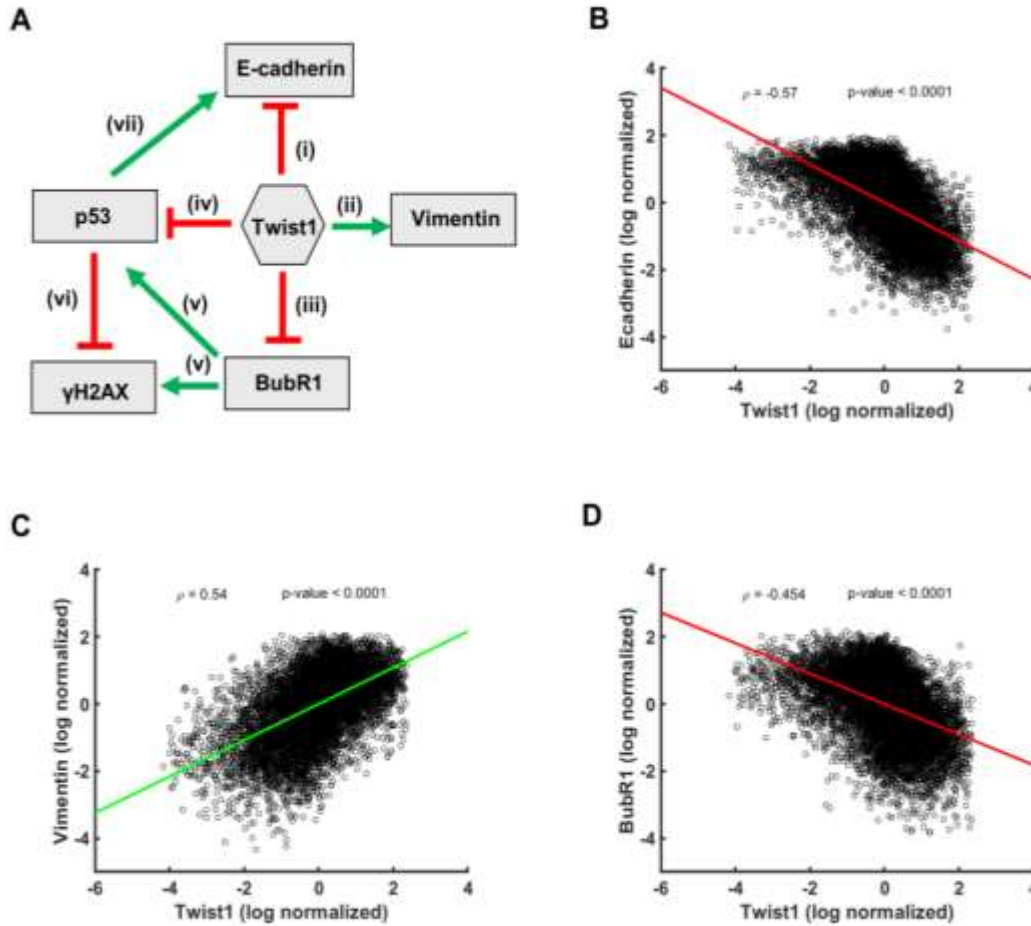
Fig.4.24: RACIPE analysis for correlation between levels of Twist1, EMT and CIN factors (A) Network depicting the interactions among Twist1, EMT and CIN genes. Correlation plots of the log normalized gene expression values of **(B)** Twist1 and E-cadherin **(C)** Twist1 and Vimentin **(D)** Twist1 and BubR1. (ρ = Pearson correlation coefficient, P-values show the significance of Pearson correlation).

Across the sampled parameter sets, we observed a significant negative correlation for Twist1- E-cadherin **(Fig.4.24B),** and positive correlation for Twist1-Vimentin **(Fig.4.24C)**. This suggests that although the extent of EMT induction via Twist1 is heterogeneous across single cells, an ensemble behavior shows robust induction of EMT by Twist1 by altering the levels of E-cadherin and Vimentin. Notwithstanding the intrinsic heterogeneity across cells, a systems biology approach corroborates our experimental data which shows that Twist1 and BubR1 expression levels are negatively correlated **(Fig.4.24D)**.

4.2.9 Twist1 overexpression positively correlates with EMT and CIN- The Cancer Genome Atlas (TCGA) analyses

We sought to ask if the expression levels of Twist1 correlate with levels of (i) EMT associated genes (ii) chromosomal instability (CIN) genes (iii) DNA double-strand break (DSB) genes, and (iv) tumor mutation burden and copy number alterations (CNA) in cancer patients. We analyzed gene expression data and somatic mutations of 30 distinct cancer types from The Cancer Genome Atlas (TCGA). While Twist1 expression positively correlates with primary tumors, their level varies within and between cancer types, likely due to the cell-of-origin and tumor stage differences **(Fig.4.25A)**. For example, liver hepatocellular carcinoma (LIHC) and kidney cancers (KIPAN), which originates from epithelial cells, showed the least Twist1 expression, whereas sarcoma (SARC) and Uterine carcinosarcoma (UCS) which originates from mesenchymal cells showed higher expression, as previously shown at the level of EMT gene signatures (Gibbons and Creighton, 2018). Colorectal cancers (COADREAD) also showed a high expression of Twist1, especially in the late stage tumors **(Fig.4.25B)**.

Fig.4.25
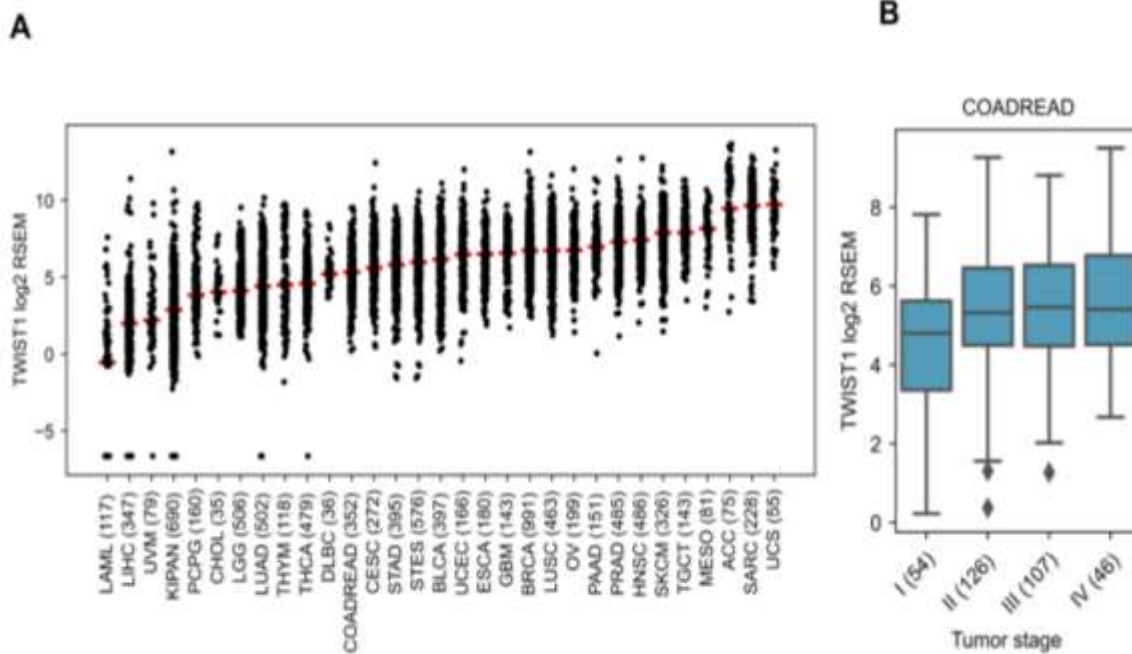
Fig.4.25: TCGA analyses for Twist1 expression levels across cancer subtypes (A) TWIST1 gene expression (log2 RSEM + 0.01) level within and across cancers of TCGA. Each dot represents a tumor sample and the horizontal red bar indicates the median expression value within that cancer cohort. The cancer type abbreviations are shown below. **(B)** TWIST1 expression in colorectal cancers (COADREAD) stratified by tumor stages.

Furthermore, Twist1 expression positively correlates with EMT associated genes (CDH1, OCLN, TJP1, CDH2, FN1, SNAI1, and VIM) in various cancers **(Fig.4.26)**. A significant and positive correlation was detected between expression of CIN genes and certain tumor types which includes kidney cancers, lower-grade glioma (LGG) and lung adenocarcinoma (LUAD); whereas, in stomach adenocarcinoma (STAD) showed a significant negative correlation between Twist1 and CIN genes. However, consistent with experimental data **(Fig.4.19A&C)**, colorectal cancers showed a moderate negative correlation between Twist1 and MAD2L1- a mitotic spindle assembly checkpoint protein. AURKC expression showed a marginal increase with Twist1 expression **(Fig.4.26)**. However, we found a significant correlation with DSB genes in only a few tumor types (for example, STAD showed a strong negative correlation similar to CIN genes).

Figure 4.26
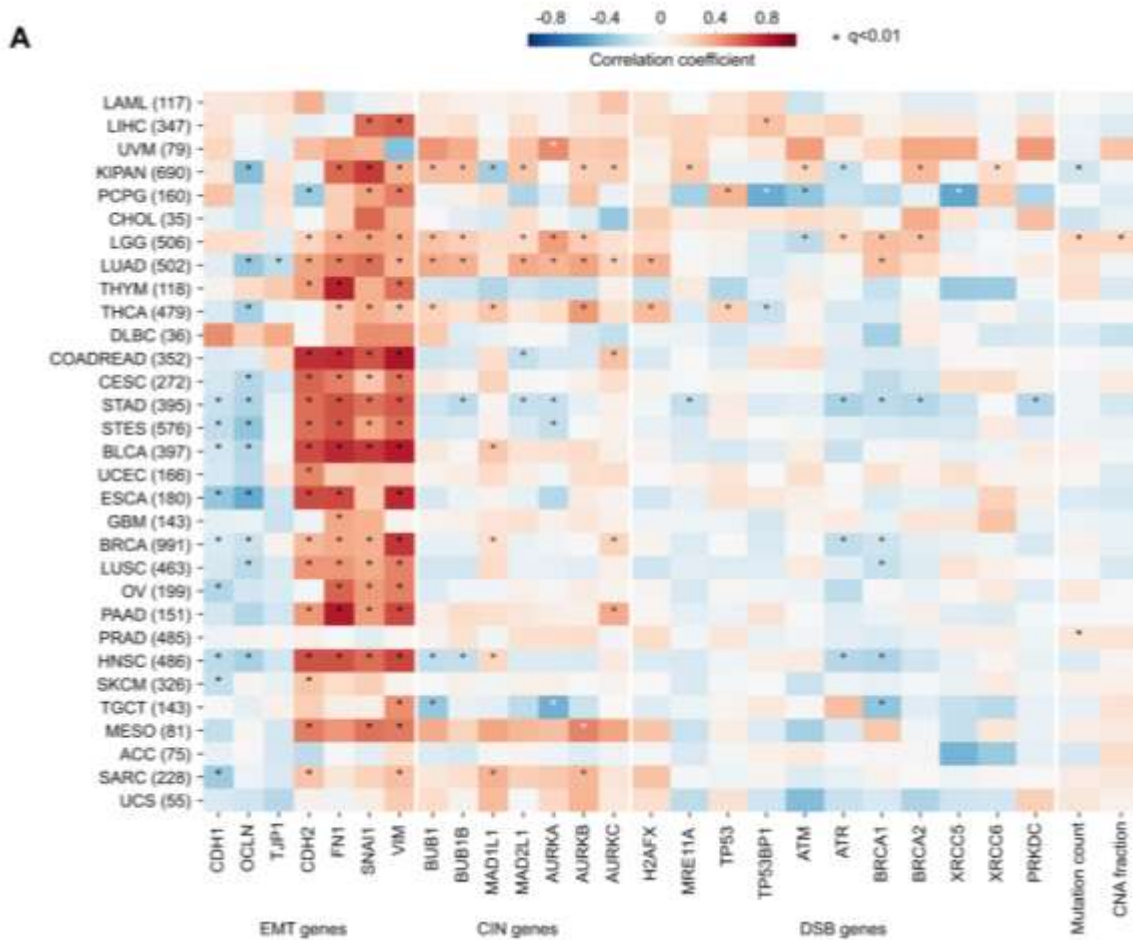
Fig.4.26: TCGA analyses for correlation between levels of TWIST1, EMT and CIN across human cancers (A) Heatmap representing the correlation coefficient (and its significance) between TWIST1 expression and (i) EMT, (ii) CIN, (iii) DSB gene expression, (iv) mutation count and (v) copy number alteration (CNA) fraction for all 30 distinct cancers, computed using an iteratively reweighted least-squares approach. Color coding indicates the correlation coefficient, ranging from −1 to +1, where −1 being strong negative correlation (dark blue), 0 for no correlation (white) and +1 strong positive correlation (dark red). Significant correlations (q < 0.01) are marked with an asterisk (∗). Cancer type abbreviations: ACC: adrenocortical carcinoma, BLCA: bladder urothelial carcinoma, BRCA: breast invasive carcinoma, CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma, CHOL: cholangiocarcinoma, COADREAD: colorectal adenocarcinoma, DLBC: diffuse large B-cell lymphoma, ESCA: esophageal carcinoma, GBM: glioblastoma multiforme, HNSC: head and neck squamous cell carcinoma, KIPAN: pan kidney carcinomas, LAML: acute myeloid leukemia, LIHC: liver hepatocellular carcinoma, LGG: brain lower grade glioma, LUAD: lung adenocarcinoma, LUSC: lung squamous cell carcinoma, MESO: mesothelioma, OV: ovarian serous cystadenocarcinoma, PAAD: pancreatic adenocarcinoma, PCPG: pheochromocytoma and paraganglioma, PRAD: prostate adenocarcinoma, SARC: sarcoma, SKCM: skin cutaneous melanoma, STAD: stomach adenocarcinoma, TGCT: testicular germ cell tumors, THCA: thyroid carcinoma, THYM: thymoma, UCS: uterine carcinosarcoma, UCEC: uterine corpus endometrial carcinoma, UVM: uveal melanoma.

We also examined the correlation between EMT and CIN .We compared Twist1 expression with tumor mutation burden and copy number alterations (CNA) **(Fig.4.26)**. The total number of somatic point mutations and the fractions of the genome with amplifications or deletions were considered as tumor mutation burden and CNA events, respectively. The cancer types LGG (Brain), KIPAN (Pan Kidney) and PRAD (prostate) showed a significant positive correlation with Twist1 expression for somatic mutation and for CNA events, in LGG **(Fig.4.26)**. In particular, LGG showed a significant positive correlation between TWIST1 with (i) EMT (ii) CIN (iii) DSB gene expression, and (iv) mutation burden and CNA event. Taken together, these results suggest that Twist1 expression correlates with EMT in various cancers. However, the correlation between TWIST1 with CIN and DSB genes is cancer subtype specific.

4.3 Discussion

Epithelial to mesenchymal transition (EMT) is a remarkable model of cellular plasticity. While EMT was characterized as binary states, namely epithelial and mesenchymal, recent studies show that this transition is a continuum that involves intermediate states and that are referred to as E/M hybrids (Nieto et al., 2016). The incomplete EMT with intermediate states is called partial EMT (Jolly et al., 2015). Cells of epithelial cancers can exhibit both mesenchymal and epithelial characteristics, that is, the hybrid E/M phenotype. Colorectal cancer cell lines DLD1 and SW480, although both of epithelial origins, are at extremes of the EMT spectrum. DLD1 is more epithelial while SW480 is more mesenchymal-like (Sacchetti et al., 2020). This is an important determinant of how Twist1 induces EMT in these cell lines.

Here we show that Twist1 overexpression leads to chromosomal instability in both DLD1 and SW480 cells **(Fig.4.9)**. DLD1 cells show both losses and gains of chromosomes, while SW480 cells predominantly show losses. A previous TCGA analysis of colorectal carcinoma patients has revealed that tumorigenesis requires gains of human chromosomes Chr.7, 1q, 8q, 13q and 20q and a loss of Chr.4, 1p, 5q, 8p, 14q, 15q, 17p & 18q (Cancer Genome Atlas Network, 2012). We asked if progression of CRC induced by EMT affected the ploidy of specific chromosomes in the two cell lines. Array CGH analysis reveals that no specific chromosomes are lost or gained in the two cell lines, upon Twist1 overexpression **(Fig.4.10)**. In an independent study, array CGH analyses performed on colorectal cancer patient tumors identified copy number aberrations and exhibited sub-chromosomal deletions in human chromosomes 4, 8, and 18 (Camps et al., 2006; Douglas et al., 2004), (Douglas et al., 2004). In a comparable manner, DLD1 and SW480 cell lines showed sub-chromosomal deletions in chromosomes 4, 8, 10, 18 and X, and chromosomes 3, 4, 10, 13, 18 and X, respectively. Twist1 overexpression also induces CIN in MCF-7 breast cancer cells. Spectral karyotyping (SKY) analyses show tetrasomy (~4 copies) of most chromosomes except human Chr.2, 3, 12, 18 and 21, upon Twist1 overexpression (Vesuna et al., 2006).

Furthermore, Twist1 overexpression is shown to increase DNA double strand breaks, as seen by increase in γH2AX foci **(Fig.4.14)**. DNA double strand breaks contribute to chromosome missegregation events and increased subchromosomal deletions (Bakhoum et al., 2014). Both DLD1 and SW480 cells show higher frequency of deletions upon Twist1 overexpression

(Fig.4.10). Chromosomal instability may further drive phenotypic switching between epithelial and mesenchymal fates, by deletions and amplifications of genes associated with these two cell states. For example, loss of regions of chromosomes like chromosome 16q, which harbours the *E-cadherin* gene, can induce mesenchymal phenotype, whereas loss of chromosome arm 10p on which *Zeb1* gene is present, induces an epithelial phenotype (Gao et al., 2016).

Twist1 associated EMT enhances nuclear deformations such as blebs and micronuclei **(Fig.4.6).** Micronuclei are commonly detected in cancer cells and contribute to genomic instability (Hatch et al., 2013; Zhang et al., 2015)). Previous studies show that micronuclei during interphase, result in loss of nuclear envelope integrity, further contributing to DNA double strand breaks (DSBs) and chromothripsis, potentially facilitated by the entry of cytoplasmic content in the micronucleus (Hatch et al., 2013; Zhang et al., 2015). The altered micronucleus DNA is then re-integrated into the nucleus during subsequent cell divisions, contributing to genomic instability (Zhang et al., 2015). Reduced lamin levels induced by Twist1 overexpression **(Fig.4.17)**, weaken the nuclear envelope, leading to nuclear deformation. In addition, decrease in B-type lamins is an independent mechanism that contributes to CIN in colorectal cancer cell lines. Lamin B2 localizes outside the spindle poles during mitosis and has a critical role in preventing CIN in colorectal cancers, by maintaining spindle pole stability and spindle assembly (Kuga et al., 2014). We surmise that a Twist1-mediated decrease in lamin levels is an indirect means of contributing to CIN in colorectal cancers.

In addition, lamin loss affects chromatin organization and gene expression across cell types (Kuga et al., 2014). Furthermore, decreased levels of lamin A/C are associated with cancers (Wu et al., 2009). Therefore, decrease in nuclear lamins through Twist1 overexpression may impact chromatin organization and gene expression. Stem cells and undifferentiated cells are characterized by relatively reduced lamin levels and 'floppy' chromatin (Melcer et al., 2012). Excess levels of Twist1, and the concomitant reduction in lamin levels may induce stemness in transformed cells and create 'founder' populations of cancer stem cells, with elevated genomic instability and resilient sub-populations of cancer cells (Greaves, 2013).

Cells with lowered levels or lacking Mad2 can proliferate with inactivated p53 and tolerate high levels of genomic instability (Burds et al., 2005; Foijer et al., 2017). This may be why DLD1 and SW480 cells which lack a functional p53 can accommodate genomic instability, in spite of

lowered Mad2 levels by Twist1 **(Fig.4.19 & Table 4.1)**. Furthermore, the status of p53 is potentially an important determinant of CIN, since cells with mutant p53 are associated with CIN, while cells with wild type p53 show significantly reduced CIN in cancer cells (Ahmed et al., 2013) (Hanel and Moll, 2012). This is recapitulated by the fact that Twist1 cannot induce CIN in HCT116 cells, which have wild type p53. Future studies may be directed towards investigating if overexpressing wild type p53 in DLD1 and SW480 in the background of Twist1 overexpression can rescue CIN. However, in certain cases, CIN is also induced in cells with wild type p53, suggesting that p53 is not the sole determinant and the overall genetic background is critically linked to CIN induction (Vesuna et al., 2006) (Hanel and Moll, 2012).

The Cancer Genome Atlas (TCGA) is a useful repository of cancer patient data. It allows for various molecular correlates like expression status, mutations, and copy number alterations across cancer subtypes (Cancer Genome Atlas Research Network, 2014). Furthermore, mathematical modelling and simulations have the power to compute and predict the potential outcome of novel molecular interactions, and their pathways involved in actively promoting cancers. It is therefore beyond any doubt that an interdisciplinary approach of studying theoretical and experimental paradigms is essential for cancer intervention.

Taken together, these studies suggest that in addition to inducing EMT, Twist1 also enhances nuclear and mitotic aberrations that further contribute to chromosomal instability in colorectal cancers **(Fig.4.27)**. This is largely mediated by a collective decrease in levels of key checkpoint and genomic stability factors, underscoring the mechanistic involvement of Twist1 with CIN during EMT.

Figure 4.27
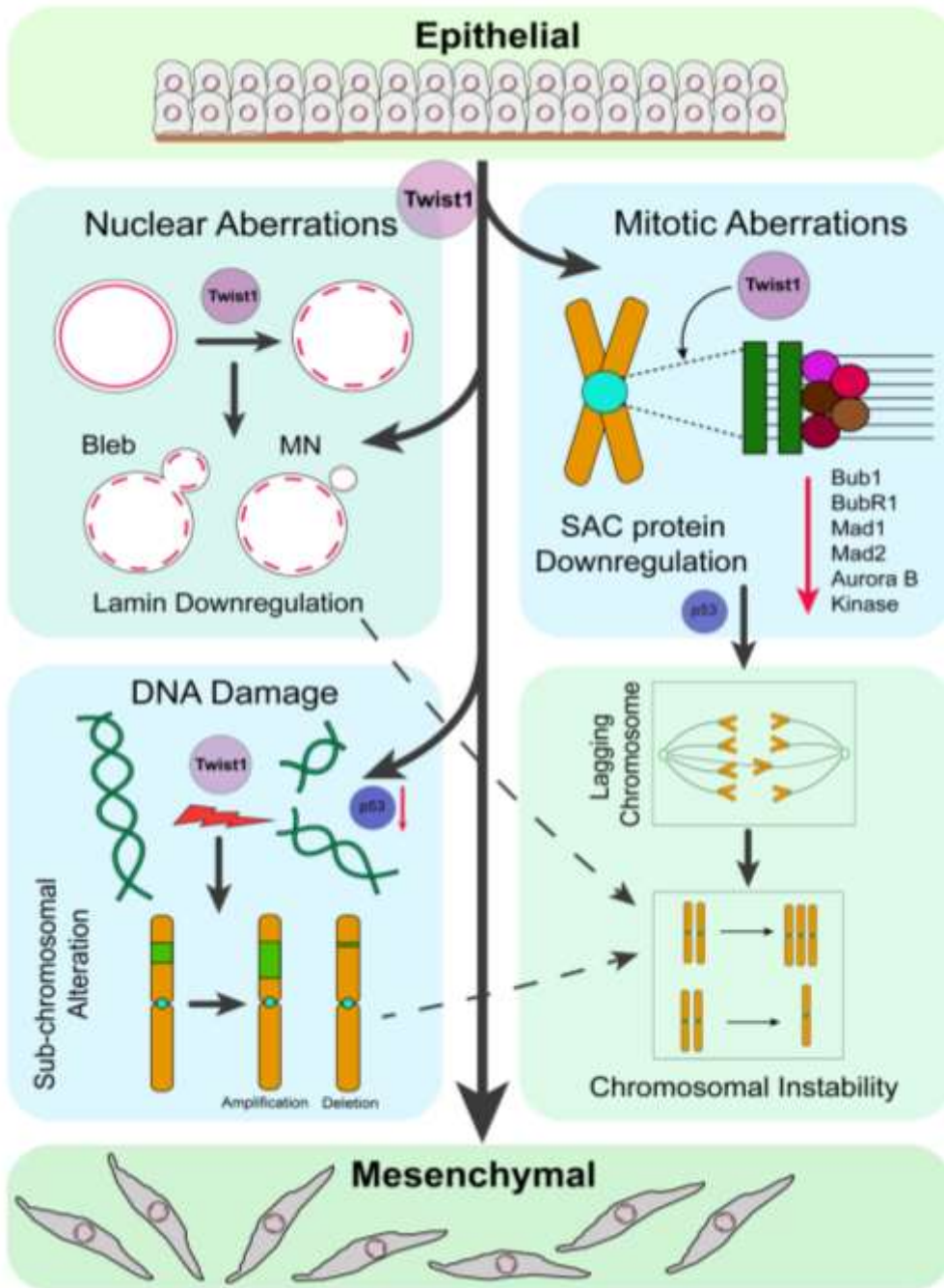
Fig.4.27: Speculative model suggesting a novel role for Twist1 overexpression in inducing CIN in colorectal cancer cells during EMT. Twist1 overexpression in colorectal cancer cell lines (i) induces EMT (ii) downregulates nuclear envelope proteins Lamin A/C, B1 and B2, associated with nuclear aberrations and CIN (iii) induces DNA double strand breaks that result in enhanced sub-chromosomal alterations and CIN, potentially via p53 and (iv) downregulates cell cycle regulators Bub1, BubR1, Mad1, Mad2 and Aurora B Kinase leading to mitotic defects, that contribute to enhanced CIN. In summary, Twist1 overexpression enhances CIN in the context of EMT, which further contributes to cellular heterogeneity and cancer progression.

Chapter 5 Discussion and Future Perspectives

Cancer metastasis is the leading cause of cancer-associated deaths (Mehlen and Puisieux, 2006). Epithelial to Mesenchymal Transition (EMT) contributes to metastasis and is associated with stemness and resistance to therapy (Gooding and Schiemann, 2020). Here, we demonstrate that overexpression of Twist1, a known EMT inducer, exhibits chromosomal instability (CIN) in colorectal cancer cells **(Fig.4.9)**. CIN is a hallmark of cancers (Vargas-Rondón et al., 2017). Chromosomal imbalances impact expression levels of most genes on that chromosome and across the genome, which impacts cell physiology (Dürrbaum and Storchová, 2016). Aneuploidy may improve fitness in specific microenvironments during cancer progression (Simonetti et al., 2019). Twist1 induced CIN further contributes to cellular heterogeneity that provides a survival advantage to migrating cells.

The spindle assembly checkpoint (SAC) complex regulates cell cycle progression and ensures chromosome alignment at the metaphase plate (Pachis and Kops, 2018). If the SAC proteins are compromised, cells undergo premature anaphase resulting in chromosomal missegregation contributing to aneuploidy (Simonetti et al., 2019). Twist1 overexpression downregulates SAC proteins further enhancing CIN in colorectal cancer cells **(Fig.4.9).** Introducing extra copies of human chromosomes 3, 5, 8, 13, 18, or 21 in the colorectal cancer cell line HCT116, showed increased migration. Furthermore, trisomy of Chr. 5 specifically induced partial EMT in these cells (Vasudevan et al., 2019). Along with whole chromosome losses and gains, Twist1 also shows copy numbers variations **(Fig.4.10).** Amplification and deletion of specific genes can also cause phenotypic switches. For example, loss of regions of chromosomes like Chr.16q, which harbours *E-cadherin* gene, can induce mesenchymal phenotype whereas loss of Chr.10p on which *Zeb1* gene is present induces epithelial phenotype (Gao et al., 2016).

We analyzed TCGA gene expression datasets and somatic mutations of 30 different cancer types for correlation between Twist1 expression and expression of genes involved in EMT, CIN and DNA DSB repair. We observed that Twist1 shows a positive correlation with EMT across cancers while the correlation of CIN and DNA DSBs genes are cancer subtype-specific **(Fig.4.26)**. We compared this analysis with another EMT inducer used in this study, - TGF-β1 **(Fig.5.1)**. Most cancer subtypes showed a positive correlation with EMT even for TGF-β overexpression. TGF-β overexpression showed a strong correlation with EMT in pan kidney carcinomas (KIPAN), bladder urothelial carcinoma (BLCA) and stomach adenocarcinoma

(STAD). Similarly, Twist showed a correlation in esophageal carcinoma (ESCA), bladder urothelial carcinoma (BLCA), stomach adenocarcinoma (STAD) among others.

Figure 5.1
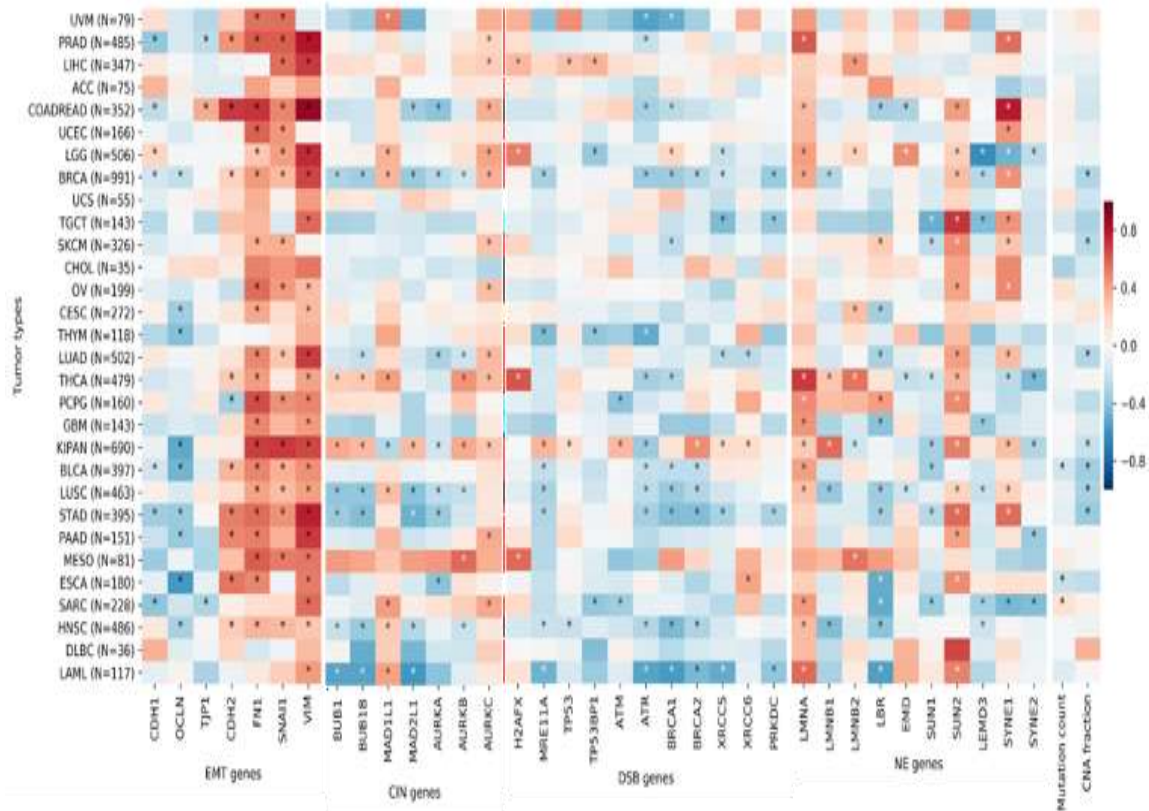
Fig.5.1: TCGA analyses for correlation between levels of TWIST1, EMT and CIN across human cancers. Heatmap representing the correlation coefficient (and its significance) between TGF-β expression and (i) EMT, (ii) CIN, (iii) DSB gene expression, (iv) mutation count and (v) copy number alteration (CNA) fraction for all 30 distinct cancers, computed using an iteratively reweighted least-squares approach. Color coding indicates the correlation coefficient, ranging from −1 to +1, where −1 being strong negative correlation (dark blue), 0 for no correlation (white) and +1 strong positive correlation (dark red). Significant correlations (q < 0.01) are marked with an asterisk (∗). Cancer subtype abbreviations are listed in **Fig 4.26.** Data analyses by Elangoli Ebrahimkutty Faseela and Radhakrishnan Sabarinathan (NCBS).

A significant and positive correlation was detected between CIN and certain tumor types which includes pan kidney cancers (KIPAN), lower-grade glioma (LGG) and lung adenocarcinoma (LUAD); whereas, stomach adenocarcinoma (STAD) showed a significant negative correlation between Twist1 and CIN genes. Interestingly, TGF-β showed a correlation of CIN in a wider range of cancer subtypes. It showed a correlation with breast invasive carcinoma (BRCA), pan kidney carcinomas (KIPAN), and lung squamous cell carcinoma (LUSC). We also examined if

there exists a correlation between TGF-β or Twist1 expression and tumor mutation burden and copy number alterations (CNA). The cancer types LGG (Brain), KIPAN (Pan Kidney) and PRAD (prostate) showed a significant positive correlation with Twist1 expression for somatic mutation and for CNA events in LGG **(Fig.4.26)**. Interestingly, TGF-β showed a positive correlation for somatic mutations only in SARC (sarcoma). Taken together, these results suggest that Twist1 and TGF-β expression correlates with EMT in various cancers while correlation with expression levels of CIN genes, DSB genes, somatic mutations and copy number alterations is cancer subtype specific and specific to the EMT inducer.

Nuclear and mitotic aberrations are precursors to CIN (Hatch et al., 2013). Pathologists have observed nuclei of cancer cells for change in shape and size as diagnostic and prognostic markers (Chow et al., 2012). Nuclear envelope proteins regulate nuclear shape and size (Webster et al., 2009). Research into alterations of nuclear envelope proteins has established their role in a variety of diseases collectively known as 'envelopathies'. Envelopathies range from premature ageing to muscle dystrophies. Although nuclear envelope proteins are ubiquitously expressed in all cell types, the disease phenotypes are tissue-specific in nature (Chi et al., 2009). Therefore, this highlights the importance of cellular context - a preeminent consideration in cancers. Deregulation of nuclear envelope proteins is also seen across cancer subtypes (Chow et al., 2012). Critical roles for nuclear envelope proteins are emerging in cancer-associated phenotypes such as cell migration and metastasis. Deregulation of levels of nuclear envelope proteins impacts the rigidity of the nucleus. It affects the deformability of the nucleus which modulates its ability to squeeze through constricted spaces for metastasis (Friedl et al., 2011). Overexpression of Lamin A increases the migration of a colon cancer cell line SW480 and upregulates the actin-binding protein - T-plastin and the down-regulation of E-cadherin (Willis et al., 2008). Our analyses of TCGA data also revealed deregulation of lamin levels across cancers **(Fig.3.17)**.

Lamin B2 depletion impacts chromosomal stability in colorectal cancer cells. Twist1 overexpression also induced aneuploidy and downregulated nuclear lamins **(Fig.4.17).** Therefore lamin downregulation may in turn contribute to Twist1-induced CIN. Furthermore, reduced levels of Lamin B1 shows micronuclei formation (Kuga et al., 2014). Interestingly, the activation of the cGAS/STING cytosolic DNA sensing pathway is also associated with micronuclei formation in cancer cells along with the activation of NFκB signaling, EMT and increased

cellular invasiveness (Bakhoum et al., 2018). In summary, the role of nuclear envelope proteins underlies the mechanistic underpinnings for a wide variety of cancer phenotypes. Nuclear envelope proteins function as a signaling platform for various cancer-inducing pathways. Nuclear envelope proteins modulate TGF-β signaling, an inducer of EMT. Nuclear envelope proteins at the nuclear periphery MAN1 and Lamin A/C sequester Smads, the effectors of TGF-β signaling (Lin et al., 2005). Mouse embryonic fibroblasts (MEF) lacking LMNA show increased proliferation upon TGF-β1 treatment (Van Berlo et al., 2005). Furthermore, cells with Nesprin2 knockdown showed altered translocation kinetics of pSmad2/3 into the nucleus (Rashmi et al., 2012). While TGFβ1 treatment shows a cytostatic role in normal cells, however, the regulation of TGFβ1-induced EMT by nuclear envelope proteins is largely unexplored. Recent work has shown the involvement of Lamin B1 in inducing EMT. Reduced levels of Lamin B1 promote cell migration, tumor growth, metastasis and EMT in lung cancer patients mediated by altering H3K27me3 occupancy by the polycomb repressive complex 2 (PRC2) (Jia et al., 2019). Furthermore, reduced Lamin B1 levels dampened TGFβ1-induced EMT in NMuMG cells as well (Pascual-Reguant et al., 2018).

In addition to regulating cell signaling, the nuclear envelope proteins are primarily involved in genome organization. Recent work has highlighted the impact of signaling pathways on 3D genome organization (D'Ippolito et al., 2018). Furthermore, the 3D topology of the genome is involved in promoter-enhancer contacts that further regulates transcription (Schoenfelder and Fraser, 2019). Architectural proteins such as lamins, CTCF, cohesins, NUPs shape the 3D organization of the genome (Gómez-Díaz and Corces, 2014). Depletion of Rad21 (a subunit of cohesin) in breast cancer cell line, induces EMT by transcriptional activation of *TGFB1* and *ITGA5*. Reduced levels of Rad21 also disrupt intrachromosomal chromatin interactions of the *TGFB1* and *ITGA5* loci (Chromatin conformation capture (3C) experiment) leading to transcriptional activation. In contrast, overexpression of Rad21 in mesenchymal cancer cells induces MET (Yun et al., 2016). Here, we examined the effect of perturbing levels of two genome organizers, Lamins and CTCF independently, on TGFβ1-induced EMT in A549 cells. While the downregulation of either of the nuclear lamins did not affect TGFβ1-induced EMT, CTCF depletion also does not impact EMT induction but altered the expression profiles of genes upregulated upon TGFβ1 treatment **(Fig.3.18, 3.21, 3.22)**. We surmise that the depletion of CTCF followed by chromosome capture assays will be essential for unravelling the mechanistic

underpinnings of the role of CTCF in EMT. In addition, TGFβ1-induced EMT increased chromatin accessibility upon EMT which was reversed upon MET in MCF10A breast cancer cells. Interestingly, regions of increased accessibility showed enrichment for AP1 and Smad binding sites and a decrease in CTCF binding. However, CTCF knockdown did not affect EMT induction in MCF10A cells (Johnson et al., 2020).

TGF-β1-induced EMT in NMuMG cells was the first comprehensive study on 3D genome organization in EMT. LaminB1 occupies the euchromatic LADs (eLADs) which are dynamic during EMT. HiC data revealed eLADs as a part of the 'A compartment' during EMT initiation, while novel eLADs are in 'B compartment' during EMT progression, suggesting the impact of EMT on genome organization (Pascual-Reguant et al., 2018).

While the role of nuclear envelope proteins in TGFβ1-induced EMT is being studied, their role in Twist1 induced EMT remains unclear. Here, we report a novel molecular cross-talk between Twist1 and nuclear lamins, which impacts chromosomal stability in colorectal cancer cells. However, the mechanistic basis remains unclear. Decreased lamin levels impact the 3D organization of the genome (Kim et al., 2019). However, how Twist1 impinges on genome organization in the context of EMT is unknown.

EMT is accompanied by the secretion of factors such as TGF-β, Wnt, HGF and EGF by the tumor microenvironment (Thiery et al., 2009). Intravital microscopy on xenografted tumor cells reveals transient secretion of TGFβ. Nonetheless, circulating tumor cells (CTCs) retain their EMT status much after they have migrated away from the primary tumor. This suggests mechanisms that not only initiate but also maintain EMT (Giampieri et al., 2009). Diverse EMT inducers converge on the activation of a subset of EMT-Transcription factors (EMT-TFS) such as Zinc finger transcription repressors - Snail1, Slug, and ZEB1/2, and basic helix-loop-helix (bHLH) proteins, Twist1/2, that directly bind to E-boxes in promoters (Peinado et al., 2007). Here, we studied two models of EMT induction- (i) TGF-β1 induced EMT in lung adenocarcinoma cell line and (ii) Twist1 induced EMT in colorectal cancer cell lines. Our studies reveal that Twist1 mediated EMT induces CIN in colorectal cancer cell lines, consistent with CIN induction in the breast cancer cell line MCF7 (Vesuna et al., 2006). While the ploidy of A549 cells remained unaltered upon TGFβ1 induced EMT, TGF-β treatment of MCF10A cells

for 6 days induces CIN, accompanied by the down-regulation of a large number of nuclear envelope proteins. Lamin B1 downregulation phenocopies the effects of TGFβ1 on ploidy (Comaills et al., 2016) in MCF10A cells. We surmise that in addition to the genetic background, the extent of EMT induction is an important determinant of effects such as CIN, that accompanies EMT.

TGF-β1 induced EMT induces expression of Snail1 and Twist1 in multiple normal epithelial and carcinoma cell lines including A549 (Tran et al., 2011). This results in the temporal co-operation of Snail1 and Twist1. Snail1 is uniquely required for the initiation of EMT and the levels of Snail1 diminished over time. At the same time, Twist1 begins to increase to maintain late EMT. Twist1 shows late EMT-associated cell proliferation arrest by inducing a low ERK:p38 activity ratio (Tran et al., 2011). This Snail1–Twist1 temporal cooperation is also observed in human breast cancer metastasis. Snail1 level in the primary tumor is a prognostic marker for metastasis whereas Twist1 level and Twist1: Snail1 expression ratio is strongly prognostic of relapse (Tran et al., 2011). Therefore, time and/or EMT stage (initial versus maintenance) is an important factor that influences EMT associated phenotypes.

Various EMT inducers elicit transcriptional changes. A core set of genes are commonly deregulated by most EMT inducers (Gröger et al., 2012). It is unclear as to how different EMT inducers impact genome organization. We surmise that either a core set or context-specific genomic alterations serve as signatures unique to each EMT inducer.

EMT is a reversible process as Mesenchymal to Epithelial Transitions (MET) (Kalluri and Weinberg, 2009). While certain transcriptional changes are potentially reversible, it is unclear if this is accompanied by associated changes in chromatin organization. Furthermore, while Twist1 induces aneuploidy in colorectal cancer cells, the impact of such an aneuploid state on genome organization is unclear. A study on a cell line established from healthy colon mucosa with a normal karyotype (46, XY) and its isogenic cell line with an extra copy of chromosome 7 (47, XY, +7) demonstrated changes in chromatin organization changes with trisomy of Chr.7. Furthermore, A/B compartmentalization and TAD boundaries were altered as compared to their normal counterpart. In addition, Chr.4 showed large-scale changes in chromatin organization, while Chr.14 showed a massive switch from the A to B compartment, along with changes in

gene expression levels from genes on Chr.14. However, positions of chromosome 7 territory were unaltered in these cells as examined by 3D-FISH. RNA seq and protein profiling of these cells showed enrichment of HGF/MET-axis involved in malignant transformation (Braun et al., 2019). In summary, chromosomal aneuploidies impact nuclear organization and function. This study highlights the impact of aneuploidy on genome organization in cancers.

Adding to the complexity of EMT is the novel understanding that EMT is not a binary process but a continuum composed of E/M hybrids. Of note, cell lines used in this study are at different stages within the EMT spectrum. In addition, heterogeneity of intermediate states further impacts cell migration, invasion, metastasis and resistance to therapy (Jolly et al., 2015). It is noteworthy that no detailed study has been performed to examine the nuclear landscape in these hybrid cellular states. An immuno-FISH analysis for chromosome territory organization with staining for EMT markers will form the basis for characterizing the contribution of cellular and genomic heterogeneity associated with these hybrid cell types. Similarly, immuno-FISH for gene loci associated with EMT along with RNA-FISH will establish the nuclear structure-function relationships in such a dynamic process. Single cell Hi-C combined with single cell RNA-seq will provide a comprehensive understanding of epithelial-mesenchymal plasticity.

In summary, while our research has contributed to the understanding of nuclear structure-function relationships in EMT, what we know is only the tip of the iceberg and much remains to be explored. Recent research efforts are geared towards developing new tools and technologies to map genome structure in space and time (4D). The goal of such studies is to gain deeper mechanistic insights into how the nucleus is organized in health and disease. The dynamic model of EMT provides a suitable paradigm for the comprehensive mapping of chromatin organization that contribute to the development of novel diagnostic and prognostic markers for therapeutic intervention.

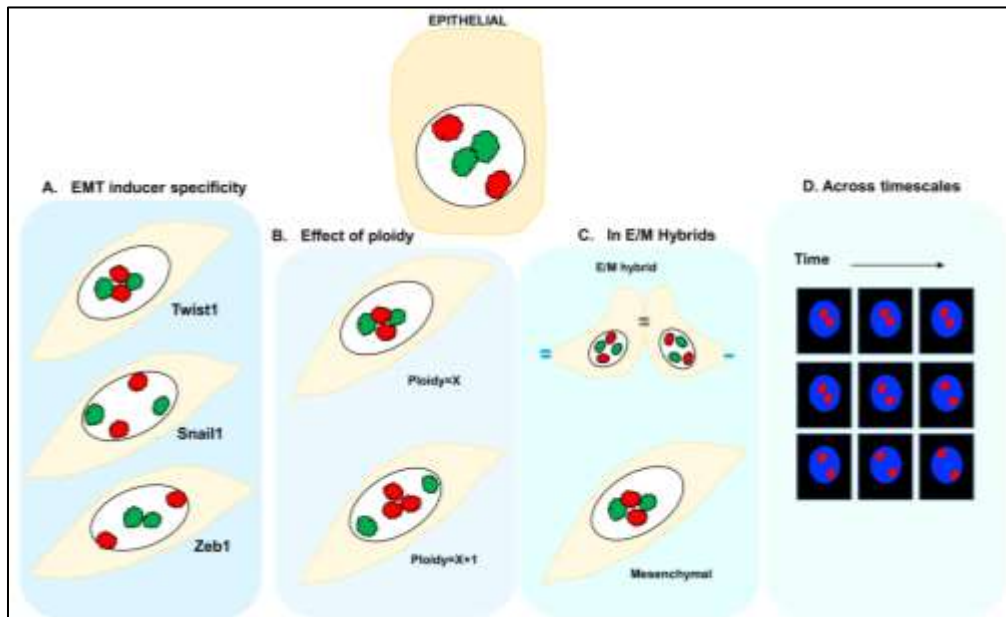**Figure 5.2**

Fig.5.2: Future directions in understanding nuclear structure-function relationships in EMT (A) Specificity of EMT inducer on 3D genome organization **(B)** Aneuploidy and genome organization **(C)** Genome organization in EMT hybrids **(D)** 3D genome organization across timescales. Chromosome Territories in red and green are used as representation of 3D genome organization.

References

1. Aaronson, R.P., and Blobel, G. (1975). Isolation of nuclear pore complexes in association with a lamina. Proc. Natl. Acad. Sci. USA *72*, 1007–1011.

2. Agapova, L.S., Ilyinskaya, G.V., Turovets, N.A., Ivanov, A.V., Chumakov, P.M., and Kopnin, B.P. (1996). Chromosome changes caused by alterations of p53 expression. Mutat. Res. *354*, 129–138.

3. Aghdassi, A., Sendler, M., Guenther, A., Mayerle, J., Behn, C.-O., Heidecke, C.-D., Friess, H., Büchler, M., Evert, M., Lerch, M.M., et al. (2012). Recruitment of histone deacetylases HDAC1 and HDAC2 by the transcriptional repressor ZEB1 downregulates E-cadherin expression in pancreatic cancer. Gut *61*, 439–448.

4. Ahmed, D., Eide, P.W., Eilertsen, I.A., Danielsen, S.A., Eknæs, M., Hektoen, M., Lind, G.E., and Lothe, R.A. (2013). Epigenetic and genetic features of 24 colon cancer cell lines. Oncogenesis *2*, e71.

5. Baghdassarian, N., and Ffrench, M. (1996). Cyclin-dependent kinase inhibitors (CKIs) and hematological malignancies. Hematol. Cell Ther. *38*, 313–323.

6. Bakhoum, S.F., Kabeche, L., Murnane, J.P., Zaki, B.I., and Compton, D.A. (2014). DNA-damage response during mitosis induces whole-chromosome missegregation. Cancer Discov. *4*, 1281–1289.

7. Bakhoum, S.F., Ngo, B., Laughney, A.M., Cavallo, J.-A., Murphy, C.J., Ly, P., Shah, P., Sriram, R.K., Watkins, T.B.K., Taunk, N.K., et al. (2018). Chromosomal instability drives metastasis through a cytosolic DNA response. Nature *553*, 467–472.

8. Ballabio, E., Cantarella, C.D., Federico, C., Di Mare, P., Hall, G., Harbott, J., Hughes, J., Saccone, S., and Tosi, S. (2009). Ectopic expression of the HLXB9 gene is associated with an altered nuclear position in t(7;12) leukaemias. Leukemia *23*, 1179–1182.

9. Belt, E.J.T., Fijneman, R.J.A., van den Berg, E.G., Bril, H., Delis-van Diemen, P.M., Tijssen, M., van Essen, H.F., de Lange-de Klerk, E.S.M., Beliën, J.A.M., Stockmann, H.B.A.C., et al. (2011). Loss of lamin A/C expression in stage II and III colon cancer is associated with disease recurrence. Eur. J. Cancer *47*, 1837–1845.

10. Bendinelli, P., Maroni, P., Matteucci, E., and Desiderio, M.A. (2015). HGF and TGFβ1 differently influenced Wwox regulatory function on Twist program for mesenchymal-epithelial transition in bone metastatic versus parental breast carcinoma cells. Mol. Cancer *14*, 112.

11. Van Berlo, J.H., Voncken, J.W., Kubben, N., Broers, J.L.V., Duisters, R., van Leeuwen, R.E.W., Crijns, H.J.G.M., Ramaekers, F.C.S., Hutchison, C.J., and Pinto, Y.M. (2005). A-type lamins are essential for TGF-beta1 induced PP2A to dephosphorylate transcription factors. Hum. Mol. Genet. *14*, 2839–2849.

12. Boudaoud, A., Burian, A., Borowska-Wykręt, D., Uyttewaal, M., Wrzalik, R., Kwiatkowska, D., and Hamant, O. (2014). FibrilTool, an ImageJ plug-in to quantify fibrillar structures in raw microscopy images. Nat. Protoc. *9*, 457–463.

13. Boveri, T. (1909). Die Blastomerenkerne von Ascaris megalocephala. Archiv für Zellforschung.

14. Braccioli, L., and de Wit, E. (2019). CTCF: a Swiss-army knife for genome organization and transcription regulation. Essays Biochem *63*, 157–165.

15. Brackley, C.A., Johnson, J., Michieletto, D., Morozov, A.N., Nicodemi, M., Cook, P.R., and Marenduzzo, D. (2018). Extrusion without a motor: a new take on the loop extrusion model of genome organization. Nucleus *9*, 95–103.

16. Braun, R., Ronquist, S., Wangsa, D., Chen, H., Anthuber, L., Gemoll, T., Wangsa, D., Koparde, V., Hunn, C., Habermann, J.K., et al. (2019). Single Chromosome Aneuploidy Induces Genome-Wide Perturbation of Nuclear Organization and Gene Expression. Neoplasia *21*, 401–412.

17. Briand, N., and Collas, P. (2020). Lamina-associated domains: peripheral matters and internal affairs. Genome Biol. *21*, 85.

18. Buchwalter, A., Kaneshiro, J.M., and Hetzer, M.W. (2019). Coaching from the sidelines: the nuclear periphery in genome regulation. Nat. Rev. Genet. *20*, 39–50.

19. Burds, A.A., Lutum, A.S., and Sorger, P.K. (2005). Generating chromosome instability through the simultaneous deletion of Mad2 and p53. Proc. Natl. Acad. Sci. USA *102*, 11296–11301.

20. Burk, U., Schubert, J., Wellner, U., Schmalhofer, O., Vincan, E., Spaderna, S., and Brabletz, T. (2008). A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. EMBO Rep. *9*, 582–589.

21. Burke, D.J., and Stukenberg, P.T. (2008). Linking kinetochore-microtubule binding to the spindle checkpoint. Dev. Cell *14*, 474–479.

22. Camps, J., Armengol, G., del Rey, J., Lozano, J.J., Vauhkonen, H., Prat, E., Egozcue, J., Sumoy, L., Knuutila, S., and Miró, R. (2006). Genome-wide differences between microsatellite stable and unstable colorectal tumors. Carcinogenesis *27*, 419–428.

23. Cancer Genome Atlas Network (2012). Comprehensive molecular characterization of human colon and rectal cancer. Nature *487*, 330–337.

24. Cancer Genome Atlas Research Network (2014). Comprehensive molecular profiling of lung adenocarcinoma. Nature *511*, 543–550.

25. Cannan, W.J., and Pederson, D.S. (2016). Mechanisms and Consequences of Double-Strand DNA Break Formation in Chromatin. J. Cell Physiol. *231*, 3–14.

26. Capell, B.C., and Collins, F.S. (2006). Human laminopathies: nuclei gone genetically awry. Nat. Rev. Genet. *7*, 940–952.

27. Capo-chichi, C.D., Cai, K.Q., Simpkins, F., Ganjei-Azar, P., Godwin, A.K., and Xu, X.-X. (2011). Nuclear envelope structural defects cause chromosomal numerical instability and aneuploidy in ovarian cancer. BMC Med. *9*, 28.

28. Cardenas, H., Vieth, E., Lee, J., Segar, M., Liu, Y., Nephew, K.P., and Matei, D. (2014). TGF-β induces global changes in DNA methylation during the epithelial-to-mesenchymal transition in ovarian cancer cells. Epigenetics *9*, 1461–1472.

29. Carmona, F.J., Davalos, V., Vidal, E., Gomez, A., Heyn, H., Hashimoto, Y., Vizoso, M., Martinez-Cardus, A., Sayols, S., Ferreira, H.J., et al. (2014). A comprehensive DNA methylation profile of epithelial-to-mesenchymal transition. Cancer Res. *74*, 5608–5619.

30. Cassar, L., Nicholls, C., Pinto, A.R., Chen, R., Wang, L., Li, H., and Liu, J.-P. (2017). TGF-beta receptor mediated telomerase inhibition, telomere shortening and breast cancer cell senescence. Protein Cell *8*, 39–54.

31. Chaffer, C.L., Marjanovic, N.D., Lee, T., Bell, G., Kleer, C.G., Reinhardt, F., D'Alessio, A.C., Young, R.A., and Weinberg, R.A. (2013). Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. Cell *154*, 61–74.

32. Chandra, S.H.V., Wacker, I., Appelt, U.K., Behrens, J., and Schneikert, J. (2012). A common role for various human truncated adenomatous polyposis coli isoforms in the control of beta-catenin activity and cell proliferation. PLoS One *7*, e34479.

33. Chang, A.T., Liu, Y., Ayyanathan, K., Benner, C., Jiang, Y., Prokop, J.W., Paz, H., Wang, D., Li, H.-R., Fu, X.-D., et al. (2015). An evolutionarily conserved DNA architecture determines target specificity of the TWIST family bHLH transcription factors. Genes Dev. *29*, 603–616.

34. Chen, Z.F., and Behringer, R.R. (1995). twist is required in head mesenchyme for cranial neural tube morphogenesis. Genes Dev. *9*, 686–699.

35. Chen, D.-H., Yu, J.-W., and Jiang, B.-J. (2015). Contactin 1: A potential therapeutic target and biomarker in gastric cancer. World J. Gastroenterol. *21*, 9707–9716.

36. Chen, M.-W., Hua, K.-T., Kao, H.-J., Chi, C.-C., Wei, L.-H., Johansson, G., Shiah, S.-G., Chen, P.-S., Jeng, Y.-M., Cheng, T.-Y., et al. (2010). H3K9 histone methyltransferase G9a promotes lung cancer invasion and metastasis by silencing the cell adhesion molecule Ep-CAM. Cancer Res. *70*, 7830–7840.

37. Chen, Y., Zhang, Y., Wang, Y., Zhang, L., Brinkman, E.K., Adam, S.A., Goldman, R., van Steensel, B., Ma, J., and Belmont, A.S. (2018). Mapping 3D genome organization relative to nuclear compartments using TSA-Seq as a cytological ruler. J. Cell Biol. *217*, 4025–4048.

38. Chi, Y.-H., Chen, Z.-J., and Jeang, K.-T. (2009). The nuclear envelopathies and human diseases. J Biomed Sci *16*, 96.

39. Chow, K.-H., Factor, R.E., and Ullman, K.S. (2012). The nuclear envelope environment and its cancer connections. Nat. Rev. Cancer *12*, 196–209.

40. Clements, L., Manilal, S., Love, D.R., and Morris, G.E. (2000). Direct interaction between emerin and lamin A. Biochem. Biophys. Res. Commun. *267*, 709–714.

41. Cochrane, D.R., Cittelly, D.M., Howe, E.N., Spoelstra, N.S., McKinsey, E.L., LaPara, K., Elias, A., Yee, D., and Richer, J.K. (2010). MicroRNAs link estrogen receptor alpha status and Dicer levels in breast cancer. Horm. Cancer *1*, 306–319.

42. Comaills, V., Kabeche, L., Morris, R., Buisson, R., Yu, M., Madden, M.W., LiCausi, J.A., Boukhali, M., Tajima, K., Pan, S., et al. (2016). Genomic Instability Is Induced by

Persistent Proliferation of Cells Undergoing Epithelial-to-Mesenchymal Transition. Cell Rep. *17*, 2632–2647.

43. Cremer, T., and Cremer, M. (2010). Chromosome territories. Cold Spring Harb. Perspect. Biol. *2*, a003889.

44. Cremer, M., von Hase, J., Volm, T., Brero, A., Kreth, G., Walter, J., Fischer, C., Solovei, I., Cremer, C., and Cremer, T. (2001). Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. Chromosome Res. *9*, 541–567.

45. Cremer, M., Küpper, K., Wagler, B., Wizelman, L., von Hase, J., Weiland, Y., Kreja, L., Diebold, J., Speicher, M.R., and Cremer, T. (2003). Inheritance of gene density-related higher order chromatin arrangements in normal and tumor cell nuclei. J. Cell Biol. *162*, 809–820.

46. Cremer, M., Grasser, F., Lanctôt, C., Müller, S., Neusser, M., Zinner, R., Solovei, I., and Cremer, T. (2008). Multicolor 3D fluorescence in situ hybridization for imaging interphase chromosomes. Methods Mol. Biol. *463*, 205–239.

47. Crisp, M., Liu, Q., Roux, K., Rattner, J.B., Shanahan, C., Burke, B., Stahl, P.D., and Hodzic, D. (2006). Coupling of the nucleus and cytoplasm: role of the LINC complex. J. Cell Biol. *172*, 41–53.

48. Di Croce, L., and Helin, K. (2013). Transcriptional regulation by Polycomb group proteins. Nat. Struct. Mol. Biol. *20*, 1147–1155.

49. Croft, J.A., Bridger, J.M., Boyle, S., Perry, P., Teague, P., and Bickmore, W.A. (1999). Differences in the localization and morphology of chromosomes in the human nucleus. J. Cell Biol. *145*, 1119–1131.

50. D'Ippolito, A.M., McDowell, I.C., Barrera, A., Hong, L.K., Leichter, S.M., Bartelt, L.C., Vockley, C.M., Majoros, W.H., Safi, A., Song, L., et al. (2018). Pre-established Chromatin Interactions Mediate the Genomic Response to Glucocorticoids. Cell Syst. *7*, 146–160.e7.

51. Dalton, W.B., Yu, B., and Yang, V.W. (2010). p53 suppresses structural chromosome instability after mitotic arrest in human cells. Oncogene *29*, 1929–1940.

52. Darwanto, A., Kitazawa, R., Maeda, S., and Kitazawa, S. (2003). MeCP2 and promoter methylation cooperatively regulate E-cadherin gene expression in colorectal carcinoma. Cancer Sci. *94*, 442–447.

53. Davidson, I.F., Bauer, B., Goetz, D., Tang, W., Wutz, G., and Peters, J.-M. (2019). DNA loop extrusion by human cohesin. Science *366*, 1338–1345.

54. Dekker, J., Marti-Renom, M.A., and Mirny, L.A. (2013). Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat. Rev. Genet. *14*, 390–403.

55. Demmerle, J., Koch, A.J., and Holaska, J.M. (2012). The nuclear envelope protein emerin binds directly to histone deacetylase 3 (HDAC3) and activates HDAC3 activity. J. Biol. Chem. *287*, 22080–22088.

56. Dhasarathy, A., Phadke, D., Mav, D., Shah, R.R., and Wade, P.A. (2011). The transcription factors Snail and Slug activate the transforming growth factor-beta signaling pathway in breast cancer. PLoS One *6*, e26514.

57. Díaz-López, A., Moreno-Bueno, G., and Cano, A. (2014). Role of microRNA in epithelial to mesenchymal transition and metastasis and clinical perspectives. Cancer Manag Res *6*, 205–216.

58. Diogo, V., Teixeira, J., Silva, P.M.A., and Bousbaa, H. (2017). Spindle assembly checkpoint as a potential target in colorectal cancer: current status and future perspectives. Clin Colorectal Cancer *16*, 1–8.

59. Ditchfield, C., Johnson, V.L., Tighe, A., Ellston, R., Haworth, C., Johnson, T., Mortlock, A., Keen, N., and Taylor, S.S. (2003). Aurora B couples chromosome alignment with anaphase by targeting BubR1, Mad2, and Cenp-E to kinetochores. J. Cell Biol. *161*, 267–280.

60. Dong, C., Wu, Y., Yao, J., Wang, Y., Yu, Y., Rychahou, P.G., Evers, B.M., and Zhou, B.P. (2012). G9a interacts with Snail and is critical for Snail-mediated E-cadherin repression in human breast cancer. J. Clin. Invest. *122*, 1469–1486.

61. Dong, C., Wu, Y., Wang, Y., Wang, C., Kang, T., Rychahou, P.G., Chi, Y.I., Evers, B.M., and Zhou, B.P. (2013). Interaction with Suv39H1 is critical for Snail-mediated E-cadherin repression in breast cancer. Oncogene *32*, 1351–1362.

62. Dong, P., Kaneuchi, M., Watari, H., Hamada, J., Sudo, S., Ju, J., and Sakuragi, N. (2011). MicroRNA-194 inhibits epithelial to mesenchymal transition of endometrial cancer cells by targeting oncogene BMI-1. Mol. Cancer *10*, 99.

63. Dorner, D., Gotzmann, J., and Foisner, R. (2007). Nucleoplasmic lamins and their interaction partners, LAP2alpha, Rb, and BAF, in transcriptional regulation. FEBS J. *274*, 1362–1373.

64. Douglas, E.J., Fiegler, H., Rowan, A., Halford, S., Bicknell, D.C., Bodmer, W., Tomlinson, I.P.M., and Carter, N.P. (2004). Array comparative genomic hybridization analysis of colorectal cancer cell lines and primary carcinomas. Cancer Res. *64*, 4817–4825.

65. Du, L., Yamamoto, S., Burnette, B.L., Huang, D., Gao, K., Jamshidi, N., and Kuo, M.D. (2016). Transcriptome profiling reveals novel gene expression signatures and regulating transcription factors of TGFβ-induced epithelial-to-mesenchymal transition. Cancer Med *5*, 1962–1972.

66. Dürrbaum, M., and Storchová, Z. (2016). Effects of aneuploidy on gene expression: implications for cancer. FEBS J. *283*, 791–802.

67. El Ghouzzi, V., Lajeunie, E., Le Merrer, M., Cormier-Daire, V., Renier, D., Munnich, A., and Bonaventure, J. (1999). Mutations within or upstream of the basic helix-loop-helix domain of the TWIST gene are specific to Saethre-Chotzen syndrome. Eur. J. Hum. Genet. *7*, 27–33.

68. Elowe, S. (2011). Bub1 and BubR1: at the interface between chromosome attachment and the spindle checkpoint. Mol. Cell. Biol. *31*, 3085–3093.

69. Fang, Y., Liu, T., Wang, X., Yang, Y.M., Deng, H., Kunicki, J., Traganos, F., Darzynkiewicz, Z., Lu, L., and Dai, W. (2006). BubR1 is involved in regulation of DNA damage responses. Oncogene *25*, 3598–3605.

70. Fedorova, E., and Zink, D. (2008). Nuclear architecture and gene regulation. Biochim. Biophys. Acta *1783*, 2174–2184.

71. Feng, X.-H., and Derynck, R. (2005). Specificity and versatility in tgf-β signaling through Smads. Annu. Rev. Cell Dev. Biol. *21*, 659–693.

72. Foijer, F., Albacker, L.A., Bakker, B., Spierings, D.C., Yue, Y., Xie, S.Z., Davis, S., Lutum-Jehle, A., Takemoto, D., Hare, B., et al. (2017). Deletion of the MAD2L1 spindle assembly checkpoint gene is tolerated in mouse models of acute T-cell lymphoma and hepatocellular carcinoma. Elife *6*.

73. Foisner, R., and Gerace, L. (1993). Integral membrane proteins of the nuclear envelope interact with lamins and chromosomes, and binding is modulated by mitotic phosphorylation. Cell *73*, 1267–1279.

74. Förthmann, B., van Bergeijk, J., Lee, Y.-W., Lübben, V., Schill, Y., Brinkmann, H., Ratzka, A., Stachowiak, M.K., Hebert, M., Grothe, C., et al. (2013). Regulation of neuronal differentiation by proteins associated with nuclear bodies. PLoS One *8*, e82871.

75. Friedl, P., and Alexander, S. (2011). Cancer invasion and the microenvironment: plasticity and reciprocity. Cell *147*, 992–1009.

76. Friedl, P., Wolf, K., and Lammerding, J. (2011). Nuclear mechanics during cell migration. Curr. Opin. Cell Biol. *23*, 55–64.

77. Fukagawa, A., Ishii, H., Miyazawa, K., and Saitoh, M. (2015). δEF1 associates with DNMT1 and maintains DNA methylation of the E-cadherin promoter in breast cancer cells. Cancer Med *4*, 125–135.

78. Fulka, H., and Aoki, F. (2016). Nucleolus precursor bodies and ribosome biogenesis in early mammalian embryos: old theories and new discoveries. Biol. Reprod. *94*, 143.

79. Gao, C., Su, Y., Koeman, J., Haak, E., Dykema, K., Essenberg, C., Hudson, E., Petillo, D., Khoo, S.K., and Vande Woude, G.F. (2016). Chromosome instability drives phenotypic switching to metastasis. Proc. Natl. Acad. Sci. USA *113*, 14793–14798.

80. Gebeshuber, C.A., Zatloukal, K., and Martinez, J. (2009). miR-29a suppresses tristetraprolin, which is a regulator of epithelial polarity and metastasis. EMBO Rep. *10*, 400–405.

81. Gerace, L., and Huber, M.D. (2012). Nuclear lamina at the crossroads of the cytoplasm and nucleus. J. Struct. Biol. *177*, 24–31.

82. Gerlich, D., Beaudouin, J., Kalbfuss, B., Daigle, N., Eils, R., and Ellenberg, J. (2003). Global chromosome positions are transmitted through mitosis in mammalian cells. Cell *112*, 751–764.

83. Gesson, K., Vidak, S., and Foisner, R. (2014). Lamina-associated polypeptide (LAP)2α and nucleoplasmic lamins in adult stem cell regulation and disease. Semin. Cell Dev. Biol. *29*, 116–124.

84. Gesson, K., Rescheneder, P., Skoruppa, M.P., von Haeseler, A., Dechat, T., and Foisner, R. (2016). A-type lamins bind both hetero- and euchromatin, the latter being regulated by lamina-associated polypeptide 2 alpha. Genome Res. *26*, 462–473.

85. Giampieri, S., Manning, C., Hooper, S., Jones, L., Hill, C.S., and Sahai, E. (2009). Localized and reversible TGFbeta signalling switches breast cancer cells from cohesive to single cell motility. Nat. Cell Biol. *11*, 1287–1296.

86. Gómez-Díaz, E., and Corces, V.G. (2014). Architectural proteins: regulators of 3D genome organization in cell fate. Trends Cell Biol. *24*, 703–711.

87. Gonzalez, D.M., and Medici, D. (2014). Signaling mechanisms of the epithelial-mesenchymal transition. Sci. Signal. *7*, re8.

88. González-Aguilera, C., Ikegami, K., Ayuso, C., de Luis, A., Íñiguez, M., Cabello, J., Lieb, J.D., and Askjaer, P. (2014). Genome-wide analysis links emerin to neuromuscular junction activity in Caenorhabditis elegans. Genome Biol. *15*, R21.

89. Gooding, A.J., and Schiemann, W.P. (2020). Epithelial-Mesenchymal Transition Programs and Cancer Stem Cell Phenotypes: Mediators of Breast Cancer Therapy Resistance. Mol. Cancer Res. *18*, 1257–1270.

90. Gotzmann, J., and Foisner, R.P. (2013). Lamins and Emerin in Muscular Dystrophy: The Nuclear Envelope Connection - Madame Curie Bioscience Database - NCBI Bookshelf.

91. Greaves, M. (2013). Cancer stem cells as "units of selection". Evol. Appl. *6*, 102–108.

92. Greenburg, G., and Hay, E.D. (1982). Epithelia suspended in collagen gels can lose polarity and express characteristics of migrating mesenchymal cells. J. Cell Biol. *95*, 333–339.

93. Gregory, P.A., Bert, A.G., Paterson, E.L., Barry, S.C., Tsykin, A., Farshid, G., Vadas, M.A., Khew-Goodall, Y., and Goodall, G.J. (2008). The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. Nat. Cell Biol. *10*, 593–601.

94. Gregory, P.A., Bracken, C.P., Smith, E., Bert, A.G., Wright, J.A., Roslan, S., Morris, M., Wyatt, L., Farshid, G., Lim, Y.-Y., et al. (2011). An autocrine TGF-beta/ZEB/miR-200 signaling network regulates establishment and maintenance of epithelial-mesenchymal transition. Mol. Biol. Cell *22*, 1686–1698.

95. Gröger, C.J., Grubinger, M., Waldhör, T., Vierlinger, K., and Mikulits, W. (2012). Meta-analysis of gene expression signatures defining the epithelial to mesenchymal transition during cancer progression. PLoS One *7*, e51136.

96. Grover, S., and Mujib, B.R.A. (2017). Scoring criteria for the evaluation of micronuclei in oral exfoliated cells. South Asian J Cancer *6*, 89.

97. Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W., et al. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature *453*, 948–951.

98. Guerrero-Martínez, J.A., Ceballos-Chávez, M., Koehler, F., Peiró, S., and Reyes, J.C. (2020). TGFβ promotes widespread enhancer chromatin opening and operates on genomic regulatory domains. Nat. Commun. *11*, 6196.

99. Haensel, D., and Dai, X. (2018). Epithelial-to-mesenchymal transition in cutaneous wound healing: Where we are and where we are heading. Dev. Dyn. *247*, 473–480.

100. Hahn, S.A., Schutte, M., Hoque, A.T., Moskaluk, C.A., da Costa, L.T., Rozenblum, E., Weinstein, C.L., Fischer, A., Yeo, C.J., Hruban, R.H., et al. (1996). DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. Science *271*, 350–353.

101. Hamamori, Y., Sartorelli, V., Ogryzko, V., Puri, P.L., Wu, H.Y., Wang, J.Y., Nakatani, Y., and Kedes, L. (1999). Regulation of histone acetyltransferases p300 and PCAF by the bHLH protein twist and adenoviral oncoprotein E1A. Cell *96*, 405–413.

102. Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674.

103. Hanel, W., and Moll, U.M. (2012). Links between mutant p53 and genomic instability. J. Cell Biochem. *113*, 433–439.

104. Hao, S., He, W., Li, Y., Ding, H., Hou, Y., Nie, J., Hou, F.F., Kahn, M., and Liu, Y. (2011). Targeted inhibition of β-catenin/CBP signaling ameliorates renal interstitial fibrosis. J. Am. Soc. Nephrol. *22*, 1642–1653.

105. Hatch, E.M., Fischer, A.H., Deerinck, T.J., and Hetzer, M.W. (2013). Catastrophic nuclear envelope collapse in cancer cell micronuclei. Cell *154*, 47–60.

106. Hay, E.D. (1990). Role of cell-matrix contacts in cell migration and epithelial-mesenchymal transformation. Cell Differ. Dev. *32*, 367–375.

107. Heldin, C.-H., and Moustakas, A. (2016). Signaling Receptors for TGF-β Family Members. Cold Spring Harb. Perspect. Biol. *8*.

108. Hetzer, M.W., Walther, T.C., and Mattaj, I.W. (2005). Pushing the envelope: structure, function, and dynamics of the nuclear periphery. Annu. Rev. Cell Dev. Biol. *21*, 347–380.

109. Hinz, B., and Gabbiani, G. (2003). Cell-matrix and cell-cell contacts of myofibroblasts: role in connective tissue remodeling. Thromb. Haemost. *90*, 993–1002.

110. Ho, C.Y., and Lammerding, J. (2012). Lamins at a glance. J. Cell Sci. *125*, 2087–2093.

111. Ho, C.Y., Jaalouk, D.E., Vartiainen, M.K., and Lammerding, J. (2013). Lamin A/C and emerin regulate MKL1-SRF activity by modulating actin dynamics. Nature *497*, 507–511.

112. Højfeldt, J.W., Agger, K., and Helin, K. (2013). Histone lysine demethylases as targets for anticancer therapy. Nat. Rev. Drug Discov. *12*, 917–930.

113. Holaska, J.M., Kowalski, A.K., and Wilson, K.L. (2004). Emerin caps the pointed end of actin filaments: evidence for an actin cortical network at the nuclear inner membrane. PLoS Biol. *2*, E231.

114. Huang, B., Lu, M., Jia, D., Ben-Jacob, E., Levine, H., and Onuchic, J.N. (2017). Interrogating the topological robustness of gene regulatory circuits by randomization. PLoS Comput. Biol. *13*, e1005456.

115. Ibarra, A., Benner, C., Tyagi, S., Cool, J., and Hetzer, M.W. (2016). Nucleoporin-mediated regulation of cell identity genes. Genes Dev. *30*, 2253–2258.

116. Jan, Y.N., and Jan, L.Y. (1993). HLH proteins, fly neurogenesis, and vertebrate myogenesis. Cell *75*, 827–830.

117. Jia, Y., Vong, J.S.-L., Asafova, A., Garvalov, B.K., Caputo, L., Cordero, J., Singh, A., Boettger, T., Günther, S., Fink, L., et al. (2019). Lamin B1 loss promotes lung cancer development and metastasis by epigenetic derepression of RET. J. Exp. Med. *216*, 1377–1395.

118. Johnson, K.S., Hussein, S., Song, S., Chakraborty, P., Jolly, M.K., Toneff, M., Lin, Y., and Taube, J. (2020). Gene expression and chromatin accessibility during progressive EMT and MET linked to dynamic CTCF engagement. BioRxiv.

119. Johnson, V.L., Scott, M.I.F., Holt, S.V., Hussein, D., and Taylor, S.S. (2004). Bub1 is required for kinetochore localization of BubR1, Cenp-E, Cenp-F and Mad2, and chromosome congression. J. Cell Sci. *117*, 1577–1589.

120. Jolly, M.K., Boareto, M., Huang, B., Jia, D., Lu, M., Ben-Jacob, E., Onuchic, J.N., and Levine, H. (2015). Implications of the hybrid epithelial/mesenchymal phenotype in metastasis. Front. Oncol. *5*, 155.

121. Jones, J., Wang, H., Zhou, J., Hardy, S., Turner, T., Austin, D., He, Q., Wells, A., Grizzle, W.E., and Yates, C. (2012). Nuclear Kaiso indicates aggressive prostate cancers and promotes migration and invasiveness of prostate cancer cells. Am. J. Pathol. *181*, 1836–1846.

122. Jones, J., Wang, H., Karanam, B., Theodore, S., Dean-Colomb, W., Welch, D.R., Grizzle, W., and Yates, C. (2014). Nuclear localization of Kaiso promotes the poorly differentiated phenotype and EMT in infiltrating ductal carcinomas. Clin Exp Metastasis *31*, 497–510.

123. Jürgens, G., Wieschaus, E., Nüsslein-Volhard, C., and Kluding, H. (1984). Mutations affecting the pattern of the larval cuticle inDrosophila melanogaster : II. Zygotic loci on the third chromosome. Wilhelm Roux'. Archiv. *193*, 283–295.

124. Kalluri, R., and Neilson, E.G. (2003). Epithelial-mesenchymal transition and its implications for fibrosis. J. Clin. Invest. *112*, 1776–1784.

125. Kalluri, R., and Weinberg, R.A. (2009). The basics of epithelial-mesenchymal transition. J. Clin. Invest. *119*, 1420–1428.

126. Khanna, N., Hu, Y., and Belmont, A.S. (2014). HSP70 transgene directed motion to nuclear speckles facilitates heat shock activation. Curr. Biol. *24*, 1138–1144.

127. Kim, T., Veronese, A., Pichiorri, F., Lee, T.J., Jeon, Y.-J., Volinia, S., Pineau, P., Marchio, A., Palatini, J., Suh, S.-S., et al. (2011). p53 regulates epithelial-mesenchymal transition through microRNAs targeting ZEB1 and ZEB2. J. Exp. Med. *208*, 875–883.

128. Kim, Y., Zheng, X., and Zheng, Y. (2019). Role of lamins in 3D genome organization and global gene expression. Nucleus *10*, 33–41.

129. Kimura, H. (2013). Histone modifications for human epigenome analysis. J. Hum. Genet. *58*, 439–445.

130. Kong, W., Yang, H., He, L., Zhao, J., Coppola, D., Dalton, W.S., and Cheng, J.Q. (2008). MicroRNA-155 is regulated by the transforming growth factor beta/Smad pathway and contributes to epithelial cell plasticity by targeting RhoA. Mol. Cell. Biol. *28*, 6773–6784.

131. Korpal, M., Ell, B.J., Buffa, F.M., Ibrahim, T., Blanco, M.A., Celià-Terrassa, T., Mercatali, L., Khan, Z., Goodarzi, H., Hua, Y., et al. (2011). Direct targeting of Sec23a by miR-200s influences cancer cell secretome and promotes metastatic colonization. Nat. Med. *17*, 1101–1108.

132. Kuga, T., Nie, H., Kazami, T., Satoh, M., Matsushita, K., Nomura, F., Maeshima, K., Nakayama, Y., and Tomonaga, T. (2014). Lamin B2 prevents chromosome instability by ensuring proper mitotic chromosome segregation. Oncogenesis *3*, e94.

133. Kumarswamy, R., Mudduluru, G., Ceppi, P., Muppala, S., Kozlowski, M., Niklinski, J., Papotti, M., and Allgayer, H. (2012). MicroRNA-30a inhibits epithelial-to-mesenchymal transition by targeting Snai1 and is downregulated in non-small cell lung cancer. Int. J. Cancer *130*, 2044–2053.

134. Labade, A.S., Karmodiya, K., and Sengupta, K. (2016). HOXA repression is mediated by nucleoporin Nup93 assisted by its interactors Nup188 and Nup205. Epigenetics Chromatin *9*, 54.

135. Lamouille, S., Xu, J., and Derynck, R. (2014). Molecular mechanisms of epithelial-mesenchymal transition. Nat. Rev. Mol. Cell Biol. *15*, 178–196.

136. Latil, M., Nassar, D., Beck, B., Boumahdi, S., Wang, L., Brisebarre, A., Dubois, C., Nkusi, E., Lenglez, S., Checinska, A., et al. (2017). Cell-Type-Specific Chromatin States Differentially Prime Squamous Cell Carcinoma Tumor-Initiating Cells for Epithelial to Mesenchymal Transition. Cell Stem Cell *20*, 191–204.e5.

137. Lee, K.K., and Workman, J.L. (2007). Histone acetyltransferase complexes: one size doesn't fit all. Nat. Rev. Mol. Cell Biol. *8*, 284–295.

138. Leggett, S.E., Sim, J.Y., Rubins, J.E., Neronha, Z.J., Williams, E.K., and Wong, I.Y. (2016). Morphological single cell profiling of the epithelial-mesenchymal transition. Integr Biol (Camb) *8*, 1133–1144.

139. Leshner, M., Devine, M., Roloff, G.W., True, L.D., Misteli, T., and Meaburn, K.J. (2016). Locus-specific gene repositioning in prostate cancer. Mol. Biol. Cell *27*, 236–246.

140. Di Leva, G., Gasparini, P., Piovan, C., Ngankeu, A., Garofalo, M., Taccioli, C., Iorio, M.V., Li, M., Volinia, S., Alder, H., et al. (2010). MicroRNA cluster 221-222 and estrogen receptor alpha interactions in breast cancer. J. Natl. Cancer Inst. *102*, 706–721.

141. Li, F., Huarte, M., Zaratiegui, M., Vaughn, M.W., Shi, Y., Martienssen, R., and Cande, W.Z. (2008). Lid2 is required for coordinating H3K4 and H3K9 methylation of heterochromatin and euchromatin. Cell *135*, 272–283.

142. Liang, Y., Franks, T.M., Marchetto, M.C., Gage, F.H., and Hetzer, M.W. (2013). Dynamic association of NUP98 with the human genome. PLoS Genet. *9*, e1003308.

143. Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289–293.

144. Lillie, F.R. (1908). The development of the chick.

145. Lin, F., Morrison, J.M., Wu, W., and Worman, H.J. (2005). MAN1, an integral protein of the inner nuclear membrane, binds Smad2 and Smad3 and antagonizes transforming growth factor-beta signaling. Hum. Mol. Genet. *14*, 437–445.

146. Liu, Y., and Bodmer, W.F. (2006). Analysis of P53 mutations and their expression in 56 colorectal cancer cell lines. Proc. Natl. Acad. Sci. USA *103*, 976–981.

147. Liu, S., Ye, D., Guo, W., Yu, W., He, Y., Hu, J., Wang, Y., Zhang, L., Liao, Y., Song, H., et al. (2015). G9a is essential for EMT-mediated metastasis and maintenance of cancer stem cell-like characters in head and neck squamous cell carcinoma. Oncotarget *6*, 6887–6901.

148. Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402–408.

149. Logarinho, E., Bousbaa, H., Dias, J.M., Lopes, C., Amorim, I., Antunes-Martins, A., and Sunkel, C.E. (2004). Different spindle checkpoint proteins monitor microtubule attachment and tension at kinetochores in Drosophila cells. J. Cell Sci. *117*, 1757–1771.

150. Lombaerts, M., van Wezel, T., Philippo, K., Dierssen, J.W.F., Zimmerman, R.M.E., Oosting, J., van Eijk, R., Eilers, P.H., van de Water, B., Cornelisse, C.J., et al. (2006). E-cadherin transcriptional downregulation by promoter methylation but not mutation is related to epithelial-to-mesenchymal transition in breast cancer cell lines. Br. J. Cancer *94*, 661–671.

151. Lund, E.G., Duband-Goulet, I., Oldenburg, A., Buendia, B., and Collas, P. (2015). Distinct features of lamin A-interacting chromatin domains mapped by ChIP-sequencing from sonicated or micrococcal nuclease-digested chromatin. Nucleus *6*, 30–39.

152. Luo, W.-R., Wu, A.-B., Fang, W.-Y., Li, S.-Y., and Yao, K.-T. (2012). Nuclear expression of N-cadherin correlates with poor prognosis of nasopharyngeal carcinoma. Histopathology *61*, 237–246.

153. Ma, L., Young, J., Prabhala, H., Pan, E., Mestdagh, P., Muth, D., Teruya-Feldstein, J., Reinhardt, F., Onder, T.T., Valastyan, S., et al. (2010). miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis. Nat. Cell Biol. *12*, 247–256.

154. Macías-Silva, M., Abdollah, S., Hoodless, P.A., Pirone, R., Attisano, L., and Wrana, J.L. (1996). MADR2 is a substrate of the TGFbeta receptor and its phosphorylation is required for nuclear accumulation and signaling. Cell *87*, 1215–1224.

155. Majello, B., De Luca, P., and Lania, L. (1997). Sp3 is a bifunctional transcription regulator with modular independent activation and repression domains. J. Biol. Chem. *272*, 4021–4026.

156. Malouf, G.G., Taube, J.H., Lu, Y., Roysarkar, T., Panjarian, S., Estecio, M.R., Jelinek, J., Yamazaki, J., Raynal, N.J.-M., Long, H., et al. (2013). Architecture of epigenetic reprogramming following Twist1-mediated epithelial-mesenchymal transition. Genome Biol. *14*, R144.

157. Mani, S.A., Guo, W., Liao, M.-J., Eaton, E.N., Ayyanan, A., Zhou, A.Y., Brooks, M., Reinhard, F., Zhang, C.C., Shipitsin, M., et al. (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell *133*, 704–715.

158. Mao, Y.S., Sunwoo, H., Zhang, B., and Spector, D.L. (2011). Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. Nat. Cell Biol. *13*, 95–101.

159. Markiewicz, E., Tilgner, K., Barker, N., van de Wetering, M., Clevers, H., Dorobek, M., Hausmanowa-Petrusewicz, I., Ramaekers, F.C.S., Broers, J.L.V., Blankesteijn, W.M., et al. (2006). The inner nuclear membrane protein emerin regulates beta-catenin activity by restricting its accumulation in the nucleus. EMBO J. *25*, 3275–3285.

160. Markowitz, S., Wang, J., Myeroff, L., Parsons, R., Sun, L., Lutterbaugh, J., Fan, R.S., Zborowska, E., Kinzler, K.W., and Vogelstein, B. (1995). Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. Science *268*, 1336–1338.

161. Martello, G., Rosato, A., Ferrari, F., Manfrin, A., Cordenonsi, M., Dupont, S., Enzo, E., Guzzardo, V., Rondina, M., Spruce, T., et al. (2010). A MicroRNA targeting dicer for metastasis control. Cell *141*, 1195–1207.

162. Massagué, J. (1992). Receptors for the TGF-beta family. Cell *69*, 1067–1070.

163. Massagué, J., Seoane, J., and Wotton, D. (2005). Smad transcription factors. Genes Dev. *19*, 2783–2810.

164. Massari, M.E., and Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. Mol. Cell. Biol. *20*, 429–440.

165. McDonald, O.G., Wu, H., Timp, W., Doi, A., and Feinberg, A.P. (2011). Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. Nat. Struct. Mol. Biol. *18*, 867–874.

166. Meaburn, K.J., and Misteli, T. (2008). Locus-specific and activity-independent gene repositioning during early tumorigenesis. J. Cell Biol. *180*, 39–50.

167. Meaburn, K.J., Cabuy, E., Bonne, G., Levy, N., Morris, G.E., Novelli, G., Kill, I.R., and Bridger, J.M. (2007). Primary laminopathy fibroblasts display altered genome organization and apoptosis. Aging Cell *6*, 139–153.

168. Mehlen, P., and Puisieux, A. (2006). Metastasis: a question of life or death. Nat. Rev. Cancer *6*, 449–458.

169. Melcer, S., Hezroni, H., Rand, E., Nissim-Rafinia, M., Skoultchi, A., Stewart, C.L., Bustin, M., and Meshorer, E. (2012). Histone modifications and lamin A regulate

chromatin protein dynamics in early embryonic stem cell differentiation. Nat. Commun. *3*, 910.

170. Meng, J., Chen, S., Han, J.-X., Qian, B., Wang, X.-R., Zhong, W.-L., Qin, Y., Zhang, H., Gao, W.-F., Lei, Y.-Y., et al. (2018). Twist1 Regulates Vimentin through Cul2 Circular RNA to Promote EMT in Hepatocellular Carcinoma. Cancer Res. *78*, 4150–4162.

171. Meuleman, W., Peric-Hupkes, D., Kind, J., Beaudry, J.-B., Pagie, L., Kellis, M., Reinders, M., Wessels, L., and van Steensel, B. (2013). Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence. Genome Res. *23*, 270–280.

172. Miettinen, P.J., Ebner, R., Lopez, A.R., and Derynck, R. (1994). TGF-beta induced transdifferentiation of mammary epithelial cells to mesenchymal cells: involvement of type I receptors. J. Cell Biol. *127*, 2021–2036.

173. Miralles, F., Posern, G., Zaromytidou, A.-I., and Treisman, R. (2003). Actin dynamics control SRF activity by regulation of its coactivator MAL. Cell *113*, 329–342.

174. Mironchik, Y., Winnard, P.T., Vesuna, F., Kato, Y., Wildes, F., Pathak, A.P., Kominsky, S., Artemov, D., Bhujwalla, Z., Van Diest, P., et al. (2005). Twist overexpression induces in vivo angiogenesis and correlates with chromosomal instability in breast cancer. Cancer Res. *65*, 10801–10809.

175. Miyazono, K. (2009). Transforming growth factor-beta signaling in epithelial-mesenchymal transition and progression of cancer. Proc. Jpn. Acad. Ser. B. Phys. Biol. Sci. *85*, 314–323.

176. Mizuguchi, Y., Specht, S., Lunz, J.G., Isse, K., Corbitt, N., Takizawa, T., and Demetris, A.J. (2012). Cooperation of p300 and PCAF in the control of microRNA 200c/141 transcription and epithelial characteristics. PLoS One *7*, e32449.

177. Mongroo, P.S., and Rustgi, A.K. (2010). The role of the miR-200 family in epithelial-mesenchymal transition. Cancer Biol. Ther. *10*, 219–222.

178. Moustakas, A., and Heldin, C.-H. (2007). Signaling networks guiding epithelial-mesenchymal transitions during embryogenesis and cancer progression. Cancer Sci. *98*, 1512–1520.

179. Murre, C., McCaw, P.S., Vaessin, H., Caudy, M., Jan, L.Y., Jan, Y.N., Cabrera, C.V., Buskin, J.N., Hauschka, S.D., and Lassar, A.B. (1989). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell *58*, 537–544.

180. Murre, C., Bain, G., van Dijk, M.A., Engel, I., Furnari, B.A., Massari, M.E., Matthews, J.R., Quong, M.W., Rivera, R.R., and Stuiver, M.H. (1994). Structure and function of helix-loop-helix proteins. Biochim. Biophys. Acta *1218*, 129–135.

181. Németh, A., Conesa, A., Santoyo-Lopez, J., Medina, I., Montaner, D., Péterfia, B., Solovei, I., Cremer, T., Dopazo, J., and Längst, G. (2010). Initial genomics of the human nucleolus. PLoS Genet. *6*, e1000889.

182. Nie, D., Fu, J., Chen, H., Cheng, J., and Fu, J. (2019). Roles of MicroRNA-34a in Epithelial to Mesenchymal Transition, Competing Endogenous RNA Sponging and Its Therapeutic Potential. Int. J. Mol. Sci. *20*.

183. Nieto, M.A., and Cano, A. (2012). The epithelial-mesenchymal transition under control: global programs to regulate epithelial plasticity. Semin. Cancer Biol. *22*, 361–368.

184. Nieto, M.A., Huang, R.Y.-J., Jackson, R.A., and Thiery, J.P. (2016). EMT: 2016. Cell *166*, 21–45.

185. Nojadeh, J.N., Behrouz Sharif, S., and Sakhinia, E. (2018). Microsatellite instability in colorectal cancer. EXCLI J. *17*, 159–168.

186. Nüsslein-Volhard, C., Wieschaus, E., and Kluding, H. (1984). Mutations affecting the pattern of the larval cuticle inDrosophila melanogaster : I. Zygotic loci on the second chromosome. Wilhelm Roux'. Archiv. *193*, 267–282.

187. Van Oorschot, B., Oei, A.L., Nuijens, A.C., Rodermond, H., Hoeben, R., Stap, J., Stalpers, L.J., and Franken, N.A.P. (2014). Decay of γ-H2AX foci correlates with potentially lethal damage repair and P53 status in human colorectal carcinoma cells. Cell Mol Biol Lett *19*, 37–51.

188. Panda, S., Antoch, M.P., Miller, B.H., Su, A.I., Schook, A.B., Straume, M., Schultz, P.G., Kay, S.A., Takahashi, J.S., and Hogenesch, J.B. (2002). Coordinated transcription of key pathways in the mouse by the circadian clock. Cell *109*, 307–320.

189. Park, S.-M., Gaur, A.B., Lengyel, E., and Peter, M.E. (2008). The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. Genes Dev. *22*, 894–907.

190. Pascual-Reguant, L., Blanco, E., Galan, S., Le Dily, F., Cuartero, Y., Serra-Bardenys, G., Di Carlo, V., Iturbide, A., Cebrià-Costa, J.P., Nonell, L., et al. (2018). Lamin B1 mapping reveals the existence of dynamic and functional euchromatin lamin B1 domains. Nat. Commun. *9*, 3420.

191. Pastushenko, I., Brisebarre, A., Sifrim, A., Fioramonti, M., Revenco, T., Boumahdi, S., Van Keymeulen, A., Brown, D., Moers, V., Lemaire, S., et al. (2018). Identification of the tumour transition states occurring during EMT. Nature *556*, 463–468.

192. Paulin-Levasseur, M., Blake, D.L., Julien, M., and Rouleau, L. (1996). The MAN antigens are non-lamin constituents of the nuclear lamina in vertebrate cells. Chromosoma *104*, 367–379.

193. Paulsen, J., Liyakat Ali, T.M., Nekrasov, M., Delbarre, E., Baudement, M.-O., Kurscheid, S., Tremethick, D., and Collas, P. (2019). Long-range interactions between topologically associating domains shape the four-dimensional genome during differentiation. Nat. Genet. *51*, 835–843.

194. Pederson, T. (2011). The nucleus introduced. Cold Spring Harb. Perspect. Biol. *3*.

195. Peinado, H., Ballestar, E., Esteller, M., and Cano, A. (2004). Snail mediates E-cadherin repression by the recruitment of the Sin3A/histone deacetylase 1 (HDAC1)/HDAC2 complex. Mol. Cell. Biol. *24*, 306–319.

196. Peinado, H., Olmeda, D., and Cano, A. (2007). Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? Nat. Rev. Cancer *7*, 415–428.

197. Peric-Hupkes, D., Meuleman, W., Pagie, L., Bruggeman, S.W.M., Solovei, I., Brugman, W., Gräf, S., Flicek, P., Kerkhoven, R.M., van Lohuizen, M., et al. (2010). Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. Mol. Cell *38*, 603–613.

198. Piccinin, S., Tonin, E., Sessa, S., Demontis, S., Rossi, S., Pecciarini, L., Zanatta, L., Pivetta, F., Grizzo, A., Sonego, M., et al. (2012). A "twist box" code of p53 inactivation: twist box: p53 interaction promotes p53 degradation. Cancer Cell *22*, 404–415.

199. Piek, E., Moustakas, A., Kurisaki, A., Heldin, C.H., and ten Dijke, P. (1999). TGF-(beta) type I receptor/ALK-5 and Smad proteins mediate epithelial to mesenchymal transdifferentiation in NMuMG breast epithelial cells. J. Cell Sci. *112 (Pt 24)*, 4557–4568.

200. Pinto, M., Oliveira, C., Cirnes, L., Carlos Machado, J., Ramires, M., Nogueira, A., Carneiro, F., and Seruca, R. (2003). Promoter methylation of TGFbeta receptor I and mutation of TGFbeta receptor II are frequent events in MSI sporadic gastric carcinomas. J. Pathol. *200*, 32–38.

201. Pradhan, R., Ranade, D., and Sengupta, K. (2018). Emerin modulates spatial organization of chromosome territories in cells on softer matrices. Nucleic Acids Res. *46*, 5561–5586.

202. Pyrpasopoulou, A., Meier, J., Maison, C., Simos, G., and Georgatos, S.D. (1996). The lamin B receptor (LBR) provides essential chromatin docking sites at the nuclear envelope. EMBO J. *15*, 7108–7119.

203. Qiu, Y., and Huang, S. (2020). CTCF-mediated genome organization and leukemogenesis. Leukemia *34*, 2295–2304.

204. Quinodoz, S.A., Ollikainen, N., Tabak, B., Palla, A., Schmidt, J.M., Detmar, E., Lai, M.M., Shishkin, A.A., Bhat, P., Takei, Y., et al. (2018). Higher-Order Inter-chromosomal Hubs Shape 3D Genome Organization in the Nucleus. Cell *174*, 744–757.e24.

205. Rabl, and C. (1885). Uber Zelltheilung. Morphol. Jahrb.

206. Rajagopalan, H., Nowak, M.A., Vogelstein, B., and Lengauer, C. (2003). The significance of unstable chromosomes in colorectal cancer. Nat. Rev. Cancer *3*, 695–701.

207. Ranade, D., Koul, S., Thompson, J., Prasad, K.B., and Sengupta, K. (2017). Chromosomal aneuploidies induced upon Lamin B2 depletion are mislocalized in the interphase nucleus. Chromosoma *126*, 223–244.

208. Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665–1680.

209. Rashmi, R.N., Eckes, B., Glöckner, G., Groth, M., Neumann, S., Gloy, J., Sellin, L., Walz, G., Schneider, M., Karakesisoglou, I., et al. (2012). The nuclear envelope protein Nesprin-

2 has roles in cell proliferation and differentiation during wound healing. Nucleus *3*, 172–186.

210. Reinhold, W.C., Reimers, M.A., Maunakea, A.K., Kim, S., Lababidi, S., Scherf, U., Shankavaram, U.T., Ziegler, M.S., Stewart, C., Kouros-Mehr, H., et al. (2007). Detailed DNA methylation profiles of the E-cadherin promoter in the NCI-60 cancer cells. Mol. Cancer Ther. *6*, 391–403.

211. Rio, D.C., Ares, M., Hannon, G.J., and Nilsen, T.W. (2010). Purification of RNA using TRIzol (TRI reagent). Cold Spring Harb. Protoc. *2010*, pdb.prot5439.

212. Roche, J., Nasarre, P., Gemmill, R., Baldys, A., Pontis, J., Korch, C., Guilhot, J., Ait-Si-Ali, S., and Drabkin, H. (2013). Global Decrease of Histone H3K27 Acetylation in ZEB1-Induced Epithelial to Mesenchymal Transition in Lung Cancer Cells. Cancers (Basel) *5*, 334–356.

213. Rodríguez, J., Calvo, F., González, J.M., Casar, B., Andrés, V., and Crespo, P. (2010). ERK1/2 MAP kinases promote cell cycle entry by rapid, kinase-independent disruption of retinoblastoma-lamin A complexes. J. Cell Biol. *191*, 967–979.

214. Rothballer, A., Schwartz, T.U., and Kutay, U. (2013). LINCing complex functions at the nuclear envelope. Nucleus *4*, 29–36.

215. Rowley, M.J., and Corces, V.G. (2018). Organizational principles of 3D genome architecture. Nat. Rev. Genet. *19*, 789–800.

216. Sacchetti, A., Teeuwssen, M., Verhagen, M., Joosten, R., Gusinac, A., Watson, M.M., Stabile, R., Kim, W.K., Ubink, I., van Werken, H.J.G., et al. (2020). Phenotypic plasticity and partial EMT underlie local invasion and distant metastasis in colon cancer. BioRxiv.

217. Samwer, M., Schneider, M.W.G., Hoefler, R., Schmalhorst, P.S., Jude, J.G., Zuber, J., and Gerlich, D.W. (2017). DNA Cross-Bridging Shapes a Single Nucleus from a Set of Mitotic Chromosomes. Cell *170*, 956–972.e23.

218. Schoenfelder, S., and Fraser, P. (2019). Long-range enhancer-promoter contacts in gene expression control. Nat. Rev. Genet. *20*, 437–455.

219. Shachar, S., Voss, T.C., Pegoraro, G., Sciascia, N., and Misteli, T. (2015). Identification of Gene Positioning Factors Using High-Throughput Imaging Mapping. Cell *162*, 911–923.

220. Shi, Y., and Massagué, J. (2003). Mechanisms of TGF-beta signaling from cell membrane to the nucleus. Cell *113*, 685–700.

221. Simonetti, G., Bruno, S., Padella, A., Tenti, E., and Martinelli, G. (2019). Aneuploidy: Cancer strength or vulnerability? Int. J. Cancer *144*, 8–25.

222. Simpson, P. (1983). Maternal-Zygotic Gene Interactions during Formation of the Dorsoventral Pattern in Drosophila Embryos. Genetics *105*, 615–632.

223. Smith, E.M., Lajoie, B.R., Jain, G., and Dekker, J. (2016). Invariant TAD Boundaries Constrain Cell-Type-Specific Looping Interactions between Promoters and Distal Elements around the CFTR Locus. Am. J. Hum. Genet. *98*, 185–201.

224. Smith, E.R., Capo-Chichi, C.D., and Xu, X.-X. (2018). Defective nuclear lamina in aneuploidy and carcinogenesis. Front. Oncol. *8*, 529.

225. Solovei, I., Wang, A.S., Thanisch, K., Schmidt, C.S., Krebs, S., Zwerger, M., Cohen, T.V., Devys, D., Foisner, R., Peichl, L., et al. (2013). LBR and lamin A/C sequentially tether peripheral heterochromatin and inversely regulate differentiation. Cell *152*, 584–598.

226. Stinson, S., Lackner, M.R., Adai, A.T., Yu, N., Kim, H.-J., O'Brien, C., Spoerke, J., Jhunjhunwala, S., Boyd, Z., Januario, T., et al. (2011). miR-221/222 targeting of trichorhinophalangeal 1 (TRPS1) promotes epithelial-to-mesenchymal transition in breast cancer. Sci. Signal. *4*, pt5.

227. Stoker, M., Gherardi, E., Perryman, M., and Gray, J. (1987). Scatter factor is a fibroblast-derived modulator of epithelial cell mobility. Nature *327*, 239–242.

228. Stone, R.C., Pastar, I., Ojeh, N., Chen, V., Liu, S., Garzon, K.I., and Tomic-Canic, M. (2016). Epithelial-mesenchymal transition in tissue repair and fibrosis. Cell Tissue Res. *365*, 495–506.

229. Su, Y.J., Chang, Y.W., Lin, W.H., Liang, C.L., and Lee, J.L. (2015). An aberrant nuclear localization of E-cadherin is a potent inhibitor of Wnt/β-catenin-elicited promotion of the cancer stem cell phenotype. Oncogenesis *4*, e157.

230. Sun, H.B., Shen, J., and Yokota, H. (2000). Size-dependent positioning of human chromosomes in interphase nuclei. Biophys. J. *79*, 184–190.

231. Sun, L., Yao, Y., Liu, B., Lin, Z., Lin, L., Yang, M., Zhang, W., Chen, W., Pan, C., Liu, Q., et al. (2012). MiR-200b and miR-15b regulate chemotherapy-induced epithelial-mesenchymal transition in human tongue cancer cells by targeting BMI1. Oncogene *31*, 432–445.

232. Sun, Y., Durrin, L.K., and Krontiris, T.G. (2003). Specific interaction of PML bodies with the TP53 locus in Jurkat interphase nuclei. Genomics *82*, 250–252.

233. Szabo, Q., Jost, D., Chang, J.-M., Cattoni, D.I., Papadopoulos, G.L., Bonev, B., Sexton, T., Gurgo, J., Jacquier, C., Nollmann, M., et al. (2018). TADs are 3D structural units of higher-order chromosome organization in Drosophila. Sci. Adv. *4*, eaar8082.

234. Szczerbal, I., Foster, H.A., and Bridger, J.M. (2009). The spatial repositioning of adipogenesis genes is correlated with their expression status in a porcine mesenchymal stem cell adipogenesis model system. Chromosoma *118*, 647–663.

235. Taimen, P., Pfleghaar, K., Shimi, T., Möller, D., Ben-Harush, K., Erdos, M.R., Adam, S.A., Herrmann, H., Medalia, O., Collins, F.S., et al. (2009). A progeria mutation reveals functions for lamin A in nuclear assembly, architecture, and chromosome organization. Proc. Natl. Acad. Sci. USA *106*, 20788–20793.

236. Thiery, J.P. (2002). Epithelial-mesenchymal transitions in tumour progression. Nat. Rev. Cancer *2*, 442–454.

237. Thiery, J.P., Acloque, H., Huang, R.Y.J., and Nieto, M.A. (2009). Epithelial-mesenchymal transitions in development and disease. Cell *139*, 871–890.

238. Thisse, B., el Messal, M., and Perrin-Schmitt, F. (1987). The twist gene: isolation of a Drosophila zygotic gene necessary for the establishment of dorsoventral pattern. Nucleic Acids Res. *15*, 3439–3453.

239. Tiwari, N., Gheldof, A., Tatari, M., and Christofori, G. (2012). EMT as the ultimate survival mechanism of cancer cells. Semin. Cancer Biol. *22*, 194–207.

240. Tornesello, M.L., Faraonio, R., Buonaguro, L., Annunziata, C., Starita, N., Cerasuolo, A., Pezzuto, F., Tornesello, A.L., and Buonaguro, F.M. (2020). The Role of microRNAs, Long Non-coding RNAs, and Circular RNAs in Cervical Cancer. Front. Oncol. *10*, 150.

241. Tran, D.D., Corsa, C.A.S., Biswas, H., Aft, R.L., and Longmore, G.D. (2011). Temporal and spatial cooperation of Snail1 and Twist1 during epithelial-mesenchymal transition predicts for human breast cancer recurrence. Mol. Cancer Res. *9*, 1644–1657.

242. Trelstad, R.L., Hay, E.D., and Revel, J.D. (1967). Cell contact during early morphogenesis in the chick embryo. Dev. Biol. *16*, 78–106.

243. Tretbar, S., Krausbeck, P., Müller, A., Friedrich, M., Vaxevanis, C., Bukur, J., Jasinski-Bergner, S., and Seliger, B. (2019). TGF-β inducible epithelial-to-mesenchymal transition in renal cell carcinoma. Oncotarget *10*, 1507–1524.

244. Trojer, P., and Reinberg, D. (2007). Facultative heterochromatin: is there a distinctive molecular signature? Mol. Cell *28*, 1–13.

245. Tumbar, T., Sudlow, G., and Belmont, A.S. (1999). Large-scale chromatin unfolding and remodeling induced by VP16 acidic activation domain. J. Cell Biol. *145*, 1341–1354.

246. Valcourt, U., Kowanetz, M., Niimi, H., Heldin, C.-H., and Moustakas, A. (2005). TGF-beta and the Smad signaling pathway support transcriptomic reprogramming during epithelial-mesenchymal cell transition. Mol. Biol. Cell *16*, 1987–2002.

247. van Koningsbruggen, S., Gierlinski, M., Schofield, P., Martin, D., Barton, G.J., Ariyurek, Y., den Dunnen, J.T., and Lamond, A.I. (2010). High-resolution whole-genome sequencing reveals that specific chromatin domains from most human chromosomes associate with nucleoli. Mol. Biol. Cell *21*, 3735–3748.

248. Vander Ark, A., Cao, J., and Li, X. (2018). TGF-β receptors: In and beyond TGF-β signaling. Cell Signal. *52*, 112–120.

249. Varga, T., and Aplan, P.D. (2005). Chromosomal aberrations induced by double strand DNA breaks. DNA Repair (Amst.) *4*, 1038–1046.

250. Vargas-Rondón, N., Villegas, V.E., and Rondón-Lagos, M. (2017). The role of chromosomal instability in cancer and therapeutic responses. Cancers (Basel) *10*.

251. Vasudevan, A., Baruah, P.S., Smith, J.C., Wang, Z., Sayles, N.M., Andrews, P., Kendall, J., Chunduri, N.K., Levy, D., Wigler, M., et al. (2019). Single chromosome gains can function as metastasis suppressors and metastasis promoters. BioRxiv.

252. Vesuna, F., Winnard, P., Glackin, C., and Raman, V. (2006). Twist overexpression promotes chromosomal instability in the breast cancer cell line MCF-7. Cancer Genet Cytogenet *167*, 189–191.

253. Vesuna, F., van Diest, P., Chen, J.H., and Raman, V. (2008). Twist is a transcriptional repressor of E-cadherin gene expression in breast cancer. Biochem. Biophys. Res. Commun. *367*, 235–241.

254. Vićovac, L., and Aplin, J.D. (1996). Epithelial-mesenchymal transition during trophoblast differentiation. Acta Anat. (Basel) *156*, 202–216.

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

256. Wang, F.E., Zhang, C., Maminishkis, A., Dong, L., Zhi, C., Li, R., Zhao, J., Majerciak, V., Gaur, A.B., Chen, S., et al. (2010). MicroRNA-204/211 alters epithelial physiology. FASEB J. *24*, 1552–1571.

257. Watson, M.L. (1955). The nuclear envelope; its structure and relation to cytoplasmic membranes. J Biophys Biochem Cytol *1*, 257–270.

258. Webster, M., Witkin, K.L., and Cohen-Fix, O. (2009). Sizing up the nucleus: nuclear shape, size and nuclear-envelope assembly. J. Cell Sci. *122*, 1477–1486.

259. Weidner, K.M., Sachs, M., and Birchmeier, W. (1993). The Met receptor tyrosine kinase transduces motility, proliferation, and morphogenic signals of scatter factor/hepatocyte growth factor in epithelial cells. J. Cell Biol. *121*, 145–154.

260. Willis, N.D., Cox, T.R., Rahman-Casañs, S.F., Smits, K., Przyborski, S.A., van den Brandt, P., van Engeland, M., Weijenberg, M., Wilson, R.G., de Bruïne, A., et al. (2008). Lamin A/C is a risk biomarker in colorectal cancer. PLoS One *3*, e2988.

261. Wong, T.-S., Gao, W., and Chan, J.Y.-W. (2014). Transcription regulation of E-cadherin by zinc finger E-box binding homeobox proteins in solid tumors. Biomed Res. Int. *2014*, 921564.

262. Wu, Z., Wu, L., Weng, D., Xu, D., Geng, J., and Zhao, F. (2009). Reduced expression of lamin A/C correlates with poor histological differentiation and prognosis in primary gastric carcinoma. J Exp Clin Cancer Res *28*, 8.

263. Xia, J., Benner, M.J., and Hancock, R.E.W. (2014). NetworkAnalyst--integrative approaches for protein-protein interaction network analysis and visual exploration. Nucleic Acids Res. *42*, W167–74.

264. Xu, J., Zhu, W., Xu, W., Yao, W., Zhang, B., Xu, Y., Ji, S., Liu, C., Long, J., Ni, Q., et al. (2013). Up-regulation of MBD1 promotes pancreatic cancer cell epithelial-mesenchymal transition and invasion by epigenetic down-regulation of E-cadherin. Curr. Mol. Med. *13*, 387–400.

265. Yan, F., Wang, Y., Wu, X., Peshavariya, H.M., Dusting, G.J., Zhang, M., and Jiang, F. (2014). Nox4 and redox signaling mediate TGF-β-induced endothelial cell apoptosis and phenotypic switch. Cell Death Dis. *5*, e1010.

266. Yang, J., Mani, S.A., Donaher, J.L., Ramaswamy, S., Itzykson, R.A., Come, C., Savagner, P., Gitelman, I., Richardson, A., and Weinberg, R.A. (2004). Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. Cell *117*, 927–939.

267. Yun, J., Song, S.-H., Kim, H.-P., Han, S.-W., Yi, E.C., and Kim, T.-Y. (2016). Dynamic cohesin-mediated chromatin architecture controls epithelial-mesenchymal plasticity in cancer. EMBO Rep. *17*, 1343–1359.

268. Zhang, Y.E. (2009). Non-Smad pathways in TGF-β signaling. Cell Res. *19*, 128–139.

269. Zhang, J., and Ma, L. (2012). MicroRNA control of epithelial-mesenchymal transition and metastasis. Cancer Metastasis Rev. *31*, 653–662.

270. Zhang, C.-Z., Spektor, A., Cornils, H., Francis, J.M., Jackson, E.K., Liu, S., Meyerson, M., and Pellman, D. (2015). Chromothripsis from DNA damage in micronuclei. Nature *522*, 179–184.

271. Zhang, W., Duan, N., Zhang, Q., Song, T., Li, Z., Chen, X., and Wang, K. (2018). The intracellular NADH level regulates atrophic nonunion pathogenesis through the CtBP2-p300-Runx2 transcriptional complex. Int. J. Biol. Sci. *14*, 2023–2036.

272. Zhang, Z., Liu, S., Shi, R., and Zhao, G. (2011). miR-27 promotes human gastric cancer cell metastasis by inducing epithelial-to-mesenchymal transition. Cancer Genet *204*, 486–491.

273. Zhao, Z., Rahman, M.A., Chen, Z.G., and Shin, D.M. (2017). Multiple biological functions of Twist1 in various cancers. Oncotarget *8*, 20380–20393.

274. Zhou, B., Liu, Y., Kahn, M., Ann, D.K., Han, A., Wang, H., Nguyen, C., Flodby, P., Zhong, Q., Krishnaveni, M.S., et al. (2012). Interactions between β-catenin and transforming growth factor-β signaling pathways mediate epithelial-mesenchymal transition and are dependent on the transcriptional co-activator cAMP-response element-binding protein (CREB)-binding protein (CBP). J. Biol. Chem. *287*, 7026–7038.

275. Zierhut, C., Jenness, C., Kimura, H., and Funabiki, H. (2014). Nucleosomal regulation of chromatin composition and nuclear assembly revealed by histone depletion. Nat. Struct. Mol. Biol. *21*, 617–625.

276. Zink, D., Fischer, A.H., and Nickerson, J.A. (2004). Nuclear structure in cancer cells. Nat. Rev. Cancer *4*, 677–687.

277. Zorn, C., Cremer, T., Cremer, C., and Zimmer, J. (1976). Laser UV microirradiation of interphase nuclei and post-treatment with caffeine. A new approach to establish the arrangement of interphase chromosomes. Hum. Genet. *35*, 83–89.

Publication

Khot, M., Sreekumar, D., Jahagirdar, S., Kulkarni, A., Hari, K., Faseela, E.E., Sabarinathan, R., Jolly, M.K., and Sengupta, K. (2020). Twist1 induces chromosomal instability (CIN) in colorectal cancer cells. Hum. Mol. Genet. *29*, 1673–1688

GENERAL ARTICLE

Twist1 induces chromosomal instability (CIN) in colorectal cancer cells

Maithilee Khot[1], Dyuthi Sreekumar[1,‡], Sanika Jahagirdar[1,‡],
Apoorva Kulkarni[1], Kishore Hari[2], Elangoli Ebrahimkutty Faseela[3],
Radhakrishnan Sabarinathan[3], Mohit Kumar Jolly[2] and
Kundan Sengupta[1,*,†]

[1]B-216, Chromosome Biology Lab (CBL), Indian Institute of Science Education and Research (IISER), Dr Homi Bhabha Road, Pashan, Pune 411008, India, [2]Center for BioSystems Science and Engineering, Indian Institute of Science, Bengaluru 560012, India and [3]National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bengaluru 560065, India

*To whom correspondence should be addressed at: B-216, Chromosome Biology Lab (CBL), Indian Institute of Science Education and Research (IISER), Dr Homi Bhabha Road, Pashan, Pune 411008, India. Tel: +91 20 25908071; Fax: +91-20-20251566; Email: kunsen@iiserpune.ac.in

Abstract

Twist1 is a basic helix-loop-helix transcription factor, essential during early development in mammals. While Twist1 induces epithelial-to-mesenchymal transition (EMT), here we show that Twist1 overexpression enhances nuclear and mitotic aberrations. This is accompanied by an increase in whole chromosomal copy number gains and losses, underscoring the role of Twist1 in inducing chromosomal instability (CIN) in colorectal cancer cells. Array comparative genomic hybridization (array CGH) analysis further shows sub-chromosomal deletions, consistent with an increased frequency of DNA double strand breaks (DSBs). Remarkably, Twist1 overexpression downmodulates key cell cycle checkpoint factors—Bub1, BubR1, Mad1 and Mad2—that regulate CIN. Mathematical simulations using the RACIPE tool show a negative correlation of Twist1 with E-cadherin and BubR1. Data analyses of gene expression profiles of patient samples from The Cancer Genome Atlas (TCGA) reveal a positive correlation between Twist1 and mesenchymal genes across cancers, whereas the correlation of TWIST1 with CIN and DSB genes is cancer subtype-specific. Taken together, these studies highlight the mechanistic involvement of Twist1 in the deregulation of factors that maintain genome stability during EMT in colorectal cancer cells. Twist1 overexpression enhances genome instability in the context of EMT that further contributes to cellular heterogeneity. In addition, these studies imply that Twist1 downmodulates nuclear lamins that further alter spatiotemporal organization of the cancer genome and epigenome. Notwithstanding their genetic background, colorectal cancer cells nevertheless maintain their overall ploidy, while the downstream effects of Twist1 enhance CIN and DNA damage enriching for sub-populations of aggressive cancer cells.

Introduction

Twist1 is a basic helix-loop-helix (bHLH) transcription factor that is essential for normal vertebrate development, but is overexpressed in cancers of the breast, prostate and stomach, including melanomas, gliomas and osteosarcomas (1,2). Increase in Twist1 levels is implicated in dissemination of tumorigenic cells and chemoresistance (3). Twist1 is a master regulator of epithelial-to-mesenchymal transition (EMT) (4) and promotes stemness (5)—a characteristic feature of EMT (6–8). Twist1 binds to the promoter of the E-cadherin gene (that encodes for a cell adhesion protein) and suppresses its expression (9). Decrease in E-cadherin levels reduces the cobblestone morphology of epithelial cells, also facilitating their dissemination (10). Consistently, a subpopulation of breast, colorectal, prostate and lung carcinomas shows Twist1 expression, typically at the invasive edge of cells (11). As Twist1 drives tumor progression, its contribution to EMT is extensively studied across cancers (4). However, the impact of Twist1 overexpression on chromosomal stability in the context of EMT in cancer cells remains unclear.

Twist1 overexpression induces chromosomal instability (CIN) in cancers of the breast (12). Spectral karyotyping (SKY) analyses of metaphases derived from Twist1 overexpressing MCF-7 (breast cancer cell line) showed an increase in chromosomal aberrations such as aneuploidy and translocations (13). Consistent with this observation, the stroma of colorectal tumors shows a positive correlation between Twist1 positive cells and CIN (14). However, the underlying mechanisms of Twist1-induced CIN remain elusive.

Another interesting vignette in our understanding of the mechanistic basis of CIN also has its origins in the maintenance of the morphology and function of the nucleus by the type V intermediate filament proteins—Lamins A/C, B1 and B2 that are localized at the inner nuclear envelope (15,16). Mutations or loss of lamins strikingly alter nuclear shapes resulting in aberrant nuclei, nuclear blebs and micronuclei, which are precursors of CIN (17). Lamin loss also impacts the cellular transcriptome (18). Interestingly, Lamin B2 knockdown shows chromosomal gains in the otherwise diploid colorectal cancer cells (DLD1) (19). Furthermore, Lamin B2 depletion shows chromosomal imbalances in colorectal cancer cells and associates with the spindle machinery, further suggesting the role of lamins in chromosome segregation in mitotic cells (20). However, the mechanisms underlying lamin functions in chromosomal stability in cancer cells are unclear.

Colorectal cancers show microsatellite instability (MSI), characterized by the insertion of repetitive nucleotide stretches, typically corrected by proteins of the mismatch repair system (MMR) such as MSH2, MSH6, MLH1 and PMS2 (21). Colorectal cancers that are mismatch repair-deficient (MMR−) show high microsatellite instability (MSI+), while mismatch repair-proficient (MMR+) colorectal cancers do not show microsatellite instability, but show elevated levels of CIN (21).

The cell cycle checkpoint and tumor suppressor protein p53 is essential for the maintenance of chromosomal stability across cancers (22,23). Furthermore, the status of p53 is potentially an important determinant of CIN, since cells with mutant p53 are associated with CIN, while cells with wild type p53 show significantly reduced CIN in cancer cells (22,24). Evidence of CIN induction exists even in the presence of wild type p53, suggesting alternate pathways of CIN induction in cancer cells (13,25). Reduction in p53 levels also enhances the susceptibility of cells to DNA damage, as ascertained by an increase in γH2AX foci (26).

With the wealth of patient data available from The Cancer Genome Atlas (TCGA)—various molecular correlates that range from mutations, copy number alterations, and expression status among others—can be attributed to target genes in specific cancer subtypes (27). Furthermore, mathematical modelling and simulations have the power to compute and predict the potential outcome of novel molecular interactions and their pathways involved in actively promoting cancers. It is therefore beyond any doubt that an interdisciplinary approach of studying theoretical and experimental paradigms is essential for cancer intervention.

Here, we show that Twist1 overexpression induces EMT to varying extents in the two colorectal cancer cell lines. Furthermore, Twist1 overexpression significantly increases nuclear and mitotic aberrations, accompanied by an increase in CIN. In addition, Twist1 induces sub-chromosomal deletions, consistent with an increase in DNA double strand breaks, as revealed by an increase in γH2AX foci. Twist1 overexpression showed a significant decrease in the levels of Spindle Assembly Checkpoint (SAC) proteins such as Bub1/R1, Mad1/2 and Aurora B Kinase, and the p53 oncoprotein, underscoring their collective role in regulating chromosomal stability in colorectal cancer cells. This was also corroborated by mathematical simulations, which showed a negative correlation between the levels of Twist1 and BubR1. Taken together, our studies suggest an overarching role of Twist1 in modulating chromosomal stability in colorectal cancer cells.

Results

Twist1 overexpression shows differential induction of EMT in colorectal cancer cells

The role of Twist1 is well established in EMT during early development and cancer progression. Transient overexpression was preferred over stable expression as a model to mimic the heterogeneous upsurge in the levels of Twist1 during cancer progression (28). We therefore studied the effect of transiently overexpressing Twist1 in two independent colorectal cancer cell lines—(i) DLD1—a near diploid, mismatch repair-deficient cell line and (ii) SW480—aneuploid, mismatch repair proficient cell line. We independently transfected these two cell lines with Twist1 and examined Twist1 protein levels by immunoblotting, which showed a significant increase in both cell lines (Fig. 1A). Consistent with previous results, Twist1 overexpression showed a significant decrease in the levels of the epithelial marker—E-cadherin (~30%), and an increase in the expression levels of the bonafide mesenchymal marker—Vimentin (~43%) in DLD1 cells (Fig. 1B). In contrast, the hyperdiploid colorectal cancer cell line SW480, showed a marked decrease (~64%) in E-cadherin levels, but only a marginal increase in Vimentin levels (Fig. 1C). We also examined the status of EMT induction at the single cell level by performing immunofluorescence staining. E-cadherin levels showed a significant decrease in both DLD1 (~50%) and SW480 cells (~45%) (Fig. 1D–F), underscoring that the decrease in epithelial mark(s) is an important event in EMT. Furthermore, these cells showed an increase in aspect ratio (DLD1 ~ 30%, SW480 ~ 24%)—a characteristic feature of cell elongation as quantified from phalloidin labelled cells (Fig. 1G–I). In summary, colorectal cancer cells exhibit EMT to varying extents upon Twist1 overexpression.

Nuclear and mitotic aberrations are enhanced in colorectal cancer cells upon Twist1 overexpression

Aberrant nuclear morphologies such as nuclear blebs and micronuclei are enhanced in cancers and characterize cancer
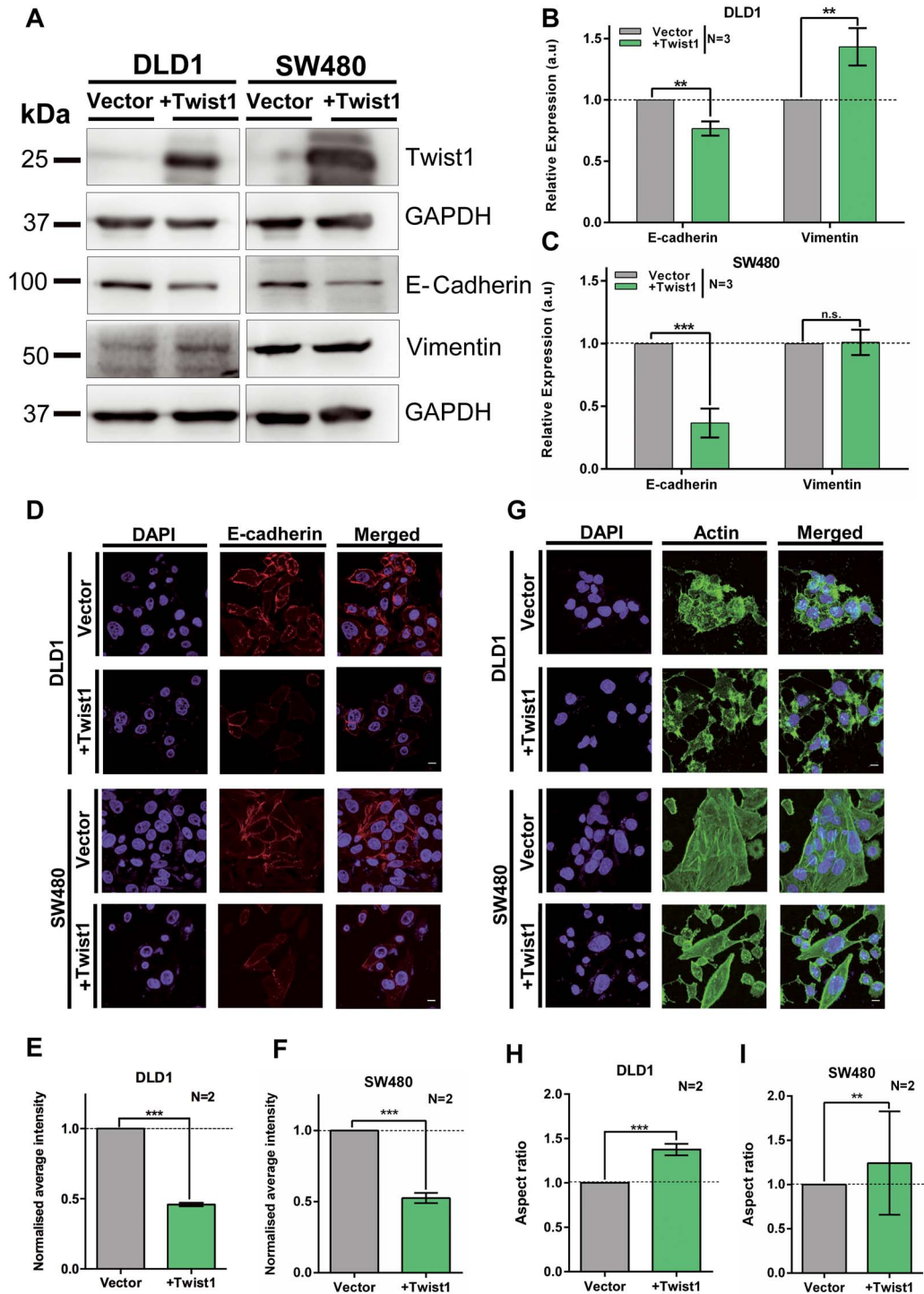
Figure 1. Differential induction of EMT upon Twist1 overexpression in colorectal cancer cells. (**A**) Representative immunoblot showing Twist1 overexpression with a concomitant decrease in E-cadherin and marginal increase in Vimentin levels in DLD1 and SW480 cell lines. (**B** and **C**) Quantification of band intensities of E-cadherin and Vimentin protein levels in DLD1 and SW480 cell lines upon Twist1 overexpression. Data from three independent biological replicates normalized to GAPDH (unpaired t-test, $N = 3$, mean $\pm$ SD, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$). (**D**) Representative mid-optical sections from confocal z-stacks of DLD1 and SW480 cells immunostained for E-cadherin, scale bar $\sim$10 µm. (**E** and **F**) Normalized fluorescence intensity of E-cadherin for vector and Twist1 overexpressing cells. Data from two independent biological replicates for DLD1 and SW480, respectively, (Mann–Whitney test, $N = 2$, $n > 60$, mean $\pm$ SD, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$). (**G**) Immunostaining for actin showing an elongated and spindle-shaped morphology upon Twist1 overexpression, scale bar $\sim$10 µm. (**H** and **I**) Quantification of aspect ratio of cells (Mann–Whitney test, $N = 2$, $n > 40$, mean $\pm$ SD, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$). N: number of independent biological replicates, n: number of cells.

progression (29). The frequency of such aberrant nuclear morphologies are diagnostic features, quantified in histopathological analyses of tissue biopsy samples (30). In addition to inducing EMT, Twist1 is also an oncoprotein (3). Here we overexpressed Twist1 and determined the frequency of nuclear blebs and micronuclei in colorectal cancer cells (Fig. 2A and B and Supplementary Material, Table S1). While there was an increase (∼5%) in the frequency of micronuclei and nuclear blebs (∼9%) in DLD1 cells, SW480 cells hardly showed an increase in these aberrations (Fig. 2B). Since aberrant nuclei are also precursors of CIN, we determined the extent of mitotic aberrations upon Twist1 overexpression (17,31). Interestingly, near diploid DLD1 cells showed a significant increase in the extent of mitotic aberrations than SW480 cells (Fig. 2C and D). Furthermore, DLD1 cells showed an increase in anaphase bridges (∼23%), lagging chromosomes (∼14%) and tripolar spindles (∼9%), while SW480 cells showed a decrease (∼3%) in anaphase bridges, accompanied by an increase in lagging chromosomes (∼16%) and tripolar spindles (∼14%), respectively (Fig. 2D and Supplementary Material, Table S2).

Having found a significant increase in mitotic aberrations associated with Twist1 overexpression, we asked if Twist1 induces CIN in colorectal cancer cells. We first analyzed the ploidy of cells upon Twist1 overexpression by flow cytometry. Neither DLD1 nor SW480 cells showed changes in their overall ploidy, upon Twist1 overexpression for ∼72 h (Fig. 2E and Supplementary Material, Fig. S1E and F).

However, we detected a significant increase in the number of cells showing whole chromosomal gains (∼23%) and losses (∼16%), upon Twist1 overexpression in DLD1 cells (Fig. 2F and G). In contrast, there was a significant increase in the number of SW480 cells, showing whole chromosomal losses (∼32%), but a decrease in cells with whole chromosome gains (∼7%) upon Twist1 overexpression (Fig. 2F and H and Supplementary Material, Fig. S2C and D, Table S3).

We asked if Twist1 overexpression also induces CIN in another near diploid colorectal cancer cell line—HCT116 (i) wild type for p53 and (ii) shows microsatellite instability (MSI+). Remarkably, HCT116 cells did not show any change in their modal chromosome numbers of 42–43, upon Twist1 overexpression (Supplementary Material, Fig. S3). This is consistent with an overarching role for wild type p53 protein in the maintenance of chromosomal stability in colorectal cancer cells.

Nuclear lamins (Lamin A/C, B1 and B2) localized at the inner nuclear envelope maintain nuclear structure and function (16). Lamins also modulate chromosomal stability in colorectal cancer cells (19,20). While immunoblotting assays show comparable levels of all three subtypes of nuclear lamins in DLD1 cells, in contrast, SW480 cells show reduced levels of endogenous B-type lamins (Fig. 2I). Interestingly, lamin levels decreased in both cell lines, with B-type lamins showing a further decrease in SW480 as compared to DLD1 cells upon Twist1 overexpression (Supplementary Material, Fig. S2E and F). These results suggest that Twist1 overexpression decreases B-type lamins levels, consistent with an increase in aberrant nuclear shapes and CIN upon loss of B-type lamins in colorectal cancer cells.

In summary, CIN is induced in a differential manner in the two colorectal cancer cell lines, upon Twist1 overexpression, since DLD1 cells exhibit both whole chromosomal gains and losses, while SW480 cells predominantly show whole chromosomal losses (Fig. 2F–H). Taken together, Twist1 overexpression induces and enhances levels of nuclear aberrations and CIN in colorectal cancer cells.

Genome-wide increase in sub-chromosomal deletions upon Twist1 overexpression

Our studies unravel a positive correlation between Twist1 overexpression and CIN in colorectal cancer cells. We therefore performed array-comparative genomic hybridization (array CGH) as an independent approach to determine the extent of amplifications and deletions at the sub-chromosomal level across the genome (Fig. 3A and B). Cells were subjected to EMT induction upon Twist1 overexpression, followed by array CGH analyses, while cells transfected with the corresponding empty vector served as reference. Analysis of array CGH data revealed sub-chromosomal amplifications and deletions across the genome (Fig. 3C–F). Sub-chromosomal deletions were more prevalent upon Twist1 overexpression (Fig. 3C and D). Surprisingly, human Chr.4, Chr.10, Chr.18 and Chr.X showed a significantly greater extent of sub-chromosomal deletions in DLD1 cells as compared to other chromosomes (Fig. 3E). SW480 cells on the other hand showed a larger repertoire of sub-chromosomal deletions that predominantly map to human Chr.3, Chr.4, Chr.6, Chr.10, Chr.13, Chr.18 and Chr.X (Fig. 3F). Deletions occurred primarily in chromosomes 4, 10, 18 and X consistently in both cell lines upon Twist1 overexpression (Fig. 3E and F). Of note, the extent of sub-chromosomal deletions was considerably elevated in SW480 than DLD1 cells. An independent array CGH analyses performed on colorectal cancer patient tumors identified copy number aberrations and exhibited sub-chromosomal deletions in human chromosomes 4, 8, and 18, respectively (32). In summary, array CGH analyses revealed a significant increase in the frequency of sub-chromosomal deletions upon Twist1 overexpression—an additional contributor of CIN (33).

Twist1 overexpression induces DNA damage and downregulates p53

A noteworthy finding from genome wide array CGH analyses was the striking increase in sub-chromosomal deletions across the genome upon Twist1 overexpression. Since chromosomal aberrations such as deletions are consequences of DSB formation (34,35), we sought to examine whether Twist1 overexpression induces DNA DSBs in colorectal cancer cells (Fig. 3G and H). We monitored the number of γH2AX foci in single cells as a marker of DNA Double Strand Breaks (DSBs) upon Twist1 overexpression and upon cisplatin treatment (50 μm), by immunofluorescence assays. Interestingly, Twist1 overexpression in DLD1 cells showed a significant increase in the number of γH2AX foci in the interphase nucleus (Fig. 3I). However, cisplatin treatment in the background of Twist1 overexpression did not alter the number of DNA damage foci (Fig. 3I). In contrast, Twist1 overexpression in SW480 cells showed a significant increase in γH2AX foci independently and in the presence of cisplatin (Fig. 3J). Taken together, this suggests that Twist1 overexpression induces and enhances DNA double strand breaks in colorectal cancer cells (Fig. 3G–J).

Since p53 is a master regulator of genome integrity in mammalian cells (36,37), we determined the effect of Twist1 overexpression on the levels of p53. Interestingly, Twist1 overexpression showed a decrease in p53 levels in DLD1 and a marginal decline in SW480 cells (Fig. 3K and L). Taken together, this suggests that the decrease in p53 levels potentially predisposes cells to elevated levels of DNA damage in cancer cells.
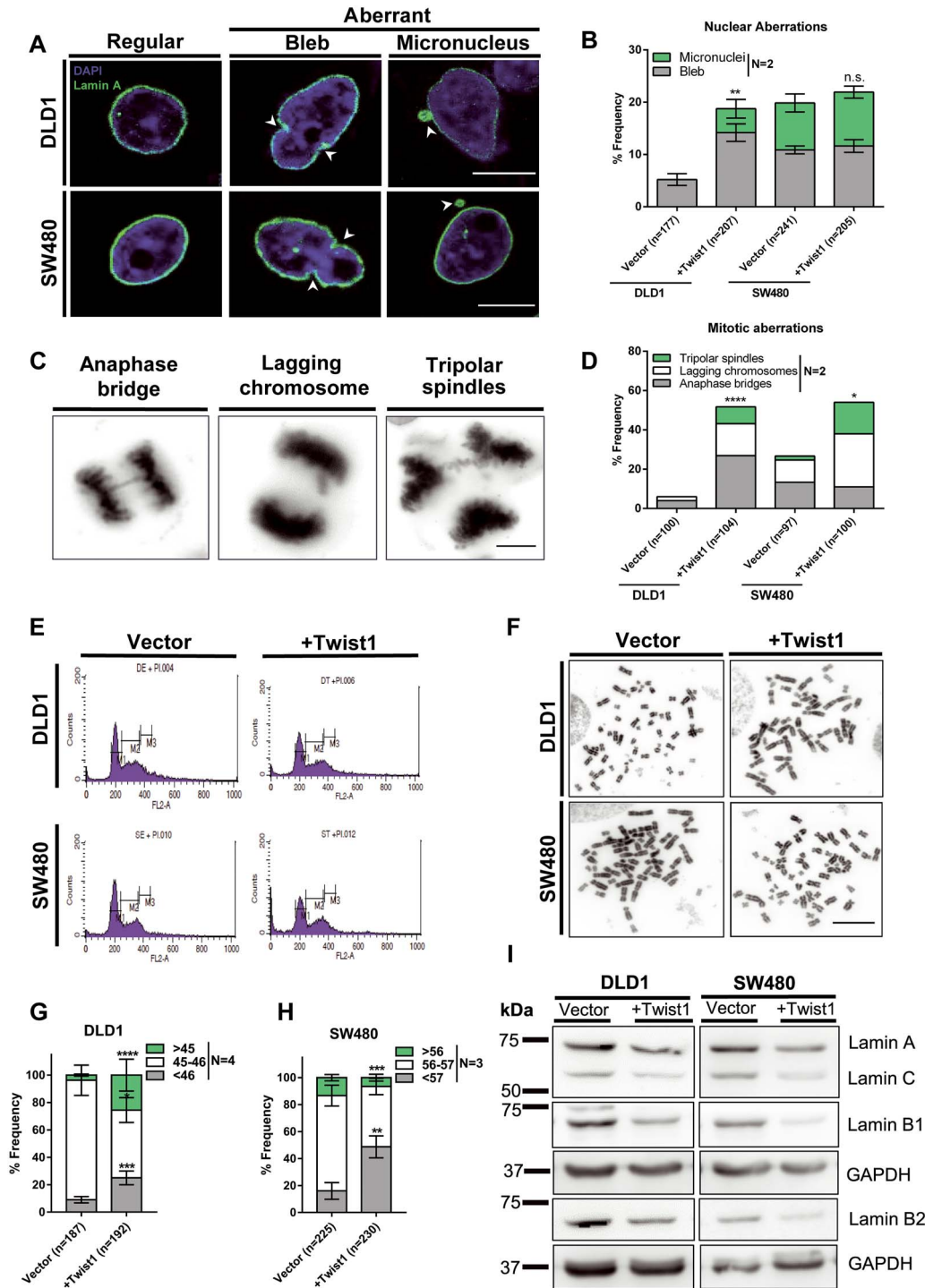
Figure 2. Twist1 overexpression enhances nuclear and mitotic aberrations in colorectal cancer cells. (**A**) Representative confocal images of nuclei upon Twist1 overexpression in DLD1 and SW480 cells immunostained with Lamin A showing nuclear blebs and micronuclei, scale bar ~10 μm. (**B**) Quantification of number of cells showing aberrant nuclei upon Twist1 overexpression. Data quantified from two independent biological replicates (Chi-square test, $N = 2$, mean with range, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$). (**C**) Representative images of mitotic aberrations showing anaphase bridges, lagging chromosomes and tripolar spindles, scale bar ~10 μm. (**D**) Quantification of the number of mitotic aberrations upon Twist1 overexpression. Data quantified from two independent biological replicates (Chi-square test, $N = 2$, mean, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$). (**E**) Representative flow cytometry profiles for ploidy analysis of vector and Twist1 overexpressing cells ($N = 2$). (**F**) Representative images of metaphase chromosome spreads derived from DLD1 and SW480 cell lines upon Twist1 overexpression, scale bar ~10 μm. (**G** and **H**) Quantification of whole chromosomal gains and losses for DLD1 and SW480 cells, respectively (data quantified from $n > 180$ independent metaphase spreads collected from $N = 3$ independent biological replicates, Z-test of proportions, mean ± SEM, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$). (**I**) Representative immunoblot showing downregulation of lamin levels upon Twist1 overexpression ($N = 2$).
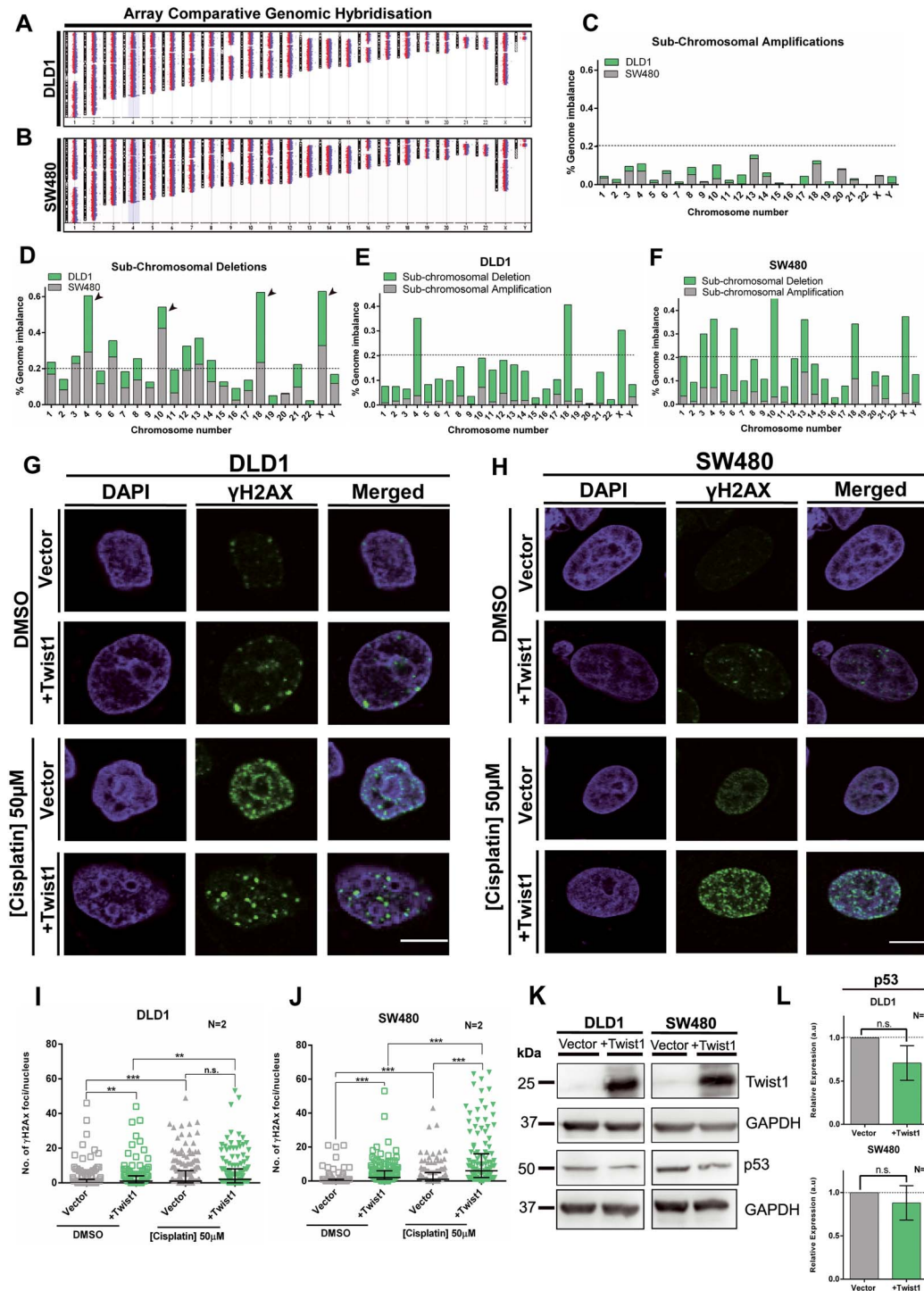
Figure 3. Twist1 overexpression induces sub-chromosomal deletions. (**A** and **B**) Chromosomal ideograms showing sub-chromosomal deletions and amplifications derived from array comparative genomic hybridisation (array CGH) for Twist1 overexpressing DLD1 and SW480 cells, respectively. (**C**) Sub-chromosomal amplifications quantified for each chromosome normalized to its total DNA content for DLD1 and SW480 cells. (**D**) Sub-chromosomal deletions quantified for each chromosome normalized to its total DNA content for DLD1 and SW480 cells. (**E**) Sub-chromosomal amplifications and deletions quantified for each chromosome normalized to its total DNA content for DLD1. (**F**) Sub-chromosomal amplifications and deletions quantified for each chromosome normalized to its total DNA content for SW480. The array CGH was from two independent biological replicates ($N = 2$, mean). (**G**) Representative mid-optical sections of DLD1 cells immunostained for γH2AX foci upon Twist1 overexpression upon DNA damage induction upon cisplatin treatment, vehicle control (DMSO), scale bar $\sim$10 µm. (**H**) Representative mid-optical sections of SW480 cells immunostained for γH2AX foci upon Twist1 overexpression upon DNA damage induced upon cisplatin treatment, vehicle control (DMSO), scale bar $\sim$10 µm. (**I** and **J**) Quantification of γH2AX foci in (I) DLD1 and (J) SW480 cells, respectively (Mann–Whitney test, $N = 2$, $n > 130$, Median-IQR, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$, n: number of nuclei). (**K**) A representative immunoblot showing p53 levels upon Twist1 overexpression in DLD1 and SW480 cells. (**L**) Quantification of p53 protein levels from band intensities normalized to GAPDH (unpaired t-test, $N = 4$, mean $\pm$ SD, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$), N: number of independent biological replicates.
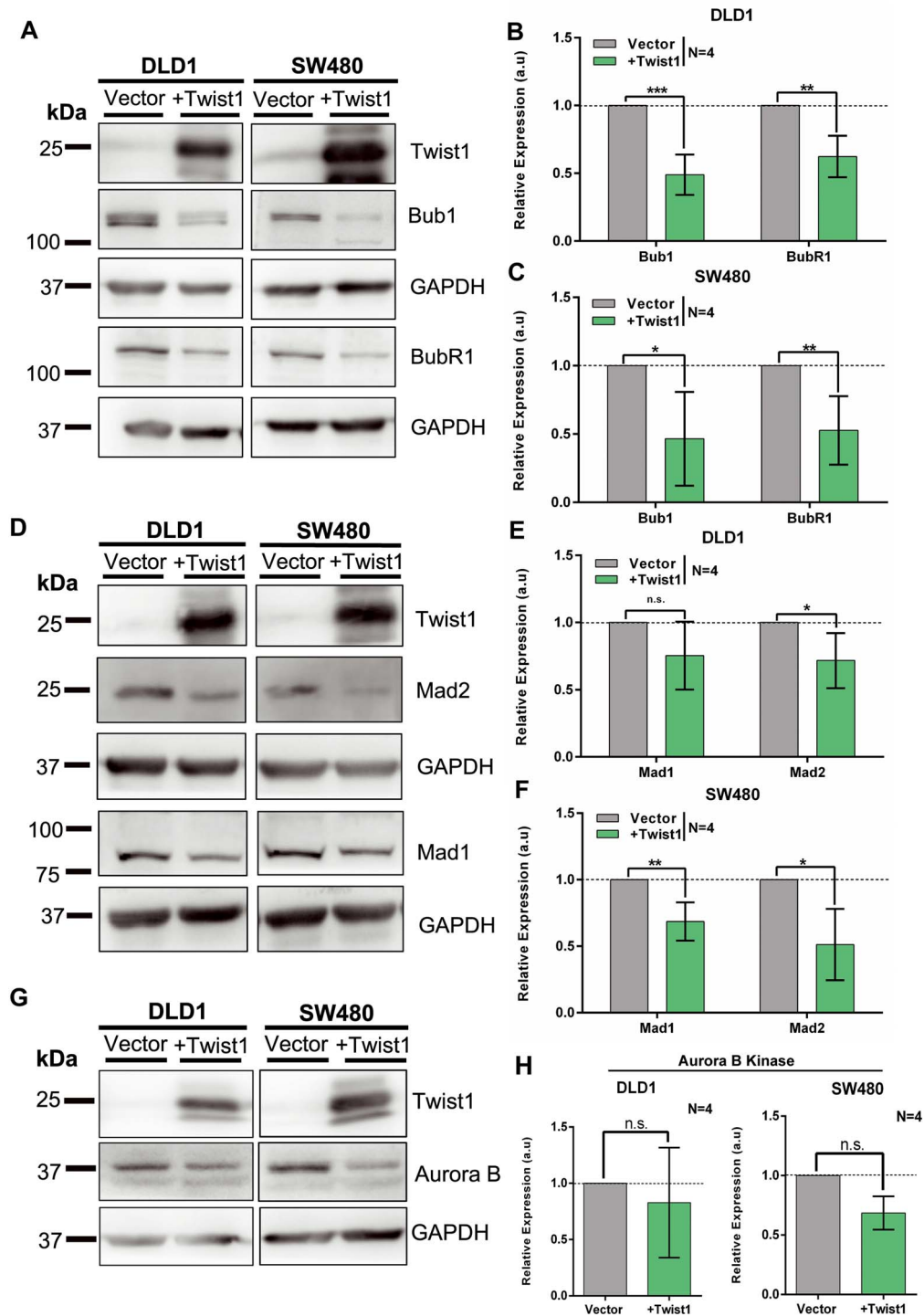
Figure 4. Twist1 overexpression shows a decrease in checkpoint proteins. (**A**) A representative immunoblot showing Twist1 overexpression, accompanied by a significant decrease in Bub1 and BubR1 levels in DLD1 and SW480 cells. (**B** and **C**) Quantification of Bub1 and BubR1 protein levels from (B) DLD1 and (C) SW480 cells, calculated from band intensities normalized to GAPDH (unpaired *t*-test, $N = 4$, mean $\pm$ SD, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$). (**D**) A representative immunoblot showing Twist1 overexpression, accompanied by decrease in mitotic checkpoint proteins Mad1 and Mad2 in DLD1 and SW480 cells, respectively. (**E** and **F**) Quantification of Mad1 and Mad2 protein levels from (E) DLD1 and (F) SW480 cells, calculated from band intensities normalized to GAPDH (unpaired *t*-test, $N = 4$, mean $\pm$ SD, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$). (**G**) A representative immunoblot showing Twist1 overexpression and decrease in Aurora B Kinase levels, in DLD1 and SW480 cells. (**H**) Quantification of Aurora B Kinase protein levels (unpaired *t*-test, normalized to loading control GAPDH, $N = 4$, mean $\pm$ SD, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ and $****P < 0.0001$).

Twist1 overexpression downregulates checkpoint proteins

We sought to address the underlying mechanisms leading to CIN upon Twist1 overexpression. As we detected a significant increase in mitotic defects and whole chromosomal aberrations, we monitored the levels of Spindle Assembly Checkpoint (SAC) proteins namely, Bub1, BubR1, Mad1, Mad2 and Aurora B Kinase (38–42). We overexpressed Twist1 independently in the two colorectal cancer cell lines and performed immunoblotting on whole cell extracts derived from these cells. Remarkably, the levels of the Bub1 and BubR1 proteins of the SAC showed a significant and comparable decrease in both colorectal cancer cell lines upon Twist1 overexpression (Fig. 4A–C). Mad1 and Mad2—components of the mitotic checkpoint complex—also showed a decrease upon Twist1 overexpression (Fig. 4D–F). In addition, Aurora B Kinase—a part of the chromosome passenger complex—also showed a decrease in protein levels (Fig. 4G and H). CIN is a consistent feature associated with the deregulation of Bub1/BubR1 levels (43–45), and decrease in their levels further affects the levels of downstream proteins such as Mad1/2 (38,46). In summary, Twist1 overexpression shows a decrease in the levels of CIN regulators, which further underscores the contribution of Twist1 to CIN in colorectal cancer cells.

We asked if the decreased levels of checkpoint proteins upon Twist1 overexpression were also elicited at the transcript level. We therefore performed RT-PCR analyses of checkpoint genes upon Twist1 overexpression. Interestingly, BUB1, BUBR1, MAD2L1 and AURKB showed a significant decrease in their transcript levels in both DLD1 and SW480 cell lines (Supplementary Material, Fig. S4B, C, E and F). In contrast, MAD1L1 showed a differential response as it was downregulated in DLD1 but significantly upregulated in SW480 cells upon Twist1 overexpression (Supplementary Material, Fig. S4D). In summary, Twist1 overexpression represses checkpoint genes at the transcriptional level, otherwise required for the maintenance of chromosomal stability of colorectal cancer cells.

Twist1 impinges on CIN regulation

To address the potential crosstalk between Twist1 and the regulators of chromosomal stability, we performed network analyses of (i) generic protein–protein interactions (47), (ii) transcription factor (TF)–gene interactions (48) of Twist1 and factors associated with EMT, CIN and DNA damage using NetworkAnalyst—a visual data analytics platform (Supplementary Material, Fig. S5A and B). From protein–protein interaction network analyses, p53 emerges as a major hub through which Twist1 regulates CIN factors, since Twist1 affects the DNA-binding activity of p53, thereby impairing its function (11). In addition, p53 directly interacts with Aurora B Kinase and Bub1, while interacting with Mad2 via FZR1 (Supplementary Material, Fig. S5A). Bub1 also interacts with Mad1 and BubR (39,46). Twist1 shows a potentially indirect interaction with Lamins via p53 and SUMO1—a posttranslation protein modifier, and CDK1—a cell cycle regulator (Supplementary Material, Fig. S5A).

ENCODE ChIP-Seq data for transcription factor (TF) enrichment on target genes (Supplementary Material, Fig. S5B) show that Twist1 may modulate TP53 activity via the histone modifier SUZ12- part of the polycomb repressive complex 2 (PRC2). Furthermore, Twist1 indirectly modulates Aurora B Kinase activity via SP3—a transcriptional repressor/activator (49). Additionally, Twist1 modulates CDH1 (E-cadherin) via EZH2 (component of the PRC2) (50). Twist1 differentially regulates lamins (LMNA, LMNB2)

through the transcriptional repressor CTBP2 (51). Transcription factor—RFXANK is enriched on LMNB2 and MAD2L1 genes. Also, NR4A1—a nuclear transcription factor—emerges as a modulator of VIM (Vimentin) and BUB1. In summary, while Twist1 functions as a transcription factor, its protein–protein interaction analyses highlight p53 as a central node, further suggestive of the role of Twist1 in modulating CIN via p53.

A simulation-based approach shows negative correlation between Twist1 and E-cadherin and BubR1 levels

We next constructed a regulatory network by integrating our experimental data with known interconnections among Twist1, E-cadherin, Vimentin, BubR1, γH2AX and p53 (Fig. 5A). (i) Twist1 overexpression downregulates E-cadherin by binding to its promoter (9) (ii) Twist1 overexpression upregulates Vimentin and is mediated by CUL2 (52) (iii) Twist1 overexpression in colorectal cancer cells downregulates BubR1 (Fig. 4A–C) (iv) Twist1 downregulates p53 levels (11) (v) BubR1 levels positively correlate with levels of p53 and γH2AX in response to DNA damage (53) (vi) p53 is required for repair of DNA damage and shows a negative correlation with γH2AX levels (26). We therefore sought to identify the robust dynamic features emerging from these interconnections. We simulated the network using the RACIPE tool (54) that models a given network using a system of ordinary differential equations (ODE). Each equation in the system represents the dynamics of a node in the network. The ODEs are then simulated for multiple parameter sets chosen randomly from a pre-defined, biologically relevant parameter space. This way, the tool allows us to capture the dynamics of the network while recapitulating the omnipresent cell-to-cell variability. The output of these simulations is the steady state values of each node, i.e. gene expression levels (Fig. 5B–D).

Across the sampled parameter sets, we observed a significant negative correlation for Twist1- E-cadherin (Fig. 5B) and positive correlation for Twist1-Vimentin (Fig. 5C). This suggests that although the extent of EMT induction via Twist1 is heterogeneous across single cells, an ensemble behavior shows robust induction of EMT by Twist1 by altering the levels of E-cadherin and Vimentin. Notwithstanding the intrinsic heterogeneity across cells, a systems biology approach corroborates our experimental data, which shows that Twist1 and BubR1 expression levels are negatively correlated (Fig. 5D).

Twist1 overexpression positively correlates with EMT and CIN: The Cancer Genome Atlas (TCGA) analyses

We sought to ask if the expression levels of Twist1 correlate with levels of (i) EMT-associated genes, (ii) CIN genes, (iii) DNA double-strand break (DSB) genes and (iv) tumor mutation burden and copy number alterations (CNAs).

We analyzed the gene expression data and somatic mutations of 30 distinct cancer types from The Cancer Genome Atlas (TCGA). This shows that Twist1 expression is evident in many primary tumors, and that their level varies within and between cancer types, likely due to the cell-of-origin and tumor stage (Fig. 6A). For example, liver hepatocellular carcinoma (LIHC) and kidney cancers (KIPAN), which originate from epithelial cells, showed the least Twist1 expression, whereas sarcoma (SARC) and uterine carcinosarcoma (UCS) that originate from mesenchymal cells showed higher expression, as previously shown at the level of EMT gene signatures (55). Colorectal cancers (COADREAD) also showed a high expression of Twist1,
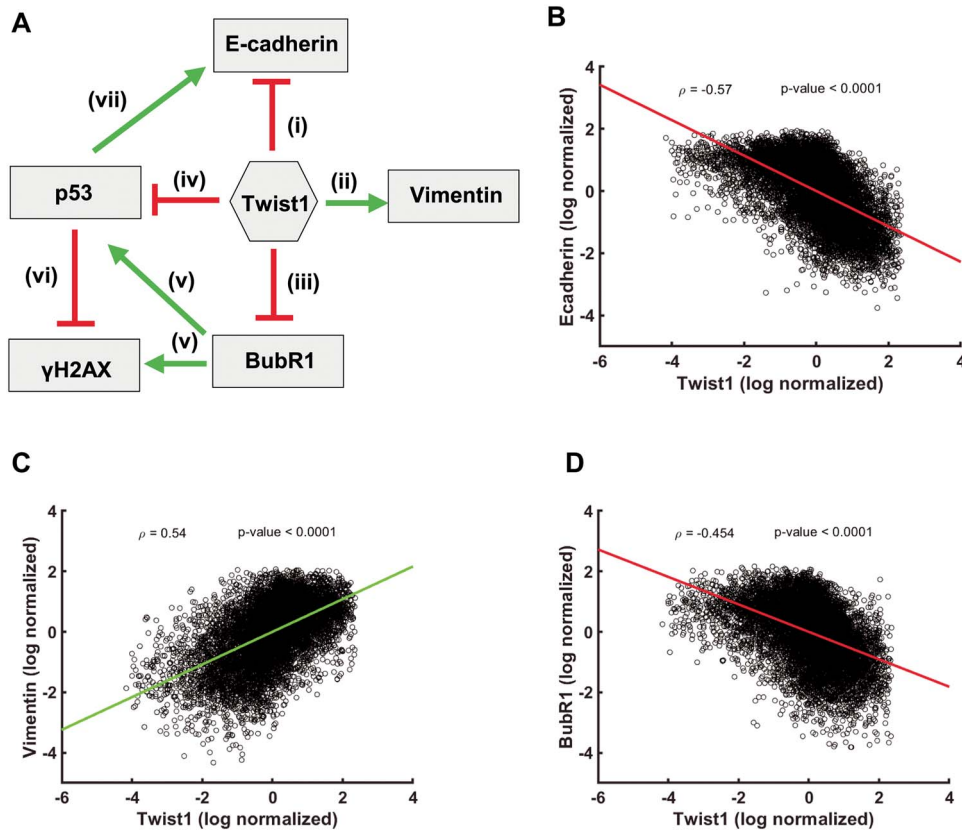
Figure 5. Correlation between levels of Twist1, EMT and CIN factors. (**A**) Network depicting the interactions among Twist1, EMT and CIN genes. Correlation plots of the log normalized gene expression values of (**B**) Twist1 and E-cadherin (**C**) Twist1 and Vimentin (**D**) Twist1 and BubR1, ρ = Pearson correlation coefficient, *P*-values show the significance of Pearson correlation.

especially in the late stage tumors (Fig. 6B). Furthermore, Twist1 expression positively correlates with EMT genes (CDH1, OCLN, TJP1, CDH2, FN1, SNAI1 and VIM) in various cancers (Fig. 6C). With CIN genes, we observed a significant positive correlation in certain tumor types, which include kidney cancers, lower-grade glioma (LGG) and lung adenocarcinoma (LUAD); whereas, in stomach adenocarcinoma (STAD) a significant negative correlation was observed between Twist1 and CIN genes. However, consistent with experimental data (Fig. 4D), colorectal cancers showed a moderate negative correlation between Twist1 and MAD2L1—a mitotic spindle assembly checkpoint protein. AURKC expression showed a marginal increase with Twist1 expression (Fig. 6C). However, we found a significant correlation with DSB genes in only a few tumor types (e.g. STAD showed a strong negative correlation similar to CIN genes).

We also examined if a correlation exists between Twist1 expression and CIN for which we compared Twist1 expression with tumor mutation burden and copy number alterations (CNAs). The total number of somatic point mutations and the fraction of genome with amplifications or deletions were considered as tumor mutation burden and CNA events, respectively. The cancer types LGG (Brain), KIPAN (Pan Kidney) and PRAD (prostate) showed a significant positive correlation with Twist1 expression for somatic mutation and for CNA events in LGG (Fig. 6C). In particular, LGG showed a significant positive correlation between TWIST1 with (i) EMT and genomic instability markers such as CIN and DSB gene expression and (ii) mutation burden and CNA event. Taken together, these results suggest that Twist1 expression is highly correlated with EMT in

various cancers. However, the correlation between TWIST1 with CIN and DSB genes is cancer subtype-specific.

Discussion

Twist1 is essential for the induction of EMT during the normal process of gastrulation during early development (4,56). However, Twist1 is overexpressed across cancers with a well-established role in metastasis (57). Here we show that Twist1 induces chromosomal and genomic instability in colorectal cancer cells (Fig. 2F–H). Twist1-induced CIN is characterized by both losses and gains of chromosomes in the near diploid DLD1 (CIN-) colorectal cancer cells, while the aneuploid SW480 (CIN+) colorectal cancer cells show chromosomal losses. Interestingly, a high-resolution approach of array comparative genomic hybridisation (array CGH) in addition to chromosomal imbalances also reveals extensive deletions at the sub-chromosomal level (Fig. 3A–F). Consistent with our results, Twist1 overexpression also induces CIN in MCF7 breast cancer cells. Spectral karyotyping (SKY) analyses show tetrasomy ($\sim$4 copies) of most chromosomes except human Chr.2, 3, 12, 18 and 21 upon Twist1 overexpression (13).

Twist1 overexpression shows a significant increase in nuclear aberrations such as nuclear blebs and micronuclei (Fig. 2A and B), consistent with decreased levels of nuclear lamins A/C and B2 (Fig. 2I). In addition, decrease in B-type lamins also induces CIN in colorectal cancer cell lines (19,20), suggestive of the involvement of lamins in the mechanistic basis of CIN induction. Lamin B2 localizes outside the spindle
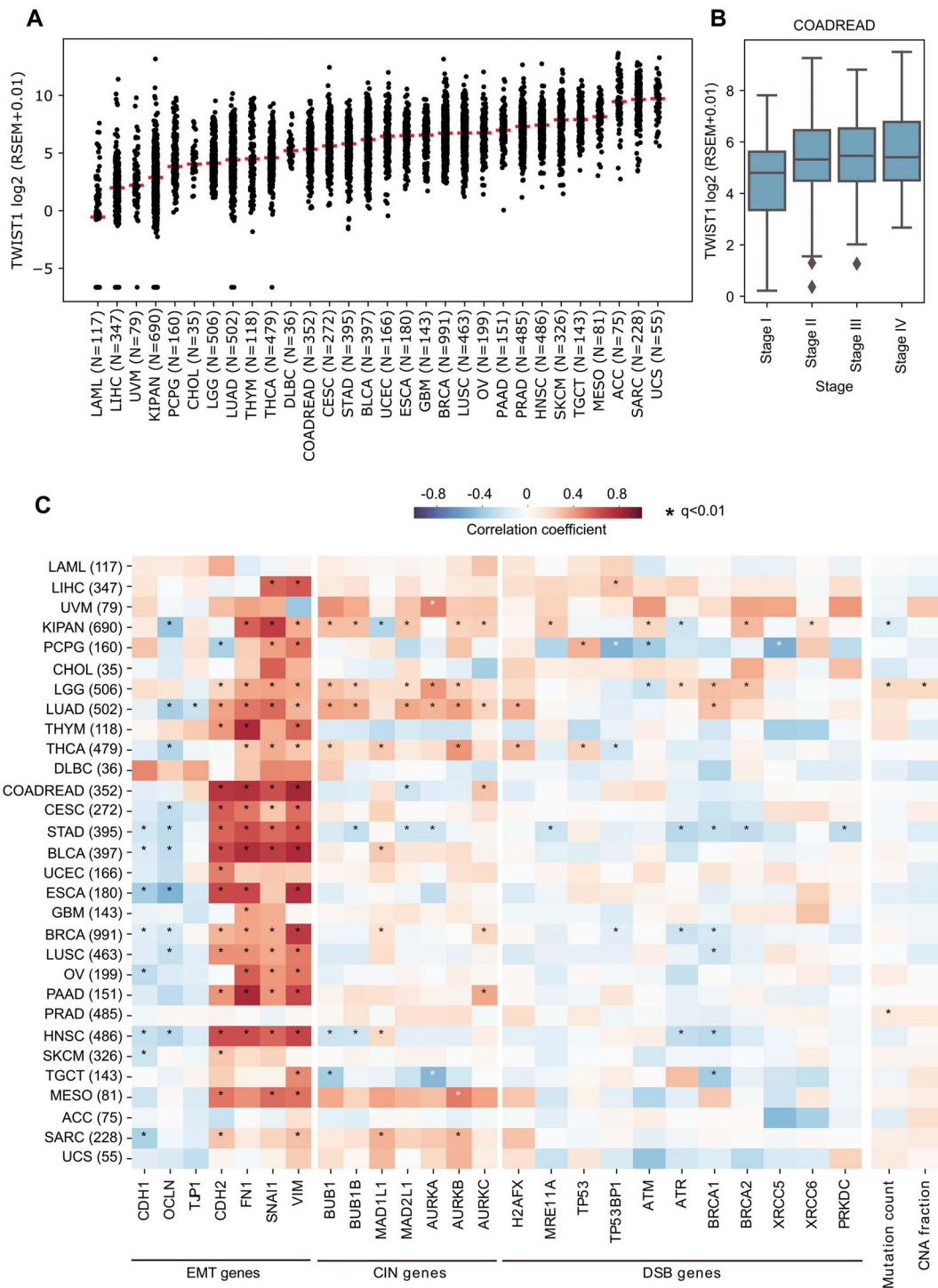
Figure 6. Correlation between gene expression levels of TWIST1, EMT and CIN across human cancers. (**A**) TWIST1 gene expression (log2 RSEM + 0.01) level within and across cancers of TCGA. Each dot represents a tumor sample and the horizontal red bar indicates the median expression value within that cancer cohort. The cancer type abbreviations are shown below. (**B**) TWIST1 expression in colorectal cancers (COADREAD) stratified by tumor stages. (**C**) Heatmap representing the correlation coefficient (and its significance) between TWIST1 expression and (i) EMT, (ii) CIN, (iii) DSB gene expression, (iv) mutation count and (v) copy number alteration (CNA) fraction for all 30 distinct cancers, computed using an iteratively reweighted least-squares approach. Color coding indicates the correlation coefficient, ranging from −1 to +1, where −1 being strong negative correlation (dark blue), 0 for no correlation (white) and +1 strong positive correlation (dark red). Significant correlations ($q < 0.01$) are marked with an asterisk (*). Cancer type abbreviations: ACC: adrenocortical carcinoma, BLCA: bladder urothelial carcinoma, BRCA: breast invasive carcinoma, CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma, CHOL: cholangiocarcinoma, COADREAD: colorectal adenocarcinoma, DLBC: diffuse large B-cell lymphoma, ESCA: esophageal carcinoma, GBM: glioblastoma multiforme, HNSC: head and neck squamous cell carcinoma, KIPAN: pan kidney carcinomas, LAML: acute myeloid leukemia, LIHC: liver hepatocellular carcinoma, LGG: brain lower grade glioma, LUAD: lung adenocarcinoma, LUSC: lung squamous cell carcinoma, MESO: mesothelioma, OV: ovarian serous cystadenocarcinoma, PAAD: pancreatic adenocarcinoma, PCPG: pheochromocytoma and paraganglioma, PRAD: prostate adenocarcinoma, SARC: sarcoma, SKCM: skin cutaneous melanoma, STAD: stomach adenocarcinoma, TGCT: testicular germ cell tumors, THCA: thyroid carcinoma, THYM: thymoma, UCS: uterine carcinosarcoma, UCEC: uterine corpus endometrial carcinoma, UVM: uveal melanoma.

poles during mitosis and has a critical role in preventing CIN in colorectal cancers by maintaining spindle pole stability and spindle assembly (20). We surmise that Twist1-mediated decrease in lamin levels is an indirect means of contributing to CIN in colorectal cancers. It is well established that lamins maintain structural and functional integrity of the nucleus in eukaryotic cells (58,59). In addition, lamin loss affects chromatin organization and gene expression across cell types (19), while decreased levels of lamin A/C are also associated with cancers (60,61). Therefore, decrease in nuclear lamins through Twist1 overexpression may impact chromatin organization and gene expression. Stem cells and undifferentiated cells are characterized by relatively reduced lamin levels and 'floppy' chromatin (62). Furthermore, the increase in lamin levels correlates with differentiation (63,64). Excess levels of Twist1 and the concomitant reduction in lamin levels may induce stemness in transformed cells and create 'founder' populations of cancer stem cells, with elevated genomic instability and resilient sub-populations of cancer cells (65). This is consistent with a marked increase in Twist1 levels across cancers as well as in the aggressive colorectal cancers inferred from TCGA patient datasets (Fig. 6A and B).

The finding that Twist1 overexpression consistently down-modulates levels of checkpoint regulators further underscores the role of Twist1 in aggravating CIN in cancers (Fig. 4) (44). While reduced levels of lamins and checkpoint factors independently induce CIN, the underlying mechanisms of how lamins crosstalk with regulators of CIN remain unanswered. Notwithstanding a striking downregulation of the checkpoint factors at the transcript (Supplementary Material, Fig. S4B–F) and protein levels (Fig. 4), CIN− (DLD1) or the CIN+ (SW480) colorectal cancer cells nevertheless resist an increase in their overall ploidy levels upon Twist1 overexpression (Fig. 2E, G–H).

Twist1 overexpression also induces EMT in human mammary epithelial cells (HMLE) (66). Analyses of Twist1 occupancy from ChIP-Seq datasets available from HMLE cells (66) shows Twist1 enrichment on E-cadherin (CDH1 gene) at −29 kbp, +20 bp and +49 kbp from the TSS (Supplementary Material, Table S5). However, Twist1 does not show promoter occupancy on kinetochore associated genes that we examined. Gene ontology analysis of genes that show promoter binding (−1 to +1 kb from TSS) of Twist1 enriched for the p53 signaling pathway (Supplementary Material, Fig. S6A) (67). Interestingly, genes of the p53 pathway—MDM2, CHEK2 and CCNB1—show Twist1 occupancy on their respective promoters (Supplementary Material, Fig. S6B). Thus, Twist1 potentially modulates the p53 signaling pathway via MDM2, CHEK2 and CCNB1, suggestive of Twist1-dependent and Twist1-independent transcriptional modulation of genes that maintain chromosomal stability (Supplementary Material, Fig. S6, Table S5). Furthermore, Twist1 showed a relatively proximal occupancy to the TSS of the Lamin A/C and Lamin B2 promoters, consistent with their repression upon Twist1 overexpression (Supplementary Material, Table S5). Since Twist1 is also enriched at diverse distant sites with respect to the TSS, this suggests a hitherto undiscovered role of Twist1 in the regulation of long-distance chromatin interactions that further impinge on chromosomal and genomic stability in cancer cells.

Interestingly, analyses of transcription factor (TF)-gene interactions from ChIP-Seq data further implies that Twist1 potentially modulates the occupancy of histone modifiers (EZH2 and SUZ12) and transcription factors (SP3 and CTBP2) that collectively impinge on the factors regulating CIN and DNA damage, in the context of EMT (Supplementary Material, Fig. S5B).

Twist1 overexpression was associated with numerical chromosomal gains and losses and an increase in cellular heterogeneity with sub-populations consisting of both CIN− and CIN+ cells (Fig. 2F–H). Of note, the status of p53 is an important determinant of CIN regulation in colorectal cancers since DLD1 and SW480 cells with mutant p53 manifest CIN, while HCT116 cells, wild type for p53, do not show CIN upon Twist1 overexpression (Supplementary Material, Fig. S3). This contrasts with MCF7 breast cancer cells (wild type for p53) (25) which show CIN upon Twist1 overexpression (13). It is noteworthy that Twist1 overexpression showed a decrease in the levels of mutant p53 protein (Fig. 3K and L), the significance of which in the context of CIN is unclear. Interestingly, analyses of protein–protein interaction networks also reveal the impact of Twist1 and its interactors that impinge on the p53 signaling pathway (Supplementary Material, Figs S5A and S6). Taken together, these evidences imply p53-dependent and p53-independent regulation of CIN upon Twist1 overexpression in the context of the genetic background of cancers of diverse origin.

The observed increase in DNA double strand breaks marked by a higher frequency of γH2AX foci upon Twist1 overexpression is a case in point that further corroborates the increase in sub-chromosomal deletions, as revealed by array CGH analyses (Fig. 3A–J). Furthermore, DNA double strand breaks are precursors of chromosomal missegregation events, which potentially contribute to Twist1 induced CIN (68,69). Of note, both the colorectal cancer cell lines show amplifications and deletions with a higher preponderance of sub-chromosomal deletions (Fig. 3D). Genomic instability may further drive phenotypic switching between epithelial and mesenchymal fates by deletions and amplifications of genes associated with these two cell states (70). We speculate that Twist1-induced genome instability potentially drives EMT and therefore cancer progression. Furthermore, while Twist1 overexpression shows a positive correlation with the expression levels of EMT genes, on the contrary, it shows a cancer-specific correlation with DNA DSBs or CIN-associated genes (Fig. 6C).

Taken together, these studies suggest that in addition to inducing EMT, Twist1 also enhances nuclear and mitotic aberrations and DNA double strand breaks that further contribute to genomic instability (Fig. 7). This is largely mediated by a collective decrease in levels of key checkpoint and genomic stability factors, underscoring the mechanistic involvement of Twist1 with CIN during EMTs.

Materials and Methods

Cell line validation

Human colorectal adenocarcinoma cell lines DLD1 and SW480 cells were validated by karyotyping and were kind gifts from the laboratory of Thomas Ried (NCI/NIH, Bethesda, USA). HCT116 colorectal cancer cells were from Mayurika Lahiri, IISER-Pune. The karyotypes of these cell lines are stable as they did not vary across passages. This was validated by analyses of metaphase spreads across passages that consistently showed a modal number of 45–46 chromosomes for the DLD1, 42–43 modal number chromosomes for HCT116 cells and a modal number of 56–57 chromosomes for the SW480 cells (Supplementary Material, Fig. S1A–D). These cells were free of mycoplasma contamination.

Cell culture and transfection

DLD1 cells were cultured in RPMI1640 (Gibco, 11875), while HCT116 and SW480 were cultured in DMEM (Gibco, 11995) media, supplemented with heat inactivated 10% FBS (Gibco, 6140) and
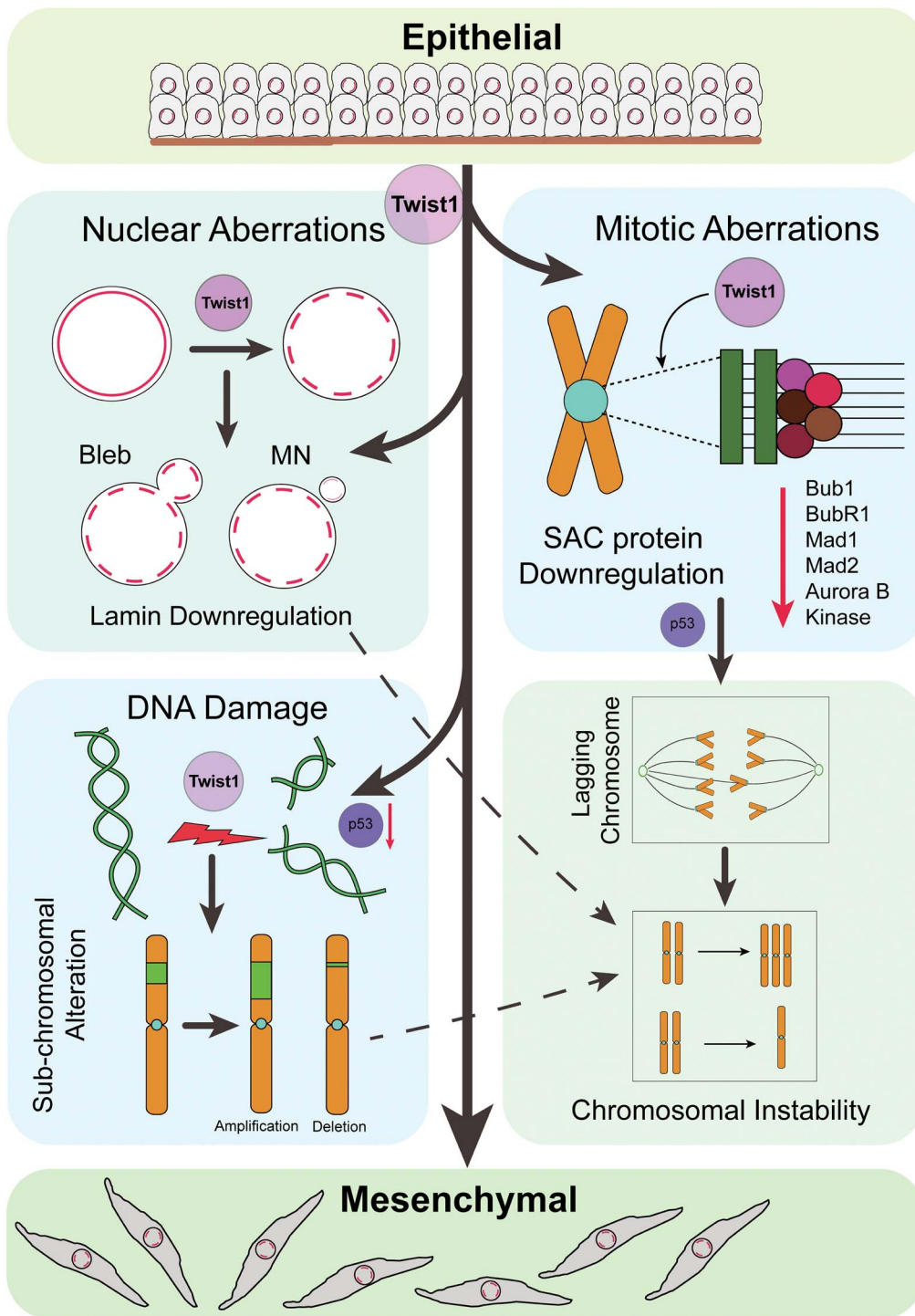
Figure 7. Speculative model suggesting a novel role for Twist1 overexpression in inducing CIN in colorectal cancer cells during EMTs. Twist1 overexpression in colorectal cancer cell lines (**i**) induces EMT, (**ii**) downregulates nuclear envelope proteins Lamin A/C, B1 and B2, associated with nuclear aberrations and CIN, (**iii**) induces DNA double strand breaks that result in enhanced sub-chromosomal alterations and CIN, potentially via p53 and (**iv**) downregulates cell cycle regulators Bub1, BubR1, Mad1, Mad2 and Aurora B Kinase leading to mitotic defects that contribute to enhanced CIN. In summary, Twist1 overexpression enhances CIN in the context of EMTs, which further contributes to cellular heterogeneity and cancer progression.

penicillin (100 units/ml)/Streptomycin (100 µg/ml) (Gibco, 15070-063). Cells were maintained in 5% CO_2 at 37°C. DLD1 and SW480 and HCT116 cells ($\sim$0.4 $\times$ 10^6) were transfected with $\sim$2 µg of pBp-mTwist1 vector (Gift from Annapoorni Rangarajan, IISc Bengaluru, India and Robert Weinberg, MIT, USA) using LTX and

PLUS (Invitrogen 15338100) reagents, with pBp-Empty vector as control. Cells were transfected with Twist1 for 24 h, and 1 µg/ml and 0.8 µg/ml puromycin (Gibco A11138), respectively, were added to select for transfected cells and cultured for another 48 h. All experiments were performed for 72 h.

Western blotting

Protein lysates were prepared by scraping cells in ice cold RIPA buffer (pH = 7.2, 50 mм Tris Cl, 150 mм NaCl, 0.1% SDS, 0.01% sodium azide, 0.5% sodium deoxycholate, 1 mм DTT, 1% NP40) containing 1X protease inhibitor cocktail (Roche, 4693116001). This was followed by centrifugation at 300 g at 4°C for 10 min. Protein estimation was performed using BCA kit (Thermo, Pierce) and an equal amount of protein was loaded onto an SDS-PAGE gel. Proteins thus resolved were transferred onto a PVDF membrane (Millipore). Immunoblots were blocked using 5% non-fat milk prepared in 1× TBST (pH 7.4). Immunodetection was performed by adding primary antibodies against Twist1 (ab50887), 1:500; E-cadherin (ab1416), 1:1000; Vimentin (Sigma, V2258), 1:500; Lamin (A + C) (ab108595), 1:1000; Lamin B1 (ab16048), 1:1000; Lamin B2 (ab8983), 1:500; Bub1 (ab54893), 1:1000; BubR1 (ab54894), 1:1000; Mad1 (ab126148), 1:3000; Mad2 (ab24588), 1:500; Aurora B Kinase (CST3094), 1:1000; Aurora B Kinase (ab2254), 1:1000; p53 (ab28) 1:1000; GAPDH (Sigma, G9545), 1:10 000. Secondary antibodies used were sheep anti-mouse-HRP (Amersham, NA9310V), 1:10 000; donkey anti-rabbit-HRP (Amersham, NA9340V) and goat anti-rat-HRP (Amersham, NA935) 1:10 000, for 1 h at RT. Between incubation, blots were rinsed thrice with 1× TBST for 10 min each at RT. Chemiluminescent substrate ECL Prime (Amersham, 89168-782) was used to develop immunoblots and imaged with ImageQuant LAS4000.

Immunofluorescence assay

$\sim 0.4 \times 10^6$ cells/well were seeded onto glass coverslips and transfections were performed as described previously. Cells were washed with 1× PBS and fixed with 4% paraformaldehyde (Sigma, P6148) prepared in 1× PBS, pH 7.4 at RT for 10 min, washed thrice in 1× PBS (5 min each). Fixation was followed by permeabilization in 0.5% Triton-X-100 prepared in 1× PBS at RT for 10 min. Cells were blocked in 1% bovine serum albumin (BSA) (Sigma, A2153) prepared in 1× PBS, for 30 min and washed three times with 1× PBS. Incubation with primary antibodies was performed in 0.1% BSA for 90 min at RT and with secondary antibodies for 60 min at RT, with washes in between using 1× PBS. Primary antibodies used were E-cadherin (ab1416) 1:500; Phalloidin Alexa Fluor 488 (A12379) 1:100; Lamin A (ab26300) 1:1000; γH2AX (ab26350) 1:750. Secondary antibodies were diluted in 1× PBS + 0.1% Triton X-100 (PBST): Goat anti Rabbit-Alexa 488 (Invitrogen, A11034), 1:1000; Goat anti Rabbit-Alexa 568 (Invitrogen, A11011), 1:1000; Goat anti mouse-Alexa 488 (Invitrogen, A11029), 1:1000; Goat anti mouse-Alexa 568 (Invitrogen, A11004), 1:1000. Cells were washed thrice in 1× PBST. Cells were counterstained with 4′,6-diamidino-2-phenylindole (DAPI) (Invitrogen, D1306) for 2 min at RT, washed in 1× PBS for 5 min and mounted in Slowfade Gold Antifade (Invitrogen, S36937). Cells were imaged on a Zeiss LSM710 confocal microscope with 405, 458 and 561 nm laser lines, using a 63× oil immersion objective, NA 1.4 at 1× digital zoom. X–Y resolution was 512 × 512. Confocal z-stacks were collected at intervals of 0.34 μm.

Flow cytometry

The empty vector and Twist1 transfected DLD1 and SW480 cells were trypsinized, washed with 1× PBS and then fixed in chilled 70% ethanol. Ethanol was added dropwise to the pellet while vortexing. This ensured fixation of all cells and minimized clumping. After chilling on ice for 15 min, the cells were centrifuged at

200 g for 10 min. The pellet was resuspended in 1× PBS, subjected to RNase (Sigma) (10 μg) treatment at 37°C for 45 min. Further, propidium iodide (Sigma) (10 μg) was added to the samples. Cell suspensions were subsequently run on FACSCalibur (BD Biosciences) and analyzed using Cell Quest Pro software.

Metaphase spread preparation

Colcemid (Roche) (1% v/v) was added to cells (empty vector, Twist1 transfected) at ∼60–70% confluency and incubated for 90 min at 37°C. The media was collected, and the cells were washed with 1× DPBS and trypsinized. Cells were centrifuged at 200 g at 4°C for 10 min. The pellet was resuspended in 4 ml pre-warmed 0.075 M KCl and incubated at 37°C for 30 min. ∼4–5 drops of fixative [methanol:acetic acid (3:1)] were added, and cells were centrifuged at 200 g at 4°C for 10 min. The supernatant was discarded, followed by two more washes in fixative. Cells were finally resuspended in ∼100–200 μl of fixative as per the volume of the pellet, followed by dropping on clean glass slides.

Array comparative genomic hybridisation (array CGH)

High-quality genomic DNA was extracted from DLD1 and SW480 cells transfected with pBp-Empty as a control and pBp-mTwist1. DNA was fragmented using restriction digestion. The control sample was labelled with Cy3 and Twist1 samples using Cy5. The DNA was hybridized on the Agilent Human 1X1M array (Agilent 073558). Image analysis was performed using Agilent Feature Extraction and Agilent CytoGenomics 3.01.1 software. Copy number alterations (CNAs) were mapped to the genome build GRCh38/hg19 for analysis and interpretation. Detailed protocol for array CGH is provided as Supplementary Information S1. We acknowledge Genotypic Technology Private Limited, Bengaluru, India for sample processing and data analysis.

RT-PCR

Total RNA was extracted using the Trizol method (71) from DLD1 and SW480 cells transfected independently with vector control and Twist1. cDNA was synthesized from 1 μg of total RNA with the Verso cDNA kit (AB-1453/B) using Olido(dT) primers. cDNA was used as a template, and RT-PCR was carried out using primers designed to span intron-exon junctions (Supplementary Information S2). GAPDH was used as internal control. Real-time quantitative PCR was performed in 5 μl reaction mixture containing KAPA SYBR FAST qPCR Master Mix (2×) (KK4602, Merck) and 2 μм each of the forward and reverse primer using the Bio-Rad RT-PCR instrument (CFX96 Touch). Fold change in expression was calculated by double normalization of Ct values to the internal control (GAPDH) and empty vector control by the $2^{-\Delta\Delta Ct}$ method (72).

Statistical analysis and graphs

A minimum of 30 cells were analyzed for each biological replicate. All experiments were performed in at least $N = 2$ independent biological replicates. The number of technical replicates (n) differs for each experiment. Statistical analysis was performed, and graphs were plotted with GraphPad Prism 6 software.

Image processing and analysis

Images were quantified using ImageJ software. E-cadherin levels were measured by tracing out E-cadherin staining manually, and intensity was measured along the traced line. For actin staining,

aspect ratio was calculated as a ratio of major axis/minor axis. For analyses of γH2AX foci, thresholding was performed for each nucleus counterstained with DAPI, and the 'find maxima' function was used to enumerate the number of γH2AX foci per nucleus.

Mathematical modeling

The network was simulated using the tool 'RAndomized CIrcuit PErturbation (RACIPE)' (Supplementary Information S3) (54). RACIPE models a given regulatory network using a system of Ordinary Differential Equations (ODEs) and samples multiple parameter sets randomly via a uniform distribution from a predefined range of parameters. For each parameter set, the system of ODEs representing the interactions in the network is simulated at multiple initial conditions to identify the number of steady states. For the current analysis, 10 000 parameter sets were sampled, and 100 random initial conditions were chosen for each parameter set. The ODEs were integrated using Euler's method of numerical integration. Linear regression was used to fit the gene expression data obtained from RACIPE to a line. Corresponding P-value ranges are reported.

TCGA expression analysis

Gene expression (RSEM gene-normalized, version 2016_01_28) and somatic mutation data (MC3) of TCGA samples ($n = 8657$) across 30 tumor types were downloaded from Firebrowse server (http://firebrowse.org). The correlation coefficient between TWIST1 expression and other gene expressions/mutation burden/copy number alterations and its significance were computed using iteratively reweighted least-squares approach. To adjust for multiple hypothesis testing, Bonferroni correction on P values per gene set was performed, and $q < 0.01$ was considered as significant. All plots were generated using the Seaborn package in Python. The results shown here are in whole or part based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

Supplementary Material

Supplementary Material is available at *HMGJ* online.

Acknowledgements

We thank IISER Pune for equipment, microscopy and flow cytometry facilities. We thank Dr Nagaraj Balasubramanian, Biology, IISER Pune for reagents. We thank Ajay Labade, Devika Ranade and Jiffin Benjamin, Chromosome Biology Lab (CBL) for critical comments. Constructs were gifts from Prof. Annapoorni Rangarajan, IISc Bengaluru and Prof. Robert Weinberg, MIT, USA. We thank Prof. Annapoorni Rangarajan for useful suggestions. Genotypic Technology Private Limited, Bengaluru, for sample processing and analyses of array CGH data. We acknowledge F1000, Google, Skype and Zoom for facilitating our seamless interaction while writing this manuscript.

Conflict of Interest Statement. The authors declare no conflict of interest.

Funding

Wellcome Trust, Department of Biotechnology India Alliance Intermediate Fellowship (Grant #500164/Z/09/Z); Department

Author Contributions

The study was designed by M.K., D.S., S.J., A.K. and K.S. Experiments were performed by M.K., D.S., S.J. and A.K. Computational analyses were performed by K.S.H., F.E.E., M.K.J. and R.S. The manuscript was written by M.K., D.S., S.J., M.K.J., R.S. and K.S.

References

1. Vesuna, F., Lisok, A., Kimble, B. and Raman, V. (2009) Twist modulates breast cancer stem cells by transcriptional regulation of CD24 expression. *Neoplasia*, **11**, 1318–1328.
2. Vesuna, F., Bergman, Y. and Raman, V. (2017) Genomic pathways modulated by twist in breast cancer. *BMC Cancer*, **17**, 52.
3. Zhao, Z., Rahman, M.A., Chen, Z.G. and Shin, D.M. (2017) Multiple biological functions of Twist1 in various cancers. *Oncotarget*, **8**, 20380–20393.
4. Yang, J., Mani, S.A., Donaher, J.L., Ramaswamy, S., Itzykson, R.A., Come, C., Savagner, P., Gitelman, I., Richardson, A. and Weinberg, R.A. (2004) Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell*, **117**, 927–939.
5. Beck, B., Lapouge, G., Rorive, S., Drogat, B., Desaedelaere, K., Delafaille, S., Dubois, C., Salmon, I., Willekens, K., Marine, J.-C. *et al.* (2015) Different levels of Twist1 regulate skin tumor initiation, stemness, and progression. *Cell Stem Cell*, **16**, 67–79.
6. Mani, S.A., Guo, W., Liao, M.-J., Eaton, E.N., Ayyanan, A., Zhou, A.Y., Brooks, M., Reinhard, F., Zhang, C.C., Shipitsin, M. *et al.* (2008) The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell*, **133**, 704–715.
7. Morel, A.-P., Lièvre, M., Thomas, C., Hinkal, G., Ansieau, S. and Puisieux, A. (2008) Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PLoS One*, **3**, e2888.
8. Jolly, M.K., Huang, B., Lu, M., Mani, S.A., Levine, H. and Ben-Jacob, E. (2014) Towards elucidating the connection between epithelial-mesenchymal transitions and stemness. *J. R. Soc. Interface*, **11**, 20140962.
9. Vesuna, F., van Diest, P., Chen, J.H. and Raman, V. (2008) Twist is a transcriptional repressor of E-cadherin gene expression in breast cancer. *Biochem. Biophys. Res. Commun.*, **367**, 235–241.
10. Brableltz, T., Hlubek, F., Spaderna, S., Schmalhofer, O., Hiendlmeyer, E., Jung, A. and Kirchner, T. (2005) Invasion and metastasis in colorectal cancer: epithelial-mesenchymal transition, mesenchymal-epithelial transition, stem cells and beta-catenin. *Cells Tissues Organs*, **179**, 56–65.
11. Piccinin, S., Tonin, E., Sessa, S., Demontis, S., Rossi, S., Pecciarini, L., Zanatta, L., Pivetta, F., Grizzo, A., Sonego, M. *et al.* (2012) A "twist box" code of p53 inactivation: twist box: p53 interaction promotes p53 degradation. *Cancer Cell*, **22**, 404–415.
12. Mironchik, Y., Winnard, P.T., Vesuna, F., Kato, Y., Wildes, F., Pathak, A.P., Kominsky, S., Artemov, D., Bhujwalla, Z., Van Diest, P. *et al.* (2005) Twist overexpression induces in vivo

angiogenesis and correlates with chromosomal instability in breast cancer. *Cancer Res.*, **65**, 10801–10809.

13. Vesuna, F., Winnard, P., Glackin, C. and Raman, V. (2006) Twist overexpression promotes chromosomal instability in the breast cancer cell line MCF-7. *Cancer Genet. Cytogenet.*, **167**, 189–191.

14. Celesti, G., Di Caro, G., Bianchi, P., Grizzi, F., Basso, G., Marchesi, F., Doni, A., Marra, G., Roncalli, M., Mantovani, A. *et al.* (2013) Presence of Twist1-positive neoplastic cells in the stroma of chromosome-unstable colorectal tumors. *Gastroenterology*, **145**, 647–657.e15.

15. Gerace, L. and Huber, M.D. (2012) Nuclear lamina at the crossroads of the cytoplasm and nucleus. *J. Struct. Biol.*, **177**, 24–31.

16. Goldman, R.D., Gruenbaum, Y., Moir, R.D., Shumaker, D.K. and Spann, T.P. (2002) Nuclear lamins: building blocks of nuclear architecture. *Genes Dev.*, **16**, 533–547.

17. Hatch, E.M., Fischer, A.H., Deerinck, T.J. and Hetzer, M.W. (2013) Catastrophic nuclear envelope collapse in cancer cell micronuclei. *Cell*, **154**, 47–60.

18. Zheng, X., Hu, J., Yue, S., Kristiani, L., Kim, M., Sauria, M., Taylor, J., Kim, Y. and Zheng, Y. (2018) Lamins organize the global three-dimensional genome from the nuclear periphery. *Mol. Cell*, **71**, 802–815.e7.

19. Ranade, D., Koul, S., Thompson, J., Prasad, K.B. and Sengupta, K. (2017) Chromosomal aneuploidies induced upon Lamin B2 depletion are mislocalized in the interphase nucleus. *Chromosoma*, **126**, 223–244.

20. Kuga, T., Nie, H., Kazami, T., Satoh, M., Matsushita, K., Nomura, F., Maeshima, K., Nakayama, Y. and Tomonaga, T. (2014) Lamin B2 prevents chromosome instability by ensuring proper mitotic chromosome segregation. *Oncogene*, **3**, e94.

21. De' Angelis, G.L., Bottarelli, L., Azzoni, C., De' Angelis, N., Leandro, G., Di Mario, F., Gaiani, F. and Negri, F. (2018) Microsatellite instability in colorectal cancer. *Acta Biomed*, **89**, 97–101.

22. Hanel, W. and Moll, U.M. (2012) Links between mutant p53 and genomic instability. *J. Cell. Biochem.*, **113**, 433–439.

23. Yeo, C.Q.X., Alexander, I., Lin, Z., Lim, S., Aning, O.A., Kumar, R., Sangthongpitag, K., Pendharkar, V., Ho, V.H.B. and Cheok, C.F. (2016) p53 maintains genomic stability by preventing interference between transcription and replication. *Cell Rep.*, **15**, 132–146.

24. Ahmed, D., Eide, P.W., Eilertsen, I.A., Danielsen, S.A., Eknæs, M., Hektoen, M., Lind, G.E. and Lothe, R.A. (2013) Epigenetic and genetic features of 24 colon cancer cell lines. *Oncogene*, **2**, e71.

25. Leroy, B., Girard, L., Hollestelle, A., Minna, J.D., Gazdar, A.F. and Soussi, T. (2014) Analysis of TP53 mutation status in human cancer cell lines: a reassessment. *Hum. Mutat.*, **35**, 756–765.

26. Van Oorschot, B., Oei, A.L., Nuijens, A.C., Rodermond, H., Hoeben, R., Stap, J., Stalpers, L.J. and Franken, N.A.P. (2014) Decay of γ-H2AX foci correlates with potentially lethal damage repair and P53 status in human colorectal carcinoma cells. *Cell. Mol. Biol. Lett.*, **19**, 37–51.

27. Cancer Genome Atlas Research Network (2014) Comprehensive molecular profiling of lung adenocarcinoma. *Nature*, **511**, 543–550.

28. Gajula, R.P., Chettiar, S.T., Williams, R.D., Nugent, K., Kato, Y., Wang, H., Malek, R., Taparra, K., Cades, J., Annadanam, A. *et*

al. (2015) Structure-function studies of the bHLH phosphorylation domain of TWIST1 in prostate cancer cells. *Neoplasia*, **17**, 16–31.

29. Smith, E.R., Capo-Chichi, C.D. and Xu, X.-X. (2018) Defective nuclear lamina in aneuploidy and carcinogenesis. *Front. Oncol.*, **8**, 529.

30. Zink, D., Fischer, A.H. and Nickerson, J.A. (2004) Nuclear structure in cancer cells. *Nat. Rev. Cancer*, **4**, 677–687.

31. Capo-chichi, C.D., Cai, K.Q., Simpkins, F., Ganjei-Azar, P., Godwin, A.K. and Xu, X.-X. (2011) Nuclear envelope structural defects cause chromosomal numerical instability and aneuploidy in ovarian cancer. *BMC Med.*, **9**, 28.

32. Ashktorab, H., Schäffer, A.A., Daremipouran, M., Smoot, D.T., Lee, E. and Brim, H. (2010) Distinct genetic alterations in colorectal cancer. *PLoS One*, **5**, e8879.

33. Rajagopalan, H., Nowak, M.A., Vogelstein, B. and Lengauer, C. (2003) The significance of unstable chromosomes in colorectal cancer. *Nat. Rev. Cancer*, **3**, 695–701.

34. Varga, T. and Aplan, P.D. (2005) Chromosomal aberrations induced by double strand DNA breaks. *DNA Repair (Amst)*, **4**, 1038–1046.

35. Cannan, W.J. and Pederson, D.S. (2016) Mechanisms and consequences of double-strand DNA break formation in chromatin. *J. Cell. Physiol.*, **231**, 3–14.

36. Agapova, L.S., Ilyinskaya, G.V., Turovets, N.A., Ivanov, A.V., Chumakov, P.M. and Kopnin, B.P. (1996) Chromosome changes caused by alterations of p53 expression. *Mutat. Res.*, **354**, 129–138.

37. Giono, L.E. and Manfredi, J.J. (2006) The p53 tumor suppressor participates in multiple cell cycle checkpoints. *J. Cell. Physiol.*, **209**, 13–20.

38. Ditchfield, C., Johnson, V.L., Tighe, A., Ellston, R., Haworth, C., Johnson, T., Mortlock, A., Keen, N. and Taylor, S.S. (2003) Aurora B couples chromosome alignment with anaphase by targeting BubR1, Mad2, and Cenp-E to kinetochores. *J. Cell Biol.*, **161**, 267–280.

39. Burke, D.J. and Stukenberg, P.T. (2008) Linking kinetochore-microtubule binding to the spindle checkpoint. *Dev. Cell*, **14**, 474–479.

40. Diogo, V., Teixeira, J., Silva, P.M.A. and Bousbaa, H. (2017) Spindle assembly checkpoint as a potential target in colorectal cancer: current status and future perspectives. *Clin. Colorectal Cancer*, **16**, 1–8.

41. Logarinho, E., Bousbaa, H., Dias, J.M., Lopes, C., Amorim, I., Antunes-Martins, A. and Sunkel, C.E. (2004) Different spindle checkpoint proteins monitor microtubule attachment and tension at kinetochores in drosophila cells. *J. Cell Sci.*, **117**, 1757–1771.

42. Elowe, S. (2011) Bub1 and BubR1: at the interface between chromosome attachment and the spindle checkpoint. *Mol. Cell. Biol.*, **31**, 3085–3093.

43. Nicholson, J.M., Macedo, J.C., Mattingly, A.J., Wangsa, D., Camps, J., Lima, V., Gomes, A.M., Dória, S., Ried, T., Logarinho, E. *et al.* (2015) Chromosome mis-segregation and cytokinesis failure in trisomic human cells. *elife*, **4**, 1–23.

44. Baker, D.J., Jin, F., Jeganathan, K.B. and van Deursen, J.M. (2009) Whole chromosome instability caused by Bub1 insufficiency drives tumorigenesis through tumor suppressor gene loss of heterozygosity. *Cancer Cell*, **16**, 475–486.

45. Lee, H. (2003) Impaired phosphorylation and mislocalization of Bub1 and BubR1 are responsible for the defective mitotic checkpoint function in Brca2-mutant thymic lymphomas. *Exp. Mol. Med.*, **35**, 448–453.

46. Johnson, V.L., Scott, M.I.F., Holt, S.V., Hussein, D. and Taylor, S.S. (2004) Bub1 is required for kinetochore localization of BubR1, Cenp-E, Cenp-F and Mad2, and chromosome congression. *J. Cell Sci.*, **117**, 1577–1589.

47. Xia, J., Benner, M.J. and Hancock, R.E.W. (2014) NetworkAnalyst–integrative approaches for protein-protein interaction network analysis and visual exploration. *Nucleic Acids Res.*, **42**, W167–W174.

48. Xia, J., Lyle, N.H., Mayer, M.L., Pena, O.M. and Hancock, R.E.W. (2013) INVEX–a web-based tool for integrative visualization of expression data. *Bioinformatics*, **29**, 3232–3234.

49. Majello, B., De Luca, P. and Lania, L. (1997) Sp3 is a bifunctional transcription regulator with modular independent activation and repression domains. *J. Biol. Chem.*, **272**, 4021–4026.

50. Malouf, G.G., Taube, J.H., Lu, Y., Roysarkar, T., Panjarian, S., Estecio, M.R., Jelinek, J., Yamazaki, J., Raynal, N.J.-M., Long, H. *et al.* (2013) Architecture of epigenetic reprogramming following Twist1-mediated epithelial-mesenchymal transition. *Genome Biol.*, **14**, R144.

51. Zhang, W., Duan, N., Zhang, Q., Song, T., Li, Z., Chen, X. and Wang, K. (2018) The intracellular NADH level regulates atrophic nonunion pathogenesis through the CtBP2-p300-Runx2 transcriptional complex. *Int. J. Biol. Sci.*, **14**, 2023–2036.

52. Meng, J., Chen, S., Han, J.-X., Qian, B., Wang, X.-R., Zhong, W.-L., Qin, Y., Zhang, H., Gao, W.-F., Lei, Y.-Y. *et al.* (2018) Twist1 regulates vimentin through Cul2 circular RNA to promote EMT in hepatocellular carcinoma. *Cancer Res.*, **78**, 4150–4162.

53. Fang, Y., Liu, T., Wang, X., Yang, Y.M., Deng, H., Kunicki, J., Traganos, F., Darzynkiewicz, Z., Lu, L. and Dai, W. (2006) BubR1 is involved in regulation of DNA damage responses. *Oncogene*, **25**, 3598–3605.

54. Huang, B., Lu, M., Jia, D., Ben-Jacob, E., Levine, H. and Onuchic, J.N. (2017) Interrogating the topological robustness of gene regulatory circuits by randomization. *PLoS Comput. Biol.*, **13**, e1005456.

55. Gibbons, D.L. and Creighton, C.J. (2018) Pan-cancer survey of epithelial-mesenchymal transition markers across the cancer genome atlas. *Dev. Dyn.*, **247**, 555–564.

56. Thisse, B., el Messal, M. and Perrin-Schmitt, F. (1987) The twist gene: isolation of a drosophila zygotic gene necessary for the establishment of dorsoventral pattern. *Nucleic Acids Res.*, **15**, 3439–3453.

57. Eckert, M.A., Lwin, T.M., Chang, A.T., Kim, J., Danis, E., Ohno-Machado, L. and Yang, J. (2011) Twist1-induced invadopodia formation promotes tumor metastasis. *Cancer Cell*, **19**, 372–386.

58. Dechat, T., Adam, S.A. and Goldman, R.D. (2009) Nuclear lamins and chromatin: when structure meets function. *Adv. Enzym. Regul.*, **49**, 157–166.

59. Shimi, T., Pfleghaar, K., Kojima, S., Pack, C.-G., Solovei, I., Goldman, A.E., Adam, S.A., Shumaker, D.K., Kinjo, M., Cremer, T. *et al.* (2008) The A- and B-type nuclear Lamin networks: microdomains involved in chromatin organization and transcription. *Genes Dev.*, **22**, 3409–3421.

60. Wu, Z., Wu, L., Weng, D., Xu, D., Geng, J. and Zhao, F. (2009) Reduced expression of Lamin A/C correlates with poor histological differentiation and prognosis in primary gastric carcinoma. *J. Exp. Clin. Cancer Res.*, **28**, 8.

61. Belt, E.J.T., Fijneman, R.J.A., van den Berg, E.G., Bril, H., Delis-van Diemen, P.M., Tijssen, M., van Essen, H.F., de Lange-de Klerk, E.S.M., Beliën, J.A.M., Stockmann, H.B.A.C. *et al.* (2011) Loss of Lamin A/C expression in stage II and III colon cancer is associated with disease recurrence. *Eur. J. Cancer*, **47**, 1837–1845.

62. Melcer, S., Hezroni, H., Rand, E., Nissim-Rafinia, M., Skoultchi, A., Stewart, C.L., Bustin, M. and Meshorer, E. (2012) Histone modifications and Lamin A regulate chromatin protein dynamics in early embryonic stem cell differentiation. *Nat. Commun.*, **3**, 910.

63. Mahajani, S., Giacomini, C., Marinaro, F., De Pietri Tonelli, D., Contestabile, A. and Gasparini, L. (2017) Lamin B1 levels modulate differentiation into neurons during embryonic corticogenesis. *Sci. Rep.*, **7**, 4897.

64. Bermeo, S., Vidal, C., Zhou, H. and Duque, G. (2015) Lamin a/C acts as an essential factor in mesenchymal stem cell differentiation through the regulation of the dynamics of the Wnt/β-catenin pathway. *J. Cell. Biochem.*, **116**, 2344–2353.

65. Greaves, M. (2013) Cancer stem cells as "units of selection". *Evol. Appl.*, **6**, 102–108.

66. Chang, A.T., Liu, Y., Ayyanathan, K., Benner, C., Jiang, Y., Prokop, J.W., Paz, H., Wang, D., Li, H.-R., Fu, X.-D. *et al.* (2015) An evolutionarily conserved DNA architecture determines target specificity of the TWIST family bHLH transcription factors. *Genes Dev.*, **29**, 603–616.

67. Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, **4**, 44–57.

68. Bakhoum, S.F., Kabeche, L., Murnane, J.P., Zaki, B.I. and Compton, D.A. (2014) DNA-damage response during mitosis induces whole-chromosome missegregation. *Cancer Discov.*, **4**, 1281–1289.

69. Burrell, R.A., McClelland, S.E., Endesfelder, D., Groth, P., Weller, M.-C., Shaikh, N., Domingo, E., Kanu, N., Dewhurst, S.M., Gronroos, E. *et al.* (2013) Replication stress links structural and numerical cancer chromosomal instability. *Nature*, **494**, 492–496.

70. Gao, C., Su, Y., Koeman, J., Haak, E., Dykema, K., Essenberg, C., Hudson, E., Petillo, D., Khoo, S.K. and Vande Woude, G.F. (2016) Chromosome instability drives phenotypic switching to metastasis. *Proc. Natl. Acad. Sci. U. S. A.*, **113**, 14793–14798.

71. Rio, D.C., Ares, M., Hannon, G.J. and Nilsen, T.W. (2010) Purification of RNA using TRIzol (TRI reagent). *Cold Spring Harb. Protoc.*, **2010**, pdb.prot5439.

72. Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(−Delta Delta C(T)) method. *Methods*, **25**, 402–408.